DHSI Colloquium 2016
Book of Abstracts

dhsicolloquium.org

Please note that scheduling is subject to change

Colloquium Chairs

James O'Sullivan @jamescosullivan
University of Sheffield

Mary Galvin @galvinmary
University College Cork

Colloquium Assistant

Lindsey Seatter @lindseyseatter
University of Victoria

Colloquium Session 1, June 7 4:15-5:30pm

Working at the Intersection of Renaissance Studies and DH: an Update on Iter Initiatives

Randa El Khatib (University of Victoria)
Lindsey Seatter (University of Victoria)
Bill Bowen (University of Toronto, Scarborough)
Daniel Powell (King's College London)
Ray Siemens (University of Victoria)

The objective of this presentation is to familiarize, or refamiliarize, DHSI participants with recent developments of Iter Community (itercommunity.org) and to provide a summary of the advancements of Renaissance Knowledge Network (ReKN; rekn.itercommunity.org), an Iter Pilot Project.

The Iter Community is a platform embedded in Iter (itergateway.org; a web-based study and teaching resource for the Middle Ages and Renaissance) that facilitates scholarly communication, interaction, and engagement. ReKN is a major initiative, associated with the Advanced Research Consortium (ARC), which follows the lead of NINES and 18Connect by aiming to create an integrated online research environment tailored to the needs of scholars studying the Renaissance.

Working at the intersection of the Renaissance and digital humanities, these projects champion and foster social knowledge creation as an integral part of modern humanities scholarship. One manifestation of this set of values is the creation of online spaces where scholars can connect, create, and collaborate on projects -- fundamentally disrupting the traditional image of the isolated humanities scholar tinkering away in a siloed discipline.

Our presentation will construct a virtual and visual tour of these interconnected initiatives. We will begin with Iter Community, which, since 2014, has been undergoing a dramatic redesign. We will explore the Iter Commons and foreground the importance of this initiative as a space that connects scholars and facilitates a productive, online research community.

The presentation will culminate with an in-depth discussion of one of Iter's affiliated projects: ReKN. After outlining the vision of ReKN, we will turn to our attention to recapping the completed efforts of the project's first phase: an annotated bibliography (completed by Randa El Khatib and Lindsey Seatter) and environmental scan of the intersections between Renaissance studies and digital humanities (led by Daniel Powell with William R. Bowen and Raymond G. Siemens). In order to demonstrate the breadth and depth of these resources, we will showcase a variety of tools, online projects, and foundational publications that feature or address the following topics: large-scale databases, online artifacts, digital resources, spatial humanities projects, and scholarly events or organizations. To conclude, we will discuss the future directions and goals of the ReKN project as it enters its second phase.

Biographical Notes

Randa El Khatib (U Victoria; khatib@uvic.ca): Randa is a first-year English doctoral student at the University of Victoria. Her research interests include early modern literature, spatial humanities, and social knowledge creation. She is also on the organizing team of the DHSI unconference, and a Research Assistant at the Electronic Textual Culture Lab where she works on the Open Social Scholarship and ReKN projects.

Lindsey Seatter (U Victoria; lseatter@uvic.ca): Lindsey is a second-year, English doctoral student at the University of Victoria. Her dissertation research is focused on exploring the narrative patterns exhibited across Jane Austen's print and manuscript novels through distant and digital techniques. She is also a Research Assistant in the Electronic Textual Culture Lab where she works on the ReKN project, the Open Social Scholarship initiative, and as the DHSI 2016 Colloquium Assistant.

William R. Bowen (U Toronto Scarborough; william.bowen@utoronto.ca): William is an Associate Professor and Chair of the Department of Arts, Culture and Media. His research interests lie in early speculative musical thought and in the facilitation of research and teaching through the development of digital resources and platforms. He is the founding director of Iter: Gateway to the Middle Ages and Renaissance, editor of FICINO and Renaissance and Reformation / Renaissance et Réforme and co-editor with Raymond G. Siemens of the book series New Technologies in Medieval and Renaissance Studies

Daniel Powell (King's College London; daniel.j.powell@kcl.ac.uk) Daniel Powell is a Marie Skłowdowska-Curie Fellow in the Digital Scholarly Editions Initial Training (DiXiT) Network, an Action funded by the European Commission's 7th Framework Programme for Research and Technological Development (FP7). He is based in the Department of Digital Humanities at King's College London and affiliated with the Electronic Textual Cultures Lab and Department of English at the University of Victoria, with research interests in the digital humanities, social knowledge creation, scholarly communications, media archaeology, graduate education in the humanities, cyberinfrastructure, and early modern culture. His work has appeared in Digital Studies / Le champ numérique, Renaissance and Reformation / Renaissance et Reforme, Scholarly and Research Communication, and Religion and Literature, as well as in volumes published by the Modern Language Association, NeDiMAH, and the International Journal of Learning and Media.

Ray Siemens (U Victoria; siemens@uvic.ca): Ray is a Distinguished Professor in the Faculty of Humanities at the University of Victoria, in English and Computer Science, and past Canada Research Chair in Humanities Computing (2004-15). He is founding editor of the electronic scholarly journal Early Modern Literary Studies, and his publications include, among others, Blackwell's Companion to Digital Humanities (with Schreibman and Unsworth), Blackwell's Companion to Digital Literary Studies (with Schreibman), A Social Edition of the Devonshire MS (MRTS/Iter, and Wikibooks), and Literary Studies in the Digital Age (MLA, with Price). He directs the Implementing New Knowledge Environments project, the Digital Humanities Summer Institute, and the Electronic Textual Cultures Lab, recently serving also as Vice President / Director of the Canadian Federation of the Humanities and Social Sciences for Research Dissemination, Chair of the MLA Committee on Scholarly Editions, and Chair of the

international Alliance of Digital Humanities Organisations.

The NYU Libraries Web Hosting Pilot: An Update and Lessons Learned

Zach Coble (New York University)

This paper will give an update on a year-long initiative at NYU Libraries to provide the NYU community with access to a robust and flexible web hosting service. This initiative, called the Web Hosting Pilot, uses Reclaim Hosting to provide users with web hosting accounts, which includes a cPanel, command line and FTP access, and options for customizable URLs as well as 1-click installations for Omeka, WordPress, and Scalar. The pilot is offered from January through December 2016, and will then be reviewed to determine whether to offer it as a proper, full scale service.

While other universities are providing similar services, this service is unique as it is institutionally housed within the library and offered by its digital scholarship department, and serves the entire global community (including NYU campuses in Abu Dhabi and Shanghai).

The paper will discuss
- How and why the service originated
- The planning stages, including a needs gap analysis (internal) and environmental scan (external)
- Preliminary data regarding usage, service, and technical support
- An overview of our concerns around scalability, policy, and intellectual property

One of the most important outcomes of the pilot is determining the level of staffing and support required for this service. The pilot is supported by two librarians, with limited support from two others. How can such a small group provide a service that provides technical customization and flexibility to the entire NYU global community? Who is responsible for helping clients resolve PHP errors? What about questions around how to integrate Omeka into a course? We will provide an update on our strategy to partner with NYU Central IT to provide basic (i.e. "tier 1") support.

More importantly, we will share the data we have gathered thus far. Coming at the half-way point of the pilot, this data will provide insight on the number of user sites, the types of technology used on those sites, and the level of technical and service support needed.

Biographical Note

Zach Coble is the Digital Scholarship Specialist at New York University Libraries, where I work in the Digital Scholarship Services (DSS) department. DSS provides consultation, project management, and other services for NYU faculty and students interested in incorporating digital humanities and publishing methods in their research and teaching. He also co-edits dh+lib and is a master's student in NYU's ITP program.

Exploring Place in the French of Italy

Heather Hill (Fordham University)

Beginning in the mid-thirteenth century and continuing through the end of the fifteenth century, many writers in the Italian peninsula created and copied prose and verse works in the French language. Approximately two hundred of these texts remain, forming the "French of Italy" corpus. Traditionally, scholars of the French of Italy have used close reading of individual texts to explore this corpus with the assumption of a geographic orientation toward France. New scholarship challenges this notion, arguing that these texts taken together focus toward the east. Rather than copying the French texts verbatim and maintaining a French perspective, Italian writers reclaimed French texts into the French of Italy corpus by changing the geographic orientation.

Fordham University's Center for Medieval Studies' digital project, "Exploring Place in the French of Italy" (EPFOI), uses geospatial visualizations to support the idea that French of Italy texts focus geographically on more than just France. These visualizations display identifiable place names from a sample of sixteen texts from the corpus and reveal trends that not only indicate an eastern gaze in literature but also variations in the authenticity of place names throughout different genres. The EPFOI project website displays these visualizations for scholars to consider for their own research.

In addition to the core mapping visualizations, the project website includes essays investigating what EPFOI does for the greater study of the French of Italy. It also contains what we call "micro essays," or short blurbs that describe trends we have seen in the visualizations and possible reasons why these trends occurred. Although they are not fully fleshed out, these micro essays offer scholars and students potential avenues for further research. The Center for Medieval Studies further invites scholars to use our data from EPFOI for their own research and offers it as a model for other digital medievalists interested in using mapping and text mining for their projects.

This paper will describe the "Exploring Place in the French of Italy" project, especially in terms of the process it took for Fordham's Center for Medieval Studies to complete it. It will also discuss how EPFOI might serve as a model for other projects and what it has done for both the study of the French of Italy and the community of digital medievalists. Finally, this paper will consider the future of the project itself and how it affects other projects we are currently doing at Fordham's Center for Medieval Studies.

Biographical Note

Heather Hill is currently finishing her second year of her master's program in Medieval Studies at Fordham University. She has contributed to several digital projects at Fordham's Center for Medieval Studies, including "Exploring Place in the French of Italy," "Oxford Outremer Map Project," and the upcoming "Independent Crusaders Mapping Project." Heather will be attending the Masters in Library Science program at Pratt Institute in the fall, pursuing a career as a digital humanities librarian.

Digital Representations of Petrarch's "Rerum Vulgarium Fragmenta"

Isabella Magni (Indiana University)

This proposal originates with my collaboration with Prof. Wayne Storey and Prof. John Walsh's Petrarchive (temporary link of the updated website available at: http://dcl.slis.indiana.edu/petrarchive/newindex.php - the updated website will soon be moved to: http://petrarchive.org), a digital edition based on a new "rich-text" approach to one of the icons of both Italian and Western Medieval literature: Petrarch's songbook Rerum Vulgarium Fragmenta. I will show how innovative uses of digital tools (from advanced uses of TEI-based encoding to deep maps) can improve our understanding of both the text and its material containers, the manuscripts. Digital tools allow more accurate visual representations of the Fragmenta's different layers, with its continuous erasures and palimpsests (example: charta 26r in which the digital interface allows for the first time to visualize the relationships among multiple texts in the Rvf and the palimpsest Donna mi uene spesso ne la mente, erased and overwritten by Or uedi amor by the poet himself). This new digital edition also offers users more authentic ways of displaying and reading Petrarch's attentive visual poetics, transcriptional design and material medieval forms. The project's editorial choices incorporate facsimile images of the main witness (manuscript Vaticano Latino 3195), along with diplomatic and edited transcriptions of the text, and a rich apparatus, including a glossary and three indexes. The Petrarchive project (new updated release in November 2015) encourages us to rethink how we read and interpret Petrarch's masterpiece.

