Citizen Science System Assemblages: Understanding the Technologies that Support Crowdsourced Science

Nathan R. Prestopnik
Syracuse University
337 Hinds Hall
Syracuse, NY 13244
315-443-5509
napresto@syr.edu

Kevin Crowston
Syracuse University
348 Hinds Hall
Syracuse, NY 13244
315-443-1676
crowston@syr.edu

ABSTRACT

We explore the nature of technologies to support citizen science, a method of inquiry that leverages the power of crowds to collect and analyze scientific data. We evaluate these technologies as system assemblages, collections of interrelated functionalities that support specific activities in pursuit of overall project goals. The notion of system assemblages helps us to explain how different citizen science platforms may be comprised of widely varying functionalities, yet still support relatively similar goals. Related concepts of build vs. buy and web satisfiers vs. web motivators are used to explore how different citizen science functionalities may lead to successful project outcomes. Four detailed case studies of current citizen science projects encompassing a cross-section of varying project sizes, resource levels, technologies, and approaches to inquiry help us to answer the following research questions: 1) What do typical system assemblages for citizen science look like? 2) What factors influence the composition of a system assemblage for citizen science? 3) What effect does the assemblage composition have on scientific goals, participant support, motivation, and satisfaction? and 4) What are the design implications for the system assemblage perspective on citizen science technologies?

Categories and Subject Descriptors
H.1.2 [**Information Systems**]: Models and Principles - Human factors, Human information processing, Software psychology

General Terms
Design Human Factors

Keywords
Citizen Science, System Assemblages, Socially Intelligent Computing, Web Technology.

1. INTRODUCTION

Citizen Science is a term used to describe research collaborations that enable non-scientist members of the public to assist with scientific investigations [3, 17]. Sometimes described as crowdsourced science, citizen science projects include (for example) those that ask participants to classify photographs of galaxies taken by space telescopes, report bird sighting data for ornithological research, or plant sunflowers and observe bee pollination at various times during the day or year.

While their scientific areas of inquiry vary widely, these projects adopt a common approach to data collection: large numbers of individuals are asked to submit individual observations or analyses via web- or mobile-based technologies. Despite this commonality, citizen science projects vary widely in the specific functionalities employed. Some projects utilize off-the-shelf software while others hire contract or in-house developers to build custom platforms. Supporting scientific goals is critical for virtually all citizen science projects, but some projects do this better than others, and some fall short in various ways.

No one functionality or technology can adequately support every one of a citizen science project's goals. Citizen science project developers ultimately mix and match, choosing different functionalities and technologies for different purposes. Furthermore, though project goals are often similar across citizen science projects, the specific functionalities and technologies that support them are often chosen for pragmatic reasons rather than scientific ones. To analyze this situation, we draw on the notion of a system assemblage [7-9].

Viewing the system as a *system assemblage*, a collection of interrelated functional components and social activities, is a valuable way of conceptualizing and understanding citizen science projects. The connections among the elements of the assemblage may be loose, as when a variety of off the shelf software systems are used separately to support various project goals, or tight, as in a project website that presents all functionalities (data collection tools, visualization components, blogs, news feeds, email, forums, project information, etc.) in a unified way. There is no ideal assemblage to aspire to; a multitude of parameters make each assemblage unique. Nonetheless, achieving a smoothly functioning system assemblage is a key requirement for any successful citizen science project.

Two theoretical dimensions are of particular interest in the system assemblage viewpoint we have adopted: “build vs. buy” and “motivation vs. satisfaction.” In forming a system assemblage to support citizen science, developers and scientists will inevitably encounter a fundamental question: should the system under development be built from scratch, or should it be composed of off-the-shelf components? Variables that enter into this decision include cost, development time, and technical proficiency of project members, and the answer will have important implications for the character of the final assemblage and for project goals. In any citizen science project, participation is also a key variable: high levels of participation will typically lead to more and more rapidly collected data, greater public exposure, increased project resources, and enhanced potential to develop new and more ambitious scientific goals. The composition of the system

Conference'12, February, 7-10 2012, Toronto, Ontario, Canada.

assemblage, as well as the way it was developed, will have repercussions for motivating participation, as well as effects on how satisfied users are with a given project. Both motivation and participation are deeply important to a citizen science project, since both can ultimately impact participation and project success.

We explore the system assemblages of four citizen science projects through two separate data collection efforts. The first is a functionality review of 27 current citizen science projects. The second is a more detailed comparative case study evaluation of four of these. In our review of 27 citizen science projects, we evaluate each one for the functionality it contains. For each case study, we evaluate the assemblage in terms of its overall composition, plus build vs. buy decision-making, motivation, and satisfaction. We address the following research questions: 1) What do typical system assemblages for citizen science look like? 2) What factors influence the composition of a system assemblage for citizen science? 3) What effect does the assemblage have on participant support, motivation, and satisfaction? and 4) What are the design implications for the system assemblage perspective on citizen science technologies?

