First Monday, Volume 20, Number 1 - 5 January 2015

Surveying the citizen science landscape

by Andrea Wiggins and Kevin Crowston

Abstract

Citizen science has seen enormous growth in recent years, in part due to the influence of the Internet, and a corresponding growth in interest. However, the few stand-out examples that have received attention from media and researchers are not representative of the diversity of the field as a whole, and therefore may not be the best models for those seeking to study or start a citizen science project. In this work, we present the results of a survey of citizen science project leaders, identifying sub-groups of project types according to a variety of features related to project design and management, including funding sources, goals, participant activities, data quality processes, and social interaction. These combined features highlight the diversity of citizen science, providing an overview of the breadth of the phenomenon and laying a foundation for comparison between citizen science projects and to other online communities.

Contents

Introduction

Citizen science — research projects engaging the public as contributors to science in fields like astronomy, ecology, and even genomics — has grown in popularity and received a lot of media attention lately. However, much of the attention has focused on a few projects that do not necessarily provide the most appropriate models for those seeking to study or start a citizen science project. This paper provides an empirical foundation for characterizing the diverse phenomenon that is citizen science and describes a breadth of citizen science practice that is much richer than is generally recognized.

Public awareness of citizen science today is generally limited to a few projects. In this, the state of citizen science is much like the early days of open source software research when most people believed that all OSS projects resembled Linux, a project we now recognize as unique in its structure and operation. Similarly, citizen science is currently perceived as science-oriented crowdsourcing that involves either image recognition or reporting wildlife sightings, but in reality, the diversity of the field is much greater.

Broadly defined, citizen science is a form of research collaboration that involves volunteers in producing authentic scientific research. Drastically simplified, citizen science is research accomplished by engaging humans as “sensors” to collect scientific data or as “processors” to manipulate scientific data or to solve data analysis problems. Involving people beyond professional scientists in a project enables unique new research, cumulating the contributions of many individuals (Bonney, *et al.*, 2014) and taking advantage of human competencies that can be substantially more sophisticated than machines. To produce sound science, citizen science projects incorporate means of assuring the quality of the research activities according to accepted scientific norms (Wiggins, *et al.*, 2011), which distinguishes them from many other forms of collective content production.

A variety of labels have been used to describe citizen science and related forms of public participation in scientific research. Among these, volunteer monitoring, participatory action research, civic science, action science, and community science have been used most often. Each label is specific to a particular set of research disciplines and emphasizes different features of public engagement in science. For example, *civic science* (Irwin, 1995) focuses on public participation in making decisions about science, whereas *community science* (Wandersman, 2003) describes participatory community-centered social science. The variety of related terms describing the larger space of public participation in science demonstrates the diversity of both the research disciplines in which citizen science is applied (*e.g.*, ecology, astronomy, conservation biology, climatology, earth sciences, geography), and also those interested in understanding citizen science (*e.g.*, science and technology studies, Internet studies, science communication, human-computer interaction). Increasingly, the term citizen science is used to refer to all of these variations on scientific collaboration with the public.

Though interest in citizen science has recently increased, the phenomenon is not new. And like most social phenomena, citizen science is more complicated than it may first appear and than is typically described in publications, popular and scientific. A few especially successful ventures have gained broad recognition, such as eBird (Sullivan, *et al.*, 2009) and the Zooniverse projects, such as GalaxyZoo (Raddick, *et al.*, 2009), which respectively focus on bird observations and (primarily) astronomical image classification. However, the citizen science landscape includes many other projects with characteristics that do not entirely match the patterns of these notable exemplars (*i.e.*, they are the Linuxes of the citizen science world). Citizen science projects also vary substantially in their complexity and the degree to which volunteers are engaged in the research (Bonney, *et al.*, 2009).

Given the common reliance on technology, citizen science could be viewed as yet another form of Internet-enabled collaboration or peer production, but that would be an oversimplification; part of the appeal of citizen science is its novelty and diversity. Much citizen science could be (and until recently was) successfully conducted in a completely "analog" environment. However, the increasing availability of information technologies, the Internet in particular, has made it easier to productively involve many more people, dramatically expanding the scale and scope of citizen science. Projects that task volunteers with data processing are fundamentally reliant on technologies that enable participation, and most projects that engage volunteers in collecting data are increasingly dependent on technologies as well. Rather than a single style of online collaboration that we can easily compare to wikis and open source software, citizen science represents a proliferation of inherently varied collaboration strategies.

And yet, citizen science shares a number of features with other open contribution systems. As a phenomenon, it overlaps with crowdsourcing and is sometimes called "crowdsourcing for science", but in fact only some citizen science projects are akin to crowdsourcing, while others are not due to scale or structure. These and other features of citizen science makes it an interesting context for studying such wide-ranging topics as participation, distributed citizenship, motivation, sociotechnical controversies, online communities, volunteerism, science learning, science identity, collective action, expertise, socialization, coordination, social and environmental justice, and many more. It has also proven to be a fruitful space for methodological development in areas such as participatory research protocols, high performance computing, dynamic task allocation, ground-referencing remote sensing data, and sophisticated statistical modeling.

The primary contribution of this research is describing and documenting the breadth and variety of citizen science. This documentation helps to clarify the boundaries of the phenomenon, and also indicates directions for future research and exploration. For this research, citizen science projects were identified as those actively engaging members of the public as volunteers in collecting or working with scientific data, at scales ranging from one volunteer to a quarter of a million individuals. The research results are summarized in Table 1 in the style of a rubric of project features, with several categories related to project design and management. The project characteristics in the second column of Table 1 emerged from analysis of survey responses to questions about the rubric categories, as described below.

Table 1: Summary of citizen science project features.

Category	Characteristic	Mountain Watch	Citizen Sort	eBird	Galaxy Zoo
Funding	1. Sustainable mix 2. Grants & membership 3. Private donations 4. Entrepreneurial	1, 2	2	1, 2, 3, 4	2
Goals	1. Resource management & conservation 2. Scientific knowledge 3. Education	1, 2, 3	2	1, 2	2
Participation activities	1. Natural history observation 2. Environmental quality monitoring 3. Content processing	1, 2	3	1	3

Data quality processes	1. Observational data 2. Measurement data 3. Replication	2, 3	3	1	3
Communication media	1. Science & data 2. Basic coordination 3. Social networking	1, 2	2	1	1, 2, 3
Rewards	1. None 2. Competitive participation 3. Volunteer appreciation	1	2	2	1
Social opportunities	1. In person 2. Distributed socialization 3. Formal education	1	2	N/A	2

Cases

We use two contrasting projects as running examples throughout this paper to provide points of comparison and illustration that help demonstrate the diversity of citizen science projects. We will briefly introduce these cases here, referring back to them to illustrate the results and discussion points. Table 1 includes the observed data for the two cases as examples to demonstrate the wide variety of project characteristics.