In this presentation I will share both my personal experience in collaborating to the digital project (from the first steps into the world of encoding - 2013 - to my new role of investigator and lead encoder - current) and innovative technical features of the edition.
I would love to present the Petrarchive project either in the 10 minutes or in the 5 minutes Paper Presentations. I will attend the DHSI both weeks.

Biographical Note

Isabella Magni earned her Bachelor and Master's Degrees at the Università Cattolica del Sacro Cuore in Milan (Italy), with a focus on French and British literature, with a second Master's Degree in Italian and Medieval studies from Indiana University. She is currently ABD finishing her PhD in Italian with a focus on Petrarch, early Italian and Occitan literature and digital humanities. For the past three years she has served as the principal assistant on the Petrarchive.org project, collaborating with Wayne Storey and John Walsh in the production of the new rich-text digital edition of the Rerum Vulgarium Fragmenta (a project funded by IU's New Frontiers initiative (2013–14) and by the NEH in 2014–2017).

Voices of Southern Patagonia

Gustavo Navarro (Universidad Nacional de la Patagonia Austral)

Currently, both information and artistic output are carried out mainly in digital forms. The preservation of this material sets out new problems to keep it usable and available, since technologies for storing it are replaced by new generations that are more powerful and often

incompatible with their predecessors. These discussions about developments for the preservation of digital heritage have taken place in certain regions in the world, mostly English-speaking areas, for some years while developments in Latin America are still scarce.

My work lies in the elaboration of strategies aimed at preserving the sound digital heritage of the Patagonia Argentina region, taking into account the degree of urgency, local circumstances, available resources, and future predictions.

This work in collaboration with Municipal Archives of the province of Santa Cruz provides, in addition to knowledge obtained in the exchange, potential benefits from sustainability on the basis of inter-institutional agreements, and the setting up of shared standards, either to achieve interoperability between programs or to engender a common understanding of practices that help to achieve the desired objectives of preservation.

<u>Biographical Note</u>

Gustavo Navarro is the Director of project Archivo Memoria de la Patagonia Austral. Koluel. http://www.koluel.org. Phd Universidad de Las Palmas de Gran Canaria. Centro de Innovación para la Sociedad de la Información. CICEI and Professor Universidad Nacional de la Patagonia Austral. Unidad Académica San Julián. Santa Cruz. Argentina Labor of issues: Research in the areas of preserving cultural heritage and scientific digital objects, humanities informatics, digital humanism, sound

Colloquium Session 2, June 8 4:15-5:30

Transforming Bad OCR into Useful Metadata? Exploring NLP Possibilities

Evan Williamson (University of Idaho)

An ideal digitization project scans the best available witnesses at the highest possible resolution, enabling the creation of quality access derivatives. In the case of document materials, detailed access to the resource can be enhanced by full text transcripts generated by Optical Character Recognition (OCR). Unfortunately in the real world it doesn't always work out smoothly. Limitations in time, labor, or access to the originals might prompt a more pragmatic approach. One area where this often happens is the digitization of microfilm collections. Already one step removed from originals, and often worn from heavy use, microfilm is a less than ideal witness. However, some vendors provide fast and reasonable digitization service which may be expedient to provide digital access to otherwise inaccessible materials.

University of Idaho's student newspaper, The Argonaut, has been in print since 1889, a fascinating record of more than a hundred years of student history. UI Library Digital Initiatives received approximately 6,000 issues digitized from microfilm held by the Idaho State Historical Society. An index used to access print copies in Special Collections provides quality metadata for approximately half the issues. However, the remainder had only a date plus a "full text" field extracted from the OCR transcript.

This full text is particularly challenging as metadata. The low resolution and variable condition

of the microfilm, combined with the complex layouts and fonts of the newspaper, result in digital images difficult for OCR engines to parse. A large percentage of the resulting transcript is gibberish, full of extra white space, odd punctuation, and junk characters:

T I b VOLUME~II t .I r--——— = ':-'.I:::I ~!':: ~!!:II ' l~:,:--.i! 'v""j -4. 'g '...I r r,'="""''': UNIVERSITY- OF IDAHO, b]OSCOW)i JANUARY ""5. NO. -13 19'l C.TORIOUS TONET. T..CROOKS,"".',"".'"','""'.'"','"',,,'"".:,,-',,'.'"":,,"".,,';, RECEPTlOB BT BOISE,U. OFL have passed the age limit. next Wins All--Contests on Southern I'opula'r "ear. Graduate of'Class or"; .. I lclaho Trip. 19orI jjecejves miocles C jdat Work I' Fu 'o n us sm

Furthermore, the garbled transcript often exceeds the content management system's character limit for a metadata field. This situation left me frustrated and asking questions: Can any meaning of value be salvaged from this ugly transcript to provide better access to the materials?

Recent articles about using Natural Language Processing (NLP) to help automate the creation of access points from archival finding aids inspired me to explore methods to extract usable information. I approach the OCR transcripts as messy data that may be cleaned with further processing using NLP and machine learning techniques. First, I used Python NLTK to efficiently clean the OCR transcript creating metadata files compatible with the limitations of our content management system. Second, I compared options for extracting named entities and keywords using commercial APIs and NLTK. Finally, I tested the possibilities of machine learning classification using items in the existing index as training data. While it is clear that starting with the ideal of higher quality images (thus better OCR) and hand crafted metadata yields higher quality access points, these techniques highlight some efficient automated methods of providing better access to large and messy collections in the real world.

Biographical Note

Evan is a Digital Infrastructure Librarian at University of Idaho Library. He is also a recent graduate of iSchool@UBC with Dual Master of Archival Studies and Library and Information Studies with a background in Classical Studies and Art History.

Piloting Linked Open Data for Artists' Books at University of California, Irvine

Emily Mathews (University of California, Irvine)

Artists' books are a common component of many art libraries, and are of great interest to artists and art historians because of their highly visual, interactive and sculptural qualities. However, many of these art-like qualities remain underdescribed when only represented in the library catalog. UCI Libraries are working on a NEH grant to extend interoperability and discoverability of artists' books. We have identified processes of transforming legacy metadata from our Library catalog to linked open data while enhancing records with VRA Core elements. In addition to publishing linked open data with digital surrogates of artists' books in our special collections, we are building a prototype visualization tool which will allow researchers to traverse relationships within and between the works, discovering connections between artists, genres, techniques, and materials.

The end users of our artists' book pilot are first and foremost in our minds as we not only develop the pilot tool, but also strategize the implementation of the linked open data workflow. The successful model developed by UCLA's Center for Primary Research and Training (CPRT) serves as an inspiration for us: we are engaging with faculty and graduate students who use the artists' book collection in the curriculum to justify further directions in enhancing the rest of the collection using our pilot workflows.

VRA Core is an essential tool for describing works of art; however, further enhancements through advances in linked open data schema and visualization technologies takes our pilot to the next level of discoverability. Visual Resources Association (VRA) members have clearly articulated the value of linked open data to cataloging cultural objects. While much has been written, and many are working on tools to help streamline the data transformation, there is still a lot of uncertainty and no clear workflows, tools, or best practices identified for the community. Meanwhile, the issue of interoperability between libraries, archives, and museums is a huge issue in the cultural institution community. Our goal is to build an extensible work process others may readily adapt. In this article, we will describe behind-the-scenes challenges in creating a one-size-fits-all recommendation. We will also provide suggestions for meeting the difficulties and scaling the work to integrate other artists' books collections as "linked" collections.

Our pilot project will also be justified for further implementation through business case planning based on the Ithaka report Sustaining the Digital Humanities: Host Institution Support Beyond the Startup Phase, as well as the CLIR-DLF report Fit for Purpose: Developing Business Cases for New Services in Research Libraries. We will share the data gathering mechanisms and background research used to advocate for our continued work on this important collection.

Biographical Note

Emilee Mathews has been the Research Librarian for Visual Arts at University of California, Irvine since 2013. Before then, she served as the Interim head of the Fine Arts Library at Indiana University, Bloomington, where she also obtained her two master's degrees, in library science and art history. She is currently engaged in an NEH grant-funded project to pilot linked open data on UCI's collection of artists' books. She also researches and writes on convergences of digital humanities, new media art, and information science.

Four Words: The Role of Theory in Digital Humanities

Grant Glass (University of North Carolina at Chapel Hill)

We often hear that Digital Humanities should always root its work in traditional humanities methods. Yet, increasingly Digital Humanities has borrowed language from scientific and other traditions. This paper argues that this blurring of scientific and humanistic vocabularies is appropriate to the blended work of Digital Humanities. What we cannot see is the possibility that Digital Humanities is in itself a post-modern discourse technology. It contains the progression of Literary and Cultural theory. The influences of Quantum Theory are also spread throughout Digital Humanities. It is not apparent on face value, but the sematic implications of the language used within Digital Humanities allow us to see these loose connections and this paper unpacks

these connections to find how to more closely describe Digital Humanities.

Blurring the line between theorized and method, this paper proceeds by way of through both a close reading of key Humanities, Scientific, Sociological, and Digital Humanities text and by tracing the use of theory within Digital Humanities works through text mining. By using both a diachronic study, asking historical questions, and synchronic, asking semantic ones, this paper arrives at an understanding of the complexity in theorizing Digital Humanities.
Essentially this entire paper and project utilize the tools already at play within digital humanities scholarship, asking both historical questions and sematic questions about how theory is used and executed. These practices are not experimental, they are bread and butter and therefore demand a blended vocabulary. Can you use the tools of Digital Humanities to study Digital Humanities, or does it inhibit your ability to be critical in your reflection? The tools of DH have the potential to impose a particular perspective, but these same tools can reveal the potential for new modes of inquiry and theorization about Digital Humanities. It is a necessary research experience to implore the tools of Digital Humanities when looking at it, in order to find these hidden implications and discourses.