2. BACKGROUND

2.1 System Assemblages

A basic feature of our analysis is consideration of the technology support for a citizen science project as an assemblage of loosely coupled systems rather than a single monolithic system. We draw from Kling's [7-9] approach to computer technology evaluation that he calls "web models." Web models "conceive of a computer system as an assemblage of equipment, applications, and techniques with identifiable information-processing capabilities" [8]. This approach is distinguished from "engineering models," or "discrete entity" analyses, which emphasize equipment and information processing capabilities only. Kling emphasizes the importance of human, social and organizational elements in these models [7-9], and we similarly see these elements as critical to a citizen science instantiation. We refer to the group of technologies, functionalities, social and communication elements, and organizational decisions developed to support citizen science as a "citizen science system assemblage."

Most citizen science system assemblages leverage multiple technologies to support multiple functionalities. Rather than deploying single, self-contained systems, citizen science projects are more likely to choose from a "menu" of functionalities. One project may opt to include a web forum to support participant collaboration, while another may view this functionality as unnecessary. Because of the social nature of citizen science, many of the functionalities that make up a citizen science assemblage have some communication role: a forum enhances communication between participants and project scientists, content pages provide information or standardize training, blogs and news feeds reach out to participants, and contact forms allow participants to address members of the project team.

The social nature of the citizen science system assemblage places people at the center of any citizen science project. Generating adequate levels of participation is a key goal for project scientists, since participation levels have direct effects on scientific goals. The citizen science system assemblage must therefore be developed in a manner that supports participation by motivating users to join the project, by satisfying users so that they continue to participate, or, ideally, both. Whether a system assemblage is built from off-the-shelf components or custom-built will have an important impact on how motivating and satisfying it is.

2.2 Build vs. Buy

One important organizational issue in developing a citizen science assemblage is the build vs. buy [4, 10] decision. Some take a "build" mindset, developing most of their system components in-house, thereby retaining a great deal of control over component functionality and the ability to integrate components. This degree of control usually comes at the cost of increased development time and more money spent on professional design expertise. Forgoing this expertise may save on expenses, but can cost a project in other ways. For example, projects developed by volunteer, non-expert labor may, depending on the capabilities of the volunteer, have limited or less reliable data collection instruments. Similarly, they may suffer through ineffective or unusable design, impacting participant motivation and satisfaction. Tailored approaches to development may also require that system developers spend a large portion of build time on features other than data collection; basic add, edit, and delete functionality for information pages, as well as participant management features that must often be built from scratch.

Some citizen science projects instead seek to leverage technologies developed by others, either by purchasing them or by finding free or open-source solutions. This "buy" mentality can result in inexpensive and high quality support for some assemblage components, particularly those common to many web applications like participant management, content management, or site navigation. Much basic web functionality comes prepackaged with bought systems, so developers can dedicate their time to creating data collection instruments and other novel features, rather than recreating industry-standard functionality like participant account or content management tools. However, bought systems sometimes place restrictions on the augmenting features that may be designed into them. Some systems are not open technology, making additions very difficult or even impossible; more often, such systems are either too complex or too limited in their capabilities, increasing the difficulty of appending advanced functionality onto them or seamlessly integrating the various assemblage components.

The build vs. buy decision shapes the final character of a system assemblage, particularly as it relates to the support of participants. Built systems tend to allow a level of flexibility and integration that applies well to unique features that may be highly motivating for participants. Bought systems tend to emphasize conformity to existing standards, and often address more standardized functionality such as managing user accounts, maintaining web content, and communicating with participants.

2.3 Satisfiers vs. Motivators

Because citizen science efforts are predicated on the contributions of a large number of participants, motivation is a core issue. If too few participants are motivated to join a citizen science project or if too many lose interest, the project will fail. Motivation to participate is multi-faceted: participants join citizen science projects for a variety of reasons, including their interest in the subject of scientific inquiry, the relevance of data collection efforts to particular interests or hobbies, how inherently fun a project is, or for altruistic reasons [2, 6, 12, 13, 16].

Zhang and von Dran [18] have argued that website features can be grouped into two distinct categories: *satisfiers* and *motivators*. Satisfying features are those whose absence will cause a participant to experience dissatisfaction with a website. Motivating features are those which add value to a website. Participants expect websites to maintain a certain level of quality, and substantial reduction of that quality will cause dissatisfaction.

At the same time, the motivation to use and revisit a website comes from a potentially different set of features; in other words, “being usable and being likable are two different goals,” [18].

In their research, Zhang and von Dran [18] identified a variety of common web features and clustered them into satisfier and motivator categories. Satisfiers include participant account and security features, basic usability, cross-browser compatibility, up-to-date information, and the quality of informational content. The absence of such features will discourage participation, so their presence is simply a requirement. Motivating features include those which produce cognitive outcomes such as learning new knowledge or skills, the fun of exploring a site, the presence of multimedia, level of challenge, the ability to control interaction, and the presence of eye-catching visuals. The presence of such features is not required for basic use, but can encourage further usage. Zhang and von Dran's [18] research did not specifically look at the citizen science context. However, the concepts of satisfier and motivator are relevant to citizen science, as good usability and highly motivating design are important to science and participation goals. Citizen science websites usually address highly specific scientific interests, so these projects must convince average citizens that site content and activities are interesting, enjoyable, and easy to use; failure to attract or retain participants will also result in the failure of ultimate scientific goals. This is to say that citizen science projects must meet a high bar in terms of both satisfiers and motivational features.

