DIGITAL

From Tagging to Theorizing: Deepening Engagement with Cultural Heritage through Crowdsourcing

MIA RIDGE

Abstract Crowdsourcing, or "obtaining information or services by soliciting input from a large number of people," is becoming known for the impressive productivity of projects that ask the public to help transcribe, describe, locate, or categorize cultural heritage resources. This essay argues that crowdsourcing projects can also be a powerful platform for audience engagement with museums, offering truly deep and valuable connection with cultural heritage through online collaboration around shared goals or resources. It includes examples of well-designed crowdsourcing projects that provide platforms for deepening involvement with citizen history and citizen science; useful definitions of "engagement"; and evidence for why some activities help audiences interact with heritage and scientific material. It discusses projects with committed participants and considers the role of communities of participants in engaging participants more deeply.

INTRODUCTION

For museums, libraries, and archives with significant backlogs of un-digitized, under-catalogued, and un-researched material, crowdsourcing is a useful framework for inviting audiences to help with the resource-intensive tasks of creating or improving content about collections. There is plentiful evidence of the productivity of crowdsourcing in digitizing content, improving metadata, or identifying specimens from cultural heritage: see the lines of corrected text in Trove; pages transcribed in Old Weather; or animals identified in Serengeti (Proctor 2013; Romeo and Blaser 2011; Causer and Wallace 2012; Holley 2010; Kosmala 2013). Echoing themes from Nancy Proctor’s article “Crowdsourcing—An Introduction: From Public Goods to Public Good” in this journal (January 2013), I argue here that participation in crowdsourcing should also be recognized as a valuable form of public engagement with cultural heritage. Well-designed crowdsourcing projects can help meet the core missions of museums (Poole 2013) by connecting people, culture, history, and collections while providing the public with platforms for enjoyable, meaningful activity.

This article reviews the evidence for a relationship between active participation through crowdsourcing and engagement with cultural heritage objects and knowledge. Its inception lies a few years in the past, when I was inspired by the early steve.museum crowdsourcing project to make crowdsourcing games based on "difficult" objects from science and social history museum collections. While evaluating the games, I noticed that participants were talking about the objects after they finished playing the games (Ridge 2011b). I discovered that two art historian friends, who had never shown any interest in astronomy collections before joining this crowdsourcing activity, were discussing the difference between heliocentric and geocentric
astrolabes on Facebook. My curiosity was piqued. What was it about playing a crowdsourcing game that engaged them when other encounters with the collections had not? Then, in mid-June 2013, the *Oxford English Dictionary*—itself an example of proto-crowdsourcing—included the term “crowdsourcing” for the first time. The *OED* definition: “The practice of obtaining information or services by soliciting input from a large number of people, typically via the Internet and often without offering compensation.” This definition is adequate for commercial crowdsourcing, but, in implying tangible “compensation,” it underestimates the value for participants of engaging in cultural heritage crowdsourcing projects.

In this article, I provide a brief overview of content, tasks, motivations, and participants typical of crowdsourcing projects in galleries, libraries, archives, and museums (or “GLAMs”). I consider how crowdsourcing in museums differs from more traditional forms of user-generated content. I present some design techniques for encouraging initial, on-going, and more engaged participation; these are drawn from successful GLAM crowdsourcing projects, casual game design, and scaffolding theory. I review the evidence for crowdsourcing as a form of productive engagement with cultural heritage and consider the role of project structures and communities of practitioners in supporting the development of skills and deeper engagement with cultural heritage.

**DEFINING CROWDSOURCING IN CULTURAL HERITAGE**

“Crowdsourcing” was coined in 2006 when Jeff Howe and Mark Robinson riffed on the term “outsourcing” to describe “the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call” (Howe 2006b), or more simply, using “the spare processing power of millions of human brains” (Howe 2006a). However, the discomfort of cultural heritage and academic institutions with both the terms “crowd” and “outsourcing” is apparent in many online and in-person discussions of crowdsourcing. Commenters often question the size and composition of the “crowd” and express discomfort with the perceived threat of replacing digitization and research staff with volunteers. However, the term has enough traction to provide a convenient short-hand name for a particular type of participatory activity. As Estelles-Arolas and Gonzalez-Ladron-de-Guevara point out, crowdsourcing is evolving to the extent that the label may be applied to almost any Internet-based collaborative activity (2012). Lines are also blurred between related terms: for example, “cognitive surplus” (Shirky 2011); “human computation,” in which human effort is used for “tasks that computers cannot yet perform” such as complex visual or semantic processing (Law and von Ahn 2009); and “social computing” and “collective intelligence” (Quinn and Bederson 2011).

