Museum Crowdsourcing—Detecting the Limits: eMunch.no and the Digitisation of Letters Addressed to Edvard Munch

Author: Joanna Iranowska (University of Oslo, Norway)

Keywords: digitisation, crowdsourcing, text collections, online participation, Edvard Munch.

Biography

Joanna Iranowska is a PhD candidate in Museum Studies at the University of Oslo. Between 2012 and 2015, she worked as a research assistant in the project ‘Edvard Munch’s Writings’. She is a part of the research project ‘Museum: A Culture of Copies’ and is working on her dissertation on digital reproductions within the art museum context, with a specific focus on the collection of the Munch Museum in Oslo, Norway.

Introduction

This chapter provides insights from the process of running a crowdsourced manuscript project in a middle-sized, one-artist museum: the Munch Museum in Oslo, Norway. This museum is devoted to Edvard Munch (1863–1944), the prominent Norwegian expressionist artist and painter of the famous The Scream. The museum team conducted the project, which aimed to produce an electronic scholarly edition of the artist’s letters, literary works, notes and diaries, today available at www.eMunch.no. The ‘Edvard Munch’s Writings’ research project was established in August 2007. ¹ Four years later, on 20 January 2011, the digital archive ‘eMunch.no’ was launched. The digital archive presents digital facsimiles and encoded transcriptions of the manuscripts as a flexible and searchable historical source for scholars, researchers, students and others. Between 2011 and 2012, a selection of texts was translated into English, German and French. The aim of the third phase (2012–2016), which is the focus of this case study, was to publish the artist’s incoming correspondence (Bøe, 2015). Considering the substantial workload, employees’ limited capacity and the newest trends in digitisation of the cultural heritage (personal

¹ The project was founded by The Freedom of Expression Foundation (Fritt Ord), Arts Council Norway (Norsk Kulturråd), The City of Oslo Art Collections’ Institute for Scholarly Research (Kunstsamlingenenes Institutt for Vitenskapelig Forskning), Ministry of Foreign Affairs, Tate Modern and the Munch Museum.
communication, Digital Collection Manager Hilde Bøe, 27 June 2017), the project editors decided to experiment with asking the audience for help with proofreading and encoding a collection consisting of nearly 6,000 letters. These letters were received by the painter between 1876 and 1944, sent by his friends, family and acquaintances, among them lovers, patrons, doctors and many prominent figures of European modernism. The museum established a *wiki workplace*, a digital Wikimedia platform that was open between May 2014 and June 2016, which was expected to accelerate the proofreading and encoding process and minimise production expenses for the online edition. However, the project gathered fewer volunteers than expected and did not deliver spectacular outcomes in terms of the number of completed transcriptions. This is hardly an isolated incident when experimenting with crowdsourcing within galleries, libraries, archives and museums (GLAMs). Therefore, I believe that investigating this new model of collaboration from the theoretical viewpoint of actor network theory (ANT) as a trail of associations between heterogeneous elements (Latour, 2007) can offer new, valuable insights concerning the challenges that GLAMs are frequently faced with in relation to crowdsourcing. Thus, in this short case study, I read the eMunch.no crowdsourcing project as a group of human and non-human actors assembled and mobilised to produce an online edition of Munch’s writings. To explore that network, I followed and characterised the crucial actors who performed in the making of the scholarly edition: eMunch.no, the letters, editors, wiki workplace and crowd. Using ANT as an analytical tool enabled me to trace relationships among the actors and shed light on the threats and limitations of engaging volunteers in a scholarly edition transcription project.

To follow the actors, I used different investigation techniques: interviews with the editors, analysis of documents, participant observation of the volunteers, online questionnaires and technical walkthroughs of the wiki workplace and eMunch.no, which resulted in hundreds of screenshots. In addition, between 2012 and 2015, I was an eMunch.no research assistant, which allowed me an insider’s perspective.

