

Citizen Science and Memory Institutions: Opportunities and Challenges

Milena DOBREVA

University of Malta, Msida, Malta

milena.dobreva@um.edu.mt

Anna DEVRENI-KUTSUKI

South-Western University in Blagoevgrad, Bulgaria

annadevreni@gmail.com

INFORUM 2015: 21st Annual Conference on Professional Information Resources
Prague, May 26-27, 2015

Abstract

This paper presents some initial findings of the EC-funded Civic Epistemologies project which help to understand better the current attitudes and challenges memory institutions are facing in launching citizen science projects as new ways of engagement with the academics and the general public. In Section 1 it introduces the concept of citizen science; Section 2 is analysing citizen science similarities and differences with crowdsourcing; then Section 3 addresses the issues of typical tasks citizens are involved in, and Section 4 provides a brief overview of some of the user studies conducted within the Civic Epistemologies project – focus groups held in Malta, Sweden, and Spain which aimed to understand the points of views of policy makers and citizen researchers, and activists. Finally the summarises some of the possible ways forward wider use of citizen science projects within the memory institutions' context.

1. Introduction to Citizen Science

Citizen science is a contemporary reinvention of some research practices of the past when ‘unprofessional’ researchers contributed to scientific projects led by academics; a worth-noting peak of research undertaken in this paradigm had been observed in the 19th century. In the 21st century, citizen science mostly resides in digital environments and depends upon infrastructures which not only provide citizens with access to research data management, but also play the role of novel scientific communication tools aiming to engage and support citizens in their research contribution. The Green Paper on Citizen Science commissioned by the EC (2013) defines citizen science as “general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their tools and resources”. Indeed, the practice of involving citizens in research in domains such as astronomy, lexicography and biology was well established in the 19th century; the phenomenon is currently studied in depth within the AHRC-funded project ‘Constructing Scientific Communities: Citizen Science in the 19th and 21st Centuries’ (2014) based in the Universities of Oxford and Leicester in the UK; sometimes these practices are called “proto-crowdsourcing” (see Ridge 2014, p. 5 – however all historical examples provided by Mia Ridge can in fact be considered citizen science in the sense that the contribution of volunteers is coordinated and integrated into a research activity by an academic or curator). One example of a long-running study integrating citizen science is the Christmas Bird Count (n.d.) which started in 1900 and is still continuing; this effort aims to gather data on amounts and types of birds across different geographic areas and involves volunteer birdwatchers. Yet another wide-ranging effort is the creation of a dictionary of Mediaeval Latin which took 101 years to complete. This project produced seventeen lexicographic volumes the first of which was published in 1975 and the last one in 2013; however the contribution on them launched as early as in 1913 (Coughlan, 2014).

The advancement of ICT, Internet and mobile technologies opens new prospects for bringing together different communities unified by their interest to contribute to research. This resulted in a rapid growth of the citizen science initiatives around the globe, and subsequently in an increased body of academic publications discussing various aspects of it as demonstrated in (Dobрева, Azzopardi 2014). The current technological infrastructures facilitate two dimensions of citizen science: *scale* and *substance* of tasks performed. The current social media culture makes it easy to bring together big groups of people but also the modern technology offers mobile devices and a wide range of tools which could engage citizens in a variety of research-related tasks. Thus it is not coming as a surprise that the number of projects experimenting with citizen involvement across various sciences constantly grows. The most typical scenario is the one of citizens directed by professional researchers in studies which revolves mostly around observation of natural phenomena and notation in multiple locations or across longer time spans.

The interest to such projects grew to the extent that specialised platforms which allow to define research tasks and involve users had been created; e.g. Zooniverse (Smith et al., 2013), Curio (Law et al., 2013), and CrowdCrafting (2013) developed in a collaboration between the Citizen Cyberscience Centre and the Open Knowledge Foundation (OKF). These platforms are used for research in different domains, but mostly in the Sciences with few implementations in the Humanities.

2. Citizen Science and Crowdsourcing

Citizen science is often used as a synonym for *crowdsourcing*. Indeed, there are significant similarities in both domains, including the participation of the citizens and the technological infrastructures used. The use of the term “Citizen science” however is justified when the effort involving citizens is aiming a research project guided by an academic. When we consider the digital cultural heritage domain, crowdsourcing is still more popular and as a form which can be used to understand better the patterns of engagement, tasks for volunteers and benefits, our work on requirements included crowdsourcing as a more familiar concept.

Although the application of citizen Science in the field of Humanities has been less common than in the sciences, there are a number of examples of crowdsourcing projects which recently had been presented in an edited collection (Ridge 2014) and in overview of activities in the British Library (Ellis, 2014). In this section we will be looking at major characteristics of crowdsourcing in the memory institutions taking them as an inspiration for citizen science initiatives.

