

Managing Uncertainties: Small-scale Crowdsourcing of Author Letters

Jenny Bergenmar [0000-0002-6761-3544]¹

¹ University of Gothenburg, Box 200, 405 30 Gothenburg, Sweden
firstname.lastname@lir.gu.se

Abstract

The increased digital accessibility of cultural heritage materials has brought to the fore the need to make these resources searchable and useful, for example by engaging the public in transcription and data tagging. However, engaging the public in the collection of cultural heritage material is more uncommon. This paper presents the project *Remembering Selma Lagerlöf*, a crowdsourcing initiative to collect the letters Lagerlöf sent to her audience and publish them in a digital exhibition, along with letters from the public to Lagerlöf. The methodological challenges connected with crowdsourcing projects that involve the collection of cultural heritage material are discussed. This paper particularly focuses on issues concerning public engagement and the relation between researchers and contributors, as well as on issues concerning the material, such as contextualization and ethical considerations. The letters form a microhistory of the individuals who, for various reasons, contacted Lagerlöf; the contributors submitting the letters have also proven to be an important source of factual information about the context of the letters. The letters to Lagerlöf from the public can be described as a passively stored memory, which is included in the archive of memory, but not in the canon of public knowledge. However, in many cases, these letters may be actively remembered within the families they concern, who preserve the personal contexts of the letters and consider them relevant.

Keywords: Citizen humanities, Letters, Selma Lagerlöf, crowdsourcing, cultural heritage

1 Introduction

The project *Remembering Selma Lagerlöf* began after the completion of an earlier project that examined the public's letters to Swedish author and Nobel Prize laureate Selma Lagerlöf during her literary career from 1890 to 1940. Thousands of letters were written to Selma Lagerlöf by her audience, commenting on their reading of her works as well as on countless other subjects; the letters thus reflect both historical and societal change

during fifty years, and writing and reading practices of that time. In some cases, the letters also reveal that Lagerlöf replied to them. However, in contrast to the letters to Lagerlöf from the public, which are kept by the National Library of Sweden, her replies have not been collected or preserved in any archive. Rather, these letters are often privately owned by the families that originally received them. In other cases, they have been lost and then found by chance, or have circulated through the commercial antiquarian market. In order to discover how Lagerlöf answered the people who wrote to her, and what characterized her interactions with the public, the aim of the project *Remembering Selma Lagerlöf* is to engage the public in collecting these letters and to make them accessible in a digital exhibition.

The increased digital accessibility of cultural heritage materials has brought forward the need to make these resources searchable and useful, for example through document transcription and data tagging. Citizen science has long engaged volunteers in the sciences, but the number of projects inviting non-experts to contribute to the use and exploration of cultural heritage materials is also increasing [5]. As Jennings, Dobрева and Devreni-Koutsouki have pointed out, there is “a range of terms applied for similar activities but appealing to different audiences: citizen humanities, citizen history, citizen archiving” [4]. Crowdsourcing has become a widely used term, which Terras has described as “the practice of using contributions from a large online community to undertake a specific task, create content, or gather ideas” [11]. However, the term is somewhat inapt in this case, since the project in question has relatively few participants. It may be described as a small-scale crowdsourcing project involving basic digitization. Unlike many other citizen humanities initiatives, this project did not start from a collection previously digitized by a cultural heritage institution. Instead, its purpose was to collect material that was unknown and dispersed across many private owners. In the classification of crowdsourcing activities by Oomen and Aroyo, which comprises correction and transcription, contextualization, complementing of collections, classifications, co-curation, and crowdfunding, this project can best be classified under complementing of collections, with elements of contextualization [8]. The aim of this paper is to discuss the methodological challenges connected with crowdsourcing involving the collection of cultural heritage material, with a specific focus on issues concerning public engagement and the relation between researchers and contributors, and on issues concerning the material itself, such as contextualization and ethical considerations.

