

2. EMERGING METHODS AND GENRES



HUMANITIES KNOWLEDGE USED TO HAVE A RECOGNIZABLE FORM.

WE KNEW IT WHEN WE SAW IT BECAUSE IT HAD LOOKED THE SAME FOR CENTURIES: PRINTED PAGES WITH LINEAR PROSE AND A BIBLIOGRAPHIC APPARATUS WRITTEN BY AN AUTHOR AND PUBLISHED IN THE FORM OF AN ARTICLE OR BOOK. THOUGH THE FORMAT COULD VARY WIDELY, FROM MATERIALS TO LAYOUT, SUCH DESIGN DECISIONS WERE RARELY CARRIED OUT BY THE SCHOLARS WHO CREATED THE CONTENT. WITH FEW EXCEPTIONS, THE HUMANITIES HAVE ADOPTED HOMOGENEOUS APPROACHES TO PRODUCING SCHOLARLY RESEARCH. YET ARTIFACTS CREATED BY DIGITAL TECHNOLOGIES THAT “LIVE” IN DIGITAL ENVIRONMENTS ARE COMPARATIVELY DIFFERENT—IN TERMS OF MATERIAL COMPOSITION, AUTHORSHIP, MEANING-MAKING, CIRCULATION, READING, VIEWING, NAVIGATION, EMBODIMENT, INTERACTIVITY, AND EXPRESSIVITY—FROM ARTIFACTS CREATED BY THE WORLD OF PRINT.



DIGITAL MEDIA are not more “evolved” than print media nor are books obsolete; but the multiplicity of media and the varied processes of mediation and remediation in the formation of cultural knowledge and humanistic inquiry require close attention. We strongly believe that humanists need to apply the same kind of rigorous media-specific, social, cultural, and economic analyses that we have honed to study print culture to understand the specificity and affordances of digital culture and to interrogate the status of knowledge, the concept of culture, and the redefinition of the social in our global information age. We also believe that humanists must actively engage with, design, create, critique, and, finally, hack the environments and technologies that facilitate this research as we render this world-as-a-world to help us produce knowledge about who we are, where we live, and what that means.

The purpose of this chapter is to provide a field map of the experimental forms and different “knowledge models” emerging in the Digital Humanities. This is not meant to be an exhaustive or definitive list of new methods and genres but rather a conceptual and theoretical introduction to emergent practices of scholarly inquiry. We move from examining the impact of technology on the most established of humanistic practices—the decisions about what constitutes a text and its variants—to positing that enhanced critical curation of those texts makes possible augmented editions and fluid textualities that rely on the affordances of digital environments. This fluidity allows digital humanists to play with scale, both in terms of how they approach data and how they model their results. Toggling between distant and close, macro and micro, and surface and depth becomes the norm. Here, we focus on the importance of visualization to the Digital Humanities before moving on to other, though often related, genres and methods such as locative investigation, thick mapping, animated archives, database documentaries, platform studies, and emergent practices like cultural analytics, data-mining, and humanities gaming. All of these are then situated within a technological matrix that almost demands the repurposing and remixing of cultural content. We conclude this chapter by considering the utopian prospect that the massive spread of shared knowledge across networks could give rise to a state of “ubiquitous scholarship,” of ever-more interconnected, publicly engaged, participant citizens.

It is a tremendously exciting time for the humanities, as knowledge not only looks and sounds markedly different than it once did, but also feels different because it is experienced in new contexts and environments and created in collaborative spaces that involve communities who rarely, if ever, had the chance previously to participate in the scholarly enterprise.

The chapter is divided into emerging areas of experimentation, as represented in the index below. A Portfolio of Case Studies follows, providing concrete examples of these emerging genres and methods in application. Through linkages both graphical and conceptual, they combine to provide a lively and practical demonstration of Digital Humanities theory in practice.

ENHANCED CRITICAL CURATION	→
AUGMENTED EDITIONS AND FLUID TEXTUALITY	→
SCALE: THE LAW OF LARGE NUMBERS	→
DISTANT/CLOSE, MACRO/MICRO, SURFACE/DEPTH	→
CULTURAL ANALYTICS, AGGREGATION, AND DATA-MINING	→
VISUALIZATION AND DATA DESIGN	→
LOCATIVE INVESTIGATION AND THICK MAPPING	→
THE ANIMATED ARCHIVE	→
DISTRIBUTED KNOWLEDGE PRODUCTION AND PERFORMATIVE ACCESS	→
HUMANITIES GAMING	→
CODE, SOFTWARE, AND PLATFORM STUDIES	→
DATABASE DOCUMENTARIES	→
REPURPOSABLE CONTENT AND REMIX CULTURE	→
PERVASIVE INFRASTRUCTURE	→
UBIQUITOUS SCHOLARSHIP	→

ENHANCED CRITICAL CURATION

digital collections multimedia critical editions object-based argumentation expanded publication
experiential and spatial mixed physical and digital

Collection-building and curation have always defined humanistic learning: so much so that even the most ancient literary forms adopt listing, cataloging, and inventorying as key features of poetic communication. Inventories abound in Hesiod, Aeschylus, Sophocles, and Euripides. The epigrammatist Callimachus composed the 120 tablets of the *Pinakes* in which the entire holdings of the Alexandrian Library were enumerated. Apollonius of Rhodes opens the lead section of his *Argonautica* with a panoramic listing of the Argonauts. And the sheer proliferation of catalogs in the Homeric epics shows how cataloging can put into play a vivid mode of representation that is neither that of the laundry list with its skeletal seriatim organization nor that of an exhaustive, didactically intentioned inventory, but rather a composition that treads along the boundary line between verbal and visual figuration in the pursuit of informational concision and compactness. A poem within the poem, a condensation of names, actions, and things, the catalog is an art of memory that is also an art of data compression and of performance. In short, *katalégein* designates poetic composition as a compression algorithm and audience reception as a decompression tool some 26 centuries before the word “digital” began to refer to 0’s and 1’s.

Awareness of this historical background is crucial in order to understand how collection-building and curation have remained constants of humanistic knowledge production from remote antiquity through early modern courts to the academies of the Baroque era to late 19TH century universities where chairs were typically associated with the research collections. These domains became disjoined from the mainstream of scholarly practice only during the late print era, and are once again becoming integral to many forms of Digital Humanities practice.

The accumulation and care of knowledge were paramount within classical, pre-modern, and early modern regimes of data scarcity. In those eras, bits of information were valuable *a priori* and therefore either preserved, relayed, or reused, irrespective of whether they could be integrated into a cohesive structure or system of belief. Copiousness and copying were understood as inherently good under such circumstances, and collecting served the project of gathering, conservation, and retrieval. With the spread of print and the rise of modern institutions of memory (with their systematic approaches to collection and conservation), a new regime arises within which there exist such proliferations of historical information and

ENHANCED CRITICAL CURATION

cultural material that data from the past can no longer be assumed to possess *a priori* value. They become supports for the production of knowledge, knowledge's precondition but not its substance.

Informed critical judgments regarding the relationship between originals and copies, the greater or lesser authority of a given object or set of objects, and the work's meaning all become far more significant than the mere fact of accumulation. Following in the footsteps of the courtly patrons and collectors of the early modern period, new professional figures emerge alongside scholars by the end of the 19TH century, entrusted with guardianship over the remains of the past and armed with a battery of scientific and analytical techniques: archivists, museum curators, catalogers and librarians. The division between these figures and professional scholars is never absolute (as evidenced, for instance, by the role of attribution studies in art history and critical editions in the evolution of literary studies). But two parallel institutional worlds emerge that the digital revolution is reconnecting under transformed circumstances.

The reconnection assumes multiple forms: digital collection-building and curation on the part of individual scholars within and outside existing digital repositories as a form of scholarly practice; multimedia modes of argumentation that are object-based rather than discursive; conjugations of visible or audible digital media with physical objects in experiential exhibition spaces; the expanded publication of interpretive research results flanked by the archival documents and data sets that support them; large-scale collaborations that result in geospatially organized scholarly work; and critical editions of media artifacts that surround a primary object with multimedia objects rather than adopting only a text-based annotation system. It is nourished by the drive for research innovation; by the vastly expanded range of cultural materials now being produced and collected by institutions of memory, as well as by individuals and corporations; by the enhanced accessibility of these collections to both specialists and non-specialists alike thanks to their dissemination in digital form via the Internet; and by the crisis of print-based scholarly publishing and the potential for print-plus and post-print models that operate on scales unthinkable under the regime of print. *see* CASE STUDY 2 → 64 | 65; CASE STUDY 3 → 66 | 67; CASE STUDY 5 → 70 | 71

The Library of Alexandria is said to have held roughly half a million scrolls, representing works numbering in the tens of thousands. Twenty centuries later, Google Books has scanned, to date, around 14 million of the estimated 130 million printed books housed in physical libraries worldwide. What this means is that a

contemporary scholar has at his or her fingertips access to 500 times the entire corpus of knowledge seemingly available in the ancient world without even consulting a second literary database or scanning the stacks of a major research library. The figure expands exponentially when we turn our attention away from “books” or “works” to other categories: print artifacts, letters, sound recordings, paintings, photographs, objects, telegrams, Web pages, email messages, blogs, tweets....

