The Experience of the Citizen Scientist

Edited interview with Amy Clotworthy, Ph.D. – independent researcher, Danish National Archives

Amy Clotworthy: I did a project in collaboration with the Danish National Archives, which has been conducting the country's longest-running crowdsourcing project. It was started in 1992, and I was brought in because they didn't have very much information about the volunteers who were involved with the crowdsourcing. They wanted to get a little bit more information about the practices and processes, and what exactly it meant to be a volunteer at the National Archives. The crowdsourcing project was basically started because a number of amateur historians and genealogists were interested in making the records that they were already working with available to other people, so they wanted to make them available to other historians and other researchers in an academic context, as well as the general public and other genealogists. So, it was really sort of a ground-up project in that the genealogists who are already doing a lot of this work started to initiate interest in making everything digitally available. And so, they gathered together a number of resources – a number of people – and in collaboration with the Danish National Archives, they developed the crowdsourcing database, which is called the Danish Demographic Database. And that led to a more formal crowdsourcing portal, which is simply called Crowdsourcing Portal. That was started in I believe 2014, in order to really organise and compile lot of the digital information that the crowdsourcing volunteers were working with.

You weren't involved in the project in its initial stages, but do you know how the project recruited their participants, their citizen scientists?

Yeah, I believe it was really quite a snowball effect – that this particular group of amateur historians and geologists were quite – they showed a lot of interest in making sure that these records were digitally input into the computer systems. So, they had a good relationship with the employees and the people who were in charge of the organisation at the Danish National Archives, and they really initiated the project on their own volition, you could say. So, they started organising it in a very informal way, and then the process became a lot more organised as the years went on. They developed a number of different templates – a number of different platforms that the volunteers could use to put this information into the computer system and digitalise everything. This was the photography and the actual information from the archival material. And it was really the volunteers themselves who initiated it and, through their own genealogical networks, they also recruited additional volunteers. So it's really been a process of outreach from the ground level, you could say, among the volunteers – among the genealogical researchers themselves. It's only been organised in the last few years in a more formal way through this crowdsourcing portal. There are employees at the Danish National Archives who are responsible for recruiting and retaining the volunteers they have, and there are several different ways that they do that, mainly through social media – they have a

Facebook page – and also through a lot of the different events that they host at the various locations of the National Archives.

What kind of unexpected outcomes did you find in your own research in terms of the participants' feelings towards this project?

Yeah, that's an interesting question. My background as an ethnologist is looking a cultural processes – how people have certain practices in their everyday lives that may impact culture, and how culture impacts practices. I was interested in looking at how those particular practices (of the volunteers) might not necessarily conform exactly to what the Danish National Archives was expecting from their work. In other words, there are there certain challenges that the volunteers experience when they're doing this type of work, and I conducted the short project to look at their experiences – to look at their perspectives on doing this type of digitalisation. And I found that it was really interesting because they were incredibly motivated; they were incredibly invested in the outcome of this work. But because we're humans, and the volunteers are humans, they don't necessarily conform to a particular template, and they often make little interpretations of the material in a certain way. So, even though the National Archives will tell them, "You can only put in the name exactly as it's written – even if there's a misspelling, you have to capture the exact name it's written", the volunteers might say, "Well, I know that this person is called Petersen (not Peterson), and I'm going to spell it the correct way because I know from my other research that there might be a connection to this other Petersen". So, there's this process of adaptation that takes place and interpretation. And that, of course, is part of historical research – meaning, that everybody gets involved, and there may be interpretations of a date or place or a practice in a death certificate, so something is written specifically. But the volunteer knows from their other research, "Well, this isn't quite right", so they may make a certain adaptation. So, I think that was really interesting to find out through the project. I think it just goes to show how engaged they are, and how interested they are in making sure that things are easier for future researchers who come in and search for records.

And the volunteers themselves – were you aware of any expectations they might have had before they joined the project, and do you know if they changed or have they realised those expectations whilst doing this work?

I think the volunteers expected that they could come into the crowdsourcing project – I'm thinking specifically of the people who are transcribing the records, at this point – and I think they had an expectation that they could come in and just easily find a name or birthplace or a death certificate. (They could find) all the information that's included in those records, and that it would go into this online template. But what I think was a particular challenge for them was actually reading some of the historical documentation because, in Denmark, so much of the handwriting from the 1700s and 1800s is this very specific kind of Germanic text and Germanic handwriting. The Danish National Archives actually offers courses for the volunteers in reading this very specific kind of handwriting, so that they can more easily go through the records. I think that was really interesting for them to realise that, "Oh, it's not exactly the language (we use) in our current day-to-day Danish" – it's actually a very specific form of, not only

handwriting, but a way of forming sentences in the information (which makes it) very specific to that historical time.

What was the demographic of the people who volunteered to take part in this project?

Yeah, one of the reasons I was brought into the project is because I have expertise in working with older citizens. The demographic profile of these volunteers is that most of them are over age 65. (Most of them have) retired from work – some of the people I talked to were also involved in some consulting or part-time work, but most of them had already gone on pension and were just interested in volunteering because they had a little bit more free time.

