

Introduction to the special issue on digital scholarship

Stephen Griffin

Published online: 24 February 2015
© Springer-Verlag Berlin Heidelberg 2015

Digital scholarship, or “cyberscholarship”—that based on data and computation—is radically reshaping knowledge discovery, creation, analysis, presentation and dissemination in many scholarly domains. Distinguishing features of digital scholarship are multi-stage workflows that often involve cross-disciplinary collaborations, use of a large variety of information objects from multiple sources, new research methodologies, innovative data analytics and multiple forms of presentation of research outcomes. The enabling environment for digital scholarship is a rapidly expanding global digital ecology composed of richly annotated datasets, open source tools and a growing appreciation of open access digital publication of text and data as a measure and driver of scholarly productivity.

This special issue contains papers that report on research related to the broad set of activities that enable digital scholarship. For digital scholarship to flourish, consideration must be given to the entire data lifecycle. The digital libraries community has laid the foundation for digital scholarship by developing information environments and resources and by exploring new interdisciplinary problem domains. As large volumes of “born digital” data are created and legacy collections are converted to digital form, new possibilities for scholarly work appear. To the degree that repositories can be interlinked at the data element level, interoperability and functionality is greatly increased. The result is knowledge infrastructures capable of supporting a broad spectrum of scholarly activities. Semantic web and linked open data activities are developing standards, protocols and best practices for achieving this.

Michael Lesk’s introduction to the Issue offers a rich historic perspective on digital libraries research and digital scholarship. He reminds us that the current information environment is a result of the efforts of many people from many kinds of organizations. Often they were focused on different goals. What is seen as essential is the communication of ideas and this we must continue to value and improve through new models of scholarly communication.

The special issue contains six papers that describe work that describe outstanding examples of contemporary digital scholarship. They are:

- 1 Digital field scholarship and the liberal arts: Results from a 2012 to 2013 sandbox, James Proctor (corresponding author)

In this paper, the results of a sandbox (or collective on-line experiment) involving Digital Field Scholarship (DFS) is discussed. DFS is defined as scholarship for which field-based research, concepts and methods are significant. The availability to easily incorporate spatial and temporal indexed data into DFS opens new opportunities for disciplinary and multidisciplinary work. It also greatly expands problem space of in a broad set of topical domains in the sciences digital humanities.

- 2 Exploring publication metadata graphs with the LOD-milla browser and editor, Andras Micsik, (corresponding author)

Linked open data (LOD) and semantic web are revolutionizing repository interoperability. By structuring and creating relationships at the data level, altogether new functionalities result, transforming browsing, search and retrieval across web sites. In this paper, generic functions for browsing LOD repositories are discussed as well as

S. Griffin (✉)
School of Information Sciences, Pittsburgh, USA
e-mail: sgriffin@pitt.edu

new search strategies and data requirements necessary to meet the needs of digital scholars.

- 3 Towards robust tags for scientific publications from natural language processing tools and Wikipedia, Michał Lopuszyński (corresponding author)

Text mining and natural language processing have long been central research areas in digital libraries research. Given the vast amount of internet accessible textual content, tools to enhance the means to extract value from stored textual information improves the ability for the content to be searched, organized, presented and used. This paper describes research in methods of tagging scientific publications with labels gleaned from Wikipedia and the ArXiv preprint collection. The resulting analysis yields insight into the effectiveness of the different approaches and suggests what utility the methods may have for digital scholarship.

- 4 VisInfo: A digital library system for time series research data based on exploratory search, Jurgen Bernard (corresponding author)

Increasingly, non-textual knowledge and data are the basis for much scholarly work. Images, audio, multimedia, computational models and a wide variety of other data assemblages have proven capabilities for advancing digital scholarship. This paper introduces “VisInfo”, a web-based exploratory search system for time series research data. VisInfo is the product of collaborative efforts by data providers, digital librarians and computer scientists. The result is a powerful suite of tools for scholars dealing with large data sets covering extensive time periods.

- 5 What lies beneath?: Knowledge infrastructures in the subseafloor biosphere and beyond, Peter T. Darch (corresponding author)

This paper reports the findings of qualitative case studies of data practices in four multidisciplinary scientific collaborations—two large collaborative projects and two small projects. It examines an increasingly important aspect of digital scholarship—data workflow and the artifacts that naturally result from working with very large

datasets. The context of the paper are changes in research practices as a result of the “data deluge” over the past decade what that portends for research and scholarship as viewed from multiple perspectives. Contemporary digital scholarship is in a constant state of flux as data becomes inextricably entwined in the social, cultural and practical dimensions of research practice. This paper touches on these issues and others of significant and immediate import.

- 6 Image restoration; Photonegative digitization; historical documents, George V. Landon

Central to digital scholarship in the humanities and social sciences are collections of printed manuscripts, photographic images, audio recording, handwritten documents and more. Conversion of these often requires preliminary restoration work to bring them to a condition where accurate digitization can be done. Photographs on acetate-based film causes special challenges because of rapid deterioration over time. Yet, collections of photographs are among the most important source of historical content for digital scholarship in many disciplines. This paper describes the complex technical dimensions of automated image restoration and shows promising outcomes for use in memory institutions that are holders of important historic collections.

The papers give a glimpse of the continuing movement toward data-centered approaches to inquiry that now have now become a staple of research and scholarship in almost every disciplinary domain. Digital scholarship will continue to proliferate as network-centric models of scientific communication become the norm and the future reporting of scholarly work accommodates dynamic and recombinant document models capable of fully reporting scholarly research across all workflow stages.

Much of the seminal work in developing the information environments and resources that support digital scholarship can be linked directly to digital libraries research—past and present. Digital libraries research will continue to inform and provide the essential information resources for scholars in the years to come.