The scale and scope of these “collections,” not to mention the accelerating creation of multimedia document collections in the present, so far exceeds the capabilities of traditional institutions of memory, not to mention the potential reach of scientific conservation methods, that enormous backlogs have become the norm. Despite sometimes heroic efforts to contain them, the amount of unprocessed or inaccessible materials in basements or off-site storage facilities are certain to grow within a setting in which information overload, the need to sift through and navigate vast data sets, and the proliferation of data trash are all givens. Critical curation is an essential scholarly practice in the print-plus and post-print world.

In common parlance, curation refers to the supervision and organization of preserved or exhibited physical items, although the term has origins in the theological domain, as in curates of the church who helped care for the souls of the dead. The term has exploded onto the contemporary scene, even invading business parlance: Slogans such as “ours is the age of curation” or “why calling yourself a curator is the new power move” are proclaimed in business blogs and reviews. What they are pointing to is the same urge animating the work of digital humanists: that the mere existence of vast quantities of data, artifacts, or products is no guarantee of impact or quality. To curate is to filter, organize, craft, and, ultimately, care for a story composed out of—even rescued from—the infinite array of potential tales, relics, and voices. In the Digital Humanities, curation refers to a wide range of practices of organizing and re-presenting the cultural record of humankind in order to create value, impact, and quality.

AUGMENTED EDITIONS AND FLUID TEXTUALITY

structured mark-up natural language processing relational rhetoric textual analysis
variants and versions mutability

Critical editions—accurate and reliable versions of a text with an apparatus that presents and analyzes the evidence and source material to reconstruct and explain the original—have been a central part of the humanities for centuries. Some of the earliest critical editions were produced for the *Hebrew Bible* and the *New Testament* and aimed to become definitive editions for a wide readership. As the genre emerged and textual criticism matured, critical editions sought to foreground instabilities and uncertainties in the reconstructive process, providing ample room for commentary and annotation, the collation of textual variants, stemmatics, and studies of authorship, editing, transcribing, and translation practices.

Digital environments provide the ability to pull together many versions of a single work, tracking its development, noting its variants, and presenting the whole comparative array of witnesses. The basic tools for migrating texts into a digital environment are well-suited to such editorial tasks. The use of structured and/or tagged approaches to identify persons, themes, places, or features of a text provides a way to maximize the intellectual investigation of documents and to display these interpretations. As standards for mark-up (the tagging process used in transcriptions) have extended and improved, many nuances in textual analysis have become part of the set of interpretive elements. Not only can we identify what something is, but we can characterize its relation to other elements or entities (part of, derived from, a cousin of, a version of, and so forth). Under digital conditions, the very same procedures may now be extended to other categories of cultural objects, such as sound recordings, video, and film. *see* CASE STUDY 2 → 64 | 65

Fluid textuality refers to the mutability of texts in variants and versions whether these are produced through authorial changes, editing, transcription, translation, or print production. In a fundamental sense, then, texts have always been fluid and modular. But the advent of word processing drew intensified attention to this aspect of textuality. Writers were thrilled with the experiences of cutting and pasting whole portions of texts without retyping. The notion of transforming a work by changing its format and typographic font with the strokes of a few keyboard commands excited critical and creative imaginations. When it first appeared, hypertext was a foreign and intriguing concept, with nodes, links, and forking paths structured to create a multifaceted text in ways that had been tried in print formats but that took on an aura of novelty and promise in new media.

New dimensions of fluidity allowing for manipulation and machine processing of textual elements were introduced through natural language processing (NLP) and other tools for textual analysis. Global changes, searches, substitutions, counting, listing, reordering—these and other activities can be carried out through commands that treat a text as an object on which to perform operations that are somewhat at variance with conventional reading. In its most fluid state, a text file can be used to generate a nonverbal outcome. An ASCII string, or keyboarded text, can be turned into output in musical or visual form, or used to make a three-dimensional print, a pattern, or a design that serves as a template for another project in a medium far from that of verbal language. Texts are constantly flowed and reflowed, repurposed and reworked for different output streams and audiences.

With the increased fluidity of texts we see a corresponding change in authorial identity. We are witnessing a shift from the age of the individual voice to that of the collaborative, collective, and aggregated voice of the fluid text. Work in digital media frequently involves a composite warp and woof whose choral “textuality” reconnects the term with its origins in the world of textures and textiles.

With the emergence of the augmented edition, the fields of analysis and editing have grown and skills for creating scholarly editions are in increasing demand. The corpus of texts and other artifacts that comprise the record of human thought and culture is migrating from print and manuscript into digital forms, and, as that process advances, the need for editing and the opportunities for critical analysis continue to expand. Imagine that the surface of the screen becomes a deep space, and what appears at first to be a single page of a text or object extends through a multiplicity of embedded layers, each displaying a different facet of an argument or history of a work’s production. Reader-viewers tier down and tier out, sifting through and engaging with, for example, a single word usage throughout a text, across a corpus, and across every book published in a given year. *see* CASE STUDY 2 → 64 | 65

The editing practices that extend into the augmented edition add other dimensions as well, allowing for a work to be understood within its larger field of cultural production; placed into the constellation of other productions and publications or artifacts of material culture; or situated within the documented events of an era. An augmented edition supports an array of arguments, with materials marshaled in demonstration of interpretations from a range of viewpoints or along a host of different lines of thought. The organization of argument in digital space creates new modes of content that are relational in their rhetoric. Sequence, juxtaposition, ordering, navigating, and analyzing are all features of the augmented edition. *see* CASE STUDY 1 → 62 | 63; CASE STUDY 3 → 66 | 67

SCALE: THE LAW OF LARGE NUMBERS

quantitative analysis text-mining machine reading digital cultural record algorithmic analysis

Although the Internet (meaning the technological infrastructure for transferring data over a distributed network of computers) is barely half a century old and the World Wide Web (meaning hyperlinked, hypertextual documents viewable in browsers) has only been around since the 1990s, it is striking to ponder the sheer volume of data they have “produced.” Statistics vary. Some sources suggest there are more than 21 billion indexed Web pages, but the number of URLs indexed by Google is over one trillion; Google has scanned, cataloged, and made searchable more than 14 million books; Technorati has indexed well over a hundred million blog records since 2002, according to its annual report on the “state of the blogosphere”; JSTOR has over 7 million articles from more than 1,000 publishers; Facebook’s repository is growing at a rate of 5 billion pieces of content per week, ranging from photographs and videos to news stories and blog posts; Twitter users produce, according to cofounder Biz Stone, one billion tweets per week. And these statistics do not even take into consideration the scope of other content produced and shared on the Web, such as email, not to mention content produced through participation in online community forums, chat groups, Instant Messaging, multi-player gaming, and mixed reality environments such as Second Life. We are producing, sharing, consuming, and storing exponentially more cultural material—including texts, images, audio, and time-based data—than ever before. We are producing data at a rate that already outstrips our ability to store them and outpaces our ability to catalog, analyze, and archive these data in meaningful ways.

The humanities have historically been the province of close analysis of limited data sets: a literary study of a novel or poem, an art history monograph about a painter’s *oeuvre*, an architectural critique of a Peruvian village’s building styles. There have long been, of course, historical, generic, and stylistic studies with a broad scope. But for the most part there is a significant divide between the ways in which the humanities approach subject matter and the ways that more quantitatively inclined disciplines approach data. In the sciences, one of the key determinants is the law of large numbers, which states that the more times a researcher repeats a given experiment, the closer that researcher comes to determining an average value that defines the results of that experiment. Translated, this is simply a way of expressing confidence that ever-larger data sets will offer ever-more verifiable conclusions. Certain physical sciences deal with extremely large numbers such as atmospheric concentrations of carbon dioxide for climatology or genetic sequencing in biology. Until recently, such was rarely the case in the humanities.

When digital technologies allow for the storage and analysis of millions of books, billions of tweets, and hundreds of billions of interactions, the ways in which we can query and comprehend the cultural record explodes. Concepts, trends, actions, and the flow of human communication come into view at a macro scale. For instance, when huge numbers of related images can be parsed by machine, and the images themselves carry massive amounts of metadata, new questions can be asked about our relationships to the visual world. How do markets set the value of images? How do images in free circulation differ from those that carry price tags? Can we detect global patterns and regional differences in the ways that societies absorb and regenerate visual culture?