There are many works that have tirelessly looked at how to define the Digital Humanities, however only needed to look at the semantic and historical aspects of the discipline through this type of analysis to arrive at what makes Digital Humanities unique. It is a new discipline, emerging from both a historical Humanities perspective, but also a new scientific and sociological standpoint. Rather than moving its theoretical grounding onto another borrowed discourse, Digital Humanities found its home in its practice and became a truly unique discipline. Digital Humanities does not exist without practice, it needs to build, to create in order to theorize and argue. Marking up a text using TEI XML causes one to ask theoretical questions about what a text is and how one approaches that text. Instead of looking at other technologies or discourse for guidance, we only need the technology of Digital Humanities. It is so close to what we have seen in media studies: "post-modern discourse technologies such as hypertext afford us the possibility of creating new kinds of syntagmatic meanings by linking across the re-contextualized elements of traditional genres and forms" (Lemke 216). Digital Humanities is our new post-modern discourse technology.

Biographical Note

Grant Glass is a first year graduate student at the University of North Carolina Chapel Hill in the English and Comparative Literature Department. He studies British and European Romanticism and Literary Theory. Grant recently graduated with a degree in Digital Humanities from King's College London. His interests include textual analysis aided by digital technology and material culture of the book. He has been involved with a number of digital projects including electronic textual editing, digital critical editions, and the history of digital humanities. Currently he is researching the effects of new technology on innovation within the humanities.

Rethinking the Exhibition Catalogue: Documentation, Curation, and the Digital Humanities Project

Julia Polyck-O'Neill (Brock University)
Aleksandra Kaminska (Simon Fraser University)

Catalogues are important tools in the documentation of exhibitions, and have remained stable and relevant archival forms in the face of ephemeral art practices, from performance to digital and media art. The stability of its paper form has also meant that documentation has been flattened to fit the page. As evidenced by recent projects such as the Getty Foundation's Online Scholarly Catalogue Initiative, the conventions of the exhibition catalogue are being critically revised and reimagined. While the digital tools and methods of digital humanities and digital art history can augment the form of the catalogue, they also raise new questions, not only about the relationship between text and image, author and reader, but also about the very role and place of the catalogue in the methodologies of the curator and the art historian. Indeed, the conversion of the catalogue into online space increasingly blurs the lines between exhibition and documentation space, perhaps even making the catalogue obsolete in its traditional role as an object with archival value. This presentation considers the catalogue's transformation in digital space—from linear and flat to interactive and modular, archival to ephemeral, static to active— from future-oriented document, to part of the curatorial process and project itself. In considering this hybrid digital exhibition-documentation space and the way it is reshaping digital art history, what can we learn, moreover, from the archival practices of (new) media art and digital art history? What are the implications of distinguishing the online catalogue as a project of digital humanities vs one of digital art history? By looking at specific examples of online catalogues and the experiences of the authors producing catalogues in this changing landscape, this paper considers how the move to digital form is shifting the catalogue's relationship to curation, the nature and practices of documentation and archiving, and ultimately the modes of art (historical) research.

Biographical Notes

A PhD Candidate in Brock University's Interdisciplinary Humanities program, Julia Polyck-O'Neill is a curator, visual artist, research associate with Brock's Center for Digital Humanities (currently a Contributing Editor for Digital Scholarship Ontario), HASTAC Scholar, and fellow of the Editing Modernism in Canada project. Her SSHRC-supported research examines historic and contemporary conceptualisms in Vancouver visual arts and literature.

Aleksandra Kaminska is a Mitacs Elevate Postdoctoral Fellow at Simon Fraser University and a Research Associate at Sensorium, Centre for Digital Arts and Technology at York University. She is also the Managing Editor of the journal Public: Art Culture Ideas.

Digital Storytelling for Digital Humanities

John F. Barber (Washington State University Vancouver)

The humanities focus on telling stories about what it means to be human. Digital Humanities (DH) uses research and practice at the intersection of the disciplines of the humanities and computing technologies to influence creation, dissemination, preservation, research, and teaching activities associated with such stories. At first thought, combining computational technologies with storytelling techniques may sound a bit much. But digital storytelling may extend the ability of DH to engage academic research with creative practice and may promote new ways of

teaching and learning as well as creating, critiquing, and consuming humanities research and scholarship.

Biographical Note

John F. Barber teaches in the Creative Media & Digital Culture program at Washington State University Vancouver, Vancouver, Washington, USA. He is the developer and curator of Brautigan Bibliography and Archive (www.brautigan.net), an online, interactive information structure known as the preeminent resource on the life and writings of American author Richard Brautigan. Barber also developed and maintains Radio Nouspace (www.radionouspace.net), a curated listening gallery/virtual museum and practice-based research and creative practice space. Both are inspired by the radio medium with its emphasis on sound as the basis for narrative and storytelling. His radio+sound art work has been broadcast internationally, and featured in juried exhibitions in America, Canada, Germany, Portugal, and Macedonia.

Colloquium Session 3, June 9 4:15-5:30
Special Session: Gender and the Digital

A Textual Analysis of Female Renaissance Playwrights using R
Elizabeth Ramsay (Trent University)

Using Topic Modelling I hope to prove that female Renaissance playwrights such as Mary Sidney Herbert and Lady Mary Wroth used a more domestic lexicon in their writing than their male counterparts. They wrote using more private forms of writing, in general, than men: women wrote closet dramas and men wrote plays for the public stage. This is not of course a clear cut boundary as individual cases blur these boundaries, but it can be taken as an overall guideline to show how women negotiated the "male" field of published writing. A comparison of a corpus of female Renaissance playwrights' texts and males' writing can show the different topics they employ in their writing. Word clouds allow this to be seen visually and for similar topics to be found across copra and compared. This comparison will in turn show how women writers re-created writing as a domestic space in order to enter the field by highlighting their use of private language.

Once women were able to establish themselves as writers, their writing gained popularity. This can be seen by tracing the types of writing seen throughout the Renaissance period. Early in the period, female playwrights were publishing in manuscript form and sharing their writing only with friends and family, such as Queen Elizabeth's translations, but by the end of the period publishers are seeking them out to publish their work. An example of this is Margaret Cavendish when she published her plays in 1662 because she was asked to do so by her publisher. This change shows that women were successful in breaking into the public forum of the publishing industry as there was now a public demand for their writing. This study seeks to follow one way in which the private and public spheres of the Renaissance period became interwoven.

Biographical Note

Elizabeth just finished my M.A. in English Literature (Public Texts program) at Trent University this September. She is a Renaissance and Classics scholar with an interest in Book History. She

has a background in feminist criticism and am developing an interest in the Digital Humanities. Elizabeth's Master's thesis explored the place of Shakespeare's Classical Women (Cleopatra, Imogen, and Cressida) in Renaissance society.

The Comparison of Human-Reading and Machine-Reading in Le Système de la Nature

Maryam Mozafari (Simon Fraser University)

My thesis is about d'Holbach's philosophy of Materialism in Le Système de la Nature. The purpose of my thesis is to compare human-reading with machine-reading. To achieve this, I used two methods: As one method, I used a qualitative data analysis software program (NVivo 10), in order to do the thematic analysis of the main themes of this book. This software gave me the opportunity to thematically code my corpus and run different queries such as text search or matrix coding query. Text search query searches for words or phrases that we look for in the corpus. Matrix coding query can be used to find patterns in data, and in this query pairs of items are cross-tabulated. Thus, with this software, initially I ran text search query to identify each theme. Then, I ran the matrix coding query to do the comparative analysis of the related themes in order to examine the density of the appearance of each theme together. As the second method, I used R Mallet package for topic modeling which allowed me to compare the thematic structure of the corpus in terms of co-occurrence of the main themes. My results reveal that human-reading is different from machine-reading with regards to the main themes and their co-occurrence. In another words, in human reading, I considered four main themes and I compared the degree of appearance of Naturalism and Materialism with these themes. In this type of reading proximity of the words is considered. While, in Machine-Reading, topic models identify correlated themes regardless of their proximity in the text. In topic modeling, algorithms recognize co-occurrence of the words which are semantically related and this could build on human-reading. Additionally, I found that the usage of both software packages is different: NVivo 10 as a mixed-method of human-reading and machine-reading allowed me to examine my hypothesis about the main themes and their co-occurrence. My theory was that in Le Système de la Nature, Naturalism and Materialism appear more with the theme Religion and Moral, however quantitative analysis of the text with NVivo 10 showed that Moral appears the most. Furthermore, with the visualization and analytical tools of this software, I was able to visualize my findings. On the other hand, R-mallet package, as a method for machine-reading, helped me to bring into attention other main topics of this corpus of eighteenth century text in terms of examining thematic trends such as Homme and Bonheur.

Biographical Note

Maryam is a M.A. candidate in the department of French at Simon Fraser University.

Does Gender Affect How Genre-Conformingly Writers Write?

Sayan Bhattacharyya (University of Illinois, Urbana-Champaign)
Ted Underwood (University of Illinois, Urbana-Champaign)

Literary genre is a somewhat subjective category, but treating it as an empirical social phenomenon permits text-analysis-based exploration of testable claims concerning genre,

through the building of predictive (rather than explanatory) models [1]. Using such a model created through applied machine learning and based on page level genre-classification of English-language volumes from the HathiTrust Digital Library's collection [2, 3], we address the following question: Does author gender influence the extent to which authors conform to a genre's conventional expectations?

Each page of a sampling of 414 volumes from the eighteenth and nineteenth centuries was classified, on a page-by-page basis, by the genre categories of fiction, poetry, biograpy and drama by a pool of human readers. For each volume, the overall classification accuracy of the by-page genre classifications was computed (Footnote 1). We then assigned a gender category to each volume based on subjectively inferring the author's gender from the author's name (Footnote 2). Each volume was classified into one of three literary genre categories: drama, fiction and poetry (Footnote 3). Genre classification accuracy, overall, for male authors was very slightly higher than that for female authors. For fiction and drama, accuracy for female authors was slightly less than that for male authors, but, for poetry, was slightly higher than that for male authors.

Our analysis suggests that women writers wrote may have written more "genre-conforming" poetry, but less "genre-conforming" fiction and drama, than male writers. Women writers accommodating to gender-stereotyping expectations framing poetry as a more feminine genre is a possible explanation for this observation. Thus, our work is a proof-of-concept that computational analysis can be useful for evaluating claims concerning literary genre and gender, such as Stuart Curran's observation that "while the rise of the novel… sealed off [poetry]... as a male, upper-class fiefdom, [b]y the 1820s… poetry… [had become] a woman's genre" [4].

References

[1] Shmuéli, Galit. "To Explain or to Predict?" _Statistical Science_, 25(3): 289-310, 2010.

[2] Underwood, Ted, Shawn Ballard, Michael L Black & Boris Capitanu."Mapping Genre at the Page Level in English-Language Volumes from HathiTrust, 1700-1899". _Digital Humanities 2014_. Lausanne, July 2014.

[3] Underwood, Ted. _Understanding Genre in a Collection of a Million Volumes: Interim Report_. University of Illinois, Urbana-Champaign, December 29, 2014. https://figshare.com/articles/Understanding_Genre_in_a_Collection_of_a_Million_Volumes_Interim_Report/1281251

[4] Curran, Stuart. "Women Readers, Women Writers," in _The Cambridge Companion to British Romanticism_, ed. Stuart Curran, Cambridge, 1993, p. 182.

