Open archaeology: a survey of collaborative software engineering in archaeological research

Zack Batist[1,*]   Joe Roe[2,*]

[1] University of Toronto
[2] University of Bern

[*] Correspondence: Zack Batist <z.batist@mail.utoronto.ca>, Joe Roe <joe@joeroe.io>

Surveying the first quarter-century of computer applications in archaeology, Scollar (1999) lamented that the field relied almost exclusively on "hand-me-down" tools repurposed from other disciplines. 25 years later, this is no longer the case: computational archaeologists often find themselves practicing the dual roles of data analyst and "research software engineer" (Baxter et al. 2012), developing and applying new tools that are tailored specifically to archaeological problems and archaeological methods. Though this trend can be traced to the very earliest days of the field (Cowgill 1967), its most recent manifestation is distinguished by its apparently embrace of practices from free and open source software (FOSS). Most prominently, since around 2015, there has been a rapid uptake of workflow tools designed for open source development communities, such as the version control system `git` and associated online source code management platforms (e.g. GitHub, GitLab). These tools facilitate collaboration among developers and users of FOSS software using patterns that can diverge quite radically from conventional scholarly norms (Tennant et al. 2020).

In this paper, we investigate modes of collaboration in this emerging community of practice using "open-archaeo" (https://open-archaeo.info/), a curated list of 500+ active open projects in archaeology, augmented where available with data on their activity on GitHub. We apply network analysis to map the networks of collaboration that have emerged among archaeologists using GitHub, and conduct a quantitative exploratory analysis to characterize the nature and intensity of these collaborations. Our results show an uneven adoption of open source collaborative practices beyond the basic use of `git` as a version control system and GitHub to host source code. A majority of repositories have 1–3 contributors, with only a few projects distinguished by an active and diverse developer base. More community-oriented features, such as GitHub's "issues" and "comments", are used in only a minority of repositories. Preliminary findings from ongoing work also shows how collaborative behaviours vary across different subcommunities. By analyzing interactions among archaeologists using the GitHub platform, including the ways they contribute to collective code bases, we demonstrate how archaeological software engineering is beginning to foster new kinds of collaborative commitments that blend open source protocols, DIY work ethic, conventional forms of scholarly communication, and cultural practices unique to archaeology.

References

Baxter, Rob, N. Chue Hong, Dirk Gorissen, James Hetherington, and Ilian Todorov. 2012. "The Research Software Engineer." In *Digital Research Conference, Oxford*, 1–3.

Cowgill, George L. 1967. "Computer Applications in Archaeology." In *Proceedings of the November 14-16, 1967, Fall Joint Computer Conference*, 331–37. AFIPS '67 (Fall). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1465611.1465654.

Scollar, Irwin. 1999. "25 Years of Computer Applications in Archaeology." In *Archaeology in the Age of the Internet. Caa97: Computer Applications and Quantitative Methods in Archaeology, Proceedings of the 25th Anniversary Conference, University of Birmingham, April 1997*, edited by L. Dingwall, S. Exon, V. Gaffney, S. Laflin, and M. van Leusen, 5–10. BAR International Series 750.

Tennant, Jonathan, Ritwik Agarwal, Ksenija Baždarić, David Brassard, Tom Crick, Daniel J. Dunleavy, Thomas Rhys Evans, et al. 2020. "A Tale of Two 'Opens': Intersections Between Free and Open Source Software and Open Scholarship." SocArXiv. https://doi.org/10.31235/osf.io/2kxq8.