To answer such questions, we will have to design and employ new tools to thoughtfully and meaningfully sift through, analyze, visualize, map, and evaluate the deluge of data and cultural material that the digital age has unleashed: tasks that will require humanists to contend with text-mining tools, machine reading, and various kinds of algorithmic analyses. One way of navigating this process is through distant reading, a form of analysis that focuses on larger units and fewer elements in order to reveal patterns and interconnections through shapes, relations, models, and structures. It is a term that is specifically arrayed against the deep hermeneutics of extracting meaning from a text through ever-closer, microscopic readings. But, beyond distant reading, the time has come to entertain the possibility of machine reading, in which trends, correlations, and relationships are extracted through computational methods. Because information is being produced on a scale that far exceeds the faculties of human comprehension, it has become impossible to read, comprehend, and analyze the digital cultural record without the assistance of digital tools and methods. *see* CASE STUDY I → 62 | 63

To cite an example whose ethical stakes are high: What would it mean to subject the 52,000 Holocaust video testimonies in the USC Shoah Foundation Institute archives to machine reading and algorithmic analyses? Averaging two hours apiece, it would take a person 24 years to watch them all, assuming he or she watched 12 hours every day of the year. There is simply no way we can process and make sense of the volume of cultural data—including traditional printed materials—without the help of a computer to process, index, select, and cluster data on a comprehensible scale. But what are the implications of turning Holocaust testimonies into units of data, statistical analyses, and compact visualizations? Does this sort of quantitative analysis not inevitably, or perhaps by definition, subject the victims to further objectification, another dehumanizing process? Might there be an “ethics

DISTANT / CLOSE, MACRO / MICRO, SURFACE / DEPTH

of the algorithm” that could mediate between the ethical demands of listening to individual Holocaust testimonies and the macrocosmic view enabled by a statistical representation of the total event? It is here that we need digital humanists to bring together the tools of technological analysis and the values, critical skills, and historical knowledge that animate the humanities disciplines.

DISTANT / CLOSE, MACRO / MICRO, SURFACE / DEPTH

large-scale patterns fine-grained analysis close reading distant reading differential geographies

Within the humanities, close reading has been a central practice that is premised on careful attention to features contained in a text, as well as its variations, history, transmission, possible meanings, and range of nuances. Close reading has its roots in the philological traditions of the humanities, but for more than a generation has often been equated with deep hermeneutics and exegesis, techniques in which interpretations are “excavated” from a text through ever-closer readings of textual evidence, references, word choices, semantics, and registers. The growth in size and accessibility of digital databases and concurrent advances in text-mining and what Lev Manovich has called cultural analytics have opened up new ways of creating meaning through distant reading. In the Digital Humanities, distant reading explicitly ignores the specific features of any individual text that close reading concentrates on in favor of gleaning larger trends and patterns from a corpus of texts. Distant reading is therefore not just a “digitization” or “quickener” of classic humanities methodologies. It is, rather, a new way of doing research wherein computational methods allow for novel sets of questions to be posed about the history of ideas, language use, cultural values and their dissemination, and the processes by which culture is made. Distant reading is almost not reading at all, but rather engages the abilities of natural language processing to extract the gist of a whole mass of texts and summarize them for a human reader in ways that allow researchers to detect large-scale trends, patterns, and relationships that are not discernable from a single text or detailed analysis. *see CASE STUDY 2* → 64 | 65

Rather than pitting distant reading against close reading, what we are seeing is the emergence of new conjunctions between the macro and the micro, general surface trends and deep hermeneutic inquiry, the global view from above and the local view on the ground. The digital humanist is capable of “toggling” between views of the data, zooming in and out, searching for large-scale patterns and then focusing in on fine-grained analysis. While distant reading may be “new” insofar as computational techniques are involved in sifting through, organizing, and visualizing

multitudes of data, it is worth remembering that humanities scholarship has long oscillated between and sometimes even conjoined these two approaches. After all, census data provide an overall statistical picture of demographics but tell nothing of the individuals who live in a given census tract; it is the task of oral histories, biographies, and psychological analyses to delve into the depths of the self. Similarly, a view from above in Google Earth allows a researcher to quickly pan over large regions of the Earth in order to discern surface structures of the built and natural environment, as well as overlay imagery and data sets such as National Oceanic and Atmospheric Administration satellite photographs of areas affected by storms, geographic information system (GIS) data relating to demographics, traffic trend data, and so forth. But these data become more meaningful when yoked with the stories of the people who actually live and have lived there, allowing researchers to not only “skim over” the surface of the Earth but also “drill down” into the micro-level temporal layers comprising the histories of every neighborhood, block, and home. Radically innovative approaches to mapping could emerge from within the Digital Humanities to create environments for exploring differential geographies and delving into heterogeneous geospatial representations, beyond simply registering the phenomenological aspects of space on conventional maps. It remains a challenge how to conceive, design, and implement such platforms. *see* CASE STUDY 1 → 62 | 63; CASE STUDY 2 → 64 | 65

CULTURAL ANALYTICS, AGGREGATION, AND DATA-MINING

parametrics cultural mash-ups computational processing composite analysis algorithm design

The field of cultural analytics has emerged over the past few years, utilizing the tools of high-end computational analysis and data visualization to dissect large-scale cultural data sets. Such data sets might include historical data that have been digitized, such as every shot in the films of Vertov or Eisenstein, the covers and content of every magazine published in the United States in the 20TH century, or the collected works of Milton, not to mention contemporary, real-time data flows such as tweets, SMS text messages, or search trends. Based on assumptions that meaning, argumentation, and interpretive work are not limited to the “insides” of texts or necessarily even require “close” readings, cultural analytics proposes that computational tools be used to enhance literary and historical scholarship. But creating models, visualizations, maps, and semantic webs of data that are simply too large to read or comprehend using unaided human faculties brings other questions. What parameters are used to incorporate cultural artifacts into data sets? Any

conclusions based on these techniques are necessarily shaped, even determined, by these initial choices (e.g. if the gender categories for a census are only male and female, then how can we assess the percentage of transgendered populations?). Cultural analytics does not analyze cultural artifacts, but operates on the level of digital models of these materials in aggregate. Again, the point is not to pit “close” hermeneutic readings against “distant” data mappings, but rather to appreciate the synergistic possibilities and tensions that exist between a hyper-localized, deep analysis and a macrocosmic view.

Cultural analytics also broadens the canon of objects and cultural material under consideration by humanities scholars: Traditionally thought-of cultural objects are now digitized, marked-up, accessible, and shareable in multiple formats and on a variety of platforms, while “born digital” objects—whether tweets, blogs, videos, Web pages, music, maps, photographs, or hypermedia artifacts that combine many different media types—provide data for analysis and populate new forms of knowledge creation and curation. The “data” of cultural analytics are exponentially expanding in terms of volume, data type, production and reception platform, and analytic strategy, making it all-the-more important that humanists are engaged with the design of algorithms, mining and visualization tools, and archiving techniques that foreground questions of value, interpretation, and meaning.

Aggregation of large-scale amounts of information allows data or files to be merged and then outputted into displays that highlight distinctive features such as data points, clusters, and trends. Structured data lend themselves to this processing, since one might easily take dates, places, quantitative information, names, or other elements from a set of files and create an analysis of its contents. Tracking network traffic, or money flows, or resource depletion, or economic trends works well in aggregate. In text processing, looking at word frequency and use (the n-gram approach) is a way of aggregating information and data. The aggregate subsumes individual instances, extracting information from the whole. Cultural mash-ups often aggregate materials in novel ways that allow digital manipulation to repurpose the sources.

Composite analysis preserves individual elements but uses the patterns among them to show something about the whole set of discrete elements. The information and data remain linked to the individual instance rather than being extracted from it into a larger whole. The affordances of large-scale displays, in which thousands of individual images or artifacts can be shown and accessed creates a composite environment. Use of computational methods to discern patterns among such large corpora is essential, though figuring out what the particular purposes or research

questions are that can be answered by such techniques necessarily depends upon recognizing that analysis and processing follow from the fundamental decisions about what constitutes the data and the ways in which these data are structured.

Finally, data-mining is a term that covers a host of techniques for analyzing digital material by “parameterizing” some feature of information and extracting it. This means that any element of a file or collection of files that can be given explicit specifications, or parameters, can be extracted from those files for analysis. The “mining” of these data often depends on creating a display of the results as statistics, texts, or in an information graphic known as a data visualization. Understanding the rhetoric of graphics is another essential skill, therefore, in working at a scale where individual objects are lost in the mass of processed information and data. To date, much humanities data-mining has merely involved counting. Much more sophisticated statistical methods and use of probability will be needed for humanists to absorb the lessons of the social sciences into their methods. *see* CASE STUDY 2 → 64 | 65

VISUALIZATION AND DATA DESIGN

data visualization mapping information design simulation environments spatial argument
modeling knowledge visual interpretation

In recent years humanists have become increasingly involved in what is often referred to as the “visual turn” in scholarship, sometimes correlated with the “spatial turn” that has favored mapping. As digital tools have become prevalent, the interest in “reading” the visual has extended to “authoring” the visual—using visual means to express intellectual concepts. What might it mean to make a visual argument, for instance, or to shape a concept through graphical means?

Currently, visualization in the humanities uses techniques drawn largely from the social sciences, business applications, and the natural sciences, all of which require self-conscious criticality in their adoption. Such visual displays, including graphs and charts, may present themselves as objective or even unmediated views of reality, rather than as rhetorical constructs. Much could be learned from the visual languages and semiotic critiques of art, architecture, and design. Visual special effects, which add more to spectacle than to legibility, are suspect, and information graphics conceived without some professional competence in their design are often unintentionally misleading. By the same token, visualizations designed to specifically address the communication needs of humanities research will only be created if humanists become actively engaged in their design.

The visualization identified here refers primarily to graphical or rendered visual interpretations rather than photographs or films, which have their own unique history and rhetorical qualities. Visualization is intellectually distinct from illustration, the employment of a graphical feature, photograph, map, or other representational device to elucidate, explain, or show something in a text. In the latter case, the text still assumes priority, and the illustration is meant to summarize an argument, provide a reference point, or corroborate the text. While visualizations may illustrate data through processes of aggregation and distillation, visualization in the Digital Humanities takes several different forms, all of which are arguments in themselves and must be evaluated in terms of the rhetorics of information design and display.