Museums sometimes conflate crowdsourcing with “user-generated content” projects on online platforms or in-gallery interactives. This habit can cause confusion when trying to understand motivations for participation and the value subsequently placed on the content that has been created. While crowdsourcing is closely related to Web 2.0-style user-generated content (UGC) projects, and raises similar issues about the validation and incorporation of audience-created content into collections documentation and other museum informatics systems, crowdsourcing projects only offer tasks that will contribute to the specific, shared, and substantial goals of the project. The Citizen Sci-
ence Alliance requires that all their crowdsourcing projects answer “a real scientific research question” and “must never waste the ‘clicks,’ or time, of volunteers” (Romeo and Blaser 2011). Dunn and Hedges state that “humanities crowdsourcing” requires “a clearly-defined humanities direction and/or research question” (2012). Peter Samis’s memorable phrase summarizes the fate of many UGC projects: “We opened the door to let visitors in… then we left the room.” The value for the writer of a comment and the value of that content for a later reader in simple “have your say” activities is often not commensurate (Shirky 2012; Clari and Graham 2012). Holley differentiates between social engagement—“giving the public the ability to communicate with us and each other”—as a method, and crowdsourcing as the process through which a group working collaboratively achieves a “shared, usually significant, and large goal” (2010). To summarize these definitions and the related literature, crowdsourcing is emerging as a form of engagement with cultural heritage that contributes toward a shared, significant goal or research area, by asking the public to undertake tasks that cannot be done automatically, in an environment where the tasks, goals (or both) provide inherent rewards for participation.

A number of related terms describe crowdsourcing projects based in humanistic or scientific disciplines. Citizen science, where “volunteers from the general public assist scientists in conducting research” (Raddick et al. 2010), is well established. Citizen history (Frankle 2011) and the U.S. National Archives term “citizen archivist” are gaining ground. The long tradition of volunteering in cultural heritage encompasses both citizen science and citizen history (Proctor 2013). The role of traditional volunteer bureaus in matching people to opportunities has been supplemented by online citizen science portals. As an evolution of citizen participation in collection, research, and observation, crowdsourcing as we know it has been transformed by technology, but not created by it. Digital technology is able to provide almost instantaneous data-gathering and feedback, automatic validation, and the ability to reach both broad and niche groups through loose networks. For museums, technology has also helped manage the limitations of physical space, conservation, location, and opening hours, all of which previously affected access to collections.

SOURCE MATERIALS, TASKS, OUTCOMES AND PLATFORMS

Tasks and source materials in cultural heritage crowdsourcing projects tend to fall into common groups: the collection, description, transcription, or specialist digitization of material culture, natural history, and historic documents. These contribute to the transformation and gathering of knowledge about cultural heritage. Crowdsourcing projects often divide up large tasks (like digitizing an archive) into smaller, more manageable tasks (like transcribing a name, a line, or a page); this method has helped digitize numerous primary sources. In cultural heritage crowdsourcing, easy-to-complete “microtasks,” or “one-off tasks requiring minimal effort” (McGonigal 2008) enable a broad base of potential participants who would find more complex or time-consuming tasks less satisfying. Tasks can be described as the “atoms” of crowdsourcing, and can be linked together to form larger actions that contribute to project goals.

Oomen and Aroyo categorized crowdsourcing within cultural heritage in terms of “tangible outcomes,” including “correction and transcription,” “contextualization,” the collection of objects, “classification,” and “co-curation” (2011). Task types and outcomes are often
closely linked—for example, the act of connecting museum objects to each other, or to authority records, creates relationship data—but the link is not always straightforward. Tags generated by tagging tasks can be divided into three general classes (Sen et al. 2006) and applied to the tagging of art works. Arends et al. distinguished between subjective tags (opinions and interpretations, useful for self-expression), personal tags (useful for content organization), and factual tags (2012). Other content created through cultural heritage crowdsourcing includes quality ratings, subjective rankings, spatial coordinates, extended creative or factual descriptions, personal experiences or memories, family records, and even game levels (Preloaded undated); through tasks as varied as tagging, transcription, geo-location, and creating or sharing media such as text, images, audio, and video.