Footnote 1: This consisted of determining the proportion of the total number of words in the volume that was correctly classified by the predictive model, correct classification being considered as: the page having been classified by the predictive model to be of the same genre as indicated by the ground-truth, human-made classification of the page.

Footnote 2: We used the author's name when it was provided in the HathiTrust Digital Library's biographical metadata for the volume. A small number of volumes for which author gender could not be thus inferred, and those volumes which were collections of writings by various authors, were not used in the analysis.

Footnote 3: This classification was based on which category among these three had been assigned by the predictive model to the most number of pages within the volume.

Biographical Notes

Sayan Bhattacharyya is currently a Council on Library and Information Resources (CLIR) Postdoctoral Fellow at the HathiTrust Research Center, Graduate School of Library and Information Science, University of Illinois, Urbana-Champaign. He earned his PhD in Comparative Literature from the University of Michigan, Ann Arbor.

Ted Underwood holds the position of Professor, and Liberal Arts and Sciences (LAS) Centennial Scholar in English, at the Department of English, University of Illinois, Urbana-Champaign.

Poster Session, June 10 5:00-6:00

English 4300: A Testing Ground for DH Practicums

Greg Chan, James Hospedales, Steven H. Lee, and Jennifer Wiens (Kwantlen Polytechnic University)

This presentation showcases the works-in-progress of students enrolled in KPU's English 4300 - Writing and Persuasion Beyond the Classroom: Apprenticeship in the Digital Humanities, a service learning course where the digital (physical computing) meets the humanistic (rhetorical theory). As part of this summer's DH iteration of the course, each student is in the process of completing a 12-hour, community-based practicum in which he or she is authoring content and assisting in the development of a digital object using open-source software. These DH projects include the design of a website and social media presence for the KPU Criminology department's Social Justice Centre; the editorial support of Mise-en-scène: The Journal of Film and Visual Narration, an open-access film studies journal sponsored by KPU's English department premiering this fall; the production of a promotional video for KDocs, Kwantlen's official documentary film festival; and the management and marketing of the KDocs Outreach Program in its visits this summer to Guildford Park Secondary, Pathways Aboriginal Centre, and Vancouver's City Hall. Each practicum is coordinated by English 4300 instructor Greg Chan, in partnership with project supervisors Janice Morris (KDocs Outreach), Manon Boivin (KDocs video), and Michael Ma (Social Justice Centre).

The English 4300 poster aims to illustrate the value of the following:
• Civic engagement. As participating students are working in partnership with community groups throughout the semester, they are gaining experience with non-profits, social justice issues, and volunteerism.
• Experiential learning. Rhetorical theory within an academic context only can be restrictive;

however, when students are able to apply in to digital computing activities with practical outcomes, they get the best of both worlds.

•   Mentorship/Apprenticeship. Students in the cohort are working closely with their instructor, their peers, and their practicum supervisors – relationships that depend on collaborative learning and mutual support.

•   Field study. The students representing the English 4300 cohort applied for a student-led research grant to fund their participation in the DHSI colloquium, which means they have demonstrated the course teachings on persuasive writing to present at their first academic conference. They will be live Tweeting and Storify-ing their field study experience at the #DHSI2016 to document their growth as digital humanists.

•   Liberal Arts degree. English 4300 is an atypical course offering, as it is the only service learning/digital humanities course offered by KPU's Faculty of Arts at this point. As an experimental iteration, it is showcasing how a liberal arts course—or even an entire program—can be both theoretical and practical. The English department aims to develop a digital humanities stream of undergraduate courses as a follow up to this experimental run.

Biographical Note

Greg Chan is a faculty member in the Department of English at Kwantlen Polytechnic University, where he teaches composition, literature, digital humanities, and film studies courses. He is the Editor-in-chief of Mise-en-scène: The Journal of Film & Visual Narration, an open-access film studies journal published by Simon Fraser University, and a board member/community outreach coordinator for KDocs, Kwantlen's official documentary film festival.

The HathiTrust Research Center: Supporting Large-Scale Analysis of the HathiTrust Digital Library

Sayan Bhattacharyya (University of Illinois)
Nicholae Cline (Indiana University)
Eleanor Dickson (University of Illinois)
Leanne Mobley (Indiana University)

The HathiTrust Research Center (HTRC) is the research arm of the HathiTrust. It facilitates scholarly research using the large-scale HathiTrust Digital Library (HTDL) by providing mechanisms for researchers to access HathiTrust content and study it using computational tools for text analysis. These mechanisms include a suite of tools and services that create opportunities for non-consumptive textual analysis, which will enable researchers to conduct analyses despite copyright restrictions. HTRC development is guided by research initiatives, such at the HTRC Data Capsule and the User Requirements Study, that explore both cutting edge infrastructure as well as users' needs in order to find solutions for enabling large-scale text analysis. Additionally, the HTRC works directly with scholars through a program called Advanced Collaborative Support, by which awardees receive development time and resources to aid their text analysis research. The HTRC is further engaged in supporting computational text analysis through training initiatives, including a 3-year train-the-trainer program aimed at building digital humanities competencies in academic librarians. The HathiTrust Research Center is a shared project of the University of Illinois and Indiana University. The HathiTrust is a consortium of

over 100 partner institutions that contribute digitized library content to the HathiTrust Digital Library, which is nearly fourteen million items in size.

Biographical Notes

Sayan Bhattacharyya is the CLIR Postdoctoral Fellow at the HathiTrust Research Center in the Graduate School for Library and Information Science at the University of Illinois.

Nicholae Cline is the Digital Research Librarian at Indiana University.

Eleanor Dickson is the HathiTrust Digital Humanities Specialist at the University of Illinois.

Leanne Mobley is the Scholarly Technologies Librarian at Indiana University.

All are members of the joint Scholarly Commons group in the HathiTrust Research Center.

Digital Scholarship in the Institutional Repository

Jeri Wieringa (George Mason University)

Comprised of multiple layers of information — content shaped by the code that drives the interface; the underlying data and its structure; external resources pulled into the argument — digital scholarships presents conceptual and technical challenges to institutional repository systems optimized for the static, and stable, PDF. This poster will present the approach being investigated at the George Mason University Libraries for ingesting and preserving digital scholarly objects, particularly dissertations, in our DSpace-powered institutional repository. In this poster I will present the different strategies considered — including library hosting of digital projects, archiving via Archive-it and WebArchive files, hosting within the repository, archiving of component parts, and combinations thereof — the trade-offs associated with each, and the strategy we have chosen to pursue.

The primary component parts of digital scholarly projects that we have considered for archival purposes are: the data on which analysis and interface are based, the processes used to parse that data, the infrastructure used to present the work, and the experience of the work at the time of publication. While not all digital projects involve all of these layers, where they exist each is vital to the argument being made and must be taken into account for archiving and preservation. We have chosen not to pursue the transfer of digital scholarly projects to library servers as the primary strategy for preservation. Long term maintenance of websites as "working archives" may be desirable and even necessary for press publications or other official digital products; however, such an approach requires large investments of time and infrastructure. That level of investment is not feasible for every digital project, particularly considering the variety and number of projects that may be submitted. Instead, we have chosen to focus on archiving the component parts of digital dissertations and digital scholarship generally. This approach allows the intellectual work of the digital project to be documented and, if need be, reassembled and interrogated, but in such a way that is manageable for libraries with minimal infrastructure and technical support. In this poster, I will present our research in developing this approach and

lessons learned from early attempts to put it into practice. We hope to generate feedback and conversation regarding our approach.

Biographical Note

Jeri Wieringa is the digital publishing production lead with Mason Publishing, part of the George Mason University Libraries, and a Ph.D. candidate in history at George Mason.

Canada's Early Women Writers and all their relatives

Karyn Huenemann (Simon Fraser University)

One of the great outcomes of the digital age is the interconnectedness that has become possible across disciplines, not only within the Academy, but also in the world at large. As academics, we now have untapped resources to mine: not only texts and archives that are being digitized and digitally catalogued, but boxes in people's basements and attics and records in people's minds that can be shared with us through online conversations.
One of the mandates of the Canadian Writing Research Collaboratory (CWRC), run out of the University of Alberta and headed by Dr. Susan Brown, is to construct a repository of information that will be accessible and useable by anyone who has access to the Internet. Our project, Canada's Early Women Writers (CEWW), is currently constructing a bio-bibliographical database of all Canadian women who published in English—in any genre, in any forum—before 1950; this data will be housed in CWRC's repository and accessible through the CWRC interface.

This poster will present our project to the DHSI participants, focusing on our contribution to and interaction with CWRC and highlighting the interdependency of the DH research community and the general public.
CWRC is amassing extensive data from a multitude of research projects, including CEWW. CWRC's linked dataset will be utilized by customized embedded applications providing powerful, innovative tools for analysis. Scholars will thus have access to data from other projects, facilitating cross-discipline research and analysis at an unprecedented level.

CEWW is engaged with the acquisition of data for use within the CWRC project, and our goal is to provide as comprehensive a set of data as possible. Information for the original database at SFU, created in the 1980s, was acquired through written letters and visits to archives and libraries across Canada. The accessibility of digital information has permitted the expansion of the original database of 471 authors, so that now our list of Canadian women writers exceeds 4000.

In addition to being an invaluable source of data for scholars of Canadian literature, CWRC promises greater collaboration between the research community and members of the general public; one of the criteria we use in deciding which of the 4000 women to create entries for, is the interest shown in the author by members of the general public.

CEWW has been maintaining an online presence since early 2011. Every week or so, we post a poem or article by one of our more obscure authors; as well, the site hosts the extensive lists of

names of women we know published something. Not infrequently, we are contacted by individuals from around the world with questions or information about the women on these lists. It often takes only a small clue to help us unravel the details of the author's life: and from a name, and possibly a date, with the help of the digital community, a full entry is born. The construction of our database is truly a community project, which, to us, is one of the powers of the digital world: the sharing of data makes all of us intellectually richer.

Biographical Note

Karyn Huenemann is the Project Manager for the Canada's Early Women Writers project at Simon Fraser University. This project was begun in the 1970s, pre-digital, by Dr. Carole Gerson. Karyn was hired at the outset to oversee the technical aspects of the updated project, a role that provides gratifying recourse into both the digital and the human.