2.4 Build for Motivation, Buy for Satisfaction

In [10] the build vs. buy decision is summarized as follows: “Build business value, buy the basics.” In a business setting, basics might include things like HR tools [10], but for a citizen science system assemblage, basics will be those elements required to keep the system running and functioning smoothly: satisfier elements such as user account tools, basic usability, or security features. On the other hand, in a citizen science system assemblage, “business value” will include the ability to motivate participation and ensure that adequate scientific data is collected. Furthermore, [4] suggests that cost, time, specifications, and resources will be key factors in the build vs. buy decision. The low cost and ability to rapidly implement will make buy decision-making attractive to projects with lower resources, but highly specific project requirements may push an assemblage toward built components. That is, lower resourced projects may tend to favor bought solutions unless science goals mandate a more tailored approach.

Accordingly, we suggest that projects where a build mindset is adopted will have more unique or custom functionalities, and that these will favor motivational elements over satisfiers. Attention on these projects will be directed toward science support features particularly, as these offer the most opportunity for the kinds of innovation favored by build approaches. Such features might include highly visual interfaces, games or game-like features, social functionality, high scores, participant statistics, or data visualizations. We further suggest that projects where a buy mindset is adopted will have correspondingly strong satisfier functionalities, but may see less adoption of motivator functionalities if project resources force constraints on tailored development. In general, we posit that for a given citizen science system assemblage, a build mentality will generate more motivational elements, while a buy mentality will generate more satisfier elements.

3. CITIZEN SCIENCE WEBSITE REVIEW

Twenty-seven separate citizen science websites were reviewed for the features and functionalities that they contained. These websites came from the sample developed by Wiggins and Crowston [17] using the landscape sampling method described in [1]. A list of key functionalities was iteratively generated by visiting each of the 27 citizen science websites and comprehensively evaluating their publically available content. The resulting list of 33 functionalities included items like “project information,” “forums,” “registration forms,” “submit text data,” blogs,” etc. Websites were flagged as either containing or lacking these features.

Identified features were organized into a continuum from virtually ubiquitous to highly optional. The following table shows this feature set, organized by overall count within the 27 reviewed projects:

Table 1. Citizen Science Website Features

Citizen Science Features	Count
Submit Data	27
Instructions	26
Project Information	26
Contact Information	23
Scientific Information	21
Registration	20
Collect Participant Information	18
Education	18
Project Data	17
Affiliates and Sponsors	16
Team/Staff Information	16
FAQ	15
Links	13
Sign In	13
News Feed	12
Blogs	11
Alerts	9
Forum	9
Donate	8
Practice and Testing	7
Email List	5
Participant Scores and Stats	5
Photo/Image Gallery	5
Published Papers	5
Calendar	2
Customized Participant Experience	2
Page Translation	2
Sales/Store	2
Contests	1

This list contains a mix of satisfier and motivator functionalities, with satisfiers being particularly well represented. To more deeply evaluate the composition of citizen science system assemblages, we developed detailed comparative case studies on four of the projects in this review.

4. CASE STUDIES

We used a qualitative case study approach to explore the relationship between satisfiers vs. motivators and build vs. buy in citizen science system assemblages. Case study projects included *eBird*, *Galaxy Zoo*, the *Great Sunflower Project*, and *What's Invasive*, citizen science undertakings in the scientific domains of ornithology, astronomy, apiology, and phenology respectively. The chosen case sites span a range of technological sophistication and development approaches.

Scientists and system developers were interviewed about project histories, scientific and outreach goals, supporting technologies, and the citizen science phenomenon as it relates to their particular research efforts. Ten separate interviews of nine scientists and system developers were conducted in total. Interviews lasted for approximately one hour each. The websites and mobile applications for the four projects were also scrutinized in detail for their specific functionalities and design approaches.

5. CASE 1: *eBIRD*

eBird is joint project of the Cornell Lab of Ornithology and National Audubon Society. It is a citizen science project that engages a global network of birdwatchers to submit their observations (2 million per month) via web-based tools.

5.1 Build vs. Buy in *eBird*

The current instantiation of *eBird* is a hybrid of build vs. buy, with some content portions of the project website running off a CMS system, while data collection interfaces and the database itself are built in-house and are highly customized. The hybrid approach has helped *eBird* to address the various challenges associated with the development of a citizen science system assemblage. Using off the shelf CMS software allows for easy management of basic content, freeing developer time for more complex challenges in database design, data collection, and providing feedback to participants. *eBird* has adopted the strategy that basic functions can be bought, while more advanced tools should be custom.

eBird supports its scientific goals through tailored applications that allow participants to submit locations and bird observations. These tools also give participants the ability to view their prior submissions, maintain "life lists" of birds, and organize their work in various ways. The built tools on *eBird* support a variety of motivational functions. However, the CMS system is targeted at basic participant functionality, such as communicating with participants and organizing site content.