2 Engaging the Public

It is a challenge for all citizen humanities projects to find ways to promote user engagement. One shortcut to make a project known to volunteers is to use an infrastructure such as *Zooniverse*, which gives project leaders the opportunity to create and list a project page on its site. Participants can then choose to become involved in different projects, and the projects benefit from an already existing community of potential contributors. However, for projects using less common languages, *Zooniverse* may not be as useful, since the majority of the registered cultural heritage projects involve English material. Independent web platforms such as *Transcribe Bentham* at University College

London (UCL) and – of particular interest in this case – *Letter of 1916* at Maynooth University are also available.¹ However, both these projects started from digitized collections at cultural heritage institutions, which was not the case with *Remembering Selma Lagerlöf*. Since the latter project is independent of cultural heritage institutions, it would have been ideal to integrate the project webpage with the digital infrastructure of the University of Gothenburg, in order to make the academic context clear. However, this was not possible due to the university's IT policy. Instead, a simple project webpage was created on WordPress (minnenavselma.org). This webpage serves three different purposes: (1) to provide an interface for submitting material; (2) to provide an interface for displaying and contextualizing this material; and (3) to provide an interface for displaying letters to Lagerlöf from the public (i.e. the results from a previous project investigating Lagerlöf's letters held by the National Library), in addition to those she wrote in return (i.e. the current project). Social media in the form of a Facebook page and Facebook campaigns was used to communicate the project and disseminate the website link. The Swedish Literature Bank provided a link to the webpage on its well-established and active Facebook page, and the University of Gothenburg advertised the webpage on its departmental and faculty webpages.

The project budget also included costs for traditional newspaper advertising in order to make the project more widely known. However, when we received funding at the end of 2016, the story of the project was picked up by newspapers that ran articles on it, leading to several interviews on local radio. This news coverage proved effective, as it immediately resulted in contact with people owning letters. Later, a press release was issued by the University of Gothenburg, which led to more press and radio coverage. Contact was also made with organizations that were likely to be interested in material of this kind, including the Swedish Association for Genealogists (*Sveriges släktforskarförbund*) and local history societies. Projects like this, which complement an existing collection (in this case, the Selma Lagerlöf collection at the National Library of Sweden), stand or fall on the possibility of reaching a specific group of volunteers (here, the owners of letters from Lagerlöf) and on their willingness to share the letters with a larger audience. In this, the project is unlike many other citizen humanities projects, which involve tasks such as working with transcription and annotation. Thus, contributors cannot be recruited through workshops in schools, for example, or other face-to-face activities. Most of the people who have contributed material (i.e., the owners of letters from Selma Lagerlöf) are related to the addressee, and the letters have been preserved within their families. On the other hand, this project holds an advantage: the problem of engaging and training volunteers over time in order to obtain good results, which is common in transcription projects, is not an issue here. In this project, one-time participation is sufficient.

¹ <https://www.zooniverse.org>, <http://blogs.ucl.ac.uk/transcribe-bentham/> and <http://letters1916.maynoothuniversity.ie>

3 Interaction with Contributors

An important feature of the project webpage is the web form, which was created for the submission of material. The web form includes simple instruction on how to photograph and upload the letters, and how to provide basic information about the letter such as the name of the addressee, the year, the place and a short description of the letter or addressee. An email address is also required, along with confirmation that the contributor permits the letter to be made public on the webpage. The web form was intended to be both a practical solution for receiving contributions and a form of “scaffolding” [9], by clearly showing participants what information about the material is required. However, only about two thirds of the contributors used the web form to submit material. Instead, some contributors emailed the material directly to the project members, while others sent paper copies by mail. With a project such as this one, in which many of the owners of the material are older, use of a digital medium can be an obstacle. Nevertheless, seniors have been surprisingly active on Facebook in this project, both in sending messages to the project and in reacting to new material on the webpage, which is advertised on Facebook. Email and mail also support a more personal contact than a web form, in which only an automatic confirmation of the submission is issued. Since the majority of the contributions are letters that have been inherited within families, there is often a family history to be told, which is connected to a letter to Lagerlöf written by some member of the family. These are in most cases kept in the National Library. Thus, one of the central motivations for the contributors is to make their family’s history known, and one of the primary motivations of the project members is the same: to know more about the context of the persons who wrote to Lagerlöf and about her answers to them. While the main motivations of the participants seem to involve reciprocity – or, more exactly, an acknowledgement and confirmation of the personal histories the letters contain – the main motivation for this project, which takes an institutional point of view, is similar to that of many citizen humanities projects in that it is factual (i.e., aims to gain information on specific historical material) and contains an opportunity to share and communicate knowledge [3].