In this particular study, I was interested in speaking to a range of different people from all sorts of backgrounds. Because my project was a short-term project, I was really only able to speak to people who were working as volunteers at the office in Copenhagen. The Danish National Archives has several locations around Denmark but, with the time constraints, I was unable to go out to some of these other offices. But I did speak to the volunteers in Copenhagen and Odense, which are the two primary offices. And all of them are retired people, and all of them have a variety of different backgrounds in terms of academic experience – some of them have PhDs, some of them just went to trade school. They've had a variety of different careers in their previous working lives – some of them had been schoolteachers, some of them had been IT specialists. So they really, I think, reflected a broad range of characteristics in terms of representing the Danish population, you could say, and I spoke to a mix of men and women who were involved with the different volunteer functions that the Danish National Archives offers. So I think, in terms of the profile of the volunteers, it was a wide range of different tasks that they had (at the Archives). They were varied in terms of their gender, but in terms of their age profile, they were (mostly) over 65.

Did you speak to any of the project leads at the Danish National Archives about their experiences with this crowdsourcing project?

Yes, throughout the short project that I conducted, I had fairly close contact with Katrine Tovgaard-Olsen, Archivist and crowdsourcing coordinator, who's directly responsible for organising all the crowdsourcing activities. So she was the one who helped me get access to the volunteers initially so I could make contact with them, and she made introductions of me and my project. So the volunteers knew exactly what I was looking for, and what I was hoping to find out with this particular project. She was really my gatekeeper in terms of providing access.

And there were several other people – historians who I have contact with, and who I knew in advance – who helped me get into the project initially, so I had a lot of good contact with people directly responsible for the crowdsourcing initiative at the Archives. They provided me with a lot of insight in terms of the background, and in terms of their expectations for the project.

And what was their overall feeling about the crowdsourcing project once it had been finished?

Well, the crowdsourcing project is ongoing. There's no lack of historical records that need to be digitalised, so there's no end in sight, really. The volunteers (digitalise) nearly one million records per year, and they're still only up to I think the 18th century[1].

Denmark, as you know, it's an extremely small country, so it's quite impressive that so many records even exist. They have records that are not only birth and death certificates, but all the church records – and there were thousands, millions of church records, especially in earlier times when most people were attending church on a regular basis. Plus, they have a lot of war records – things having to do with the Danish West Indies in historical times – so there's quite a lot of variation; there's a large variety of historical material that needs to be digitalised.

What did the participants gain from taking part in this crowdsourcing project?

For the volunteers themselves, they feel a real obligation – not only to the Danish National Archives because they've collected all these materials for as long as they have, and have allowed access to the general public to use these records and just do research with these records. But the volunteers also feel a very strong sense of obligation to future generations of researchers. So, for them, what they get out of it is this sense of connection with all the researchers and genealogists who have come before them, and a sense of connection to the people who will come after them, and who they are helping to make these records available. So, it's really a sense of pride that they have in being part of the whole process. It's not just, "Well, I've photographed this one record" or "I've transcribed the information from this one record" – it's really, "I'm part of this whole community of people who are engaged with historical research and engaged with digital humanities". And for them it's really very life-affirming in a lot of ways. I think again, because my background is in health and ageing studies, we can look at this as being really important to older volunteers' quality of life. They're getting a sense of community; they're getting a sense of connection with not only people or other researchers from the past but also other researchers in the future, and they have a particular engagement with the actual people represented in the historical records. So, all of these factors, I think, are directly related to helping them feel a sense of community and connection with the people that are implied in a lot of these records. So, this whole project is really affirming for them and really helps keep them engaged in a particular way. I think they would say that the benefit that they get is feeling part of this whole process.

---

[1] The official count from the Danish National Archives: "Since 1992, more than 1,000 volunteers have entered more than 20 million records in DDD. In 2018 alone, the volunteer photographers digitalised more than 1,000,000 records, which can be viewed on https://www.sa.dk/en/ "

Finally, having conducted your own research into the volunteers and the practices there, what do you think works in a crowdsourcing project like this – and what doesn't necessarily work quite so well?

That's a good question. I think one of the most important things in terms of retaining volunteers for the transcription in particular is to make sure they have the support they need, and to make sure that they feel recognised and appreciated for their efforts. And I think in terms of the project that I conducted, where I was specifically talking to all sorts of different volunteers – the counsellors, the photographers, and the transcriptionists – about their experiences doing the digitalisation of the historical material, working with members of the public, helping to develop the research involved with Digital Humanities, I think it was important for them to feel like their contributions have made a difference. And that me coming in as a qualitative researcher and asking them about their experiences was also another form of acknowledgement for the work that they're doing. The people who I spoke with feel a strong sense of obligation, as I mentioned before, and they don't necessarily think about it in terms of "Well, this is something special that I'm doing". They just feel, "Well, I'm interested in genealogy, and this is helping other people, so this is important to me".

But I think it's important in terms of recruiting and retaining future volunteers for crowdsourcing projects to acknowledge that they have a human value – they have their own lives that go beyond these projects. They may have health problems, they may have problems in their relationships with their children or grandchildren; they have all sorts of different issues that could impact how much they're able to produce in terms of the output of the digital records. So, for any organisation that's looking to recruit new volunteers for crowdsourcing projects, I think it's really important for them to recognise the fact that people have their own lives, and that it is a real gift of people's time and effort and energy to get involved with these projects.

It's (also) important to have a good structure in place – to have some sort of organised process in terms of what the expectations are, what materials are available for the crowdsourcing, and what the (process should be). And once those things are clear, then to make sure that there's an open dialogue that's very iterative to hear the volunteers' concerns and problems, so that the process becomes smoother for everybody involved. And, as the project develops over time, to keep that pipeline open to make sure that the expectations are still clear, and that everybody is satisfied with what's taking place.

Amy Clotworthy is an independent researcher at the Danish National Archives. She was interviewed by Vicky Garnett on behalf of PARTHENOS in April 2019.

PARTHENOS is a Horizon 2020 funded project.