# VOSS: Theory and design of virtual organizations for citizen science

This proposal presents a three-phase theory-based study of virtual organizations (VOs) that enable massive virtual collaboration in scientific research. The projects to be studied take the form of a VO with a core of scientists and project leaders coordinating the work of a larger number of volunteer contributors, a format called citizen science. The proposed study is grounded in small group theory and rooted empirically in case studies, surveys and action research in citizen science projects. It is directed at advancing the understanding of what constitutes effective citizen science VOs and under what conditions citizen science VOs can enable and enhance scientific and education production and innovation. The specific goal of the proposed research is to identify key lever points in work design for enabling citizen science VOs to involve distributed, diverse volunteers in producing large scale, high quality, valued scientific research in an organizationally sustainable fashion. It addresses the following two research questions:

- What technological and social arrangements support intellectual production and innovation in virtual organizations of citizen scientists?
- What are the social and technological barriers to and enablers of participation in a virtual organization of citizen scientists?

*Expected intellectual merits.* The proposed study will make the following theoretical and practical contributions.

- *Theoretical contributions.* This research will add to understanding of organizational design from a new perspective, focusing on the sociotechnical structures and processes involved in production of scientific knowledge in projects comprising professionals and volunteers. The study will contribute to theory by refining and validating a conceptual framework to describe the relationships between organizational structure, work design and cyberinfrastructure technologies in use.
- *Practical contributions.* The research will indicate opportunities for employing citizen science in scientific research, which could lead to novel implementations of citizen science in other areas of scientific and engineering research and education. Results will aid scientists and project leaders in identifying appropriate project structures and best practices to employ when revising current citizen science projects or launching new ones.

*Expected broader impacts.* The project will benefit society by:

- investigating how involving the public in scientific research can advance science directly, in addition to goals of outreach or informal learning;
- generating and disseminating insights directly applicable to improving the design and conduct of citizen science projects, thereby improving the experiences of participants and the quality of the research results;
- determining the conditions under which citizen science VOs provide solution for large-scale data collection, as well as opportunities to leverage public interest in other aspects of scientific knowledge production; and
- exploring how citizen science VOs can be employed to extend the kinds of data that can be collected for scientific research, e.g., in social sciences or small sciences.

As well, the project will contribute to the education of doctoral, masters and undergraduate students who will learn about research and cyberinfrastructure development through their participation in the proposed project.

# VOSS: Theory and design of virtual organizations for citizen science

Virtual organizations are increasingly central to science and engineering projects, including those funded by the National Science Foundation [7, 49, 59, 84]. Much of the research on virtual organizations (VOs) for scientific work has focused on distributed collaboration among scientists and their students. This focus has led to a rich stream of research on scientific collaboratories [e.g., 44, 45, 68]. However, the widespread deployment of information and communication technologies (ICTs) has enabled new options for distributed collaboration. Phenomena such as free/libre open source software (FLOSS), Wikipedia and other forms of online interaction [e.g., 1, 27, 29, 48, 83, 110] prompt us to consider the potential of VOs for supporting massive, distributed and heterogeneous participation in scientific projects. Our study examines the phenomenon of citizen science, that is, research projects involving "partnerships between volunteers and scientists that answer real-world questions"[1] [10, 19, 21, 102]. These project-based partnerships are a form of VO, fitting the NSF's definition of "a group of individuals whose members and resources may be dispersed geographically, but who function as a coherent unit through the use of cyberinfrastructure"[2]. However, there are few studies of this form of VO, and its potential benefits for science are still being established.

The proposed research project, grounded in theories of small group behavior, has the goal of understanding what constitutes effective VOs for citizen science and under what conditions citizen science VOs can enable and enhance scientific and education production and innovation.

## 1.1 Citizen science

Citizen science is related to long-standing programs employing volunteer monitoring for natural resource management [4, 22, 46], and is often employed as a form of education and outreach to promote public understanding of science [6, 12, 69, 88, 99]. However, citizen science projects are increasingly focused on benefits to the scientific research as well [5, 9, 79]. The evidence is clear that in the right circumstances, citizen science can work on a massive scale and is capable of producing high quality data [11, 47, 102] as well as unexpected insights and innovations [72, 96].

Public contributions to scientific research can take a variety of forms, with participation ranging from nearly passive to deep engagement in the full process of scientific inquiry. Diverse volunteer populations can contribute to scientific research through a variety of activities, from primary school students engaging in structured classroom projects, to families volunteering together in "bioblast" one-day organism census events, to geographically-distributed individuals monitoring wildlife populations over time. In the biological and environmental sciences, citizen science projects have focused primarily on observation of ecosystems and wildlife populations (e.g., monarch butterflies, birds, reef fishes), where volunteers form a human sensor network for data collection. By contrast, in projects organized by researchers in astronomy, such as NASA's Clickworkers [51], volunteers apply superior human perceptual capacities to computationally difficult image recognition tasks, providing an important service in data analysis. The level of cyberinfrastructure support also varies, from simple data collection to more sophisticated task support. A goal of the proposed research is to identify, develop and test additional forms of useful cyberinfrastructure support.

This type of organizational and work design is not new to science (e.g., the Audubon Christmas survey of birds started in 1900), but we are now reaching the point where cyberinfrastructure

[1] from Citizen Science Central `http://www.birds.cornell.edu/citscitoolkit/`

[2] from the Virtual Organizations as Sociotechnical Systems (VOSS) Program Solicitation, NSF 09-540

and ubiquitous computing make broad participation of the public in scientific work a realistic research strategy in an increased variety of projects [85, 86]. VOs involving volunteers also show potential as a solution for data analysis in sciences drowning in data [67, 98]. The potential benefits of citizen science are beginning to be realized more widely, particularly when coupled with traditional scientific studies [42], leading to an increasing number of projects (the blog `http://citizensci.com/` lists dozens of examples). Bhattacharjee [8] notes that use of citizen science research methods were previously seen as a barrier to obtaining research funding from NSF, but more recently NSF has started funding such projects based on their scientific merit. A search of NSF-funded projects identified 28 that explicitly involve citizen science, with a total of $21.8 million in funding over the last eight years. Many of these projects incorporate citizen science solely as a form of outreach or education ($10.7 million for 10 projects), while six grants for planning efforts have come to $2.5 million and nine grants for project implementation and tool development have received $6.9 million in funding. As citizen science projects receive increased attention, the amount of funding per project is on the rise, with past projects averaging $636,000 of funding and current projects funded at an average of $865,000. To date, these projects have focused primarily on the natural sciences, specifically environmental and climate change ($9.2 million), as well as ornithology and entomology ($4.4 million each). By contrast, just under $200,000 has been granted for social studies of citizen science, for two small-scale case studies, suggesting future opportunities in this area. The increasing scale of citizen science projects, some of which involve tens of thousands of members of the public in distributed data collection, suggests a need for additional research. In particular, further study is needed to understand the effects of organizational and work design on the scientific outcomes of citizen science projects, and to determine if this research approach can be applied more broadly and more rigorously.

In response to this need, the current proposal presents a three-phase, theory-based study of the sociotechnical aspects of VOs that enable massive virtual collaboration of volunteer participants in scientific research, specifically cyberinfrastructure-supported citizen science research projects. Such VOs are similar in some respects to massive virtual collaborations such as FLOSS or Wikipedia, but are united by scientific goals that pose particular constraints on task design. For example, reliability of data collection is critical to the value of a scientific project, but not something that can necessarily be left to the "wisdom of crowds". Including volunteers in scientific research projects also results in very different distributed organizational structures than those of scientific collaboratories, raising new challenges for scientists to manage. For example, the design of scientific collaboratories may tacitly assume that participants have comparable and high levels of skill and will contribute relatively equally. This is rarely the case for citizen science volunteers, who may have widely varying levels of skill or knowledge, and contribute at levels differing by orders of magnitude. Combined, these factors raise unique concerns for the development of citizen science VOs as sociotechnical systems. Citizen science VOs thus represent a potentially important part of the VO design space, but one that has not yet been addressed by studies of scientific collaboratories.

The proposed study is grounded in small group theory and rooted empirically in case studies, surveys and action research in citizen science projects. The specific goal of the proposed research is to identify key strategies in task design for enabling citizen science VOs to involve distributed, diverse volunteers in producing large scale, high quality, valued scientific research in an organizationally sustainable fashion. The study addresses the following specific research questions:

- What technological and social arrangements support intellectual production and innovation in virtual organizations of citizen scientists?
- What are the social and technological barriers to and enablers of participation in a virtual organization of citizen scientists?

*Expected intellectual merits.* The proposed study will make the following theoretical and practical contributions.