5.2 Motivation vs. Satisfaction in *eBird*

eBird leverages an existing hobby—birding—to capture and retain a pool of willing and motivated participants. Thus, a high base level of motivation is already present among participants of the system. Many citizen science projects attempt to turn mundane tasks into fun activities through games or other mechanisms. *eBird* takes the opposite stance, using an already fun activity that is enjoyed by millions to generate large amounts of scientific data. Capitalizing on this built-in motivational effect, the *eBird* website uses custom interfaces to provide participants with exactly the kinds of tools they desire for viewing, visualizing, and exploring the data that they and others have submitted. *eBird* is a highly motivating web experience for birders because custom development has made it their ideal, and this approach was highly intentional. According to one *eBird* developer, "I think, at least all of us are feeling like we would get a lot more mileage out of giving people, giving our users actual tools to play around with their data more, which is outside of sort of content management stuff…. We're not going to find a CMS that's going to be able to do that sort of stuff."

At the same time, the *eBird* designers have also emphasized satisfaction oriented features such as informative content, adequate privacy management, and well-designed functionality that works properly and as expected. This need for participant satisfaction has been emphasized both in their own custom design work and in the CMS system they use to augment it.

6. CASE 2: *GALAXY ZOO*

Galaxy Zoo is a citizen science project that has individuals undertake online annotation of millions of astronomical photographs collected by the Hubble Space Telescope, Sloan Digital Sky Survey telescopes, and others. *Galaxy Zoo* asks individuals to identify various galaxy features (the number of spiral arms, how round or elliptical they are, etc.) in these photographs, and this information is used to paint a more detailed picture of the universe we live in. *Galaxy Zoo* is a project within the *Zooniverse*, a collection of citizen science projects maintained and developed by the *Citizen Science Alliance*.

6.1 Build vs. Buy in *Galaxy Zoo*

Most *Zooniverse* projects are annotation based; instead of having participants go into the field to collect and report observations, these projects have participants work with a computer interface to generate metadata (annotations) about assets (photographs of galaxies, scans of old ship's logs, photos of moon craters, etc.). Specialized interface functionality is thus required to present the participant with an asset and capture the user-generated metadata. *Galaxy Zoo* uses a visual decision tree format to elicit information from participants. The project *Moon Zoo* has participants draw outlines around moon craters. *Old Weather* asks participants to highlight sections of a scanned ship's log and type in the information it contains. Each of these systems has a unique interface, though other web elements such as user account management are similar across projects.

To support the specialized interface, *Zooniverse* system assemblages are more integrated than most other citizen science efforts. They are built, rather than leveraging off the shelf software, and their designers spend a great deal of time on the design process. The build approach allows *Zooniverse* system assemblages to feel completely integrated at the website level, even as some features are used on multiple websites while others are deployed only in one place. This approach requires that the *Zooniverse* design team spend extra effort on basic usability, management of content, and management of users.

The build decision is directly related to creating a highly motivating experience for participants. According to one developer, "We had pretty grand designs when I joined, and we still do.... *Galaxy Zoo 2* was the first of a new breed for us. We're now eight or nine projects down since then. I think going the bespoke route allowed us to really be pretty opinioned about what we thought was the right approach.... Trying to run a *Drupal* CMS and just kind of hacking in some code to make a decision tree appear on the page just wouldn't have felt right. It wasn't grand enough for our plans."

Traditional programming practice dictates that code be reused as much as possible, to save time on new development. Rather than sharing code from prebuilt tools, an extensive code library was developed during the *Galaxy Zoo* implementation that can be used in subsequent projects. However, the code library became a source of tradeoffs very similar to those found in the build vs. buy decision. In essence the code library became a "bought" system that was complex and sometimes time consuming for newly hired programmers to use effectively. It offered a great deal of functionality, but sometimes more than was needed. Interestingly, it was difficult for newer developers to gain a sense of ownership over their work when the code they were working with came largely from this library.

The *Zooniverse* design team adopted a new development paradigm, where previous code is retained or centralized for identical or highly similar tasks (i.e. satisfier functionalities such as participant management), but each *Zooniverse* site is thereafter approached like an independent design project. This approach allows developers to direct their time toward developing motivating functionalities that support scientific goals, while simultaneously ensuring that previously developed and successful functionalities with a satisfier role are effectively reused.

6.2 Motivation vs. Satisfaction in *Galaxy Zoo*

One important motivational element identified in the *Galaxy Zoo* case study is the notion of "rules for ethical citizen science." These rules suggest that citizen science participants 1) be treated as collaborators, 2) not have their time wasted, and 3) not be asked to undertake tasks that could be better undertaken by a computer. These rules have influenced many of the components included on *Zooniverse* system assemblages, including tools for participant collaboration, web sections to release scientific data and papers, and acknowledgements and the assignment of co-authorship on various academic papers. These rules impact the kinds of data that are collected and, notably, the custom interfaces by which they are submitted.

