Describing Public Participation in Scientific Research

Andrea Wiggins
Syracuse University
337 Hinds Hall
Syracuse, NY 13244
awiggins@syr.edu

Kevin Crowston
Syracuse University
348 Hinds Hall
Syracuse, NY 13244
crowston@syr.edu

ABSTRACT

We report the results of a descriptive survey of citizen science projects, a form of scientific collaboration engaging members of the public with professional researchers. This phenomenon has seen explosive growth in recent years and is garnering interest from a broadening variety of research domains. However, the lack of adequate description of this diverse population hinders useful research. To address this gap, we conducted a survey of citizen science projects. We present a description of the phenomenon to establish a basis for sampling and evaluation of research on citizen science, including details on project resources, participation, technologies, goals, and outcomes. We then reflect on several points of potential development, including technologies to support participation, potential for expanding engagement, and data policies. The diverse organizational and functional arrangements in citizen science projects suggest a variety of areas for future research.

Keywords

citizen science, participation, organizational design

1. INTRODUCTION

Citizen science is a form of research collaboration in which professional scientists engage with members of the public in order to accomplish scientific research [2]. This is a form of distributed scientific collaboration, but instead of engaging only professional scientists and technicians, as in the majority of prior research [16], citizen science also involves non-professionals in the research process. This phenomenon has exploded in recent years, due in part to the influence of the Internet, and as a result, researchers from a variety of research domains are becoming interested in this phenomenon as a subject of study. Citizen science is a promising area for research, particularly for topics such as motivation, collaboration, coordination, social computing, and collective action [11].

iConference 2012 Toronto, Ontario, Canada

In other contributory contexts, such as open source software development, research has fixated on a few projects that are meaningfully different from the bulk of the population, e.g. Linux and Mozilla [4], partly due to their overwhelming success and resultant visibility. Emerging research on citizen science has the potential to follow this pattern and generate a skewed view of the phenomenon; well-known projects such as eBird [15, 18] and Galaxy Zoo [12], for example, are well known but not representative of the vast majority of citizen science. Several typologies of citizen science demonstrate that there are a wide variety of approaches to structuring participation [1, 5, 10, 6], but there is relatively little reflection on other descriptive characteristics of these projects. As in open source software, we expect that there is no one "typical" citizen science project, so the goal of this paper is to support the development of systematic knowledge about citizen science projects.

We provide factual description of a phenomenon of increasing interest that can aid in the development and evaluation of future research on citizen science. We considered a variety of dimensions of projects to understand the diversity of the phenomenon, and conducted a survey of citizen science projects to learn about their organizational characteristics and aspects of participation. Our results demonstrate that there are currently many ways that this approach to science is being applied in the field.

2. BACKGROUND

Research across disciplines conducted following the citizen science model typically focuses on either data collection, such as eBird and Monarch Watch, or data reduction, such as Stardust@Home and Galaxy Zoo. Monitoring and observation oriented projects are centered on collecting data from contributors, often at larger temporal and geographic scales than are otherwise possible, while data reduction projects leverage human perceptual capacities and problem-solving skills to accomplish analysis tasks that are currently too difficult for computers [3]. The design of research within each of these types is highly variable, a characteristic that complicates meaningful comparison across projects.

These projects are increasingly enabled by and take advantage of information and communication technologies to advance scientific research [13]. They are often considered a type of crowdsourcing, a term referring to a set of distributed production models that make an open call for contributions from a large, undefined network of people [8, 17]. This conceptualization may only be applicable to a limited portion of the population of projects, however, leading us to seek a better understanding of the diversity of the phenomenon.

3. METHODS

In this section, we describe the survey instrument, sample, and response rate for a survey of citizen science projects.

3.1 Survey Instrument

A survey instrument was composed to directly elicit selected descriptive characteristics of projects. It was presented as a two-part questionnaire: first, a brief project profile and second, a separate, lengthier survey.