Recently, Noordegraaf et al (2014) suggested a model for crowdsourcing in the memory institutions context which explores six pillars, namely: institution, collection, goal, crowd, infrastructure, and evaluation. The rationale is that considering crowdsourcing should start with the major institutional dimensions (listed under *Institution*), then be narrowed down looking at the *Collection* pillar, and so on. This model does not elaborate in depth on the various stakeholders which should be in charge for various [components of] the pillars under consideration.

It is worth noticing the components of the Infrastructure pillar which include complexity of the task, evolution of the task, level of scaffolding (identified as “limitation in variability in response (through the implementation of pull-down menus rather than open-text fields, for instance)”, and generic platform or devoted project site). The authors summarise:

“Therefore, the infrastructure of a project must be designed with concern for the variables [...] the complexity of the task being asked of the crowd, whether it can be broken down further into components, whether or not the user interface should be scaffolded to encourage members of the crowd, and whether a generic platform should be used to host the project are all questions that arise in the design of the project infrastructure.”

(Noordegraaf et al. 2014)

However, one should bear in mind that the specific solutions are still quite volatile and there are more questions than answers particularly in the memory institutions domain:

“How can our tools act as scaffolds to help make the most of user efforts? What expertise can we embed inside the design of our tools to magnify user efforts? How can our tools put a potential user in exactly the right position, with the right knowledge, just at the moment he or she needs it, to accomplish a given activity?”

(Owens 2013)

We would like to conclude this section with an observation that emphasizes on and deepens the insight into knowing the citizens who would be potentially able and willing to contribute to citizen science projects.

“This is one of the places where libraries, archives, and museums have the most to offer. As stewards of cultural memory, our institutions have a strong sense of purpose and their explicit mission is to serve the public good. This notion of motivation prompts further key questions for projects: Whose sense of purpose does this project connect to? What identities are involved? What kinds of people does this project matter to? And how can we connect with and invite the participation of those people?”

(Owens 2013).

3. Citizens and their Contribution: Science vs Humanities

While we were not able to identify larger-scale surveys on citizen science applications in cultural heritage institutions, such work was systematically done in the last years by Angela Wiggins and Steven Crowston from the Syracuse University in the USA. Wiggins and Crowston (2012a and 2012b) summarise results from 63 surveys completed as a result of 840 emailed requests for participation which were used to create 128 project profiles.

The range of activities to which unprofessional researchers contribute in citizen-science projects as suggested by Wiggins and Crowston (2012b) include the following:

1. Define question
2. Gather information
3. Develop hypothesis
4. Design study
5. Data collection

6. Analyse sample
7. Analyse data
8. Interpret data
9. Draw conclusions
10. Disseminate results
11. Discuss results and ask new questions

Those activities assume different levels of creativity. The tasks of transcribing historical letters or providing geolocations would normally be considered to be quite trivial and are from the contributive type of citizen involvement in the memory institutions domain as defined in (Bonne, 2009). Thus one research question for the future is how citizens involved in Humanities research could contribute to creative rather than trivial tasks?

In their further study (Wiggins and Crowston 2015) revise the granularity of their previous classification of activities and arrive to the following structure of activities:

“The main research activities open to participation in the responding projects were observation, data entry, and species identification. This reflects the fact that most of the responding projects focused on data collection, frequently for observational data. The next most common tasks were measurement, site selection and/or description, and photography. These tasks are specific to certain types of field-based participation that can also include observation.

Additional activities reported by respondents were diverse, primarily scientific tasks related to specific project requirements, and occasionally tasks related to stewardship and communication. These participant activities aligned with some of the primary goal areas discussed earlier.

- *Scientific tasks*
 - *Posing new questions, literature reviews, paper writing, etc.*
 - *Videography*
 - *Monitoring*
 - *Insect rearing*
 - *Identifying animal tracks*
 - *Creating maps*
 - *Stewardship*
 - *Organization and landowner coordination*
 - *Manual labor, habitat construction, shell recycling*
 - *Communication*
 - *Communication with other participants and with scientists*
 - *Sharing observations and findings at meetings of related groups.”*
- (Higgins and Crowston, 2015)

It is very informative to explore the evolution of the concept of activities and tasks in the works of Wiggins and Crowston as well as to compare their list of activities with evidence from the memory institutions and Humanities domains.

A popular classification of typologies of crowdsourcing projects in the memory institutions domain had been made by Oomen and Aroyo (2011). They suggest six different typologies of projects; each one linked to a different kind of study, and respectively tasks:

1. Correction and transcription – the citizen is given access to a database (this is usually a text-based database like scanned manuscripts) and then he gets the task of transcribing or making corrections to the text which was already transcribed electronically via a computer programme.
2. Contextualization – Citizens submit data such as letters, stories, films, photographs or other documentary material in order to gather a meaningful context.
3. Complementing Collection – Citizens are asked to add data into databases with the ultimate aim of completing them or making the collection grow.