The rules for ethical citizen science stem from a desire for scientific outreach and education. *Galaxy Zoo* and the other *Zooniverse* projects have a core desire to teach participants what it is to do science. This outreach effort emphasizes scientific process—the scientific method and how scientific inquiry, regardless of topic, is conducted. Such outreach and education efforts are considered motivating by *Galaxy Zoo* participants [13]. The motivational aspects of outreach are compounded by tailored tool design to support independent scientific inquiry outside the aegis of the established *Galaxy Zoo* research questions. A system developed for *Moon Zoo* and intended for use on other *Zooniverse* sites allows participants to collaborate with each other on independent lines of scientific inquiry.

7. CASE 3: *GREAT SUNFLOWER PROJECT*

The *Great Sunflower Project* is a citizen science project to explore how ecosystem services like pollination are affected by environmental factors. Participants in this project use a *Drupal*-based website to submit information about sunflower gardens that they have established, as well as the number of times bees visit sunflowers within a given a period of time. Aggregated, this data shows where bees are providing good pollination service, as well as places where pollination may be suffering.

7.1 Build vs. Buy in *Great Sunflower Project*

Limited project resources were highly influential in the build vs. buy decision for the *Great Sunflower Project*. While other case study projects in this research had the resources to attempt custom development, the *Great Sunflower Project* adopted a buy mentality and used the open source CMS *Drupal* to deploy a system assemblage at much lower cost. The developer for the *Great Sunflower Project* was highly supportive of open-source, CMS-based development generally, saying, "When you embrace a project like a content management system, you get the benefit of years of different people working on different sites, and all the experience behind that built into your project. Frankly, I would not embark upon a project like this building it from scratch."

Data collection for the *Great Sunflower Project* takes place through forms programmed by its developer within the *Drupal* framework. Visual design was kept as simple as possible because of the expense of custom artwork. The *Great Sunflower Project* has also deployed only those functionalities that are required to support a narrowly defined set of project goals; resources and personnel are not available to support a wide variety of functionalities that may not provide much additional value to the project. This has had the effect of making technology-based support for *Great Sunflower Project* participants both inexpensive and easy to deploy. The *Drupal* CMS handles user account management, website content, and a host of other support tasks. A free forum is used to manage participant communication. Email and web comments are also used for some communication tasks. As in the other case studies, the *Great Sunflower Project* incurred most of its development-related expenses in the support of built systems, namely in hiring a programmer to develop the forms used to collect data, where custom development was required.

7.2 Motivation vs. Satisfaction in *Great Sunflower Project*

Despite adopting a buy mentality and correspondingly lacking very many motivational elements, the *Great Sunflower Project* teaches that designing for participant satisfaction is deeply worthwhile, especially in a citizen science system assemblage where improved usability can lead to beneficial outcomes: heightened participant retention and improved data validity. The *Great Sunflower Project's* programmer and its principal investigator concentrated on usability issues surrounding data collection forms and noted specific improvements in the quality of submitted data when those issues were resolved. The participant population for this project is predominantly in an older and less web-experienced age demographic, so good usability ensures that participants can find the information they need and easily navigate the project website. This has direct implications for participation in the project, which is high (approximately 80,000 registered users) despite resource limitations.

Designing satisfying interfaces even for simple tasks can be more challenging than it seems; principal investigators on citizen science projects rarely have expertise in information architecture. They are often unsure how to create the easiest, simplest, and most usable experience for their participants. This places project staff in the awkward position of rapidly learning a discipline's worth of design knowledge that typically falls well outside their field of expertise. Learning curves can be steep, and avoidable mistakes are often made. This is particularly problematic since in a citizen science system assemblage, if usability is not at its best, data may be called into question and participation may suffer.

Adopting a buy mentality helped the *Great Sunflower Project* to address these challenges, as off the shelf software ensured good assemblage usability. CMS systems and web forms benefit from many product development cycles and standardized web form specifications issued by the W3C. The *Great Sunflower Project* has leveraged such forms into a highly successful and highly cost effective approach to data collection.

8. CASE 4: *WHAT'S INVASIVE*

What's Invasive is a citizen science project to collect information about invasive species developed by the *Center for Embedded Network Sensing* (CENS) at UCLA. Participants in this project use mobile devices to collect and submit information about the species they are observing, particularly their geographic location.

8.1 Build vs. Buy in *What's Invasive*

What's Invasive uses mobile technology and a website to introduce people to the project and allow visitors to view and visualize various data. Data from the mobile application is submitted to a database which feeds back to mobile devices and to the website. This assemblage was developed in the build tradition, and the website, database, and mobile component are all built in a custom manner. In fact, the CENS mission is to develop custom systems. According to one developer, "CENS has been innovating software for mobile devices through work by Computer Science and Electrical Engineering graduate and undergraduate students for the last five years or so." In this sense, the build vs. buy decision was simple for *What's Invasive*, since a build mentality is core to the group that developed the technology used in the system assemblage.