Darwin's Semantic Voyage

Jaimie Murdock (Indiana University)
Colin Allen (Indiana University)
Simon DeDeo (Indiana University)

New advances in the digitization of historical archives allow us to construct biographical datasets to study how a single individual, over the course of a lifetime, explores and synthesizes the work of contemporaries and predecessors. In this work, we explore the reading habits of one of the most successful and celebrated scientists of the modern era: Charles Darwin. Darwin was a meticulous record-keeper—starting in April 1838, he kept a notebook of "books to be read" and "books read". These records span the 23 years from 1837–1860, a critical period which culminated in the publication of The Origin of Species. We link his records with the full text of the original volumes, and then use probabilistic topic models to represent these texts as mixture of topics. We use information theory to measure the surprise, or unpredictability, of the next text that Darwin chose to read, compared to his past history of reading. This allows us to investigate his shifts between reading within a given domain and switching to new domains. Darwin's behavior shifts from exploitation to exploration on multiple timescales, and at the longest timescale these shifts correlate with major intellectual epochs of his career. Finally, we show that Darwin's consumption of the texts is more exploratory than the culture's production of them. In contrast to previous studies using topic modeling to analyze the large-scale structures of scientific disciplines and the humanities, our focus on the role of individual reading behavior allows us to see what Tria et al. (2014) call "the interplay between individual and collective phenomena where innovation takes place". Our work reveals an important distinction between these two levels of analysis; underneath gradual cultural changes are the long leaps and explorations comprising an individual's consumption, combination, and synthesis. Using these same techniques, we demonstrate relationships among his writings and readings, confirming his own pronouncements about intellectual influences, and indicating how such evidence may help resolve scholarly questions about such influences.

Biographical Notes

Jaimie Murdock is a PhD student at Indiana University in Informatics and Cognitive Science, researching creativity and the digital humanities.

Colin Allen is Provost Professor of Cognitive Science and History & Philosophy of Science.

Simon DeDeo is an Assistant Professor at the School of Informatics and Computing at Indiana University and External Professor at the Santa Fe Institute.

Possible Spanish Idiom In A Name At Nootka

Paula Johanson (U Victoria)

The name of Maquinna, leader of the Mowachat people, was recorded in journals by Cook (in 1778) and Vancouver (in 1792) when these explorers visited the harbour they called Nootka. Maquinna's name is in use today among his descendants, including Chief Earl Maquinna George. It is interesting to study this name's origin by reading not only his descendant's autobiography but records in English written by explorers and colonists, and records in Spanish as well. An Internet search for mentions of Maquinna in Spanish turns up some idioms appropriate for study. First Nations names and words, or phonetic transliterations in English, are often slightly changed from the original by English speakers unfamiliar with the original language; similar phonetic changes happen in Spanish as well. While there are two words similar to Maquinna in First Nations vocabularies recorded in the 1860s by English-speaking colonists (and kept in the online catalog for the Heritage Room of Greater Victoria Public Library), there are several Spanish idioms accessible online which use the identical word. A study of these idioms suggests the name Maquinna is not an English transliteration, but a Spanish version of the original name of the Mowachat leader.

Biographical Notes

Paula Johanson (MA in Canadian Literature) is in the Digital Humanites Graduate Certificate program in the English Department at University of Victoria. She was a Community Fellow at UVic's Centre for Cooperative and Community-Based Economies. Her published fiction includes a novel, short stories, and poetry, while her published nonfiction includes thirty books on science, health, and literature for educational publishers. She has taught writing workshops, worked an organic-method small farm, and raised gifted twins. Twice she was shortlisted for the Prix Aurora Award for Canadian science fiction, and she served on the 2015 jury for the Sunburst Award for Canadian speculative fiction. Check out her author website at http://paulajohanson.blogspot.ca – and ask how kayaking and hearing aids have improved her studies!

Colloquium Session 4, June 14 4:15-6:00
Special Session: Building an Inclusive DH Community

Working in the Digital Humanities - An exploration of scholarly practices for early-career academics

Steven G. Anderson (University of California, Riverside)

Matt Bouchard (University of Toronto)
Andy Keenan (University of Toronto)
Lee Zickel (Case Western Reserve University)

In this project, we explore how different ways of working impact scholarship in the digital humanities. While there have been many discussions defining digital humanities, we believe there is an interesting space for exploration around how digital humanities scholars work. We evaluated the different practices that each of us use to accomplish our particular goals. All four of us are early career academics, in or approaching the candidacy stages of our PhDs.

We are building on the work started by Powell et al. (2012) to take a snapshot of current practices in DH and continue (re)defining of PhD DH work called for by Berman (2011) and Smith (2012). Our addition to the community of scholarship with this paper is the intersection between praxis and poesis, or knowing through process.

This project focuses on four major areas for scholarly production: 1. Time management - how to manage the competing demands of the academic life; 2. Creativity - how to create new ideas, new directions, or new research interests; 3. Tools / Software - how to select the appropriate workflow for a successful project; 4. Production - how to spend the hours allotted to the particular project, which is often the most difficult part.

Based on these four areas, our group generated a series of questions to help elicit responses. We each answered these questions and created a framework for our scholarly work practices.

In this presentation, we will report our findings from these four areas and discuss the implications for our scholarly output. We hope to contribute insights for different ways to "do DH". Drawing on the suggestions from Ramsay that DH is about making/building/programming/participating (Ramsay 2013), we will examine exactly how this "doing DH" happens. As most DH scholars are expected to participate in multiple roles, we will share our practices for balancing these many competing priorities.

References

Berman, Russell. 2011. "Reforming Doctoral Education: The Sooner the Better." In MLA Newsletter. http://www.mla.org/blog&topic=143.

Powell, D. J., Bouchard, M., Dalgleish, M., Keenan, A., McLeod, A., & Thomson, T. (2012). Conversation, Collaboration, Credit: The Graduate Researcher in the Digital Scholarly Environment. Digital Studies/Le champ numérique.

Ramsay, S. (2013). On building. Vanhoutte, E., Nyhan, J., & Terras, M. (Eds.). Defining digital humanities: a reader. Ashgate Publishing, Ltd, 243-45.

Smith, Sidonie. 2012. "Toward a Sustainable Humanities: Reconceptualizing Doctoral Education for the 21st Century." Paper presented at the Congress 2012 of the Humanities and Social Sciences, Canadian Federation for the Humanities and Social Sciences, University of Waterloo

and Wilfrid Laurier University, Waterloo-Kitchener, Ontario, May 26 - June 2.
http://www.youtube.com/watch?v=LcMSqrBYp8Y&feature=plcp

Biographical Notes

Steven G. Anderson is a PhD candidate in 20th Century American History at the University of California, Riverside.

Matt Bouchard is a PhD candidate in the Faculty of Information at the University of Toronto.

Andy Keenan is a PhD candidate in the Faculty of Information at the University of Toronto.

Lee Zickel is a PhD student in Design and Innovation at the Weatherhead School of Business Management at Case Western Reserve University.

DH Internships: Building Digital Humanities Capacity with Care at Emory University

Alan G. Pike (Emory University)

Graduate students have long played an integral role in the production of Digital Humanities scholarship in the US and elsewhere. They work in labs; as research assistants and fellows; as principal investigators and grantees. Graduate students work in digital scholarship with and without pay, in all kinds of institutional settings and arrangements. In her 2015 blog post about digital humanities and graduate education in the United States, Bethany Nowviskie asked,

How might taking care—and taking the concept of care more seriously in graduate education and cultural heritage infrastructure-building—serve to expand our scope?

Nowviskie argued that, as we continue to build individual, institutional, and even national capacity for digital scholarship in higher education, we should make an "ethic of care" the foundation upon which we work. This presentation will discuss how the Emory Center for Digital Scholarship's recently launched Digital Scholarship Internship Program is designed to build digital scholarship capacity at Emory through with the "ethic of care" for which Nowviskie so vigorously argues. The Digital Scholarship Internship Program presents a unique opportunity for graduate students to learn digital scholarship skills, tools, and methods relating to research and pedagogy, all while giving them practical experience in digital scholarship work. The training and professional development that make up this program are designed to prepare all ECDS graduate students to be successful in careers within, alongside, and beyond the academy.

This presentation will offer an overview of the structure of the program, and a discussion of the challenges and opportunities created in and through the program during its pilot year. The program has welcomed 5 new students this year, bringing the ECDS graduate student staff to almost 30. It is structured around 3 tiers of skills acquisition, expertise, and participation in project work and faculty support. This structure is supported by a system of microcredentials to mark progress through the program and skills acquisition. The presentation will also address the underlying questions about capacity and care that inspired the program's design, and which should drive the larger DH community in its efforts to include graduate students in its work,

improve the career prospects of its students, and refine the nature of collaborative work in the field.

Biographical Note

Alan G. Pike is the Digital Scholarship Training Coordinator at Emory University's Center for Digital Scholarship. He runs the ECDS Digital Scholarship Internship Program for graduate students, handles center operations and project management work, and consults with faculty, staff, and students on digital research and pedagogy.

Mapping German Prisoners of War in the Soviet Gulag, 1945-1956

Susan Grunewald (Carnegie Mellon University)

There is no historical event so central to Russian identity today as the "Great Patriotic War," or World War II. My project examines German prisoners of war in the Soviet Union from 1941-1956. The Soviet government held roughly 1.5 million German POWs in the Gulag system after the end of the war, the largest and longest held group of prisoners kept by any of the victor nations. My project centers on the following questions: Why did the Soviets keep the German POWs for so long? Was their incarceration politically or economically motivated? Finally, how did they fit into the Gulag forced labor system? In answering these questions, I aim to also discuss the extent to which German POWs contributed to postwar reconstruction, and whether their experiences were unique or shared by other prisoners in the Gulag.

Based on preliminary research, I argue that the Soviet government detained the POWs primarily as a needed labor source to aid in post-war reconstruction. The war did tremendous damage in the Soviet Union: the country lost 26 million soldiers and civilians, two of the six million Jews murdered by the Nazis were killed on Soviet soil, and the industrial and agricultural infrastructure in the occupied territories was destroyed. The Soviet Union emerged from the war with a dire shortage of labor and a ruined economy. Due to this damage, the Soviet government chose to harness the labor of able-bodied German POWs for as long as possible to aid in the postwar reconstruction. Archival evidence shows that the state was eager to release prisoners who were ill or injured. Treatment of the POWs was not especially punitive, and they were not singled out among other prisoners. Indeed, they received roughly the same treatment as Soviet convicts in the Gulag system. Like Soviet convicts, their conditions improved or worsened depending on the overall conditions in the nation. As provisioning and consumption gradually improved after the war, so too did conditions in the camps.

A significant part of my project is mapping. Using Russian archival sources and the program ArcGIS, I aim to answer key questions related to the German POW Gulag experience including: First, were the main POW camps located in rural or in urban areas, and how isolated were they from the Soviet prison population and the free population? Second, what contribution did the POWs make to reconstruction? Were they primarily used for resource extraction? To what extent did they have a hand in reconstructing the massive destruction of cities? Third, how did they fit into the Soviet Gulag forced labor system? Were they primarily distributed to POW only camps or were they integrated into existing ones for political prisoners and convicts? Finally, how did

their geographical location shape the experiences of POWs?