4. Classification – Citizens tag the data, or label it, in order to easily group similar data and make the information more easily retrievable in the future.
5. Co-curation – This practise occurs mostly with projects involving the aesthetic arts. Citizens interact with institutions and voice their opinions when it comes to choosing articles or items for publication.
6. Crowdfunding – Citizens are asked to gather together money and/or resources in order to support efforts initiated by others.

A different approach is proposed by Tobias Blanke and Mark Hedges (2013) within the context of Humanities e-Science; while their paper is not focused on citizen science it identifies some typical scenarios and illustrates how all of them are integrating a number of scholarly primitives, namely collecting, discovering, comparing, delivering, and collaborating. While it would require an additional study to justify the use of the same or different set of primitives in citizen science, this is an approach which introduces different levels of granularity with the primitives as the smallest building blocks of more complex activities.

4. Evidence from the User Studies within Civic Epistemologies project

In order to gather evidence on opinions of citizen scientists as well as major stakeholders within the citizen science domain the Civic Epistemologies project funded in 2014-2015 by the EC organised focus groups covering different geographic regions and bringing together these three groups of stakeholders (policy makes, citizen activists, and citizen scientists). The project also gathered data via an online survey.

As one major outcome from this project – and still unique for the citizen science domain, we would like to mention that the evidence gathered during focus groups was used to develop three personae, of Theresa (citizen scientist), Mark (policy maker) and Stina (activist) which are discussed in more detail in Deliverable 2.2 of the project. The method of developing personae is a summative way to describe key characteristics of a typical user; the development of personae requires substantial evidence and the three personae presented below as examples of typical stakeholders based on the feedback provided during the three focus groups are initial attempts to capture the most evident characteristics of the representatives of the groups.

Here we are presenting as an example the description of Mark.

| | |
|---|--|
| <p style="text-align: center;"><i>The policy maker,</i> Mark</p> <hr style="width: 20%; margin: auto;"/>  | <p><i>Mark is a 40-year old CH professional from Malta with a role in defining the policies of his institution. He regularly uses CH collections not only for professional reasons but also because he has strong personal interest in the area. Mark is not quite sure how to use the digital collections of his institution for artistic purposes.</i></p> <p><i>He is not that familiar with citizen science and has not played an active role in such projects but could be interested to try it in the future. Mark sees a range of benefits from using citizen science – mostly related to an improved relation and services offered to the general public but also to the visibility of his institution.</i></p> <p><i>Mark is convinced that the main benefit from citizen science is not cutting any costs but better engagement with the general public.</i></p> |
|---|--|

5. Conclusions

This paper explored the domain of citizen science comparing its application within the domains of Sciences and the Humanities - the latter being of great relevance to memory institutions. The paper also looked at the similarities and differences of crowdsourcing and citizen science, arguing that the successful crowdsourcing implementation within the cultural heritage sector can be particularly useful to find successful approaches for introducing citizen science projects.

What factors would facilitate the introduction of citizen science initiatives in the memory institutions? The first one is definitely wider awareness on the **ways to organise citizen science projects**, on the **existing tools and infrastructures** and on **strategies to attract groups of contributors**.

Another essential consideration for successful introduction of citizen science is the **granularity of tasks** entrusted to the citizens; further essential characteristic of the tasks in the memory institution domain is to what extent they are **engaging for the participants**:

“It isn’t about Sisyphean tasks; it is about providing meaningful ways for the public to enhance collections while more deeply engaging and exploring them.” (Owens 2013)

Possibly tasks which appear either too overwhelming or boring would disengage participants very quickly but it is a complex question what are the individual interests, knowledge and perseverance of individual contributors. One interesting recent example of an individual citizen science project was the work done by a single volunteer from Australia who painstakingly entered the data from over 50.000 digitized passport applications from the National Archives of Malta thus creating a database which is now used for various historical and sociological research tasks, in particular in Maltese diaspora studies (Caruana 2015). May be this could be taken as a rare or extreme case of long-term individual engagement, which constitutes a huge contrast with the notoriously difficult attention capture of the “google generation”. However citizen science should offer ways for engagement of volunteers of different personal profiles and for the time being the mechanisms for assisting longer engagement are mostly tried but need further research.

Further work on the use of digital libraries for citizen research will be done within the framework of the ACoMin (“Advanced Computing for Innovation”) project.

While this paper addressed a number of current issues related to the use of citizen science in the memory institutions, this is a domain which still requires deeper insights and more support for the cultural heritage sector. The work of the Civic Epistemologies project which seeks to develop a roadmap for citizen science applications within the digital cultural heritage institutions will help to identify trends as well as ways forward in this interesting area.