8.2 Motivation vs. Satisfaction in *What's Invasive*

The *What's Invasive* development team has directed more attention toward motivational elements than satisfying ones, in large part because of the emphasis on mobile technology in this project. Mobile technologies are still relatively novel, and therefore have the potential to be highly motivating. Motivational attributes are compounded by the wide variety of clever and engaging functionalities supported. However, factors such as network connectivity, battery life, GPS limitations, and environmental conditions require much planning and preparation ahead of time if mobile devices are to be used as a motivator in a system assemblage. To achieve their potential as a motivating tool for citizen science, mobile devices must be able to communicate effectively with databases and various other technologies despite these limits. In addition, adverse environmental conditions such as wet weather, cold, or heat can depress these devices' value as motivators. It is important to consider whether the scientific goals of a project may expose an individual's personal mobile device to damage, thereby inadvertently creating obstacles to participation.

Simply releasing a mobile application is not enough to motivate participants by itself. One early project developed by CENS, *HAB Watch* (intended to monitor harmful algal blooms), clearly demonstrated how an application that is well designed from a technical and scientific standpoint may still fall short if functionality to attract and retain participants is not included. *HAB Watch* concentrated entirely on data collection, but did not give participants other reasons for downloading or using the app. Vehicle parking information, news feeds, or games, for example, might have made this app more enticing to prospective users. The *What's Invasive* mobile application has a broader base of support and has not run into such serious difficulties, but motivating participation remains an important goal. Currently *What's Invasive* is undergoing updates to include game-like elements, scores, stats, and social activities for participants in an effort to enhance motivation and improve recruitment and retention of participants. Both the mobile application and the website are also undergoing usability and other satisfier-oriented improvements.

Games may be an especially powerful way to successfully invert the *eBird* approach of capitalizing on an existing activity. For projects with no hobbies to leverage, games can produce the same sense of fun and enjoyment as "real world" activities, while still successfully linking participants to data collection efforts. For example, *What's Invasive* doesn't have any inherent interest group associated with it other than the relatively small number of people who are genuinely enthusiastic about invasive species reduction and park ecology. However, a large pool of potential participants for *What's Invasive* may be found among park-goers who are interested in hiking or walking, but do not have any particular interest in or awareness of ecology, biology, or park preservation. Mobile platforms can support pervasive gaming, where a game is infused into the physical environment. Such games, if tied to a citizen science project like *What's Invasive*, could successfully wed outdoor enthusiasts to a citizen science project like *What's Invasive*. Just as with other motivational functionalities, mobile or other games virtually require a build mentality over a buy one.

9. DISCUSSION

9.1 Factors Influencing the Composition of System Assemblages

The build vs. buy paradigm has been discussed at length as one key influence on the composition of a system assemblage for citizen science. Project resources and practical concerns often dictate this decision [4, 10] and thus the final shape of the system assemblage. The principal investigator for the *Great Sunflower Project* related that the web developer for the project, a professional programmer in the *Drupal* community, is a family relation; thus his services were accessible and obtained at a reasonable price. The developer himself indicated that his preference is for open platforms like *Drupal*. However, *Drupal* was specifically adopted not because he "evangelizes" for it in particular, but because it is the system that he was most familiar with. The *Great Sunflower Project's* path from scientific questions to technology implementation is similar to many successful citizen science deployments, in which the technologies of the system assemblage are chosen for largely pragmatic or convenience-based reasons, while functionality is purposed around broader project goals.

Similarly, *Zooniverse* developers explain that when approached about designing new sites, they make pragmatic decisions about whether a project will be "done on the cheap" or "be a full project." This determination impacts the technologies that will be used, even if project goals are already defined.

For *What's Invasive* and other projects developed by CENS, a balance must be achieved between computer science students who seek development projects, and project scientists, who seek to collect environmental data using mobile technologies. Science and technology goals are not always complementary, so a fair amount of pragmatism is introduced into the project selection and development process to ensure that both computer science and environmental science goals are being met.

9.2 Composition of System Assemblages

Our case studies showed that participant motivation is often a key consideration during the design of a system assemblage. The question of how to get adequate participation is one that principal investigators and developers fixate upon. Possibly because of their interest in data, most scientists frame their concerns about motivation on data collection interfaces, which may lead them along more tailored, motivation-centered lines of thinking. However, this raises several problems to be avoided.

First, considering data collection tools as the only kind of viable motivator for a citizen science project can be misleading; in fact, potential participants may be far more motivated by other kinds of activities. For example, a long line of scholarly inquiry shows the potential for games and gaming as motivational tools in various contexts [e.g. 5, 11, 14, 15]. In the 27 websites we reviewed, few projects used games in any meaningful way to motivate participation. The possibility exists that games have been tried and have failed as citizen science motivators. However, projects like *Fold.It* (http://fold.it) use games as an effective motivator. Others, like *Stardust@Home* (http://stardustathome.ssl.berkeley.edu/) are able to motivate with individual scores and achievements similar to those found in games. For this reason, we believe it is more likely that games are an untapped resource for motivating participation in citizen science endeavors. In fact, it may be beneficial to implement an array of motivational systems within a citizen science assemblage, including games, visualization tools, meaningful feedback to users, and more. Some of these will directly impact data collection, and some will more generally create excitement and enthusiasm for continued participation.