- *Theoretical contributions.* This research will add to our understanding of organizational design from a new perspective, focusing on the sociotechnical structures and processes involved in production of scientific knowledge in projects comprising professionals and volunteers. The study will contribute to theory by refining and validating a conceptual framework of the relationships between organizational structure, work design and cyberinfrastructure technologies in use.
- *Practical contributions.* The research will indicate the best areas of opportunity for employing citizen science in scientific research, which could lead to novel implementations of citizen science in other areas of scientific and engineering research and education. Results will aid scientists and project leaders (the "practitioners" of citizen science) in identifying appropriate project structures and best practices to employ when revising current projects or launching new ones.

*Expected broader impacts.* The project will benefit society by:

- investigating how involving the public in scientific research can advance science directly, in addition to goals of outreach or informal learning;
- generating and disseminating insights directly applicable to improving the design and conduct of citizen science projects, thereby improving the effectiveness of VOs in terms of the experiences of participants and the quality of research outcomes;
- determining the conditions under which citizen science VOs provide a solution for large-scale data collection, as well as opportunities to leverage public interest in other aspects of scientific knowledge production; and
- exploring how citizen science VOs can be employed to extend the kinds of data that can be collected for scientific research, e.g., in social sciences or "little sciences" [40].

As well, the project will contribute to the education of students who will learn about research and cyberinfrastructure development through their participation in the proposed project.

The remainder of this proposal is organized into five sections. In section 2, we develop a conceptual framework for our study. In section 3, we present the study design, with details of the data collection and analysis plans. In section 4, we present the project management plan. We conclude in section 5 by reviewing the intellectual merits and expected broader impacts of our study and results of prior NSF support.

## 2 Conceptual development

In this section we develop the preliminary conceptual framework for our study. For this project, we have chosen to analyze citizen science projects as a kind of small group, specifically, a work team. Guzzo and Dickson [52, p. 308] defined a work team as "made up of individuals who see themselves and who are seen by others as a social entity, who are interdependent because of the tasks they perform as members of a group, who are embedded in one or more larger social system (e.g., community, or organization) and who perform tasks that affect others (such as customers or coworkers)". A team differs from a community of practice because members have a shared output whereas in communities of practice (e.g., the copier repairmen studied by Orr [87]), members share common practices, but are individually responsible for their own tasks. Members of a citizen science project share a goal and social identity, and the members perform interdependent tasks that affect others. A difference is that project members may vary greatly in their degree of identification with and contribution to the project, so the importance of these factors will be explored explicitly rather than taken for granted. Adopting this perspective allows us to draw from the extensive research on small groups, thus providing a strong theoretical starting point for our project.

For the current study, our conceptual framework draws on work in the small group literature [e.g., 54, 75, 80], incorporating concepts and relationships from the literature on organizational design, job design, volunteerism and participation in virtual communities, at both individual (i.e., volunteer, staff member) and organizational/project levels. Synthesizing elements from organizational design, sociology and studies of nonprofit management with small group theory strengthens our conceptual framework for understanding the antecedents of scientific knowledge production through massive virtual collaboration. Given the similarity of citizen science VOs to other forms of massive virtual collaboration such as FLOSS, we draw in particular on our prior research on FLOSS teams. Figure 1 shows the initial version of our framework, which is adapted from one we developed from a review of literature on FLOSS development [38], and which extends the input-process-output (IPO) framework that was the basis for our earlier NSF-funded research [31].

We organize our conceptual framework as an input-mediator-output-input (IMOI) model [63]. Inputs are the starting conditions of a team, which includes member characteristics and project/task characteristics [54]. Mediators represent factors that mediate the influence of inputs on outputs and are further divided into two categories: processes and emergent states. Processes represent dynamic interactions among team members as they work on their projects, leading to the outputs. Emergent states are constructs that "characterize properties of the team that are typically dynamic in nature and vary as a function of team context, inputs, processes and outcomes" [75, p. 357]. Outputs are the task and non-task consequences of a team functioning [77]. For example, outcomes for a citizen science project can include scientific data collected (a task output) as well as the volunteer

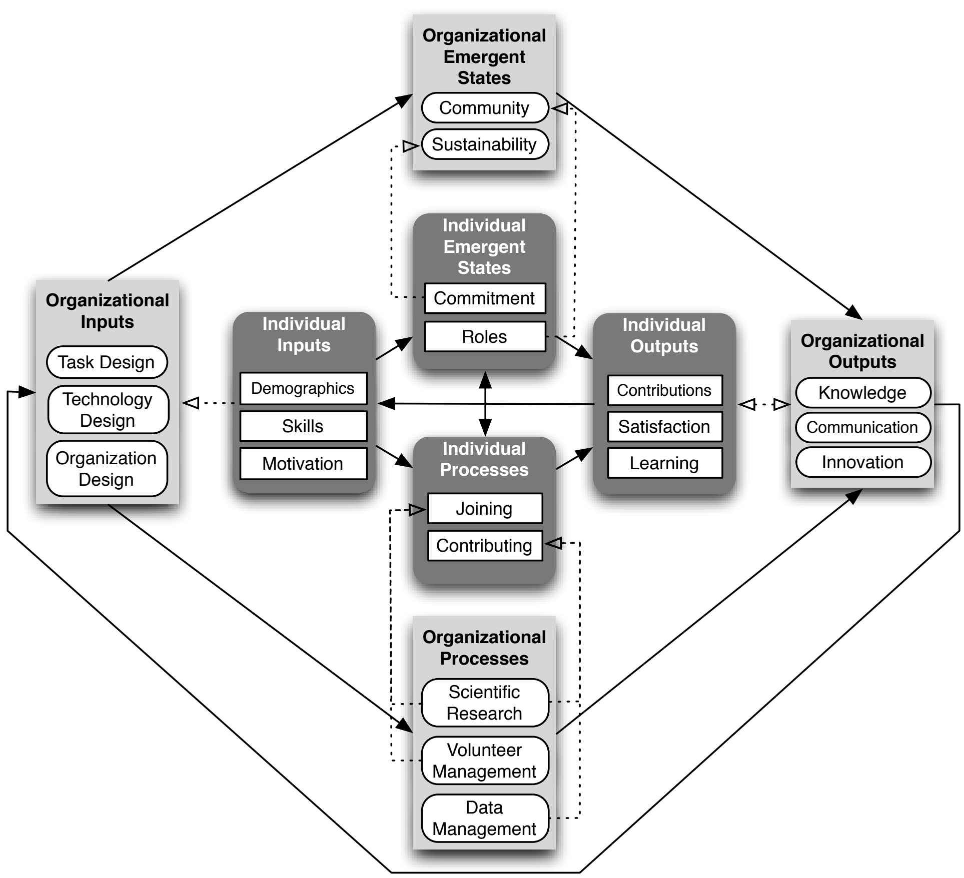


Figure 1: Initial conceptual framework: a multilevel input-mediator-output-input model

learning about the science (a non-task output from the point of view of the research itself). Finally, the framework includes feedback loops between outputs and inputs, treating outputs also as inputs to future group processes and emergent states [63]. In the remainder of this section, we briefly describe each of the elements of the framework and relations among them.

## 2.1 Inputs

Inputs are the initial conditions of a project, drawn from the surrounding environment and affected by prior project outputs. We include both individual level characteristics (the volunteers and the project staff) and project/task characteristics. At the individual level, staff and volunteers come to the project with diverse demographics, levels of skill, and motivations for participation that affect their individual contributions to the project. While demographics and skills will vary among volunteers involved in different projects, both practical reports and academic theory suggest a number of common motivators for volunteerism, which may have differential effects on individual experiences and performance in citizen science projects [18, 71, 97].

At the organizational level, we will examine the effects of organizational, task and cyberinfrastructure technology design. Organizational design is a key point of differentiation between citizen science VOs and other scientific collaboratories. The configuration and geographical distribution of participants can vary widely, as can the size of the core research group, which can range from a single PI with a research assistant or two to an interorganizational network of governmental agencies, scientific researchers and nonprofit organizations, each with different interests to fulfill and resources to contribute. However, the overall structure of the projects seems likely to parallel the "onion" structure that describes many FLOSS projects: a core of highly involved project leaders, surrounded by a larger group of active volunteers and a still larger group of occasional contributors [31]. One important difference in citizen science projects is that there are often formal status differences that separate these groups, e.g., most core participants likely have graduate training and formal roles as staff or advisors to the projects, while other participants are lay volunteers.

The second organizational input, "task design", encompasses several related concepts, including the research design for the study, the job design for volunteers and researchers and the task design for citizen science protocols. Citizen science as a mode of production is likely suited only to particular types of research, so research designs and protocols must reflect careful consideration of job design and task design [10, 21, 90, 93, 108]. Some tasks may be feasible and interesting for volunteers, with proper design, while others may have to be reserved for paid professional staff. While some researchers have carefully honed research and protocol design configurations for effective data collection by volunteers [39, 47], it is not clear whether the willingness, interest and diverse skills of the volunteers are fully engaged [42]. This observation suggests that there may be room for volunteers to contribute productively to additional aspects of scientific research, within the appropriate enabling structures.

