Accessioning the Digital Humanities: Report from the 1st Archival Education and Research Institute

Sarah Buchanan <sarahab_at_ucla_dot_edu>, The Meadows School

Abstract

Within the field of archival science, recent attention has been paid to identifying scholastic practices that will ensure the development of innovative research as well as the preparedness of future archival educators. Information science, long the academic frame of archival programs in the U.S. and elsewhere, currently allows for considerable co-expansion with digital humanities innovations when we consider the possibility of digital libraries, digital archives, and web-based collections integrating a humanist approach to display and users' interactivity with cultural objects. This paper reviews a workshop dedicated to charting the relationship between digital humanities and archival scholarship, as well as the opportunities to refine curricular and theoretical development in these two disciplines. The concepts expressed here would facilitate the efforts of practitioners to critically examine issues of pedagogy, practical training, and disciplinary alignment under a goal of sustaining the extraordinary expansion of applied humanist theory exemplified in research ventures we have seen thus far.

The Archival Education and Research Institute held in July 2009 represented the inaugural event of a four-year project to develop strategies and examine current issues in archival science. Towards the study of records with enduring value, created in many contexts and preserved in various formats, this Institute explored varied issues shaping archival work and professional manuscript curation, and offered a unique forum for participants to share research ideas which extend into the purview of the digital humanities and cultural heritage preservation. Aspects of the Institute included curriculum and pedagogy development, methodology workshops, research presentations, and interactivity with the local community. The efforts of the Institute — part of "Building the Future of Archival Education and Research," an IMLS-funded initiative [AERI 2009] — are extended to a global academic community, incorporating diverse perspectives from several continents regarding the structure and nature of archival education in universities, and archivists' relationship to the information science infrastructure. A workshop, "Curriculum Development in Digital Humanities and Archival Studies" was held that focused on digital humanities scholarship, providing one of the first platforms for digital humanists and archival researchers to congregate in a non-virtual meeting space and collectively outline future directions for digital humanities research. [1]

Conveners of the workshop brought perspectives that highlighted recent advances in digital library and digital collection management. Stephen Davison, head of the Digital Library at the University of California, Los Angeles (UC Los Angeles), illustrated the changes wrought when traditional collection development shifts from physical accumulation of objects and documents, to digital projects. Archivists are familiar with the concept of appraisal, having been exposed to theories and practices regarding the evaluation of a document's "enduring value," viability to the institution's focus, and scholarly relevance. [2]

But how does a digital archivist appraise a collection which consists of potentially thousands of digital objects, and produce a navigable, attractive digital library which is accessible to non-expert users? Davison suggested that digital archivists may be guided in their decisions by conducting an evaluation of the collection's projected "usefulness," so as to ensure that the archivist's efforts are relevant to the greatest number of users (though niche projects also serve the needs of specialized scholars, such as the Rossetti Archive at http://www.rossettiarchive.org/). Davison cited the [3]

success of projects within Classics, Medieval, and Tibetan studies — attributable in part to the clearance of these materials from copyright restrictions (still applicable to numerous 20th century collections), the depth of host institutions' collections, and recognition that researchers otherwise face the difficulties of travel to access remote materials in-person. Davison also spoke about developing metadata standards which serve to guide the project team throughout the initiative, providing uniformity and overall cohesion. In the case of the Encyclopedia of Egyptology (http://escholarship.org/uc/nelc_uee), for example, the migration of cuneiform is faculty-directed and peer-reviewed, and the production of digital surrogates is overseen by a technical team. Several other Digital Library projects have relied on guidelines denoting the type and repeatability of certain metadata fields and vocabularies [Riemer et al. 2009]. Collaboration remains a key component of digital initiatives for several reasons: 1) original source documents must continue to be preserved apart from digital surrogates, 2) digital archives which are "spun-off" from parent library or archival institutions should contribute added value to the institution, and 3) from digitization to server maintenance, funding and infrastructure partners provide necessary sustainability.