A second problem to avoid is thinking that motivators are more important to consider than satisfiers. For two of our four case study projects, participants signed up much more rapidly and in greater numbers than expected. The *Great Sunflower Project* had 25,000 participants register in just two weeks. *Galaxy Zoo* saw such interest on its first day that the servers couldn't handle the load. This circumstance was the result of good publicity more than it was the result of specific motivational elements on either project's website, and it emphasizes the need for strong satisfier elements in addition to powerful motivators. Good publicity and well-designed and implemented motivators can turn heads and attract attention to a project, especially long term through word-of-mouth promotion. However, motivated participants will eventually leave if more basic satisfier requirements such as reliable functionality, easy-to-use navigation, and basic security are not fulfilled.

Fixation on motivators over satisfiers may result from relatively limited web design knowledge held by many project scientists. One experienced *Galaxy Zoo* developer made the interesting point when discussing project planning for new citizen science websites, that, "People typically think far more conservatively than you might imagine." This developer went on to explain that project scientists, with limited knowledge of the web's potential and its limitations, typically make one of two mistakes: they either assume that relatively achievable interactions with participants will be impossibly difficult, or they devise complex interactions that are not very feasible. It is rare for project scientists to truly understand what can be supported and what cannot. Thus, project scientists may worry that particularly motivating functionalities will be too complex to implement while forgetting that many satisfiers are well understood and routine to deploy. A more even-handed emphasis on both motivators and satisfiers will produce a balanced web experience that keeps participants engaged, and, importantly, does not drive them away.

This point is reflected in our review of 27 citizen science assemblages. The list we developed includes many satisfiers such as clear task instructions, project information, contact information, staff bios, registration forms, etc. From this, one may conclude that once a project is deployed and in use, project scientists and developers rapidly come to recognize the importance of satisfier elements and content, even if they are not contemplated as much as motivational features during early phases of design. For example, the principal investigator for the *Great Sunflower Project* states, "We are trying to think through what pages we would like to have and how to make them more accessible and apparent to participants," and the web developer on that project adds, "Simplifying forms lowers barriers to entry for someone who wants to participate." Though both of these individuals were worried about motivating participation earlier in the project timeline and, in fact, assumed participation would be relatively low, these later statements reflect a newfound concern for satisfiers as the *Great Sunflower Project* has progressed.

9.3 Design Implications

Citizen science projects have scientific goals that must be supported through technology. At the same time, these goals can become subordinate to a variety of factors, including project resources and the need for participants. Resource limitations force practical considerations to the forefront: to learn X, we must implement Y which will cost Z. Finding ways to implement Y at the lowest possible cost may involve searching for open-source or free software, using volunteer effort, or even revising scientific or project goals. It can be easy to believe that a cheaper component that is similar to Y may also do an adequate job of supporting research question X, but this is not always a safe assumption. Limited resources should be balanced against scientific goals to ensure that the science mission of a citizen science assemblage is not being compromised by tradeoffs in quality over cost.

Scientific outcomes are also impacted by participation. For science goals to be achieved, participants must be attracted to a project, motivated to involve themselves with it, and successfully managed so that their time is well spent. However, motivating, attracting, and retaining participants is a much larger design job than building a simple interface to collect data. Some functionalities support motivation, while others best support satisfaction. Planning an assemblage that can support both is highly important. Project scientists and developers should consider the full scope of a system assemblage to support citizen science, rather than concentrating only on its data collection components. Most of the project members we spoke with indicated that recruitment and managing participants occupies a much greater portion of their time than they had expected. Most had either already implemented or were seeking ways to implement various technologies to reduce the time spent on participant management duties or to more effectively attract participants to their project.

It is also important for project scientists and developers to acknowledge the importance of motivational and satisfier elements in a citizen science system assemblage. In the citizen science domain, motivating participation is valued; scientists who manage citizen science projects are very interested in understanding more about motivation and participation. However, there is much less overt enthusiasm for satisfier elements such as good usability, good organization, adequate privacy controls, or responsive communication to participants. These factors are recognized as necessary, but are sometimes undervalued in comparison to more motivational elements. They are not always planned for, and may be added over time as the realities of a project force their inclusion into the assemblage. The very fact that citizen science projects are system assemblages, however, means that emphasizing satisfiers is at least as important as emphasizing motivation. A poorly integrated assemblage that is difficult to use or has technical problems will likely fail no matter how motivating certain of its components are. Attracting participants to a project, only to have them grow unhappy and leave, provides little benefit to the project or participants and does nothing to serve science goals. We suggest that satisfier elements should be considered as early as possible, especially on projects with limited resources, as they constitute a healthy foundation upon which motivational components can eventually flourish.

Finally, consideration of the assemblage as a whole is important. Many project scientists and developers, especially those combining various bought or off-the-shelf systems, think deeply about specific functionalities, but miss the bigger picture of assemblage integration. Often integration is dictated by what specific technologies—a CMS, a visualization module, a data collection form—will allow. Not all technologies selected for inclusion in an assemblage come with serious constraints, but thinking early and often about how the components of an assemblage should work together will have positive effects on how easy the assemblage is to use and how well it satisfies and retains participants.