Finally, technology design and use is of particular interest given the potential of cyberinfrastructure to support citizen science VOs, data management in particular. The potential for economies of scale increases with the diffusion of consumer technologies (e.g., PCs, high-speed Internet, GPS, digital cameras) and new modes of interaction that facilitate wider participation. The Appalachian Trail (AT) MEGA-Transect Project's Citizen Science Working Group report recommends that project partnerships include a scientist and an educator to address the scientific and educational goals of the project, as well as a technologist to help address the potentially substantial data management and information systems challenges [3]. When considering how organizational design and task design interact with cyberinfrastructure in the context of scientific VOs, the entire research process must be examined. For citizen science VOs, concerns over volunteers' ability to use instrumentation and the usability of data reporting forms (and subsequent usability of the data they capture) has prompted careful attention to usability testing of technologies designed for volunteers [3, 93]. Understanding the range of interactions between such diverse end users and technologies that support the scientific research is important to creating usable, robust systems for collecting usable independent contributions by distributed volunteers [16].

## 2.2 Processes

In the IMOI model, the inputs described above are conceptualized as affecting the effectiveness of projects through two sets of moderators, processes (described in this section) and emergent states (described in the following section). Processes are the dynamic interactions among group members as they work on their projects, leading to the outputs. Understanding these work practices is key to answering our first question regarding the technological and social arrangements that support intellectual production and innovation in virtual organizations of citizen scientists. At the individual level, processes of interest include joining a group, participating and making contributions [5, 103].

Individuals can participate in and contribute to a project at a range of levels. A minimal level involves simply providing computing resources (e.g., SETI@home) or serving as a subject for research (e.g., by joining a subject pool for online surveys). Even in these cases, participants may benefit by learning about the scientific goals and methods of the project. A higher level of participation may involve data collection or monitoring (e.g., the Audubon Society's Christmas Bird Count), usually about the volunteers' local environment. Similarly, participants might do basic data analysis (NASA's Clickworkers), contributing their human perceptual and knowledge organization capabilities. With training, some volunteers participate at still higher levels, contributing to research through problem formulation or hypothesis testing [102], assisting in running projects, e.g., by supervising or training other volunteers and even by making novel contributions [93, 96].

At the organizational level, processes of interest include the process of scientific research itself, throughout the data lifecycle. A key issue here is the nature of the science being done, the kinds of data and analysis required and the mapping of tasks to different actors, e.g., volunteers or staff. Similarly, the processes employed by the project for data management will have a significant impact on the project outcomes [3]. A particular concern is the practices employed by interorganizational projects to ensure reliability of data created by volunteers. Finally, a unique aspect of citizen science projects is the applicability of volunteer management practices more often associated with nonprofit organizations, e.g., recruitment, selection, orientation, training, supervision, evaluation, recognition and retention of volunteers [43, 50, 55, 109]. In addition to determining conditions under which desirable outcomes can be obtained, a goal of the study is to understand how design requirements for the organization, the scientific work and the cyberinfrastructure interact in order to optimize participation and the scientific value of the work [15, 60, 101].

## 2.3 Emergent states

Emergent states are dynamic properties of the group that vary as a function of inputs and processes; past research suggests a number of potentially relevant emergent states. These include task-related factors that describe the state of the group in terms of its progress on the scientific task, as well as social factors that describe social states of the group that enable that work. At the level of the project, research on other kinds of VOs has identified the importance of factors such as trust, cohesion, conflict and morale that affect the feelings of community in the group, and thus its long term sustainability [2, 65, 66, 76, 78, 81, 89]. In our work, we will seek to determine the role these factors may play in citizen science VOs.

At the individual level, we are particularly interested in the evolution of volunteers through different roles in the group, from initial volunteer through sustained contributor, and potentially to more central roles [24, 92]. A closely related concern is volunteers' level of commitment to the project and how it influences their task performance [14, 20]. Understanding these is a key step in answering our second question regarding the social and technological barriers to and enablers of participation in a virtual organization of citizen scientists. For example, in [24] we described various motivational factors at play as participants move from curious initial participants to sustained contributors to meta-contributors, whose efforts help structure and thus ease the contributions of others. We view the decision to make an initial contribution as largely curiosity-driven ("testing the waters"), driven by project visibility and facilitated by the expected ease of joining and participating and the contributor's having available time and some level of expertise, domain interest and self-efficacy. By contrast, we expect that the decision to continue contributing derives from the contributor's feelings of commitment to the project and its goals, the intrinsic motivation of the task and feedback from the task and other participants. Finally, we suggest that the decision to meta-contribute is driven by a sense of group membership, leading to feelings of obligation to the group, as well as by the intrinsic motivation of the task. In the proposed study, we will explore whether these stages of engagement and the theorized motivations are useful in describing the participation of contributors to citizen science projects.

In the IMOI model, processes and emergent states are conceptualized as moderating the relation between inputs and outputs of the project. At the individual level, the input elements of organizational design, task design and technology design affect motivation and participation of distributed volunteers, thus affecting the outputs [13, 64, 71, 100, 104]. For example Bussell and Forbes [13] describe a variety of ways in which people can volunteer and note the importance of carefully designed processes to retain volunteers, both indicating the importance of task design. At the project level, the inputs transform the means of production of scientific knowledge [70], shaping the demand for supporting cyberinfrastructure [82, 86] and potentially transforming the organizational design at the organizational level.

## 2.4 Outputs

Finally, outputs represent task and non-task consequences of a functioning group [77] that lead to the project's effectiveness. At the individual level, task outputs for a scientific project are contributions, often raw or processed data although other contributions are possible depending on the project. Important measures of this output (and thus, the overall effectiveness of the VO) include the quantity and especially the quality of the data, analysis and findings. In addition to the individual-level outputs, a citizen science VO will have outputs at the project level, such as the scientific knowledge created from the data. Innovative findings, processes and tools can also emerge from involving the public in scientific research. For example, a new astronomical body, now called *Hanny's Voorwerp*, was discovered by a Dutch elementary school teacher volunteering with the GalaxyZoo project [17]. Finally, at the societal level, the success of a project may affect public participation in and perception of science [19], create informal learning opportunities [69, 73], and provide the mechanisms for knowledge production at an unprecedented pace and scale [8, 41].

To further conceptualize VO effectiveness, we draw on Hackman's [53] model of group effectiveness. In addition to task completion, as outputs Hackman also includes the satisfaction of group members' individual needs, which includes aspects such as individual learning and personal satisfaction. These measures of effectiveness relate closely to the educational mission of many citizen science projects. Finally, Hackman includes the importance of the continued ability of the group to work together, speaking to the sustainability of the project, in terms of both the task and the social structure of the group. In other words, a VO is not effective if it achieves its goal once but drives away participants in the process.

A key point of the IMOI model is that outputs themselves become future inputs to the process. Positive personal outcomes can lead to increased motivation for future participation, and individual learning can increase a member's capability to work on additional tasks. At the project level, learning may lead to innovation in research approach, resulting in changes in the task design and thus the group processes. Positive project outputs may lead to increased interest among practitioners in applying this mode of research, as well as increased visibility for the project, helping to recruiting and retain additional volunteers.

In summary, synthesizing elements of prior research on small groups with contextually relevant theory provides a solid theoretical grounding for studying the organization of large numbers of virtual volunteers for scientific research and the context, structure and function of citizen science VOs. However, the several differences in settings between prior work and the current study suggests the need to both validate the applicability of this body of theory, and also search for possible extensions [43]. Therefore, we next present the design of our proposed research project, in which we first refine this conceptual framework in an exploratory study, before using the refined framework for further research and to guide cyberinfrastructure design.

# 3 Research design

In this section, we discuss the design of the study. In this section, we address the overall research strategy, methodological integration and concepts to be examined, deferring to subsequent sections the details of the proposed data collection and analysis techniques. We will model our study after our successful study of work practices for FLOSS teams (NSF HSD 05–27457), with appropriate tailoring for the context of this research setting. Our overall plan is to use our framework to guide an initial round of exploratory research to identify the factors and relations that are most important for overall effectiveness. In subsequent phases, the refined framework will be tested and used to suggest potentially beneficial technologies to support these projects. We envision our entire research project as having three overlapping phases. Each phase will last roughly a year, though the transition between these phases will be gradual, as the start of one phase overlaps the completion of the previous stage. The timeline for the research is shown in Figure 2, and the synthesis of methods in Figure 3.