**eMunch.no**

The online edition of Munch’s writings, eMunch.no (see Figure 1), is my point of departure, since it emerged as an agent mobilising the editors, and the establishment of a wiki workplace and crowd for the realisation of the crowdsourcing project. The aim of the crowdsourcing was to supplement eMunch.no with facsimiles of letters addressed to Munch, accompanied by transcripts and equipped with editorial commentaries (Bøe, 2015). The eMunch.no was where most of the crowd met Munch’s related documents for the first time; it was the gate to the wiki workplace in which crowdsourcing took place.
In contrast to hardcopy books, eMunch.no is a dynamic digital publication which is ever growing and allows new research to be continually included as we add reference data and new documents that have been unknown or unavailable to us’ (Henrichsen & Ydstie, 2011, p. 5). The digital versions of letters established multiple connections with each other, both within the archive via hyperlinks (they are searchable and encoded according to names, places and dates) and outside the archive (to the editors, volunteers, software and hardware). Every month, about 1,574 users visit the archive.²

[Figure 1. eMunch.no. A letter from Emanuel Goldstein to Edvard Munch (MM.K.1497-01).]

Being the result of a collaborative effort, the transcripts embody what textual scholars and book historians propose to call ‘social text’ or ‘social edition’ (Robinson, 2016; see also McKenzie, 1996; and McGann, 1983). The transcripts in the digital archive are shaped and transformed by many other factors than Munch’s sole authorship. Moreover, ‘the social’, in the light of the material semiotics and ANT, is composed of multiple human and non-human actors—the editors, volunteers, underlying technology and negotiations between them—that together have brought eMunch.no into existence.

The Letters

The letters’ cultural importance prompted the launching of the ‘Edvard Munch’s Writings’ project. Following theorists interested in ANT and materiality, I read the letters as objects with a certain degree of agency (Petch, Larson & Gosden, 2007; see also Latour, 2007) that were capable of assembling an actor network of museum professionals, volunteer workers, hardware and software around them. Throughout the years, the letters provoked researchers’ interest, encouraged research grant proposals, the hiring of specialists and finally, the establishing of a wiki workplace. As expressed by the Munch Museum’s director and chief curator Ingebjørg Ydstie, ‘Munch’s letters, notes, journals and a number of other documents have always been a key wellspring for the comprehension of his art’ (Henrichsen & Ydstie, 2011, p. 5).

The painter had a habit of preserving papers, letters and documents. In 1943, a year before he died, in a letter to a good friend he explained how he kept most letters while living a nomadic life: ‘I have never used a waste bin but my suitcases’.³ Today, the archive in the Munch Museum in Oslo houses and

² Statistics for 1 January–31 December 2017.
³ A letter from Edvard Munch to Christian Gierlof, 1943, MM N 3027, Munch Museum. Author’s translation.
preserves about 90 per cent of the writings related to the painter. A considerable part of the collection is the correspondence that Munch received—nearly 6,000 letters from 556 senders—between 1876 and 1944. This large and heterogeneous collection consists of long letters, postcards, picture postcards, greeting cards, envelopes, invitations, telegrams, visit cards and so on. Hundreds of different handwriting styles and many languages are represented (Norwegian, German, Danish, Swedish, French and English). The letters provide unique insights into the artist’s life and career. They:

contain information about pictures that he was working on, commissions, exhibitions and sundry information regarding the artworks. They occasionally give us a good overview of where he was staying and the people he met. (Woll, 2011, p. 67)

The collection of letters informs the research of art historians, philologists, cultural historians and other scholars. In addition to the original manuscripts, the archive maintains typed versions of the letters that have been transcribed systematically from the 1940s until today by museum librarians (personal communication, librarian Inger Engan, 20 March 2018).

