The Digital Humanities Summer Institute and Extra-Institutional Modes of Engagement

Voytek Bialkowski, Rebecca Niles, Alan Galey

Voytek Bialkowski is completing his second year of studies at the Faculty of Information, University of Toronto. His current research interests include interface design, reading media, ethnographic approaches to reading, and historical reading practices. Voytek participated in the DHSI 2010 workshop SEASR in Action: Data Analytics for Humanities Scholars.

Rebecca Niles is a first-year Master of Information student at the Faculty of Information, University of Toronto. She has an academic background in English literature, book history, and textual scholarship. Her current research interests include image-based digital representation of books as cultural and technological artefacts, as well as developing digital resources for the study of books by textual scholars and bibliographers. Rebecca participated in the DHSI 2010 workshop Digitization Fundamentals and their Application.

Alan Galey is Assistant Professor in the Faculty of Information, University of Toronto, where he also teaches in the collaborative graduate program in Book History and Print Culture. His research focuses on intersections between textual scholarship and digital technologies, especially in the context of theories of the archive and the history of scholarly editing. Information about his teaching and research may be found at http://individual.utoronto.ca/alangaley.

Abstract

This paper discusses the experiences of two participants in the Digital Humanities Summer Institute (DHSI), a week-long workshop and lecture series designed to provide an intensive environment for learning and collaboration in the field of digital humanities outside of regular institutional curricula. The DHSI presents an opportunity to consider how digital humanities scholars define themselves as a methodological and theoretical community, given that the field is composed of scholars from a diverse range of backgrounds who may disagree strongly on the nature of the relationship between humanities and information studies that defines digital humanities as a discipline. Two major paradigms emerge, one which sees digital humanities as the application of digital tools to humanistic topics, and another which privileges critical reflection on how digital modes of writing, reading, and scholarship impact our understanding of humanistic inquiry. This paper functions as both a reflection on the DHSI as a form of extra-curricular scholarly engagement, as well as an inquiry into broader theoretical and disciplinary problems within the digital humanities.

Introduction

Digital humanities is a confusing term - literally - in that it fuses together two ideas, implying a prior separation. This article explores the consequences of that perceived separation for the digital humanities curriculum, and takes the authors’ experiences with a recent Digital Humanities Summer Institute as an example of
curricular negotiation of a fundamental question: in order to become part of an emerging, diverse, and interdisciplinary field like digital humanities, how does one learn what one needs to know? From this seemingly straightforward question emerge others, with implications both institutional and epistemological. Is this knowledge best imparted through academic or professional degree programs, or through summer institutes and THATCamps, or through self-directed learning? What characterizes the pedagogical culture of digital humanities, and how does that culture structure the experiences of newcomers?

The perceived need to distinguish digital humanities from implicitly non-digital humanities manifests itself in definitions of the field that rely upon metaphors of impact or intersection. For digital technology and the humanities to be intersecting implies that they were (and still are) travelling in different directions. Similarly, impact metaphors imply the foreignness of the impacting term, and the passiveness of the impacted. Rarely do we hear of the impact of the humanities on information technology; the determinism almost always runs in the opposite direction, with technology as the active principle and the humanities as the passive one. These metaphors are artefacts of the twentieth century, and embody notions of disciplinary division (and hierarchy) that emerged from the same postwar climate that produced C.P. Snow’s encomium to science, *The Two Cultures* (1959/2007), and W.H. Auden’s satire against the marginalizing of arts education, “Under Which Lyre” (1946/1998). An alternate way of construing the meaning of *digital humanities* might be to read the word *digital* as a reminder of a technological engagement within humanities scholarship which may have been forgotten over the past century, but which was never truly lost. The present generation of humanities students and faculty working to apply their digital knowledge - not merely commenting on computing from a distance - reminds us that humanists have been information technologists from the beginning.

The question facing these re-awakening humanists is how to acquire digital skills and genuine code-level literacy. This is a difficult question, since it means more than simply downloading and using a digital tool in one’s research. A perennial debate in the field concerns whether (and how) humanists should learn the programming skills that would take them beyond mere tool-use (Rockwell 2003, Cohen et al. 2008, Kirschenbaum 2009, Reside 2010, and Galey 2011). Examples of resources for self-directed learning of fundamental skills and principles include *The Programming Historian* (http://niche-canada.org/programming-historian), and TEI [Text Encoding Initiative] *By Example* (http://tbe.kantl.be/TBE/). As useful as these kinds of practical resources may be, they also remind us that the question cannot be reduced to a simple matter of knowledge transfer.

