Managing Citizen Science in the Humanities: The challenge of ensuring quality

Prats Lopez, M.

2017

document version
Publisher's PDF, also known as Version of record

1

Introduction

"All for one, one for all."
Alexandre Dumas (1802-1870)

Citizen science is a participative form of organizing and conducting scientific research, usually in the form of a project, where researchers and citizens work together to answer empirical questions (Cohn, 2008; Riesch & Potter, 2014; Wiggins & Crowston, 2011). In other words, citizen science involves professional and citizen scientists co-producing knowledge. Though citizen science projects can be initiated by public organizations, scientists or citizens (Wiggins & Crowston, 2011), the term 'citizen science' is usually associated with projects initiated by scientists who involve citizens in the collection and processing of data through the Internet (European Commission DG RTD and DG CNECT, 2014).

Citizen science is not a new phenomenon. The professionalization of science occurred in the 19th century (Olesko, 2008), and before then the people we now call 'scientists' were citizens with a wide range of professions (Silvertown, 2009). As such, citizens have been asked to help in solving organizational, societal and scientific problems for centuries. The idea behind citizen science is that, people can accomplish more by joining forces than alone. Our history is full of examples of people collaborating and combining their knowledge to solve scientific problems. A well-known example is the competition launched by the British Parliament in 1714 to find a way to determine longitude at sea (Afuah & Tucci, 2012; Jeppesen & Lakhani, 2010). The competition lasted for several years, was open to people of all nationalities and offered different prizes depending on the accuracy of the solutions contributed by participants (Johnson, 1989). The highest prize of £20,000 was awarded to John Harrison, a carpenter and clockmaker (Royal Museums Greenwich, 2015). Another example of a historical collaborative enterprise is the Oxford English Dictionary. In 1857, the Philological Society of London announced its intention to collect materials for a supplement to improve the dictionaries that existed at the time, and called for volunteers to read books and supply quotations to the editors (Gilliver, 2012). A year later, the editors decided to write a whole new dictionary and issued further open calls for volunteers.

Back then, participation in such projects involved volunteers writing and posting their paper-based contributions. Technological developments of the 20th century, and in particular the Internet, have opened up the range of possibilities for engaging a greater number of diverse people to come up with solutions to problems or to support collective endeavours. In 2014, about 78% of households in the European Union had Internet access through broadband (Eurostat, 2015). As the Internet and other technological artefacts have become cheaper and more accessible to both organizations and the wider public, the possibilities for distributed work have grown and new forms of organizing, such as citizen science, have emerged (Puranam, Alexy & Reitzig, 2014). Following on from the idea that larger groups of people can do more than one person or one organization on its own, an increasing number of individuals and institutions have taken advantage of the Internet to call for help in solving scientific problems. The Internet allows individual researchers or organizations to reach, coordinate and integrate the contributions of many people at a relatively low cost (Afuah & Tucci, 2012). New technologies are therefore an essential part of modern citizen science.

Though the ideas behind citizen science existed long before the Internet, it is because of such technological developments that the number of citizen science projects has grown in recent years (Wiggins & Crowston, 2015). The Internet is indeed increasingly used to support citizen science projects in different disciplines because it allows to reach a greater number of potential citizen participants. For instance, citizens participate in astronomy projects to understand the universe, by classifying images of galaxies (e.g.GalaxyZoo.org); they also contribute to the conservation of nature, by monitoring the number and distribution of birds (e.g. eBird.org); they support the advancement of medical research, by performing game-like activities to increase understanding of the structure of proteins (e.g. Fold.it); and make historical materials widely accessible and searchable, by transcribing handwritten letters and manuscripts of well-known people and making them available online (e.g. Transcribe Bentham). Yet, the increasing use of technology is a tendency of our society in general (Fiol & O'Connor, 2005) and not an attribute of any specific type of citizen science (Wiggins & Crowston, 2011), because all citizen science projects use technology and the Internet, albeit with different intensity and for different purposes.

