From Conservation to Crowdsourcing: A Typology of Citizen Science

Andrea Wiggins
School of Information Studies
Syracuse University
337 Hinds Hall, Syracuse, NY 13244 USA
Email: awiggins@syr.edu

Kevin Crowston
School of Information Studies
Syracuse University
346 Hinds Hall, Syracuse, NY 13244 USA
Email: crowston@syr.edu

***Abstract*—Citizen science is a form of research collaboration involving members of the public in scientific research projects to address real-world problems. Often organized as a virtual collaboration, these projects are a type of open movement, with collective goals addressed through open participation in research tasks. Existing typologies of citizen science projects focus primarily on the structure of participation, paying little attention to the organizational and macrostructural properties that are important to designing and managing effective projects and technologies. By examining a variety of project characteristics, we identified five types—Action, Conservation, Investigation, Virtual, and Education—that differ in primary project goals and the importance of physical environment to participation.**

I. Introduction

Citizen science is a form of research collaboration involving members of the public in scientific research projects to address real-world problems [1]. Active engagement in scientific work differentiates citizen science from other forms of public participation in scientific research where volunteers take less active roles, such as providing computing resources for projects like SETI@home or participating as a subject in a research study. Citizen science is related to long-standing programs employing volunteer monitoring for natural resource management [2], and is often employed as a form of informal science education or outreach to promote public understanding of science [3].

Experience with this model of science shows that with thoughtful study design and under the right circumstances, citizen science can work on a massive scale, generating high quality data that lead to reliable, valid scientific outcomes as well as unexpected insights and innovations [4]. These successes motivate scientists to explore how members of the public might contribute in their projects. However, the space of projects is vast and as yet not well described. As a result, it is difficult to cumulate findings or to determine how or if one project's experiences will be relevant to another.

Further complicating this picture, the phenomenon of citizen science has evolved over time. For example, a number of the citizen science projects that emerged over the past two decades place more emphasis on scientifically sound practices and measurable goals for public education [5]. Virtual modes of contribution make it possible for a broader audience to engage in scientific work, as in other open content projects. An increasing number and variety of citizen science projects are taking advantage of the affordances of technology to advance scientific research [6]. The forms of participation usually involve contributing data according to an established protocol, or completing structured recognition, classification, or problem-solving tasks that depend on human competencies [7].

Citizen science typologies to date have focused primarily on the integration of public participation in different steps of scientific research, with little attention to sociotechnical and macrostructural factors influencing the design and management of participation. Our goal in developing this typology is to generate a more comprehensive description of the landscape of citizen science by examining common characteristics of projects, grouping similar projects that share necessary conditions for successful research employing this mode of production. This typology will also serve the secondary purpose of establishing a basis for theoretical sampling to guide future research and cyberinfrastructure development.

II. Related Work

Citizen science project share characteristics with other kinds of open communities: there are similarities to peer production, open data is relatively common, and open participation is nearly universal. Substantial differences make existing typologies for these related phenomena unsatisfactory for describing citizen science; however, research on these related phenomena is useful as background to our conceptualizing. We also discuss existing typologies of citizen science projects.

Citizen science projects that are entirely mediated by information and communication technologies (ICTs) are often considered a form of crowdsourcing applied to science. Crowdsourcing is an ill-defined but common term referring to a set of distributed production models that make an open call for contributions from a large, undefined network of people [8], [9]. Such projects also represent a type of distributed scientific collaboration, also known as collaboratories [10], but until recently, most of this research has focused on projects in which all contributors are scientists or supporting professionals. Emerging scientific cyberinfrastructure projects are a related phenomenon that bear stronger similarity to citizen science than do the collaboratories discussed in the literature to date [11], [12].

Citizen science is clearly a different way of organizing online contribution than has been previously analyzed in the literature [13]. Unlike most online communities that have been studied, these projects are not self-organizing [14], [15], [16]. Citizen science does not represent peer production in the same sense as seen in prior work because the power structure of these projects is usually hierarchical. Furthermore, citizen science is not necessarily "open science," a term that refers to open source-like practices in formal scientific research settings. Many citizen science projects share data, but may not make the full research process publicly viewable for comment and discussion.

The structure of tasks is very similar to those of peer production, and the existing literature is helpful for understanding key aspects of citizen science [17], [18]. The hierarchical form of most projects differs from peer production, however, and likely creates a different sense of community with respect to authority, leadership, decision-making and sustainability [19]. Finally, there are strong similarities with respect to issues of motivation and progressive engagement that bear a striking resemblance to virtual communities or networks of practice, albeit with scientists as overseers of the community's practices [20], [21]. These prior models from studies of online communities of practice provide insight into the design of tasks and technologies to support citizen science communities.

In addition to considering the prior research on open communities, we review several typologies of public participation in research. In considering the wide range of roles for the public in scientific research, we examine three typologies from the ecological sciences in Table I, which describe several approaches to involving the public in research. The distinction between experts and non-experts, generally based on the acquisition of formal scientific training and credentials, is at the heart of most differentiations between forms of research practice. Large-scale citizen science projects present an interesting challenge to the dominant view of scientific expertise, as the assumption that volunteers are always non-experts can be faulty.

These typologies examine the participation of the public by focusing on engagement in different steps of scientific research. The level of detail in these analyses differs, as do their final categorizations, yet they are largely in alignment. Table I lists the different steps in scientific inquiry that are considered in each of the typologies. The table also includes the definitions for three classes drawn from a recent report from the Center for the Advancement of Informal Science Education [22], which makes the most comprehensive assessment of participation models in citizen science.