The first portion of the questionnaire was a project profile, allowing projects to opt-in for listing on several cooperating websites that provide listings of citizen science projects, and update existing project profiles based on data provided with the sampling frame or create a new project profile. The project profile included 23 items that would be considered useful by potential participants; the second portion of the questionnaire was the project survey, which asked for additional details in several categories. The full survey included 57 items, with free-response spaces for each structured item. There were no required fields, so each item had a variable response rate. The items covered several categories, but those reported in this paper focused on data validation methods.

3.2 Sample

The sampling frame was composed of projects listed on Cornell Lab of Ornithology's citizen science email list and in the now-defunct Canadian Citizen Science Network. These are the most comprehensive sources of contacts for North American citizen science projects. Approximately 60 additional contacts were manually mined from the online community directory at http://www.scienceforcitizens.net to extend the disciplinary diversity of the sample.

These sources provided a combined set of approximately 840 contacts after removing duplicates and bad addresses. These contacts are individuals who had self-identified as responsible for or interested in the management of citizen science projects. Approximately 280 projects were identified in this process, and another 560 individuals who may be connected with additional projects were also invited to participate.

3.3 Response Rate

In response to approximately 840 emailed requests for participation, 128 project profiles were created or updated. 73 surveys were initiated and 63 fully completed, for a participation rate of 15% and a response rate of approximately 8%. The surveys and profiles were combined for analysis. The response rate is low, though not atypical for such a survey. However, it should be noted that contacts were asked to report on projects, and the number of projects is smaller than the number of contacts, meaning that the response rate for projects (our unit of analysis) is better than it appears. As noted above, we were able to identify approximately 280 projects, which would lead to a response rate of about 22% rather than 8%; the actual response rate lies somewhere in between these two figures.

Most of the responses came from small-to-medium sized projects, based in the United States, with several Canadian projects reporting, and three from the UK; a handful of projects are organized by research teams that span international boundaries. The sample best represents medium-sized North American citizen science projects, and nearly all responding projects are of the monitoring and observation types. The sample is also subject to self-selection bias, such that projects interested in attracting more participants through a directory listing were more likely to respond than those that may selectively engage contributors, for example, based on known subject expertise. However, despite these limitations, we believe that the resulting sample is generally representative of the population of citizen science projects. Independent review of the response pool characteristics by staff at the Cornell Lab of Ornithology, who have conducted numerous similar surveys, suggested that the responses provide a fairly representative sample of the larger community.

4. RESULTS

Several results from the survey are reported here, spanning the broad categories of inputs, processes, and outputs.

4.1 Inputs

The inputs we consider here are project resources, such as staffing, budget, and funding sources, and tools and technologies, including future plans.

4.1.1 Project Resources

To better understand the resources that projects are able to devote to various aspects of development, implementation, and improvement, we asked about levels of staffing, funding, and sources of funding.

Project Staffing & Budget.

50 responding projects had between zero and over 50 paid full-time equivalent employees (FTEs). The frequency distribution is a near normal distribution, 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, with an average of $104,882 but with a median of $35,000 and a mode of $20,000. Budgets had a multi-modal distribution, with the majority of project budgets falling below $50,000 per year.

52 projects included the year founded in their responses. Responding projects were widely variable with respect to the age or duration of the project. A few projects were not yet operational, and one was over 100 years old. The average age of currently operational projects is 13 years, while the median is 9 years and the mode is 2 year: most of the responding projects were started in the last 10 years.

Funding Sources.

We asked projects about the types of funding sources that they use to support their projects. Most projects rely primarily on federal or other grants, followed by in-kind contributions and private donations. Relatively few projects are funded by sponsorships, memberships, merchandise sales, or service fees. For 64 responding project, the reported sources of funding were:

- Participant fees: 11
- Federal grants: 34
- Other grants: 34
- Private donations: 23

- Sponsorship: 3
- Service fees: 1
- Memberships: 4
- Merchandise sales: 4
- In-kind contributions: 31
- Not sure/don't know: 3

Additional and more specific funding sources named included state appropriations, private foundations, and government agencies. Projects employed up to five different funding sources to meet their expenses; however, several projects reported that they currently operate unfunded or by leveraging other revenue streams for related projects, with comments suggesting that startup funding is easier to acquire than support for ongoing operations.

