**VOSS — Crowdsourcing**

<0 images>

**From Crowdsourcing to Knowledge Communities: Creating Meaningful Scholarship through Digital Collaboration**

Voss, J., Wolfenstein, G., Frank, Z., Heuser, R., Young, K. and Stanhope, N.

Over the last three years, Stanford University’s Center for Spatial and Textual Analysis (CESTA) has jointly led a research project in partnership with Historypin to examine the potential for leveraging crowdsourced information about photographic, map, and textual content for humanities research. The project was funded by the Andrew W. Mellon Foundation and sought to study elements of user interface, user interaction, community engagement, and collaborative partnerships.

Our work mirrors the findings of others utilizing crowdsourcing in cultural heritage, in that there is an important distinction within the realm of crowdsourcing between more straightforward knowledge gathering and more complex and collaborative tasks and processes that revolve around what we termed ‘knowledge communities’: small networks of neighbors or enthusiasts representing a group of people that could be systematically organized to share and participate in research for a common aim. Engaging with these communities often requires longer time frames than simpler task-driven crowdsourcing may allow, and is necessarily much more collaborative than extractive. Furthermore, there are important social benefits for the institutions and communities alike that fulfill the new missions of 21st-century cultural heritage and academic institutions.

Throughout our research, we found that not only will there be a much more collaborative future among academic humanists, memory institutions, and knowledge communities, but that only through such collaborations can some questions relative to the humanities be explored.

In this presentation, we summarize the methodology of the project and present key learnings and examples that we hope will assist academic and cultural heritage institutions that are interested in engaging with knowledge communities. In addition, we will provide a way to practically approach such projects with realistic expectations and ideas for successful implementation.

**Methodology**

In this section we describe in some detail the methodology used in each of the three projects and the different approaches and communities we were working with. We had three sub-projects: the Year of the Bay, Living with the Railroads, and the Emotions of London. Together, these projects allow us to offer a much broader range of findings and practices than we could have with a single one.

**Crowdsourcing and Knowledge Communities**

We’ve seen that digital-transcription projects on the humanities side, or ‘citizen science’ on the scientific side, tend to blend a combination of ‘crowd’ and ‘community’ (Clauser and Wallace, 2012). In the former, loosely associated individuals perform autonomous and relatively simple tasks, while the latter often requires collaboration and a more social and coordinated element (Haythornthwaite, 2009). In Living with the Railroads and Year of the Bay, we were not dealing with questions of transcription or necessarily easily defined tasks, but rather more complex questions of contextualization and even contribution with a specific goal in mind, still meeting definitions of crowdsourcing but leaning toward more complex tasks requiring an engaged community (Dunn and Hedges, 2012). In both projects, there was a collaborative approach between the participating institutions and the individual participants toward a common aim, however broadly defined—be it exploring the human and ecological history of the bay or tracing western expansion through life around the railroads. This also speaks to the potential for strengthening and measuring community ties through this work, which Nick Poole and Nick Stanhope describe as a new approach not only to digitization, ‘but of a new role for museums as places of engagement and participation in which the disciplines or curatorship, digitisation, collections management and documentation are shared with the user in a joint effort to develop shared cultural capital’ (Poole and Stanhope, 2012).

As these projects progressed, it was helpful for us to understand and define the various communities we were increasingly in partnership with, not as users or as a crowd, but as ‘knowledge communities’. Whether they were local residents or history enthusiasts who were knowledgeable about their area, or in some cases obsessed with solving a history mystery that was presented to them, or who had specific topical knowledge about details pertaining to one railroad line, these small networks of neighbors or enthusiasts represented a group of people who could be systematically organized to share and participate in research for a common aim. After two years of working with specific individuals and groups of individuals, we came to see this as an important distinction within the realm of crowdsourcing, and an approach that is most importantly collaborative in intent and design rather than extractive and focused on the needs of an institution or a researcher. This approach has also been termed community-sourcing (Sample Ward, 2011).

Literature on crowdsourcing generally supports the process (which remains frustratingly variably defined) on the basis of whether some sort of utilitarian goal is met: entries are added (Wikipedia), materials are transcribed (Transcribe Bentham), data is classified (Galaxy Zoo). In this, understandings of crowdsourcing, both definitional and in terms of results, remain wedded to a market model. Can you get data or analysis? Can you get it more quickly or more cheaply? These are the primary questions. The question that motivated our project—whether crowdsourcing can be useful to humanities researchers pursuing humanities projects—found a very different answer, though utilitarian goals remain. Our three sub-projects engaged three different types of crowds: the amorphous public of people interested in the Bay Area, the expert community of train enthusiasts, and the paid community of Amazon Mechanical Turkers. All three were, unsurprisingly, only partially successful in terms of the research questions initially posed. There is usable and interesting data (images and metadata uploaded to Historypin for the first two, analysis of literary passages as to emotionality in the latter). But it isn’t perfect, and in some cases doesn’t quite meet the goals of the various PIs. But the conclusions we have drawn thus far ask us—and, by extension, other researchers—to look beyond the data to the process itself.

