

RESEARCH PAPER

A Knowledge Perspective on Quality in Complex Citizen Science

Montserrat Prats López*, Maura Soekijad†, Hans Berends† and Marleen Huysman†

This article examines how project leaders in complex citizen science projects ensure the quality of project outcomes given the challenges of involving citizens, whose knowledge is diverse and unknown beforehand. To this aim, a qualitative multiple-case study was carried out to compare the practices of five collaborative online citizen science projects in the humanities in which citizens transcribe, translate, and annotate handwritten manuscripts from the 17th, 18th, and 19th centuries. The choices made to access knowledge, thus to recruit participants either through open or targeted calls, seem to be fundamental for the different configurations of knowledge management activities that project leaders apply to ensure quality outcomes. Other factors influencing these knowledge management configurations are citizens' proximity, knowledge characteristics, technology affordances, and the extent to which project leaders are aware of citizens' backgrounds and skills. This study adds to earlier frameworks proposed to advise the design and management of citizen science projects. By taking a knowledge perspective, this article provides practical directions for project leaders involved in citizen science and highlights the need to put time and effort in managing knowledge processes.

Keywords: citizen science; knowledge management; humanities; manuscript transcription

Introduction

The phenomenon of online citizen science, defined as a participative way of running scientific research projects, in which researchers and citizens work together through the Internet primarily collecting, processing, and/or analysing data (Wiggins and Crowston 2011; Riesch and Potter 2014), has become a topic of research in itself. Research about citizen science has mainly focussed on two key areas of interest for project leaders: volunteer engagement and the quality of project outcomes.

We focus on quality, as this is a key concern for project leaders striving for high-quality project outcomes (Riesch and Potter 2014) that depend on volunteers. The quality of citizen science outcomes is essential because the reliability of research depends on it. Concerns about quality in citizen science (Oomen and Aroyo 2011; Wiggins et al. 2011; Sheppard, Wiggins and Terveen 2014; Riesch and Potter 2014; Bonney, Cooper and Ballard 2016) are not surprising given the involvement of (usually unknown) diverse and distributed citizens with different levels of expertise versus the complexity of research tasks for which academics have been trained for years (Miller 2001).

In many well-known citizen science projects, citizens usually perform straightforward tasks, such as classifying images based on predefined categories as in project Galaxy Zoo, or transcribing structured information like ships' logbooks in Old Weather (Dunn and Hedges 2014; Ponciano and Brasileiro 2014; Mitchell, Crowston and Østerlund 2018). The quality of those types of tasks is mainly ensured by aggregating multiple contributions or comparing them with a gold standard (Brumfield 2012; Law et al. 2017). However, not all problems and activities are suitable to division into simpler tasks (Afuah and Tucci 2012). In complex citizen science, tasks are less modularizable, and more knowledge-intensive and time-consuming. Examples of complex citizen science include the transcription, translation, and contextualisation of handwritten manuscripts commonly used in the humanities, in particular in historical and literary research (Dunn and Hedges 2014). Given the knowledge-intensity of complex tasks, the quality of their outcomes are usually difficult to evaluate (Alvesson 2001).

Quality is essential for the outcomes of citizen science, and earlier research has recommended keeping outcomes in mind when designing projects (Shirk et al. 2012), but we still know little about how quality is ensured, especially in complex citizen science (Kittur et al. 2013); thus, the aim of this study is to understand how project leaders in complex citizen science ensure the quality of project outcomes.

Recent studies indicate that the feasibility of delegating research tasks to (usually) unknown citizens through the

* Open Universiteit, NL

† Vrije Universiteit Amsterdam, NL

Corresponding author: Montserrat Prats López
(montserrat.pratslopez@ou.nl)

internet depends on the type of knowledge needed to perform such tasks and the quality requirements of resulting outcomes (Law et al. 2017). Because scientific research is the knowledge-creating process par excellence, and quality is a characteristic of knowledge as input to task performance and of the outcome of knowledge work (Haas and Hansen 2007), we take a knowledge perspective to understand how quality is ensured in complex citizen science projects. Knowledge management refers to a set of processes that facilitate knowledge creation, finding and connecting knowledge that is geographically, structurally or functionally scattered, to support innovation and improve performance (Davenport and Prusak 2000; Hislop, Bosua and Helms 2018).

In this paper, we focus on the humanities and use the word science to refer to all fields of academic research, hence our consideration of crowdsourcing projects in the humanities as citizen science (Dunn and Hedges 2013). The humanities are a suitable setting for studying complex citizen science because they involve knowledge-intensive tasks, such as the transcription of manuscripts, in which citizens contribute to the interpretation and processing of textual data (de la Flor et al. 2010). Quality of manuscript transcriptions is essential because transcriptions are used as input for linguistic, literary, or historical research.

In the remainder of this article, we review citizen science literature and identify knowledge management activities in these types of projects. We describe the research setting and methods, and compare the knowledge management activities used in five collaborative online citizen science projects in the humanities. This study expands prior frameworks for the design and implementation of citizen science projects (Shirk et al. 2012; West and Pateman 2016) by examining in detail the activities used to manage knowledge work and to address quality issues in complex citizen science.

A Knowledge Perspective on Quality in Citizen Science

Citizen science projects involve knowledge work. First, tasks contribute to the scientific research process (Cooper et al. 2007) and hence to knowledge creation. Second, these tasks depend on human skills (Wiggins and Crowston 2011) involving creativity and leading to unique outcomes (Hislop 2008) to support research objectives. And third, unique outcomes entail the integration and application of knowledge (Hislop 2013) of both professional researchers and citizen participants.

The characteristics of citizen science projects, and the importance of knowledge in it, lead to two knowledge challenges. The first challenge is the uncertainty resulting from the diversity and geographical distribution of citizens, who are a priori unknown to scientists leading a project (Franzoni and Sauermann 2014). This results in uncertainty about citizens' knowledge and time availability for projects (Law et al. 2017). The second challenge is the weaker control that project leaders have on knowledge flows, as citizens are not employed by the research organization (Kittur et al. 2013; Simula 2013) and are thus not subject to formal supervision (Sheppard et al. 2014).

Given the characteristics and knowledge-related challenges of citizen science, a knowledge perspective seems a suitable lens for studying how project leaders ensure quality in this context. Therefore, we review and integrate the citizen science literature from a knowledge management point of view.

Accessing versus acquiring knowledge

Knowledge acquisition refers to the activities used to obtain new knowledge for an organization. Traditionally, new knowledge can be acquired through research and learning (Huber 1991) and/or by hiring new employees (Davenport and Prusak 2000) with distinct knowledge and expertise, evidenced by their curriculum and qualifications. In citizen science, however, recruiting can be either to acquire new knowledge and learn from the public (Afuah and Tucci 2012; Oosterman et al. 2014) or to access and use specific skills of the public.

In citizen science, project leaders do not select and hire employees (Simula 2013). Instead, citizens voluntarily decide whether to participate in a project, which is referred to as self-selection (Afuah and Tucci 2012; Franzoni and Sauermann 2014), which leads to the abovementioned challenge of knowledge uncertainty. To avoid such uncertainty and to ensure quality, the literature suggests selecting participants a priori with specific qualifications as a prerequisite for participation (Sodré and Brasileiro 2017) or by testing their skills (Wiggins et al. 2011). However, given the voluntary nature of citizen science, such practices raise questions about how to deal with participants whose knowledge and skills do not fit expectations. Moreover, since only a small group of people contribute regularly to such projects (Sauermann and Franzoni 2015), it is unclear how selecting participants could still result in time and resource efficiency, key benefits of citizen science (Franzoni and Sauermann 2014).

