Introducing the Anthology for Computers and the Humanities

Taylor Arnold[1] ⓘ, Maria Antoniak[2] ⓘ, Miguel Escobar Varela[3] ⓘ, Marie Puren[4] ⓘ, Mila Oiva[5] ⓘ, Amanda Regan[6] ⓘ, Lauren Tilton[7] ⓘ, and Melanie Walsh[8] ⓘ

[1] Data Science and Statistics, University of Richmond, U.S.A.
[2] Computer Science, University of Colorado Boulder, U.S.A.
[3] Faculty of Arts and Social Sciences, National University of Singapore
[4] Laboratoire de Recherche de l'EPITA, Paris, France
[5] History and Archaeology, University of Turku, Finland
[6] History and Geography, Clemson University, U.S.A.
[7] Rhetoric and Communication Studies, University of Richmond, U.S.A.
[8] Information School, University of Washington, U.S.A.

Abstract

We present the *Anthology of Computers and the Humanities*, a new publication designed to host open-access, peer-reviewed proceedings from digital humanities (broadly defined) conferences and workshops. The *Anthology* provides a home for conference proceedings, offering essential structures such as DOIs and searchability, while being explicitly attuned to the needs of research situated in humanities disciplines.

Keywords: digital humanities

1 Background

The Anthology for Computers and the Humanities builds on developments across academia. Papers presented at academic conferences and workshops are frequently made publicly available in the form of proceedings. Proceedings can be published as a special issue or section of an existing journal, as a printed book, or—as is increasingly the case in many fields—in special-purpose digital publication venues. Proceedings enable the rapid dissemination and citation of work that has been presented to and reviewed by experts within a field, both for those attending the event in person and for others who read the proceedings after the conference has concluded. Papers within proceedings are frequently peer-reviewed as part of the conference itself and are written in formats that can be easily adapted to archival forms, such as LaTeX converted to PDF or Markdown converted to PDF or HTML. Papers typically do not go through extensive third-party copyediting. These features make proceedings ideal for quickly disseminating research in an open platform.

The role of conference proceedings differs significantly across fields [5]. In many social sciences domains, proceedings are considered precursors to longer papers in journals, with proceedings treated as working drafts rather than independent publications [6]. Proceedings from mathematics and statistics conferences, when produced, are often treated as supplementary details that contain technical proofs and information that cannot be included in the presentations themselves. Within some subfields of philology and linguistics, proceedings are edited, published, and "counted" as equivalent to chapters in edited volumes. In computer science, proceedings have

taken a particularly prominent role in their respective fields [8]. Across almost every subfield of computer science, papers published in conferences have become the de facto form of publication over the past several decades.

Scholarly associations frequently host open platforms for conference proceedings, providing access to cutting-edge research before, during, or right after the conference. The Institute of Electrical and Electronics Engineers (IEEE), for example, hosts a Conference Proceedings platform that hosts proceedings dating back to 1936. The Association for Computing Machinery (ACM) also manages the International Conference Proceeding Series, which is reported to receive an average of 4.4 million page views and 1.75 million downloads per month. The AIP Conference Proceedings, published by the American Institute of Physics, and the Association for Computational Linguistics' ACL Anthology provide similar structures within their respective fields. Platforms provided by scholarly associations typically offer many of the same bibliometric tools provided by journals, such as DOIs, a top-level ISSN, and inclusion in research search aggregator tools [1]. A limitation of these platforms hosted by scholarly associations, however, is that their disciplinary focus may make them less suitable for research situated outside their core fields.

As an alternative to dedicated platforms for publishing proceedings, many conferences post PDF files of papers directly on their website or link to general-purpose archival platforms such as arXiv or Zenodo.[1] These options are most popular with fields that treat proceedings as grey literature, which are distinct from research published in journal articles and printed books. Publishing directly on a conference website offers an easy and inexpensive way of making conference proceedings available without the support of a scholarly association. However, this comes along with many downsides. The conference site may only continue to exist for a short time, making papers only temporarily available, and does not come along with the kinds of bibliometric tools that are important to making research findable, and are often essential for research to "count" as published research for the authors. While augmenting these direct links with general-purpose archival sites such as Zenodo or arXiv provides some level of long-term access and persistent identifiers, the fact that anyone can upload their materials to these sites without any form of peer review causes challenges for considering such work as formal research outputs. The Anthology for Computers and the Humanities builds on the opportunities afforded by models across academia to support peer-reviewed, open access research published on a timeline that aligns with computationally-infused research to maximize our scholarship's impact [2, 7].