Cultural heritage crowdsourcing platforms range from games for creating metadata about objects (Ridge 2011a; Flanagan and Carini 2012) to projects based on existing social platforms like Flickr Commons, online newspaper archives (Holley 2010), or specialist sites for transcribing hand-written correspondence (Moyle, Torra, and Wallace 2011), each of which is designed to suit particular motivations for and types of participation. It is outside the scope of this essay, but crowdfunding is a related dynamic field which will yield useful case studies for future research.

PARTICIPANTS

Unlike commercial crowdsourcing, participation in cultural heritage crowdsourcing is driven by pleasure, not profit. Rather than monetary recompense, GLAM projects provide an opportunity for altruistic acts, activated by intrinsic motivations, applied to inherently engaging tasks, encouraged by a personal interest in the subject or task. In order to understand how projects can create deeper involvement with cultural heritage, it is necessary to consider why participants start and continue contributing to crowdsourcing projects.

We might categorize participants in cultural heritage crowdsourcing projects groups in two ways: those who are intentionally participating in crowdsourced tasks for the intrinsic rewards; and those whose contributions are a side effect of their participation in other core activities. Intentional participants could be considered “digital volunteers” and include people who are passionate about the relevant subject, people who like doing the task offered in the project or the source material used, and people who are unable to volunteer in venue opening hours or locations.

Museums can find the open nature of crowdsourcing calls for participants to be challenging, because tasks are undertaken online by possibly anonymous participants. To counter this, museums use terms such as “community-sourcing,” or working with people who already have a relationship with an institution (Phillips 2010), and “nichesourcing,” where tasks are “distributed amongst a small crowd of amateur experts” (de Boer et al. 2012). Some projects cannot be as “open to all” as museums might like, because participants are limited to those who have or are willing to learn skills such as paleography or georectification.

While some participants in crowdsourcing projects are new to GLAM activities and resources, others are experts in the relevant field. Citizen historians are perfect examples of “professional amateurs” (professional amateurs): “self motivated, enthusiastic, and dedicated” amateurs who work to professional standards (Leadbeater and Miller 2004; Terras 2010). As an example of the blurred lines between professional and amateur historians, people interviewed for my current research have included trained historians who
work in other fields while enjoying hobbies that let them keep up their historical research skills; academics who learned their historical research skills on the job; and self-taught researchers with decades of experience as practicing historians.

**DESIGNING FOR PARTICIPATION**

The following section outlines some design considerations for encouraging and deepening participation through museum crowdsourcing projects, including design techniques such as microtasks, scaffolding, and related concepts such as casual game design and “flow.”

As crowdsourcing becomes more popular, a growing body of literature discusses the challenges of engaging mass audiences while competition for participants increases. Museums find it useful to provide easy ways to begin, so that potential participants who come across a project are immediately engaged. Designing a seductive initial task “that can be accomplished quickly and easily” is key: “It is less important at the onset to make something interesting or challenging than it is to make something easy” (McGonigal 2008). Snoek et al. described how the audio-visual archive-based crowdsourcing game *Waisda?* designed microtasks that led users to increase their level of activity by providing feedback “just by clicking buttons,” or validating another user’s tag with a “thumbs up.” Interactions were designed to entice the user into increasing the level of participation; users who press the “thumbs-down button” are asked to correct the label (2010).

**SCAFFOLDED DESIGNS IN CULTURAL HERITAGE CROWDSOURCING**

It may sound paradoxical, but constraints in participatory interfaces can encourage engagement. Scaffolding supports novices “by limiting the complexities of the learning context” and gradually “fading” or removing those limits (Dabbagh 2003). Nina Simon observes that the best participatory experiences in museums are “scaffolded to help people feel comfortable engaging in the activity” by building their confidence. In museum projects, scaffolding supports increased participation by providing “clear roles and information about how to participate” (Simon 2010). Scaffolding is designed for face-to-face educational environments where teachers can monitor student performance and adjust lessons accordingly. However, some principles can be extrapolated to guide the design of audience experiences online.

The New York Public Library’s hugely successful crowdsourcing project *What’s on the Menu?* focuses on digitizing its collection of historic menus. The value in reducing task complexity is evident on its closely scaffolded interface designed around the key tasks of transcribing menu items and prices. By removing any uncertainty about how to fill in the two text boxes, the interface design reduces cognitive overhead, making the task simpler and more enjoyable. The whole site is designed to minimize barriers and encourage participation in clearly defined tasks.