There is increasing segmentation of crowdsourcing approaches in cultural heritage (Ridge, 2014), with great import for the academy as well as memory institutions. Crowdsourcing offers academic humanists a different way of engaging the public, especially in collaboration with non-academic organizations. We suggest that, if researchers have flexibility in their projects, building or collaborating with knowledge communities can result in original research, and an engagement with the community that is typically missing from university-based research. Without this flexibility, crowdsourcing possibilities become more limited, though still possible, as our work with the Turkers demonstrated.

Furthermore, we feel that museums, libraries, archives, and academic institutions alike increasingly embrace the opportunities they have to positively impact communities, not just in terms of cultural programming, but also as partners in strengthening a range of societal measures (Van Thoen, 2014). Meaningful engagement of knowledge communities can have a range of outcomes in addition to the traditional measurements of crowdsourcing for the institutions as well as the communities, such as increasing interest in and sense of ownership of institutions, strengthening community ties and associational life, and decreasing isolation among frequently marginalized sectors of society such as seniors (Thomson and Chatterjee, 2013).

**Practical Applications**

This section will be presented in detail, as it’s the heart of the practical takeaways for institutions interested in engaging knowledge communities. These are broken into three sections, including (a) design and expectations; (b) methods of engagement; and (c) skills, training, and technology.

**Future Research**

Here we will look at key areas of future research that we’ll be pursuing to continue our work in this field.

**Acknowledgements**

We have so many people to thank for their roles in this project over the last year that it’d be impossible to name them all. The Crowdsourcing for Humanities Research project was funded by a grant from the Andrew W. Mellon Foundation and led by Zephyr Frank at Stanford University, who served as the principal investigator. You can see all of the people who worked on the project on our behind-the-scenes blog,http://crowdsourcingthehumanities.blogspot.com. We’d also like to thank the growing community of practitioners and scholars sharing their experience in this rapidly evolving field, not all of whom are cited below.

**References**

**Clauser, T. and Wallace, V.** (2012). Building a Volunteer Community: Results and Findings from Transcribe Bentham. *Digital Humanities Quarterly,* **6**(2), http://www.digitalhumanities.org/dhq/vol/6/2/000125/000125.html.

**Dunn, S. and Hedges, M.** (2012). Crowd-Sourcing Scoping Study: Engaging the Crowd with Humanities Research. AHRC report, http://crowds.cerch.kcl.ac.uk/wp-content/uploads/2012/12/Crowdsourcing-connected-communities.pdf.

**Haythornthwaite, C.** (2009). Crowds and Communities: Light and Heavyweight Models of Peer Production. *Proceedings of the 42nd Hawaiian Conference on System Sciences*, Waikola, Hawaii, IEEE Computer Society, pp. 1–10.

**Poole, N. and Stanhope, N.** (2012). The Participatory Museum. Collections Trust, http://www.collectionstrust.org.uk/past-posts/the-participatory-museum (modified 3 July 2014).

**Ridge, M.** (2014). Crowdsourcing Our Cultural Heritage: Introduction. In Ridge, M. (ed.), *Crowdsourcing Our Cultural Heritage*. Ashgate, pp. 1–13.

**Sample Ward, A.** (2011). Crowdsourcing vs Community-Sourcing: What’s the Difference and the Opportunity? http://amysampleward.org/2011/05/18/crowdsourcing-vs-community-sourcing-whats-the-difference-and-the-opportunity/.

**Thomson, L. and Chatterjee, H.** (2013). UCL Museum Wellbeing Measures Toolkit. AHRC and University College London, http://www.ucl.ac.uk/museums/research/touch/museumwellbeingmeasures/wellbeing-measures/UCL\_Museum\_Wellbeing\_Measures\_Toolkit\_Sept2013.pdf. (See also UCL Museums and Collections, Touch and Wellbeing, http://www.ucl.ac.uk/museums/research/touch.)

**Van Thoen, L.** (2014). Museums Make You Happier and Less Lonely, Studies Find. Freelancers Broadcasting Network, https://www.freelancersunion.org/blog/2014/03/26/museums-good-for-you-happier-community/.