## 3.1 Phase One: Investigating Issues and Practices

In the first phase, we will build on prior efforts to catalog examples of citizen science projects, e.g. `http://citsci.com` and `http://birds.cornell.edu/citscitoolkit/projects/alphabetical`,

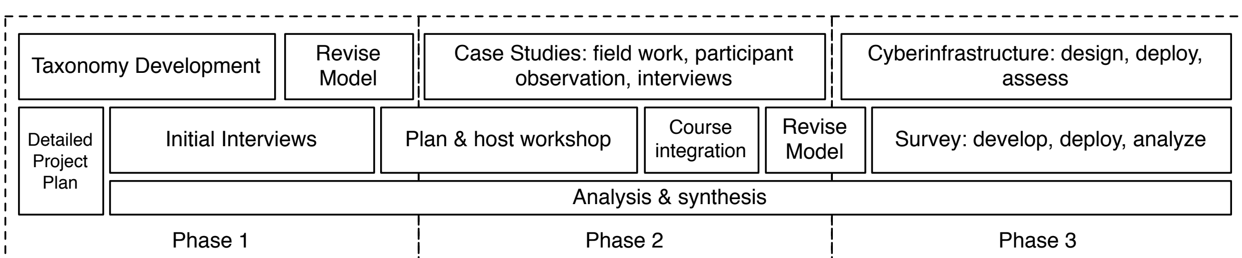


Figure 2: Timeline of proposed study activities.

by creating a taxonomy of citizen science VOs. Developing a taxonomy based upon our initial conceptual framework and interviews will generate a comprehensive description of the landscape of citizen science research and will also serve the secondary purpose of establishing a basis for theoretical sampling in later phases. A key goal in developing a taxonomy is to identify relevant theoretical dimensions and values that characterize different kinds of citizen science projects. As such, the taxonomy building will also contribute to building a richer conceptual framework to guide subsequent data collection and analysis. Finally, by examining the common characteristics of citizen science projects, we will make a preliminary identification of the necessary conditions, as opposed to best practices, for successful research projects employing this mode of production.

The taxonomy will be based on data collected from the Internet and from academic publications, as well as interviews with citizen science practitioners. These interviews will focus on illuminating issues and relevant dimensions of practice and will help identify other citizen science projects for our taxonomic development efforts. Interviews are also expected to establish relationships with project leaders that will provide access for later stages of research. In addition, we plan to incorporate the ideas of citizen science, e.g., by developing a system that allows practitioners to add their own projects to the catalog, described along the dimensions of the taxonomy.

### 3.2 Phase Two: Detailed Case Studies and Stakeholder Involvement

In the second phase, spanning year two, our primary focus will be development of detailed comparative case studies exploring the constructs from our conceptual framework and their relations for a small number of projects. For each case, we will examine the organizational, task and technology designs as inputs, the individual and project level processes and emergent states, and the outputs and effectiveness of the project, both for individual participants and overall. The case study protocol will be aimed at understanding the technological and social arrangements that support production and the social and technological barriers to and enablers of participation in the selected citizen science VOs (the study's research questions). The insights drawn from these in-depth case studies will contribute to further refinements of our conceptual framework, through which we will identify specific mechanisms employed in task design to support research quality and ongoing participation.

The depth of inquiry for these case studies will necessarily restrict the number of cases so we expect to select three or four citizen science projects for this phase of research. Theoretical sampling of cases will be guided by our findings from the first phase of research, but will likely reflect such factors as scientific field of inquiry, integration of citizen scientists as research collaborators, project scale and scope, overall effectiveness and use of ICT. For example, case selection might include a small project with one PI and a few students using open source software to collect data for a single study (e.g., The Great Sunflower Project); a large research group with multiple projects that employ centralized cyberinfrastructure to engage public participation at a national level over a period of years (e.g., Cornell Laboratory of Ornithology); and an interorganizational network of researchers, governmental agencies, and public sector groups with federated information systems supporting the collection of multiple types of environmental monitoring data for long-term ecological research and natural resource management (e.g., the Appalachian Trail MEGA-Transect).

To develop our case study methodology, in parallel with the taxonomy development in year 1, we will conduct an initial case study with one citizen science project. The AT MEGA-Transect's phenology collaboration[3] has informally agreed to serve as a case site during their 2009–2010 citizen

[3] Partners include the USA National Phenology Network, the National Parks Service's Northeast Temperate Network (NETN), Acadia National Park, Appalachian National Scenic Trail, Appalachian Mountain Club, Appalachian Trail Conservancy, and The Wildlife Society; `http://usanpn.org/?q=nps`

science program pilot to develop standardized protocols and regionally coordinated monitoring. In developing our methodology, we will again explore the potential of citizen science approaches for collecting data on a broader range of projects than we can address directly.

In addition to developing detailed case studies, the second phase will also include a research workshop to bring together researchers using and studying citizen science approaches to discuss the issues identified throughout the first and second phase of research. The workshop will provide an opportunity for mid-stream validation of the research findings, intensive involvement of stakeholders in developing specific aspects of the third phase of research inquiry and peer networking for participants. Funding has been requested to support travel and sustenance for a meeting of 15 participants at Syracuse University's Minnowbrook facility, as well as for modest administrative support for planning. A date for the workshop will be selected to be convenient for participants given the constraints of their fieldwork. Participants will be selected based on leading projects identified in the Phase I, suggestions from Phase II case site participants and an open call for participation. Particular attention will be paid to participation of members of underrepresented groups, and funding will be made available to enable participation by graduate students. The workshop will result in a scientist-focused report presenting our taxonomy of citizen science VOs and identifying common issues and solutions.

### 3.3 Phase Three: Surveys and Cyberinfrastructure

For the third phase, roughly in year three, we will build on the prior two phases and the workshop to test our refined framework in two ways. First, we will develop and deploy web-based surveys to assess the performance of the revised conceptual framework for describing citizen science research practices, challenges and outcomes, particularly as they relate to data quality and ongoing participation. There will be several targets for the surveys: 1) the scientists directing (or potentially directing) projects that currently or potentially use citizen science approaches (the "practitioners"), 2) current volunteers in projects and 3) members of the general public who might become volunteers. The survey will examine perceptions of the inputs, processes, emergent states and relevant outcomes and the relations among them to identify technological and social arrangements that support production and the social and technological barriers to and enablers of participation. Funding has been requested to support the survey, primarily for payment of incentives for participation.

The third phase will also include an action research and design science element, in which undergraduates enrolled in existing information management courses will develop and test prototypes of cyberinfrastructure to support selected citizen science projects, using design insights from our case studies and based on our conceptual framework. For example, the AT MEGA-Transect phenology study plans to collect phenology observations (i.e., times of recurring natural phenomena) via brochures made available to hikers on the trail. A simple piece of cyberinfrastructure would be a web-based reporting site to replace paper data forms. A more ambitious project might use the camera and GPS capabilities of smart phones to upload geotagged pictures of events to a system that enables volunteer analysis, e.g., of species and stage of development. A competition will be held to select the two most promising designs for implementation as field experiments in the Spring of the third year (we anticipate being able to complete at least a pilot test during the funding period). Our project plan includes training the research groups to maintain the technology. The seasonal nature of data collection in most citizen science projects allows for a natural annual cycle of implementation and evaluation.

Finally, we will synthesize the data from each phase of research for analysis of the impacts of organizational design, work practices and cyberinfrastructure on organizational and individual processes, emergent states and outputs in existing citizen science VOs. From this multi-faceted perspective, we will propose a framework to establish more general applicability of the citizen science mode of VO design for scientific research. Analysis of the conditions under which this organizational structure can provide increased organizational capacity and generate large-scale data sets suitable for scientific research will provide insight into other potential areas of inquiry that can benefit from this mode of organizing. We are particularly interested in the applicability of participatory modes of research for social science and "little science" research [40].

### 3.4 Data collection

To explore the constructs identified in the conceptual development section of this proposal (Figure 1), we will collect and analyze a range of data. Data collection specifics for case studies will vary based on the organizational characteristics, but each case will include documentation of data and volunteer management practices, research and protocol designs, participant observation and interviews with researchers and citizen scientist volunteers. In the remainder of this section, we will briefly review each kind of data and our plans for data collection.

*Interviews.* Interviews will be conducted in phases I & II of the project, face-to-face when feasible (e.g., for local sites and during observational periods) and by phone in other cases. Initial interviews with project managers will focus on understanding the dimensions of citizen science projects. Interviewing will continue until we reach saturation on the theoretical dimensions of the typology. In the second phase, interviews will be conducted with project staff and volunteers to develop a deeper understanding of the structural relationships affecting research and participation outcomes and to uncover issues and challenges faced by researchers managing citizen science VOs.