Mountain Watch (http://www.outdoors.org/conservation/mountainwatch/) engages hikers in monitoring alpine plant life cycles (phenology) in the White Mountains of New Hampshire. Mountain Watch was developed by the Appalachian Mountain Club to answer research questions about the effects of climate change on alpine plants, and to create an opportunity for public engagement.

Citizen Sort (http://citizen-sort.org/) is an online video game platform designed in collaboration with biologists and naturalists to support image classification for scientific research (specifically, identifying the species of animals shown in collections of photos). The games vary in the degree to which they make the scientific work the focus of the game, and were specifically created to appeal to different participant interests. Citizen Sort was developed by information scientists to study how games engage participants and the influence of design on data quality.

For comparison, we also include the well-known projects eBird (http://ebird.org) and Galaxy Zoo (http://galaxyzoo.org) in Table 1, both of which operate at substantially larger scales than most citizen science projects. While the two projects focused on collecting data (Mountain Watch and eBird) have many similarities, they also display meaningful differences, and likewise with the two projects focused on processing data (Citizen Sort and Galaxy Zoo). Less evident are the differences that are not easily encapsulated in the table, such as scale of operations.

Methods

To develop a broad understanding of the features of citizen science projects, we surveyed citizen science project leaders to learn more about the way these projects are organized, how they work, and what they produce. The sampling frame was drawn from several citizen science e-mail lists, as well as online project directories. Most of the responses came from small-to-medium sized projects. The majority are U.S.-based, with several Canadian projects reporting and two from the U.K. In addition, the majority of respondents represented projects focused on data collection, with very few projects focused on data analysis as a primary task, so this sample best represents mid-sized U.S.-based data collection projects. For more details about the survey methods, see Wiggins and Crowston (2012).

The survey questions and subsequent feature categories were derived from a theoretical model developed to direct organizational research on citizen science projects (Wiggins and Crowston, 2010). The model included elements to represent the inputs, processes, emergent states, and outputs of citizen science projects. Each category in Table 1 represents a single questionnaire item; for example, the “Participation Activities” category summarizes the responses to the question, “What are the primary types of activities for people contributing to the project? Please check all that apply.” Respondents were also provided opportunity to make free text responses, in case the provided items did not adequately describe their project. In the results, we also discuss findings from several related questions that provided additional descriptive context.

We report the responses we received from a total of 77 projects. All responses were aggregated anonymously; free text responses were paraphrased. Because most of the survey questions were presented as categorical multiple-select items, we also discuss our findings from quantitative analyses. We used factor analysis techniques to identify several sets of naturally associated characteristics that commonly co-occurred in these projects, resulting in the groupings shown in the “Characteristic” column of Table 1, and giving us several models of shared features among citizen science projects. We organize these findings into three themes: project characteristics, science practices, and participant engagement.

Project characteristics

To better understand the resources that citizen science projects can access, we asked about project demographics, such as year founded, levels of staffing and funding, and sources of funding. Responding projects were widely variable with respect to the age or duration of the project. A few projects were not yet operational; one was over 100 years old. Most of the responding projects were started in the last 10 years.

Staffing and funding sources

Fifty responding projects had between zero and over 50 paid full-time equivalent employees (FTEs), with the majority of projects employing 1-1.5 FTEs. Several noted that this allocation of staffing was spread across numerous individuals, each contributing only a small fraction of their time. Annual budgets ranged from $125 to $1,000,000 (USD or equivalent); 43 projects responded with estimated annual budgets averaging around $105,000, but the majority of project budgets fell well below $50,000 per year.

Most projects relied primarily on grants, followed by in-kind contributions and private donations. Relatively few projects were supported through sponsorships, memberships, merchandise sales, or service fees. Project organizers employed up to five different types of funding sources to meet their expenses. However, several project organizers reported that they were operating unfunded. They also felt that startup funding was easier to acquire than support for ongoing operations.

Four dominant models of funding emerged from the analysis of these responses. Projects funded by the combination of non-federal grants, private donations, and in-kind contributions pull together funding from a variety of sources. While challenging to manage, this funding model is fairly *sustainable* as reliance on a single source of funding is minimized.

Federal grants and memberships also defined a distinct funding category; this suggests strong institutional support (as might be expected of a project headquartered at a membership organization or academic institution) that can be leveraged to obtain larger sums. *Private donations* were another primary funding arrangement as well as *entrepreneurial* sources, including participant fees and merchandise sales.

Mountain Watch could be considered a sustainable project because although funding was always tight, the project organizers acquired resources from a variety of charitable foundations, organizational partners, and other sources. Citizen Sort, on the other hand, was initially supported by a grant from the U.S. National Science Foundation that enabled the substantial initial investment in developing a technology platform, but had not yet developed a source of ongoing support.

Estimating project “size”

We attempted to evaluate the relative size of project operations through several measures. Many projects require participants to register an online account; the number of people who register is one potential measure of the size or scale of a project. However, not all people who register will participate, and the rate of “conversion” from registrant to contributor varies by project. Another way to understand project size and scale is the number of contributors, or individual volunteers. A third way to understand the scope of projects is to look at the number of contributions they have received. Notably, however, each type of contribution (and even the same type of contribution, such as observations, but in different projects) involves different requirements of contributors.

The distributions of these numeric measures of project size were highly skewed. While mean values for registrations and contributors were over 1,000 individuals, there were substantially lower median values, indicating that a few very large projects were skewing distributions that were otherwise populated by small-to-medium sized projects. There are several reasons for this, such duration and geographic scale of the projects, for example, and the fact that some responding projects were not yet in operation at the time of the survey. In addition, some projects were designed for one-time contribution, while others expected participants to make multiple contributions, making it difficult to adequately contextualize these figures.