Sharing and integrating knowledge

Knowledge sharing refers to the communicative activities by which individuals make part of their knowledge available to others (Berends 2005). Traditionally, the knowledge management literature distinguishes different ways to communicate and share knowledge, depending on the distribution of people (Greenberg and Roseman 2003) and the characteristics of knowledge (Alavi and Denford 2011). Activities, such as storing and distributing documented knowledge through websites, are suitable for transferring explicit knowledge, whereas personal interactions are better for sharing tacit knowledge (Alavi and Denford 2011). Knowledge sharing through face-to-face interaction is easier among people collocated in space and time, while asynchronous communication tools are more suitable for sharing knowledge between distributed people (Greenberg and Roseman 2003). Citizen science participants are geographically dispersed and mainly contribute online, which could mean that only explicit knowledge is shared through asynchronous means, raising questions about the type of knowledge and how it is shared.

Coordinating and applying knowledge

Knowledge has value for an organization only when it is applied in action (Alavi and Denford 2011). In citizen science this refers to the research-related tasks outsourced to and performed by the public. Scholarly research includes tasks with varying degrees of knowledge tacitness, which can make the tasks complex. In online citizen science, tasks are usually simplified (Riesch and Potter 2014) or modularized (Afuah and Tucci 2012) to allow their online performance. This usually results in a pooled interdependence (Haythornthwaite 2009), as tasks are performed independently from each other and integrated into one final outcome. Yet we know little about how knowledge-intensive tasks are coordinated and knowledge is applied in complex citizen science (Mitchell et al. 2018). Moreover, whereas in research organizations work is allocated on the basis of employees' skills and expertise, in citizen science, people self-allocate or decide which tasks they want to perform (Puranam, Alexy, and Reitzig 2014). This brings us back to the question of how to deal with knowledge uncertainty resulting from self-allocation.

In knowledge management, the integration of effort refers to the embedding of expert knowledge into routines, rules, and procedures, so that non-experts can perform tasks without having to learn (Grant 1996). Similarly, citizen science projects use protocols or standardized procedures, project plans, training, and supervision (Wiggins et al. 2011; Bordogna et al. 2014; Riesch and Potter 2014; Freitag, Meyer and Whiteman 2016). However, unlike employees who receive monetary rewards for applying their knowledge, citizen volunteers are not paid nor subject to contractual agreements (Simula 2013; Franzoni and Sauermann 2014) or to the rules of scientific practice (Hedges and Dunn, 2018). This makes them free to follow or ignore any plans, procedures, or training. Altogether, this raises questions about how to manage knowledge flows and reduce knowledge uncertainty.

Assessing knowledge

Knowledge assessment has barely been examined in the knowledge management literature. In citizen science, knowledge assessment is usually associated with the evaluation of citizens' contributions (Wiggins et al. 2011). Suggested assessment activities include comparing contributions with existing scientific literature or professional observations (Riesch and Potter 2014). This implies that literature or observations already exist (Freitag et al. 2016), but it is unclear how quality is assessed for new research topics.

Another way to assess contributions is through multiple-keying with voting; that is, having multiple participants perform exactly the same task and assume that what the majority inputs is correct (Brumfield 2012). This is usually used for modular and structured tasks (Brumfield 2012; Law et al. 2017), but as tasks become more complex it is more difficult to validate the quality of their results (Alvesson 2001). Other means to assess citizen contributions is through expert reviews (Wiggins et al. 2011) or citizen peer-reviews (Brumfield 2012). However, expert reviews require time investments that citizen science was

supposed to reduce, and citizen peer reviews raise again questions about the selection of participants (Dow et al. 2012). Therefore, it is still unclear how the results of knowledge work are assessed in complex citizen science.

Based on this review, it is still unclear how leaders of complex citizen science projects deal with diverse and distributed knowledge and ensure quality outcomes. We therefore carried out empirical research to understand how citizens are recruited, knowledge is shared, tasks coordinated and performed, and outcomes assessed in complex citizen science.

Research Methods

This research focuses on the transcription of old handwritten manuscripts, their translation into modern language, and contextual annotation—tasks that are knowledge-intensive, time-consuming, and prone to errors. These tasks are knowledge-intensive because they involve diverse and hard-to-decipher handwriting styles (de la Flor et al. 2010), and require the recognition and interpretation of words and abbreviations based on the context of the manuscript, the peculiarities of the author's handwriting, and the historical period. These tasks are also time-consuming because manuscripts vary in length and condition of the paper; that is, it takes time to complete a transcription or translation and to indicate which parts are unreadable because of damaged paper or smudged ink. Moreover, manuscripts contain vast amounts of textual data at many levels (i.e., characters, words, sentences, paragraphs, and pages), increasing the likelihood of errors. Transcribing manuscripts is susceptible to human mistakes because it is easy to skip a line while reading, and people tend to finish off sentences or words before they actually read them completely.

To understand how knowledge work is managed and quality ensured in these complex projects, we conducted a qualitative multiple-case study (Eisenhardt 1989, 1991), with the citizen science projects as cases and the activities performed by project leaders as focus of analysis. We purposefully contacted three core organizations in the field of cultural heritage and humanities research in the Netherlands: The Cultural Heritage Agency, the Meertens Institute, and the Huygens Institute for Netherlands History. We then used theoretical sampling (Eisenhardt 1989) to select projects involving citizen participants who performed knowledge-intensive tasks through the Internet. The following five projects (**Table 1**) were examined and compared to build an explanation (Yin 2014) of how quality is assured in citizen science.

The first project, *Letters and Correspondents around 1900*, started in 2009. The project leader was a scholar in a Dutch research institute and the transcriptions were meant to be used by other humanities researchers. Participants included volunteers with a literature background and literature students at a Dutch university. Together they constituted a small community of about 10 to 20 people. They used a web-based tool to integrate transcriptions. The scans and transcriptions have been available online since November 2016. The second project, *Digitizing Belle van Zuylen's Correspondence*, started in 2010 and was

Table 1: Overview of the cases at the time of the study.

	Letters and Correspondents around 1900	Digitizing Belle van Zuylen's Correspondence	Sailing Letters	Gouda on Paper	Transcribe Bentham
Start of project	2009	2010	November 2011	November 2011	September 2010
End of project	November 2016	Ongoing	October 2012	Ongoing	Ongoing
No. of recruited (registered) citizens	20	7	100	60	3.000
No. of active* citizen participants	20	5	100	50	11
Type of documents	Letters	Letters	Letters	Books and Manuscripts	Manuscripts and letters
Scope of project **	1,912 letters	1,762 letters	5,862 letters	1,000 pages	15,634 pages
Type of tasks	Adding metadata Transcribing Annotating	Modernizing Annotating	Adding meta-data Transcribing	Transcribing Translating	Transcribing Encoding

* Refers to people who have been actively engaged in the project during the period of study.

** Number of letters or pages transcribed or translated in the project up to March 2016.

also led by a professional researcher and a research assistant from the same institute. Citizen contributors were, at the time of writing, members of an association interested in the work of this 18th-century writer. They use e-mail as means of communication and a web-based tool to integrate all contributions into one searchable online edition. The third project, *Sailing Letters*, started in 2011 and took just over a year to transcribe about 5,800 scans of handwritten documents from the 17th and 18th centuries. Participation was open to everyone who felt capable of carrying out this task. About 100 citizen volunteers contributed to the project.