In 2012, in the initial stage of the third phase, the fragile and vulnerable manuscripts of letters to Munch were handled carefully by research assistants. They were taken out of a filing cabinet, unfolded and examined. They were given museum registration numbers, recorded in the museum’s database and placed in protective acid-free sleeves and folders. The folders were moved from an archive study room to the repository of prints and drawings. Almost 70 years after the painter’s death, the letters became officially registered museum objects—their cataloguing resulted in stabilisation of the chaotic archival material and transformation of its ontology. It was estimated that the collection consisted of about 3,000 letters addressed to Munch (Haugsland, 2011); as it turned out, when the registration was complete, it was twice as many. All 6,000 letters were scanned and saved on the Munch Museum’s hard drive. On a screen, letters were flattened into two-dimensional images and deprived of their material features. They could not be touched, smelled or physically turned around; their ‘materiality is translated into a sequence of zeroes and ones’ (Müller, 2010, p. 300). After the digitisation, the volunteers and editors started to rely mostly on the digital versions of the letters and only rarely consulted the originals resting in the archive. The digital versions took over.

Next, the digital facsimiles were uploaded to wiki workplace. As a result of turning the letters into digital facsimiles and embedding them in the global internet network, new actors, the networked objects, were brought into being (Cameron & Mengler, 2011). The networked objects juxtaposed digital images with OCR-transcripts and metadata (see Figures 1 and 2).
We are used to thinking about a museum as containing a collection of objects gathered by people, but, as Petch et al. (2007) interestingly observed, it is possible to reverse this thinking: ‘Certain kinds of objects attract certain people’ (p. 65). In the case at hand, ‘Since the Munch Museum opened in 1963, the Museum’s professionals have continually worked with the extensive handwritten material’ (Henrichsen & Ydstie, 2011, p. 5). The creation of the online edition of letters mobilised the forming of a ‘text project’\(^4\) group consisting of an edition philologist (the project leader), a curator, a research librarian, a librarian and one or two newly recruited research assistants with backgrounds in philology or literature (one of them the author).

Between 2012 and 2016, the project group held weekly meetings. However, only two project workers worked full-time with proofreading and encoding, and the two librarians were contributing when their duties allowed them. Since the museum keeps typed transcriptions of the letters from the 1970s, the crowdsourced task did not require transcribing the manuscripts from scratch. It consisted of comparing the digital facsimiles of letters with text captures (optical chapter recognition [OCR]-scanned transcriptions from the 1970s), inserting of break lines and extensible mark-up language (XML) code (encoding tags, elements and attributes) so that the text layout would resemble the original manuscript. Typically, in each OCR-scanned page, approximately 10 to 15 mistakes were found (e.g. ‘!’ scanned as ‘l’, ‘rn’ as ‘m’, or ‘0’ as ‘O’). Statistically, one project worker could process 10 to 20 medium-long letters a day, which is quite limited considering the total workload of 6,000 letters.

Although the technology did a considerable amount of the work recognising the text and creating text files, the time-consuming quality control was necessary: ‘Handwritten records in particular are currently difficult or impossible to digitise and process using computers, so human input may be necessary to transcribe or verify their content’ (Blaser, 2014, p. 49). Despite the technological progress, the making of an online scholarly edition takes thousands of hours of human labour (Eggert, 2016). Traditionally, proofreading used to be conducted by textual scholars, but considering the vast amount of material, limited funding and the newest trends in digital humanities (personal communication, Digital Collection

\(^4\) An informal name of the project used frequently by the museum staff.
Manager, 27 June 2017), the project group decided to establish an online transcription desk, modelled after the broadly known and prominent ‘Transcribe Bentham’ project. Reliance on voluntary labour potentially meant a power shift in editors/museum public relations and encouraged a change of attitude towards traditional museum custodianship. Editors managed the collections at large and arranged the letters into categories: ‘family’, ‘friends and enemies’, ‘portraits and models’, ‘biographers, art historians and critics’ and ‘artists’. Displaying digital images of letters was a form of online curation.