4.1.2 Tools and Technologies

To learn about the range of information and communication technologies supporting citizen science, we asked about tools for communication with project participants, as well as technology plans, both for the immediate future and long-term interests.

Communication Tools.

Several technologies are used for communication among project organizers and between project and participants, with websites and email being the most common by a large margin. 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 used between one and twelve tools to communicate, with an average of 5.4 and a median and mode of 5. Print publications are not substituted by electronic publications, but are complementary. The responding 67 projects used the following tools for communication:

- Website: 64
- RSS: 4
- Email: 61
- Conference calls or webinars: 10
- Print publications: 27
- Research articles: 25
- Blogs: 19
- Forums: 15
- Photo galleries: 17
- Maps: 35
- Graphs and charts: 25
- Animated or interactive data visualizations: 12
- Data querying and summary tools: 27
- Social media (e.g., Twitter, Facebook): 24

Other communication tools that were reported included one-on-one phone calls, e-newsletters, classroom presentations, posters in facilities, private wikis for teams, and presentations at scientific meetings.

Technology Plans.

We asked what new technologies, besides communication tools, are planned for implementation in the next few years. Some of these plans included:

- New or additional data analysis tools (6)
- Smartphone/mobile apps (8)
- New or improved websites (9)
- Video for training (3)
- Online data entry
- Facebook accounts (2)
- Mapping capabilities (2)
- Database improvements (2)
- Support materials (2)
- None (4)

While these responses represent the more common technology plans, there were numerous variations specific to different projects needs and stages of development. For example, new or improved websites were a priority for new projects with no website as yet, mature projects with outdated websites, and projects that seek to enhance participation with new tools and visualizations.

Future Technologies.

We also asked what new technologies or improvements to current technologies projects would like to implement in the future, beyond what is currently planned. Some of the desired tools included:

- Mobile applications for data entry (7)
- Real-time, interactive, and dynamic data visualizations (4)
- Animated and interactive maps (3)
- Use of GPS units by most participants
- Decision-support recommendations for management activities
- Google Earth/3D technology
- Complete revision of project database, website, and data entry application
- Web-based analysis tools for digital photos

Related comments on technologies were similarly diverse. Several projects commented that they would like to add mobile technologies but do not have adequate funding. One project mentioned that it would be helpful to have open-source mobile apps for standard point-count protocols, while another envisioned a "map server" for environmental monitoring projects to permit personalized maps for any project. An entomology-oriented project noted that cell phones cannot record certain insect calls because the devices are not sensitive to a wide enough audio range. Multiple projects mentioned that they have difficulty acquiring funding to develop tools and technologies, indicating this is a substantial hurdle. This is a matter of particular concern for a number of projects, as these technologies are critical infrastructure for supporting project goals. We also observed that the projects with the most funding also had the most sophisticated technologies, as well as the highest reported participation rates, suggesting that the investment in tools provides substantial return on investment.

4.2 Processes

We examined several aspects of project processes, including the type of contributions and participation details. To better understand the design of project participation, we asked respondents about the types of activities that project participants engage in, any explicit rewards they receive for participation, and the social opportunities that may be available to participants.