Of the two models for organizing participation and production on the Internet that Haythornewaithe has identified – “[a] crowdsourcing model based on micro-participation from many, unconnected individuals, and a virtual community model, based on strong connections among a committed set of connected members” [2] – *Remembering Selma Lagerlöf* relies on micro-participation from unconnected individuals, although not very many (47 individual contributors to date). To some extent, the project draws on communities that have already formed around interests such as genealogy and local history, but these communities are not virtual ones. Furthermore, this project does not function as a space for peer production, since there is usually no contact between individual contributors. Since the contributors are individual and no learning process is required to participate, the project employs what Haythornewaithe calls “light-weight peer production”; that is, no long-term commitment is needed to the task or to any community [2]. However, personal contact with the individual contributors has proven to be a more “heavy-weight” production. This contact fulfils different needs, including: (1) confirmation to publish letters to Lagerlöf from deceased relatives, which involves

contacting contributors to obtain their agreement to the publication of the original letter answered by Lagerlöf, so as to display the whole correspondence; (2) the need, in many cases, for additional information, since the letters submitted by email or mail do not always include the information required by the web form; and (3) contact with all contributors once the material they submitted is published on the webpage, which involves giving them a chance to voice their opinion. The latter is especially important when the letters concern family members.

Another aspect of the interaction between experts and contributors in this project is that the material that has been published often gives rise to questions about the context. These questions usually appeared on Facebook, either as comments to the posts or in messages. Like Nourse, Insole and Warren, we found that “contributors are motivated by the personal knowledge they gain by taking part” [7]. This has been true both for people who did not submit material but followed the project on Facebook and asked questions, and for contributors, who often want to know what their relatives wrote to Lagerlöf.

4 Unknown Letters, Unknown Contexts

The letters to Selma Lagerlöf from the public can be divided into various categories. Important categories include: letters asking for economic assistance, letters from teachers and schoolchildren, letters about the reading of her works, and letters that reflect Lagerlöf’s position as a celebrity. Needless to say, there are also letters from Lagerlöf’s friends, family and publishers, along with various letters about translation and other business. These are the type of letters to an author that literary scholars usually focus on, since they can yield important clues about the author’s life and literary production. However, in this project, our main interest was in Lagerlöf’s answers to the ordinary people who wrote to her. These letters remain a forgotten heritage in more than one sense: the letters to Lagerlöf from the public are practically unexplored by research, and her answers are completely unknown. Even though we have been able to match the original letters with Lagerlöf’s answers in many cases, representative answers from Lagerlöf to letters in all the important categories have not been submitted, while some categories are over-represented. None of Lagerlöf’s answers to letters requesting economic support have been submitted, although it is clear both from comments in letters and from subsequent letters thanking her for money that she answered many of them. Perhaps these letters were preserved to a lesser extent due to the poor life conditions of the receiver. Or, they may have been purposely destroyed by supplicants or their relatives who found it shameful to have asked for support (and who were perhaps rejected). It is also possible that these letters have been preserved but not submitted by the owners, who are often relatives of the original writer.

A disproportionate number of answers to letters from senders who aspired to be authors were submitted. Here, the situation is reversed: the letters are often confidently and fluently written, and even when the answers from Lagerlöf reject their requests for comments or recommendations, her answers can still be seen as an acknowledgement,

since they often include some words of encouragement. In some cases, these aspiring authors also ended up becoming published, with or without Lagerlöf's help.

The project also received interesting contributions in which Lagerlöf answers letters from teachers and schoolchildren, questions about her literary works (in particular, their historical background), invitations and appeals to engage her in different kinds of social work, and requests to publish a text in various publications (often for charitable causes). In a strict sense, other contributions fall outside of the scope of this project; rather than consisting of her answers to the public, these letters were written by Lagerlöf to various addressees for business purposes, were related to her estate Mårbacka, or were addressed to her publisher or to friends, colleagues and acquaintances.

The contributions are interesting documents in themselves, in many cases, and provide answers to the question of how Lagerlöf dealt with letters from unknown persons. It is clear that Lagerlöf considered answering to be a part of her profession and that she put effort into her answers. Furthermore, the contributions provide information on who the letter writers were. In many cases, the letters to Lagerlöf from the public contain a micro-autobiography that was used rhetorically to frame the question or reason for contacting the author. Even so, new contexts emerge when additional information about the situation of the letter writer is provided. These letters shift the focus from Lagerlöf and her role as an author, to more quotidian realities and broader historical questions. The letters about the estate give insight into the families working there, and aspects of farm life and business. The letters from aspiring authors tell the tale of a Sweden before the welfare state, where education and access to the written word was a luxury that was not for everyone. The histories of authors who are forgotten today emerge, and it is sometimes clear that Lagerlöf's support was what helped them to gain access to a literary life.