Some of the less successful projects I have analyzed have failed in part because their initial task was too complex or required too much domain-specific knowledge, necessitating awkward and ineffective pre-task tutorials or instruction sheets, or the project was too loosely defined, with no clear feedback when a task was completed successfully.

**THE INTERSECTION BETWEEN SCAFFOLDING AND DESIGN FOR CASUAL GAMES**

Games can also be effective drivers of participation in museums (Birchall et al. 2012).
Games provide useful demonstrations of the power of scaffolded interactions. Crowdsourcing games, or Games with a Purpose (GWAP), in which “players perform a useful computation as a side effect of enjoyable game play,” proved that games could bring mass audiences to computational problems such as describing the content of images with tags (von Ahn and Dabbish 2004; 2008). Flanagan and Carini found that GLAM crowdsourcing games could generate more content per participant than non-game interfaces (2012). Other crowdsourcing games in this area include games about art (Brooklyn Museum’s Tag! You’re It! and Freeze Tag! [Bernstein 2008; 2009]); contemporary audio-visual material (Waisda? [Oomen and Aroyo 2011; Snoek et al. 2010]); and historic newspapers (DigitalKoot). Currently, most successful crowdsourcing games are focused on microtasks like tagging, validating data, or transcribing small sections of content within larger collections.

The precepts for casual game design, which are drawn from years of practical experience in the gaming sector, encapsulate some of the principles of design for increased participation that have evolved from scaffolding theory. Casual games are “games with a low barrier to entry that can be enjoyed in short increments” (Casual Games SIG 2009); these include puzzles, word games, board games, card games, and trivia games. Features of casual games like Solitaire and Angry Birds include easy-to-learn gameplay, simple controls, addictive and “forgiving” gameplay with a low risk of failure, and inclusive, accessible themes (Casual Games SIG 2009); these characteristics make games ideal for crowdsourcing (Ridge 2011a; 2011b). A key design principal—carefully managed complexity levels with a shallow learning curve and guidance through early levels—is clearly related to scaffolding. For museums with limited design budgets and large collections to cover, the International Game Developers Association recommends favoring “a variety of content over a variety of mechanics in a single game”; it reports that adding similar content to the same game structure leads the player to “greater feelings of mastery” (Casual Games SIG 2009). An additional benefit is that growing feelings of mastery can help participants stay engaged.

Crowdsourcing games can build tutorials for new skills into the gameplay itself at the point where they are needed (Ridge 2011a). It seems that casual game design operationalizes scaffolding theory in ways that usefully inform design for cultural heritage crowdsourcing.

**MOTIVATIONAL FRAMEWORKS FOR PARTICIPATION IN CROWDSOURCING**

Understanding why people participate in crowdsourcing is important in designing for participant recruitment and retention. Project marketing and instructional messages that match participant motivations have enhanced “persuasive impact” and help volunteers find more enjoyable and satisfying roles that match their motivations (Clary et al. 1998; Romeo and Blaser 2011). In order to find research on motivations for participation in crowdsourcing projects that did not offer monetary recompense, I turned to research into citizen science and other “community-based peer-production projects” where people participate in collaborative efforts to create “publicly available knowledge-based products” such as open source software or the collaboratively written website Wikipedia (Nov 2007; Nov, Arazy, and Anderson 2011). So why do people participate when they are not being paid? Brabham reported that several studies have located the primary motivator of participation in open source projects in
“the pleasure found in doing hobbies” (2008). Research for citizen science projects discovered that the most important motivations for voluntary participation were the collective (“the importance attributed to the project’s goals”) and intrinsic fun, or “the enjoyment associated with participation in the project” (Nov, Arazy, and Anderson 2011). Raddick et al. identified the main benefits for participants in the citizen science project Galaxy Zoo: enjoyment, community, the ability to participate in real science, and recognition for their participation (2009). Oomen and Aroyo discussed two groups of “motivational factors” in cultural heritage crowdsourcing: “connectedness and membership,” and “sharing and generosity” (2011). Many researchers have found that community and social interactions are important motivators for participation (Nov, Arazy, and Anderson 2011; Clary et al. 1998). Looking to pre-online volunteering, Holmes found that “social opportunities” and “colleagues” were important motivations for continued volunteering in the heritage sector (2003). These various motivations can be grouped into altruistic, intrinsic, and extrinsic motivations.