The scientific problems that citizens are asked to help with are not only broad in terms of the types of topics, but also differ in the types of tasks: information gathering and processing, or ideation and experimentation (Boudreau & Lakhani, 2013; Brabham, 2013). The advantages of involving large numbers of citizens for ideation and experimentation-based projects lie in the access to a great diversity of knowledge that otherwise could not be reached; while the advantages for projects aimed at gathering and processing data are the time and resource efficiency with which these tasks can be achieved (Brabham, 2013; Franzoni & Sauermann, 2014; Shirk et al., 2012).

Research organizations can benefit from engaging in citizen science through the potential time and resource efficiency that results from involving large numbers of people to voluntarily collect, process or analyse information, perform a task or solve a problem (Brabham 2013; Franzoni & Sauermann, 2014). Moreover, the recent economic downturn has increased the pressure on public universities and research institutes to find diverse ways

to fund research projects (Estermann & Claeys-Kulik, 2013). The advantage of engaging in citizen science is that, in view of the limited resources of public research institutes, these projects allow them to perform empirical research that could not have been done without citizens contributing their time and knowledge to science (Riesch & Potter, 2014). Citizen science also allows a greater diversity of knowledge, the possibility of reaching individuals with specialized or rare knowledge, and the chance to detect errors, verify results and build upon each other's contributions (Franzoni & Sauermann, 2014; Jeppesen & Lakhani, 2010; Wexler, 2011).

Citizen science is also a means to improve the public understanding of science. Given the widespread access to information through the Internet, the general public is increasingly questioning the claims of scientific experts (European Commission, 2008). Educating the 'lay' public through unidirectional communication does not seem to be effective in improving public understanding of and engagement with science (European Commission, 2008; Haywood & Besley, 2014). Citizen science projects are not only initiated to support specific research questions, but often also include educating the public as one of their implicit or explicit objectives (Wiggins & Crowston, 2011).

1.1 The need for quality

Despite the importance and benefits of citizen science for researchers and society in general, the quality of citizen science outcomes remains a point of concern (Oomen & Aroyo, 2011; Riesch & Potter, 2014; Sheppard, Wiggins, & Terveen, 2014; Wiggins, Newman, Stevenson, & Crowston, 2011). This concern derives from the need for quality of information inherent to scientific research. High quality information is essential for science, because the reliability of theories and their application depends on the quality of data and their interpretation. Moreover, quality is of vital importance for research organizations to keep their long-developed expertise and trusted authority.

Empirical scientific research involves different activities, such as: choosing a topic; defining the research question; reviewing existing literature; developing hypotheses; designing the study; collecting, analysing and interpreting data; drawing conclusions and communicating results (Wiggins & Crowston, 2011). In the traditional research process, the quality of research outcomes is guaranteed by the rigorous training that researchers undergo (i.e. PhD or doctoral schooling), and by the established system of peer-reviews. Given the years of dedication that researchers spend on studying specific topics, it is obvious that there is a certain *knowledge divide* between academics and the general public (Miller, 2001).

In contrast, in the case of citizen science, it is a priori not clear what kind of knowledge and skills participants have, and it is also not possible to submit voluntary citizens to demanding and lengthy trainings. Specifically, the open and voluntary nature of participation in citizen science projects, means that there is no employment contract between the research organization and citizen contributors (Simula, 2013; Wexler, 2011). Citizens are free to decide whether and when to participate, which results in *participation and retention uncertainty*. There is no guarantee about the number of people who will enrol in a citizen

science project, the time they will make available and the effort they will put into performing a task. Moreover, the openness of citizen science also means that anyone can contribute to a project, bringing highly diverse knowledge and skills. Though this diversity is one of the benefits of citizen science, it also creates *knowledge uncertainty*. In conclusion, project leaders responsible for organizing a citizen science project do not know a priori who will participate (Franzoni & Sauermann, 2014), nor the type and level of knowledge they will bring to the project. Given these uncertainties, the concern about the quality of citizen science outcomes and the efforts to ensure it seem well justified.