The use of graphs, charts, diagrams, and other visualization techniques is often associated with data visualization, the expression of quantifiable or quantitative information in graphic form. But the models of statistical expression, such as bar and pie charts, came from the world of 18TH century “political arithmetic” and provided a rich and much developed legacy that extended the vocabulary of much older visual forms of diagrams, grids, and trees. Business, governments, and administrative organizations all made use of these forms to express quantitative analyses in legible formats. Informational and statistical visualizations engender the rhetoric of clarity, precision, and fact, though they are, of course, constructed interpretations. When done well, they can make persuasive visual arguments, allow something new to emerge, or even be subverted for poetic effect.

Visualizations of data that are produced computationally tend to be derived from large-scale data sets such as social networks, digitized corpora, and demographic data. The visualizations, either custom-built (e.g. network analysis diagrams) or created for use in an online environment (e.g. Many Eyes), may be used as analytical and interpretive tools—to reveal patterns or anomalies or concurrences—or they may be produced to illustrate findings or serve as the distillation of an argument. Of course, the structures of the data and the questions that are asked of them will, inevitably, determine the visualization produced and the answers obtained. Perhaps it is of little surprise, then, that data visualizations tend to take the established forms mentioned earlier—charts, diagrams, grids, or trees—although we are increasingly immersed in a world of graphical possibilities that have yet to be realized. *see* CASE STUDY 3 → 66 | 67

Mapping is a distinct form of visualization built on the history of cartography; ideologies of discovery, ownership, and control; levels of abstraction; scale; relations between the real and representation; symbology; visual signposting; perspective;

and coordinate systems. Mapping in the humanities ranges from historical mapping of “time-layers” to memory maps, conceptual mapping, community-based maps, and forms of counter-mapping that attempt to un-ontologize cartography and imagine new worlds. In the 1950s, the members of the experimental Situationist group developed an approach to experiencing urban spaces that they termed “psycho-geography.” The immersive and experiential wandering advocated by the approach gave rise to a handful of maps that suggested flows and movement through space as a ludic, exploratory exercise that could result in a new critical awareness of urban environments. Similarly speculative, cognitive maps are used to model experience in many domains of human life where qualitative properties are given dimension and formal value in visual form without any need to ground them in quantitative information. *see* CASE STUDY 1 → 62 | 63

A human life may have many such experiential dimensions in which affective properties shape the intellectual argument and give rise to a graphic form that shows the size, scale, proportions, orientation, direction, or figurative shape of knowledge unfolding over time. Maps, animations, and visual images from the vast inventory of human imaginings have much to offer to contemporary scholars re-imagining their own concepts of intellectual argument. The pictorial conventions of visual representation may well be repurposed, just as the organization of cabinets of curiosity, antiquated libraries, personal spaces of study, commonplace books, or other instruments of memory, argument, and rhetoric are finding their place again within the broader understanding of how we produce and represent intellectual arguments and model knowledge.

Experiential visualization uses movement through the time and space of a three-dimensional world as the primary mode of engagement. Historical simulation environments can take a viewer into an immersive environment (or, at least a virtual one), creating the experience of walking through, for example, a Chinese farming village during the Han dynasty or the Roman Forum in late antiquity. Historical simulation environments don’t represent the past “as it really was”; instead, they foreground interpretation, analysis, and experimentation, allowing new research questions to be asked and hypotheses to be tested using a wide range of variables. For instance, one may employ time-sliders to visualize when and where certain buildings came into existence or to investigate kinetic aspects of events in time-space environments, such as parades, funeral processions, orations, and protests. Experiential visualization is not a simple mimesis or positivistic reconstruction of a historical reality, nor is it a simple augmentation of a real-world site, but rather an investigation of a state of knowledge. *see* CASE STUDY 4 → 68 | 69

Visualization can be used in many other ways to sketch out an argument or to map its constituent parts or even to model its initial formulation. Visualization has the power to unleash imaginative and conceptual potential. By identifying elements of a system and thinking about how they relate to each other sequentially, or hierarchically, or relationally, humanists discover ways of modeling knowledge that were not part of their textual training. As with so many aspects of digital work, the strengths of these techniques are amplified when they are in dialogue with, rather than opposed to or exclusive of, traditional methods. The use of visualization or distant reading, for example, in concert with attention to individual texts, or aggregation techniques in dialogue with studies of outliers and anomalies can provide valuable contributions to the discussion of meaning-production that could not be obtained using only one or the other method. Knowing what to read and visualize as well as how to read and visualize forms is at the basis of digital literacy and the assessment of meaning in these new formats. *see* CASE STUDY 5 → 70 | 71

LOCATIVE INVESTIGATION AND THICK MAPPING

spatial humanities digital cultural mapping interconnected sites experiential navigation
geographic information systems (GIS) stacked data

Traditional scholarship in the humanities moved among a few select sites for research and teaching: the library, the archive, and the classroom. The “holdings” of the library were just that: holdings held for the initiated who had the privilege of access and use. Scholars made pilgrimages to special collections to view artifacts or read rare books, often examining these objects under guarded conditions that were established to limit access and thereby preserve the safety—and aura—of the original. The seminar room or the lecture hall was the primary site for the transmission of knowledge mostly in a single direction: professors professed knowledge and students consumed it.

Today, the boundaries of the library, the archive, and the classroom have become more porous, interconnected, and globally extensible. Countless new sites for knowledge creation and dissemination have emerged, bringing scholarship into communities and communities into the academy. Libraries have allowed millions of their volumes to be digitized and have opened up their collections via Web services, making them available to the digitally enfranchised public. Some archives are following suit, removing physical and virtual walls that once restricted use. The traditional learning space of the classroom has been rethought in ways that promote interactivity, discovery, and co-creation, often through real-time feedback

mediated by social technologies that blend physical and virtual spaces. Courses have been taught, for example, in virtual worlds with avatars participating from around the globe, connecting the physical space of the classroom with the infinitely expansive and fluid realms of cyberspace. Webcast or web-linked teaching is now commonplace.

But a few caveats are in order. The networked world is a patchwork, very much marked by social and economic inequalities; access and participation are hardly open to everyone. Different zones are governed by distinct attitudes toward cultural property, licensing, and pressures of sustainability. Cost-recovery models exist even in the most elite sectors while many individuals and communities have limited connectivity. Just as in the realm of bricks-and-mortar education, inequities abound. Diverse and competing interests will continue the struggle for control.

As the contours of scholarship are undergoing a fundamental remapping through collaborations in which researchers can curate, narrate, annotate, and augment physical landscapes, the boundaries of inside and outside have become fluid. The interior realm where curators make arguments in space through the meticulous staging of physical objects, supported by labels, wall text, and installation architectures can now be enriched by media that draw the outside world into the gallery. Likewise, visitor itineraries can now be extended out into the surrounding landscape in ways that apply traditional curatorial skills to the shaping of paths through the physical world. Data landscapes can be curated in the physical space of a city, allowing a user with a GPS-enabled mobile device, for example, to listen to geo-coordinated soundscapes curated by musicians while walking down a sidewalk or to follow in the footsteps of the dead and hear stories told by generations of immigrants about a neighborhood. Such locative investigations bring together the analytical tools of geographic information systems (GIS), the structuring and querying capacities of geo-temporal databases, and the delivery interfaces on GPS-enabled mobile devices. *see CASE STUDY 5* → 70 | 71

This attention to place has resulted in the emergence of a significant sub-field of the Digital Humanities variously called “Digital Cultural Mapping” or “Spatial Humanities.” It is here that geographic analysis, digital mapping platforms, and interpretive historical practices come together to form richly textured, multidimensional investigations of place. Unlike conventional approaches to mapping, which tend to be positivistic and mimetic, these practices of thick mapping in the Digital Humanities place a primacy on experiential navigation, epistemologies of representation, and the rhetorics of visualization. After all, a map is a visualization or

representation of a group of relations (and structuring assumptions) that present a state of knowledge. The map may or may not have a referent in the “real world,” but it does make an argument, and in the digital realm it becomes an interactive site for creating, representing, and navigating knowledge. Digital maps are essentially navigable layers of spatial data rendered visually, ranging, for example, from demographic and census data to location-specific video histories, Twitter streams, and historical map layers.

Such maps are not meant to be static representations or accurate reflections of a physical reality; instead, they function as stacked representations in which one representation is linked or keyed to another. Within a dynamic, ever-changing environment, new data sets can be overlaid, new annotations can be added, new relationships among maps can be discovered, and, perhaps most importantly, missing voices can be returned to specific locations through “writerly” projects of memory that the participatory architecture of Web 2.0 applications has made possible. Thick mapping thus enables an unbounded multiplicity of participatory modes of storytelling and counter-mapping in which users create and delve into cumulative layers of site-specific meaning. Far from the Apollonian eye looking down from a transcendental view, thick maps betray the contingency of looking, the groundedness of any perspective, and the embodied relationality inherent to any locative investigation. *see CASE STUDY I* → 62 | 63

THE ANIMATED ARCHIVE

user communities permeable walls active engagement bottom-up curation multiplied access
participatory content creation

Derived from ancient Greek ἀρχεῖον (*government*) and the late Latin word *archivum*, the English derivative *archive* now refers not just to public administrative records but also to the entire corpus of material remains that the past, whether distant or close, has bequeathed to the present: artifacts, writings, books, works of art, personal documents, and the like. Its semantic field also encompasses the institutions that house and preserve such remains. In all of these meanings, archive connotes a past that has severed its ties with the present and has entered the crypt of history only to resurface under conditions of restricted access.