Much depends upon the answer to the question “how do we learn what we need to know?” This essay describes one of the field’s most successful responses, the Digital Humanities Summer Institute at the University of Victoria. A week-long event that has run every spring since 2004, the DHSI combines the best aspects of a skills workshop, international conference, and summer camp. Participants spend five days attending plenary lectures and pursuing their own projects in courses on topics such
as text encoding, digitization, multimedia design, text analysis, visualization, and theoretical and disciplinary perspectives. These courses are the core of the DHSI curriculum, offering students the opportunity to learn in small, collegial groups at the beginner, intermediate, and advanced levels - and indeed offering faculty the opportunity to be students again for a week. That levelling spirit is reinforced by other aspects of the Institute which bring the various courses together. At the beginning and end of each day, all DHSI participants attend plenary lectures by leading practitioners in the field, which brings all participants together in the same room to consider questions that all digital humanists face (such as the nature of the academic job market, or lessons to be learned from particular projects). In recent years the morning lectures have showcased short presentations by graduate students in the field, a symptom of how student-driven the field has become even during the seven years since the DHSI began.

For a field that emphasizes the value of formalizing one’s research questions through computational processes, the DHSI makes a surprising virtue of informality. Participants usually dress casually, stay in campus residences, and spend a great deal of time socializing together in the evenings. These rituals mark the experience as being continuous with one’s everyday scholarly and social life, rather than an exception from normality, as with most academic and professional conferences. This strategic informality pervades much of the field, and partly structures how knowledge is developed and imparted in the digital humanities. A parallel example is the emerging type of event known as a THATCamp (The Humanities and Technology Camp), an unconference begun in 2008 at George Mason University and repeated in many other places, including the University of Toronto in 2010. While THATCamps represent genuine crowdsourcing of conference themes, with sessions treated as emergent semi-improvised structures based on the declared interests of participants, the DHSI offers a slightly more planned and predictable experience in the form of a curriculum. The spread of both approaches to training digital humanists testifies to their relative merits, and the importance of community-based approaches to training digital humanists.

However, a curriculum in any form presupposes what a given community needs to know, and how they should learn it. THATCamps and summer institutes alike are emphatically welcoming, and strive to avoid the in-or-out cultures that have developed in more traditional fields. But treating digital humanities as a field constituted by whoever shows up does not dispel the difficult questions that necessitated these sorts of events in the first place. As Patrik Svensson (2009) puts it in his recent overview of the field, we cannot overlook the possibility that being a digital humanist means “being in between, having multiple identities, lacking a stable identity, and engaging richly but not unproblematically with other disciplines” (para. 60); this differs markedly from the relatively stable identities of many professional graduate programs. If the identity of a digital humanist is defined by the educational investments he or she decides to make - such as learning a new programming language, or reading deeply in the theoretical traditions - then which kinds of investments tend to have the greatest effects? To what extent should skills be
differentiated from knowledge, if at all? To employ another foundational concept in
digital humanities, that of productive constraint, the making of a curriculum forces a
community to make choices, to invest in some areas of knowledge and not others. An
event like the DHSI is therefore an intervention in a more complex history than it
might appear on the surface, and brings into question the relationship between
technology and the humanities, both past and future.

With these questions in mind, we offer a report on two of the 2010 DHSI
course offerings, in which two of this article’s authors participated (Bialkowski and
Niles, respectively; Galey is a past instructor and plenary speaker at the DHSI). Our
primary purpose is to provide a sampling of the digital humanities skills training that
tends to be offered outside of traditional university programs. However, the DHSI and
events like it are not simply venues for knowledge transfer; they also prompt critical
reflection on the nature of a field still in formation, both in relation to other knowledge
domains and to its own internal culture. In that spirit, we close with a discussion of
the reversibility of the impact metaphor discussed above: if digital humanities limits
itself to the application of digital tools to humanities problems, what becomes of the
humanistic impulse to problematize the tools and the epistemic structures they
assume?

SEASR in Action: Data Analytics for Humanities Scholars

Our first example is the DHSI intermediate workshop on the Software
Environment for the Advancement of Scholarly Research (SEASR), a powerful visual
programming environment developed by a team of researchers and programmers at
the University of Illinois at Urbana-Champaign. The SEASR workshop has been in the
DHSI’s intermediate-level program for the past two years. It is an intensive week-long
introduction to developing and using text analysis programs within a novice-friendly
programming environment. The course topics included specific attention to a range of
elements and functions of the software: its analytics capabilities, its “Community
Hub” (a repository for facilitating the sharing of scholarly work), its interoperability
with other programs, and development goals for future iterations. At its core, SEASR
aims to facilitate the development, sharing, and documentation of digital humanities
projects. Scholars are able to mobilize SEASR’s vast array of processing functions on
their own textual (or multimedia) data. SEASR currently works with text as well as
audio data, while image-handling functionality is expected in upcoming releases. The
primary focus of SEASR, however, is its application for textual projects. In the realm of
text analysis, SEASR has been adapted as back-end analytics software for the MONK
project (Metadata Offer New Knowledge), an environment for analyzing texts from
multiple collections. Likewise, SEASR has also been extended to communicate with
various other applications, such as the popular Zotero application.