In any given year, Mountain Watch engaged between several hundred and two thousand people, but over time had accumulated data contributed by over 15,000 individuals; most participants report data for one or two (up to six) locations. Citizen Sort games have involved over 4100 people, each of whom completed approximately 80 contributions or decisions, although the participation distribution is also highly skewed, with many participants doing only a few. Citizen Sort participants spent seven seconds on average per contribution, and at the time of writing had generated 330,000 decisions. In contrast, Mountain Watch participants required several hours on average to generate each data point because participation involved hiking to alpine elevations. Such strong contrasts in these projects' data collection approaches illustrates the difficulty of comparing project based on these measures. Comparisons to other projects would also be similarly difficult unless they shared similar protocols and participation structures.

Project goals

We asked about project goals to assess their influence on other project characteristics. Recognizing that most projects have multiple goals, we provided a list of several of the most common goals mentioned in prior surveys (science, monitoring, education, conservation, stewardship, outreach, action, discovery, management, restoration), and asked organizers to rank each goal independently according to its importance to the project.

Project goals clearly shaped the design of participation tasks and related project structures. In prior work, we identified several goal-based categories of citizen science projects. The analysis reported here confirmed prior analyses that used data mining techniques (Wiggins and Crowston, 2012), and highlighted three common sets of goals.

The first set of interrelated goals included management, action, conservation, restoration, and stewardship. These goals are congruent with citizen science projects focused on *resource management and conservation*, often operated by governmental agencies or non-profits. These may be relatively short-term projects with specific goals focused on decision-making, *e.g.*, establishing damage from a marine oil spill and evaluating the effectiveness of remediation efforts. Land management and conservation projects are not necessarily driven by hypotheses or publication demands, but rather a need to establish a scientific basis for decisions.

By contrast, the goals of science and monitoring were frequently related, representing projects focused primarily on contributing to *scientific knowledge.* Although these goals are also relevant to land management and conservation, they may describe longer-term academic research projects centered on scientific studies that used monitoring data for publication rather than decision support. Within the field of ecology, monitoring is a distinctive form of research that is focused primarily on knowledge production. We would therefore expect that these projects may to yield more academic publications than the other two categories.

The final category combines the goals of education, outreach, and discovery; although discovery can be a scientific goal, comments from respondents clearly indicated that they associated this concept with participant learning. *Education* focused projects, which sometimes include formal classroom-based participation, aim to get people more involved in science, provide learning experiences, and may also involve more youth.

In our example cases, Mountain Watch shared goals from all of the categories we discussed, primarily science, conservation, and education; these goals were inherited in part from their parent organization's mission, which focused on conservation, education, and recreation. From the project organizers' standpoint, however, scientific knowledge creation was the primary goal, with conservation expected to follow from the science. Similarly, scientific knowledge production is also the main goal of Citizen Sort, which seeks to better understand how games can benefit citizen science while also providing a platform to aid researchers with data processing.

Doing science

We next consider the different ways that the projects do science by examining the types of research activities open to participants, how the projects measured contributions (in data points or other units), and the data quality strategies they used.

Participant research activities

The main research activities open to participation (shown below) in the responding projects were observation, data entry, and species identification. This reflects the fact that most of the responding projects focused on data collection, frequently for observational data. The next most common tasks were measurement, site selection and/or description, and photography. These tasks are specific to certain types of field-based participation that can also include observation.

Additional activities reported by respondents were diverse, primarily scientific tasks related to specific project requirements, and occasionally tasks related to stewardship and communication. These participant activities aligned with some of the primary goal areas discussed earlier.

- Scientific tasks
  - Posing new questions, literature reviews, paper writing, etc.
  - Videography
  - Monitoring
  - Insect rearing
  - Identifying animal tracks
  - Creating maps

- Stewardship
  - Organization and landowner coordination
  - Manual labor, habitat construction, shell recycling
- Communication
  - Communication with other participants and with scientists
  - Sharing observations and findings at meetings of related groups

The types of tasks that co-occurred in citizen science projects were fairly consistent. Our analysis generated three participation models that can be summarized as natural history observation; environmental quality monitoring; and classification and site-based data collection.

