

Hitoshi Kamada

Digital humanities

Roles for libraries?

How has electronic information changed scholarly research in the humanities? Academic librarians may first think of electronic journals and books they deal with daily and the digitization projects being done to preserve and archive print information. But besides creating a digital version of scholarly publications and primary resources, electronic environment is presenting new opportunities to analyzing information for humanities research with the help of computational media. The term *digital humanities* is being referred to more and more, as the crossroad of information technologies and traditional humanities research. In my short definition, it is the application of information technologies to analyzing humanities as well as many interdisciplinary subjects. How might this research approach impact libraries?

The 2009 Digital Humanities Conference's papers touched on many topics related to libraries, including metadata, digital archives, XML, and e-books.¹ These are familiar terms among librarians when talking about digitization, data curation, and digital preservation. I see significant crossovers here. Others recognize this as well. The Digital Humanities Summer Institute² held in Victoria, British Columbia, Canada, this past summer provided a forum for the arts, humanities, and library communities. Collaborations are happening on a larger scale, such as NEH's partnership with IMLS and other agencies on Digital Humanities Initiative.³

Application of information technologies to the study of humanities is spreading

across the world, even if it is not necessarily called digital humanities. In Japan, several universities have created new departments or research programs dedicated to this field. One of the researches there used a computer application to examine grammatical structure of a Japanese literary classic, *The Tale of Genji*, written by Murasaki Shikibu, to suggest that someone other than Shikibu played a part in writing the last part of this classic.

Computer assisted analysis allows researchers to gather and analyze a large amount of data, which would otherwise be too time-consuming by human work, and can help find unexpected relationships among variables in this data, which might have been overlooked by manual analysis. It allows researchers to look at data in different ways. Text, which is meant to be read from beginning to end, can be sliced and diced to be examined from any angle. And, the digital application is not just about text data. Other media such as visual objects and data mapped on Geographic Information Systems can be electronically analyzed, synthesized, and presented, or combined with other types of data, for many disciplines, such as art and archeology.

Text mining in the study of literature is one of the most popular computer applications for humanities research. Text

Hitoshi Kamada is associate librarian, research support services team, at the University of Arizona, e-mail: kamadah@u.library.arizona.edu

© 2010 Hitoshi Kamada

mining software used to be very expensive or had to be developed in-house. But as more affordable text mining software becomes available for individual users, not just for large scale enterprise, there will be more potential use for researchers in the humanities. With computer assistance, textual information is not just something that researchers read through but can become a subject of systematic data analysis. And, this application is not limited to the humanities. Text on Web sites, blogs, and other social network media can become the data for social or cultural studies research on current social phenomena. Matthew Hurst' blog *Data Mining: Text Mining, Visualization and Social Media*⁴ provides a lot of interesting examples of how to look at social media.

These kinds of research involve not just the knowledge of relevant computer applications but also often skills and knowledge in collecting and organizing data, in which librarians have unique training and background. What is the best possible way to extract relevant text information on the Internet and organize and convert it to the format suitable for an analysis? How are elusive conversations on social networks captured? What is the best approach to digitizing this classical document for a particular kind of analysis? What would be possible metadata to effectively classify or describe art objects for this research project?

In some types of digital humanities research, I see collecting, repurposing, and organizing data or information for computer-assisted analysis as an integral part of scholarly research. And my own research based on text mining showed me that this research does not have to be large-scale, institution-wide projects. Researchers may choose to engage in smaller digital projects designed to support their work without a large investment in resources. Large scale digital projects by libraries or universities are still important. But I would like to see how librarians can help researchers with

their digital research projects on an individual basis. Can librarians provide support for collecting and organizing data in a way suited to a particular research project and perhaps even help put the collected data into digital archives in a reusable format and possibly facilitate collaboration with other researchers?

Researchers, especially in the humanities, may still think it is important for their libraries to have as many books and journals as possible, and some of them may prefer the print format. But they are realizing that many of these materials are now available electronically or via ILL/document delivery. It is more critical for researchers to have access to primary resources for their research. Research libraries are realizing they have to play a pivotal role in access to primary resources if they want to stay relevant in academic research in the current information-rich environment.

With new research approaches, such as digital humanities, researchers dig in primary resources or data not only in print but in a variety of media. In the digital research environment, primary resources can be of various types and media, and researchers may want them in different formats repurposed for the analysis they want to do.

Maybe it is time to think about how libraries can provide or help customize electronic information or data in a more interactive way as primary source data for individual computer-assisted research projects, in addition to continuing our effort to digitizing resources in larger scale.

Notes

1. Digital Humanities 2009, www.mith2.umd.edu/dh09/.
2. Digital Humanities Summer Institute, www.dhsi.org/.
3. Bruce Cole, "The Humanities in the Digital Age" *First Monday* 13:8-4 (2008), firstmonday.org/.
4. Matthew Hurst, *Data Mining: Text Mining, Visualization, and Social Media*, datamining.typepad.com/data_mining/. 