Becoming acquainted with a software suite, especially one as powerful and
extensible as SEASR, is not an easy task. However, the format and general pedagogy
of the course was well-suited to the subject matter and to the task of learning a new
software environment. It was beneficial to work through the entire textual analysis
cycle firsthand (including the often laborious process of cleaning data by eliminating so-called textual noise such as misspellings, acronyms, and paratextual information). By working closely with texts in this way, researchers are able to gain not only a better understanding of the processes undertaken in the field, but also, and more importantly, valuable insight into the design decisions and motivations of SEASR’s design and research team.

Just as important as SEASR’s value as a tool for analysing a corpus of documents is SEASR’s existence as a historically situated piece of technology that encompasses multiple traditions in reading, scholarship, and textuality. From this vantage point, the extensibility of the SEASR infrastructure is among its most compelling features, and the ability to experiment with some of these extensible elements is an illuminating experience. One of SEASR’s features allows users to create mashups wherein multiple visualization elements can be displayed on the same output webpage. A textual analysis of a given work or corpus may yield multiple visualizations, such as entity networks, tag clouds, and date-line visualizations. SEASR allows the user to determine how these visualizations are to be juxtaposed, and in this sense allows for quite a bit of interpretative freedom. This sort of modularization of textual analysis components has gained significant ground in the digital humanities scholarship of the last several years. Both the Mandala browser (http://www.mandala.humviz.org) and the related ‘data-mining droplet’ interface (Ruecker, Sinclair, & Radzikowska, 2009) conceptualize the user’s intervention into the text as a modular activity, wherein droplets or anchors constitute multiple points of contact between scholar and text. This modular, mashup affordance of the software was not only particularly fun to experiment with, but also provided an important avenue for engaging with the conceptual layers of the software.

That engagement goes beyond tool-use to include, reciprocally, the study of digital tools as artefacts. Many scholars working within the area of digital humanities consider related tools and prototypes as being worthy of study in their own right as historically situated objects. As Johanna Drucker (2005) writes, “preservation of cultural heritage, as well as other patterns of access and use, will be carried out through the electronic instruments we are currently making” (para. 7). The University of Toronto’s T-Space is just one example of a digital preservation project that aims to make scholarly work accessible. Another preservation effort, albeit in a different realm, is presented by the website Home of the Underdogs (http://www.homeoftheunderdogs.net), which aims to preserve unsupported, incomplete, or otherwise abandoned video game programs. The preservation of unwanted and abandoned video games reveals significant correlates with digital humanities projects, where completing and subsequently disseminating projects presents inherent difficulties for scholars (see Kirschenbaum, 2009a). Here, SEASR provides a valuable space not only for collaboration, but also for the documentation and digital preservation of nascent tools (or even modular components) that might otherwise be discarded or forgotten. Collaboration and preservation then, in SEASR’s “Community Hub,” form a coherent and valuable collaborative record. In this way, the SEASR workshop presents us with
a paradoxical yet ubiquitous scenario in the digital humanities, in which one’s tools for analysis become, simultaneously, objects of analysis themselves.

Digitization Fundamentals and Their Application

A gentle introduction to studying digital humanities, and into the digital humanities community in general, was the beginner workshop group entitled “Digitization Fundamentals and Their Application.” The focus of this workshop was to develop a functional knowledge of different methods of acquiring, refining, processing, and utilizing information pertaining to artefacts, aural or visual, static or animated. The course outlined how to plan successful digitization projects, develop an organizational structure to manage large caches of data, select appropriate devices and formats for input, and create platforms for display and dissemination of output.

Each day was dedicated to a specific element of digitization - usually a medium, such as audio or video, but occasionally on a form of output, such as how to host digitization projects on the web. The mornings were generally spent acquiring the foundational knowledge needed to plan and implement a digitization project in that day’s medium, and in the afternoons participants were given free access to a wide range of equipment to help put the morning’s fundamentals into practice. This workshop allowed participants to practice digitization both in the lab and in the wild, as they were able to choose to work within one of the University of Victoria’s well-appointed computer labs or take equipment to a nearby site of their choice, such as the University of Victoria’s McPherson Library and its rare book room.