If the outcomes of citizen science are to be valuable for scientific organizations, it is essential that the data resulting from such projects fulfil scientific quality standards. Concerns about quality in the scientific community could offset the benefits of citizen science and its acceptance as an effective means to organize the collection and analysis of data for research purposes (European Commission, 2015; Riesch & Potter, 2014). Quality is thus seen as a challenge and a priority among professional scientists who lead citizen science projects (Riesch & Potter, 2014). However, the literature on citizen science is diverse and scattered across disciplines, and it is not clear how scientists' need to obtain high-quality information affects the organization and working practices of citizen science (Preece, 2016). Therefore, in this dissertation I seek to answer the following research question:

How is the need for quality addressed in the organization of citizen science projects?

To answer this question, I examine how project leaders and citizens deal with the need to ensure the quality of citizen science project outcomes, and how technology is used for this purpose. First, I investigate how scientific quality requirements shape the organization of citizen science projects. Second, I look at the role of citizens and how project leaders' strive for quality affects the efforts and learning process of citizens. And finally, I examine the use of technology resulting from the need to standardize and integrate citizens' contributions in order to evaluate and ensure quality.

In the following pages, I first present the research approach taken in this dissertation, the research context, the cases of my empirical study and the research methods; then I introduce the three chapters that form the core of my research.

1.2 Research approach

The purpose of this dissertation is to gain a better understanding of the phenomenon of citizen science and to explain how the need for quality is addressed in citizen science projects. The most appropriate way to examine quality in the context of citizen science is to study projects as they happen in real life (Edmondson & McManus, 2007; Yin, 2011). To this end, I have taken a case study approach that has allowed me to consider the different perspectives of professional and citizen scientists as they engage in citizen science projects (Yin, 2011).

I have carried out an extensive qualitative multiple-case study concerning quality in the relatively new and complex phenomenon of citizen science, for which little theory exists (Edmondson & McManus, 2007). Studying multiple citizen science projects means that findings are more generalizable and better substantiated than with a single-case study (Eisenhardt & Graebner, 2007). Examining multiple projects has also allowed me to zoom into specific aspects of citizen science by selecting the cases in which these aspects are most apparent (Eisenhardt, 1989). In total, I have followed and compared five projects, giving equal attention to the social and material aspects that characterize their organizational practices (Leonardi, 2012).

1.2.1 Research context

Research on the phenomenon of citizen science has thus far mainly focused on projects in the natural sciences (Wiggins & Crowston, 2011), and less is known about public participation in other scientific disciplines. Given my interest in the humanities and the growing number of citizen science projects in this area, I have chosen to focus on five citizen science projects in the fields of history, language and literature. It is important to note that, in this dissertation, the terms scientist, scholar and researcher are used interchangeably, and that the word *science*[1] is used to refer to all fields of academic research.

Focusing on the quality of citizen science outcomes in the humanities is particularly interesting, because of the different criteria and means to justify knowledge compared with the natural sciences. While in the natural sciences knowledge is inferred from empirical observations, in the humanities knowledge is commonly more descriptive and symbolic in nature (Muis, Bendixen & Haerle, 2006). This has consequences for the way that knowledge is assessed and accepted. In the natural sciences, the consistency and reliability of information is important (Muis et al., 2006), which in citizen science can be achieved by comparing multiple observations of the same research object or phenomenon. This is how the quality of citizen contributions is ensured in projects such as Galaxy Zoo[2]. In contrast, the referent or reality written in historical or literary works lies in the past or future, and cannot be verified by experiments, observations and analysis of samples (Broudy, 1977). Therefore, in the humanities, knowledge is mainly evaluated in terms of textual and

[1] The word *science* comes from the Latin word *scientia*, which means knowledge or expertise, regardless of the field. Source: http://www.etymonline.com/index.php?allowed_in_frame=0&search=science

[2] https://www.galaxyzoo.org/#/story

historical accuracy (Donald, 1990) and authenticity (Broudy, 1977), which requires assessment by experts who look at the plausibility of data and findings (Muis et al., 2006).

As well as having different epistemologies, the academic fields also differ in their relationship to and use of technology. Compared to other disciplines, such as the natural and medical sciences, the impact of technological developments in the humanities has led to a shift in the traditional way of working (Arthur, 2009; Europeana, 2014; European Commission, 2014). This shift is characterized by a move from relatively solitary towards more collaborative research; by an increased use of online databases; and by changes in communication and information sharing with the public, from physical places to virtual spaces (Arthur, 2009).