4.2.1 Participation Activity Types

The main participation activities for responding projects were observation, data entry, and species identification. This reflects the fact that most of the responding projects focus on data collection, frequently for ecological observations. The next most common tasks were measurement, site selection and/or description, and photography; these tasks are more specific to certain types of participation protocols. For 64 responding projects, the participation activities identified included:

- Observation: 63
- Species identification: 50
- Classification or tagging: 16
- Data entry: 58
- Finding entities (in images or natural habitats): 22
- Measurement: 31
- Specimen/sample collection: 21
- Sample analysis: 14
- Site selection and/or description: 27
- Geolocation: 19
- Photography: 25
- Data analysis: 22

Additional activities that were reported focus mainly on scientific tasks related to specific project requirements. These participant activities included:

- Posing new questions, lit reviews, paper writing, etc.
- Videography
- Monitoring
- Insect rearing
- Organization and landowner coordination
- Identifying animal tracks
- Manual labor, habitat construction, shell recycling
- Creating maps
- Communication with other participants and scientists
- Sharing observations and findings at meetings of related groups

The diversity of these participation activities clearly demonstrates that the nature of the contribution that participants make can varies substantially from one project to the next.

4.2.2 Project Contributions

As one indicator of the relative size of projects, we asked for several measures of participation, such as number of project contributions and contributors. We also asked projects to define the unit of contribution for their project to put these measures into context. Among 60 respondents to this item, 41 defined the unit of contribution as observations. Other units of contribution included:

- specimens or samples (12)
- images or photographs (10)
- classifications (7)
- blog posts and forum comments (6)
- mentoring (4)
- writing, reporting, presenting (5)
- counts (3)
- protocols (2)
- volunteers (2)
- hours of effort (2)
- days of records per season
- GIS-like data
- measurements
- participating schools
- sample analysis

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 the types of 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 made comments to this effect.

Other comments regarding contributions included several notes that contributions are difficult to quantify or assess due to program breadth, longevity, or lack of funding. Notably, however, asking about the unit of contribution quite clearly demonstrated that each type of contribution (and even the same type of contribution, but in different projects) involves different requirements of contributors. While we collected quantitative data on the numbers of contributors and contributions, for all of the above-mentioned reasons, we do not summarize the numeric responses, as the results are difficult to interpret at best and potentially misleading at worst.

For example, some projects are designed for one-time contribution, and others expect participants to make multiple contributions, which substantially influences the interpretation of relative numbers of contributors and contributions. While a one-to-one ratio may appear to be poor performance compared to projects in which participants contribute tens, hundreds, or thousands of observations, it is the expected level of contribution in others and can represent a successful outcome. Likewise, some projects are seasonally constrained, geographically localized in scope, or selective about participants due to expertise requirements, and therefore aim to engage only a few dozen individuals, while others are global in scale and have over a quarter million active contributors.

4.2.3 Rewards to Contributors

In most contexts, voluntary work is motivated by a range of incentives; we inquired about explicit rewards that may be relevant to motivation. The most common explicit rewards for contributors were public acknowledgement and volunteer appreciation events, followed by free equipment, supplies, or training. A substantial number of projects (15) provide no explicit rewards to participants, relying on participant interests and values. 65 projects responded with the following rewards:

- None: 15
- Free equipment/supplies/training: 17
- Certificate: 8
- T-shirts: 3
- Promotional items, i.e. stickers, pins, keychains: 9
- Top contributor listings: 6
- Personal performance ratings: 5
- Public acknowledgement: 36
- Role advancement: 4
- Editor/moderator privileges: 1
- Naming privileges: 1
- Co-authorship privileges: 5

- Volunteer appreciation events: 20

Additional rewards reported reflect a range of organizational resources beyond basic funding sources. These forms of acknowledgement include special experiences or privileges as well as more standard intangible and tangible rewards. Among these were:

- Travel funds to attend workshops
- Access to restricted areas
- Personal thank-you or feedback (3)
- Newsletter (2)
- Education, Knowledge (3)
- Recognition awards, e.g. Volunteer of the Year (2)
- Letter of acknowledgement
- Hats with project logo
- Public display of photos
- Social rewards
- Plants for creating a butterfly habitat
- Top contributors receive mileage reimbursement when funds allow
- Drawings, contests, giveaways (2)
- Reference materials and CDs with species sounds
- Borrowing privileges for equipment
- 128-page annual project summary publication