The fourth project, *Gouda on Paper*, was initiated and led by one expert volunteer with transcribing experience and an educational background in language and literature, and a professional archivist from the regional archive. The call for participants took place in November 2011 through local media. Participation is open to anyone who feels capable of performing the proposed tasks. During our study, the project counted 50 active participants. They use various technologies to support their tasks: e-mail, Dropbox, and a web-based tool to integrate transcriptions. Finally, the *Transcribe Bentham* project of University College London started in 2010 with the aim of transcribing and encoding (TEI-compliant XML) the handwritten original work of Jeremy Bentham, to support existing research projects (Causser, Tonra, and Wallace 2012; Causser, Grint, Sichani and Terras 2018.). Participation is open to everyone, and from October 2012 to June 2014, about 400 people had transcribed or partially transcribed at least one manuscript, and of these 11 had transcribed 100 folios or more. The project uses an online transcription environment (based on open source software) where all (diplomatic) contributions are posted and integrated. All these projects are collaborative (Shirk et al. 2012) because they were designed by project leaders of different institutions and citizens volunteered to transcribe,

translate, or annotate data—tasks that entail the analysis and processing of textual data.

Data collection

We followed the activities in these five projects over a period of more than two years. Data were collected by the first author between December 2012 and December 2015. Our data (**Table 2**) consist of: semi-structured interviews (over 26 hours) with project leaders and volunteer citizens, observations of meetings and training sessions (45 hours), and documents including project manuals, screenshots of website pages, news articles, and other project-related documents (83 documents). Several interviews were conducted via Skype, follow-up information and clarifications were obtained via e-mail and telephone, and numerous documents were gathered to complement and triangulate findings.

Semi-structured interviews allowed for consistency in the topics covered across cases and for flexibility in adjusting questions depending on the type of interviewees and the flow of the conversation (Weiss 1994; Patton 2002). The length of the interviews was about one hour on average. We interviewed project leaders and through them gained access to citizen participants. Project leaders explained the ways of working, provided supporting documentation, and allowed observations of project meetings and training sessions. Interviews with project leaders took place in their offices, while interviews with citizens were held in their homes or via Skype. Most of the interviews (27) were taped (upon prior informed consent) and transcribed verbatim, and notes were taken of the informal conversations.

Observations of meetings and trainings facilitated our understanding of the activities and the dynamics of collaboration among participants and with project leaders. In general, the first author took an observer-as-participant role (Gold 1958). Project members knew about her

Table 2: Data sources.

Source	Letters and Correspondents around 1900	Digitizing Belle van Zuylen's Correspondence	Sailing Letters	Gouda on Paper	Transcribe Bentham
Interviews (formal and informal)	9	7	6	7	9
Observations*	–	6	–	15	–
Documents					
Manual (versions)	5	5	6	12	3
Website/blog	1	5	1	4	2
Other project documents	1	1	4	12	2
News articles (incl. recruiting open calls)	–	1	2	4	1
Minutes of meetings	–	1	–	7	2

* Projects with no observations had no training sessions during the research period.

presence, which allowed her to freely observe and ask questions to get to know participants, to clarify and understand their activities and their use of the online tools. For the Transcribe Bentham project, given the lack of training sessions and the broad geographical distribution of participants, the first author examined the transcription desk and performed a few partial transcriptions, which helped us to understand the type of task, its complexity, and the tool used to perform it. Since the focus of the research was understanding project leader activities in real-life projects, we did not quantitatively measure quality, nor was this possible because only one project kept track of the differences between contributed transcriptions and reviewed versions.

Data analysis

Data analysis was aimed at explanation building (Yin 2014) by identifying, describing, and connecting the activities performed by project leaders to manage knowledge work and ensure quality. First, through an iterative process, we coded the activities or work practices within these projects, including the main task and supporting activities. A distinction was made between activities performed by citizens and by project leaders. We also coded communication activities among citizens and between citizens and project leaders, and identified the technology and manuals used in each project.

Second, activities across projects were compared in terms of their purpose from a knowledge management perspective. For instance, training sessions and manuals were seen as different ways to share project leaders' expert knowledge with citizens. Third, we grouped activities with the same purpose and assessed similarities/differences across projects (Eisenhardt 1989). We compared the different ways (i.e., patterns) in which participants were recruited, how project leaders shared knowledge, how tasks were performed, how quality was assessed, and the roles of technology. Finally, we looked for similarities and differences (Eisenhardt 1989, 1991) among the cases to explain (Yin 2014) how and why specific combinations of activities were chosen on a case-by-case basis.

Findings

In the studied projects, textual data are interpreted and processed by citizen volunteers. The expected outcomes are transcriptions of such quality that they can be used in further academic research. In this context, quality involves two essential characteristics: accuracy, the match between processed information and the original object of research; and uniformity, the standardized way in which data is presented.

When asked what a good quality transcription entailed, one project leader described accuracy as, "...all the letters in the old form are converted into letters in the modern form."

Similarly, one citizen volunteer explained: "In the transcription you have these really old textual characters, so you convert them into modern day writing... so, that strange curl, is it an 'L', is it a 'B'? In the end there is only one character. It's about finding the right letter."

Another citizen said, "For instance, the word "immediately," that's with double "m" but [author] writes it with one "m" and you can think that you know how it should be, but it's not how it's written."

Accurate transcriptions require, at least, having the basic cultural competence to recognize handwritten characters and the structure of sentences. Moreover, understanding the language of the period in which manuscripts are written (e.g., Latin, 17th-century Dutch, 18th-century French) helps in the interpretation of textual characters, and can contribute to a greater accuracy.

Uniformity refers to the presentation of textual information in a standard manner. When explaining what makes a good transcription, a project leader said, "following the guidelines. Because you must, of course if you work with lots of different people, well... stick to the agreements. So, for example, what do you do with indentation? And underlining? And what do you do with words that you cannot read?"

Standardization or uniformity requires citizen volunteers to be aware of the rules of the field and the project, and to be thorough in applying them consistently.

Standardization is important for quality because information needs to be searchable and allow aggregation (i.e., into periods, authors, location) to support research analyses.

In the subsections that follow, we discuss the knowledge-management activities used to facilitate quality contributions.

Accessing knowledge: recruiting participants

We distinguished two types of recruiting approaches in the studied projects: an open call, in the form of a public announcement of the project; and a targeted call, where the invitation to participate was directed to only a specific group of people (Table 3). We also found that some projects used both types of call. More importantly, our findings indicate that the different ways of recruiting participants influenced the number and type of people, and the knowledge they brought into the project.

The projects Gouda on Paper and Transcribe Bentham used a true open call to recruit participants. In both cases, project leaders announced the project through various media and set no restrictions for participation, thus creating a greater pool of potential participants. In contrast, the projects Letters and Correspondents around 1900 and Digitizing Belle van Zuylen’s Correspondence recruited people within the networks of their respective project leaders. That is, they targeted people who they thought would be interested or who they knew had relevant knowledge to perform the main task. The project Letters and Correspondents around 1900 targeted primarily university students with a history and literature background. Digitizing Belle van Zuylen’s Correspondence recruited participants among the members of the long-established association dedicated to the work of this female author. Finally, the Sailing Letters project targeted citizens who had participated in a previous citizen science project, namely transcribing a 17th-century bible, but additional people joined after hearing about the project in the media or through the project leader’s network.

At the time of our study, the projects Gouda on Paper, Sailing Letters, and Transcribe Bentham had reached about 50, 100, and 400 contributors respectively with their open call approach, whereas the targeted call of Digitizing Belle van Zuylen’s Correspondence and Letters and Correspondents around 1900 enabled them to recruit 7 and 20 participants respectively.

Project leaders and citizen participants were very much aware of self-selection. For instance, one project leader said, “Volunteers are not selected by me [...] everyone who wants to contribute can do that [...] though they should believe that they can do it.” Volunteers explained their decision to participate with comments such as, “I have gained a lot of experience in these 20 years [...] most people who enrol [in project] are very interested, they are well-educated. So, most of them know that they can handle this [task].” Another volunteer commented, “And because we are very interested in [author] and because we thought that we had some knowledge that could be useful, we said: let’s do it!”. That is, they referred to the fit between the project and their knowledge and interests.