Besides evaluating the stages of scientific inquiry in which the public is involved, Cooper et al. [23] include additional details of research, education, and management goals, which are contrasted in a framework for integrating individual property owners in monitoring and active conservation efforts in residential areas. The research models in the typology demonstrate the interplay of scientist and landowner roles in adaptive management practices, which apply scientifically informed natural resource management strategies in an iterative process of intervention, evaluation, and revision.

Stage of Inquiry	**Cooper et al.**	**Wilderman**	**Bonney et al.**	*Contributory*	*Collaborative*	*Co-created*
Define question	✓	✓	✓			X
Gather information			✓			X
Develop hypotheses			✓			X
Design study	✓	✓	✓		(X)	X
Data collection	✓	✓	✓	X	X	X
Analyze samples		✓	✓		X	X
Analyze data	✓		✓	(X)	X	X
Interpret data	✓	✓	✓		(X)	X
Draw conclusions	✓		✓		(X)	X
Disseminate results			✓	(X)	(X)	X
Discuss results & ask new questions			✓			X

TABLE I

VOLUNTEER INVOLVEMENT IN ENVIRONMENTAL SCIENCE TYPOLOGIES, WITH DEFINITIONS OF PARTICIPATORY SCIENCE MODELS. ✓= INCLUDED IN MODEL; X = PUBLIC INCLUDED; (X) = PUBLIC SOMETIMES INCLUDED.

Distinguishing *community science* from *citizen science* based on the community control of the inquiry, Wilderman [24] proposes an alternate typology that includes community consulting, community-defined research, community workers, and community-based participatory research. These categories are congruent with those presented elsewhere with some relatively small variations, and the author differentiates between two forms of community workers models based on whether or not analysis activities are exclusive to scientists, also adding a category in which the community is engaged in a consultative capacity, represented as "science for the people". Contrasting this practice against "science by the people" casts the typical scientist-initiated project model in a negative light; a more neutral perspective might suggest "science with the people" as another potential characterization.

A comprehensive, educationally-focused technical report that summarizes many of these views discusses *contributory*, *collaborative*, and *co-created* projects, shown in Table I, synthesizing many of the prior typologies [22]. The authors examined case study projects with a rubric-based evaluation to make a multi-faceted assessment of outcomes in several key focus areas. The final framework resembles a simpler variation on the other models, but includes more detail with respect to the steps of scientific inquiry in which volunteers may be included, moving the sophistication of the typology up a level despite its apparent simplicity.

These and other prior efforts targeted at understanding the defining features of citizen science projects have focused almost exclusively on the types of scientific tasks performed by volunteers, leading to a limited though functional view of participation. The typologies to date have given little consideration to the alignment of the organizational characteristics, enabling technologies, and goals of projects. Building on this work, we examine a much broader range of characteristics to develop a typology focused instead on project goals and uses of technology to overcome the limitations of virtuality, which can help inform the development of cyberinfrastructure to support citizen science.

III. Methods

In this section, we describe the process we used to construct an empirically-grounded typology of citizen science projects. Our approach started with landscape sampling of projects, which were then coded on 80 facets for inductive, qualitative clustering based on subsets of these dimensions.

A. Data Set and Sampling Methods

Drawing on existing directories of citizen science projects[1], we examined a substantial number of additional features of these projects beyond those documented in the existing directory descriptions. The definition of citizen science that we employed for sampling is a project in which a professional researcher collaborates with volunteers in scientific research, meaning that these are frequently (but not always) scientist-initiated projects. As a sampling frame, the online directories of projects allowed us to select projects that are considered to be citizen science initiatives by an independent third party.

B. Landscape Sampling

Citizen science is a relatively new phenomenon, despite its roots in decades of volunteer monitoring practices. As a result, our sampling method was purposive, seeking to identify a sample of projects that is comprehensive in type rather than frequency. We employed the landscape sampling method discussed in [25], with sample selection resembling a form of ecological population study in which the researcher attempts to identify all unique types in an environment prior to assessing frequency of occurrence. This selection of projects specifically focused on identifying as broad a range of project types as possible, concluding when the addition of more projects no longer yielded particularly distinctive characteristics or combinations of features.

Landscape sampling provided an excellent approach for inductive typology development, but has inherent limitations as well. While it is useful for describing and categorizing a broad range of citizen science projects, we cannot use this sample to draw statistically representative conclusions about citizen science projects. The typology presented here does discuss the relative frequency of some facets according to type where these were consistent with other project characteristics, but does not discount the possibility (or likelihood) of counterexamples for the trends identified within each class. Table III-B provides a brief overview of the sampled projects that we evaluated for our typology.

[1]Citizen Science Toolkit, http://www.citizenscience.org, and Science for Citizens, http://www.scienceforcitizens.net

C. Analysis

The analysis process began with identifying a wide range of facets to describe citizen science projects. The facets were drawn from a conceptual model we constructed to describe citizen science projects [26], including inputs, processes and outputs at both the project and participant level. Given the exploratory nature of this work, the intent was to employ a very broad and inclusive set of characteristics for later revision with the addition of empirical evidence.