Acknowledgements. The authors acknowledge the support of the EC-funded projects Civic Epistemologies (*Development of a Roadmap for Citizen Researchers in the age of Digital Culture*) funded under FP7 grant agreement 632694 and ACoMin “Advanced Computing for Innovation”, Grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions). The authors would like to thank in particular their colleagues from the

References

- Blanke, T., M. Hedges (2013). Scholarly primitives: Building Institutional Infrastructure for Humanities e-Science. In: Future Generations Computer Systems 29, 654-661.
- Bonne, R., et al. (2009) Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education. A CAISE Inquiry Group Report. Online: <http://files.eric.ed.gov/fulltext/ED519688.pdf>
- Caruana, M. (2015) Analysis of data from Maltese Passport Applications held at the National Archives of Malta: A new digital resource. In: Proceedings of the International conference The Commonwealth and its People: Diasporas, Identities, Memories, 24-26 June 2015, Valletta (forthcoming).
- Christmas bird count (n.d.). Online: <http://birds.audubon.org/christmas-bird-count>
- Constructing Scientific Communities: Citizen Science in the 19th and 21st Centuries (2014). Online: <http://conscicom.org/>

- Coughlan, S.: Dictionary reaches final definition after century (2014) Online: <http://www.bbc.co.uk/news/education-28952646>
- CrowdCrafting. (2013) Online: <http://blog.okfn.org/2013/09/17/crowdcrafting-putting-citizens-in-control-of-citizen-science>.
- Dobrev, M., and D. Azzopardi (2014) Citizen Science in the Humanities: A Promise for Creativity. In: G. Papadopoulos (ed.) Proceedings of the 9th International Conference on Knowledge, Information and Creativity Support Systems, Limassol, Cyprus, November 6-8, 2014, ISBN: 978-9963-700-84-4, pp. 446-451.
- Dobrev, M., and D. Azzopardi. (2015) D2.1 Results of the workshop on requirements. Civic Epistemologies project deliverable. http://www.civic-epistemologies.eu/wp-content/uploads/2014/08/CIVIC-EPISTEMOLOGIES_D2.1-Results-of-the-workshop-on-requirements_V1.0-FINAL.pdf
- Ellis, S. (2014) A History of Collaboration, a Future in Crowdsourcing: Positive Impacts of Cooperation on British Librarianship. In: Libri, 64(1), 1-10.
- Green Paper on Citizen Science (2013). European Commission. <http://www.societize.eu/sites/default/files/Green%20Paper%20on%20Citizen%20Science%202013.pdf>
- Law, E., Dalton, C., Merrill, N., Young, A., Gajos, K.: Curio (2013) A Platform for Supporting Mixed-Expertise Crowdsourcing. Human Computation and Crowdsourcing: Works in Progress and Demonstration Abstracts. AAAI Technical Report CR-13-01. An Adjunct to the Proceedings of the First AAAI Conference on Human Computation and Crowdsourcing. 99-100.
- Noordegraaf, J., Bartholomew A., Eveleigh, A. (2014) Modeling Crowdsourcing for Cultural Heritage. Museums and the Web, In: N. Proctor & R. Cherry (eds). Museums and the Web. (2014) Online: <http://mw2014.museumsandtheweb.com/paper/modeling-crowdsourcing-for-cultural-heritage/>
- Oomen, J., L. Aroyo (2011) Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges. Proceedings of the 5th International Conference on Communities and Technologies (C&T'11). ACM, New York, NY, USA, 138-149.
- Owens, T. (2013) Digital Cultural Heritage and the Crowd. Curator: The Museum Journal. Vol 56. Iss: 1.
- Ridge, M. (2014) Crowdsourcing Our Cultural Heritage: Introduction. pp.1-13. Ashgate.
- Smith A.M., Lynn, S., Lintott, C.J. (2013) An Introduction to the Zooniverse. Crowdsourcing: Works in Progress and Demonstration Abstracts. AAAI Technical Report CR-13-01.
- Wiggins, A, K. Crowston (2012a) Goals and tasks: Two typologies of citizen science projects. System Science (HICSS), 2012 45th Hawaii International Conference on, 3426-3435.
- Wiggins, A., K. Crowston. (2012b). Describing Public Participation in Scientific Research, iConference 2012 Toronto, Ontario, Canada. Available: <http://crowston.syr.edu/system/files/iConference2012.pdf>
- Wiggins, A., K. Crowston. (2015). Surveying the citizen science landscape. First Monday, Volume 20, Number 1 - 5 January 2015, <http://firstmonday.org/ojs/index.php/fm/article/view/5520/4194>