Biographical Note

Susan Grunewald is a Ph.D. candidate in History at Carnegie Mellon University in Pittsburgh, PA where she is currently an A.W. Mellon Fellow in Digital Humanities. Her research focuses on the Soviet Union in WWII and the immediate postwar era with a concentration on German Prisoners of War in the Soviet Union.

Undergraduate Contributorship Models with TaDiRAH

Aaron Mauro (Penn State, Behrend)

Digital scholarship is often a deeply collaborative and networked enterprise, one which involves multiple practitioners from a variety of academic, #altac, and non-academic contexts. Our current concerns about digital research in the humanities often revolve around promotion, tenure, and the crisis in publishing. These concerns are of little interest to our undergraduates. Nevertheless, validating and evaluating digital scholarship is crucial to the future of scholarly communication and also holds the promise of greater public and student engagement. Altmetrics are helping to redefine what we consider to be a research contribution and how we measure its impact. A range of knowledge stakeholders are redefining how we measure contributions from students, staff, faculty, and community stakeholders. Collaboration, I argue, is a fundamental digital humanities metric of success for digital projects, and community building is a contribution that blends expert opinion with local knowledge and community participation.

As co-chair of the new Digital Arts, Media, and Technology (DIGIT) undergraduate major at Penn State, I have worked to develop a DH intensive curriculum that makes collaboration and community formation a core goal. Project based pedagogies inform how students work in teams to solve humanities problems with digital tools. I teach four core competencies for the major, including our Introduction to Digital Humanities, Programming for Humanists, Digital Archiving, and Project Development. As a regular part of my teaching and pedagogy, I use TaDiRAH (Taxonomy of Digital Research Activities in the Humanities) to have students conduct self-assessments of their collaborative abilities (https://github.com/dhtaxonomy/TaDiRAH).

This presentation will share the results of this use of TaDiRAH in teaching with the first cohort of students entering the major. The data comprises the self-identified roles of the freshman in the Introduction to Digital Humanities course during the completion of their final projects. As undergraduate digital humanities classes and programs begin to be taught more broadly, it is important for us to understand the predilections and habits of mind that shape new researchers. Surprisingly, students' perceived technical abilities at the opening of the course does not strongly correlate to their later, self-defined roles in TaDiRAH. As technical proficiency grows, students appear to have a "tipping point" after which self-identification as a technical leader replaces previous reservations. The correlation between grading and this technical tipping point will be the focus of my analysis. My final assessment will weigh the value of modeling research activities through a taxonomy at the undergraduate level.

Biographical Note

Dr. Aaron Mauro is Assistant Professor of Digital Humanities and English at Penn State Erie, The Behrend College.

The Anti-MOOC: A Synchronous Small Seminar Format for Distance Mentoring and Digital Public History Projects

Cathy Kroll (Sonoma State University)

This presentation reports on the results of a synchronous interdisciplinary distance mentoring course—Festivals: Culture in the Making—team-taught this year by a professor of history in Texas (Dr. Whitney Snow) and a professor of English in California (Dr. Cathy Kroll). We taught our seven students from Council of Public Liberal Arts Colleges (COPLAC) around the U.S. via web conferencing hosted by the University of North Carolina, Asheville. In this course, students selected a local festival to research over the course of the semester. With the aid of our combined expertise in archival research, ethnography, and digital humanities, students learned to formulate original research questions, to conduct and digitally record oral interviews and festival videos, to use ethnographic research methods and analysis, to undertake archival research, and to design websites showcasing their research results. In our twice-weekly meetings, I demonstrated DH tools such as mapping and timeline creation, as well as principles of web design. Our students' finished digital public history projects are now helping to build a permanent record of their communities' local festival traditions and their communities' collective memories.

As professors for the course, we made certain shifts in pedagogical design in this digital environment, emphasizing individual and group mentoring. This course was distinguished by its multi-regional, multicultural, multi-tech, and interdisciplinary features. Students may have learned as much about bridging wide differences in regional culture, political philosophies, and local traditions as they did about how to conduct an ethnography of a festival and digitally craft their research results.

Biographical Note

Cathy Kroll is Professor of English at Sonoma State University in northern California. She teaches courses in nineteenth-century British and postcolonial literature, digital humanities, and writing. Her recent scholarship focuses on digital public history projects and digital pedagogy.

From Silent Films to Digital Storytelling in the Undergraduate Classroom

Mia Tootill (Cornell University)

This presentation details my experience teaching a course on "The Sounds of Silent Film," in which my students used digital tools for multimodal creation and critique. Aimed at undergraduate non-majors, the course traced a path from nineteenth-century musical and visual technological innovations to the emergence of silent film. Students were asked to consider questions such as: Why use film to tell stories? What were the limitations of written text—and even staged plays—that provoked the need for cinema? At the same time, they completed assignments that rejected

the traditional written paper format and required them to engage with technology in new ways. These included: 1) Weekly video annotations using VideoAnt, 2) A scoring project in which they used video editors to synchronize pre-written music to silent films, and 3) A final Scalar project on the intersections between musical culture and silent film in our local town of Ithaca in the 1910s.

Silent film can seem antiquated to young people today and my students were no different. Over the course of the semester, they used these digital tools and explored recent interactive projects on silent film in an effort to understand the novelty of the medium for audiences at the turn of the twentieth century. I used this digital pedagogy approach to help demonstrate that the films they studied emerged from a similar desire for multimodal engagement as their love of all things digital today. Furthermore, these assignments sought to solve accessibility issues—both in terms of the challenges of teaching music to non-majors and also the barriers to younger audiences engaging with silent film revivals today.

Biographical Note

Mia Tootill is a Ph.D. candidate in musicology at Cornell University. She currently serves as a digital humanities intern in the Cornell Library and as the assistant project coordinator for the GloPAC (Global Performing Arts Consortium) digital project. She is also a HASTAC scholar. Her research centers on nineteenth-century French opera and popular theater, silent film, data visualization, and community-engaged digital pedagogy. Her digital dissertation appendix will be available shortly at mappingparistheaters.com.

Colloquium Session 5, June 15 4:15-6:00

Sounds and Digital Humanities

John F. Barber (Washington State University Vancouver)

As a disciplinary area of scholarship and pedagogy, Digital Humanities (DH) seeks to engage academic scholarship with creative practice to promote critical thinking, communication, digital literacy, and civic engagement.

While DH has embraced visualization—images, animation, video, and text as image—as the basis for scholarship presentation and pedagogy, sound is, arguably, overlooked.

This presentation highlights topics to be considered during the DHSI 2016 "Sounds and Digital Humanities" course in which sound is positioned as a primary sensory input, still very powerful despite current visual dominance. Sound provides a way of knowing and being in the world. Listening to sound can invoke associations unlike any other medium. Sound provides new opportunities/approaches for scholarship, especially as scholars gain appreciation for and ability using sound(s) to enhance or ground their scholarship / pedagogy.

Soundscapes, sound maps, sound collages, and remixes; digital storytelling; aural and oral histories / biographies / documentaries; curated exhibitions / installations / performances / broadcasts; or stand alone artifacts (embedded sound, podcasts, web-based radio, archives,

curated collections) provide broad frameworks for the use of sounds in DH scholarship and pedagogy. Several will be outlined during the presentation.

In conclusion, opportunities / approaches for sound in Digital Humanities research, scholarship, teaching, and learning are promoted.

Biographical Note

John F. Barber teaches in the Creative Media & Digital Culture program at Washington State University Vancouver, Vancouver, Washington, USA. He developed and maintains Radio Nouspace (www.radionouspace.net), a curated listening gallery/virtual museum and practice-based research and creative practice space. Both are inspired by the radio medium with its emphasis on sound as the basis for narrative and storytelling. His radio+sound art work has been broadcast internationally, and featured in juried exhibitions in America, Canada, Germany, Portugal, and Macedonia.

Mapping a Digital History of Big Science

Elyse Graham (SUNY Stony Brook)

Our project involves developing a new kind of digital resource to capture the history of research at scientific facilities in the era of the "New Big Science." The phrase "New Big Science" refers to the post-Cold War era at US national laboratories, when large-scale materials science accelerators rather than high-energy physics accelerators became marquee projects at most major basic research laboratories. The extent, scope, and diversity of research at such facilities makes keeping track of it difficult to compile using traditional historical methods and linear narratives; there are too many overlapping and bifurcating threads. The difficulties of writing histories of the New Big Science are only partly due to format; there are also notable methodological difficulties. For instance, conventional research methods in the humanities, such as the "material history of objects," are designed to take account of comparatively simple objects, such as books, rather than such large and complex instruments. We are developing a new kind of digital tool, however, for this purpose. This presentation will discuss existing "paper prototype" models for a specific case project—the NSLS Digital Archive—along with a brief overview of the literature on the theory and practice of digital tool design.

Biographical Note

Elyse Graham is assistant professor of digital humanities at the State University of New York at Stony Brook and a research affiliate at the Massachusetts Institute of Technology. Her first book, on the future of reading, is under contract.

Mapping a Global Renaissance with 53,829 Texts

James Lee (Grinnell College)

"Mapping a Global Renaissance with 53,829 Texts" redefines our understanding of a squarely humanistic problem: the history of race in Shakespeare's era. By analyzing thousands of texts

beyond the scope of a single famous author, such as Shakespeare or Milton, and even the capacity of the individual reader, I tell a very different story about the multiple discourses of race that helped motivate England's earliest efforts to define its place in a global context in the era before colonialism. My primary technique, Latent Dirichlet Allocation (LDA) is a topic modeling algorithm that identifies clusters of words exhibiting a disproportionately high probability of occurring together in all texts discussing place names in the texts of the Early English Books Online corpus. My results affirm and complicate recent postcolonial accounts of Shakespeare's world by redefining the ideology of race in a more nuanced way than along the faultlines of identity politics. Through this methodology, I articulate an unfamiliar Renaissance vision of race by challenging the basic assumption that ethnic otherness has always been built upon a bedrock of bodily difference. The pre-modern idea of race was structured by a logic of place. Geography, and not skin color or anatomy, was the dominant factor in setting the terms of the debate. The primary question was how space, landscape, latitude, climate, and a location's flora and fauna shaped a culture, its people, and their bodies. These results suggest that the focus on the body as the measuring stick of racial ideologies should be understood as an invention of a colonialist worldview, and that historical alternatives to understand race in different terms exist. This study thus opens up a critical space to reanimate historically significant but unfamiliar models of race that we have lost in the postcolonial world.

Biographical Note

James Lee is Assistant Professor of English at Grinnell College. His first book is under contract with Northwestern University Press. His second project is a book-length digital history of Renaissance globalism, race, and geography. It uses topic modeling, network analysis, and mapping to visualize how Renaissance England began to imagine its place in a global frame of trade and diplomacy in the texts of the EEBO / TCP corpus.