We then assessed each project across approximately 80 facets. These facets included details on project demographics (e.g., age, geographic range, research discipline, stated goals), organizational features (e.g., affiliations, funding sources, tax status), participation design (e.g., task types, skills or tools required), educational features (e.g., informal learning resources, curricular materials), outcomes (e.g., publications, protocol revisions, innovations), technologies (e.g., communication tools, web site features), processes (e.g., data validation, volunteer management, communication) and data management (e.g., data sharing, ownership, stewardship). In addition to several interviews with project leaders, data were drawn primarily from project websites or other published descriptions, which poses some limits on data quality and completeness. For example, it was frequently difficult to find concrete details on outcomes, such as annual contribution rates, revisions to the project design, or changes to the size of the contributor base. The data collection process also resulted in many refinements to the list of facets—some facets were split into multiple parts, while others were abandoned as either uninteresting in light of the evidence, or impossible to adequately assess with these data collection methods.

Finally, the set of projects was manually clustered using an inductive, qualitative approach. The facets with the most complete and unambiguous data were employed for clustering, with nominal coding of each facet followed by grouping and sorting projects according to related sets of facets. For example, a set of four facets including project funding sources and specific sustainability measures (donations, sponsorships, and commerce) were grouped together for consideration and comparison against other sets of facets, such as five common types of organizational affiliations (academic, community/NGO, federal, state/municipal, other).

These groupings of facets, along with stand-alone facets like research discipline and geographic scope, were then carefully examined for commonalities across projects. The dominant goals of the project, combined with the role of the physical environment, generated a clear separation of project types that revealed other patterns of values that tended to co-occur, providing further evidence for the utility of this characteristic as a distinguishing feature. This division of projects into types is not the only possible organization, a point we will return to in the discussion. Further, due to the relatively small number of projects and the nominal or qualitative data for many facets, we did not attempt a quantitative cluster analysis, which we will undertake in future work.

Project	URL	Type	Description
ReClam the Bay	www.reclamthebay.org	Action	Restoring local bay's clams and oysters
Shermans Creek Conservation Association	www.shermanscreek.org	Action	Protecting local creek
Did You Feel It?	earthquake.usgs.gov/eqcenter/dyfi	Conservation	Collecting earthquake intensity data
Twitter Earthquake Detection Program	recovery.doi.gov/press/us-geological-survey-twitter-earthquake-detector-ted/	Conservation	Collecting real-time earthquake data
Missouri Stream Team Program	www.mostreamteam.org	Conservation	River conservation
Spotting the Weedy Invasives	www.rci.rutgers.edu/~trails	Conservation	Locating invasive plants
Invasive Plant Atlas of New England	nbii-nin.ciesin.columbia.edu/ipane	Conservation	Creating regional invasive plant database
Northeast Phenology Monitoring	www.usanpn.org	Conservation	Monitoring phenology (seasonal life cycles)
What's Invasive	www.whatsinvasive.com	Conservation	Locating invasive plants
Monarch Larvae Monitoring Project	www.mlmp.org	Investigation	Collecting monarch butterfly distribution data
Who's Whoo-ing	www.mianus.org/owlcall	Investigation	Mapping suburban owl habitats
Community Collaborative Rain, Hail and Snow Network	www.cocorahs.org	Investigation	Collecting precipitation data
Great Sunflower Project	www.greatsunflower.org	Investigation	Collecting pollinator service (bee) data
Firefly Watch	www.mos.org/fireflywatch	Investigation	Collecting firefly distribution and activity data
Gravestone Project	www.goearthtrek.com/Gravestones/Gravestones.html	Investigation	Measuring weathering to study acid rain
SnowTweets	www.snowtweets.org	Investigation	Mapping snow depth
eBird	www.eBird.org	Investigation	Collecting bird observations
The Lost Ladybug Project	www.lostladybug.org	Investigation	Collecting data about ladybug distribution
Bay Area Ant Survey	www.calacademy.org/science/citizen_science	Investigation	Collecting data on local ants
FoldIt	www.fold.it	Virtual	Proving human superiority at protein folding
The Open Dinosaur Project	opendino.wordpress.com	Virtual	Creating dinosaur limb bone measurement database
Stardust@home	stardustathome.ssl.berkeley.edu	Virtual	Finding interstellar dust particles
Galaxy Zoo	www.galaxyzoo.org	Virtual	Classifying images of galaxies
Project Implicit*	implicit.harvard.edu/implicit/research	Virtual	Examining hidden biases
The Smell Experience Project*	psych-institute.med.nyu.edu/research/submit-story	Virtual	Collecting stories about changes in sense of smell
Perfect Pitch Test*	perfectpitch.freehostia.com/info_eng.html	Virtual	Determining whether perfect pitch differs by timbres
What on Earth	www.whatonearth.org.uk	Education	Collecting images of organisms for identification
Radio Jove Project	radiojove.gsfc.nasa.gov	Education	Learning about radio astronomy
Fossil Finders	www.fossilfinders.org	Education	Learning about Devonian fossils
Globe at Night	www.globeatnight.org	Education	Learning about light pollution

*These projects are not considered citizen science because participants are subjects rather than collaborators; they are included for completeness.

IV. Findings: Types of Citizen Science

Based on the clustering, we identified five mutually exclusive and exhaustive types of projects, which we labelled Action, Conservation, Investigation, Virtual and Education. Action projects employ volunteer-initiated participatory action research to encourage participant intervention in local concerns. Conservation projects address natural resource management goals, involving citizens in stewardship for outreach and increased scope. Investigation projects focus on scientific research goals in a physical setting, while Virtual projects have goals similar to Investigation projects, but are entirely ICT-mediated and differ in a number of other characteristics. Finally, Education projects make education and outreach primary goals. In the remainder of this section, we describe each of these types in more detail, giving examples and discussing scientific, organizational and technological issues relevant to each type of project.