The Digital Humanities offers new challenges and possibilities for institutions of memory such as archives, libraries, and museums: process-based concepts of “living” archives of the present; approaches to conservation and preservation based upon multiplying (rather than restricting) access to the remains of the past; participatory

models of content production, research, and curatorship bringing together professional and citizen scholars in team-based projects that interpret the cultural patrimony as a public good; augmented approaches to programming and informal education that promise to expand traditional library and museum audiences and bring scholarship into public view; and enhanced means for vivifying and promoting active or experientially augmented modes of engagement with both the past and the present. Of course, the past was never really past; it always already belonged to the present. And digital toolkits and the expanded compass of humanistic scholarship provide some distinctive avenues for investing the present's stewardship of the past with the attributes of life. They hold out the promise of animating the archive.

Accumulation is no longer enough to ensure the survival of the cultural patrimony. Objects that sit in storage, though they may have a potential afterlife, disappear into the ever-expanding heap of cultural remains, entering a limbo that in no essential way differs from being lost. So the "animation" of archives stands for a series of strategies for launching that afterlife from the very moment of archival processing. This implies a user-centered approach to the construction of archives that builds a multiplicity of use-scenarios into the very architecture of the archive; breaks down partitions between collections and bricks-and-mortar institutions (through, for example, open application programming interfaces); engages real or potential user communities from the outset (in processing, tagging, and metadata development); and integrates curatorial and content-production tools into access portals. *see CASE STUDY 3* → 66 | 67

Embedded within the constellation of possibilities just evoked is a sort of Copernican revolution with respect to the roles performed by libraries and museums in the modern era. New conjugations of inside and outside, scholar and citizen, curator and viewer are emerging, with social technologies challenging conventional ideas of ownership, restricted use, storage and display, content creation, and curatorial control. With the shift in focus from data retrieval—essentially "top-down billboarding"—to bottom-up working and reworking of content, whether in the form of texts, still or moving images, audio, or other media, every library and museum becomes adjacent to a public square as big or as small as they choose. It also marks the beginning of an inversion which some will welcome and others will decry. Whereas the virtual was once subordinated to and cast in a supporting role with respect to the physical, we are now seeing new couplings in which an institution's virtual footprint may exceed its physical edifice and the community that it serves may be worldwide, overlapping only in small part with potential or actual physical visitor/user populations. This is one of the great opportunity spaces that

the Digital Humanities opens up, giving archivists, librarians, and curators a chance to not simply enlarge but completely re-envision their communities, publics, and missions. Every public institution has already been transformed into a glocal enterprise, local and global at the same time. Glocalization will only accelerate over the coming decades. *see* CASE STUDY 4 → 68 | 69

In sum, the memory palaces of the 21ST century will have much more permeable walls than their 19TH and 20TH century predecessors. This is also to say that they will be much bigger both from the standpoint of the physical territory that they cover and the corpora of information that they harbor. For example, the Digital Humanities harnesses the expressive power of worlds like Google Earth and three-dimensional virtual environments, and deploys the ever-increasing availability of wireless bandwidth to interact with ubiquitous computing devices equipped with GPS technologies that can calculate and annotate embodied, physical locations within inches. This is the future of knowledge, where culture and social and political practice will emphasize embeddings of the virtual within the real, where actual physical landscapes will be curated just as if they were an art gallery, and where we will be surrounded and enveloped by the collaborative and distributed building of annotations on, and overlays of, the physical world. This is a future that is already with us. The challenge for scholarship and institutions? To build platforms and collections out into these and other domains of intersection between the virtual and the physical in ways that reinforce not only access and outreach but also establish new models of imagination, quality, and rigor.

DISTRIBUTED KNOWLEDGE PRODUCTION AND PERFORMATIVE ACCESS

global networks ambient data
collaborative authorship interdisciplinary teams use as performance crowd-sourcing

The myth of the humanities as the terrain of the solitary genius, laboring alone on a creative work, which, when completed, would be remarkable for its singularity—a philosophical text, a definitive historical study, a paradigm-shifting work of literary criticism—is, of course, a myth. Genius does exist, but knowledge has always been produced and accessed in ways that are fundamentally distributed, although today this is true more than ever. It is not uncommon for dozens of people to work on a Digital Humanities project, each contributing domain-specific expertise that enables a research question to be conceptualized, answered, and then re-conceptualized and re-answered. A team of database developers and data management experts may come from a school of information sciences, while interface designers

may come from the arts, content developers may come from history departments, and coders may come from the computational sciences. Each member of the team works with the technical lead and project director who collaboratively articulate the technical and functional specifications for the project. In the end, when a project is deployed, there may be dozens of “authors”—ranging from professors and librarians to student programmers, interns, staff, and community members—who contributed to its development. Some long-term projects are the work of generations of students and scholars. Distributed knowledge production means that a single person cannot possibly conceive of and carry out all facets of a project.

Analogously, distributed access means that the audience for the project can engage with its content via multiple access points and platforms. In fact, every engagement is a performative instantiation of knowledge. With the surge of mobile devices and distributed computing, ambient networks present new possibilities for accessing information and interacting with knowledge. While we access most digital information on a screen, the means by which information circulates to find its connection to those screens is distributed across wired and wireless networks, with data shuttled seamlessly between the cloud and our local machines. Though our perception of them is limited to display devices, data streams fill the air. Their presence in and among the many other features of the physical environment makes their integration into lived experience a possibility. Human and cultural knowledge will interpenetrate the natural and built environments with increasing degrees of saturation. Access to interpretive materials, cultural history, geographical and geological knowledge, historical dimensions, narrative facts, biographical information, and the stories of events lived and experienced in our shared spaces will be a way to enhance the engagement with the real. Or, in other instances, they may provide solace, consolation, companionship, and fellowship through communicative exchange. What is certain is that knowledge production, access, and dissemination are becoming ever-more distributed processes across high-speed, mobile networks that operate seamlessly at all levels. *see* CASE STUDY 4 → 68 | 69; CASE STUDY 5 → 70 | 71

When knowledge exists in iterative form across global networks and local access points, with many versions and expressions of cultural information taking shape in a process whose life cycle is ongoing, then any access to that knowledge is a performance, an instantiation. Just as any reading of a book or a script or any viewing of a film or any playing of a score is a performance of that work, the same is true of digital works. In fact, every use of a file is different; no two files are ever the same, and the very act of opening and displaying a file is a performance of a work,

a unique instantiation in historical and social space. The difference between the performative reading of a work in an analog world and in a digital one registers dramatically how the lines between reading and authoring blur. When the material substrate records the performative variations of each instantiation, then the act of reading or viewing contributes materially, not just virtually, to the work. When and how such traces will be recorded has yet to be seen and will constitute fertile ground for research and publishing in the Digital Humanities, but the possibilities for crowd-sourced engagement with editing, proofreading, translation, and critical assessment are bringing this process into view. *see* **CASE STUDY 3** → 66 | 67

HUMANITIES GAMING

user engagement rule-based play rich interaction virtual learning environments
immersion and simulation narrative complexity

Imagine being on the streets of a South African township as it explodes in violence after the apartheid-era government switched the language of education from English to Afrikaans. This is the experience Hamilton College students have when they play the immersive game *Soweto '76*, one that deepens empathy and enlivens class discussions of race, power, and education. At Dartmouth, students compete furiously against each other to tag the materials they find in online archives. When these students play *Metadata Games*, they are encountering an open-source project that uses the affordances of gaming to build more robust archival data systems. King's College London students create avatars in *Second Life* and then reconstruct historical stages from the classical Roman Theater of Pompey to Shakespeare's Globe Theatre in London. At Duke and other participating universities, students play "Virtual Peace," a collaborative simulation game in which players analyze complex situations posed by international crises in order to learn how to make effective decisions. Digital Humanities gaming has begun to successfully engage with historical simulation, virtuous cycles of competition, and the virtual construction of learning environments.

But games of any type have never been held in high repute by academia. Relegated either to athletic departments as mere sports, or to the realm of leisure time as diversions, they have only recently begun to be taken seriously as both an object of study and a career for which to be trained. Gaming demonstrates a capacity that could transform Digital Humanities pedagogy. This is due to many factors, but two in particular stand out. The first is the explosion of processing power and connectivity. Not only are game-world simulations compelling visually and interactively,

they are also capable of functioning in real time with multiple participants spread around the world. This braiding of capacity and reach, made possible by ever-increasing processing speeds, the ubiquity of networks, and mobile connectivity, yields highly engaging forms of immersion and simulation. This rich interaction can be yoked to any content, from the expected adrenaline thrills of first-person shooters like Halo to detailed alternate reality games (ARGs) like World of Warcraft. Given the fungibility of content and the consistency of user engagement with well-designed games, “humane” and “serious” games are likely to keep pace with technological advancement. The second fact to consider here is the acculturation of a generation of students who have literally grown up gaming. They value interactive programs that engage their attention while at the same time deepening their understanding of meaningful subject matter.