Chairing the workshop was Dr. Joshua Sternfeld, who drew attention to the academic interest in digital libraries with regard to faculty members seeking interactivity with primary documents. Sternfeld suggested that as archivists have moved their collections to the web (or at least established an online presence), they have not simultaneously innovated practices in this regard, often choosing to "mass-digitize" and transfer bulk material that is not described at the granular level. The work of archivists is no less needed once objects are made available in digital form, as archivists provide researchers with informed knowledge of "where things are" that is absent when the collection at hand consists of a glut of undescribed, unclassified files and formats. Sternfeld challenged the workshop participants to consider how the humanist's role in constructing digital archives might be to make explicit (via descriptions and interpretation) those scholarly motivations which are implicit to humanities scholarship. For example, how does one visualize doubt? Sternfeld cited several cases of innovative digital archive projects which have balanced contributions from scholars in the humanities (history and regional studies), archival science, and technology disciplines. The History Engine (http://historyengine.richmond.edu/) permits students to search, find, and contribute narrative "episodes" of historical events which are then given subject and geographic tags to aid in retrieval. Another site, "The World at the Fair: Experiences of the 1893 Columbian Exposition" (http://uclawce.ats.ucla.edu/) is the result of a term course dedicated to charting and organizing a model digital archive on a specific subject. While these sites each represent individual coherence and focused intention, Sternfeld suggested that general guidelines must be developed for digital humanities initiatives — perhaps by the workshop participants — regarding the determination of overall structure, display requirements, retrieval capacity, and future enhancements and/or additions to digital projects. The intellectual analysis that is performed by archivists when writing finding aids or other guides to their collections provides crucial context to researchers, alerting them to the scope and content of materials. These aids should accompany digital collections and provide a navigable point of entry to the database or website.

Dr. Johanna Drucker described the work of humanist scholars as it contrasts with that of some archival and library institutions' "imperial impulse" toward automation. Drucker stated that while humanists are interested in providing interpretations which hinge on subjective analysis, end-users of digital library collections may seek objectivity. An example of this tension is a deep, structured level of textual analysis afforded to individual digitized materials (on the one hand, a digital library) and large-scale patterns visible within a corpus of works, on the other. Recently, new tools have become available which make large-scale automation possible for a variety of purposes: such as those used within the TAPER Project of Tufts University (http://dca.tufts.edu/?pid=49) which is conducting an accession of digital records, and the SpecLab of the University of Virginia (http://www.speculativecomputing.org/), a series of initiatives within book arts, temporality, and visualization. While projects may alternately display deep structure within individual items or broad coverage as a group, it is essential that digital humanities initiatives empower the individual users in some way, encouraging scholarly communication by making available an intrinsically rich collection to an online audience [Drucker 2009b]. For humanists who consider interpretation to be a primary function of their work as academics, their role as partners in digital collection development is to specify which aspects of humane study are most relevant, abstruse, or otherwise unknown to the academic community, and provide theoretical guidance to technologist-partners in this regard. Just as many historical accounts of the same event prove that there are no "historical truths," the digital humanists of today and of the future have the opportunity at the present moment to draw the frameworks which

will come to define not only humanities scholarship, but digital collection development for many years to come.

