Community-Based Open Knowledge

Paul Arthur and Lydia Hearn, Edith Cowan University, Western Australia

Extended abstract, for CAPOS 2023:

For centuries, academics, researchers, curators, and collectors have collaborated with engaged members of the community—or "gifted amateurs"—in what is now called *citizen science*. Among the most well-known cases was the British government's offer in 1714 of a monetary prize to anyone who could come up with a way of measuring a ship's longitudinal position. In 1879, the *Oxford English Dictionary* appealed to the public to supply lexicographers with the spelling, definitions, and meanings of rare words in magazines, journals, books, letters, and newspapers.

In the mid-1990s, *citizen science* was conceived of as research that is partially or wholly conducted by nonspecialist volunteers. Interaction between experts and the public grew dramatically when the internet enabled real-time exchange of ideas and content. It is instructive to recall a comment made by Tim Berners-Lee: "The Web is more a social creation than a technical one. I designed it for a social effect—to help people work together—and not as a technical toy." Related to citizen science, the term *crowdsourcing* was coined by Jeff Howe, editor of *Wired*, in 2006. He described how businesses could use the internet not only for "outsourcing" their work but also for "crowdsourcing" it.

Like citizen science, crowdsourcing projects can delegate jobs to the public such as collecting, cataloguing, and annotating, or invite input through involvement in the entire process from conceptualising, framing, and setting up a research agenda and methodology to analysing outcomes. These processes and relationships are also referred to using other terms such as *collective intelligence*, *communal and peer engagement*, *crowd wisdom*, *mass collaboration*, *participatory practice*, and *user-powered systems*.

The wealth of local and first-hand knowledge that can be crowdsourced adds depth and detail to social and cultural projects where individual and situated experience can provide crucial understanding. Enlisting virtual crowds has enabled organisations to reduce costs and enhance economies of scale through co-creation. Established online crowdsourcing services such as Amazon Mechanical Turk, Microworkers, and Zooniverse host myriad projects, with volunteers (paid and unpaid) completing tasks that technology alone cannot achieve. For the humanities, this can include participation in a range of tasks, such as transcribing handwritten text; correcting digitised content; categorising and cataloguing information with structured,

descriptive metadata; collaborative tagging; implicit and explicit linking of data; providing contextual details for artifacts; locating complementary objects to be included in an online collection; recording memories and intangible heritage; commenting and offering critical reflections; mapping visual, spatial, and cultural representations; translating content; and co-curation.

The "crowd" within humanities and cultural heritage crowdsourcing projects does not necessarily comprise large groups of people but can be a small number of interested and engaged citizens who may already have a relationship with the topic in question. "Nichesourcing" targets a niche community with identifiable proficiency or background.

The Transcribe Bentham project, started in 2010, is an early exemplar of crowdsourcing in the humanities. CrowdHeritage—an open platform assisted by the European Commission—is currently using crowdsourcing to improve the metadata of Europeana. Some of the most successful initiatives have been developed in the GLAM sector to encourage members of the public as volunteers and communities of interest to interact with, explore, and interpret, contextualise, and enrich collections. Trove, at the National Library of Australia, is regarded as a world-leading example of crowdsourcing. Planned in 2008 as a portal to the National Library of Australia's discovery services, it has become highly successful, with volunteers correcting the optical character recognition (OCR)-digitised content of Australian newspapers. To date, the public program has seen over 430 million lines of text corrected.

Along with the benefits, crowdsourcing presents known challenges. Ethical issues around "free labour" remain a concern. There are also hidden costs, as involving participants can be complex and time-consuming, requiring purpose-designed platforms and interfaces, and systems for checking, moderating, and processing contributed materials. Even so, crowdsourcing represents one of the most visible and widespread examples of two-way open humanities in action. The capacity to empower community members through seeking and utilising their input to add to or modify existing understandings sets it apart from the kind of open approach that permits entry into spaces where scholarly knowledge can be viewed but not altered. By allowing movement in both directions with the public, crowdsourcing offers one of the more free and democratic types of openness.

Community-based open knowledge, in all its forms, has particular relevance for the humanities, especially in terms of the current "impact" agenda that is encouraging researchers to respond to pressing contemporary concerns and show how their investigations are making a difference and engaging the public.

Bibliography

Berners-Lee, Tim. *Weaving the Web* (Harper Business, 2000).

Hedges, Mark, and Stuart Dunn. *Academic Crowdsourcing in the Humanities: Crowds, Communities and Co-Production* (Cambridge, MA, USA: Chandos Publishing, 2017), 33–38.

Heinisch, Barbara, Kristin Oswald, Maike Weißpflug, Sally Shuttleworth, and Geoffrey Belknap. "Citizen Humanities," in *The Science of Citizen Science*, ed. Katrin Vohland, Anne Land-Zandstra, Luigi Ceccaroni, Rob Lemmens, Josep Perelló, Marisa Ponti, Roeland Samson, and Katherin Wagenknecht (Cham, Switzerland: Springer, 2021), 97–118, https://doi.org/10.1007/978-3-030-58278-4_6.

Hirth, Matthias, Jason Jacques, Peter Rodgers, Ognjen Scekic, and Michael Wybrow et al. "Crowdsourcing Technology to Support Academic Research," in *Evaluation in the Crowd: Crowdsourcing and Human-Centered Experiments*, ed. Daniel Archambault, Helen C. Purchase, and Tobias Hoßfeld (Cham, Switzerland: Springer, 2017), 70–95.

Prpić, John, Araz Taeihagh, and James Melton, "The Fundamentals of Policy Crowdsourcing," *P & I: Policy & Internet* 7, no. 3 (2015): 340–61, https://doi.org/10.1002/poi3.102.