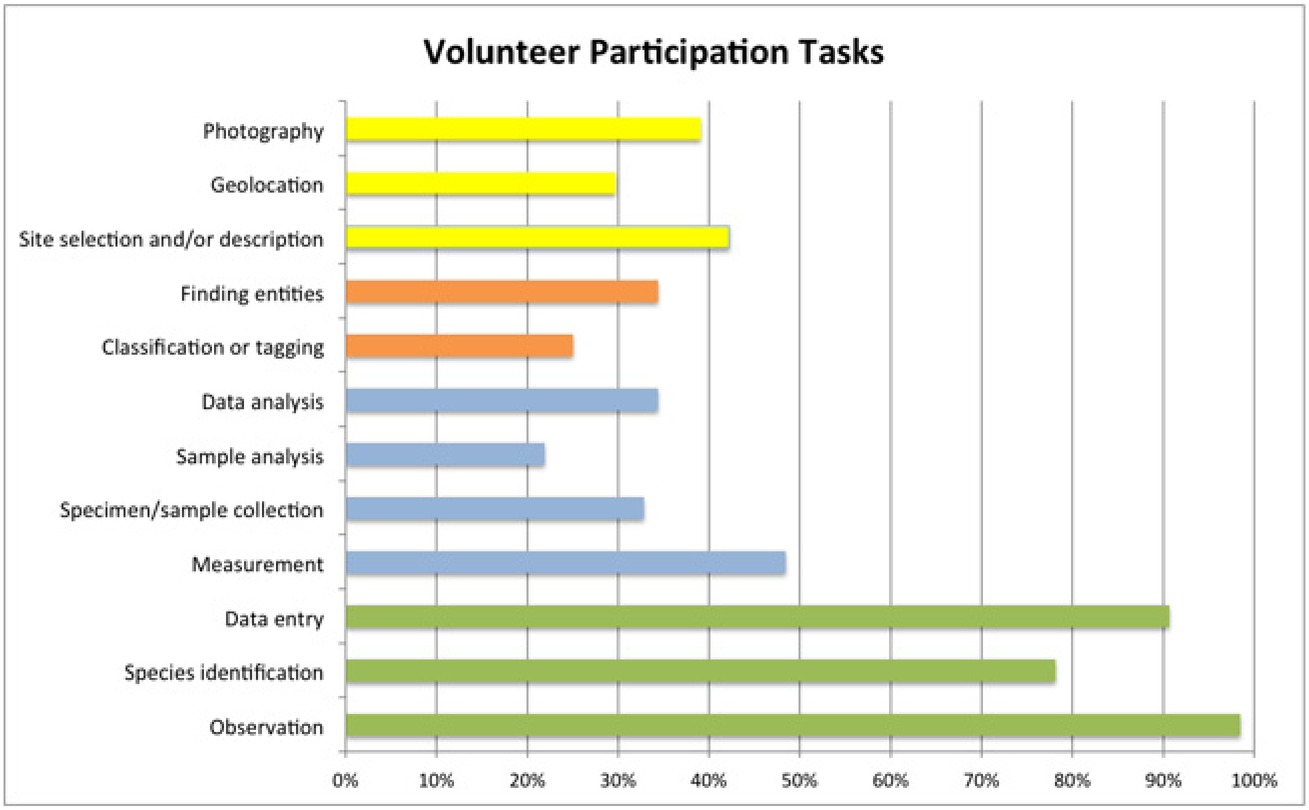

Figure 1: Volunteer participation in scientific work tasks, with observation tasks in green, measurement tasks in blue, content processing tasks in orange, and site-based observation tasks in yellow.

The *natural history observation* participation model was widespread, and included tasks of observation, species identification, and data entry. Mountain Watch is a good example of this model because participants identify and record the growth stages of specific plants. The *environmental quality monitoring* participation model involves such tasks as measurement, specimen sampling and collection, sample analysis, and data analysis. These are frequently seen in localized air and water quality monitoring projects, which often involve participants in a larger range of steps in the scientific process than in most projects.

The third model — *classification and site-based data collection* — included tasks that suggest two subsets of projects that shared the common task of detecting entities of interest, regardless of the context. The first is *content processing* projects, exemplified by the tasks of finding entities (in images) and classification of image features, as represented by the games of Citizen Sort. The second is a *site-based* participation model, a variation on the natural history observation model, wherein participants select their own research sites, geolocate the site, find targeted entities (in their surroundings), and photograph, observe, and/or measure them. Projects represented in this group include BioBlitzes and precipitation monitoring. Mountain Watch also bears some similarity to this model because in addition to monitoring plants at established locations, the data collection form accommodates making observations at any trail junction where the target species can be observed, thereby engaging participants in the tasks of site selection, geolocation, and entity finding.

Measuring contributions

We asked projects to define the unit of contribution for their project to put these measures into context. Two-thirds of respondents (41 of 60) defined the unit of contribution as observations. Other units of contribution are shown in Figure 2.

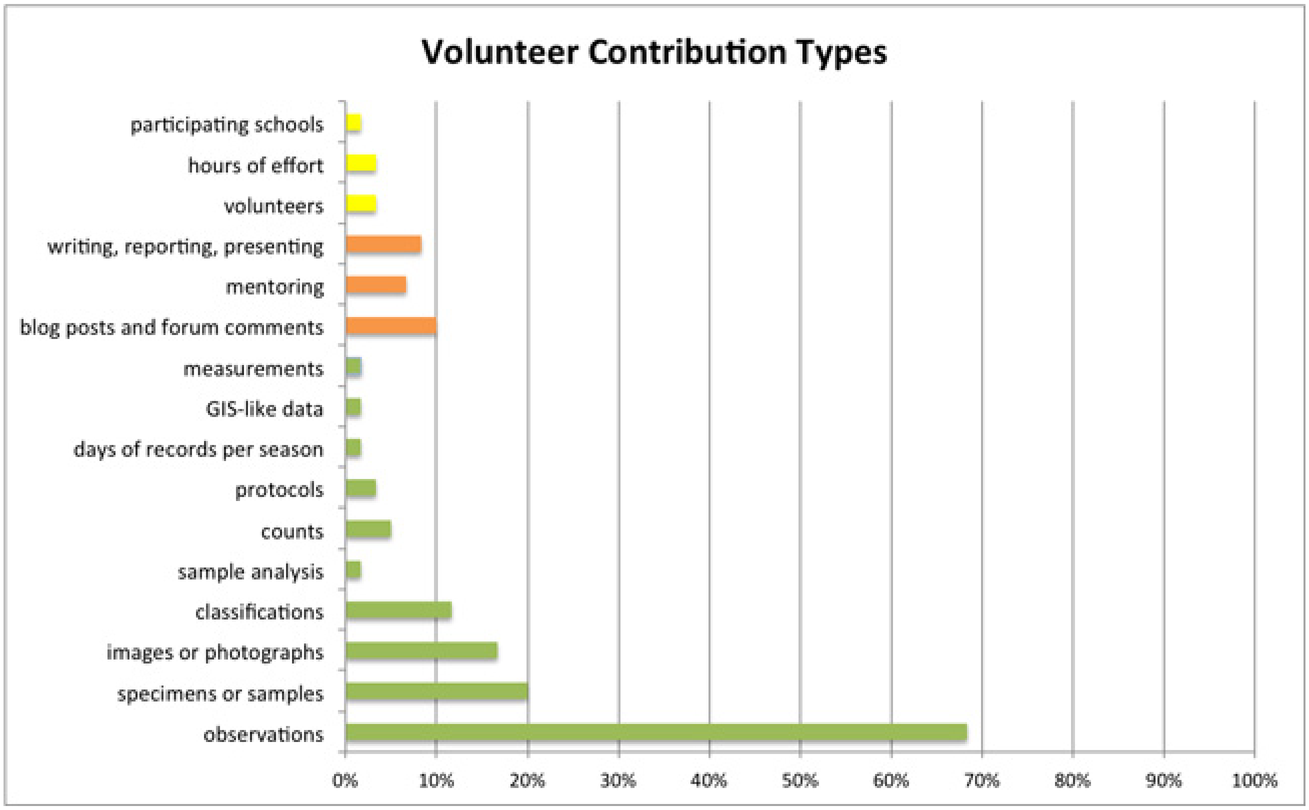

Figure 2: Types of volunteer contributions to citizen science; green represents scientific work, orange represents communication tasks, and yellow measures contributions in terms of effort or engagement.