The second half of the workshop was dedicated to discussion and group study of those skills and experiences which will best serve the development of the digital humanities field. Participants spoke provocatively about envisioning how the "archivist-humanist of the future" might perform in the context of traditional library institutions, archival repositories, and universities. In particular, workshop participants outlined training requirements, a pedagogical vision, and practical outcomes which would result from coursework in the digital humanities. Regarding pedagogy, participants noted that digital humanities research is oft content-driven, that is, motivated by scholars' expressed quest to interact with a particular cluster of humanities source material through some digital interface (i.e. literary texts in the case of digital epigraphy). Additionally, it was suggested that the digital humanities discipline can (only) define itself through the achievements and realized scope of individual projects (and therefore when seen as a whole greater than the sum of its parts, the discipline is borne). This recognition led speakers to emphasize that a pedagogical framework for the digital humanities would closely mirror that which is identified and problematized in individual humanities disciplines — and *then* realized in the digital environment (with technical partners). This entails a certain degree of creative visualization within particular disciplines, which we slated to occur well before the first web page is architected. Crafting the idea and scope of particular digital humanities initiatives was seen as eminently more valuable a pedagogical exercise than any amount of digital laying-out of content, as the latter is more easily mutable than the intellectual framework. For these group reports, participants were able to respond to and incorporate ideas generated from a bibliography of the digital humanities [Sternfeld et al. 2009], compiled by the conveners. Among twenty participants, three groups assembled within the workshop presented sets of both duties and skills which individuals saw as beneficial to digital humanities scholars. Some of the ideas proposed centered on "hard skills" such as research methodology and subject knowledge, and some ideas centered on "soft skills" such as management, communication (both within and beyond the project team), and empathy. In terms of project management, individuals asserted that the range of activities which must be performed in digital work include: selection of documents suitable for transformation, digitization, reformatting or standardizing, application of metadata standards and description, uploading of content to a stable server, promotion of the service, and maintenance and enhancement as feedback is received.

One group of participants identified three functional areas of responsibility for the "digital humanist" under which all required skills would be encompassed: management, practice, and public presentation. Management duties would include leadership and supervisory skills, and an understanding of end-user experience. Practical duties would include training the project team to perform technical work, securing the relevant local technology and daily operations, and supervising digital development and focus points. Public responsibilities would include proactive representation of the project to constituencies, outreach to community partners, a degree of networking and collaborative coordination, and periodic colloquy. Apart from the "required" skills outlined above, "preferred" skills identified by these participants included specific knowledge of established digital archival initiatives such as the Open Archival Information System (OAIS), an awareness of current trends in a rapidly progressing field, and some understanding of the human-computer interaction (HCI) discipline.

The relative youth of the digital humanities signifies that the opportunity to align archival practice with digital collections is nigh. Without some allowance of those archival specialties which include preservation considerations, selectivity and appraisal, description, and security (online), digital collections run the risks of being unable to sustain themselves and thrive in a digital world. Trained archivists bring an inherent set of graduate-level skills regarding the management and administration of collections which would provide crucial partnership in issues of long-term stewardship in the digital humanities. Additionally, the workshop participants discussed electronic finding aids as a companion tool for digital collections (specifically in the sense of whether a digital-humanist would need to have experience in composing archival finding aids — this was considered a benefit but not an essential one). In addition to discussion of the Encoded Archival Description (EAD) structured standard, participants agreed that the larger, and more looming issue is not those collections which have already been described in EAD finding aids, but the numerous collections which have *not* — and which comprise the "legacy data" backlog inherent to many repositories [Greene & Meissner 2005]. Within the archival field, archivists in the last five years have conspicuously recognized the scale of their hidden collections, and earnestly made accessible these materials by either a detailed or collection-level description. In preparing digital humanities

curricula, doctoral students and faculty participants shared the awareness that some alignment of the digital humanities with archival education will be necessary not only for reasons of shared research interests (i.e. that many skills useful to digital humanists are also valuable to archivists) but also for more obvious reasons of scalability, noting the small number of archival educators in general, and those with a humanities background. In striving for greater recognition of digital humanist endeavors, we researchers can align with the information science field, advocating our shared knowledge goals in a way that is more "Peaceable Kingdom" [Hicks 1834] than individualistically segmentary.

It is crucial that the digital humanities not only refine its extant disciplinary foci, but also begin to think generally and reflexively about its own sustainability, and that of its source data. As the digital humanities community continues growing in the direction of data collection and curation for born-digital (and not only paper-to-digital, or "digitized") materials, the field must begin to plan for regular surveys and monitoring of these valuable collections. Participants acknowledged that in many of the projects cited above, the first few years of the project have been devoted to collection-building, acquisition of objects, and interface design — but as projects enter their "second phase" of existence, now is perhaps the best opportunity to realistically chart: the quantity of materials, the level of metadata useful to the users, any digital preservation strategy that has been applied (e.g. migration or emulation), and integration of the digital collections into university scholarship.