Sharing and integrating knowledge

Following or parallel to recruiting citizens, project leaders shared their expert knowledge and facilitated communication among citizens through trainings, manuals, regular online communication, meetings, or online forums. Training was used in some projects (Table 3) to teach participants basic transcription and annotation norms, to agree on standardization rules, and to become familiar with the online tools or work environment. One project leader explained, “It is about a workshop we had twice. It was mainly technical, how it works, and after that we had one about how to actually use it. Because you transcribe, but how should you do that? A note here is different than when you put it on paper. How do you do that in the system?”

Regardless of how knowledgeable citizen participants were, project leaders provided training to ensure that transcriptions were standardized and to avoid problems

Table 3: Different configurations of knowledge management activities.

	Letters and Correspondents around 1900	Digitizing Belle van Zuylen’s Correspondence	Sailing Letters	Gouda on Paper	Transcribe Bentham
Accessing knowledge: recruiting participants	Targeted call	Targeted call	Targeted and open calls	Open call	Open call
Sharing and integrating knowledge	Training Regular comm. Manual	Training Regular comm. Meetings Manual	Regular comm. Forum Manual	Training Regular comm. Meetings Manual	Regular comm. Forum Manual
Coordinating knowledge: organizing tasks	Individual	Individual	Individual Rotation (in workflow)	Group discussions	Individual Rotation (optional)
Assessing knowledge work: evaluating contributions’ quality	Peer-reviews and Peer-expert reviews	Professional-expert review	Peer-expert reviews	Peer-expert reviews	Professional-expert reviews

with integrating multiple contributions. Training sessions were not intended to teach participants about the content or language of the text, but rather were aimed at sharing project leaders' expert knowledge about transcription conventions and using the online transcription environment. In projects with larger numbers of people, training sessions at the research institute were not organized very often; instead, project leaders chose manuals and other supporting materials that participants could use online before or while performing the task.

In fact, all projects used manuals or guidelines to share knowledge. Manuals included rules that were too broad or too specific to a transcription methodology and could not be embedded in the technology. That is, the extent to which knowledge was codified and standardized in technological artefacts (in metadata fields, encoding buttons, or drop-down lists) influenced the rules included in manuals. For instance, manuals included rules to standardize dates, spelling, and punctuation, and explained when and how to solve abbreviations. One of the manuals indicated, for example, how to enter dates: "Look whether you can find a date on the letter and fill it in, in the order: day, month, year." Similarly, in another project the guideline stated: "Date: (of the letter, this order holds for all the dates in the metadata!) 17531216 (letter 0010) yyymmdd. In case the date is not complete, then write as follow: 175312??."

Manuals could either substitute training or be an additional means to communicate project rules. They were used during training sessions and distributed to participants before engaging in the task, to ensure awareness of rules and expectations. Manuals were also used during task performance, as a reference in case of doubt and to ensure contributions fulfilled expected criteria, and after task completion, to assess and improve submitted contributions.

Manuals were revised throughout a project, on the basis of discussions during training sessions or frequently asked questions. For instance, one of the project manuals explicitly stated, "Instructions are by definition work in progress: they are modified on the basis of questions, comments, specific user cases and new insights." This was mainly the case during pilot phases of the projects, as project leaders tried to find the best way to codify knowledge and communicate field conventions and standardization rules. Another project explained the manual's work-in-progress status in its newsletter: "We are not there yet, so the manual is not final. Some issues will come from practice and they will be modified. That happened during the training evening, when we found some problems that we had not seen before. These have been added in the manual right away."

Moreover, because of the variability in citizens' knowledge and skills, some manuals offered extra supporting information. For example, the Sailing Letters manual included a detailed list of abbreviations common in 17th- and 18th-century documents, and links to specialized websites. Similarly, participants in Transcribe Bentham had access to online examples of Bentham's handwriting.

Knowledge was also shared through regular online communication. Project leaders or coordinators answered questions and resolved issues that citizens encountered while performing the task. As one participant explained, "If you are not sure about a thing or you put something, highlight it as questionable, or you do not quite understand it, you are not sure whether your reading of it was correct, you just put a little question and they will always get back to you." And one of the manuals stated, "In case of problems and special issues that are not covered in the manual, please contact the project leader."

Regular communication also included instances of feedback, about the quality of contributions and advice on how to improve them in the future. Citizen participants appreciated regular communication from project leaders, specially feedback and prompt reaction to questions, as evidenced by this participant's statement: "You can always turn to [project leader] with questions. [Project leader] answers quickly, I was really amazed, and if that is not the case then it is for a good reason and you get an answer quite soon. This is really nice, because you are busy [with task] and if you do not know something, it is really convenient that someone gives you the answer right away, then you can move on [with the task]. This is good, I like it."

Meetings and online discussion forums were organized to support knowledge-sharing and interaction among citizen participants. Projects Digitizing Belle van Zuylen's Correspondence, Gouda on Paper, and Letters and Correspondents around 1900 organized meetings most often. In Gouda on Paper, where tasks were performed in groups, proximity allowed regular meetings between representatives of each group (i.e., a group coordinators meeting). In these meetings, coordinators provided a group progress update, discussed problems, and tried to find solutions. In Sailing Letters and Transcribe Bentham, the larger number of participants and their wide geographical distribution meant that fewer face-to-face meetings were organized, and projects offered instead an online discussion forum.

Coordinating knowledge: organizing tasks

We distinguished three different approaches to organize the transcription and translation of manuscripts: individual, group discussion, and individual rotation (**Table 3**). In projects with a larger number of people and few prior knowledge requirements, such as Gouda on Paper, Sailing Letters, and Transcribe Bentham, project leaders allowed or actively encouraged the informal revision of transcriptions by rotating texts among participants or discussing them in groups. Both the group discussions and the rotation of transcripts had the same objective: having multiple people perform the main task (transcribing, translating) to improve quality. The rotation of transcripts was explained as follows:

"... as a second step we have the transcription, these have been rotated twice, still among volunteers, then is the level ... it gets better all the time [...] it can also happen that the second volunteer

is not better than the first one, so he might add little to it, but it can also be that he does actually see something... you just get the chance. After that another volunteer goes over it and then it [the text] is removed from this process."

The choice between rotating or discussing in groups was influenced by the proximity of participants and the technology used in the project. That is, if the online (transcription) tool used in the project affords versioning, this will facilitate the rotation of texts and their corresponding transcription, by tracking changes and deciding on the best transcription. If the tool does not allow versioning, rotating the transcription of texts becomes more complex and requires more coordination among participants. The studied projects had different levels of versioning, ranging from saving various versions of a transcription, keeping track of daily changes, to tracking changes at the word level. In Gouda on Paper, for instance, given the number of participants and their proximity, the project leader urged citizens to organize groups. However, technology did not afford word-level versioning; hence, groups first transcribed (or translated) the text individually in Word; then at an agreed date, participants met to compare their individual work, discuss it, produce the best transcription (or translation) possible, and add it in the online tool. In the Sailing Letters project, rotation was part of the normal workflow, organized in steps, and for each step the transcription versions were saved. Transcribe Bentham was the only case in which word-level versioning was possible. Surprisingly, despite the possibility to track and reverse word changes, very few people worked on transcriptions started by others (i.e., rotation), mostly preferring to start transcriptions from scratch. In contrast, projects based on a targeted call, such as Digitizing Belle van Zuylen's Correspondence and Letters and Correspondents around 1900, used tools that did not afford word-level versioning, and citizen participants therefore mainly worked individually.