A. Action

Action-oriented citizen science projects encourage participant intervention in local concerns, using scientific research as a tool to support civic agendas. At their core, most action projects employ participatory action research approaches, but unlike most published research using these methods, grassroots or "bottom-up" organizing is most common. These projects are not conceived or planned by scientists, but instead by citizens, and usually involve long-term engagement in local environmental concerns for which science-oriented activities are intimately linked to the physical world.

1) Example: Sherman's Creek Conservation Association (SCCA) was formed to protect a local creek and provide environmental education to the surrounding areas. The formation of SCCA was sparked by opposition to a proposed electrical generation facility that would have violated zoning regulations. Collecting 18,000 petitioner signatures—representing nearly half the population of rural Perry County, Pennsylvania—the SCCA was successful in their political action, and subsequently worked with environmental scientists to propose guidelines for revision of local zoning ordinances. The SCCA continues to engage rural citizens in watershed monitoring, stream cleanup events, and community outreach programs.

2) Scientific Issues: Action citizen science projects are most likely to engage professional researchers as consultants or collaborators rather than initiators. Their goals sometimes demand rigorous research procedures that require more substantial training and ongoing participation than many other citizen science projects. The data and research findings from Action projects are not likely to become part of the scholarly knowledge base, as the goals are strongly oriented toward providing evidence for intervention. These efforts can provide a foundation for long-term environmental monitoring in a given locale, but the wide variation in methodologies employed by each independent effort creates challenges for the aggregation of data from Action projects.

3) Organizational Issues: The bottom-up, grassroots organizing that characterizes these projects is typically only successful at a local level, and does not scale well without substantial organizational development. The Action projects identified in our sample are 501(c)(3) organizations supported mainly by their local communities and occasional small grants. The primary challenge faced by these projects, and many similar grassroots organizations, is long-term sustainability. Indeed, both examples in our sample showed substantial efforts toward ensuring project sustainability by seeking donations, sponsorships, memberships, and other fundraising initiatives.

4) Technology Issues: As in most other small nonprofit organizations, technology promises opportunity but presents substantial challenges. The Action projects in our sample made minimal use of technology to support their activities, creating simple web sites to serve primarily as outreach tools for recruiting new volunteers, listing upcoming events, thanking donors, and occasionally disseminating results. For these projects, technology is often more of a burden than a boon, as it is difficult to attract and retain volunteers to develop and maintain IT infrastructure. In addition, the local scale of participation reduces the demand for a centralized online venue for interaction or data submission, as it can be far more efficient to use lightweight technologies and rely more heavily on co-presence at meetings and events to communicate.

B. Conservation

Conservation projects support stewardship and natural resource management goals, primarily in the area of ecology; they engage citizens as a matter of practicality and outreach. Like the Action projects, they are strongly rooted in place, and volunteer engagement focuses on data collection activities. Most Conservation projects include explicit educational goals or content; the only exceptions in our sample were projects currently in a pilot phase of development. These projects tend to be regional in scope, which may be a reflection of their goals and the challenges associated with complex collaboration partnerships, as all of these projects have affiliations with larger state or federal agencies.

1) Example: The Northeast Phenology Monitoring project is a regional partnership between the US National Park Service (NPS) and the USA National Phenology Network, along with several nonprofit organizations and ecology networks. The partnership's purpose is developing a program for long-term coordinated monitoring of phenology (plant and animal life cycles) in the Northeast region. The project is currently in a pilot stage of development, refining and testing protocols intended to generate comparable data, despite the necessity of tailoring specifics to the unique features of each site. In its first year, three sites implemented different training and monitoring methods with different volunteer demographics on a small scale, with substantial NPS staff oversight. The project has subsequently increased the number of implementation sites and introduced web-based data entry to replace the provisional use of paper data forms.

2) Scientific Issues: Conservation citizen science projects focus on generating data intended primarily for resource management decision-making, and are often equally concerned with promoting volunteer stewardship and awareness. While the focus of these projects is management, careful attention is paid to scientific validity as well; projects without strong academic affiliations are typically lead by professional researchers employed in governmental organizations. The staff organizing these projects can often draw on established volunteer bases, and tend toward conservative estimates of volunteer interests and abilities based on prior experience when creating research designs. Although most of these projects generate public data sets, they are not often presented in an easily accessible format.

3) Organizational Issues: Most Conservation citizen science projects are long-term monitoring efforts, but unlike the Action projects, the projects in our sample show a very low incidence of sustainability measures. This suggests heavy dependence on federal or state funds, and grants from agencies like the United States Geological Survey and United States Department of Agriculture, which were primary sources of support for several Conservation projects in our sample. These projects include both top-down (researcher-initiated) and middle-out (management-initiated) forms of organizing [27], which appears to occur mainly in projects with federal agencies as primary partners.

4) Technology Issues: The Conservation projects in our sample demonstrate a bimodal distribution with respect to technology uses—either fairly limited, or fairly sophisticated. Two of the projects we examined are technology-driven, emerging from technology design projects or new opportunities for data collection through ICTs, one collecting earthquake reports from the Twitter social networking site, and another employing smartphone applications for submission of geotagged images of invasive plants. By contrast, while most of these projects used web technologies for data entry and access to results, two projects had not implemented online data submission.