10. CONCLUSION

The system assemblage view of citizen science projects suggests that rather than understanding technological instantiations as seamlessly integrated units, it makes more sense to think of them as collections of discrete functionalities assembled to work together in a social context toward overarching project goals. Specific decisions about what functionalities and technologies to include in an assemblage are determined by a variety of factors, including project resources, practical concerns, the desire to build or buy, support for science, support for participants, motivation, and participant satisfaction.

Typically, functionalities—the various activities that an assemblage must support—are selected based on higher order concerns such as scientific goals or participant motivation. Technologies—the specific programming languages, databases, and software suites that support selected functionalities—are chosen for more pragmatic reasons such as cost, ability to support intended activities, availability, or how well they integrate with other assemblage components.

Science and participation are critical factors for a citizen science project, and a successful system assemblage in this domain should be shaped to adequately support both. Keeping motivation and satisfaction in mind during the design of the assemblage is one way to accomplish this end, as well executed interfaces to support scientific goals are often motivating while well designed participant support features help to create a satisfying participant experience. Frequently, motivation is thought of as preeminent, but satisfaction oriented functionality is at least as prevalent on established projects. This suggests that those who intend to begin a citizen science project would do well to think through issues of participant satisfaction: high quality information architecture and navigational elements, well organized content, and careful consideration of participant privacy.

Ultimately, there is no ideal citizen science system assemblage. Rather, individual projects will have different goals, different resources, and different needs. By contemplating a project's technological instantiation as an assemblage of interrelated functionalities, however, project scientists and developers may have a better sense of the effects of various functionalities and technologies, as well as ways to identify which to include and which to avoid.

11. ACKNOWLEDGMENTS

Our thanks to members of *eBird*, *Galaxy Zoo*, the *Great Sunflower Project*, and *What's Invasive* for their assistance. This work was partially support by NSF Grants VOSS 09–43049 and SOCS 09–68470.

12. REFERENCES

[1] Bos, N., et al., *From Shared Databases to Communities of Practice: A Taxonomy of Collaboratories.* Journal of Computer-Mediated Communication, 2007. 12(2): p. 652-672.

[2] Bradford, B.M. and G.D. Israel, *Evaluating Volunteer Motivation for Sea Turtle Conservation in Florida*. 2004, University of Florida, Agriculture Education and Communication Department, Institute of Agriculture and Food Sciences: Gainesville, FL. p. 372.

[3] Cohn, J.P., *Citizen Science: Can Volunteers Do Real Research?* BioScience, 2008. 58(3): p. 192-107.

[4] Fowler, K., *Build Versus Buy.* Instrumentation & Measurement Magazine, IEEE, 2004. 7(3): p. 67-73.

[5] Garris, R., R. Ahlers, and J. Driskell, *Games, Motivation, and Learning: A Research and Practice Model.* Simulation & Gaming, 2002. 33(4): p. 441-467.

[6] King, K. and C.V. Lynch, *The Motivation of Volunteers in the nature Conservancy - Ohio Chapter, a Non-Profit Environmental Organization.* Journal of Volunteer Administration, 1998. 16(5).

[7] Kling, R., *Defining the boundaries of computing across complex organizations.*, in *Critical issues in information systems*, R.B.a.R. Hirschheim, Editor. 1987, Wiley: London.

[8] Kling, R., *Computerization and Social Transformations.* Science, Technology & Human Values, 1991. 16(3): p. 342-367.

[9] Kling, R. and W. Scacchi, *The Web of Computing: Computer Technology as Social Organization*, in *Advances in Computers*. 1982, Elsevier. p. 1-90.

[10] Knowles, J., *Buy versus build in six words.* Datamation, 1997. 43(1): p. 31-31.

[11] Malone, T., W., *What makes things fun to learn? heuristics for designing instructional computer games*, in *Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems*. 1980, ACM: Palo Alto, California, United States.

[12] Raddick, M.J., et al., *Citizen science: status and research directions for the coming decade*, in *AGB Stars and Related Phenomenastro 2010: The Astronomy and Astrophysics Decadal Survey*. 2009. p. 46P.

[13] Raddick, M.J., et al., *Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers.* Astronomy Education Review, 2010. 9(1): p. 010103-18.

[14] von Ahn, L., *Games with a purpose.* Computer, 2006. 39(6): p. 92-94.

[15] von Ahn, L. and D. Laura, *Designing games with a purpose.* Commun. ACM, 2008. 51(8): p. 58-67.

[16] Wiggins, A. and K. Crowston, *Developing a conceptual model of virtual organizations for citizen science.* International Journal of Organizational Design and Engineering, 2010. 1(1/2): p. 148-162.

[17] Wiggins, A. and K. Crowston. *From Conservation to Crowdsourcing: A Typology of Citizen Science*. in *44th Hawaii International Conference on System Sciences*. 2011. Kauai, Hawaii.

[18] Zhang, P. and G.M. von Dran, *Satisfiers and dissatisfiers: A two-factor model for Website design and evaluation.* Journal of the American Society for Information Science and Technology, 2000. 51(14): p. 1253.