Ongoing efforts in the digital humanities at the campus on which this conference was held, UC Los Angeles, included at the time of the workshop the Institute for Digital Research and Education, the Digital Humanities Reading Group, the Visualization Portal, the Experiential Technologies Center, the Center for Digital Humanities, the Keck Digital Cultural Mapping Program, and the revitalized Horn Press. The Institute for Digital Research and Education assembles faculty and researchers across the university in support of the shared development of computing initiatives. DH Reading Group discussions have thus far centered on "Pedagogical uses of real-time visual simulation: The World's Columbian Expo" by Dr. Lisa Snyder with the Urban Simulation Team (modeling off-line; static images at http://www.ust.ucla.edu/ustweb/Projects/columbian_expo.htm), and "HyperCities" with Professors Todd Presner and Phil Ethington (http://hypercities.com/). The Visualization Portal is an immersive cinema allowing researchers the capacity to display three-dimensional and historically-modeled datasets with gigabit connectivity. The Experiential Technologies Center supports the modeling of virtual reality digital models drawn from global sources (some ancient), as well as development of a graphical toolbox for sonic and visual applications. The Center for Digital Humanities provides technical support for a variety of faculty initiatives, ranging from digital library collections (e.g. "Orient North: Mapping Nordic Literary Culture": http://tango.bol.ucla.edu/orientnorth/tech.html) to departmental and academic web sites. The Keck Digital Cultural Mapping Program offers coursework in geographic and map-based technology. The Horn Press provides pedagogical learning opportunities in the form of letterpress and book arts experimentation, as well as hands-on opportunities to experience "bibliographic production" at its core. A research university that is home to both professional and academic graduate programs, this campus supports a great amount of interdisciplinary research, and the breadth of disciplines represented all but ensures that new connections will be made locally well into the future.

The challenge now is to broaden the practices that have been developed, and consider new institutional and independent partners who would benefit from collaboration with archivists and digital curators. In many areas outside of the information sciences — including urban planning, architecture, medical records, and commercial media and film — the modes of delivering information have expanded from once-purely paper-based and in-person methods to include online portals for retrieving and experiencing content. By expanding the scope of the digital humanities to include all aspects of dealing with digital objects — and preserving them for the long-term — this community has the opportunity not only to chart the direction of humanities scholarship, but also apply the access-oriented principles of the digital humanities and archival theory (as they have been thusly outlined) toward ensuring that our business and organizational colleagues develop project accountability and growth plans well into the future.

Works Cited

AERI 2009 Archival Education and Research Institute (AERI). *Building the Future of Archival Education and Research*. 2009. http://aeri.gseis.ucla.edu/.

Drucker 2009b Drucker, Johanna. "Blind Spots: Humanists must plan their digital future". *Chronicle of Higher Education*

55: 30 (2009). http://chronicle.com/free/v55/i30/30b00601.htm.

Greene & Meissner 2005 Greene, M.A., and D. Meissner. "More Product, Less Process: Revamping Traditional Archival Processing". *The American Archivist* 68: 2 (Fall/Winter 2005), pp. 208-263.

Hicks 1834 Hicks, E. *Peaceable Kingdom*. 1834.

Riemer et al. 2009 Riemer, J., et al. "Guidelines for Descriptive Metadata for the UCLA Digital Library Program". In *Digital Library Project Guidelines*. Los Angeles: 2009. http://unitproj.library.ucla.edu/cataloging/procedures/DRM_Metadata_guidelines_Version_1withUCLACoreMods.pdf.

Sternfeld et al. 2009 Sternfield, J. *Bibliography of Archives and the Digital Humanities*. *Zotero Group Library*. 2009. http://www.zotero.org/groups/archives_and_the_digital_humanities.