*Observation.* Observation will be used in Phase II. The extent of observation will vary based on access and organizational structure for each case study site. Through observation, we will explore the experiences and perspectives of both the research group members and the citizen science volunteers. For all case studies, the graduate assistant will engage in participant observation as a citizen science contributor. The independent and distributed nature of participation in many citizen science VOs is particularly convenient for participant observation. Funding has been requested to support travel to the project sites to enable observation of the core research group at work. Observation (and interviews) will be employed to document task design, focusing in particular on volunteer and data management activities and research planning. This approach will provide insight into organizational structure and context through participation in such activities as research group meetings and volunteer training events.

*Project documentation.* We will collect several forms of documentation for each case study, including documents created for citizen science volunteers and those created for organizational use. The documents will include volunteer recruitment and educational materials, research design and protocol documents, instructions for volunteers, observation reporting forms and web interfaces and published findings. Additional project documentation may include meeting minutes or other project management materials (e.g., organizational charts, planning calendars, etc.) as well as data management policies and procedures. Public-facing and internal documents are evidence of volunteer management and data management practices and provide a basis for evaluating the work structure of citizen science research in different organizational settings.

*Online community fora.* In some cases, projects communicate with volunteer participants via online discussion board or email listservs. The contents of these fora provide rich data from participants relevant to the problems they encounter, the insights they develop, the kinds of resources and contributions they can make, the volunteer roles that evolve and the way they interpret and carry out research protocols. Online communities data also provide evidence of volunteer management practices in citizen science VOs, allowing us to consider the role of organizational design for supporting large-scale volunteer contribution via online community management. This type of mechanism to support participation may be internally managed by project staff or may be delegated to a reliable group of "super volunteers" with proven expertise, as in many other online community contexts [91]. The choice of these or other solutions to address citizen science VO community management is likely to be a function of organization and task design.

*Survey.* To validate the findings from the first two phases of research, a web-based survey will be developed and deployed in the third phase. This survey will assess the extent to which the variables identified in our conceptual framework, and refined through interviews and case studies, are relevant to the broader experiences of target respondents. Items for the survey will be developed drawing where possible on prior research using the constructs of interest and pilot tested with members of the case site projects.

*Cyberinfrastructure experiments.* Cyberinfrastructure designs to enhance citizen science research quality and participation will be generated as deliverables for project-based information management courses at the Syracuse University iSchool. Organizations that serve as project sites for these courses will receive the completed design recommendations and have opportunity to provide feedback on prototypes, creating a practice-based formal learning opportunity for undergraduate students. Courses targeted for involvement include Science Data Management, Human-Computer Interaction and Website Design. We can also draw on a program currently offered by the School that employs masters students to work on development projects. Finally, we will explore involvement of students enrolled in the School's Cyberinfrastructure Facilitators certificate program. Planning of curricular integration will begin during the second year, in partnership with the course instructors, with system development to be done in the Fall of the third year.

These cyberinfrastructure designs will be tested in field experiments to examine their influence on relevant project outputs, e.g., the quantity and quality of data collected and participation. Appropriate measures of participation will vary by design solution but might include greater frequencies of observation reporting and increased levels of repeat contribution. Our analysis will also examine the informal learning benefits to volunteers, as well as the formal learning benefits of students involved in the cyberinfrastructure development. Finally, we will examine mediating variables, such as the dimensions of volunteer motivation identified in our framework.

## 3.5 Data analysis

Development of our conceptual framework for citizen science VOs requires integration of a variety of data collected through multiple methods. To accomplish this, we are planning an iterative process of data collection and conceptual framework development process, as shown in Figure 3. Each stage of data collection will contribute to ongoing analysis and will inform the subsequent research efforts.

### 3.5.1 Phase 1 Analysis

The first stage of analysis will apply content analysis techniques to project description data collected from the web, scholarly articles and interview transcripts to prepare initial findings for review and discussion at the practitioner workshop.

*Taxonomy development.* Creating the taxonomy of citizen science VOs will require content analysis to classify project characteristics from project documents and interviews. This effort will begin with Internet-based materials and will progress to include other project documentation as it becomes available through interviews. Our initial conceptual framework provides the organizational characteristics we wish to identify and, as we begin to build our taxonomy, this classification effort will inform the development of semi-structured interview schedules for case studies.

*Content analysis.* Both project documents and interviews will be subject to content analysis for identification of issues and practices related to work design and organizational design. Developing the codebook for content analysis will be a primary activity during this phase. The content analysis will be an ongoing effort that will be applied to additional data from the case studies and feedback from workshop participants.

### 3.5.2 Phase 2 Analysis

In the second phase of the analysis, we will build on the results from the first stage, our in-process detailed case studies and feedback from the workshop participants to refine our conceptual framework. In particular, we expect to further investigate the following concepts.

*Organizational design.* As we develop a more complete understanding of the variations in organizational design characteristics for citizen science VOs, our investigation will focus on identifying the antecedents to these organizational design choices. This analytical effort will contribute to developing generalized use case scenarios to document the conditions under which citizen science VOs provide a good solution to meeting research goals.

*Task design.* We expect that task design strategies will vary based on organizational design. Our analysis will extend the use case scenarios for organizational design by matching the organizational structures from our taxonomy development to the most common task design solutions. This analysis intends to provide a set of heuristics for researchers planning, managing and evaluating citizen science VOs to employ as they design and redesign research projects for this mode of engagement.

*Roles.* We plan to document citizen science VO roles using several approaches. First, we will look for evidence of explicit formal roles for staff and volunteers during our taxonomy development. Second, we will discuss project organizational structure as well as formal and informal roles of volunteers with project managers during interviews. Finally, we will look for behavioral evidence of informal roles in online communities for citizen science projects. This analysis of informal and structural roles should provide a useful tool for task design within citizen science VOs, as the ability to effectively leverage the skills and interests of a diverse pool of contributors can make an enormous

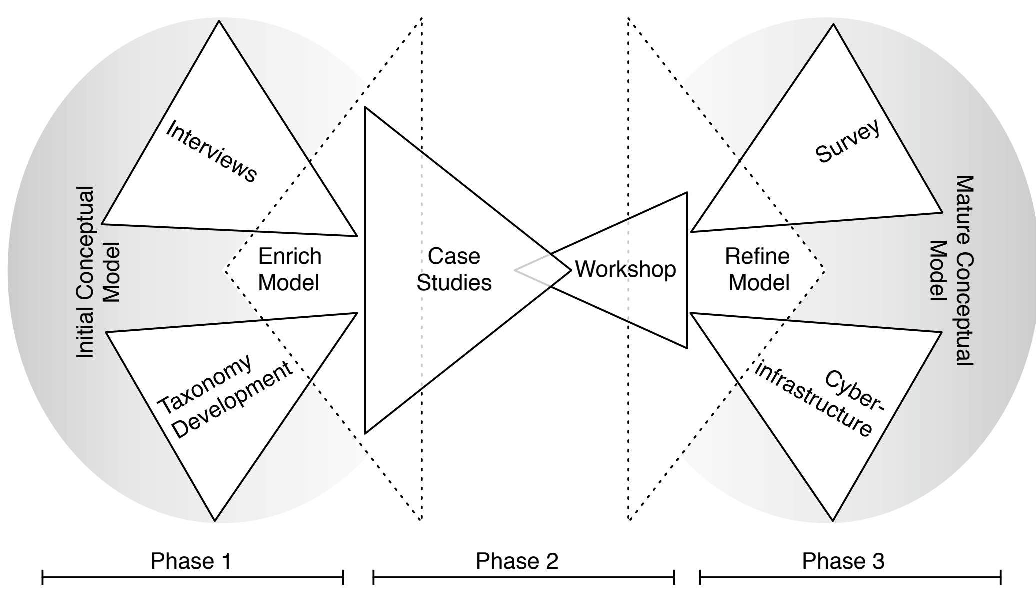


Figure 3: Planned cycle of data collection, analysis, and conceptual development.

impact on the scalability of the VOs, ongoing participation, and the quality of work.

#### 3.5.3 Phase 3 Analysis

In the third stage of analysis, we will bring together interview data with detailed case studies and survey results to answer our research questions, examining factors that enable and constrain participation, innovation and citizen science production processes. Survey and experimental data will be analyzed statistically, to examine the relation between the constructs of our framework. This aspect of our research will also examine organizational impact of cyberinfrastructure and the ways it transforms the work of scientific research. Common mechanisms for adjusting organizational structure to meet the changing demands of conducting citizen science research will be documented, as will best practices for the citizen science project types identified in the taxonomy from Phase 1.