Unlike most commercial projects, cultural heritage crowdsourcing projects are well positioned for appeals to altruism. Oomen et al. suggest that the use of specialist interfaces that reinforce the altruistic nature of the activity increases participation (2010). This has implications for the design of crowdsourcing platform functionality and interface design. Trant recommends demonstrating the use of data so that players can see the impact of their contribution (2009). In an earlier project, I learned that “validating procrastination” by offering messages supporting altruistic motivations helped participants justify their time on the activity and could increase levels of participation (Ridge 2011a). My review of crowdsourcing projects found that intrinsic motivations—an activity worth doing for its own sake (Csikszentmihalyi and Hermanson 1995)—for participating in museum crowdsourcing include fun, the pleasure in doing hobbies, enjoyment in learning, mastering new skills and practicing existing skills, recognition, community, and passion for the subject. GLAM crowdsourcing projects can appeal to extrinsic and intrinsic motivations: The same task (such as transcribing sections of a historic document) could be undertaken for differing reasons: altruistic, such as “helping to provide an accurate record” of national history as represented in an online database of digitized newspapers (Alam and Campbell 2012); intrinsic, such as solving the enjoyable puzzle of reading eighteenth-century handwriting in a correspondence; or extrinsic, as when an academic transcribes a quote from a primary source.

Trove, the National Library of Australia’s database of online resources (including digitized newspapers), offers crowdsourcing functionality that is closely aligned to the needs of users who can correct text from the digitized originals for their own uses. However, other users choose to correct incorrectly transcribed text for the intrinsic enjoyment of performing that task in that context. This echoes findings by Dunn and Hedges that the primary motivation in humanities crowdsourcing can be altruistic, extrinsic, or intrinsic, but it is nearly always strongly related to the “project or activity’s subject area” (2012).

FROM TAGGING TO THEORIZING

There is a growing body of evidence that crowdsourcing (or citizen science, or citizen history) can lead to deeper engagement with disciplines such as science or history, as well as an increase in related skills and knowledge (see for instance Dunn and Hedges 2012). Just as there
is value in a one-off contribution to a crowdsourcing project, there is also value in exercising the skills required at the initial levels of participation. The possibility of building research and synthesis skills, experience, and content knowledge through participating in crowdsourcing is an exciting opportunity for GLAMs. In this section I look for useful definitions of “engagement” and evidence for why some activities engage audiences. I present some projects with highly engaged participants who have in turn found ways to connect more deeply with the material, and I consider the role of emergent communities in these projects.

Some cultural heritage crowdsourcing projects allow participants to graduate from tightly scaffolded microtasks to higher cognitive processes like explaining, relating, and theorizing (Biggs 1999). Research into the citizen science project Galaxy Zoo has led to the development of a model: “Levels of Engagement.” At the first level, volunteers participate in simple classification tasks; at the second they participate in community discussion (for example, on a project forum or blog), and at the third and final level they move to “working independently on self-identified research projects” (Raddick et al. 2009). Raddick et al. report that citizen science research projects “have resulted in volunteers teaching themselves about scientific content, using tools of modern astronomy data, and working as scientists.”

### DEFINING “ENGAGEMENT” IN CULTURAL HERITAGE

Many definitions of “engagement” in the arts and cultural heritage simply seem to equate to physical attendance at events or venues. A more nuanced model comes from the United Kingdom’s governmental agency, Department for Culture, Media, and Sport, and its program titled “Culture and Sport Evidence” (CASE 2011). This research defines four types of engagement, each building on the previous level: 1) “Attending”—paying conscious, intentional attention to content. 2) “Participating”—an interaction that contributes to the creation of content. 3) “Deciding”—making decisions about the delivery of resources for content creation. 4) “Producing”—creating content “which has a public economic impact” (CASE 2011).

Other definitions draw directly from museology. Bitgood found that engagement involves “deep sensory-perceptual, mental and/or affective involvement with exhibit content,” possibly leading to “personal interpretation of exhibit content,”—“meaning making,”—or a “deep, emotional response” (2010). While written in the context of physical encounters, both the CASE model and Bitgood’s definition apply equally well to physical and online visits or participation, and provide an explanation for art historians’ sudden interest in astrolabes.