While many projects accept contributions of observations, a third of the responding projects defined the primary form of contribution differently, making it difficult to compare the outputs of projects. In addition, several projects listed only observations as contributions, while other projects listed several forms of contribution.

Qualitative differences in the types of contributions makes comparison of project outputs even more complex, and several respondents also commented to this effect. For example, Mountain Watch is similar to many monitoring projects in which volunteers contribute observations, while Citizen Sort participants complete classification tasks. Classifications require sitting down at a computer to view and make judgments on images, but collecting alpine plant phenology observations requires hiking steep, rocky terrain to reach observation plots at trail intersections. Each type of contribution requires specific skills, resources, and effort investment that are quite different, and this degree of divergence is characteristic of the broader variation in project complexity and volunteer commitment.

Data quality

Data quality is always a matter of concern when data are to be used for scientific research. Quality concerns are even greater when numerous people are involved in data collection and entry, and more so when the contributors are largely anonymous to the organizers. We asked project organizers to indicate which methods for ensuring data quality and validation were employed in their project. Most used at least two of the methods that we listed, shown in Figure 3.

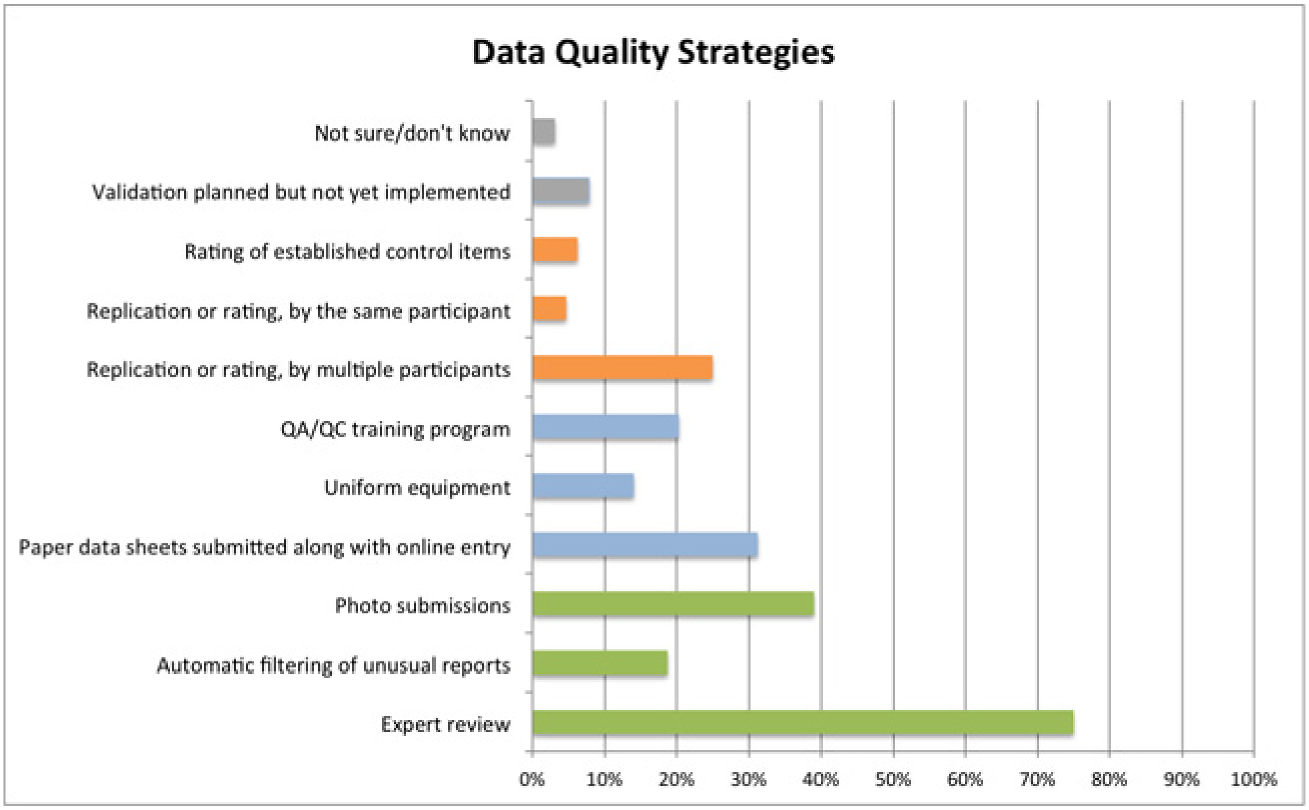

Figure 3: Data quality strategies in citizen science; green represent observational data strategies, blue represents measurement data strategies, orange represents replication-based strategies, and grey is indeterminate.

The quality of scientific data that are produced by citizen science is important to its eventual uses (Wiggins, *et al.*, 2011). Four categories of validation methods were identified in the analysis. Again, two categories were most congruent with field-based participation, while the other two were more representative of participation occurring entirely online.

Expert review, automatic filtering, and photo submissions are often used together in *observation data collection*, and sometimes at very large scales. For example, automatic filtering applied to incoming data can identify reports that merit expert review, and when available, photos can assist experts in making judgments about the validity of reported observations. These techniques are most important in situations where there is no opportunity to ground truth, replicate, or otherwise objectively verify the data, usually due to the ephemerality and contextual nature of the phenomenon being observed. Similarly, *measurement-based data collection* typically relies on paper data sheets, uniform equipment, and QA/QC training; this would likely represent most weather and environmental quality projects.

The remaining two strategies are conceptually similar because they are based on *replication of scientific results*; they are seen in both content processing and site-based tasks. First, replication of a task by multiple participants is well suited to artifact-based content processing tasks like identification, classification, and transcription that are characteristic of entirely online projects. The strategy can also apply to observations made at permanent plots in the field, with multiple participants verifying one another's work. This is a particularly convenient strategy for online projects as the number of classifications or judgments needed to validate each data point can be dynamically calculated and adjusted as needed. On the occasions when volunteers' judgments fail to converge for image analysis tasks, expert review can be applied to resolve the issue.