Nevertheless, the potential power shift is challenged by the fact that transcribing is never purely mechanical copying—there is much space for interpretation (Eggert, 2016). Thus, to obtain a high-quality edition, the project group took scrupulous control—two editors checked each delivered page. This precision, together with detailed guidelines, created a sense of authority that clashed with the flexible idea of participatory culture. While crowdsourcing usually means loosening control (Westberg Gabriel & Jensen, 2017), the editors remained facilitators of the participants’ interaction with the networked objects, which means the power shift was not as drastic as one might think. Unfortunately, my material does not illustrate how the crowd responded to this oversight. However, I did discover that some experienced volunteers were inserting encoding tags manually rather than following the guidelines and using an integrated toolbar; while some newbies found tagging too difficult and gave up quickly, not feeling ‘academic enough’ (personal communication, anonymised user, 5 July 2017). Nevertheless, the editors invested considerable time in creating the wiki workplace and building the network of volunteers. They provided feedback after each contribution and explained guidelines. However, it was not enough to sustain interest and attract enough users, which I will address in ‘The Crowd’.

The Wiki Workplace

The wiki workplace was an online platform, a workspace for open collaboration with people from outside the museum and academia. This is where facsimiles of letters were mediated via internet on the volunteers’ screens so that, one by one, they could be proofread and encoded by the crowd. The facsimiles embedded in wiki performed as material through which the volunteers learned to transcribe. Volunteers’ transcripts were proofread by the museum’s editors, who added philological commentaries and linked to people, institutions, places and dates already in the digital archive. This collaborative model was

---

5 ‘The Transcribe Bentham project deals with transcribing the difficult handwriting of the English jurist and philosopher Jeremy Bentham (1748–1832). See www.transcribe-bentham.da.ulcc.ac.uk.'
described as ‘editing without walls’ by Robinson (2010), who observed that there were ‘no walls around this workspace: its edges will merge with the whole internet’ (p. 61).

[Figure 3. Wiki workplace homepage (www.emunch.no/wiki/).]

The editors chose to rely on open-source wiki software developed by a global stakeholder, Wikimedia Foundation, which aimed to use the collaborative power of the internet and support creating and sharing knowledge of all kinds (MediaWiki, 2017). The simple graphical user interface (see Figure 3) was familiar to the millions of people who use Wikipedia every day. Once volunteers entered the page, the interface, filled with project descriptions and menu systems, guided them. The transcription desk (see Figure 2) consisted of a zoomable high-quality image of a manuscript on the right and a text data entry box with a toolbar on the left. Selecting a button on the toolbar (line break, page break, heading, paragraph, addition, deletion, questionable, illegible, marginal note, underline, superscript, unusual spelling, foreign, ampersand, long dash and comment) generated the appearance of TEI XML code. The toolbar was supposed to simplify the text-encoding process; however, occasionally, as a result of a browser error, the buttons were duplicated and shown twice (see Figure 2), which could have confused and discouraged the volunteers.

Moreover, some crucial metadata from the museum database were shown underneath the transcription box. The navigation menu on the left consisted of links to guidelines (available in video and text form) and different manuscripts. All these particular elements of the wiki workplace interface worked as Latourian mediators, providing the volunteers with agency. The letters in the form of digital facsimiles were transported from the museum’s archive to the participants’ homes, where they were mediated on their computer screens. They could be read, copied from one hard drive to another, manipulated or shared in social media with a couple of clicks. The digital facsimiles (networked objects) provided information about the original letters—in other words, they acted as informational copies (Christensen, 2017). Drawing on ANT, specifically performativity of non-human actors, Bratton (2015) suggested that ‘platforms are what platforms do’ (p. 40); they ‘centralize and decentralize at once, drawing many actors into a common infrastructure’ (p. 46). The wiki workplace acted as an online repository through which the local collection of letters was globally distributed and multiple new connections were established—
with each other, different webpages and the volunteers. It resulted in crossroads of countless attachments, impossible to trace.

The eMunch’s wiki commenced on 21 May 2014 and was open until 5 July 2016. The project leader admitted that a longer time span could possibly have improved the number of completed letters, but its duration was determined by short-term, project-based funding (personal communication, Digital Collection Manager, 27 June 2017).