C. Investigation

Investigation projects are focused on scientific research goals requiring data collection from the physical environment; these projects best fit the definition of citizen science from [1]. While education is not always an explicit goal, it is frequently a strongly valued but unstated purpose, and Investigation projects often provide educational materials or include task structures that support ongoing learning. These projects range from regional to international in scope, and can achieve very large scales of participation: two projects that we evaluated engage tens of thousands of participants and receive millions of annual observations. Most of these projects focus on biological research, but our sample also includes projects in meteorology and climatology.

1) Example: The Great Sunflower Project was created by a single scientist to study pollinator service. Participating volunteers report data on the activity of bees in their gardens, following a specific protocol for observation and reporting. Volunteers' contributions are coordinated through a simple web portal running on an open source content management system, where participants describe their gardens and make observation reports. The project has been so successful in attracting volunteer interest (close to 80,000 people registered in two years) that maintaining project sustainability has required changes to the original participation protocol, fundraising efforts outside of the usual academic sources, and additional staffing for community management.

2) Scientific Issues: Valid scientific results are a substantial concern for these projects, which are aimed at formal knowledge production and are most often organized by academics. Careful project and task design are the primary means for generating reliable results, and these projects employ a variety of validation methods (e.g., uniform equipment, entry form validation, triangulation, algorithmic flagging for expert review), though they rarely utilize direct replication or explicit volunteer screening, which might jeopardize secondary educational goals. In addition, the spatial distribution of participants is usually an asset for these projects, which typically examine geographic distribution of species or natural phenomena (e.g., precipitation), but poses concurrent challenges in sampling bias. The uneven distribution of human populations can lead to problems with representativeness of data sets covering large areas, but this drawback can be ameliorated in the research and task design, as well as by attracting larger numbers of contributors and through targeted recruitment efforts.

3) Organizational Issues: Most Investigation projects involve academics or nonprofit conservation organizations as the primary organizers, and a top-down structure of organizing is a defining characteristic. Operating at larger physical scales, these projects may attract substantially larger numbers of volunteers—several in our sample reported tens of thousands of contributors—which can quickly lead to management and sustainability challenges, as seen in the Great Sunflower Project. As a result, nearly all of these projects engage in some form of fundraising, demonstrating the most diverse set of strategies for financial sustainability among the types of projects we identified. Common measures include soliciting donations and selling promotional merchandise or tools required for participation, in addition to seeking grant funding from a variety of sources ranging from community foundations to the US National Science Foundation (NSF). The largest of the projects in this group also utilized sponsorships, sales referrals, and licensing as revenue streams.

4) Technology Issues: Similarly to the Conservation projects, the Investigation projects employ a wide variety of technologies. These range from a combination of a Yahoo! Sitebuilder web site and SurveyMonkey form for data submission to a multi-million dollar custom platform with numerous localized portals commissioned by independent organizations contributing to the project. As these contrasts suggest, the variation across Investigation project technologies is likely partially attributable to a combination of funding resources and project lifespan. All of these projects took advantage of the efficiencies offered by web-based data entry, but few provide data in readily usable formats; the eBird project is a notable exception with respect to providing data through multiple media suited to different audiences and purposes, along with the most substantive range of analytical tools of any project in our sample [28].

D. Virtual

In the science-oriented Virtual projects, all project activities are ICT-mediated with no physical elements whatsoever, differentiating them from the Investigation projects in which the physical places of volunteer participation was also important. While these projects share their goal orientation with the Investigation projects, their common features are quite distinct from all other groups of projects, and they represent a project type that has not been examined in prior typologies of citizen science. The projects in our sample came from the fields of astronomy, paleontology, and proteomics, a branch of microbiology focusing on protein structures. In addition, we classified several psychology projects listed at the Science For Citizens website, but found that they do not match our definition of citizen science because participating as a subject is not generally considered research collaboration. In most other respects, however, the features of the psychology projects were congruent with the Virtual projects, suggesting that their commonalities may be related to the influence of virtuality on participation.

1) Example: Galaxy Zoo is perhaps the best known of the Virtual projects, organized by an inter-institutional team of professional astronomers. Volunteers apply superior human perceptual capacities to computationally difficult image recognition and classification tasks, providing an important service in data reduction. The tasks are performed through a web portal that presents images of galaxies and asks volunteers to make judgments about specific characteristics, with questions such as whether the galaxy has a bulge or a bar in its center, or how many spiral arms are evident. The project's leaders ensure quality by having each image evaluated by multiple volunteers, with algorithmic flagging of low-consensus items for professional review. The site also offers a blog authored by the astronomers and forums for discussion among participants, providing multiple venues for engagement. After three years, Galaxy Zoo has classified over 56 million galaxies, and counts a growing contributor base of over a quarter of a million volunteers.

2) Scientific Issues: Like the Investigation projects, the primary challenge for Virtual projects is ensuring valid scientific results. The matter is complicated by the requirement of designing tasks suited to online participation that will elicit valuable contributions while maintaining volunteers' interest. Motivation concerns are addressed with various technological mechanisms that take advantage of natural human competitiveness, creating engaging game-like task designs, and sometimes the potential for discovery of a proverbial needle-in-a-haystack, e.g., a new type of astronomical body or a particle of space dust. In terms of validation efforts, the primary mechanism in every Virtual project is replication: multiple reviews or ratings, from two to hundreds, are combined with sophisticated algorithmic identification and prioritization of items for expert review. These methods seem to work very effectively with large numbers of volunteers, but require a critical mass of contributors.