Second, rating of established control items fits the same project scenarios as replication, and allows project leaders to assess individual performance against a known benchmark. In the case of site-based participation tasks, this practice can sometimes be used to calibrate data from individuals' own sites by adding observations of a known, shared index plot. Individual variance on the shared site can then be used to analytically correct or weight data from the self-selected sites. These two validation methods are clear complements in plot-based field research and could also be complementary techniques for assuring data quality for image classification and transcription, but they were not typically reported in combination.

Mountain Watch organizers relied primarily on replication of observations by multiple participants, including by trained staff, at the same sites. Structured data sheets and in-person training commonly seen in measurement-based data quality management were relevant as well. Similarly, Citizen Sort games relied on replication of classifications of the same image by multiple participants, looking for a consensus to determine the answer. Secondarily, computer vision tools have been tested to verify and complement the human classifications from Citizen Sort.

Engaging participants

To better understand the design of project participation, we asked about the types of communication media used for coordination and communication, any explicit rewards for participation, and the social opportunities that are available to participants.

Communication media

Several technologies were used for communication among project organizers and between projects and their participants, with Web sites and e-mail being the most common by a large margin (see Figure 4). The next most commonly-used tools were print publications, research articles, and several types of data representations, including maps, graphs, charts, and data querying and summary tools. Projects typically used about five tools, combining both print and electronic communications. The communication media and technologies used for keeping in contact with participants and other stakeholders also reflected multiple divergent paradigms. The analysis identified three categories of communication strategies.

The first group of factors was *science-focused and data-centric*, and involved use of traditional media for dissemination, including print publications, research articles, maps, graphs and charts, animated or interactive maps, and data querying tools. The availability of these tools likely coincides with more funding or technical resources; for example, data visualizations are consistently underused by projects with little funding.

A potentially similar category emerged from the *basic coordination tools* combination of Web sites, e-mail, conference calls, maps, and data querying tools. These projects were less likely to be focused on scientific publication and may have had less funding, but provided participants with some access to data.

A third group of projects focused on supporting *social participation* with blogs, forums, photo galleries, animated or interactive maps, and social media. These projects offered a variety of ways to support the social aspects of participation, but access to data was relatively limited.

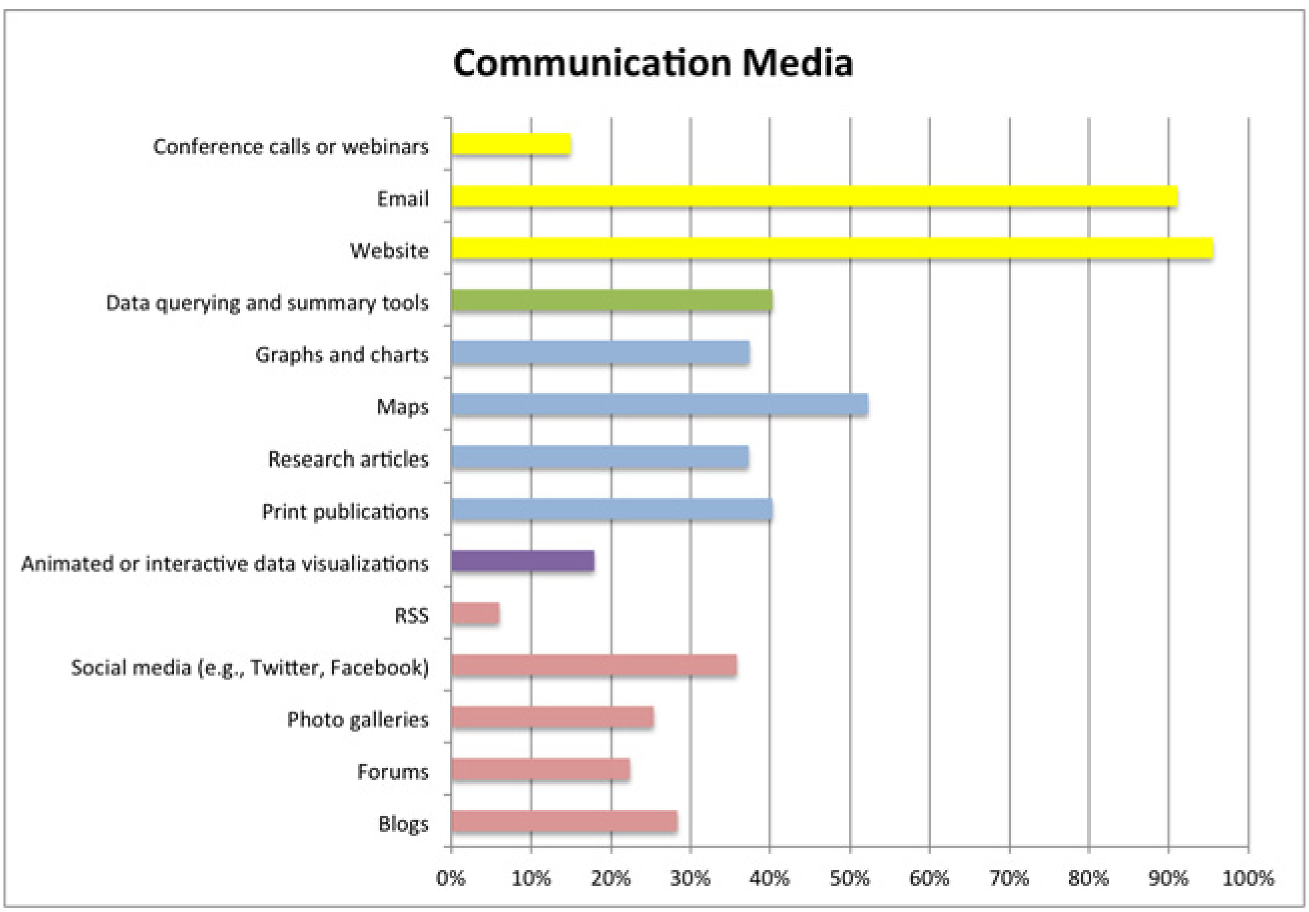

Figure 4: Communication media used in citizen science projects. Yellow represents a basic coordination tool set, blue includes research-oriented tools, and pink shows social tools.

Mountain Watch relied more heavily on in-person and on-location recruitment and communication strategies, but also leveraged email for periodic project updates and offered a Web site with an interactive map of observations. Citizen Sort depended primarily upon its websites as a point of engagement, and used few additional media for communication with participants.