### CROWDSOURCING AS PRODUCTIVE ENGAGEMENT WITH CULTURAL HERITAGE

We call what we do harvesting or cataloguing, digitizing or preserving, data visualizing or crowd-sourcing, community management or customer service, or whatever the latest round of restructuring has deemed our job to involve. But what we’re really doing is working with people to create and share back our collective and collaborative history. — Courtney Johnston (2011)

In an earlier issue of this journal, Owens concluded that crowdsourcing is a powerful tool for “deep engagement with the public” (2013). Aside from the intrinsic enjoyment discussed earlier, what is it about the scaffold-
ing, types of tasks, rewards, and community around crowdsourcing in cultural heritage that makes it so engaging? Some answers may lie in the close, active viewing of objects, scientific imagery, or historic documents, and the work required to describe, categorize, or transcribe them. Bitgood proposed a relationship between engagement and “exertion or concentration,” aided by a viewing time of “more than a few seconds” (2010). Earlier research on museum metadata games found that curiosity about presented objects “was a compelling part of the experience” (Ridge 2011a). Flanagan and Carini stated that players of the Tiltfactor metadata games “became so curious about the images they were tagging” that they added tags like “want to know more about this culture” (2012).

Deep engagement can be its own reward. “Flow” is experienced as a state of deep, enjoyable focus or engagement (Csikszentmihalyi 1990; Csikszentmihalyi and Hermanson 1995). It requires a clear goal, immediate feedback on the success of your attempts to reach that goal, and a good match between the skills of the individual and the challenges faced. Supporting flow through content and interaction design helps keep players engaged with an activity, and therefore helps crowdsourcing projects be more productive. It also points to one challenge of maintaining participation levels in crowdsourcing: When your skills are greater than the challenge, you become bored; but if you do not have the skills to meet the challenge successfully, then you experience anxiety. It can be difficult to increase the challenge and provide sufficient scaffolding so that skills and mastery grow without compromising the quality of data. Some projects have found ways to manage this increase of skills and challenges. Providing opportunities for increasing mastery is discussed below.

PROJECTS ENABLING DEEPER ENGAGEMENT

Dunn and Hedges observed that some projects support participants moving up a level to “carry out more complex tasks” (2012). Long-established or hugely popular crowdsourcing projects like Herbaria@Home (2006), FamilySearch (2007), Galaxy Zoo (2007), Founders and Survivors (2009) and Old Weather (2010) have had enough time and enrollment to demonstrate ways in which project participants can develop new skills and knowledge as a result of their growing interest in the project source material, or can graduate to more complex tasks or bigger responsibilities.

For public participation in science, Bonney et al. devised a useful model for categorizing heritage crowdsourcing projects according to the amount of control participants have over project design processes (such as defining questions for study, collecting and analyzing data, and drawing conclusions). Their model contains three categories: “contributory,” in which the public contributes data to a project designed by the organization; “collaborative,” in which the public can help refine project design and analyze data in a project led by the organization; and “co-creative,” in which the public can take part in all or nearly all processes, and all parties design the project together (Bonney et al. 2009). Most crowdsourcing projects are contributory, but ones such as Old Weather and Herbaria@Home have evolved into collaborative and at least partially co-creative projects. It may be that more collaborative project structures are a factor in successfully encouraging deeper engagement with related disciplines. Further research is needed, but one measure of long-term success in GLAM crowdsourcing may be a willingness and ability to listen to a project’s user communities and collaboratively devise new and
improved tasks and research questions. Samis and Michaelson found that “a visitor centered focus leads to organizational transformation” (2013). The impact of collaborative and co-creative crowdsourcing projects on the sponsoring organization is not yet clear, but it is probably significant that each of the projects discussed below has changed in response to participant actions and comments.

**FamilySearch**

*FamilySearch* is a genealogy site that encourages members of the public to “index” (transcribe) historic records. The site aims to get people to try a simple task—indexing records—knowing that, as transcribers are exposed to other people’s histories, they will probably be gradually interested in finding out more about their own families. This model of encouraging engagement introduces people who are not interested in family history (or who are overwhelmed by it) to the skills required in an initially closely scaffolded environment. Davis points out that transcribing historical documents “provides some much-needed, introductory, family history education” and increases the participants’ knowledge about the range of record types and genealogical information, while providing “handwriting practice” (2012).

*FamilySearch* provides a further level of involvement by inviting some established transcribers to become “arbitrators” who can review and approve the work of other transcribers. Invitations to become an arbitrator are issued after participants transcribe a certain number of records with a sufficiently high accuracy rate, or at the recommendation of a more senior participant (Anderson 2012a; Anderson 2012b).