But what exactly do such students expect and what constitutes a successful educational game? Games are rule-based. They offer copious feedback. They are essentially voluntary, running on enthusiasm and begging for engagement. Games are also quintessential delays of gratification: Give players the freedom to achieve their goals in the quickest, most rational way and satisfying game play withers. It is the obstacles overcome and the levels mastered, the reward for tasks accomplished and the rules obeyed which constitute the satisfaction of play.

Recent developments in new-media studies and narratology have removed some of the stigma that was once attached to gaming within the academy, but digital games are still considered by many in the humanities as frivolous (and monstrously violent to boot). It becomes harder to maintain this perspective as the narrative complexity, play strategy, and game “feel” (as developers call the gestalt of gamer-and-game interaction) become more developed, culturally significant, and even world-enriching. As we have seen, games in the Digital Humanities already exist that are exploring interactive models of learning and ways of critically grappling with the human experience. The challenge for the future is to take the gamesmanship of humanities research—its pursuits and pleasures, competitive drive, and seductive engagements—as the basis for games of scholarship. *see CASE STUDY 4* → 68 | 69

CODE, SOFTWARE, AND PLATFORM STUDIES

narrative structures code as text computational processes software in a cultural context encoding practices

Code studies, along with the related study of software and platforms, bring humanistic close-reading practices into dialogue with computational methods. The operations of computational media are created through the interaction of hardware and software. These work according to protocols structured into their organization as code. The study of code is driven by an interest in exposing the ways constraints make certain things possible, and exclude others. But is code a text? If so, what kind of text? Should we assess the aesthetic properties of code the same way we discern the value of any other artistic composition? Or should we condemn code work as mere craft or technique? Debates are heated, with passionate partisans on all sides. The alphanumeric system is already a code, so the heralding of a “new” field of code studies may seem inflated.

Code and software scholars begin their study with the history of encoding practices, in particular those methods that make an operation happen, such as the punch cards used to set the patterns of weaving on Jacquard looms or the programs in early computers enabled by stacks of cards whose punched openings allowed circuits to be blocked or completed. The basic binary language of digital media is the foundation of all programming code, but software and computer languages have their own history as forms with grammar and syntax. The study of software traces developments from switch settings on mid-20TH century mainframe computers to the creation of the assembly and compiling languages that underpin many of the scripting languages and much of the object-oriented software written today. The layers of software between the operations of a machine and the instructions given it by an operator offer a fascinating archaeological study, with cultural conventions often holding as much weight as technological advantage.

Scholars fascinated by the encoded protocols and instructions that constitute the language of software also look at the cultural contexts in which business, defense, or communications industries fueled the development of increasingly sophisticated approaches to encoding. The algorithm, a set of step-by-step instructions, is the heart of software programs, but these instructions have to be translated into a binary language that the computing hardware understands. The organization of processing units, the workflow cycle through circuits and transistors, the use of active buffers and parallel processors—all of these pieces of hardware interact with software in particular ways that have affordances and hindrances that vary from platform to platform. Critical approaches allow understanding of these elements as objects of study, almost as if one were reading them as text.

A particular fascination with game engines and their narrative structures fuels one area of code studies. The analysis of narrative and multi-player activities in a complex set of chained and interdependent interactions requires chunking of game elements at critical nodes or decision points. The ways this is achieved is itself a complex process—a game of sorts—in which the skills of narratology meet the worlds of probability and possibility in a combinatoric universe that must move seamlessly from one moment of illusion to another. The way this is engineered and designed elicits a fascination akin to that of expertly constructed aesthetic artifacts in any other era, such as novels or plays. Likewise, the engagement of a substantial literary community with the poetics of code has created a body of critical work that addresses the aesthetics of programming in its own right. All of these approaches offer analytic engagement with computational processes as forms of composition, exposing their complexity through careful reading, construction, and attention to structure. *see* CASE STUDY 5 → 70 | 71

DATABASE DOCUMENTARIES

variable experience user-activated multimedia prose modular and combinatoric multilinear

Digital Humanities genres include multimedia critical editions; interpretive work with expanded data sets published alongside their interpretive outcomes; conjugations of the digital and the physical, the desktop and the streets; and expanded definitions of knowledge that exercise not just sight, but the entire human sensorium. Within this set of emerging composite forms, the database documentary occupies a central position. It is a genre that has continued to evolve in dialogue with shifts in the technology of interactive media.

Cinematic documentaries work with image and sound materials that, however mediated or massaged, claim an indexical relation to the world. That is to say that they work with “real-world” materials captured, filtered, and threaded into a linear narrative artifact in the medium of film. In order to craft such a linear narrative, large amounts of footage must be shot as part of the research and development process. By necessity, most of this footage must be thrown out or reduced to a few choice sequences, given that a small core of materials must make up the story’s backbone. Only one story can be told well, even when the intended single “story” turns out to be a densely wound skein of stories, each overlapping with the next.

The database documentary also works with materials of documentary value, but on an expanded scale. Database documentaries are modular and combinatoric, branching and hypertextual, often structured more like a multimedia prose piece than

DATABASE DOCUMENTARIES
REPURPOSABLE CONTENT AND REMIX CULTURE

a film. Consisting of a series of tracks through an actual or virtual database, the documentary can be built out of a wide range of media types: not just film and video, but also sound, static image, text, animation, actual documents (or their digital equivalents), even live or dynamic feeds from the World Wide Web. Database documentaries are multilinear. They are not watched, but rather performed by a reader/viewer who is provided with a series of guided paths; and, unlike the cinematic documentary, which is free-standing, database documentaries may be built on multiple, overlapping databases. Or they may even consist only of pointers that send out calls, through open APIs, that retrieve materials hosted externally. The paths are reversible, allowing for trackbacks to the sources from which individual documents are drawn and/or to external resources. Inclusions as well as exclusions can be exposed to view, thereby creating an experience that is dynamic, active, and user-centered. Temporal sequence, duration, and sound levels, as well as the presence or absence of elements of the critical apparatus are firmly in the reader/viewer's control. *see CASE STUDY 4* → 68 | 69

The multilinear character of database documentaries creates a different series of opportunities and challenges with respect to cinematic documentaries. Given that multiple intersecting story lines are present in database documentaries and that they are user-activated, a far greater fluidity of movement and pacing must always be presumed, much as in the case of visitation paths through physical exhibits. Conferring unity upon such a variable experience can be difficult, as can the building of cohesive story lines. This said, the possibility of marshaling crisscrossing sets of data to tell interrelated stories offers powerful new modalities of scholarly argument as well as imaginative expression. The database documentary remains one of the most venerable of new media forms, with early expressions such as the Interactive Cinema Group at MIT in the late 1980s, the Labyrinth project at the University of Southern California, and the pioneering work of United Kingdom-based Blast Theory group in the field of so-called “live” documentaries.

REPURPOSABLE CONTENT AND REMIX CULTURE

participatory Web read/write/rewrite platform migration sampling and collage meta-medium inter-textuality

The ease with which content can be repurposed in digital form extends the capacities of the medium to function as a meta-medium. Photography has that property, with its ability to record and reproduce drawing, painting, printmaking, and other visual formats. Now, the digital environment serves as the simulation machine that is able to re-create and imitate other formats. But it also allows content to be

migrated from platform to platform, to be used in a variety of outputs and for a range of readers and forums. Figuring out how to write texts in a modular manner that will allow them to be recombined for different levels of interest and readership as well as different degrees of detail and granularities of argument (not to mention output and display device) is still a challenge—and represents another fertile field for the Digital Humanities to explore. The realization that print on-demand and online access are complementary modes to traditional print rather than competitive ones is already well-recognized by the publishing community as well as readers and authors. Artists also engage opportunistically with the possibilities of different venues and formats, so that their range of expression might include gallery works that are unique, printed versions for larger distribution, and online exhibits of the same work to reach yet another audience. The work is a distributed effect of each and all of these aspects rather than being limited to any single part of this continuum. *see CASE STUDY 5* → 70 | 71

Remix culture is a hallmark of the participatory, programmable Web in which a “read-only” ethos has been surpassed by one of “read/write/rewrite.” In much the same way that early textual scholarship used citations and annotations to extend authority on copied manuscripts, remix culture uses digital sampling and collage techniques to create derivative original works with a complex trail of associations, inter-textual references, and critical trajectories. Authorship is multiplicative and dissemination happens across the Web as others add to, borrow, remix, and republish the work. Best known in music, remix culture extends to photography, film, graphic design, software development, data curation, and many other realms. In essence, with the tools of both production and consumption in the hands of the public, an ever-expanding space of design and curation allows bits of data and intellectual property to move and be remixed in creative ways. *see CASE STUDY 4* → 68 | 69

The university, however, still places a primacy on the singular nature of originality of scholarship and on clear lines of demarcation for authorship. In fact, the institutional structures for generating, evaluating, and legitimizing knowledge have barely embraced repurposable and remixable intellectual culture. Perhaps this is because the institutional frameworks in which this knowledge is produced and evaluated have hardly changed over the past century. What if departments could be remixed as easily as digital music samples? What if curricula had life cycles like software? What would an open-source humanities division look like? For one thing, disciplines, departments, and administrative structures would receive date stamps and would need to innovate in order to survive.