Rewarding participation

The most commonly mentioned rewards for contributors were public acknowledgement and volunteer appreciation events, followed by free equipment, supplies, or training. Nearly a quarter of the projects provided no explicit rewards to participants, relying on participant interests and values to motivate contribution. This is consistent with prior work that found intrinsic motivations are a powerful driver of volunteerism, while extrinsic rewards may have potential to reduce participation (Wilson, 2000).

Rewards for participation in citizen science can take several forms; although our analysis identified three models of rewards, they can be summed into two main strategies: intangible competitive participation and traditional volunteer appreciation.

Entirely online projects in which there is clear opportunity for making new scientific discoveries offered editor or moderator status and naming privileges, but were only reported by a very small number of projects. Strategies that promoted *competitive participation*, a strong motivator for some individuals, include displaying top contributor rankings on a project Web site and providing access to personal performance metrics. Both of these intangible rewards are well suited to online participation settings.

The remaining two categories of participation rewards reflected more traditional volunteer management, dominated by tangible rewards. The first included offering *certificates, T-shirts, and role advancement*, suggesting mature projects that engage volunteers in multiple roles. Similarly, giving *promotional items, public acknowledgement, and holding volunteer appreciation events* were standard volunteer management practices for off-line participation, but tended to occur separately from other tangible rewards. These are more likely to occur in localized projects where the majority of participants can attend an appreciation event. Public acknowledgement in this context was notably different from top contributor rankings, with numerous ways this gesture of appreciation was made.

While Mountain Watch's Web site made it feasible to implement features that would promote competition, the participation structure was such that the cost of development would outweigh the benefit for volunteer management. Combined with their perennial challenges in obtaining funding, the project offered no explicit rewards, relying successfully on intrinsic motivations related to project goals. Citizen Sort, however, featured games designed specifically to prompt competitive participation by offering intangible rewards in the form of points and game advancement. For many participants, the enjoyment of playing the game may also have been its own reward.

Social opportunities

The primary venues for social interaction among participants were training sessions and group participation in project activities, which are most practical in local projects where face-to-face training and group participation are possible. The next most common opportunities for social interaction were social media and e-mail listservs, which are more practical for projects where participants are distributed over substantial geographic distances.

Like rewards for participation, the social opportunities for participants were influenced by the geographic scope and context of participation. We identified three strategies for social participation: in person, distributed socialization, and formal education.

The first strategy clearly indicates activities where participants can meet *in person*: meetings, in-person training sessions, volunteer appreciation events, group participation, and classroom participation. Accordingly, the second set of social opportunities is suited to *distributed socialization*, including large-scale and entirely online projects: email lists, blogs, social media, and conference calls. The final set of social opportunities included classroom participation and forums combined in a way that suggests citizen science in *formal education* settings, potentially involving multiple distributed classrooms and student interaction with researchers on forums.

Mountain Watch participants were able to engage in in-person training sessions at the AMC facilities, and some took part in structured group participation through alpine flower tours. Citizen Sort did not offer any tools or supports for socialization, but by nature would be constrained to primarily distributed socialization strategies, as are found in many other similar projects.

Discussion

Table 1 summarized analysis results for each category of project resources and practices we examined. Two contrasting citizen science projects provided examples of how this table can be used for comparison and contextualization. The findings also suggested relationships between the more specific results discussed earlier: funding reflected both project goals and institutional settings; project goals drove the design of tasks for volunteer participation, which dictated the types of data quality strategies that were appropriate for each project; and the means for supporting social engagement among participants were largely determined by the geospatial scale of participation.

In addition, survey responses revealed substantial challenges in evaluating the "size" of citizen science projects due to wide variations in what constitutes participation or data and the effort required to contribute on a per-person or per-datum level. For comparison purposes, measuring contributions from the public in terms of volunteer hours would have greater validity for future studies, but this statistic often goes unmonitored. Related future work could also evaluate the degree to which these types of project characteristics impact scientific, management, and learning outcomes.

By examining Mountain Watch and Citizen Sort, we saw how the two projects contrasted across most categories, such as the types of activities involved in participation, but converged on such points as their shared emphasis on producing scientific knowledge and use of replication as a strategy to ensure data quality. We also saw how project features were logically linked: Mountain Watch involved collecting natural history observations and environmental quality monitoring data, for which observational data quality management and replication techniques, respectively, were well suited and aligned with project goals oriented toward conservation and knowledge production. In addition, the place-based nature of observational data collection activities created in-person socialization opportunities for Mountain Watch participants, which were not feasible in Citizen Sort.

Given the focus in Citizen Sort on understanding game-based citizen science, the content processing activities with replication to ensure data quality were clearly aligned with the project's scientific knowledge production goals, and the use of competitive participation rewards (leaderboards, gamification features, etc.) were in keeping with current practices in other content processing projects. The analysis also highlighted room for expansion and enhancement in terms of social opportunities and project communication strategies.

The primary implication of these results is that the diversity of citizen science projects makes it important to take care in drawing comparisons or using any one project as an example or proxy for such a diverse population. The typical characterization of citizen science in mass media can be misleading, particularly with the increasingly common and indiscriminate use of the term crowdsourcing to refer to all citizen science, which leads to further confusion. For example, a crowdsourcing strategy to increase participation, such as gamification, might be entirely suitable for some projects and disastrous for others.

The survey results also highlighted the fact that the Internet plays a different role in citizen science projects depending on the mix of participation activities and socialization opportunities. In some cases, technology is a simple substitution for less efficient methods of participation, *e.g.*, postal mail to submit data collection forms. In other projects, the affordances of the Internet have massively increased the scale of participation because the cost of coordination and supporting more participants has decreased. For some projects, the Internet is fundamental for enabling participation because it enables a new way of accomplishing research tasks.

Expanded participation opportunities are most characteristic of the "online-only" projects, which typically involve content processing, intangible or competition-based incentives, replication for verification, and a combination of supporting technologies and communication media determined by the project specifics. These characteristics are not exclusive to online projects, however, and not all online projects fit this stereotype; even entirely new online-only projects carry over features of historic citizen science projects that worked well in earlier efforts. By surfacing the variations in several other categories of project features and processes, this research can provide a framework for systematically exploring the potential design space for citizen science.