Digitization and Dissemination of Movable Books Data

Emily Brooks (University of Florida)

The digitization of texts is at the core of first wave digital humanities projects. Google Books has scanned over 30 million books, HathiTrust has over 5 million ebooks, and Project Gutenberg has digitized over 50,000 texts. Google Books uses automatic scanning of the physical pages whereas Project Gutenberg makes the text available in ASCII, independent from the original page layout or typography. Both of these approaches are based on two-dimensional, primarily text-oriented books. Conspicuously absent from all three collections are the category of "toy and movable books." This is likely because they do not translate fully via photographs, scans, or transcriptions. There are a few videos of pop-up books scattered across the internet, but these are usually promotional for sales or exhibitions. The technology to easily capture video and 3D models is more accessible and affordable than ever before, and yet the resource shortages that many libraries, archives, and museums are already facing means that incorporating a new method of preservation is unlikely to happen in the near future (notable exceptions include the Smithsonian Institute, the Metropolitan Museum of Art, and the British Museum). My project aims to better take into account the materiality of movable books by exploring multiple methods of digitization; finding optimal solutions based on type of movability, accessibility, and cost; and exploring best archival practices. For example, movable books incorporating pull tabs are better

digitized via a time-based medium like video while intricate pop-ups are better digitized via a spatial-based medium like 3D scanning. Furthermore, there are a multitude of scanning tools to choose from such as a lower-resolution mobile application like Autodesk's 123D Catch which uses a tablet's built-in camera and photogrammetry or a desktop application like Autodesk Memento which can use high-resolution photographs from any camera. And lastly, based on data collected in the Spring 2014 Global Attitudes Survey by the Pew Research Center, another point to consider is that in developing countries like Indonesia, Ghana, India, Nigeria, Kenya, and Uganda, a person is more likely to have a smartphone. Therefore, it's important that the data be archived online in a mobile-friendly manner so that it can be more globally accessible.

Biographical Note

Emily Brooks is a Ph.D. student in the English department at the University of Florida. She is the production editor of the open-access journal, ImageTexT, is a member the inaugural digital humanities certificate class at UF, and co-leads Arduino trainings at the Marston Science Library at UF. Previous university courses she developed and taught include Writing About Magic, Writing Through Media: The History (and Future) of the Book, and Survey of American Literature (Children's Fantasy). Her research interests are in using 3D scanning and printing technologies to enhance cultural heritage preservation of non-traditional books.

Old Testament, New Tricks: Using Biblical Translation to Examine Word Sense and Popular Belief

Zachary Bleemer (University of California, Berkeley)

One strand of recent research in digital humanities and computational linguistics uses digital tools to process the natural language of large available corpora like HathiTrust and Twitter. However, sample selection bias in the construction of those corpora—arising from the (necessary) selection of texts that discuss topics of contemporaneous public interest (and thus texts that were indeed written and published), that were repeatedly condoned or overlooked by formal or informal censors, and that have been digitized and catalogued—frustrates inference about the societies and populations purportedly represented by each corpus. I introduce a framework using digital tools and a small but highly-structured dataset—namely, 14 translations of the Old Testament—that plausibly avoids these sample selection issues, and propose a novel similarity measure over translational word choice clusterings to identify differences between a set of latent beliefs manifested by each text. Biblical translations' substantial within- and between-translation word choice variation suggest that this corpus is capable of identifying a variety of multi-dimensional latent beliefs.

In order to evaluate the practical usefulness of this framework in measuring similarities and differences between popular latent beliefs, I then present a working example analyzing differences in beliefs regarding beautiful objects across three broadly-defined populations: 16th century Great Britain, 20th century Germany, and the contemporary United States. In particular, I assume that translation teams latently categorize nice-looking objects (qua their nice-looking-ness) using known characteristics of those objects (e.g. their form or ethical stature), and that all similarly-categorized objects are translated as the same term (e.g. 'beautiful' or 'fair') when the origin text describes them with the Hebrew root 'Yaphah', which refers to objects' nice-looking-

ness. This assumption is sufficient to identify and quantify differences between translation teams in their latent aesthetic categories. I present evidence that these latent beliefs are well-defined across at least one population—the contemporary United States—with disparate translation teams having produced dissimilar text that is nevertheless shown to manifest highly similar aesthetic categories. I also present evidence of measured differences in latent aesthetic categories between the three studied populations. Future research will develop unsupervised latent models (e.g. topic models) to characterize the content of these similarities and differences across populations.

Biographical Note

Zach Bleemer is a data scientist and microeconomist studying aesthetics and young person decision-making at UC Berkeley, where he is a PhD student in economics and a Digital Humanities Fellow. His research in the digital humanities, which has been supported by the Mellon Foundation, the Folger Shakespeare Library, and DH at Berkeley, examines popular aesthetic belief since 1500 across Europe and the United States. Zach holds bachelor's degrees in philosophy, mathematics, and economics from Amherst College, where he won the Gail Kennedy Memorial Award for a philosophy thesis on Kantian aesthetics, and conducted post-baccalaureate studies in statistics at Columbia University. For more information, see his website: http://econgrads.berkeley.edu/bleemer.

Colloquium Session 6, June 16 4:15-6:00

Using DH to Increase Legal Literacy and Agency

Susan Tanner (Carnegie Mellon University)

Craig Hoffman contends in "Using Discourse Analysis Methodology to Teach 'Legal English'" that both discourse analysts and legal scholars have been too limited in their fields of study to adequately address important research questions. Each field, he argues, is particularly focused on shades of meaning and each uses scientifically rigorous methods. While this is a strength of each discipline, it also means that scholars are not currently using valuable interdisciplinary tools to answer larger questions that affect each field individually.

I argue that understanding legal precedent as a complex study in intertextuality could help non-legal scholars understand some of the more sophisticated legal nuance that run, unaddressed, through bodies of legal texts as precedent. Of equal import, legal discourse is uniquely situated to test theories of how meaning of prior discourse is shaped through context. Its relationship to precedent has implications for intertextuality that go beyond mere citation. In legal opinions, concepts are entextualized and adapted to meet new legal challenges; shades of meaning are scrutinized to provide guidance for legal outcomes. However, sophisticated intertextual analyses have, until recently, been prohibitively time-consuming, even for rhetorical scholars of the law, because of the detail necessary for micro-level analysis.

Using new digital methods, like DocuScope, complex analyses of intertextual chains can be run on hundreds or thousands of legal documents simultaneously. My research seeks to add to the

availability of sophisticated legal dictionaries to begin to answer the question: "How is legal precedent entextualized within judicial opinions and what consequences are there for legal understanding of these texts?" My presentation will highlight some of the ways scholars can employ digital tools to aid in their research and showcase a case study on intertextuality in Supreme Court privacy law cases.

<u>Biographical Note</u>

Susan Tanner is a third year PhD student in Rhetoric at Carnegie Mellon and an A.W. Mellon Digital Humanities Fellow. Her background in the law (earned a J.D. in 2012) has guided her research focus on issues of exclusivity and silence in the law.

ABCs of Gamification: How we Gamified a Social Media Course

Rob Bajko (Ryerson University)
Jaigris Hodson (Royal Roads University)

The average student enters university with at least 10,000 hours of gaming experience. Games have been shown to not only increase engagement, but also to present content in an enjoyable and interactive way which permits extensive rehearsal and practice in education (Kaufman & Sauvé, 2010; Bajko, Hodson, Seaborn, Livingstone & Fels, 2015). Work by Kaufman, in particular, showed that simulation-gamification has a strong educational rationale favoring it over conventional simulation for learning in the health professions.

Gamification is the use of game mechanics in non-game systems such as education, business and health care. Game based learning is a subset of gamification activities designed to use game elements to facilitate concept mapping, tacit knowledge acquisition, procedural learning in simulations and introduction of didactic content. Gamification, as the umbrella set of pedagogical tools in continuing education also fosters intrinsic motivation. However, the penetration of gamification-based curricula in continuing education has been limited by its perceived high demand for sophisticated game management systems and graphical interfaces. This compounds the problem for faculty who wish to explore gamification in their courses, who already must focus on innovation strategies to justify these new methodologies to both administration and learners.

The material presented here is our qualitative research data that examined an undergraduate social media courses that was gamified. In the first half of 2015, three undergraduate social media classes were gamified in the intent to increase student engagement and interest in the course. The course was gamified by providing students with the opportunity to earn experience points by completing quizzes and numerous exercises in class, as well as to earn popularity points through a series of dice rolling challenges. The purpose of our study was to answer the following three questions: (1) What is the perceived impact on course engagement and performance? (2) What are the attitudes towards the gamification of a social media undergraduate course? and (3) What are the challenges in delivering a gamified course? Students were asked to complete two online surveys, one at the midway point (i.e. week 7) of the term where quantitative data were gathered and at the end of the term (i.e. week 12) where quantitative and qualitative data were gathered through an online survey and one-on-one

interview sessions respectively. We found that students reported that they felt more engaged with the course material in both courses although there was a difference between students who considered themselves gamers versus non-gamers. Students also reported an elevated interest in the material being taught. Lastly, many students reported an increased overall enjoyment in both courses, particularly when compared to other, non-gamified classes.

In this session we will also discuss our ongoing research work in the field of gamification at the university undergraduate level, specifically a social media course. Design concepts, creation of narrative and testing of the narrative/gameplay interfaces will be reviewed and discussed, alongside an overview of how to evaluate educational games for effectiveness.

References

Bajko, R., Hodson, J., Seaborn, K., Livingstone, P., & Fels, D. (2015). Guilds, die rolls, and leaderboards: Gamification of two undergraduate multimedia and social media courses. Proceedings of the EDSIG Conference, Wilmington, NC.

Kaufman, D., & Sauvé, L. (Eds.) (2010). Educational gameplay and simulation environments: Case studies and lessons learned. Hershey, PA: IGI Global.
Biographical Notes

Rob Bajko's research interest involves the impact of mobile technologies such as laptops, tablet computers, smartphones, and cellular telephones on the cultural, business and social behaviour and attitudes in the workplace. He examined how the use of existing mobile technologies and current practices are disrupting and changing the flow and structure of delivery and management of business content, and how meeting participant attitudes are changing.

Jaigris Hodson's research specializes in using computer-assisted discourse and content analysis of large multimodal online and digital texts. She has published research in a wide range of academic publications including the Canadian Journal of Communication, Cultural Studies, Critical Methodologies and Loading… Journal of the Canadian Game Studies Association. She has also published in non-academic publications such as The Evolllution and spoke at TEDX Victoria 2012. She is currently working on two Social Sciences and Humanities Research Council funded research projects. The first examines the importance of soft skills for social science and humanities students, and the second focuses on Canadian social media use during election time.