## 4 Management plan

Based on preliminary assessment of the effort required, we are requesting funding for one graduate student and a small amount of summer support for the PI (1 summer month). The PI will work during the summer on project management and research design, and supervise the graduate student during the academic year. He will take particular responsibility for case selection, overall project design and report writing. The graduate student will support the PI in case selection, theory development, definition of constructs and variables. She will have primary responsibility for data collection and analysis, under the oversight of the PI. The budget includes travel support for case study interviews and observation of research group meetings. Volunteer participant observation work will be done independently, in the citizen scientist mode. We plan to involve colleagues teaching project-based courses to oversee the cyberinfrastructure development, though the PI and student will also be involved. An initial project activity will be the development of a more detailed timeline against which progress will be measured. The budget includes support during the summer for these activities, which include preliminary event planning for the research workshop.

## 5 Conclusion

Through the three phases of the study, described above, we will develop a better understanding of the growing phenomenon of citizen science VOs, some of which now involve tens of thousands of volunteers in scientific research. We posed the following research questions:

- What technological and social arrangements support intellectual production and innovation in virtual organizations of citizen scientists?
- What are the social and technological barriers to and enablers of participation in a virtual organization of citizen scientists?

The proposed study will answer these questions using theory from small group research and data from detailed case studies, surveys and field experiments. More specifically, we will develop an inventory and taxonomy of citizen science projects, detailed case studies of a few, surveys of the researcher community and the public, and pilot-tested technologies to support effective citizen science VOs, based on designs by undergraduate students. The proposed research will thus examine the spectrum of ways that citizens can meaningfully contribute to scientific projects while maintaining the integrity of the research and will advance our understanding of the factors that affect the effectiveness of such projects.

The project will benefit society by investigating how involving the public in scientific research can advance science as the primary goal, while still supporting outreach and informal learning; generating and disseminating insights directly applicable to improving the design and conduct of citizen science projects, thereby improving the experiences of participants, the quality of the research results, and the sustainability of the projects; determining the conditions under which citizen science VOs provide a good solution for large-scale data collection, as well as opportunities to leverage public interest in other aspects of scientific knowledge production; and exploring how citizen science VOs can be employed to extend the kinds of data that can be collected for scientific research, e.g., in social sciences or "little sciences" [40].

To ensure that our study has a significant impact, we plan to broadly disseminate results through journal and conferences publications, though a workshop with practitioners and on our Web pages. We also plan to disseminate results directly to and invite contributions from interested practitioners. Courses involving cyberinfrastructure design will also be incorporated into the curricula of the Syracuse University School of Information Studies, which was recently awarded a CI-TEAM grant to develop a model curriculum to train information professionals working with cyberinfrastructure. The cyberinfrastructure experiments in year 3 of the proposed project will provide class projects for these students. Finally, the project will promote teaching, training and learning by students involved in the research project, providing the opportunity to develop skills in data collection and analysis. In addition, curricular integration with existing information management coursework will provide undergraduates with hands-on experience designing and prototyping technologies to support citizen science projects.

## 5.1 Results from prior funding

The PI for this grant, Crowston, has been funded by several additional NSF grants within the past 48 months. The grants most relevant to the current proposal are HSD 05–27457 ($684,882, 2005–2009, with R. Heckman, E. Liddy and N. McCracken), *Investigating the Dynamics of Free/Libre Open Source Software Development Teams*, IIS 04–14468 ($327,026, 2004–2006) and SGER IIS 03–41475 ($12,052, 2003–2004), both entitled *Effective work practices for Open Source Software development* and CNS Grant 07–08437 ($200,000, 2007–2010, with M. Conklin, Elon University), for *Collaborative Research: CRI: CRD: Data and analysis archive for research on Free and Open Source Software and its development*. The first three of these grants have supported a study of the evolution of effective work practices for free/libre open source software (FLOSS) projects. Findings from these grants included a taxonomy of success measures for FLOSS projects, evidence about the structure and dynamics of projects and descriptions of key practices, e.g., for decision making. The final grant, still ongoing, is supporting the development of cyberinfrastructure to support the FLOSS research community more broadly, including a catalog of FLOSS research-related resources (`http://flosshub.org`), data (`http://ossmole.sf.net/`) and working papers (`http://flosspapers.org/`), as well as development of example workflows replicating key FLOSS papers [62]. We are also registering the FLOSSmole datasets with Digital Object Identifiers (DOIs) to facilitate citation and reuse. Overall, this work has resulted in nine journal papers [25, 26, 27, 28, 29, 33, 35, 36, 61], a book chapter [23] and multiple conference papers [23, 24, 30, 31, 32, 34, 37, 37, 56, 57, 58, 62, 62, 74, 94, 95, 95, 105, 106, 107]. These grants have supported a total of six PhD students; several other PhD and masters students have been involved in specific aspects of the work.

# References


[1] B. Thomas Adler, Luca de Alfaro, Ian Pye, and Vishwanath Raman. Measuring author contributions to the Wikipedia. In *WikiSym*, 2008.

[2] M.K. Ahuja and K.M. Carley. Network structure in virtual organizations. *Organization Science*, pages 741–757, 1999.

[3] ATC Citizen Science Working Group. Citizen science: Overview of the basics. Technical report, September 2006. URL `http://sain.utk.edu/apptrail_documents/Citizen_Science_Report_20070629.pdf`.

[4] Heidi L. Ballard, David Pilz, Eric T. Jones, and Christy Getz. *Training Curriculum for Scientists and Managers: Broadening Participation in Biological Monitoring*. Institute for Culture and Ecology, Corvalis, OR, 2005.

[5] Claudia Baretto, David E. Fastovsky, and Peter M. Sheehan. A model for integrating the public into scientific research. *Journal of Geoscience Education*, 50(1):71–75, January 2003 2003.

[6] M. W. Bauer, K. Petkova, and P. Boyadjieva. Public knowledge of and attitudes to science: Alternative measures that may end the "science war". *Science Technology Human Values*, 25 (1):30–51, Jan 01 2000.

[7] F. Berman and H Brady. Final report: NSF SBE-CISE workshop on cyberinfrastructure and the social sciences. Available at www.sdsc.edu/sbe/. Technical report, National Science Foundation, 2005.

[8] Y. Bhattacharjee. Ornithology. Citizen scientists supplement work of Cornell researchers. *Science*, 308(5727):1402–1403, Jun 3 2005.

[9] R. Bonney and M. LaBranche. Citizen science: Involving the public in research. *ASTC Dimensions*, page 13, May/June 2004.

[10] R. Bonney and J. L. Shirk. Citizen science central. *Connect*, pages 8–10, March 2007 2007.

[11] C. Brewer. Outreach and partnership programs for conservation education where endangered species conservation and research occur. *Conservation Biology*, 16(1):4–6, February 2002 2002.

[12] D. Brossard, B. Lewenstein, R., and Bonney. Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education.*, 27(9):1099–1121, 2005.

[13] Helen Bussell and Deborah Forbes. Understanding the volunteer market: The what, where, who and why of volunteering. *International Journal of Nonprofit and Voluntary Sector Marketing*, 7(3), 2002.

[14] I. Byron and A. Curtis. Maintaining volunteer commitment to local watershed initiatives. *Environmental management*, 30(1):59–67, 2002.

[15] R. Caruana, M. Elhawary, A. Munson, M. Riedewald, D. Sorokina, D. Fink, W.M. Hochachka, and S. Kelling. Mining citizen science data to predict prevalence of wild bird species. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 909–915. ACM New York, NY, USA, 2006.

[16] Marcelo Cataldo, Patrick A. Wagstrom, James D. Herbsleb, and Kathleen M. Carley. Identification of coordination requirements: Implications for the design of collaboration and awareness tools. In *Proceedings of the 20th anniversary conference on Computer-supported cooperative work (CSCW '06)*, pages 353–362, Banff, Alberta, Canada, 2006. ACM. ISBN 1-59593-249-6.

[17] A. Cho and D. Clery. International year of astronomy: Astronomy hits the big time. *Science*, 323(5912):332, 2009.

[18] J. Christie. Volunteer attitudes and motivations: Research findings and their application for sustainable community involvement programs in natural resource management. *Effective Sustainability Education: What Works? Why? Where Next? Linking Research and Practice*, 2004.

[19] F. Clark and D.L. Illman. Dimensions of civic science: Introductory essay. *Science communication*, 23(1):5, 2001.

[20] Ram A. Cnaan and Toni A. Cascio. Performance and commitment: Issues in management of volunteers in human service organizations. *Journal of Social Service Research*, 24:1–38, 1999.

[21] Jeffrey P. Cohn. Citizen science: Can volunteers do real research? *BioScience*, 58(3):192–107, March 2008 2008.

[22] Caren B. Cooper, Janis Dickinson, Tina Phillips, and Rick Bonney. Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society*, 12(2), 2007.

[23] Kevin Crowston. The bug fixing process in proprietary and free/libre open source software: A coordination theory analysis. In Varun Grover and M. Lynne Markus, editors, *Business Process Transformation*. M. E. Sharpe, Armonk, NY, 2008.