3) Organizational Issues: All of the Virtual projects were formed through top-down organizing by academics, and most projects' affiliations are exclusively academic, a characteristic somewhat ironically consistent with relatively little provision of educational materials for participants. None of the Virtual projects are operated by nonprofit organizations, relying instead on research funding. Much like the Conservation projects supported by federal and state monies, the Virtual projects also had no substantive sustainability measures beyond grants; although two projects offer t-shirts through third-party services, these sales are more likely to indirectly support volunteer recruitment and retention than bolstering project finances. By contrast, however, almost all of these projects have an indeterminate project duration, suggesting that they will persist only as long as current funding (and researcher interest) permits, leaving the matters of data and infrastructure maintenance in question.

4) Technology Issues: As suggested by the scientific issues faced by Virtual projects, the types of tasks for which this approach is suited are limited, and in most cases, executing them requires complex custom web platforms. Virtual projects take advantage of advanced technology tools to make extensive use of reputational rewards and friendly competition; they award points, display leaderboards, and offer incentives such as naming privileges or co-authorship as rewards for discovery. Some projects also take advantage of self-competition by providing participants with performance feedback, a feature that is substantially easier to implement for the style of tasks seen in the Virtual projects than for most place-based projects. Access to computer scientists may also be a significant factor enabling domain researchers to coordinate the development of impressive custom platforms for human computation.

An exception to these observations is the Open Dinosaur Project, a true open science collaboration that operated differently from all other citizen science projects that we examined. This limited-duration project used only free technologies (Wordpress and Google Docs) for coordination of volunteer efforts, none of which were dependent on special software or platforms aside from the hosted services.

E. Education

The Education projects in this typology make education and outreach primary goals, all of which include relevant aspects of place. The projects that fit this definition could be further subdivided by focus on informal versus formal learning opportunities. As all members in this grouping are explicitly education-oriented, they provide informal learning resources, with most projects also offering formal curricular materials. In addition, several projects have designed tasks that permit cumulative learning experiences, a feature shared with a number of the Investigation projects.

1) Example: Fossil Finders is a project that brings together educators, students, and researchers from the Paleontological Research Institution in Ithaca, New York. The project focuses on investigation of Devonian-age fossils, with curriculum resources and teacher development workshops targeted for elementary and middle school science classrooms. Unlike many other citizen science projects, the structured classroom environment and supervised inquiry-based format permits students to form and test their own hypotheses, using data compiled from their own work as well as other classrooms' contributions. This also permits independent expansion and extension of the project activities, such as collection, identification and description of additional fossil specimens, and examination of additional features such as encrusting organisms or evidence of predation.

2) Scientific Issues: The Education projects in our sample focus on biology, astronomy, and paleontology. Some can be considered citizen science only by virtue of including a research partner as an organizer; that is, these would otherwise be classroom projects that would not involve contribution of observations or analysis to larger scientific research efforts. The relative cost of acquiring data through formal education projects is substantially higher, particularly in costs for participation, reinforcing our assessment of the primacy of educational goals over scientific contribution. In addition, there is a wide range of scientific rigor; overall, the emphasis tends to be on outreach, learning, and developing scientific inquiry skills, rather than on generating scientifically valid results. More projects in this category included data analysis tasks than any of the other types, however, providing opportunities for developing critical thinking skills that are largely absent elsewhere.

3) Organizational Issues: The top-down organizing of Education projects is hardly surprising, and most involve multiple different types of partner organizations. All of these projects appear to have substantial funding; however, as in other categories where large grants provide the bulk of project funding, limited measures were in place to ensure an ongoing revenue stream. The Radio Jove Project is an exception on this dimension; among all of the projects we sampled, it is the only completely self-sustaining project, although it almost certainly required substantial initial funding to achieve this state. The intended duration and sustainability of these projects is therefore an open question. Since half of the projects we examined appeared to be short-term projects and half long-term, it is likely that ongoing project operation is not an intended outcome for some Education projects.

4) Technology Issues: All of the Education projects used technology to support data entry tasks, with some making fairly sophisticated uses of technology. In one case, the project requires participants to build technology resources, significantly constraining potential participation but strongly supporting the educational goals. The target audiences are varied, with several projects' websites and tools designed to appeal to young children and families, and others oriented more toward teachers and older students. The functionality and resources offered on each project website is directly related to the intended audience; sites that were clearly designed to engage children had far less content and functionality than those intended for teen and adult participants. This suggests a constraint for technology design not evident in the other project types, as explicitly supporting youth participation requires a different standard of usability and instructional clarity than is typically provided for adult volunteers.

V. Discussion

In this section, we discuss the limitations imposed by the sample, the key dimensions used in the typology development, potential uses for the typology, and future work.

A. Sample Limitations

The sample of projects selected has some limitations; these include the sample size and predominance of projects based in North America. The sample size of thirty projects is only a portion of currently active citizen science projects; the Citizen Science Central site lists approximately 120 projects and Science For Citizens includes over 150 projects, with new projects surfacing regularly. The landscape sampling method has the goal of selecting representative entities rather than exhaustiveness or statistical analysis, however, making a smaller sample size appropriate for initial typology development. As our ongoing efforts seek to develop a comprehensive database of citizen science projects, we expect to identify further nuances permitting identification of subtypes based on secondary characteristics in future work.