**Old Weather**

The *Old Weather* project aimed to digitize ship logs in order to analyze historic climatic data. However, ship logs contain enough interesting oddities that transcribers started to become interested in the voyages, events, and lives of those on the ships and in maritime history generally. Dunn and Hedges found that transcribers were following these interests and becoming expert in “specialized areas of naval history” (2012). Posts on the *Old Weather* forums are a good example of the mixture: help with paleographic queries; explanation of subject-specific jargon; and curiosity about passing references in the logs. These questions develop into a deeper interest in the topic, an evolution that characterizes citizen science and citizen history projects. Participants have theorized about questions including the relationship between deaths onboard ship and successive waves of the 1918–1919 influenza epidemic and have written detailed guidelines for others who wish to edit ship histories.¹

**Herbaria@Home**

*Herbaria@Home* aims to document historical herbarium collections in museums, based on photographs of specimen sheets supplied by museums. To date, participants have documented almost 125,000 herbarium specimens. Some transcribers also became interested in the people whose specimens they were documenting, and started collating their samples and researching their lives. To support this new research, the project expanded to include a wiki of biographies of the original collectors, along with samples of their handwriting from specimen sheets, forum posts, specimens collected, and related people, places, and dates.
The Galaxy Zoo project initially asked participants to classify images of galaxies into three simple groups, with more complex classifications added in successive stages (Raddick et al. 2009; Cardamone et al. 2009; Raddick et al. 2010; Romeo and Blaser 2011). After a burst of publicity led to an increase in enquiries, the Galaxy Zoo team launched a forum “to encourage volunteers to communicate with one another and answer each other’s questions” (Raddick et al. 2010). This turned out to have unexpected consequences. Through discussion and conjecture on the forum, project participants have devised “novel collaborative research projects,” developed their own analysis tools, and discovered “at least one truly unique object” (Raddick et al. 2010; Simpson 2013). Examples include a collaboration among members of the forum who became curious about objects they nicknamed “green peas,” collecting examples of them and campaigning to “give Peas a chance,” eventually leading to the discovery of a new class of galaxy (Sheppard 2009; Cardamone et al. 2009). The forum also contains tales of the ways in which engagement with science has increased interest and changed lives: Participants have joined local astronomical societies, re-entered formal education to study astronomy, even changed careers. The Galaxy Zoo project has recently launched software tools to help “Zooites” interact and theorize “more deeply” with the data they have helped create (Simpson 2013).

COMMUNITY AS SCAFFOLDING

The projects discussed above show the value in leaving room for curiosity to develop into deeper interest in the subject of a project (see also Csikszentmihalyi and Hermanson 1995). Providing different roles within a project—such as transcriber, data validator, or community support member—is an excellent method for dealing with challenges specific to crowdsourcing. (These challenges might include validating contributions with limited resources for community outreach, and content moderation.) It is also an excellent way to keep participants motivated and excited about new challenges and responsibilities. Projects that provide a variety of tasks and roles can support a range of different levels and types of participant skills, availability, knowledge, and experience. But designing crowdsourcing interfaces that are able to determine participant skills and motivations, and with enough flexibility to respond appropriately, is difficult with limited resources. Sharma and Hannafin write: “Software constraints often limit dynamic scaffolding to interactions that can be anticipated in advance” (2007).

However, as Dunn and Hedges observed, the “vibrant and interacting communities of contributors” that emerge around many crowdsourcing projects are a factor in their success: “Communities develop and perpetuate internal dynamics, self-correct [and] provide mutual support” (2012). An examination of participant forums shows that the community itself can produce some of the personalized scaffolding for learning or mastering skills and knowledge in subject domains that digital interfaces currently cannot support. Crowdsourcing projects that encourage community participation find that dialogue between experts and novices can provide additional scaffolding through “continuous and constructive interactions between experts and learners” (Sharma and Hannafin 2007). Looking to the future, machine learning and improved computational techniques for pattern recognition and visual processing may reduce the need for volunteers for currently popular crowdsourced tasks,
raising new ethical issues for cultural heritage institutions, including their responsibility for the communities that have developed around and care deeply about their projects.