With Assman's distinction between active and passive memory, the contexts and stories appearing in the letters to Lagerlöf and her answers can be described as parts of a "passively stored memory" [1]. According to Assman, it is "the actively circulated memory that keeps the past present as the canon and the passively stored memory that preserves the past as the archive" [1]. The letters of known authors are often included in "the canon" in Assman's sense, and are "kept present", for example, in new editions of letters or as a source for biographies on the author. However, the letters of ordinary people such as those who wrote to Lagerlöf are only passively remembered and are placed in "the archive" of memory. Lagerlöf's answers to these letters might be said to inhabit a place outside of the archive, in terms of physical space at least, since they are privately owned privately by numerous separate owners. Furthermore, in many cases, these letters may be actively remembered within the families they concern, in which the personal contexts of the letters are preserved and considered relevant.

5 Managing Memories, Managing Uncertainties

Citizen humanities projects such as this – that is, small in scale, and operating outside of cultural heritage institutions – face a number of uncertainties concerning the material

produced within the project. First, there are questions about the ownership of the material. As Jennings, Dobрева and Devreni-Koutsouki have remarked, the “ownership of materials collected within citizen science projects, authorship and copyright are all issues, which would clearly benefit from further clarification” [4]. The letters by Lagerlöf collected in this project were published with the agreement of the owners. Lagerlöf’s texts are also free of copyright, since she died in 1940. However, the conditions for the original letters to Lagerlöf (that she answered) are more uncertain. When Lagerlöf’s answers are submitted by relatives of the addressee, we ask permission to publish the matching letters to Lagerlöf from their relatives. In all cases, the contributors have agreed. However, there are also cases in which letters have been submitted by contributors with no relation to the addressee. In these cases, we still published the original letters to Lagerlöf, except when the content was sensitive in some way.

This situation leads to an even more pressing issue: the ethical considerations that arise from the material collected. Lagerlöf’s own correspondence reveals that she was conscious that her letters would be read by scholars in the future. However, the unknown people who wrote to her probably did not expect their letters to be preserved, and sometimes expressly asked for confidentiality. In many cases, they were vulnerable: poor, ill and at risk of losing their homes or even their children. As mentioned before, no answers from Lagerlöf to people asking for economic assistance have been submitted. However, one answer from Lagerlöf to a woman who was mentally ill was submitted; in this case, Lagerlöf’s addressee was the first patient admitted to the Helsingborg asylum in Skåne, in 1927. The letter to Lagerlöf is somewhat incoherent and has no evident purpose aside from narrating the woman’s thoughts and events from her everyday life. The letter was sent to Lagerlöf with an accompanying note from the doctor in charge, saying that the sender hoped greatly for an answer. Lagerlöf sent a short answer that shows that she took care to respond to the content of the letter. This exchange of letters gives insight not only into the history of mental care in Sweden, but also into the voice of the patient herself, which has rarely been the case in the history of psychiatry. In dialogue with the contributors of the letter, who were not relatives, we decided to publish the letter from the patient together with Lagerlöf’s answer. The letter had been attached to the patient’s journal; it was discovered by one of the contributors, who was employed at the asylum and had done some archival research. This letter is an interesting example, since letters from people with mental illnesses form a distinct category in the letters from the public to Lagerlöf.

As with many other small-scale digital projects, the uncertainty of long-term preservation is an important issue. After the project’s end, the Centre for Digital Humanities at the University of Gothenburg will be responsible for maintaining the webpage, although to what extent it will be updated is uncertain. The National Library expressed an interest in keeping the image files of the letters, and they have been transferred to this institution for long-term preservation.

Finally, not knowing if or to what extent letters from Lagerlöf will be submitted is a type of uncertainty that is especially common in citizen humanities projects of this kind. In any project depending on volunteers, it is impossible to predict the outcome. However, in projects that involve collecting material with the help of volunteers, participation is linked to access to material, which limits the number of possible participants.

We knew that there were answers from Lagerlöf, but not whether their owners would be willing to contribute. Likewise, the quality of the submitted material was impossible to foresee. Would the short instruction provided on the web form suffice to collect the information we sought? For display on the webpage, the quality of the images of the letters proved to be sufficient. Most of the participants have submitted digital files, and the paper copies that were mailed have been digitized by the project members.