Regardless of how tasks were organized—individually, through group discussion, or in rotation—quality was primarily accomplished through individual task performance. Some individuals proofread, assessed, and improved the quality of their own transcriptions before saving or submitting them in the online work environment. One volunteer said, "I tend to go through it two or three times to figure out what the gaps are, what have I missed out. I do that as part of a proofreading process to check it all: does it all make sense? is it something that the editors will find semi-useful at least?"

Other volunteers performed their work all in one go, very carefully, so that they felt confident enough to submit their work without proofreading, as the best they were able to do.

Assessing knowledge work: evaluating contributions' quality

In all the studied projects, citizen contributions were assessed and improved. We identified two assessing approaches: professional-expert reviews and peer-expert reviews (Table 3). Professional-expert reviews were car-

ried out in the projects Transcribe Bentham and Digitizing Belle van Zuylen's Correspondence, where tasks were performed individually online. Contributions were assessed and improved individually or by a small group of two or three professional researchers. Though the number and distribution of participants was greater in Transcribe Bentham than in Digitizing Belle van Zuylen's Correspondence, only a smaller group of citizens transcribed regularly and did not know each other. Therefore, it seemed more efficient to let citizens focus on the core task and leave the assessment and correction to the professional project staff.

Gouda on Paper initially used professional-expert reviews to assess and correct participants' contributions. Over time, however, peer-expert reviews became a better option for the project. Project leaders were not able to keep up with the high number of transcriptions, and there were also more people transcribing than translating manuscripts, which resulted in workflow disconnections. Most importantly, the need to make transcriptions and translations available online to the public meant that all contributions needed to be assessed and corrected more quickly. Therefore, committees or teams of peer-experts were organized to assess the accuracy of transcriptions, check interpretations and corresponding translations, review the language of translations, and improve the readability for present-day people. The creation of committees was explained in the project's newsletter:

"We want to ask [research institute] to publish the transcribed and translated texts in [online tool]. Before we do that, we need to thoroughly go over everything again. This should be done by people with the educational background, training or profession. We have these people in the project, spread over the different groups. We have asked them to participate in the committees that will perform this final control."

Such a comment in the newsletter indicates that project leaders were aware of the expertise level of participants.

Peer-expert reviews were an essential part of the Sailing Letters project. Participants who had a relevant educational or professional background (history, literature, linguistics) and extensive experience in transcribing were asked to review and improve preceding contributions. Peer-experts were targeted by checking their short biography, usually requested by the project leader when they joined the project, and their time availability. These assessment and correction tasks were also rotated among the peer-experts.

Finally, in Letters and Correspondents around 1900, reviews changed during the course of the project. Initially the project had three main steps: transcription, assessment, and final editing. However, transcriptions and reviews done by students were not always accurate and resulted in long discussions in the annotation field. Because of this, the project leader asked experienced volunteers to carry out a second assessment round. Hence, the project combined peer reviews with peer-expert reviews.

Discussion

We set out to investigate how citizen science projects involving complex tasks are managed and quality outcomes ensured. From a knowledge perspective, project leaders ensure quality by recruiting citizens (accessing knowledge), sharing and integrating their expert knowledge, coordinating knowledge work, and assessing and improving outcomes. Together these knowledge management processes contribute to the quality of citizen science outcomes (Figure 1).

While the citizen science literature recommends project leaders to announce projects through various communication channels to recruit people with different motivations (West and Pateman 2016), the knowledge management approach proposes other recruiting strategies based on knowledge access. That is, some project leaders access knowledge through targeted calls to reduce knowledge uncertainty and to increase the chances of quality outcomes. Targeted calls are based on the idea that only a subset of the public has the knowledge and interest to contribute to the production of scientific public goods (Wasko and Teigland 2004). These calls are based on the judgment made by professional scientists about citizens' knowledge. This assessment is influenced by prior knowledge and similar social identity (Kane, Argote, and Levine 2005; Lamb and Davidson 2005). Scientific project leaders seem to evaluate citizens on the basis of similarity between their educational and professional backgrounds, characteristics of their own social identity, to reduce knowledge uncertainty (Hogg 2001; Fiol and O'Connor 2005).

Targeted calls, however, contradict one of the main characteristics of citizen science: namely open participation or unrestricted entry (Franzoni and Sauermaun 2014), and they are not enough to guarantee quality. Whether targeted or not, citizens performing complex tasks, such as manuscript transcriptions, also need to fulfil scientific

quality standards. To facilitate this, project leaders share and integrate their expert knowledge, and coordinate and assess knowledge work. The configuration of knowledge management activities depends on the number, distribution, and knowledge diversity of recruited participants (Figure 1).

Prior citizen science research has recommended providing opportunities to learn (West and Pateman 2016), to give personalized feedback (Eveleigh et al. 2014), and to facilitate social interactions (Rotman et al. 2014) on the basis of the motivations of citizen participants. A knowledge perspective shows how these activities are related to knowledge, quality, and the choices made for recruiting participants. Knowledge-sharing activities in complex citizen science are similar to the way knowledge is shared in other organizational settings (Greenberg and Roseman 2003; Ackerman et al. 2013). Open calls make a project widely known and are likely to result in a larger number of distributed participants with diverse knowledge; hence, knowledge-sharing usually takes place online, through manuals and links to extra information sources. In contrast, targeted calls are more likely to lead to a smaller and more manageable group of participants with more relevant knowledge, who may perhaps be in closer physical proximity. In those cases, organizing face-to-face meetings and training sessions is more feasible (Figure 1). Moreover, expert knowledge is integrated in rules, standards, and routines, and is embedded in technology (Kogut and Zander 1992; Grant 1996; Davenport and Prusak 2000). But, because not all scientific knowledge is unambiguous, project leaders also share knowledge through interpersonal communication (Hislop 2013) such as meetings, trainings, and online forums. It seems that knowledge sharing in citizen science requires the combination of first-generation (i.e., manuals and standard procedures) and second-generation (i.e., learning within the

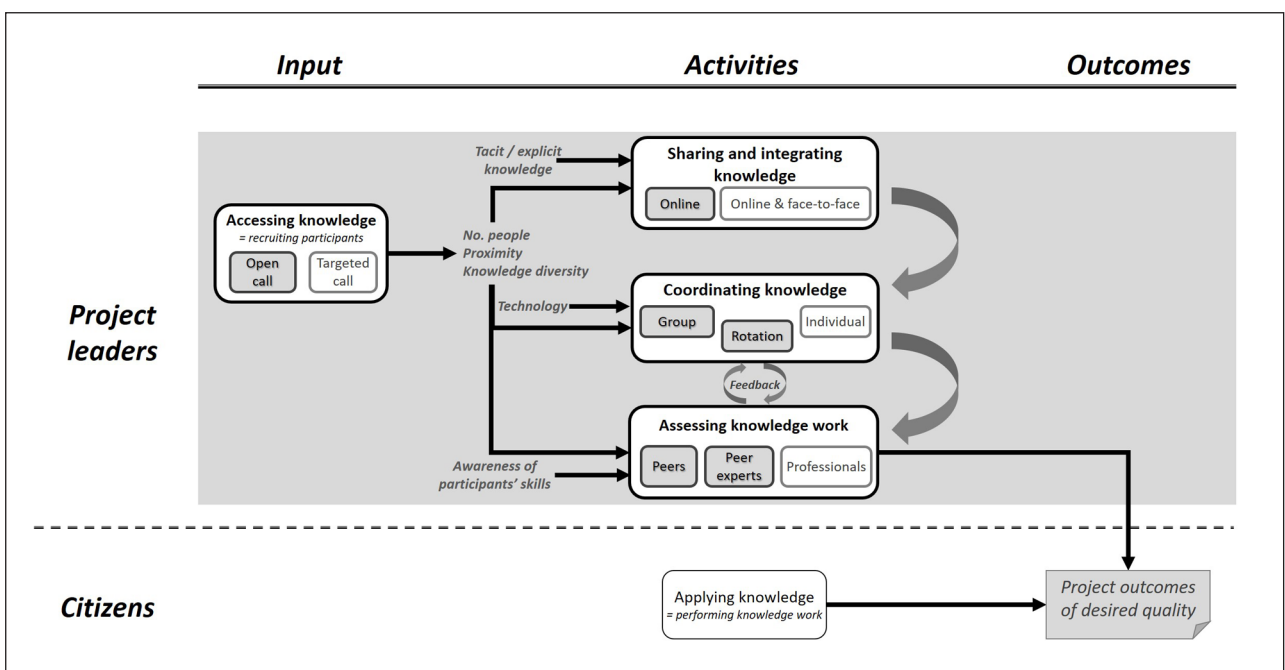


Figure 1: Possible configurations of knowledge-management activities in citizen science.

community of participants through training and meetings) knowledge-sharing practices (Ackerman et al. 2013).