One of the prevalent themes in this workshop, at the DHSI, and in the discipline of digital humanities generally, is the concept of belonging in a highly interdisciplinary community where commonalities in interests, theoretical frameworks, and methodologies can sometimes be hard to find. Introductions during the first workshop session exposed the considerable variety of backgrounds and objectives among the participants; while some members of the Digitization Fundamentals group had backgrounds relating to digital humanities, such as the four M.A. students from the University of Alberta’s Humanities Computing program, other participants were library professionals interested in improving their knowledge of digital resources and developing their digitization tool-kits. Several members of the group worked in or studied the liberal arts and were interested in learning how to create a digital platform for displaying and archiving the creative works they were involved with, while others focused on the personal and professional applications of the lessons of the workshop.

It seemed that the question of community affiliation was a common preoccupation at the DHSI 2010. One of the institute’s plenary speakers, John Unsworth (2010), dean of the iSchool at Illinois and Director of the Illinois Informatics Institute at the University of Illinois at Urbana-Champaign, delivered a talk about the state of digital humanities as an academic field as he had been observing it in the professional continuum that spans from the campus to the Twitterverse. In his lecture,
Unsworth discussed the sorts of divisions, both internal and external, which those who identify as digital humanists come across, as well as the ways in which scholars and researchers define themselves as digital humanists in distinction to others who claim that title. Unsworth registers, for instance, the tension between those who consider the use of digital tools to undertake research in otherwise conventional humanities subjects to be the mark of the digital humanist, and those who scorn this claim and insist that the truly innovative digital humanists are those whose theoretical approach to the humanities is steeped in the informational paradigm of the digital age.

This question of identity obviously affects not only the researcher’s personal identity, but also how she or he is recognized professionally, such as by funding bodies and hiring departments. Witness the recent posting of the Mellon Postdoctoral Fellowship in Digital Humanities, which is open to applicants whose research focuses on “the innovative and productive application of digital tools or resources to research questions in any subject under the Humanities Division” (Mellon Postdoctoral Fellowship in Digital Humanities, 2011). This definition favours a methodological perspective on what makes a digital humanities scholar. The major fault line, it appears, falls along whether scholars can or should be categorized by the methods they use or the theories they espouse.

Applying this contention to the Digitization Fundamentals workshop, the group of scholars and professionals participating could certainly be considered a methodological community, in that common knowledge of a range of tools and their potential applications allowed for communication on equal terrain, the sharing of ideas, and collaboration on projects, all resulting in a more sophisticated understanding of each individual’s particular areas of interest. On the other hand, this team would have been hard-pressed to find common theoretical ground that united them as digital humanities scholars, let alone discover enough commonalities in the collection of theoretical approaches to act as the foundation of a cohesive community.

This heterogeneity of theoretical approaches does not necessarily constitute a risk to the ability to collaborate in a hands-on workshop; in fact, the overall workshop experience is a prime example of what Tom Scheinfeldt (2010), of whom John Unsworth makes mention in his plenary lecture, refers to in his blog entry title as “[w]hy Digital Humanities is ‘nice.’” Scheinfeldt postulates that the congenial atmosphere often experienced in the field of digital humanities is due to the fact that digital humanists tend to collaborate on solving methodological rather than theoretical concerns. Problems of method can be settled through a combination of rational planning and experimentation, while issues of theory are endlessly debated and challenged because they are what Horst Rittel and Melvin Webber (1973) call “wicked problems” (p. 160), in which there exists neither one definite solution, nor one right approach for coming to a solution.

However, while common methodology may make for a productive workshop environment, the lively online debate about theory that Unsworth describes, as well as the discussions about theory that occurred both online and in person among DHSI
participants in response to Unsworth’s talk prove that the basis for community-building can be found in theoretical digital humanities as well. In fact, not only Unsworth’s lecture, but indeed all the plenary lectures as well as the graduate student colloquia that were presented during the week-long institute succeeded in engaging the attendees, as a community, on theoretical issues.

**Conclusion**

The DHSI at the University of Victoria is an excellent model for an academic community event that combines a highly immersive week-long program and a relaxed attitude to interaction and engagement, resulting in a unique opportunity for learning and collaboration outside of the typical academic setting. This approach diminishes the barriers typically encountered in an academic environment, including those between students and instructors, and between those who practice within academic institutions and those who practice in other sectors. The opportunity to collaborate outside of these conventional boundaries allows participants to recognize commonalities, and to develop a holistic understanding of the field in general.