2 State of Conference Proceedings in the Digital Humanities

Digital humanities (DH), a term we intend in its most inclusive and generous sense, has no broadly accepted stance on the role of conference proceedings within the field. This is not particularly surprising since DH brings together scholars and scholarship from across fields ranging from the humanities to computer science, which, as described above, have very different norms surrounding the status of conference proceedings as research. We are aware of at least four different models that DH-oriented conferences have used to disseminate conference proceedings materials online:

1. **None.** Many conferences publish only presentation titles and (possibly) abstracts on their websites, avoiding any creation of full proceedings. For example, this approach is also taken by the most recent installations of the ACH's annual conference.

2. **Self-Archiving**. Other conferences publish submissions presented at their conferences on their websites. This is the format used, for example, by recent annual ADHO conferences,

---

[1] In addition to proceedings platforms, it is also become common to publish computer science related articles on arXiv as a way of both circumventing the paywalls of some sites such as the IEEE conference proceedings, in order to get feedback before the final conference publication is published, and as a way of attempting to establish ownership over an idea ahead of other papers published at a similar time.

which creates a book of abstracts. The status of these as citable, peer-reviewed publications can be unclear.

3. **CS workshops**. Within computer science, many conferences allow proposals for co-located workshops. Papers presented at these workshops must adhere to the rules of the leading conference and can often be published through institutionally backed platforms used for the main conference proceedings. Recent examples include the LaTeCH-CLfL 2025 workshop, co-located with NAACL, which publishes proceedings to the ACL Anthology, and the VIS4DH 2025 workshop, co-located with IEEE VIS, which publishes proceedings to ….

4. **Dedicated Volumes in CS Proceedings**. The Computational Humanities Research (CHR) conference published complete conference proceedings in the CEUR Workshop Proceedings from 2020 through 2024.

5. **Special Issues.** Another format used for some conferences is to publish papers within a special issue of a journal. This seems less common in recent years because it requires a lot of work for everyone involved. Previous examples include the special issues of Digital Humanities Quarterly, which published papers from BeNeLux 2017, and the Göttingen Dialog in Digital Humanities in 2015 and 2016.

The examples above only include conference proceedings that directly publish papers as accepted to a given event. Many events subsequently invite presenters to revise their work for inclusion in related publications, such as a special issue of *Digital Scholarship in the Humanities* dedicated to an ADHO conference and issues of the *DH Benelux Journal*. These publications require additional writing and peer-review processes and therefore function more like regular journal articles than conference proceedings.

All the examples cited above are reasonable approaches in many situations, depending on the needs and goals of the conference organizers and the communities in which they are situated. For example, a workshop seeking to engage with practitioners in computer science may find co-locating with a computer science conference to be an ideal situation. Likewise, a community that encourages the rapid revision of papers for a special issue of a focused journal may decide that publishing conference proceedings is a hindrance to subsequent publications. Offering a self-archiving option with limited bibliographic support can, similarly, be a valuable compromise for making work accessible without limiting the impact of subsequent publications.

In many cases, however, the options listed above simply represent the best available alternatives among those currently accessible to the DH community. Additionally, even in cases where these options may appear promising, they may not be accessible to everyone. To be included as a computer science workshop requires a clear computer science focus and cannot be designed around a geographical community. The Computational Humanities Research team was recently informed that they would no longer be able to publish in the CEUR Workshop Proceedings. Finally, the Digital Humanities Quarterly no longer accepts proposals with a "purely event-based focus." There is a clear need for another option.

3 Goals and Structure of the Anthology

The *Anthology of Computers and the Humanities* provides the DH community with a platform for publishing conference and workshop proceedings. We are guided by a strong commitment to broad accessibility, both in terms of the content we publish and how it is accessed. The *Anthology* is accessible online with no sign-in required to access the research and no author contribution fees.