The participation of the 'crowd' or citizens in the humanities and cultural heritage covers a wide range of activities, some of which are similar to the projects done in the natural sciences (e.g. tagging of pictures, such as the project "Tag! You're it!" organized by the Brooklyn Museum[3]). However, what makes the humanities particularly interesting is the involvement of citizens in performing knowledge-intensive tasks, such as the processing and interpretation of textual data.

Humanities research is mainly based on textual data, but only an average of 23% of the collections hold by European memory institutions had been digitized up to 2015 (Nauta & van den Heuvel, 2015). Therefore an increasing number of projects in this field are aimed at supporting the wider accessibility and searchability of literary and historical manuscripts by digitally transcribing them and making them available online to other researchers and the general public. Moreover, the practice of transcribing, editing and publishing literary and historical textual materials, also called 'scholarly editing', has always been an important part of the work of humanities researchers.

An essential aspect of scholarly editing is the accurate transcription of textual (handwritten) materials, following specific rules for both transcription and annotation or commenting of the texts. Scholarly editing has traditionally been done by experts within cultural heritage organizations (archives, libraries, museums) or humanities research institutes. In the past, these experts worked mainly autonomously, with feedback from peer experts, but basically they had the knowledge and made the decisions on what to work on (manuscript), how to work (process or method), and when and how to publish (Mathijsen, 2003). Opening-up the practice of 'scholarly editing', through citizen science, means first and foremost a change in the division of labour (Puranam et al., 2014). Transcriptions that were once performed by researchers or their assistants are now carried out by citizens via the Internet.

[3] https://www.brooklynmuseum.org/community/blogosphere/2008/08/01/tag-youre-it/

1.2.2 Case selection

Through existing contacts at the KIN research group, I approached three core organizations in the field of cultural heritage and humanities research in the Netherlands, namely: the Cultural Heritage Agency of the Netherlands, the Meertens Institute and the Huygens Institute for Netherlands History. By means of familiarization interviews with managers and employees of these organizations, I learned about several Dutch and international citizen science projects supporting research in the humanities and cultural heritage.

To select projects that would help address the research question, I used theoretical sampling (Eisenhardt, 1989). I selected projects involving knowledge-intensive tasks that required considerable theoretical knowledge and intellective skills (Hislop, 2008), such as the processing and interpretation of textual and historical data. Therefore projects merely aimed at tagging or classifying images were excluded. To generalize findings and uncover issues spanning across multiple cases, I selected five projects with maximum heterogeneity (Patton, 2002), except for one common property: they all required the transcription and/or translation of handwritten manuscripts of historical value.

The projects differed in terms of: the type of manuscripts (i.e. letters, notarial deeds, inventories, chronicles); the transcription conventions applied (i.e. "diplomatic", which means transcribing the text as it is, including deletions and mistakes, versus "normalized", where the text is adjusted to make it more readable); the technology used; the way the project was organized and the characteristics of the project leaders and citizens (i.e. type and number of participants, background, experience, level of proximity, employed vs. non-employed by the organization). A short description of the projects now follows.

The project 'Letters and Correspondents around 1900' started in 2009, with the objective of transcribing the letters from and to an important 19th century Dutch writer and making them available online. These transcriptions were intended for use by other humanities researchers. The scope of the project changed slightly over the years, as the correspondence of a contemporary 19th century artist was added to the project, resulting in a total of 5.512 letters of which 1.912 were transcribed during the project. Participants in this project were both 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 project 'Digitizing Belle van Zuylen's Correspondence' started in 2010 and is still ongoing. The aim of the project is to digitize earlier paper transcriptions (published between 1979 and 1984) of the collection of private letters from and to this 18th century Dutch female writer. The participants' task involves adapting the 18th century spelling of these letters (mostly written in French) into modern-day language. Contributors to this project were, at the time of writing, members of an association interested in the work of this writer. They use e-mail as their means of communication and a web-based tool developed to integrate all contributions into one searchable online edition. They expect to make a first selection of the letters available online soon.