Frequently, shared methodological priorities among the different workshops revealed themselves. For instance, both the Digitization Fundamentals and Their Application and the SEASR in Action workshops advocate developing and maintaining best practices in data collection, storage and preservation, as well as responsible project management, so that the fruits of the digital humanist’s labour can be shared with the wider community as well as with future researchers. One of the DHSI’s plenary speakers, Stéfan Sinclair, touched on issues of preservation and documentation of digital humanities projects in his lecture. He echoed Drucker in saying that the digital tools being created now are going to be the best artefacts available for understanding the field of digital humanities, and the field of humanistic inquiry more generally, in the future (Sinclair, 2010).

This tools-centric approach is not uncommon in a field that is often defined by way of the individual projects that constitute it. Therefore, understanding the logistical demands of digital preservation and dissemination is crucial for scholars engaged in various digital humanities projects. Digitization Fundamentals and their Application, as an introductory course at DHSI geared towards not only neophyte digital humanists but also librarians, archivists, and members from the arts and culture sectors, establishes a foundation for how to collect, store and disseminate the materials which will form the building blocks of future digital humanities initiatives.

Another demand of digital humanities research is interoperability - the ability of variously-created tools to work in conjunction with one another. In the intermediate workshop, SEASR in Action, SEASR’s architects position their software suite as a possible solution to this issue of interoperability, suggesting that “SEASR eases scholars’ access to digital research materials now stored in a variety of incompatible formats” (SEASR, n.d.). However, this represents only a partial solution to the increasing problem posed by emergent and divergent file formats, databases, and other incompatible pieces of code and infrastructure. Rather, the use of applications
such as SEASR should be coupled with persistent attention to the formats and technical specifications that make up the landscape of the digital humanities. As a community-based initiative, DHSI 2010 models the interdependence of the many layers of digital humanities research; persistent emphasis, in both the beginner and intermediate workshops, on standards, responsible data collection and management, and interoperability represents not only the technical priorities common to digital humanities initiatives, but also the critical importance of collaboration as the basis for creating robust projects that will maintain their relevance over time.

The structure of the DHSI’s week-long curriculum presents a model for dealing with issues of cooperation and community in an overtly interdisciplinary field. The institute’s balance of method and theory through intermingling workshops and lecture presentations tactfully handles the heterogeneous reality of emergent disciplines, which, as Twyla Gibson (2009) notes, tend to arise initially from the coming together of individuals from various fields, which very commonly have different theoretical and methodological approaches (pp. 109-110). The DHSI provides not only the opportunity to share methods and develop a common set of tools and vocabulary, but also a forum for sophisticated the theoretical groundwork for digital humanities as a discipline through lectures, colloquia, and the resulting discursive community that emerges.

Ultimately, the DHSI, as a major community event within the field of digital humanities, offers a dual, or perhaps heterogeneous approach to the definition of digital humanities as a scholarly discipline. Activities in the DHSI workshops can very well satisfy the Mellon definition of digital humanist research as the application of innovative digital resources to issues of humanistic study. At the same time, the content of the workshops themselves also provide an opportunity to reflect on the products of digital humanities scholarship as historically situated technologies that inform our understanding of traditions in reading, scholarship, and textuality. This duality or heterogeneity is reflected in Cathy Davidson’s dichotomy between two distinct phases of digital humanities approaches, which she terms “humanities 1.0” and “humanities 2.0” (2008, pp. 711-12). As she writes, “Humanities 2.0 is distinguished from monumental, first-generation, data-based projects not just by its interactivity but also by openness about participation grounded in a different set of theoretical premises, which decenter knowledge and authority” (Davidson, 2008, pp. 711-12).

In many ways the DHSI speaks to each of these distinct epistemic approaches to the field, by supporting both the “monumental, first-generation, data-based projects” described by Davidson, while simultaneously decentering knowledge and skills acquisition outside of traditional, institutional contexts. Therefore, the DHSI, as a pedagogical model, propagates a definition of digital humanities that is inclusive of various methodological approaches using digital resources to pursue topics in the humanities, as well as theoretically-informed reflection on the ways in which perspectives on humanistic topics are shaped and altered by digital practices. Through open collaboration and discussion among both emerging and established scholars, the
DHSI promotes a re-imagining of the field wherein new perspectives on digital tools and humanities problems constantly shift the boundaries of digital humanities as a discipline.

This research was supported by funding from the Social Sciences and Humanities Research Council of Canada.

References