For individuals who are curious about citizen science, this work revealed some of the complexity behind what might initially appear to be a simple phenomenon. Like so many other open contribution systems, a successful citizen science project must balance interrelated resources and requirements in order to achieve its goals, and there is no single "right" model or style of participation. Instead, a variety of forms have evolved to meet different purposes and there are many ways that project organizers can support the interests of both researchers and participants. The type of scientific work and geographic scale of participation, however, strongly shape the strategies that a project uses to meet its goals.

For media representatives and researchers who study citizen science or related phenomena, the risk of comparing apples to oranges is prevalent. Our results speak to the importance of carefully grounding research and media reports in the most relevant context of practice, and the value of using theoretically motivated criteria for case and sample selection. This research provides foundational work for theorizing the impacts of these factors on how citizen science projects work and the outputs they produce. It offers grounding for comparison to other online collaboration formats. As the transferability of these concepts and characteristics is established, we can develop a more empirical and analytical understanding of their similarities and differences, and be better positioned to support knowledge exchange across communities of practice.

The limitations of this study are inherent in the survey methods and status of citizen science as a fast-growing and emergent phenomenon. They include the challenges with acquiring a representative sample of citizen science projects and the pace at which growth and change within the field of practice can make findings outdated. In the future, current practitioner-led efforts to make this information transparent as a part of establishing projects' legitimacy and rigor should further reduce these constraints.

Conclusion

We saw greater diversity among citizen science projects responding to our survey than is typically represented in stories about citizen science that appear in news media and popular science outlets. Although our sample included primarily observational projects in research areas related to ecology, there was an impressive range of types of participation, social opportunities, technologies in use, approaches to data validation, ways to measure contribution, and project goals.

Through this description of the increasingly popular phenomenon of citizen science, we hope to help others appreciate the diversity and complexity of citizen science. By focusing attention on a more detailed set of project features, we provide a stronger foundation for making sensible comparisons between projects. The multifaceted perspective on citizen science project characteristics and practices we outline here can support a more holistic understanding of the phenomenon, yielding more targeted participation, investment, research, and strategic management.

About the authors

Andrea Wiggins is an Assistant Professor in the College of Information Studies at the University of Maryland College Park. She studies citizen science and related open knowledge and collaboration systems.
E-mail: wiggins [at] umd [dot] edu

Kevin Crowston is a Distinguished Professor of Information Science at the Syracuse University School of Information Studies. His research interests include the ways information and communications technologies enable new kinds of organizing. Specific topics of interest include free/libre open source software development and citizen science.
E-mail: crowston [at] syr [dot] edu

Acknowledgements

This work was supported in part by U.S. National Science Foundation grants 09-43049 and 08-3094.

References

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*, volume 343, number 6178 (28 March), pp. 1,436–1,437.
doi: http://dx.doi.org/10.1126/science.1251554, accessed 18 December 2014.

Rick Bonney, Heidi Ballard, Rebecca Jordan, Ellen McCallie, Tina Phillips, Jennifer Shirk, and Candie Wilderman, 2009. *Public participation in scientific research: Defining the field and asssessing its potential for informal science education. A CAISE inquiry group report*. Washington, D.C.: Center for Advancement of Informal Science Education (CAISE), at http://informalscience.org/images/research/PublicParticipationinScientificResearch.pdf, accessed 18 December 2014.

Alan Irwin, 1995. *Citizen science: A study of people, expertise, and sustainable development*. New York: Routledge.

M. Jordan Raddick, Georgia Bracey, Pamela L. Gay, Chris J. Lintott, Phil Murray, Kevin Schawinski, Alexander S. Szalay, and Jan Vandenberg, 2010. "Galaxy Zoo: Exploring the motivations of citizen science volunteers," *Astronomy Education Review*, volume 9, number 1.
doi: http://dx.doi.org/10.3847/AER2009036, accessed 18 December 2014.

Brian L. Sullivan, Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and Steve Kelling, 2009. "eBird: A citizen-based bird observation network in the biological sciences," *Biological Conservation*, volume 142, number 10, pp. 2,282–2,292.
doi: http://dx.doi.org/10.1016/j.biocon.2009.05.006, accessed 18 December 2014.

Abraham Wandersman, 2003. "Community science: Bridging the gap between science and practice with community-centered models," *American Journal of Community Psychiatry*, volume 31, number 3, pp. 227–242.
doi: http://dx.doi.org/10.1023/A:1023954503247, accessed 18 December 2014.

Andrea Wiggins and Kevin Crowston, 2012. "Goals and tasks: Two typologies of citizen science projects," *Proceedings of the 45th Annual Hawaii International Conference on Systems Science*, at http://crowston.syr.edu/system/files/hicss-45-final.pdf, accessed 18 December 2014.

Andrea Wiggins and Kevin Crowston, 2010. "Developing a conceptual model of virtual organizations for citizen science," *International Journal of Organisational Design and Engineering*, volume 1, numbers 1-2, pp. 148–162, and at http://crowston.syr.edu/content/developing-conceptual-model-virtual-organizations-citizen-science, accessed 18 December 2014.
doi: http://dx.doi.org/10.1504/IJODE.2010.035191, accessed 18 December 2014.

Andrea Wiggins, Greg Newman, Robert D. Stevenson, and Kevin Crowston, 2011. "Mechanisms for data quality and validation in citizen science," *ESCIENCEW '11: Proceedings of the 2011 IEEE Seventh International Conference on E-science Workshops*, pp. 14–19.
doi: http://dx.doi.org/10.1109/eScienceW.2011.27, accessed 18 December 2014.

John Wilson, 2000. "Volunteering," *Annual Review of Sociology*, volume 26, pp. 215–240.
doi: http://dx.doi.org/10.1146/annurev.soc.26.1.215, accessed 18 December 2014.

Editorial history

Received 2 September 2014; revised 14 December 2014; accepted 20 December 2014.

Surveying the citizen science landscape
by Andrea Wiggins and Kevin Crowston.
First Monday, Volume 20, Number 1 - 5 January 2015
http://firstmonday.org/ojs/index.php/fm/rt/printerFriendly/5520/4194
doi: http://dx.doi.org/10.5210/fm.v20i1.5520