'Gouda on Paper' started in 2011 as a private volunteering initiative with the objective of publishing online transcriptions and translations of texts written before 1800, about or related to Gouda, which are kept in the regional archive. The first call for participants was made through local media. The project is still ongoing and more of the work is expected to become available online in 2016. Anyone can participate as long as they feel capable of performing the proposed tasks. The project has a relatively stable number of participants at about 50. They use various technologies to support their tasks: e-mail, Dropbox and a web-based tool to integrate transcriptions.

The 'Sailing Letters' project started in 2011 and took just over a year to transcribe about 5.800 scans from 17th-18th century handwritten documents. Participation was open to everyone who felt capable of carrying out this task. About 100 citizen volunteers contributed to the project. Communication took place through e-mail and an online discussion group. Individual contributions were integrated into one database through a web-based tool developed specifically for this project. The scans of the original documents and their transcriptions are available online to the general public.

'Transcribe Bentham' is a project of the University College London, which started in 2010 with the aim of transcribing the handwritten original work of Jeremy Bentham, a famous British philosopher and jurist who lived in the late 18th and early 19th centuries. Participation in this project is open to everyone. In the period of study (October 2012 to June 2014), about 3.000 people had registered; of these 400 had transcribed or partially transcribed at least one manuscript, and 11 were considered active contributors (i.e. had transcribed 100 folios or more). The project makes use of an online transcription environment (based on open source software) where all (diplomatic) contributions are posted and integrated.

1.2.3 Research methods

Focusing on citizen science projects means looking at the people who participate, the activities that take place in these projects and the technological artefacts that are used. A holistic approach of this kind is effective for theorizing about a new phenomenon like citizen science, and also for providing new insights for practitioners (Feldman & Orlikowski, 2011). By focusing on activities rather than practitioners (Nicolini, 2012), we can gain a better understanding of how the need for quality is addressed in such projects.

Data were collected in the period December 2012 to December 2015, through formal and informal interviews, observations, documents, images and quantitative data. Formal interviews followed a general interview guide and were conversational (Patton, 2002), allowing me to be flexible and adapt my questions to each person and specific conversation. This was particularly useful when interviewing volunteer citizens in their own homes. The interview guide started with an introduction, in which I asked about the background of the interviewees and how they came to be involved in the project. The interviews with professional scientists leading the projects, then included questions about setting up the project. All interviewees were asked about the activities that they and others performed in

the project, including questions about quality and the use of technology (see Appendices A and B). Interviews lasted for one hour on average, and were transcribed verbatim. I often asked interviewees, especially citizen volunteers, to demonstrate with examples how they usually performed a task. This enabled me to gain a better understanding of the task and to observe the setting in which they worked and how they interacted with the project's technological artefacts. In addition to formal recorded interviews, I also had informal conversations with project participants via telephone and Skype, of which notes were taken.

Table 1. 1. Overview of the cases

	Letters and Correspondents around 1900	**Digitizing Belle van Zuylen's Correspondence**	**Gouda on Paper**	**Sailing Letters**	**Transcribe Bentham**
Start of project	2009	2010	November 2011	November 2011	September 2010
No. of recruited (registered) citizens	20	7	60	100	3.000
No. of active* citizen participants	20	5	50	100	11
Type of documents	Letters	Letters	Books and Manuscripts	Letters	Manuscripts and letters
Scope (or size) of project**	1.912 letters	1.762 letters	1.000 pages	5.862 letters	15.634 pages
End of project	November 2016	Ongoing	Ongoing	October 2012	Ongoing

** Active citizen participants refers to people who have been engaged in the project for a longer period of time. This differs between projects.*

*** Scope refers to the number of letters or pages transcribed or translated in the project up to March 2016.*

I observed meetings, including work groups and technology training sessions. Though I mainly took an observer-as-participant role (Gold, 1958), I spent enough time in the field to develop a field relationship with some members and project leaders, allowing me to observe their activities more than once. While I did not participate in the project activities as such, I did have informal conversations with participants. During project meetings, I asked questions to clarify my understanding of what was being said, and in the training sessions, I observed how people performed their tasks, by walking around between all the participants, asking questions to get to know them and to understand the tasks and the issues they faced.