6 Conclusion

Although the name of the project is *Remembering Selma Lagerlöf*, this project is actually more about memories of the people who contacted her and who received answers from her. The primary intention behind this project was to shift the focus from the traditional way of studying an author's letters – that is, to view the letters as clues to information about the author's literary production or biography – to a microhistorical approach that would make space for voices that are seldom heard [6]. In this material, these voices are in dialogue with one of the greatest celebrities in Sweden during that period. The material reveals unknown histories in two different ways: by revealing unknown sides of the author; and by illustrating the hopes, realities and anxieties of the people who wrote to her.

Remembering Selma Lagerlöf is a digital project that requires minimal digitization of the source material, as well as a digital interface for both presentation and communication. However, in contrast to the large-scale methods that have attracted much attention in the digital humanities, this project exists on a very small scale: it engages not with a crowd but with individual contributors, and presents material that narrates the stories of individuals. Thus, this project can be described as a microhistory dealing both with a collection of letters already preserved at an institution (the letters to Lagerlöf from the public) and crowdsourced letters (her answers to them) [6]. As always with microhistorical approaches, the question is how to avoid anecdotal history writing. Paul Ricoeur's response to this question in *Memory, History, Forgetting* is that "it is the latent and dispersed properties of an available historical language – which the computer misses – that the historian brings to light and organizes into discourse" [10]. This line of reasoning certainly seems relevant for the digital humanities today, in terms of counter-balancing the exploration of history on the macroscale.

Acknowledgements

The project *Remembering Selma Lagerlöf* was funded by the Swedish Foundation for Humanities and Social Sciences, Grant No. KOM16-1302:1.

References

1. Assman, A.: Canon and archive. In: Erll, A., Nünning, A., Young, S. (eds.), *A Companion to Cultural Memory Studies. An International and Interdisciplinary Handbook*, pp. 97-108. De Gruyter, Berlin and New York (2008).

2. Haythornewaithe, C.: Crowds and communities. Light and heavyweight models of peer production, pp. 1-10. In: Proceedings of the 42nd Hawaii International Conference on System Sciences, January 5-8 (2009). DOI:[10.1109/HICSS.2009.650](https://doi.org/10.1109/HICSS.2009.650).
3. Huvila, I.: The unbearable lightness of participating. Revisiting the discourses of “participation” in archival literature. *Journal of Documentation* 71(2), pp. 358-386 (2015).
4. Jennings, E., Dobрева, M., Devreni-Koutsouki, A.: Towards user engagement models for citizen science. Initiatives in the digital cultural heritage domain. In: Ciolfi, L., Damala, A., Hornecker, E. Lechner, M., Maye, L. (eds.), *Cultural Heritage Communities. Technologies and Challenges*, pp. 78-95. Routledge, Abingdon & New York (2018).
5. Kasperowski, D., Kullenberg, C., Rohden, F.: The participatory epistemic cultures of citizen humanities. In: Pierroux, P., Hetland, P. (eds.), *A History of Participation in Museums and Archives. Traversing Citizen Science and Citizen Humanities*. Routledge, forthcoming.
6. Magnússon, S., Szijártó, I.M.: *What Is Microhistory? Theory and Practice*. Routledge, Abingdon & New York (2013).
7. Nourse, N., Insole, P., Warren, J.: Having a lovely time: Localized crowdsourcing to create a 1930s street view of Bristol from a digitized postcard collection. In: Roued-Cunliffe, H., Copeland, A. (eds.), *Participatory Heritage*, pp. 153-162. Facet Publishing, London (2017).
8. Oomen, J., Aroyo L.: Crowdsourcing in the cultural heritage domain: Opportunities and challenges. In: Proceedings of the 5th International Conference on Communities and Technologies, pp. 138-149. Association of Computing Machinery, New York, June 29-July 2 (2011).
9. Owens, T.: Digital cultural heritage and the crowd. *Curator: The Museum Journal* 56(1), pp. 125-126 (2013).
10. Ricoeur, P.: *Memory, History, Forgetting*. Blamey, K. and Pellauer, D. (transl.), University of Chicago Press, Chicago and London (2004).
11. Terras, M.: Crowdsourcing in the digital humanities. In: Schreibman, S., Siemens, R., Unsworth, J. (eds.), *A New Companion to Digital Humanities*, pp. 420-438. John Wiley & Sons, New York (2016).