Detecting Text Reuse in Nineteenth-Century Legal Codes of Civil Procedure

Lincoln Mullen (George Mason University)
Kellen Funk (Princeton University)

Beginning in the 1840s and continuing until the 1900s, most state jurisdictions in the United States adopted codes of civil procedure based on a New York code. These "Field codes" were amended by various jurisdictions to meet the legal needs of different regions. Traditional legal historians have long known that these codes were borrowed, but they have been unable to trace

the exact content of the borrowings. Using algorithms borrowed from search engines (minhash/locality sensitive hashing), gene sequencing (Smith-Waterman), and clustering (affinity propagation), we have built traced the borrowings within these codes. By tracing the borrowings we were able to identify substantive changes in the law, observe a network of families of codes as well as deviations from the typical borrowings, and depict the patterns of borrowings for each code. This method, which is broadly applicable to legal historical sources, has uncovered the 19th-century spine of modern American legal practice.

Biographical Note

Lincoln Mullen is an assistant professor in the Department of History and Art History at George Mason University, where he teaches religious history and digital history.

Kellen Funk is Porter Ogden Jacobus fellow at Princeton University and a legal historian.

Collection, Curation, and Collaboration: Representing Canadian Gay Liberationists

Raymon Sandhu (University of British Columbia, Okanagan)
Nadine Boulay (Simon Fraser University)
Stefanie Martin (Ryerson University)
Anderson Tuguinay (Ryerson University)
Seamus Riordan-Short (University of British Columbia, Okanagan)

The Lesbian and Gay Liberation in Canada (LGLC) Project reconfigures Don McLeod's monographs, Lesbian and Gay Liberation in Canada: A Selected Chronology, volume 1 (1964-1975), and volume 2 (1976-1981), as a digital resource. One facet of the project is the prosopography, a collective biography containing information on people, events, movements, and publications found in McLeod's text. Taking into account the differing skill sets of six project members, as well as institutional locations of members across Canada, this presentation will examine how collaborative methods of identity management and representation have shaped the prosopography, and the LGLC project as a whole.

The prosopography is a sparsely filled information matrix that holds and attempts to make visible historical and biographic information pertaining to the LGLC, including data on individuals involved, organizations, major events, publications, relationships, and identities. We start each prosopographic entry by recording a person's data from Don McLeod's chronology and then use external sources such as print and electronic resources, obtained through archival, textual, and online research, to add and expand on existing information.

Our divergent research practices and methods- from archival research to digitizing and encoding- shapes the different ways in which we understand the data being represented. Through our work on the project some of the challenges we have encountered include the following: 1) Considerations of what makes one source more reliable than the other: in our research, which sources do we privilege? What sources are accessible given the project timeline? And how is certain information included/excluded or categorized? 2) Through working to represent gay and lesbian liberation history from 1964-1981, how do we account for and represent changes in identity, location, and politics across time and place? Through highlighting taxonomies of

sexuality and gender identity, what other facets of identity might be obscured? What gets left out of categories of representation and identity politics for visualizing and encoding LGBTQ individual and community histories? 3) And finally, as a diverse research team situated across three separate universities, how do we engage with questions and challenges around representation and categorization through collaborating across different skill sets, institutions, and time zones? Through our polyvocal and collaborative presentation, we will examine and engage the aforementioned questions and concerns around representation in the LGLC project.

Biographical Notes

Raymon Sandhu: BA in Comp Sci at UBCO. Focus in machine learning and probability models.

Nadine Boulay: PhD at SFU, Department of Gender, Sexuality, Women's Studies- researches queer history and post-war social movements in North America.

Stefanie Martin: BA in Sociology at Ryerson. Focusing on media representation of women.

Anderson Tuguinay: MA Literatures of Modernity, Ryerson. Volunteer, Canadian Lesbian and Gay Archives.

Seamus Riordan-Short: BSc Chemistry (Analytical) at UBCO. Minor in English.

Privacy, Legality, and Feminism: How Do We Build a Feminist Politics into Open Access Data Structures?

Emily Christina Murphy (Queen's University)

In recent digital humanities research, scholars have looked to linked open data as an especially fruitful technology through which to structure open access data, and libraries and research projects alike have developed ontologies that seek to support the discovery and research processes of humanities scholars in a digital space. In this paper, I am particularly interested in what a feminist approach to data structure might look like. In particular, how can linked open data can be used to structure women's contributions to literary and cultural production in the spirit of women's cultural production historically?

While open access may seem like a natural ideological framework for modelling a feminist approach to contemporary scholarship and historical literary community, other scholarly approaches to archive and community rightfully criticize open access data as a structure that assumes a normative, settler-colonial approach to cultural access protocols (Mukurtu and the Digital Index of North American Archaeology, for example). Attending to indigenous cultural access protocols, projects may wish to preserve closed or limited access to sensitive materials or to representations of private communities, or they may wish to critique recreations of normative models of cultural production.

These critiques of normative settler colonial access protocols are instructive for feminist approaches to building, thinking through, and critiquing data structures. For example, women's writing in an early-twentieth century Anglo-American context has been characterized by private

contributions, care work and labours of "literary accouchement" (Benstock). Shari Benstock argues, for example, that modernist women like Bryher (Annie Winifred Ellerman) and H.D. (Hilda Doolittle) "lived between two worlds: the dominant heterosexual culture and a marginal homosexual culture" (Loc 6565). In living between two worlds, these women's public and private personae, and their public and private circulation of literary works are fundamental dichotomies of their contribution to the modernist literary and cultural field. Emerging projects like Linked Modernisms are developing ontologies that explicitly look to account for change and flexibility in scholarly understandings of a field and for buried assumptions and intersections of identity on the part of its subjects and its contributors. The importance of privacy in women's cultural production may prove an important cultural critical context, one that may shift either the development of linked open data ontologies, or, indeed, the increasing scholarly investment in openness itself.

This paper identifies the need for a critical perspective on the ethics of open source code work and open access data curation that aligns with an ethics of feminist scholarship. It calls to extend work on cultural access protocols and sensitivity to issues of public and private cultural production to technologies like RDF, in order to align this emergent standard within the digital humanities with work that has already been developed within and outside the digital humanities.

Biographical Note

Emily Christina Murphy is a doctoral candidate in the Department of English at Queen's University, Canada. Her dissertation project focuses on representations of female literary celebrity and mental illness in the modernist period, a project that has cultivated interests in modernist journalism, political activism, and public and private writing. She is co-editor of a forthcoming edition of Digital Humanities Quarterly on undergraduate education.

The Fiction Factory Goes Digital: Collaborating with Undergraduates on a Sweeney Todd Edition

Rebecca Nesvet (University of Wisconsin, Green Bay)

Last year at DHSI, I began building a digital documentary edition of the works of James Malcolm Rymer, Victorian inventor of the myth of Sweeney Todd. He had an immense effect upon literature and culture, creating not only Sweeney Todd but a novel that informs Dracula, an early werewolf novel, and the first prose fiction representation of the Indian Rebellion of 1857, a key event in India's struggle for independence from the British Empire. Unfortunately, because Rymer was a working-class professional writer and used a variety of pseudonyms, scholars have been slow to definitively identify and critically respond to his work. Most of his novels do not exist in any modern edition, much less any scholarly one. This is why, with fellow DHSI alumna Rae Yan, I've begun encoding and editing his fiction. This year, my students and I encoded and annotated 15 chapters of The String of Pearls. As we worked, I observed that our group dynamic and our computer lab space began to take on characteristics of Rymer's makerspace: the Salisbury Square, London "fiction factory" of his editor Edward Lloyd. My presentation will reveal how this recreated space facilitated exploration of the Salisbury Square literary community's creative and editorial practices--and my editorial team's.

Biographical Note

Rebecca Nesvet is an Assistant Professor of English at the University of Wisconsin, Green Bay. Her research on Romanticism and the penny blood author James Malcolm Rymer is published and forthcoming in Essays in Romanticism, The Keats-Shelley Journal, Women's Writing, Ethos Review, and Notes and Queries. #r3b33cca www.salisburysquare.com

Digital History and Archiving: Fostering the "Afterlife" and Accessibility of American Civil War Letters

Ashley Hughes (Texas Christian University)

Central to digital humanities (DH) is the desire to increase the public's access to print documents through digitization efforts. Although this desire to digitize print texts may seem very "first wave" DH, I believe it is a very worthwhile endeavor because it helps ensure the vitality and accessibility of texts. Such efforts help contribute to the pool of texts that digital humanists can mine for data, an important goal for DH because, as Rosanne Potter points out, "Until everything has been encoded … the everyday critic will probably not consider computer treatments of texts" (Jockers 175).

Nevertheless, even once historical documents have been digitized, they still may not be completely accessible to the public for a variety of reasons. For example, sometimes handwritten documents are difficult for modern readers, who are used to reading typed characters, to discern. In addition, individuals may not have the historical knowledge necessary to make sense of such documents.

In response to these exigencies, I developed a digital history project that provides access to selected letters from the "Squire Bosworth Papers" collection (1848-1892), recently digitalized by Virginia Tech's Special Collections. I selected these 19th century documents because they include correspondence dealing with social issues from the Civil War, such as southern women marrying Union soldiers, as well as difficulties stemming from Reconstruction.

While these documents are now available worldwide, the letters can be difficult for many people to read because they are handwritten and discuss non-contemporary topics. Therefore, this project carefully transcribes and contextualizes selected letters from this collection. My project incorporates a variety of digital tools that can help newcomers to the field, such as myself, begin to experiment with DH. Some of these tools include: Omeka, an open-source, web publishing platform for digitally archiving visual artifacts; Timeline JS, a tool for creating interactive, aesthetically pleasing timelines; and Google Maps, for plotting locations mentioned in my dataset.

My hope is that these efforts will animate the documents presented in this project, ensuring their "afterlife" once they enter the archive (Burdick et al. 48). Digital media are well-suited for facilitating a text's continued life because they incorporate a "user-centered approach" with a "multiplicity of use-scenarios" (Burdick et al. 48). These practices contribute to the accessibility of texts because they provide individuals the opportunity to add meta-data and interpretive content to raw texts.

Overall, the "Squire Bosworth Papers" offer intriguing insight into life in Randolph County, Virginia (now West Virginia) during a critical moment in American history, but these documents are not easily accessible to the public in their current state. I hope to not only make this collection accessible worldwide, but I have also couched these documents within historical data in order to contextualize them. This approach will allow users of my project to perform a close reading of these documents and "zoom out" of them in order to view their broader historical significance.

*Burdick, Anne et al. Digital_Humanities. Cambridge: The MIT Press,   2012. Print.
*Jockers, Matthew L. Macroanalysis: Digital Methods and Literary History. Urbana: University of Illinois Press, 2013. Print.

Biographical Note

Ashley is a first-year, Ph.D. student in Rhetoric and Composition at Texas Christian University.