**The Crowd**

The mobilisation of human actors was crucial for the network to function: ‘for crowdsourcing to be successful, it must rely on a robust, active, motivated crowd’ (Brabham, 2013, p. 126). As in every crowdsourcing project, the editors aimed to build a sense of community among users through delivering systematic feedback and creating room for discussion. Over one and a half years, the project gathered 342 registered users who transcribed 384 pages in total (ca. 6.6 per cent of the letter collection). Of the users, only 17 volunteers were significantly active and contributed on a daily basis, transcribing between 1 and 100 pages each. More than half of the project participants were German—the wiki workplace was available in Norwegian and German, since these two languages constituted a vast majority of the material. Consequently, the target group was smaller than those of projects directed towards English-speaking people. Another 16 users established an account but had not started transcribing. Surprisingly, as many as 296 out of 342 were spam accounts. Twenty of them were creating spam pages, which, in light of ANT, can be read as disturbing and useless connections. The spam users were probably mechanically generated and appeared in clusters, as the project leader explained. The editors had to devote time to tracing and banning the fake users and erasing the spam pages. This unexpected controversy was resolved by introducing more advanced user registration, which resulted in a new challenge—a decrease in the number of new volunteers.

Overall, the project’s *crowd* turned out to be rather tiny. The editors from ‘Munch’s Writings’ and ‘Transcribe Bentham’ admitted that their initiatives did not appeal to a crowd like the word ‘crowdsourcing’ suggests. It is more accurate to speak of ‘niche-crowdsourcing’ (personal communication, Digital Collection Manager, 27 June 2017) or ‘crowd-sifting’ (Causer & Terras, 2014, p. 73). However, participants who were involved were highly qualified and delivered good-quality transcriptions. Most were middle-aged females with higher degrees and some experience in transcribing,
and who spoke at least two foreign languages (Bøe, 2016; online questionnaire). They became involved ‘for fun’, to ‘keep their editing/text coding skills fresh’ or simply because of ‘curiosity’ (online questionnaire). In a survey, volunteers said that they felt appreciated by the museum staff and they benefited from volunteering by ‘learning about socially important persons in Munch’s times’ or by having had an opportunity to ‘work with Munch material’. Some emphasised their interest in the technological aspects of the project that offered them an opportunity to obtain ‘insight into a different transcription platform’.

The editors attempted to establish connections with volunteers through advertising the project on the museum’s social media (Facebook, Instagram, Twitter and a blog), publishing press releases (which resulted in coverage on national radio NRK P2, *Kulturnytt*) and distributing flyers. When running a crowdsourcing project, sustaining the interest is as crucial as attracting new users. The project leader from the Munch Museum described the volunteers’ tendency to lose interest after a while: ‘most people did a bit and vanished’. When asked about the reasons for dropping off, the participants mentioned work duties, illness or lack of beginners’ enthusiasm (online questionnaire). The project’s editors admitted that after one and a half years of running the project, their experience was not entirely positive:

> We worked hard and got some attention, but recruitment turned out to be difficult. We did not manage to build a community of volunteers … We have a feeling that we should have used our own time to work with the material. (Bøe, 2016)

Those who participated admitted that while they had a sense of team spirit, it was not strong. The users have not spoken or made friends (Bøe, 2016; online questionnaire).

The project leader estimated that roughly 10 systematically contributing volunteers, each delivering 250 pages, would have been sufficient to execute the project in just 15 months.

**Conclusions**

In this short case study, I followed and characterised some central human and non-human actors in the ‘eMunch.no’ crowdsourcing project. The crowdsourcing projects create a few planes of reference that

---


8 In quotes from the survey, the original syntax and spelling has been used.
are shifting out (Latour, 1999): one in which the editors are active, a second in which the action is delegated to another character, a non-human—the wiki workplace—and finally, a third in which volunteers take actions. Tracing of the online edition eMunch.no, the letters, editors, wiki workplace, crowd and associations among them revealed that when conducting a manuscript crowdsourcing project, the biggest challenge was to evenly mobilise all the actors in the actor network. Recruiting new volunteers and sustaining interest turned out to be more difficult and time-consuming than expected. As a result, the relationships between existing actors were imbalanced and disrupted. This controversy was resolved by distributing the proofreading tasks to the editors (specifically research assistants and librarians), which prevented the potential power shift and intended emancipation of the audience. Fewer transcripts were completed than had been anticipated, which means that in the case at hand, crowdsourcing did not improve effectiveness of editing and did not solve the workload problem.