[24] Kevin Crowston and Isabelle Fagnot. The motivational arc of massive virtual collaboration. In *IFIP WG 9.5 Working Conference on Virtuality and Society: Massive Virtual Communities*, Lüneberg, Germany, 1–2 July 2008.

[25] Kevin Crowston and James Howison. The social structure of free and open source software development. *First Monday*, 10(2), 2005.

[26] Kevin Crowston and James Howison. Hierarchy and centralization in free and open source software team communications. *Knowledge, Technology & Policy*, 18(4):65–85, 2006.

[27] Kevin Crowston and James Howison. Assessing the health of open source communities. *IEEE Computer*, 39(5):89–91, May 2006.

[28] Kevin Crowston and Barbara Scozzi. Open source software projects as virtual organizations: Competency rallying for software development. *IEE Proceedings Software*, 149(1):3–17, 2002.

[29] Kevin Crowston and Barbara Scozzi. Coordination practices within free/libre open source software development teams: The bug fixing process. *Journal of Database Management*, 19 (2):1–30, 2008.

[30] Kevin Crowston, Hala Annabi, and James Howison. Defining open source software project success. In *Proceedings of the 24th International Conference on Information Systems*, Seattle, WA, 2003.

[31] Kevin Crowston, Hala Annabi, James Howison, and Chengetai Masango. Effective work practices for FLOSS development: A model and propositions. In *the Hawai'i International Conference on System Science (HICSS)*, Big Island, Hawai'i, 2005.

[32] Kevin Crowston, Kangning Wei, Qing Li, U. Yeliz Eseryel, and James Howison. Coordination of free/libre open source software development. In *Proceedings of the International Conference on Information Systems*, Las Vegas, NV, USA, 2005.

[33] Kevin Crowston, James Howison, and Hala Annabi. Information systems success in free and open source software development: Theory and measures. *Software Process–Improvement and Practice*, 11(2):123–148, 2006.

[34] Kevin Crowston, Kangning Wei, Qing Li, and James Howison. Core and periphery in free/libre and open source software team communications. In *Hawai'i International Conference on System System (HICSS-39)*, Kaua'i, Hawai'i, 2006.

[35] Kevin Crowston, James Howison, Chengetai Masango, and U. Yeliz Eseryel. The role of face-to-face meetings in technology-supported self-organizing distributed teams. *IEEE Transactions on Professional Communications*, 50(3):185–203, 2007.

[36] Kevin Crowston, Kangning Wei, Qing Li, U. Yeliz Eseryel, and James Howison. Self-organization of teams in free/libre open source software development. *Information and Software Technology Journal, Special issue on Understanding the Social Side of Software Engineering, Qualitative Software Engineering Research*, 49:564–575, 2007.

[37] Kevin Crowston, James Howison, and Andrea Wiggins. Opportunities for eScience research on free/libre open source software. In *Proceedings of the Oxford e-Research 08 Conference*, Oxford, England, 11-13 September 2008.

[38] Kevin Crowston, Kangning Wei, James Howison, and Andrea Wiggins. Free/libre open source software: What we know and what we do not know. *ACM Computing Surveys*, Under review.

[39] Shane R. de Solla, Leonard J. Shirose, Kim J. Fernie, Glenn C. Barrett, Chris S. Brousseau, and Christine A. Bishop. Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. *Biological Conservation*, 121:585–594, 2005.

[40] Derek J. de Solla Price. *Little Science, Big Science*. Columbia University, New York, 1963.

[41] A. A. Dhondt, S. Altizer, E. G. Cooch, A. K. Davis, A. Dobson, M. J. Driscoll, B. K. Hartup, D. M. Hawley, W. M. Hochachka, P. R. Hosseini, C. S. Jennelle, G. V. Kollias, D. H. Ley, E. C. Swarthout, and K. V. Sydenstricker. Dynamics of a novel pathogen in an avian host: Mycoplasmal conjunctivitis in house finches. *Acta Tropica*, 94(1):77–93, Apr 2005.

[42] Caroline Dufour and Elizabeth Crisfield, editors. *The Appalachian Trail MEGA-Transect*, Harpers Ferry, WV, 2008. Appalachian Trail Conservancy.

[43] Steven M. Farmer and Donald B. Fedor. Volunteer participation and withdrawal. *Nonprofit Management and Leadership*, 9(4), 1999.

[44] Thomas A. Finholt. Collaboratories. In B. Cronin, editor, *Annual Review of Information Science and Technology*, volume 36. Information Today, Medford, NJ, 2001.

[45] Thomas A. Finholt. Collaboratories as a new form of scientific organization. *Economics of Innovation and New Technology*, 12(1):5–25, 2003.

[46] Karen Firehock and Jay West. A brief history of volunteer biological water monitoring using macroinvertebrates. *Journal of the North American Benthological Society*, 14(1):197–202, 1995 1995.

[47] Leska S. Fore, Kit Paulsen, and Kate O'Laughlin. Assessing the performance of volunteers in monitoring streams. *Freshwater Biology*, 46:109–123, 2001.

[48] Andrea Forte and Amy Bruckman. Scaling consensus: Increasing decentralization in Wikipedia governance. In *Hawai'i International Conference on System Science (HICSS-41)*, Big Island of Hawai'i, 2008. IEEE.

[49] Peter A. Freeman, Deborah L. Crawford, Sangtae Kim, and José L. Muñoz. Cyberinfrastructure for science and engineering: Promises and challenges. *Proceedings of the IEEE*, 93(3): 682–691, 2005.

[50] C. Wayne Gordon and Nicholas Babchuk. A typology of voluntary associations. *American Sociological Review*, 24(1):22–29, 1959.

[51] V. C. Gulick, G. Deardorff, B. Kanefsky, and A. Davatzes. Student and public participation in acquiring and analyzing hirizse images. In *Lunar and Planetary Science XXXVIII*, March 2007.

[52] R. A. Guzzo and M. W. Dickson. Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, 47:307–338, 1996.

[53] J. Richard Hackman. The design of work teams. In Jay W. Lorsch, editor, *The Handbook of Organizational Behavior*, pages 315–342. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[54] J. Richard Hackman and Charles G. Morris. Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz, editor, *Group Processes*, volume 8 of *Advances in Experimental Social Psychology*, pages 45–99. Academic Press, New York, 1978.

[55] J.S. Hange, B.S. Seevers, and D. VanLeeuwen. *4-H Youth Development Extension Agents' Attitudes Towards Volunteer Management Competencies*. New Mexico State University, 2002.

[56] Robert Heckman, Kevin Crowston, Qing Li, Eileen E. Allen, Yeliz Eseryel, James Howison, and Kangning Wei. Emergent decision-making practices in technology-supported self-organizing distributed teams. In *the International Conference on Information Systems (ICIS 2006)*, Milwaukee, WI, 10–13 Dec, 2006.

[57] Robert Heckman, Kevin Crowston, U. Yeliz Eseryel, James Howison, Eileen Allen, and Qing Li. Emergent decision-making practices in free/libre open source software (FLOSS) development teams. In *3rd International Conference on Open Source Software*, Limerick, Ireland, 2007.

[58] Robert Heckman, Kevin Crowston, and Nora Misiolek. A structurational perspective on leadership in virtual teams. In Kevin Crowston and Sandra Seiber, editors, *the IFIP Working Group 8.2/9.5 Working Conference on Virtuality and Virtualization*, pages 151–168, Portland, OR, 2007. Springer.

[59] T. Hey and A. E. Trefethen. Cyberinfrastructure for e-science. *Science*, 308(5723):817–821, 2005.

[60] W.M. Hochachka, R. Caruana, D. Fink, A. Munson, M. Riedewald, D. Sorokina, and S. Kelling. Data-mining discovery of pattern and process in ecological systems. *Journal of Wildlife Management*, 71(7):2427–2437, 2007.

[61] James Howison, Megan Conklin, and Kevin Crowston. FLOSSmole: A collaborative repository for FLOSS research data and analyses. *International Journal of Information Technology and Web Engineering*, 1(3):17–26, 2006.

[62] James Howison, Andrea Wiggins, and Kevin Crowston. eResearch workflows for studying free and open source software development. In *Fourth International Conference on Open Source Software (IFIP 2.13)*, Milan, Italy, 7-10 September 2008.

[63] Daniel R. Ilgen, John R. Hollenbeck, Michael Johnson, and Dustin Jundt. Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, 56:517–543, 2005.

[64] D.R. Ilgen and J.R. Hollenbeck. The structure of work: Job design and roles. *Handbook of industrial and organizational psychology*, 2:165–207, 1991.

[65] Sirkka L. Jarvenpaa and D. E. Leidner. Communication and trust in global virtual teams. *Organization Science*, 10(6):791–815, 1999.