4.3 Social Opportunities

Opportunities for social interaction can support recruitment, retention, and skill development. The primary venues for social interaction among participants are training sessions and group participation in project activities, which are most practical in projects where face-to-face training and group participation are possible. The next most common opportunities for social interaction were social media and email listservs, which are more practical for projects where participants are distributed over substantial geographic distances. 64 projects listed these opportunities for social interaction:

- None: 8
- Forums: 14
- Email listservs: 21
- Blogging and/or commenting on blogs: 14
- Social media (e.g., Twitter, Facebook, etc.): 25
- Conference calls or webinars: 4
- Meetings: 19
- Training sessions: 34
- Volunteer appreciation events: 15
- Group participation in project activities: 30
- Classroom participation: 16

Like participant rewards, the other opportunities that projects listed for social interaction reflect different resources and project designs. These included:

- Online, synchronous chat sessions with special guests
- Field activities with community members
- Videoconferencing
- Sponsored field trips
- Workshops and seminars
- Regional group participation efforts
- Full social network toolkit
- National data collection events with optional group data aggregation gatherings

Although there are some common trends among these responses related to the participation, the variety of activities, rewards, and social opportunities identified by the survey responses shows that the nature of the participation experience, even in the more common observation-oriented projects, is far from identical from one project to the next.

4.4 Outputs

We also considered the project outcomes in both intended and actual states, and data policies, as data is a key output for most projects.

4.4.1 Project Outcomes

To better understand the alignment of project plans with outcomes, we asked projects to report on types of intended and actual project outcomes. We also asked projects to rank goal areas for these outcomes, which we recently reported in a separate paper [19].

Intended Project Outcomes.

We asked projects about the outcomes they intended to produce. The responses of 64 projects show the most common are data sets, individual learning, data analysis, and academic publications and presentations. The range of intended outcomes includes:

- Data sets: 56
- Data analysis: 46
- Academic publications and presentations: 43
- Technical reports: 25
- New discoveries: 31
- New research methods: 17
- New inquiry: 21
- Policy changes: 21
- Community action: 38
- Environmental restoration: 23
- Individual learning: 47

Additional intended outcomes included stewardship, distribution, public presentation, informed land use decisions, individual behavioral change, public understanding, and use of inquiry methods in the classroom.

Actual Project Outcomes.

The results for this question require consideration of two details for interpretation of the results; one is that there were 9 fewer responses to actual outcomes than intended outcomes, and the second is that some projects were founded recently enough that they were not yet able to report on outcomes, or on the full outcomes of the project. 55 projects responded with these outcomes:

- Data sets: 48
- Data analysis: 40
- Academic publications and presentations: 33
- Technical reports: 21
- New discoveries: 22
- New research methods: 11
- New inquiry: 14
- Policy changes: 11
- Community action: 26
- Environmental restoration: 17
- Individual learning: 42

It is likely that in time, the frequencies of the outcomes reported above would appear in larger numbers. Nonetheless, respondents reported relatively high levels of success in achieving desired outcomes. In particular, many projects reported that the project had resulted in individual learning, which is frequently a project goal. At the opposite end of the spectrum, little over half of the projects interested in policy change claimed this as an actual outcome.

Additional outcomes included teachers using inquiry in their classrooms; awarding of highly competitive telescope time for follow-up on interesting discoveries; support to land use managers and decision makers; and public presentations. Naturally, as some projects are more recently founded, additional responses noted that the project was not yet in a stage that permits full evaluation of project outcomes.

4.4.2 Data Policies

As data is a key outcome for most projects, we were interested in what data policies are adopted. Data sharing and ownership are often negotiated during project founding in most academic research, but it is not clear whether collaboration with the public changes the expectations surrounding data policies. Accordingly, we asked projects about their data sharing practices and ownership policies, finding that overall, the majority of projects produce scientific data as a public good.

Data Sharing.