The objections are, of course, easy to mount: Without the long-term stability of a department, how can we prepare students for a field? How can we be sure that they have learned “the content” of a discipline, and how can we possibly credential students with degrees if they are participating in departments that will no longer exist in a matter of years? These objections, we believe, are based on assumptions that have traditionally valued “the what” (a determinate and relatively static set of knowledge objects or canon of artifacts) over “the how” (a flexible—even nimble—mode of thinking that privileges design, experimentation, risk-taking, and creative problem-solving). This is not to say that knowledge in a field is irrelevant, for the contrary is true. It is to say, however, that universities will serve their students best by credentialing the skills necessary to creatively conceptualize and solve problems: a knowledge base grounded in making and experimentation, and a social disposition that fosters collaboration with diverse partners. It is here that the core values of the humanities and the generative potential of the digital come together in the poiesis of world-making.

PERVASIVE INFRASTRUCTURE

extensible frameworks heterogeneous data streams polymorphous browsing cloud computing

With the emergence of standards-compliant Web services and dynamic cloud computing, massive data sets can be shared and accessed across networks. Web services are essentially machine-to-machine communications that allow various types of data to be accessed through specific queries. For example, a map Web service might allow a user to access census data or historical maps stored on one server from any computer able to send the appropriate query to the service; users will not receive the “actual” maps or entire copies of the data but rather access to the maps and data through calls to the service provider. The data can, then, be rendered and viewed in various interfaces, such as on a Web page, in a geo-browser, or in another visualization application. Cloud computing provides an (almost infinitely) extensible framework for massive data storage, access, and retrieval from any computer connected to the network. The metaphor of the cloud signifies the seemingly ethereal data that can be pushed and pulled through the sky, but in reality it translates into mega data centers, storage systems, and networked Web architecture to facilitate data exchange.

What does this mean for the Digital Humanities? Foremost, it means that it is now possible to share the entire data sets of research with the scholarly community and the public-at-large. In disciplines such as anthropology, archaeology, and classics,

researchers may produce millions of discrete data points over the course of a project, ranging from survey and excavation data to fieldwork documentation through integrated geographic information systems. Rather than summarizing the results of a project and drawing conclusions, researchers can make the entire data set available online, enabling other users to test hypotheses and even to add to and edit the “original” data set and accompanying metadata. Openness has benefits, but caveats about validity of data, privacy, misappropriation, and other ethical concerns are also in order. Secondly, through polymorphous browsing, users can access, manipulate, and analyze massively heterogeneous data streams, following trails of association that lead out and go deep. What this means in practice is that search and discovery tools are able to identify, aggregate, and integrate data from completely disparate sources across archives, libraries, and repositories and present these data in ways that are customizable for the needs of a given researcher. One can expand and contract, tier out and drill down through a portal that can access the world’s information regardless of where it resides. For scholars of literature, for example, it means having access to every word in every edition of every book ever published and customizing a search to answer a research question that, recursively, becomes part of the data of the system itself. *see* CASE STUDY 3 → 66 | 67; CASE STUDY 4 → 68 | 69

UBIQUITOUS SCHOLARSHIP

augmented reality web of things pervasive surveillance and tracking ubiquitous computing
deterritorialization of humanistic practice

As these emergent genres and methods illuminate, the forms that knowledge assumes can no longer be considered givens. The tools of humanistic inquiry have become as much objects of research and experimentation as have the modes of production and dissemination of knowledge. Statistical methods press against one edge of the qualitative human sciences; graphic and information design press up against another. Real time, massively participatory role-playing games create another force field exerting influence from the arts and gaming worlds. Laboratories arise with a collaborative, team-based ethos, embracing a triangulation of arts practice, critique, and outreach as they merge research, pedagogy, publication, and generative practices. The once-firm boundary lines among libraries, museums, archives, and classrooms have become increasingly porous as scholarship, no longer limited to print and the lecture hall, has started to shuttle back-and-forth between the stacks and the streets.

Location-aware smartphones and other mobile devices have a key role to play in this deterritorialization of humanistic practice. Thanks to their ubiquity, it has

become possible to couple Web-based knowledge resources to physical locations in ways that would have been hard to imagine only a decade ago. This means that scholarship in fields such as history, urban studies, architecture, art, design, and literature can now curate, narrate, annotate, and augment the physical landscape with a multitude of Web-based archival sources. Such scholarship speaks to multiple audiences and leads multiple lives. A website may be remixed as an electronic publication for use on location-aware mobile devices and later become a print artifact; the website and mobile “edition” can be built for further curation and extension on the part of end-users who can embed datascares anywhere, at anytime. Augmented reality applications allow mobile devices to combine geolocation information and enhanced imagery in a layered, site-specific presentation of events and interpretations. Imagine a time-machine application that shows your neighborhood in a fast-forward sequence from Jurassic times to the present; or think of sensors in a natural environment that expose the geological and industrial processes that formed what is before your eyes; or consider simultaneous and automatic translation applications that remove linguistic barriers to signage and information in a foreign script; or imagine the “web of things,” in which every physical entity—from the book in your hands to your hands themselves—is connected to and part of a deeply recursive information network. The growth of telecommunication and information technologies has transformed the tactical strategies for activism, protest maneuvers, community-building, and relating in the public sphere; and, at the same time, it has also transformed how we know the world, interact with one another, and generate what counts, at a given moment, as knowledge. The natural, social, and cultural worlds are interpenetrated by ever-denser technological systems and data landscapes. We live intensely intermediated lives. *see CASE STUDY 5* → 70 | 71

Ubiquitous computing, as visionary Xerox PARC researcher Mark Weiser argued, is computing that has essentially gone “invisible” precisely because it has embedded itself “in the woodwork everywhere.” Ubiquitous computing—everywhere, at anytime, in everything—is possible only when high-speed networking capacities and interoperable standards allow for constant, seamless, and infinitely deep levels of information-sharing among data centers, computers, mobile devices, physical objects, and people. Nothing exists in isolation but rather in ever-denser networks of interconnection. Of course, ubiquity has a dark side: pervasive surveillance and tracking, the colonization of everyday life by information technologies, the quantification of the biopolitical sphere into ever-smaller units of analysis and monitoring, the inability or incapacity to “de-link” or “opt-out” of these technologies.

But ubiquity also allows for the massive expansion of the scholarly enterprise through a wealth of networks, information streams, and emergent communities of practice that produce and share knowledge and culture in ways that open up opportunities for participation, dissension, and freedom. Ubiquitous scholarship is marked by an ethic of collaboration and interconnection on levels that move (almost effortlessly) between the global and the local, the library and the public square, the pen and the smartphone, the millennia-long histories of humankind and the real-time feeds of the now.

The fictional case studies that follow draw from existing projects, but are themselves imaginary, offered as descriptive rather than prescriptive models for building teams, assembling the necessary resources, and launching Digital Humanities projects into the world. The case studies provide a framework for grappling with these new domains of humanistic practice.

A PORTFOLIO OF CASE STUDIES

CASE STUDY 1

MAPPING DIFFERENTIAL GEOGRAPHIES

CASE STUDY 2

EXPANDED PUBLICATION OF A TEXTUAL CORPUS

CASE STUDY 3

AUGMENTED OBJECTS & SPACES

CASE STUDY 4

VIRTUAL RECONSTRUCTION

CASE STUDY 5

MULTI-AUTHORED LOCATIVE INVESTIGATION

CASE STUDY I MAPPING DIFFERENTIAL GEOGRAPHIES IN THE NEW WORLD ENCOUNTER

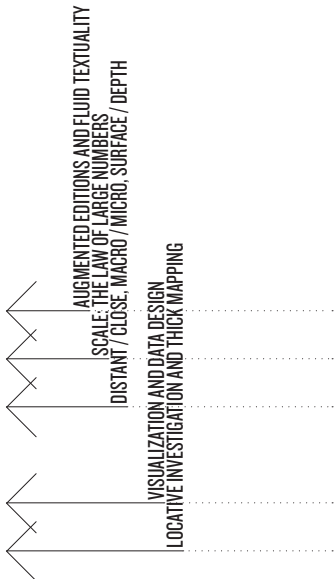
In this cartographic project, techniques of thick mapping are used in combination with text analysis, data-mining, and large-corpus natural language processing. The extended project uses a participatory architecture to support annotation, debate, and repurposing of the cartographic representations and the text visualizations. Microcosmic views of the nomenclature for different geographical features are complemented by macrocosmic views of shifts in the understanding of the shape and boundaries of geographical regions.

Scholarly attitudes toward indigenous concepts of space and geography have changed dramatically in the last two decades through the influence of post-colonial theory. The traditional narratives of “discovery” have been interrogated, qualified, and largely abandoned. Perspectives of indigenous peoples now register within the literature, but little exists of their mapping techniques, world views, and epistemologies. Many Native American techniques for understanding geography were passed on in oral description, in myths of origin and ownership, or were indicated graphically in the most ephemeral tracings of sticks in sand or dirt. Approaches to cartography have undergone their own changes during these decades, away from what historian John Rennie Short characterizes as the story of “increasing scientific rationality” and toward maps as “social constructions, stories marked by purposeful erasures and silences.”