[66] E.C. Kasper-Fuehrera and N.M. Ashkanasy. Communicating trustworthiness and building trust in interorganizational virtual organizations. *Journal of Management*, 27(3):235, 2001.

[67] Aniket Kittur, Ed H. Chi, and Bongwon Suh. Crowdsourcing user studies with Mechanical Turk. In *ACM Conference on Human-factors in Computing Systems*, pages 453–456, Florence, Italy, 2008. ACM New York, NY, USA.

[68] Richard. T. Kouzes, James. D. Myers, and William A. Wulf. Collaboratories: Doing science on the Internet. *Computer*, 29(8):40–46, 1996.

[69] Marianne Krasny and R. Bonney. Environmental education through citizen science and participatory action research. In Edward Arnold Johnson and Michael Mappin, editors, *Environmental Education and Advocacy*, chapter 13. Cambridge University Press, New York, 2005. ISBN 0521824109 (hbk.).

[70] R.N. Langlois and G. Garzarelli. Of hackers and hairdressers: Modularity and the organizational economics of open-source collaboration. *Industry & Innovation*, 15(2):125–143, 2008.

[71] A. Lawrence. 'No personal motive?' Volunteers, biodiversity, and the false dichotomies of participation. *Ethics, Place & Environment*, 9(3):279–298, 2006.

[72] Tracy Lee, Michael S. Quinn, and Danah Duke. Citizen, science, highways, and wildlife: Using a web-based GIS to engage citizens in collecting wildlife information. *Ecology and Society*, 11(1):11, 2006.

[73] B. Lewenstein, R. Bonney, D., and Brossard. Citizen science: Measuring scientific knowledge and attitudes in a new type of public outreach project. Technical Report #98-1, Texas A&M University, Center for Science & Technology Policy & Ethics, 1998.

[74] Q. Li, R. Heckman, K. Crowston, J. Howison, E. Allen, and U. Y. Eseryel. Decision-making paths in technology-supported self-organizing distributed teams. In *International Conference on Information Systems*, Paris, France, 14-17 December 2008.

[75] Mechelle A. Marks, John E. Mathieu, and Stephen J. Zaccaro. A temporally based framework and taxonomy of team processes. *Academy of Management Review*, 26(3):356–376, 2001.

[76] M.L. Markus, B. Manville, and C.E. Agres. What makes a virtual organization work? *Sloan Management Review*, 42(1):13–26, 2000.

[77] Luis L. Martins, Lucy L. Gilson, and M. Travis Maynard. Virtual teams: What do we know and where do we go from here? *Journal of Management*, 30(6):805–835, 2004.

[78] M. L. Maznevski and K. M. Chudoba. Bridging space over time: Global virtual team dynamics and effectiveness. *Organization Science*, 11(5):473–492, 2000.

[79] R. E. McCaffrey. Using citizen science in urban bird studies. *Urban Habitats*, 3(1):70–86, December 2005 2005.

[80] J. McGrath and A. Hollingshead. *Groups Interacting With Technology.* Sage, Thousand Oaks, CA, 1994.

[81] M. M. Montoya-Weiss, A. P. Massey, and M. Song. Getting it together: Temporal coordination and conflict management in global virtual teams. *Academy of Management Journal*, 44 (6):1251–1262, 2001.

[82] V. Murray and Y. Harrison. The impact of information and communications technology on volunteer management. Technical report, Canadian Centre for Philanthropy, University of Victoria, Victoria, Canada, 2002.

[83] Oded Nov. What motivates wikipedians? *Commun. ACM*, 50(11):60–64, 2007. ISSN 0001-0782.

[84] NSF Cyberinfrastructure Council. Cyberinfrastructure vision for 21st century discovery. Report, US National Science Foundation, 2008.

[85] H. Onsrud and J. Campbell. Big opportunities in access to “Small Science” data. *Data Science Journal*, 6, 2007.

[86] H. Onsrud, G. Camara, J. Campbell, and N.S. Chakravarthy. Public commons of geographic data: Research and development challenges. *Lecture Notes in Computer Science*, 3234:223–238, 2004.

[87] Julian Orr. *Talking About Machines: An Ethnography of a Modern Job.* ILR Press, Ithaca, NY, 1996.

[88] Dawn Alexandra Osborn, John S. Pearse, and Christy A. Roe. Monitoring rocky intertidal shorelines: A role for the public in resource management. In Orville T. Magoon, Hugh Converse, Brian Baird, Beth Jines, and Melissa Miller-Henson, editors, *California and the World Ocean Äô02: Revisiting and Revising CaliforniaÄôs Ocean Agenda*, pages 624–636, Reston, VA, October 27-30, 2002 2002. ASCE.

[89] Christine Overdevest, Cailin Huyck Orr, and Kristine Stepenuck. Volunteer stream monitoring and local participation in natural resource issues. *Research in Human Ecology*, 11(2): 177–185, 2004.

[90] David Pilz, Heidi L. Ballard, and Eric T. Jones. *Broadening Participation in Biological Monitoring: Guidelines for Scientists and Managers*. Institute for Culture and Ecology, Corvalis, OR, 2005.

[91] D. Powazek. *Design for community: The art of connecting real people in virtual places*. New Riders Publishing Thousand Oaks, CA, USA, 2001.

[92] Jennifer Preece and Ben Shneiderman. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 1(1), 2009.

[93] M. Prysby and P. E. Super. *Citizen Science: A Best Practices Manual (draft)*. Association of Nature Center Administrators, 2006.

[94] M. Scialdone, N. L. Li, J. Howison, R. Heckman, and K. Crowston. Group maintenance in technology-supported distributed teams. In *Best Paper Proceedings, Academy of Management Annual Meeting*, Anaheim, CA, 8–13 August 2008.

[95] Barbara Scozzi, Kevin Crowston, U. Yeliz Eseryel, and Qing Li. Shared mental models among open source software developers. In *Hawai'i International Conference on System Science*, Big Island, Hawai'i, 2008.

[96] Seattle Aquarium. "Citizen science" science report. Technical report, Seattle Aquarium, 2005.

[97] M.T. Sergent and W.E. Sedlacek. Volunteer motivations across student organizations: A test of person-environment fit theory. *Journal of College Student Development*, 31(3):255–261, 1990.

[98] A. Sorokin and D. Forsyth. Utility data annotation with Amazon Mechanical Turk. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPR Workshops 2008*, pages 1–8, 2008.

[99] M. Spiro. What should the citizen know about science? *Journal of the Royal Society of Medicine*, 97(1):50, JAN 2004.

[100] L. Sproull and S. Kiesler. Public volunteer work on the Internet. *Transforming Enterprise: The Economic and Social Implications of Information Technology*, page 361, 2005.

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

[102] Deborah J. Trumbull, Rick Bonney, Derek Bascom, and Anna Cabral. Thinking scientifically during participation in a citizen-science project. *Science Education*, 84(2):265–275, 2000.

[103] Georg von Krogh, Sebastian Spaeth, and Karim R. Lakhani. Community, joining, and specialization in open source software innovation: A case study. *Research Policy*, 32(7):1217–1241, 2003.

[104] M. Wasko, S. Faraj, and R. Teigland. Collective action and knowledge contribution in electronic networks of practice. *Journal of the Association for Information Systems*, 5(11-12): 493–513, 2004.

[105] Andrea Wiggins, James Howison, and Kevin Crowston. Replication of FLOSS research as eResearch. In *Proceedings of the Oxford e-Research 08 Conference*, Oxford, England, 11–13 September 2008.

[106] Andrea Wiggins, James Howison, and Kevin Crowston. Social dynamics of FLOSS team communication across channels. In *Fourth International Conference on Open Source Software (IFIP 2.13)*, Milan, Italy, 7-10 September 2008.

[107] Andrea Wiggins, James Howison, and Kevin Crowston. Measuring potential user interest and active user base in FLOSS projects. In Cornelia Boldyreff, Kevin Crowston, Björn Lundell, and Tony Wasserman, editors, *Proceedings of The Fifth International Conference on Open Source Systems*, Skövde, Sweden, 2009. IFIP Working Group 2.13.

[108] C. C. Wilderman. Models of community science: Design lessons from the field. In C. McEver, R. Bonney, J. Dickinson, S. Kelling, K. Rosenberg, and J. L. Shirk, editors, *Citizen Science Toolkit Conference*, Cornell Laboratory of Ornithology, Ithaca, NY, 2007. URL `http://www.citizenscience.org`.

[109] A. Wilson and G. Pimm. The tyranny of the volunteer: The care and feeding of voluntary workforces. *Management Decision*, 34(4):24–40, 1996.

[110] Bo Xu, Donald R. Jones, and Bingjia Shao. Volunteers' involvement in online community based software development. *Information & Management*, 46(3):151 – 158, 2009. ISSN 0378-7206.