A second limitation of the sample is inclusion of mostly North American projects, which does include several projects that are international in their focus and participation (organized from the US), and a UK-based project. Preliminary searches for additional projects in other countries suggests that these projects focus primarily on biodiversity monitoring (e.g., [29]) and strongly resemble those in North America; however, few examples have been found outside of North America and Western Europe.

B. Key Dimensions

As previously mentioned, we classified each project on over 80 dimensions drawn from a conceptual model based in the literature and fieldwork. The primary categories of facets included project demographics, organizational affiliations, funding sources, multiple types of outcomes, features of processes and technologies, and numerous specific aspects of project and task design. After examining several alternate arrangements, we found that grouping projects according to their explicit self-described goals produced clusters with consistent patterns evident in other sets of facets. We further divided the science-oriented projects according to whether there is a physical component to participation; strong regularities among the Virtual projects evident in no others, and the members of class of projects have previously gone unacknowledged in other typologies. We then reviewed the resulting typology with practitioners, who confirmed the intuitive fit of the typology to their experiences, as well as both the relevance and novelty of including physicality as a key point of differentiation between Investigation and Virtual projects.

All projects were coded with both a primary and secondary goal, while other facets could substantiate these explicit goals with further evidence. For example, education is both a project goal and a set of facets describing more specific details, such as whether project tasks provide structures for ongoing learning, and whether informal or formal learning resources are provided. Our evaluation is based only on the explicit goals mentioned in the project materials, which is a substantial simplification for many citizen science projects.

We note that education and outreach may be equally-weighted goals for many Conservation and Investigation projects, and this is in fact apparent in the secondary set of education facets. For example, the Missouri Stream Team Program, Monarch Larvae Monitoring Project, and eBird all showed evidence for every pro-education facet, but by contrast, Stardust@home was the only Virtual project with a positive indicator for more than one education facet. This set of observations is especially relevant to outcomes, as best practices guides for practitioners have repeatedly identified a balanced science-education approach to citizen science as most productive and successful [22].

Another theme that emerged is related to the relationship between types of affiliations and funding models that dominate each cluster. While none of the types demonstrated a single consistent strategy, the frequency of co-occurrence for particular combinations of these characteristics was striking: projects receiving substantial federal, state, or grant funding showed far less evidence of pursuing a diverse set of revenue sources that could ensure long-term sustainability. Similarly, geographic scope was linked to technological sophistication. The smaller local and regional projects had implemented the simplest technologies, while the projects operating at an international or global scale typically had the most feature-rich and carefully customized platforms to support participation by increasingly large numbers of participants.

C. Uses for the Typology

This typology is a promising complement to the existing participation-oriented typologies; it classifies projects by primary goal orientation and degree of virtuality. These features can also be compared to the participation-based typologies to identify patterns of participation style according to project goals. In addition, this typology highlights a class of projects previously overlooked by other typologies of public participation in scientific research—those entirely mediated by technology, such as Galaxy Zoo, FoldIt, and Stardust@home.

The first of several uses for this typology is guiding sampling for future research. In addition, however, we believe that the typology shows regularities among projects based on a number of readily identifiable characteristics that can be useful for project design and technology development. For example, we found that Virtual projects typically use a number of technology-based mechanisms to help motivate continued participation which are used much less frequently by other types of projects. Future technology development to support Conservation and Investigation projects might emulate the Virtual projects' successful examples, as these features have potential to improve recruitment and retention. Similarly, new projects in the early stages of development could use the typology to identify similar projects as potential sources of inspiration for research and task design. Finally, funding organizations could use the characteristics highlighted for each project type to help evaluate proposals for congruence with funding objectives.

D. Future Work

There are several potential directions for future work based on this initial typology. Better results could be achieved by directly eliciting project information via a survey, allowing us to construct a more robust classification which would almost certainly surface further nuances between project types. Classifying a larger number of projects would also help to ensure the applicability and completeness of the typology through more exhaustive sampling, permitting the use of statistical methods for clustering, and examination of the distribution of projects among the different types. In addition, this typology could be combined with prior participation-based typologies to examine the interaction of goals and participation design in practice. Finally, the typology can also be used as a foundation for further research on citizen science, providing a basis for project sampling according to readily observable characteristics for more in-depth study.

VI. Conclusion

Citizen science represents a new type of open movement, welcoming contributions to scientific research from a diverse population of volunteers. This domain of practice is rapidly expanding with the availability of enabling technologies and mounting evidence in favor of the efficacy of the research strategy. Prior citizen science typologies have focused primarily on the integration of public participation in different steps of scientific research, and this paper complements the prior work with an orthogonal classification based on project goals and virtuality. Our goal in developing this typology was to generate a comprehensive description of the landscape of citizen science to identify necessary conditions for successful research projects employing this mode of production. The typology identifies five types of citizen science projects—Action, Conservation, Investigation, Virtual, and Education—each of which is self-similar along other dimensions. This classification scheme provides insight into the organizational and macrostructural factors influencing the development of these projects, and maps the current space of citizen science more completely than prior typologies. Finally, the typology indicates potential directions for future research and cyber-infrastructure development to support public engagement in scientific research and its exciting transformative potential.

Acknowledgment

This research has been funded by US NSF OCI Grant 09-43049. The authors thank the Public Participation in Scientific Research reading group at the Cornell Lab of Ornithology for their insights and assistance with verification.

References

[1] J. P. Cohn, "Citizen Science: Can Volunteers Do Real Research?" *BioScience*, vol. 58, no. 3, pp. 192–107, March 2008 2008.