Public involvement in scientific research may create slightly different data sharing conditions than most research in the disciplinary communities for this research. We asked projects to specify with whom they currently share data, and 67 projects responded:

- Sharing with contributors: 45
- Sharing with project-affiliated researchers: 41
- Sharing with a research network or data archive: 35
- Sharing with the general public: 46
- Sharing is planned but not yet in place: 11
- Not sure/don't know: 2
- No sharing: 0

Some of these categories overlap, which may affect respondents' interpretations; for example, sharing with the general public means that contributors and researchers would also have access. For those without sharing in place (11), 5 did not elaborate; for the other 6, expectations are to share with contributors (4), affiliated researchers (2), network/archive (2), public (3). We note that in some cases, projects not sharing with affiliated researchers may not actually have affiliated researchers who will use the data. For projects where data sharing is planned but not yet in place, low staffing seems likely to be the strongest reason for not having data sharing in place. Most have 1.5 or fewer FTEs (as low as 0.05); budgets range from \$10,000 to \$100,000 annually, and the projects are up to 14 years old (median 3 years) but there is no clear association with project age or budget and lack of data sharing.

Several projects commented about other parties with which data are shared; such as sharing with researchers who are not project-affiliated but also not through a network or archive; sharing with local, state and federal agencies; and data sharing managed by a parent project.

Data Ownership.

When inquiring about data ownership, we asked projects to specify whether their policy fit one of several multiple choice options, or to describe their policy separately. 60 projects identified their data ownership policy among the following:

- No policy: 11
- Currently developing policy: 4
- Researchers own the data: 15
- Project contributors own the data: 13
- Third party owns the data: 1
- Public owns the data: 23
- Not sure/don't know: 6

As multiple selections were an option for this item, a few projects specified combinations of ownership, most commonly researchers and contributors (3), or other combinations involving the public (5), which effectively include researchers and contributors as well. Approximately 28% of these projects had no data ownership policies or the individual responding was uncertain, suggesting that best practices are yet to be established with respect to data policies in citizen science.

An additional 7 projects elaborated on different project ownership arrangements:

- "A parent project owns the data"
- "This project is officially registered as a literary work; A. Person is the author and a federation of Interest Group societies owns the monetary rights."
- "Data is owned by the Ministry of the Environment (Canada)"
- "We encourage Host Service users to share data freely, but data ownership depends on individual projects using Host Service"
- "Participants control visibility of data to people other than researchers"
- "Data is not copyrightable. Images, writing, etc. contributed to site is licensed with Creative Commons Attribution license, with attribution to the project username associated with contribution."
- "Data is made available to the NBN - UK National Biodiversity Network Data Infrastructure"

Further comments noted that state agencies or provincial governments were occasionally the project data owners, and therefore consider it publicly owned. In these cases, data sharing arrangements were not always as straightforward as the ownership policy might suggest. For example, in one case, data ownership by a state agency implied data stewardship, and although the data are considered publicly owned as well, the state agency works out arrangements to share data with researchers, demonstrating that public ownership does not equate to public access to data. Other cases were more collaborative, such as a situation in which "data is housed at State Department of Fish and Game but is shared with the State Natural Heritage Program and is uploaded annually to NatureServe."

5. DISCUSSION

We discuss observations on these results on four topics: the types of technologies which may be useful for citizen science projects, possibilities for greater citizen involvement, best practices in data policies, and areas for future research.

First, the list of types of project contributions suggests that other kinds of technologies that might be useful in supporting successful participation. Some are already under development, e.g., systems for data entry. In particular, however, species identification is a task common to such a large proportion of projects surveyed that this suggests a potential area of opportunity for developing technologies to support this task. We are currently working on a system for this purpose in another project, and we note that there are an increasing variety of mobile applications and electronic field guides that can serve this function [7, 9, 14], but that it is not yet clear how well these tools are addressing participants' needs in addition to fulfilling the functional goals set out by project organizers.