As the analysis emphasised, it was not solely the technology but rather an imbalance in the interplay between the actors that was problematic. As Latour (1996) observed, innovative technologies are weak, fragile and ‘hypersensitive to variations in its environment’ (p. 291). Perhaps the biggest challenge is that crowdsourcing is a model without clear lines of authority. By opening the archive and involving volunteers, the dynamic of the editing work dramatically changed. As this case study has illustrated, adhering to the traditional institution-dominated model (Robinson, 2010, p. 59) might delay the project or discourage the volunteers.

Several suggestions arise from this analysis that might improve participation in other projects. The first is to attempt to abandon the traditional editors–users power relations through ensuring space for creativity, comment and discussion. Second, secure enough means and resources to work with advertising, recruiting volunteers and sustaining their interest—mobilising the crowd takes more than one might initially think. Third, an interesting approach could be to consider how much quality control of the transcripts is absolutely necessary—publishing of unready, imperfect transcripts could save the editor team considerable time. Finally, long-term planning and the project’s longer duration increases the chances of its profitability. We are witnessing a participatory turn in GLAMs, and there is probably no way back. Therefore, it is crucial to learn how to use innovative and experiential modes of collaboration effectively and learn from positive and negative experiences, to which this short essay has aimed to contribute.
References


Haugslund, Å. (2011). ‘... this chaos of letters I have collected ...’. Edvard Munch, the letter writer. In M. B. Guleng (Ed.), *eMunch.no—text and image*. Oslo: AIT Otta AS.


Acknowledgements

The author wishes to thank Brita Brenna, Hilde Bøe, Janne Werner Olsrud and Thea Aarbakke for their comments on earlier versions of this essay. Thanks also to my reviewers, whose suggestions are greatly appreciated.
Kære Munch!


— — —

Jeg feder mig ned, i stedet med.

Figur 1. eMunch.no. A letter from Emanuel Goldstein to Edvard Munch (MM.K.1497-01).]
Figur 2 Wiki workplace transcription desk. Transcribing of a letter from Emanuel Goldstein (MM.K.1497-01).]
Figur 3. Wiki workplace homepage (www.emunch.no/wiki/).

Edvard Munchs tekster

Valkommen til det digitale arkivets wiki

1.1 Brevene til Munch
1.2 Dugnadshjelp!
2.1.1 Vi vil hjelpe s?l? Would you like to hel p?
2.1.1 Hvor mange brev er ferdigstilt?
2.1.1.1 Pr. 15.11.2015 er 2927 av 5838 brev ferdigstilt
2.2 Hvor mange brev er lagt inn i wikien?
3 En lekks / A Thank You Note

Valkommen til det digitale arkivets wiki

Arbeidet i wikien er nå avsluttet, og prosjektet fullførtes internett. Takk for alle gode bidrag!

Wilken er en digital arbeidsbok hvor frivillige har bidratt med konstruksjon, transkripsjon og tekster atning i Munch-museets arbeid med Edvard Munchs tekster. Wikien er en del av museets prosjekt Edvard Munchs tekster som hittil har resultert i det digitale arkivet emunch.no.

Brevene til Munch

I wikien legger vi ut brev til Munch fra venner, forstyringsdetaljer m.m. Til sammen finnes det 5838 brev til Munch i museets samling. Vi vil nå gjøre bravenes fritt tilgjengelige på nettet som lesbar tekst og foto av originalet. Brevene er på flere språk (norsk, tysk, fransk, svensk og dansk er de mest frirkorte) og fra et stort antall avsendere i mange land.

På langt nær alle brevene er «falt i det fri» enså, og derfor begynner vi arbeidet med de brevene som er det.

Dugnadshjelp!

Til dette arbeidet vil vi gleme ha hjelp!

Hvis du ønsker å bidra til arbeidet med Edvard Munchs tekster, kan du lese mer under Komme i gang og se vår video som viser hvordan du kommer i gang.