To coordinate a large number of participants, project leaders are likely to organize tasks in a collaborative manner, through group discussions or task rotation (**Figure 1**). Recent research shows that collaborative transcription methods result in better quality than the aggregated majority of multiple individual transcriptions (Blickhan et al. 2019). A collaborative approach is in line with concepts such as the wisdom of crowds and Linus' law (Raymond 1999; Surowiecki 2005) by which quality improves as more people go over the same text. Rotating tasks or discussing in groups can be seen as different ways of organizing community revision (Brumfield 2012), each depending on the proximity of participants and the technology used. If people are geographically close, they can work in groups. If technology affords versioning, tasks are easier to rotate. This confirms prior research on distributed work, as the coordination and performance of citizen science tasks depends on the type of tasks and their dependencies (Mitchell et al. 2018), as well as on the distribution of participants and the affordances of technology (Franssila et al. 2012). Coordination through rotation and discussion in groups is also intertwined with the assessment of contributions, with some form of feedback depending on the possibilities of technology and the way tasks are organized (Dow et al. 2012).

Finally, prior citizen science literature indicates the importance of monitoring and evaluating citizen science projects (West and Pateman 2016), but it does not discuss how the quality of outcomes is assessed. Our study shows different ways by which project leaders manage the assessment and correction of contributions. Knowledge assessment is performed differently depending on the field, profession, and task at hand (Robertson, Scarbrough, and Swan 2003). In complex citizen science, assessment approaches are influenced by the number of participants and the extent to which project leaders are aware of citizens' level of expertise (**Figure 1**). If the number of participants is small, project leaders tend to rely on professional-expert reviews to assess contributions. Because professional reviews do not scale well (Wiggins et al. 2011; Dow et al. 2012), projects with a large number of citizen participants are likely to use peer (-expert) reviews, as long as the project leaders know participants' level of expertise. Multiple reviews seem to be the common way by which transcriptions (as outcomes of citizen science) are assessed (Brumfield 2012), which fits the interpretative nature of humanities fields such as literary studies (Blockmans 2018), requiring various views to agree on an outcome.

This study expands earlier citizen science frameworks (Shirk et al. 2012) by examining in detail which different configurations of activities project leaders can adopt to manage knowledge flows and ensure quality. First, project leaders should be aware of the consequences of choosing between an open versus a targeted call. Open calls are likely to lead to a greater number of diverse participants, which makes task coordination (Mitchell et al. 2018) and

quality assessment (Dow et al. 2012) challenging, while targeted calls may result in more manageable projects but with slower completion pace. And second, we show how knowledge management practices play a role in ensuring quality contributions from voluntary citizens. These practices might change, however, when new technologies are integrated into projects. For instance, in the digital humanities, the application of machine learning algorithms, such as Handwritten Text Recognition software (e.g., TRANSKRIBUS), reduces the complexity of transcriptions and therefore modifies the types of tasks that citizens perform and how these are coordinated (Hedges and Dunn 2018; Brumfield 2020). The assessment of contributions might also change over time as algorithms fitting natural language are used, for example, in text similarity software to compare multiple citizen contributions (e.g., project Mutual Muses by the Getty).

Conclusion

Knowledge-management practices contribute to address quality issues in citizen science. Project leaders access and harness the knowledge of citizen volunteers by applying multiple knowledge-management activities that facilitate the performance of complex tasks and ensure quality outcomes. The way knowledge is accessed seems to lay the foundations for the different configurations of activities aimed at ensuring quality. These configurations are also influenced by citizens' proximity, the characteristics of knowledge, the affordances of technology and the extent to which project leaders are aware of citizens' skills.

Though the depth and details of this study may be limited by the scope, which covers five projects, the findings include plenty of specifics about the different ways that knowledge flows and quality are managed in complex citizen science. Since the focus of the study has been on the activities adopted by project leaders, to gain a deeper understanding of knowledge-sharing in citizen science, future research could dig deeper into the interactions between citizen volunteers.

Finally, our study has not included any quantitative measures to assess task performance and quality. Different knowledge-management configurations might require different time investments and coordination efforts, aspects which are sometimes underestimated (Riesch and Potter 2014), or could result in different levels of quality. Future research could quantitatively measure the duration of projects and the quality of crowd contributions over time, to assess the efficiency and effectiveness of different knowledge-management practices and compare them across projects. Moreover, tracking the quality of citizens' contributions over time could provide information about the learning effect that occurs when people contribute to a project for an extended period of time.

Data Accessibility Statement

Given the qualitative nature of the research and to maintain anonymity of the participants, no data from the interviews, observations, and documents is made available, other than the quotes included in the article.

Acknowledgements

We thank project leaders and citizen participants of the studied projects for the time they dedicated to our research, for answering our questions, for showing us how tasks were performed, and for giving us access to data and documents. It has been a pleasure and a privilege to meet them all. We also thank all the members of the Gouda on Paper and Transcribe Bentham initiatives, as well as the management and employees of the Huygens Institute, the Meertens Institute, and the Dutch Cultural Heritage Agency for facilitating contacts and information.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

The design of the study was developed collectively by all the authors. Montserrat Prats López took the lead in carrying out the interviews, observations, and analysis, as well as in writing and revising the paper. Maura Soekijad, Hans Berends, and Marleen Huysman contributed to the data collection process and analysis, and edited the manuscript and revisions.