In addition to interviews and observations, I also collected several types of documents, including news articles, project newsletters, minutes of meetings, presentations, various versions of the project manuals, and an archive of e-mail messages from one of the projects. Moreover, given my focus on practices, I needed to bring technology's materiality into the analysis (Nicolini, 2012, p.4). For that purpose, I was granted access to two of the online platforms used in four of these projects. This allowed me to examine the features of technology, compare them between projects, and complement interviewees' statements and descriptions concerning these technological artefacts. I therefore retrieved images from documents and took screenshots of these platforms.

Finally, I received quantitative data from the *Transcribe Bentham* project, collected in the period from 1st October 2012 to 27th June 2014 by its project leaders. This quantitative data comprised details from 4.303 pages of submitted transcriptions that had been checked and approved by *Transcribe Bentham* staff. For each transcribed page, I used data on the number of words, level of difficulty (readability) of the handwriting, number of alterations (changes) made by the project staff, and a code to identify individual transcribers (anonymized).

Table 1. 1. Data sources

Source	Letters and Correspondents around 1900	Digitizing Belle van Zuylen's Correspondence	Gouda on Paper	Sailing Letters	Transcribe Bentham	Other*	Total
Qualitative data:							
Interviews							
Informal	3	2	4	-	2	-	11
Formal (recorded)	6	5	3	6	7	8	35
Observations							
Without interview	-	5	13	-	-	-	18
With interview (recorded)	-	1	2	-	-	-	3
Documents							
Manual (versions)	5	5	12	6	3	4	35
Website / blog (incl. project news)	1	5	4	1	2	3	16
Other documents	2	1	12	4	2	-	21
News articles (incl. open call)	-	1	4	2	1	-	8
Minutes from meetings	-	1	7	-	2	-	10
e-mail notifications	-	-	-	-	4.489	-	4.489
Technology documents							
Articles	-	-	-	-	-	1	1
Online sources	-	-	-	-	-	5	5
Images of technological artefact	5	5	9	-	2	9	30
Quantitative data:							
Submitted transcriptions	-	-	-	-	4.303	-	4.303

* **Other** refers to managers, project leaders and other professionals across the different organizations directly or indirectly related to these projects.

1.3 Dissertation outline

To understand how the need to ensure quality outcomes affects the organization and working practices of citizen science projects, I focus on the three interrelated elements that characterize this phenomenon: science, citizens and technology. Specifically, from a scientific perspective, the outcomes of citizen science need to fulfil the quality standards established by scientific organizations. These outcomes are the result of tasks performed by voluntary citizens who contribute their time, knowledge and skills to these projects. And it is through the integration of citizen contributions into web-based platforms that outcomes become open to the general public or to other scientists for further research. Hence, each of the following three chapters highlights one of these elements, while the last chapter integrates them and discusses implications for theory and practice.

In Chapter 2, I examine how project leaders' need to ensure quality shapes the working practices in citizen science projects. In particular, I look at the activities used by project leaders to ensure that project outcomes fulfil scientific quality standards. Though scholars recognize the benefits of efficiency, effectiveness, reach and engagement when citizens are involved in the research process (Brabham, 2013; Franzoni & Sauermann, 2014; Shirk et al., 2012), and admit that crowds include highly skilled individuals (Brabham, 2011), there is still a general concern in the academic community about the quality of outcomes from citizen science (Oomen & Aroyo, 2011; Riesch & Potter, 2014; Sheppard et al., 2014; Wiggins et al., 2011). To achieve the potential benefits of citizen science, project leaders need to deal with the challenge of ensuring scientific quality in a context of open participation and knowledge uncertainty. I therefore seek to understand what activities project leaders organize to ensure the quality of citizen science project outcomes.

I approach this question through a knowledge management lens, because citizen science involves knowledge work, where citizens use their knowledge and skills to contribute to the scientific research process (Cooper, Dickinson, Phillips & Bonney, 2007). While traditional research ensures quality through its own specific validation practices (Jaime, Gardoni, Mosca & Vinck, 2006), it is less clear how quality is ensured in citizen science projects. Citizen scientists are not bound to the research organisation with employment contracts, nor do they necessarily aim at advancing their academic reputation by conforming to strict quality criteria and validation processes. Given that quality is an essential indicator of performance in knowledge-intensive settings (Haas & Hansen, 2007), and that performance improvement is a key aspect of knowledge management (Pfeifer, Freudenberg & Hanel, 2001; Linderman, Schroeder, Zaheer, Liedtke & Choo, 2004), a knowledge management approach is likely to help in understanding how quality is ensured in citizen science.

