

# Digital Humanities Quarterly Keywords

NULab Seedling Grant Proposal  
Julia Flanders, PI

## Background and context

Digital Humanities Quarterly is an open-access, peer-reviewed journal published by the Alliance of Digital Humanities Organizations and the Association for Computers and the Humanities, and hosted at Northeastern University. The journal has been in publication since 2007 and has published over 400 articles on a wide range of topics in digital humanities. The journal serves as a central point of access and encounter for digital humanities scholars and practitioners as well as a broader readership in the humanities and computational social sciences: a way for practitioners in different subdisciplines of digital humanities to learn about work in related fields that can be applied to their own research.

As the journal grows in size and as its topical span also grows, the problem of discoverability becomes significant. While word searches offer one way to identify relevant articles on specific topics, they're not ideal since articles often reference terms which are not central to the article's domain. The journal has been working over the past few years to develop an extensible set of topic keywords that can be included in each article's metadata, and used for retrieval and discovery. This work began with a student in the Public History MA program who undertook an internship with DHQ in which she analysed and harmonized various sets of keywords created by ADHO and ACH in support of their conference proceedings, and applied them provisionally to the DHQ articles that had been published by that point. That list was expanded and reviewed by the DHQ editorial team in summer 2020. We now request seedling funding from the NULab to finish this process, by adding the keywords to the article metadata, identifying keywords for articles that have not yet been keyworded, expanding the DHQ publication workflow to include keyword identification (working with authors), and developing discovery mechanisms so that DHQ readers can use the keywords to navigate and discover articles of interest.

The term "seedling" may seem peculiar here given that we are requesting support for the final stage of this work. However, the entire keywording process is preparatory to a larger and more significant effort to make DHQ data more usable for analysis, and to make it available as a research data set for scholars who are interested in this facet of the history of digital humanities as a field. In particular, the journal is seeking to move from being simply "open access" (in the sense that all articles are published openly under a Creative Commons license) towards a more open-source and also open-lab model. In summer 2020, the journal moved its operations onto GitHub, as a preparatory step towards making data sets derived from the journal's content more readily available to researchers for experimentation. The journal is now looking ahead to involving its reader community in developing visualizations of those data sets to help readers navigate and study the entire DHQ corpus, and towards ways of enhancing the DHQ data to create opportunities for more interesting and useful visualizations. Completing this keywording task is thus a preliminary step along that path.

## Specific goals and tasks

Under this seedling grant, DHQ would hire an MA student in early summer 2021 to accomplish the following specific tasks:

- Work with the DHQ editors to finalize the current keywords list
- Identify keywords for the recently published portion of the DHQ article corpus (i.e. those published after our previous keywording effort)
- Working with the DHQ technical editors, populate the keywords into the DHQ article metadata
- Working with the DHQ editors, develop a repeatable workflow for exporting a data set that can be shared from the DHQ GitHub repository
- Working with the DHQ editors, develop some prototype visualizations of the DHQ article corpus, using the keyword metadata; these visualizations would be published in a DHQ editorial, and would serve as a starting point for developing an eventual live visualization to be added to the DHQ interface.
- Working with the DHQ editors, develop and document the ongoing workflow of adding keywords to incoming DHQ articles as they are published.

We request \$5000 to be spent as follows: 1 or 2 MA students at \$24/hour plus benefits: 193 hours of work, approximately 20 person-hours per week for approximately 9-10 weeks. The question of whether to divide this project between two students at about 10 hours/week or hire a single student at 20 hours/week will depend on the applicant pool. This work would be planned for May and June 2021. The rough task breakdown is as follows:

<b>Task</b>	<b>Hours</b>
Orientation and training	10
Review and finalize current keywords list	20
Assign keywords to recently published articles	50
Populate the keywords into the DHQ article metadata and review	10
Develop data export workflow	30
Develop prototype visualizations	60
Develop and document ongoing workflow	13
<b>Total</b>	<b>193 hours</b>