References

- Ackerman, MS, Dachtera, J, Pipek, V and Wulf, V.** 2013. Sharing Knowledge and Expertise: The CSCW View of Knowledge Management. *Computer Supported Cooperative Work*, 22(4): 531–573. DOI: <https://doi.org/10.1007/s10606-013-9192-8>
- Afuah, A and Tucci, CL.** 2012. Crowdsourcing as a solution to distant search. *Academy of Management Review*, 37(3): 355–375. DOI: <https://doi.org/10.5465/amr.2010.0146>
- Alavi, M and Denford, JS.** 2011. Knowledge Management: Process, Practice, and Web 2.0. In: Easterby-Smith, M and Lyles, MA. (eds.), *Handbook of Organizational Learning and Knowledge Management*. Hoboken, N.J.: Wiley, pp. 105–124. DOI: <https://doi.org/10.1002/9781119207245.ch6>
- Alvesson, M.** 2001. Knowledge Work: Ambiguity, Image and Identity. *Human Relations*, 54(7): 863–886. DOI: <https://doi.org/10.1177/0018726701547004>
- Berends, H.** 2005. Exploring knowledge sharing: moves, problem solving and justification. *Knowledge Management Research and Practice*, 3 (2): 97–105. DOI: <https://doi.org/10.1057/palgrave.kmrp.8500056>
- Blickhan, S, Krawczyk, C, Hanson, D, Boyer, A, Simenstad, A, et al.** 2019. Individual vs. Collaborative Methods of Crowdsourced Transcription. *Journal of Data Mining and Digital Humanities*, Special Issue on Collecting, Preserving, and Disseminating Endangered Cultural Heritage for New Understandings through Multilingual Approaches. Available at: <https://jdmdh.episciences.org/5759> (accessed 26 May 2020).
- Blockmans, W.** 2018. Two Cultures into One? *European Review*, 26(2): 233–240. DOI: <https://doi.org/10.1017/S1062798717000631>
- Bonney, R, Cooper, C and Ballard, H.** 2016. The theory and practice of citizen science: Launching a new journal. *Citizen Science: Theory and Practice*, 1(1): 1, 1–4. DOI: <https://doi.org/10.5334/cstp.65>
- Bordogna, G, Carrara, P, Criscuolo, L, Pepe, M and Rampini, A.** 2014. A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science. *Information Sciences*, 258: 312–327. DOI: <https://doi.org/10.1016/j.ins.2013.07.013>
- Brumfield, BW.** 2012. *Quality Control for Crowdsourced Transcription*. Available at: <http://manuscripttranscription.blogspot.nl/2012/03/quality-control-for-crowdsourced.html> (accessed 3 November 2014).
- Brumfield, BW.** 2020. *The Decade in Crowdsourcing Transcription*. Available at: <https://content.fromthepage.com/decade-in-crowdsourcing/> (accessed 26 May 2020).
- Causser, T, Grint, K, Sichani, AM and Terras, M.** 2018. 'Making such bargain': Transcribe Bentham and the quality and cost-effectiveness of crowdsourced transcription. *Digital Scholarship in the Humanities*, 33(3): 467–487. DOI: <https://doi.org/10.1093/llc/fqx064>
- Causser, T, Tonra, J and Wallace, V.** 2012. Transcription maximized; expense minimized? Crowdsourcing and editing The Collected Works of Jeremy Bentham. *Literary and Linguistic Computing*, 27(2): 119–137. DOI: <https://doi.org/10.1093/llc/fqs004>
- Cooper, CB, Dickinson, J, Phillips, T and Bonney, R.** 2007. Citizen Science as a Tool for Conservation in Residential Ecosystems. *Ecology and Society*, 12(2): 11–21. DOI: <https://doi.org/10.5751/ES-02197-120211>
- Davenport, TH and Prusak, L.** 2000. *Working knowledge: How organizations manage what they know*. Boston: Harvard Business School Press.
- De la Flor, G, Jirotko, M, Luff, P, Pybus, J and Kirkham, R.** 2010. Transforming Scholarly Practice: Embedding Technological Interventions to Support the Collaborative Analysis of Ancient Texts. *Computer Supported Cooperative Work*, 19(3–4): 309–334. DOI: <https://doi.org/10.1007/s10606-010-9111-1>
- Dow, S, Kulkarni, A, Klemmer, S and Hartmann, B.** 2012. Shepherding the crowd yields better work. In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12)*. New York, NY, USA: ACM, pp. 1013–1022. DOI: <https://doi.org/10.1145/2145204.2145355>
- Dunn, S and Hedges, M.** 2013. Crowd-sourcing as a Component of Humanities Research Infrastructures. *International Journal of Humanities and Arts Computing*, 7(1–2): 147–169. DOI: <https://doi.org/10.3366/ijhac.2013.0086>
- Dunn, S and Hedges, M.** 2014. How the Crowd Can Surprise Us: Humanities Crowdsourcing and the Creation of Knowledge. In Ridge, M. (ed.), *Crowdsourcing Our Cultural Heritage*. Farnham: Ashgate Publishing Ltd, pp. 231–246.

- Eisenhardt, KM.** 1989. Building Theories from Case Study Research. *Academy of Management Review*, 14 (4): 532–550. DOI: <https://doi.org/10.5465/amr.1989.4308385>
- Eisenhardt, KM.** 1991. Better stories and better constructs: The case for rigor and comparative logic. *Academy of Management Review*, 16(3): 620–627. DOI: <https://doi.org/10.5465/amr.1991.4279496>
- Eveleigh, A, Jennett, C, Blandford, A, Brohan, P and Cox, AL.** 2014. Designing for dabblers and deterring drop-outs in citizen science. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, Toronto, Ontario, Canada, pp. 2985–2994. DOI: <https://doi.org/10.1145/2556288.2557262>
- Fiol, CM and O'Connor, EJ.** 2005. Identification in Face-to-Face, Hybrid, and Pure Virtual Teams: Untangling the Contradictions. *Organization Science*, 16(1): 19–32. DOI: <https://doi.org/10.1287/orsc.1040.0101>
- Franssila, H, Okkonen, J, Savolainen, R and Talja, S.** 2012. The formation of coordinative knowledge practices in distributed work: towards an explanatory model. *Journal of Knowledge Management*, 16(4): 650–665. DOI: <https://doi.org/10.1108/13673271211246202>
- Franzoni, C and Sauermann, H.** 2014. Crowd science: The organization of scientific research in open collaborative projects. *Research Policy*, 43(1): 1–20. DOI: <https://doi.org/10.1016/j.respol.2013.07.005>
- Freitag, A, Meyer, R and Whiteman, L.** 2016. Strategies Employed by Citizen Science Programs to Increase the Credibility of Their Data. *Citizen Science: Theory and Practice*, 1(1): 2, 1–11. DOI: <https://doi.org/10.5334/cstp.6>
- Gold, RL.** 1958. Roles in Sociological Field Observations. *Social Forces*, 36(3): 217–223. DOI: <https://doi.org/10.2307/2573808>
- Grant, RM.** 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(Winter Special Issue): 109–122. DOI: <https://doi.org/10.1002/smj.4250171110>
- Greenberg, S and Roseman, M.** 2003. Using a room metaphor to ease transitions in groupware. In: Ackerman, M. S., Pipek, V. and Wulf, V. (eds.), *Sharing Expertise: Beyond Knowledge Management*. Cambridge, MA: MIT Press, pp. 203–256.
- Haas, MR and Hansen, MT.** 2007. Different knowledge, different benefits: Toward a productivity perspective on knowledge sharing in organizations. *Strategic Management Journal*, 28(11): 1133–1153. DOI: <https://doi.org/10.1002/smj.631>
- Haythornthwaite, C.** 2009. Crowds and communities: Light and heavyweight models of peer production. In: *Proceedings of the 42nd Hawaii International Conference on System Sciences*, pp. 1–10. DOI: <https://doi.org/10.1109/HICSS.2009.137>
- Hedges, M and Dunn, S.** 2018. *Academic crowdsourcing in the humanities: crowds, communities and co-production*. Cambridge, MA: Elsevier. <https://www.science-direct.com/book/9780081009413/academic-crowdsourcing-in-the-humanities>.
- Hislop, D.** 2008. Conceptualizing Knowledge Work Utilizing Skill and Knowledge-based Concepts: The Case of Some Consultants and Service Engineers. *Management Learning*, 39(5): 579–596. DOI: <https://doi.org/10.1177/1350507608098116>
- Hislop, D.** 2013. *Knowledge management in organizations: A critical introduction*. Oxford: Oxford University Press.
- Hislop, D, Bosua, R and Helms, R.** 2018. *Knowledge management in organizations: A critical introduction*. 4th ed. Oxford: Oxford University Press.
- Hogg, MA.** 2001. Social identity and the sovereignty of the group. In: Sedikides, C and Brewer, MB (eds.), *Individual self, relational self, collective self*. London: Routledge, pp. 123–143.
- Huber, GP.** 1991. Organizational learning: The contributing processes and the literatures. *Organization Science*, 2(1): 88–115. DOI: <https://doi.org/10.1287/orsc.2.1.88>
- Kane, AA, Argote, L and Levine, JM.** 2005. Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality. *Organizational Behavior and Human Decision Processes*, 96(1): 56–71. DOI: <https://doi.org/10.1016/j.obhdp.2004.09.002>
- Kittur, A, Nickerson, JV, Bernstein, M, Gerber, E, Shaw, A, Zimmerman, J, Lease, M and Horton, J.** 2013. The future of crowd work. In: *Proceedings of the 2013 conference on Computer supported cooperative work (CSCW '13)*. New York, NY, USA: ACM, pp. 1301–1318. DOI: <https://doi.org/10.1145/2441776.2441923>
- Kogut, B and Zander, U.** 1992. Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science*, 3(3): 383–397. DOI: <https://doi.org/10.1287/orsc.3.3.383>
- Lamb, R and Davidson, E.** 2005. Information and communication technology challenges to scientific professional identity. *The Information Society*, 21(1): 1–24. DOI: <https://doi.org/10.1080/01972240590895883>
- Law, E, Gajos, KZ, Wiggins, A, Gray, ML and Williams, A.** 2017. Crowdsourcing as a Tool for Research: Implications of Uncertainty. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. New York, NY, USA: ACM, pp. 1544–1561. DOI: <https://doi.org/10.1145/2998181.2998197>
- Miller, S.** 2001. Public understanding of science at the crossroads. *Public Understanding of Science*, 10(1): 115–120. <https://journals.sagepub.com/doi/10.3109/a036859>.
- Mitchell, EM, Crowston, K and Østerlund, C.** 2018. Coordinating advanced crowd work: Extending citizen science. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*. DOI: <https://doi.org/10.24251/HICSS.2018.212>
- Oomen, J and Aroyo, L.** 2011. Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges. In: *Proceedings of the 5th International Conference on Communities and Technologies*, pp. 138–149. DOI: <https://doi.org/10.1145/2103354.2103373>
- Oosterman, J, Bozzon, A, Houben, GJ, Nottamkandath, A, Dijkshoorn, C, Aroyo, L, Leyssen, MHR and Traub, MC.** 2014. Crowd vs. experts: nichesourcing for knowledge intensive tasks in cultural heritage. In: *Proceedings of the 23rd International Conference on World*