Second, the types of project contributions mentioned by survey respondents also suggests possibilities for greater citizen involvement. For example, writing papers or technical reports is a goal for most projects; it is interesting to consider if there is a role for participant engagement in that activity. Potential benefits would include increased learning about processes of science, and further appreciation of the significance of the work when the full scientific knowledge production process is made more visible. As an exemplar of this possibility, the Zooniverse projects, which include Galaxy Zoo, maintain a blog that includes discussion of papers; in the past, the Zooniverse organizers have included some contributors as authors on published papers, acknowledging the value of their work. The Open Dinosaur project has likewise maintained an open process throughout much of their citizen science research on dinosaur limb bone measurements. On the other hand, however, engaging inexperienced contributors in writing an academic paper may represent too much overhead for most projects, given their limited resources. It is also likely that the majority of project contributors are disinterested in writing papers, as their primary motive for participation is leisure [18].

Third, the relatively inconsistent level of data policy development is surprising, given that most of the projects included data as a project output. This may be reflective of the broader situation in science more generally: data policies are still nascent. Nonetheless, the unique source of citizen science data does seem to create specific issues around data that need to be addressed, with questions around who owns contributions, how to share data with contributors, and how to ensure data quality and access. We noted that the largest projects (by any measure) all had data ownership policies, and it's possible that smaller projects may not be fully aware of data management, ownership, and intellectual property concerns. These issues may also be more carefully monitored in projects run by government agencies, as related policies such as the US Paperwork Reduction Act create a different set of constraints for enlisting public participation.

Scientific collaborations with thousands of contributors have previously occurred most often in scientific domains that are heavily reliant on major infrastructure (e.g., the Large Hadron Collider), for which conventions have been established because all parties have a vested interest in receiving credit for their work. In citizen science projects, by comparison, project organizers are more likely to come from domains where small team collaboration is most common, and many contributors may be largely indifferent to this aspect of research. This does not eliminate the potential legal and ethical considerations of scientific collaboration, but it is not clear what consequences may arise for projects that do not have established policies regarding data ownership. It may therefore be useful to share best practices in this area to support development of improved data policies across citizen science projects.

Finally, the results of our survey point to numerous potential topics for future research; we briefly outline just a few initial ideas here. Economic analysis of the value of technology funding for generating project outputs (e.g. datasets) could provide justification for investment in cyberinfrastructure to support public participation in scientific research. Research into the strategies for technology use in citizen science, and particularly analysis comparing these projects to other contexts of distributed voluntary work, could help generate new guidelines for technologies and systems to sustainable project development. Further examination of participation processes and contributions would help to identify best practices in participation task design, while also suggesting areas where systems can be developed to further support project outcomes. Research connecting rewards offered to contributors with participant motivations could inform the design of incentives to improve participant recruitment and retention. A deeper understanding of the role of social interaction in supporting project outcomes (e.g., through supporting improved recruitment, retention, and skill development) would help projects determine what level of emphasis to place on the development of opportunities for socialization, whether through technology-mediated means or by designing tasks and tools that explicitly support group participation. Connecting intended and actual outcomes through inquiry focused on understanding the barriers to achieving project goals could identify a number of ways to better support project processes and outcomes. Finally, evaluation of the costs and benefits of developing thorough data policies is clearly an area for further development. While no projects reported that they do not share data with any parties, the formality of data sharing and ownership arrangements is currently limited, and the conditions of data production are such that these questions are best addressed directly.

6. CONCLUSION

In conclusion, we see that there is great diversity among citizen science projects responding to our survey. Although the sample includes primarily observational projects in areas related to ecology, there is an impressive range of types of participation, social opportunities, technologies in use, approaches to data validation, units of contribution, and project goals. By documenting the diverse organizational and functional arrangements in citizen science projects, we contribute an initial description of the state of the population, which can support researchers in sampling for studies of citizen science, and for better understanding the broader context of citizen science when evaluating the representativeness of the results of such studies.