In Chapter 3, I consider the role of citizens and how project leaders' strive for quality affects the efforts and learning process of citizens. Citizens are individuals who are external to research organizations but connected to them through their voluntary participation in citizen science projects. The learning process of citizens and improvement of their contributions are essential for ensuring that the outcomes of citizen science projects satisfy

scientific quality standards. Therefore, in Chapter 3, I focus on the learning process of citizen scientists and how it takes place in practice.

To answer this research question, I look at citizen science from an organizational learning perspective. I examine existing organizational learning literature and build upon the multi-level organizational learning framework proposed by Crossan, Lane & White (1999). First, I suggest the new concept of 'Extra-organizational learning', to refer to the learning process of individuals and groups that are not employed by the organization but whose tasks are essential for the delivery of a product or service. And second, I examine empirically how extra-organizational learning takes place in citizen science and what research organizations do to facilitate it. For this purpose, I zoom into one of the five studied cases where learning is most visible (Eisenhardt, 1989), namely the citizen or crowd science project *Transcribe Bentham*. To understand how extra-organizational learning takes place, I use a mix of quantitative and qualitative data. The quantitative data allow measurement and visualization of participation and quality over time. The qualitative data clarify quantitative findings and tells us about the activities supporting extra-organizational learning.

In Chapter 4, I examine the use of technology resulting from the need to standardize and integrate citizens' contributions in order to ensure quality. Web-based platforms facilitate citizen science because they make it possible to integrate contributions from distributed citizens. Integration involves the 'standardization of data definitions and structures' (Goodhue, Wybo & Kirsch, 1992, p. 294), which ensure data reliability and allows their aggregation, comparison and searchability for research purposes. The use of web-based technological artefacts can also bring efficiency gains, by hosting multiple projects that attract more citizen participants (Franzoni & Sauermann, 2014). The question is whether citizen science projects using the same web-based platform actually use it in the same way, and if that's not the case, we should understand why. Since it is generally accepted that individuals often use technology in ways other than intended, my aim is to understand why citizen science platforms are used differently than envisioned by their designers.

To understand why technology is used differently than intended, I zoom into one citizen science project that exemplifies this situation. To study this phenomenon, I use the theory of affordances (Hutchby, 2001) and propose an analytical framework based on affordances and activity theory. An affordance is a relational concept that refers to the potential action that can be accomplished by using technology (Majchrzak & Markus, 2013). I propose to understand the unintended use of technology by focusing on how affordances turn into actual actions.

The following three chapters can be read as independent papers. They were written together with co-authors, so I have kept the first-person plural 'we' to refer to all the authors.

Table 1. 2. Outline of the dissertation

Chapter	Purpose & research question	Related output	Co-authors
1. Introduction	The purpose of the Introduction is to explain the challenge of ensuring quality outcomes in citizen science, and to introduce the main research question: How is the need for quality addressed in the organization of citizen science projects?	Parts of the introduction have been used during presentations and in earlier versions of Chapter 2.	-
2. Quality in the making: Managing knowledge in citizen science projects	Chapter 2 aims to examine the activities organized by project leaders to ensure the quality of citizen science project outcomes. The question to be answered is: How do citizen science project leaders ensure the quality of project outcomes?	A previous version of this chapter was peer-reviewed and presented at: OLKC conference (2016); AOM Annual Meeting (2015); 3rd European Theory Development Workshop in OMT & Strategy (2014).	Maura Soekijad, Hans Berends & Marleen Huysman.
3. Extra-organizational learning: Learning beyond organizational boundaries	The purpose of Chapter 3 is to examine the learning process of citizen scientists and how it takes place in practice. This is addressed by focusing on the concept of 'extra-organizational learning'. The research question is: How does extra-organizational learning take place in practice?	A previous version of this chapter was peer-reviewed and presented at: OLKC conference (2015).	Hans Berends, Marleen Huysman, Maura Soekijad, Tim Causer, Melissa Terras & Kris Grint.
4. The dynamics of affordances: Using an online citizen science platform	Chapter 4 aims to investigate the technology used to ensure quality. Since technology is often used in unintended ways, this chapter addresses the following question: How does the distinction between designed, perceived and actualized affordances help to explain the use of technology in unintended ways?	-	Marleen Huysman, Hans Berends & Maura Soekijad.
5. Discussion and Conclusions	This chapter summarizes the findings, and answers the main research question.	-	-