- Wide Web*, Seoul, Korea, pp. 567–568. DOI: <https://doi.org/10.1145/2567948.2576960>
- Patton, MQ.** 2002. *Qualitative research and evaluation methods*. 3rd ed. Thousand Oaks, Calif.: Sage.
- Ponciano, L** and **Brasileiro, F.** 2014. Finding Volunteers' Engagement Profiles in Human Computation for Citizen Science Projects. *Human Computation*, 1(2): 247–266. <https://arxiv.org/abs/1501.02134>.
- Puranam, P, Alexy, O** and **Reitzig, M.** 2014. What's "new" about new forms of organizing? *Academy of Management Review*, 39(2): 162–180. DOI: <https://doi.org/10.5465/amr.2011.0436>
- Raymond, E.** 1999. The cathedral and the bazaar. *Knowledge, Technology & Policy*, 12(3): 23–49. DOI: <https://doi.org/10.1007/s12130-999-1026-0>
- Riesch, H** and **Potter, C.** 2014. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Understanding of Science*, 23(1): 107–120. DOI: <https://doi.org/10.1177/0963662513497324>
- Robertson, M, Scarbrough, H** and **Swan, J.** 2003. Knowledge creation in professional service firms: Institutional effects. *Organization Studies*, 24(6): 831–831. DOI: <https://doi.org/10.1177/0170840603024006002>
- Rotman, D, Hammock, J, Preece, J, Hansen, D, Boston, C, Bowser, A** and **He, Y.** 2014. Motivations Affecting Initial and Long-Term Participation in Citizen Science Projects in Three Countries. In: *iConference 2014 Proceedings*, Berlin, Germany, pp. 110–124. DOI: <https://doi.org/10.9776/14054>
- Sauermann, H** and **Franzoni, C.** 2015. Crowd science user contribution patterns and their implications. In: *Proceedings of the National Academy of Sciences*, 112(3): 679–684. DOI: <https://doi.org/10.1073/pnas.1408907112>
- Sheppard, SA, Wiggins, A** and **Terveen, L.** 2014. Capturing quality: retaining provenance for curated volunteer monitoring data. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work and social computing*, Baltimore, Maryland, USA, pp. 1234–1245. DOI: <https://doi.org/10.1145/2531602.2531689>
- Shirk, JL, Ballard, HL, Wilderman, CC, Phillips, T, Wiggins, A, Jordan, R, Mccallie, E, Minarchek, M, Lewenstein, BV, Krasny, ME** and **Bonney, R.** 2012. Public Participation in Scientific Research: a Framework for Deliberate Design. *Ecology and Society*, 17(2): 29. DOI: <https://doi.org/10.5751/ES-04705-170229>
- Simula, H.** 2013. The Rise and Fall of Crowdsourcing? In: *Proceedings of the 46th Hawaii International Conference on System Sciences*, pp. 2783–2791. DOI: <https://doi.org/10.1109/HICSS.2013.537>
- Sodré, I** and **Brasileiro, F.** 2017. An Analysis of the Use of Qualifications on the Amazon Mechanical Turk Online Labor Market. *Computer Supported Cooperative Work*, 26(4): 837–872. DOI: <https://doi.org/10.1007/s10606-017-9283-z>
- Surowiecki, J.** 2005. *The wisdom of crowds*. New York: Anchor Books.
- Wasko, MML** and **Teigland, R.** 2004. Public goods or virtual commons? Applying theories of public goods, social dilemmas, and collective action to electronic networks of practice. *Journal of Information Technology Theory and Applications*, 6(1): 25–41. <https://aisel.aisnet.org/jitta/vol6/iss1/4/>.
- Weiss, RS.** 1994. *Learning from strangers: The art and method of qualitative interview studies*. New York: The Free Press.
- West, S** and **Pateman, R.** 2016. Recruiting and Retaining Participants in Citizen Science: What Can Be Learned from the Volunteering Literature? *Citizen Science: Theory and Practice*, 1(2): 15, 1–10. DOI: <https://doi.org/10.5334/cstp.8>
- Wiggins, A** and **Crowston, K.** 2011. From Conservation to Crowdsourcing: A Typology of Citizen Science. In: *Proceedings of the 44th Hawaii International Conference on System Sciences*, pp. 1–10. DOI: <https://doi.org/10.1109/HICSS.2011.207>
- Wiggins, A, Newman, G, Stevenson, RD** and **Crowston, K.** 2011. Mechanisms for Data Quality and Validation in Citizen Science. In: *Proceedings of the 7th IEEE International Conference on e-Science Workshops*, pp. 14–19. DOI: <https://doi.org/10.1109/eScienceW.2011.27>
- Yin, RK.** 2014. *Case Study Research: Design and Methods*. 5th ed. Los Angeles: Sage.

How to cite this article: Prats López, M, Soekijad, M, Berends, H and Huysman, M. 2020. A Knowledge Perspective on Quality in Complex Citizen Science. *Citizen Science: Theory and Practice*, 5(1): 15, pp. 1–13. DOI: <https://doi.org/10.5334/cstp.250>

Submitted: 07 May 2019

Accepted: 07 June 2020

Published: 29 July 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <https://creativecommons.org/licenses/by/4.0/>.



Citizen Science: Theory and Practice is a peer-reviewed open access journal published by Ubiq Press.

OPEN ACCESS