7. ACKNOWLEDGMENTS

This work was supported by US NSF Grant 09-43049.

8. REFERENCES

[1] R. Bonney, H. Ballard, R. Jordan, E. McCallie, T. Phillips, J. Shirk, and C. Wilderman. Public Participation in Scientific Research: Defining the Field

and Assessing Its Potential for Informal Science Education. A CAISE Inquiry Group Report. Technical report, Center for Advancement of Informal Science Education (CAISE), Washington, DC, 2009.

[2] R. Bonney, C. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. Rosenberg, and J. Shirk. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, 59(11):977–984, 2009.

[3] A. Cho and D. Clery. Astronomy Hits the Big Time. *Science*, 323(5912):332, 2009.

[4] K. Crowston, K. Wei, J. Howison, and A. Wiggins. Free/libre open source software development: What we know and what we do not know. *ACM Computing Surveys*, Forthcoming, 2012.

[5] F. Danielsen, N. Burgess, A. Balmford, P. Donald, M. Funder, J. Jones, P. Alviola, D. Balete, T. Blomley, J. Brashares, et al. Local participation in natural resource monitoring: a characterization of approaches. *Conservation Biology*, 23(1):31–42, 2009.

[6] M. Fernandez-Gimenez, H. Ballard, and V. Sturtevant. Adaptive management and social learning in collaborative and community-based monitoring: a study of five community-based forestry organizations in the western USA. *Ecology and Society*, 13(2):4, 2008.

[7] E. Graham, S. Reddy, E. Yuen, K. Mayoral, et al. Using smart phones and citizen scientists to map invasive species and track spread over time. In *California Invasive Plant Council Symposium 2009*, page 28, 2009.

[8] J. Howe. *Crowdsourcing: How the power of the crowd is driving the future of business*. Century, 2008.

[9] S. Kim, C. Robson, T. Zimmerman, J. Pierce, and E. Haber. Creek watch: pairing usefulness and usability for successful citizen science. In *Proceedings of the 2011 International Conference on Human factors in computing systems*, pages 2125–2134. ACM, 2011.

[10] E. Moore and T. Koontz. Research Note A Typology of Collaborative Watershed Groups: Citizen-Based, Agency-Based, and Mixed Partnerships. *Society and Natural Resources*, 16(5):451–460, 2003.

[11] M. Raddick, G. Bracey, K. Carney, G. Gyuk, K. Borne, J. Wallin, S. Jacoby, and A. Planetarium. Citizen Science: Status and Research Directions for the Coming Decade. In *AGB Stars and Related Phenomena, Astro2010: The Astronomy and Astrophysics Decadal Survey*, 2009.

[12] M. Raddick, G. Bracey, P. Gay, C. Lintott, P. Murray, K. Schawinski, A. Szalay, and J. Vandenberg. Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. *Arxiv preprint arXiv:0909.2925*, 2009.

[13] J. Silvertown. A new dawn for citizen science. *Trends in Ecology & Evolution*, 24:467–471, 2009.

[14] R. Stevenson, W. Haber, and R. Morris. Electronic field guides and user communities in the eco-informatics revolution. *Conservation Ecology*, 7(1):3, 2003.

[15] B. Sullivan, C. Wood, M. Iliff, R. Bonney, D. Fink, and S. Kelling. ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10):2282–2292, 2009.

[16] S. Teasley and S. Wolinsky. Scientific collaborations at a distance. *Science*, 292(5525):2254, 2001.

[17] J. Travis. Science and commerce: Science by the masses. *Science*, 319(5871):1750, 2008.

[18] A. Wiggins. eBirding: Technology adoption and the transformation of leisure into science. In *Proceedings of the 2011 iConference*, Seattle, WA, February 2011.

[19] A. Wiggins and K. Crowston. Goals and tasks: Two typologies of citizen science projects. In *Proceedings of the Fourth-fifth Hawaii International Conference on Systems Sciences (HICSS-45)*, Wailea, HI, January 2012.