1.4 Theoretical relevance

I seek to provide a greater understanding of the phenomenon of citizen science by tapping into theories and frameworks from the fields of knowledge management, organizational learning, and technology and organizing. Applying these theories and frameworks in this new context allows me to focus on concepts that have thus far been either overlooked or under-researched.

First, by taking a knowledge management perspective to quality assurance in citizen science, activities aimed at knowledge assessment or evaluation come to the fore. Accounts concerning knowledge evaluation are limited, and are scattered across the broad management literature (Rasmussen & Haggerty, 2008; Geiger & Schreyögg, 2009, 2012). Knowledge evaluation has been mentioned in conceptual papers (Brown & Duguid, 1998; Schreyögg & Geiger, 2007), in a few articles about knowledge creation (Giroux & Taylor, 2002; Robertson, Scarbrough & Swan, 2003) and in empirical papers about knowledge sharing or transfer (Haas & Hansen, 2007; Kane, Argote & Levine, 2005; Leinonen & Bluemink, 2008; Pérez-Nordtvedt, Kedia, Datta & Rasheed, 2008). In my research, the evaluation or assessment of knowledge emerges as an essential aspect of knowledge management and citizen science.

Second, in order to examine the role of citizens and how their learning process contributes to the quality of citizen science outcomes, I propose the new concept of 'extra-organizational learning'. Extra-organizational learning refers to the process by which individuals external to the organization learn and, by improving their performance, benefit the focal organization. Extra-organizational learning emerges as a process mirroring intra-organizational learning through knowledge exploitation (Crossan et al., 1999).

Third, I contribute to the literature of technology and organizing by focusing on affordances and how these turn into actual technology-mediated actions. Recent research calls for attention to the process and conditions of affordance actualization (Bygstad, Munkvold & Volkoff, 2016; Strong et al., 2014). I propose an analytical framework, based on affordances and activity theory, to explain affordance actualization and the use of technology in unintended ways.

1.5 Practical relevance

This dissertation is relevant for researchers interested in the organization of citizen science projects. The interest in citizen science gained momentum following the launch of GalaxyZoo in 2007. Since then, many more projects have been initiated in different scientific fields all over the world, see for example: scistarter, zooniverse, and iedereenwetenschapper. This growing interest (Bonney, Cooper & Ballard, 2016) led to the creation of the Citizen Science Association in 2012, followed by the development of the European and Australian associations, in 2013 and 2014 respectively, and the launch in 2016 of the peer-reviewed open access journal 'Citizen Science: Theory and Practice'. This journal and these associations are aimed at supporting the citizen science movement by facilitating knowledge

sharing and networking among project leaders (citizenscience.org, ecsa.citizen-science.net, csna.gaiaresources.com.au).

The research reported in this dissertation is also relevant for citizen science practitioners because it covers one of the most important challenges in this new field, namely: ensuring the quality of project outcomes (Bonney et al., 2016). My findings are particularly relevant for humanities researchers because, so far, the literature on citizen science has mainly focused on projects in the environmental and natural sciences (Bonney et al., 2016; Franzoni & Sauermann, 2014; Wiggins & Crowston, 2011). In recent years, the interest in citizen science has increased in other scientific fields, especially in the humanities and cultural heritage domains (Oomen & Aroyo, 2011). In this dissertation, I review and compare several citizen science projects in the humanities, focusing on the various activities aimed at ensuring the quality of project outcomes, the role of citizens in achieving quality, and the use of technology in these types of projects.