[2] K. Firehock and J. West, "A brief history of volunteer biological water monitoring using macroinvertebrates," *Journal of the North American Benthological Society*, vol. 14, no. 1, pp. 197–202, 1995.

[3] D. Brossard, B. Lewenstein, and R. Bonney, "Scientific knowledge and attitude change: The impact of a citizen science project," *International Journal of Science Education*, vol. 27, no. 9, pp. 1099–1121, 2005.

[4] D. Trumbull, R. Bonney, D. Bascom, and A. Cabral, "Thinking scientifically during participation in a citizen-science project," *Science Education*, vol. 84, no. 2, pp. 265–275, 2000.

[5] 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*, vol. 59, no. 11, pp. 977–984, 2009.

[6] J. Silvertown, "A new dawn for citizen science," *Trends in Ecology & Evolution*, vol. 24, pp. 467–471, 2009.

[7] A. Cho and D. Clery, "Astronomy Hits the Big Time," *Science*, vol. 323, no. 5912, p. 332, 2009.

[8] J. Howe, "The rise of crowdsourcing," *Wired Magazine*, vol. 14, no. 6, pp. 1–4, 2006.

[9] D. Brabham, "Crowdsourcing as a model for problem solving: An introduction and cases," *Convergence*, vol. 14, no. 1, p. 75, 2008.

[10] T. Finholt, "Collaboratories." *Annual Review of Information Science and Technology (ARIST)*, vol. 36, pp. 73–107, 2002.

[11] K. Lawrence, T. Finholt, and I. Kim, "Warm Fronts and High Pressure Systems: Overcoming Geographic Dispersion in a Meteorological Cyberinfrastructure Project," in *Hawaii International Conference on Systems Sciences*, vol. 40, no. 2. IEEE, 2007, p. 675.

[12] C. Lee, P. Dourish, and G. Mark, "The human infrastructure of cyberinfrastructure," in *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM, 2006, p. 492.

[13] D. Ellis, R. Oldridge, and A. Vasconcelos, "Community and virtual community," *Annual review of information science and technology*, vol. 38, pp. 145–188, 2004.

[14] M. Markus, B. Manville, and C. Agres, "What makes a virtual organization work?" *Sloan Management Review*, vol. 42, no. 1, pp. 13–26, 2000.

[15] M. Wasko, S. Faraj, and R. Teigland, "Collective action and knowledge contribution in electronic networks of practice," *Journal of the Association for Information Systems*, vol. 5, no. 11-12, pp. 493–513, 2004.

[16] K. Crowston, Q. Li, K. Wei, U. Eseryel, and J. Howison, "Self-organization of teams for free/libre open source software development," *Information and software technology*, vol. 49, no. 6, pp. 564–575, 2007.

[17] Y. Benkler, "Coase's Penguin, or, Linux and the Nature of the Firm." *Yale Law Journal*, vol. 112, no. 3, pp. 367–445, 2002.

[18] C. Haythornthwaite, "Crowds and Communities: Light and Heavyweight Models of Peer Production," in *Proceedings of the Hawai'i International conference on Systems Sciences, 2009*, 2009.

[19] B. Butler, "Membership size, communication activity, and sustainability: A resource-based model of online social structures," *Information Systems Research*, vol. 12, no. 4, pp. 346–362, 2001.

[20] J. Preece and B. Shneiderman, "The Reader-to-Leader Framework: Motivating Technology-Mediated Social Participation," *AIS Trans. on Hum.-Comp. Interact.*, vol. 1, no. 1, pp. 13–32, 2009.

[21] G. Fischer, "Beyond 'couch potatoes': From consumers to designers and active contributors," *First Monday*, vol. 7, no. 12-2, 2002.

[22] 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," Center for Advancement of Informal Science Education (CAISE), Washington, DC, Tech. Rep., 2009.

[23] C. B. Cooper, J. Dickinson, T. Phillips, and R. Bonney, "Citizen Science as a Tool for Conservation in Residential Ecosystems," *Ecology and Society*, vol. 12, no. 2, 2007.

[24] C. C. Wilderman, "Models of community science: design lessons from the field," in *Citizen Science Toolkit Conference*, C. McEver, R. Bonney, J. Dickinson, S. Kelling, K. Rosenberg, and J. L. Shirk, Eds., Cornell Laboratory of Ornithology, Ithaca, NY, 2007.

[25] N. Bos, A. Zimmerman, J. Olson, J. Yew, J. Yerkie, E. Dahl, and G. Olson, "From shared databases to communities of practice: A taxonomy of collaboratories," *Journal of Computer-Mediated Communication*, vol. 12, no. 2, pp. 652–672, 2007.

[26] A. Wiggins and K. Crowston, "Developing a conceptual model of virtual organizations for citizen science," *International Journal of Organizational Design and Engineering*, in press.

[27] A. Wiggins, "Organizing from the Middle Out: Citizen Science in the National Parks," in *Proceeding of iConference 2010*, 2010.

[28] 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*, vol. 142, no. 10, pp. 2282–2292, 2009.

[29] D. Schmeller, P. Henry, R. Julliard, B. Gruber, J. Clobert, F. Dziock, S. Lengyel, P. Nowicki, E. Deri, E. Budrys *et al.*, "Advantages of volunteer-based biodiversity monitoring in Europe," *Conservation Biology*, vol. 23, no. 2, pp. 307–316, 2009.