CONCLUSION

It is clear that crowdsourcing can be a productive tool for completing digitization and research tasks required by museums and cultural heritage institutions. Projects that can respond to the changing needs of their participants, by introducing new content or tasks, or by offering new roles and responsibilities, help people remain motivated to keep participating and can even generate new research questions. Crowdsourcing projects are also a powerful platform for audience interaction with museums, offering truly deep and valuable engagement with cultural heritage. As an opportunity to "exercise knowledge, skills, and abilities that might otherwise go unpracticed" (Clary et al. 1998), volunteering through GLAM crowdsourcing projects provides a platform for lifelong learning and an opportunity to engage with cultural heritage content and tasks. In the words of the project instigators of the New York Public Library’s What’s on the Menu: “We are coming to see crowdsourcing not only as a way to accomplish work that might not otherwise be possible, but as an extension of our core mission” (Lascarides and Vershbow, forthcoming). Following in the steps of traditional volunteering, but with far wider and deeper capabilities to reach and engage the public, crowdsourcing helps museums serve their core missions by providing platforms for audiences to follow their own interests and hobbies and connect with communities of practice.

ACKNOWLEDGMENTS

This article is based on a keynote lecture for the iSay conference “The Shape of Things: New and Emerging Technology-enabled Models of Participation through Visitor-generated Content” at Leicester University in February 2013, as well as on a series of workshops, and on the author's current Ph.D. research into historians and scholarly crowdsourcing. The author would like to thank the reviewers for their comments, and workshop participants for their questions, which have helped inform this essay.

NOTES


2. I am avoiding the term "gamification," which is another problematic buzzword with contested definitions. At best, it describes "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled, and Nacke 2011); at worst it talks about gimmicky or exploitative design that produces what game developer Kathy Sierra calls "a short-term sugar rush of engagement followed by a crash" (2011).


5. The original forum discussion is at http://www.galaxyzooforum.org/index.php?topic=3638.0. An outline of the community collaboration is at http://www.galaxyzooforum.org/index.php?topic=270633.msg218401#msg218401. Another example is the “Heartbeat Binary stars” discussion on the Planet Hunters citizen science project at
REFERENCES


In *Crowdsourcing Our Cultural Heritage*, Mia Ridge, ed. Ashgate.


# Proof Correction Marks

Please correct and return your proofs using the proof correction marks below. For a more detailed look at using these marks please reference the most recent edition of The Chicago Manual of Style and visit them on the Web at: [http://www.chicagomanualofstyle.org/home.html](http://www.chicagomanualofstyle.org/home.html)

<table>
<thead>
<tr>
<th>Instruction to typesetter</th>
<th>Textual mark</th>
<th>Marginal mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leave unchanged</td>
<td>• • • under matter to remain</td>
<td>(def)</td>
</tr>
<tr>
<td>Insert in text the matter</td>
<td>▲</td>
<td>(\wedge)</td>
</tr>
<tr>
<td>indicated in the margin</td>
<td>(\triangleright) through single character, rule or underline</td>
<td></td>
</tr>
<tr>
<td>Delete</td>
<td>or</td>
<td>(\triangleright) through all characters to be deleted</td>
</tr>
<tr>
<td>Substitute character or</td>
<td>(\triangleright) through letter</td>
<td>new character (\hat{}) or</td>
</tr>
<tr>
<td>substitute part of one or more word(s)</td>
<td>or</td>
<td>new characters (\hat{})</td>
</tr>
<tr>
<td>Change to italics</td>
<td>(\hat{})</td>
<td>(\hat{})</td>
</tr>
<tr>
<td>Change to capitals</td>
<td>(\hat{})</td>
<td>(\hat{})</td>
</tr>
<tr>
<td>Change to small capitals</td>
<td>(\hat{})</td>
<td>(\hat{})</td>
</tr>
<tr>
<td>Change to bold type</td>
<td>(\hat{})</td>
<td>(\hat{})</td>
</tr>
<tr>
<td>Change to bold italic</td>
<td>(\hat{})</td>
<td>(\hat{})</td>
</tr>
<tr>
<td>Change to lower case</td>
<td>(\checkmark)</td>
<td>(\checkmark)</td>
</tr>
<tr>
<td>Insert superscript</td>
<td>(\wedge)</td>
<td>(\wedge)</td>
</tr>
<tr>
<td>Insert subscript</td>
<td>(\checkmark)</td>
<td>(\checkmark)</td>
</tr>
<tr>
<td>Insert full stop</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Insert comma</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Insert single quotation marks</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Insert double quotation marks</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Insert hyphen</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Start new paragraph</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Transpose</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Close up</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
<tr>
<td>Insert or substitute space</td>
<td>linking (\bigcirc) characters</td>
<td></td>
</tr>
<tr>
<td>between characters or words</td>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td>Reduce space between characters or words</td>
<td>(\quad)</td>
<td>(\quad)</td>
</tr>
</tbody>
</table>