All events of various sizes and from any of the broadly defined areas of DH are welcome to publish proceedings. Every volume of the *Anthology* corresponds to the proceedings from a

specific event. The organizers of the event serve as, or select, the volume editors and will have complete editorial control over the content of the papers published in their volume. The volume editors are invited, but not required, to include an introductory article to serve as context for the volume. Any group can submit proposals for volumes in the anthology, provided that the papers will undergo a peer-review process and that the event has an open call for participation. We expect that all proposals for volumes that meet the stated guidelines will be accepted for inclusion in the *Anthology*.

We have designed the Anthology to maximize the findability and accessibility of all papers. Papers appearing in the *Anthology* will be published under a Creative Commons Attribution 4.0 International license (CC-BY 4.0), which permits all reuse as long as the authors are properly cited. Copyright for the paper remains with the authors. Through an institutional license from the non-profit organization Crossref, we will be able to offer dedicated DOIs for every paper and ensure inclusion in all major research aggregators [3, 4]. Additionally, we will support integration and linked data using ORCIDs and a top-level ISSN for the publication. These bibliographic tools enhance the findability of papers published in the Anthology and facilitate authors' ability to receive academic credit for their publications through their own national and institutional frameworks. Because we want papers to be easy to cite, ready-to-use citation fields are provided on every article page, along with a site-wide, comprehensive, downloadable BibTeX file containing citations to all published papers from the *Anthology*. Papers will be made available as downloadable PDFs for archival purposes and searchable HTML for findability.

Organizationally, the *Anthology* operates as an independently managed sub-entity under the Association for Computers and the Humanities (ACH). This structure follows the standard format of computer science proceedings published by scholarly associations, providing a platform to operate immediately within an existing non-profit organization. ACH has generously provided the funds and infrastructure that have allowed the *Anthology* to operate without charging any publication costs to authors or organizations. While the ACH is a US-based professional society, the *Anthology* encourages contributions from all geographic regions and welcomes papers and volumes in any language.

4   Looking Ahead

We hope that the *Anthology of Computers and the Humanities* makes an essential contribution to DH by providing all interested members of the community with access to an open publication platform, through which they can share their work. The roles of conference proceedings and even conferences themselves will, as always, continue to evolve. We are committed to the long-term preservation of published volumes in the *Anthology*. We are excited to continue adapting as we aim to help enable, acknowledge, and amplify the research being done within our community.

References

[1]   Anderson, Ashton, Jurafsky, Dan, and McFarland, Daniel. "Towards a Computational History of the ACL: 1980–2008." In: *Proceedings of the ACL-2012 Special Workshop on Rediscovering 50 Years of Discoveries*. 2012.

[2]   Bosman, Jeroen et al. "The OA diamond journals study. Part 1: Findings." Tech. rep. 2021.

[3]   Chandrakar, Rajesh. "Digital object identifier system: an overview." In: *The Electronic Library 24*., no. 4 (2006), pp. 445–452.

[4]   Hendricks, Ginny et al. "Crossref: The sustainable source of community-owned scholarly metadata." In: *Quantitative Science Studies 1*., no. 1 (2020), pp. 414–427.

[5] Kochetkov, Denis, Birukou, Aliaksandr, and Ermolayeva, Anna. "The importance of conference proceedings in research evaluation: a methodology for assessing conference impact." In: *International Conference on Distributed Computer and Communication Networks*. Springer. 2021, pp. 359–370.

[6] Lisée, Christian, Larivière, Vincent, and Archambault, Éric. "Conference proceedings as a source of scientific information: A bibliometric analysis." In: *Journal of the American Society for Information Science and Technology 59.*, no. 11 (2008), pp. 1776–1784.

[7] Suber, Peter. *Open Access*. Cambridge, MA: MIT Press, 2012, p. 256.

[8] Vrettas, George and Sanderson, Mark. "Conferences versus journals in computer science." In: *Journal of the Association for Information Science and Technology 66.*, no. 12 (2015), pp. 2674–2684.