We can now map the encounter of indigenous peoples and Europeans from different cultural perspectives, surveying incommensurate or differential geographies, explicating fundamentally distinct views of land, space, and place. This “mapping” requires careful textual analysis of the production, reception, and critical discourse around key documents in which the dialogue between indigenous and European peoples is evident. It also requires a way to produce simulations and models of a differential geography, one that would arise from the contrast of basic assumptions. For native peoples, rivers and roads were one continuous transportation route, while Europeans thought of them as land features. Europeans were focused on edges and inroads, the coastlines, harbors, bays, and means of penetrating the unknown interior. Indigenous peoples thought of land in terms of extension and activity, seasonal and tribal movements and occupation, with margins determined by social order and priority rather than physical metrics.

This project takes up the question of how, with the meager evidence before us, we can model the contrast between indigenous and European concepts of mapping at the time of early contact. Can digital means be put to the creation of an alternative view of geography and land, of spatial experience, without taking Western perspectives, epistemologies, and coordinate systems as normative? A combination of textual analysis and comparative, critical cartography will be used to explore the concept of differential geography—a mapping of space that exposes incommensurate views—and to model the changes in the historical understanding of the spaces of discovery that became the “New World.”

The project focuses on several key narratives of discovery linked to European maps, some of which relied heavily on indigenous sources of information. The texts to be used include: Christopher Columbus' letters to the king and queen of Spain, accounts of Jacques Cartier's journeys into the Gulf of St. Lawrence in the 1530s, Sir Walter Raleigh's maps and accounts of Guiana from 1595, John Smith's accounts of his capture in 1607 and the map he made in 1612 of Virginia, the Codex Nuttall, Philippe Buache's *Carte Physique de*



Terreins les plus élevés de la Partie Occidentale du Canada, printed in 1754, and Aaron Arrowsmith's 1802 map of North America, which was heavily dependent on and acknowledged Native American sources.

These primary texts will be analyzed for their use of indigenous accounts and terminology. We will create a searchable corpus that will allow text analysis of key terms whose use and meaning can be tracked through the reception history of these primary documents in the critical literature. Many of these texts are already in digital form, but they have not been analyzed for this purpose. We propose to track and visualize the changing nomenclature around a cluster of crucial concepts such as space, land, mapping, discovery, contact, nativeness, and other terms to understand how the discourse of indigenous spatial understanding contrasts with that of the Europeans. We realize that the materials for authentic indigenous voices are scant, and almost all are recorded within European texts and documents. This is not a project to recover a lost authenticity, but to analyze the shape of discursive formation.

Methods

Structured mark-up, particularly the textual analysis of terms in context, will produce a study of nomenclature shifts from first contact to the present. The reception history, citation, reuse, and repositioning of Raleigh and Nuttall within the critical literature will be used as case studies since they are long, vertical studies across several centuries of use and discussion. We will also do a lateral analysis of their presence across a corpus of crucial documents, tracking usage and changing characterizations of peoples and vocabularies.

The project poses a number of technical and conceptual challenges. While some of the primary materials are in digital form, others are handwritten manuscripts whose transcription requires specialized knowledge and skill. Nonstandard spellings, shifts in language use, and errors in Optical Character Recognition will need to be checked. The question of context as a determiner of meaning for vocabulary will need to be addressed using natural language processing (NLP) in combination with structured mark-up as a method of analysis. The NLP approach will be used to identify context-dependent features of writing while the mark-up will focus on controlled vocabulary and identifiable terminology. Both can be semi-automated, and will be complemented by the use of other digital text analysis tools that can be run on the larger corpus of secondary materials to track reception history for changing terminology and nomenclature. The scale of this second phase of textual analysis would preclude analog reading methods from being used, while the first phase of textual processing answers the demands of digital technology to make explicit the judgments of the human designers of the project. Data-mining, distant reading, and close reading will all contribute to the project.

Differential cartography will be based on contrasts among European maps, verbal descriptions and terminology that can be extracted from these as having

indigenous sources, and those few sources of indigenous cartography (verbal or visual) that exist in the record (e.g. Codex Nuttall is a pre-Conquest map). The task is to create cartographic simulations of an alternative worldview that does not reference European geospatial systems but has a consistent system grounded in indigenous experience, and then put these into contrast with the existing cartographic record of the "discovery" of the "new" world. Thick mapping techniques that layer historical materials in contrasting cartographic representations will allow us to present different views of the New World as a literal, as well as metaphoric, space of cross-cultural encounters. We are interested in imagining differential spatial systems and visualizations rooted in the worldviews and notions of proximity and distance, memory and community, duration and extension specific to indigenous cartographies. These are radically incommensurate with the projection and coordinate systems that are now naturalized features of standard mapping and GIS applications.

Work plan

- ✗ Identify sources for texts and maps
- ✗ Obtain permissions and digital versions
- ✗ Test the natural language processing analysis
- ✗ Create xml schema for indigenous vocabulary and nomenclature and for European geographical terms
- ✗ Mark texts and contrast search results with NLP analysis
- ✗ Search reception history corpus for usage change in terminology and vocabulary
- ✗ Create a list of cartographic fundamentals from indigenous perspectives
- ✗ Create simulations from these fundamentals
- ✗ Contrast with European maps of encounter
- ✗ Analyze the "differential" in these geographical and temporal attitudes and map them using a geo-temporal database that charts changes in attitude as shifting conceptions of space

Dissemination and participation

Poster and panel sessions at national professional meetings, postings on GeoBlog and cultural geography sites; invited response from senior figures in cultural geography and historical mapping; virtual roundtables organized as classroom events. The ongoing project will be supported by a participatory architecture that allows the mappings to be annotated extensively and also repurposed. Build on this material but extend to larger digitized gazetteer and cartography collections with emphasis on place names and cultural differences in geospatial features. Finally, develop prototypes for geospatial visualization engine that is conceptualized and structured according to the differential geographies embodied in indigenous worldviews.

Assessment

Peer review of data structure, credit for simulations, course evaluations from students on comparativist approaches in class and ease of use of analytic tools; scale of participation; ability to reveal both the limits and possibilities of interoperability with existing geospatial databases and other geo-browsers.



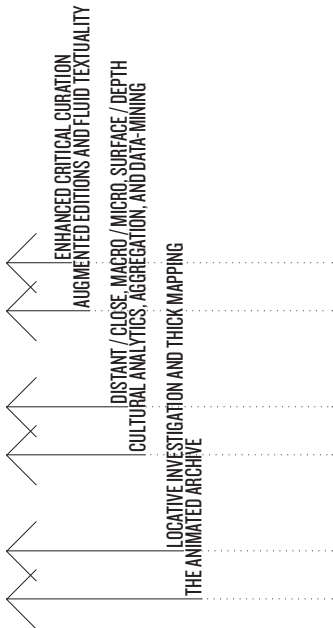
CASE STUDY 2 EXPANDED PUBLICATION OF A TEXTUAL CORPUS OF PAPYRUS FRAGMENTS FROM THE ALEXANDRIA LIBRARY

This textual corpus project will build a collection that links to existing repositories, makes use of certain text-based annotation platforms for debate about textual variants, performs some probabilistic natural language processing, uses a collation tool to study those variations, and creates an augmented critical edition of these fragments. Several different traditional and expanded publication models will be used to allow scholars with different profiles and agendas to present their findings in an appropriate fashion.

A new cache of papyrus fragments has been discovered in Alexandria, Egypt. Though considerably damaged by neglect and wear, these fragments promise to answer some long-standing questions about the spread of the Phoenician script and dates of its adoption across North Africa, particularly the coastal regions to the west, and its possible dissemination along trade routes into India. Some surprising features of these papyri make clear that they were recycled several times in the course of their use. Many are palimpsests, and some have several layers of script in varying degrees of legibility.

A host of different imaging technologies and digital platforms for integrating the data collected from the papyri are currently available. One major part of this project will be to repurpose some of the techniques that have been used successfully in other projects in Western Semitic epigraphy. However, the language in these papyri is not limited to Semitic tongues, and to the surprise of the librarians involved in the discovery, several as-yet unidentified languages seem to have been making use of scripts whose use and spread had been thought to be well-documented. One scholar has suggested the presence of Indo-European roots in the organization of the linguistic structure, which would suggest earlier contact with the Indian subcontinent and more dramatic cultural diffusion and influence than has previously been thought. Before any natural language processing can be done on the texts, they have to be deciphered (because of their poor condition). NLP techniques for ancient languages are in experimental stages, especially with fragmentary sample sets. Some speculative and probabilistic readings of the papyri and of the texts will be used.

The research problem is to identify the language groups represented in these papyri, match the script forms and use with the known corpus of Semitic epigraphy, and track the variants in a database that can support data-mining and text analysis across versions, translations, and script forms. A side benefit will be the creation of a digitized corpus of the papyri. One of the difficulties is that a major figure in early Indo-European languages is elderly, ill, and unable to travel so that his input will have to be done entirely on digital surrogates. A platform for annotating and tracking his contributions will have to be built or repurposed from existing platforms. This is an incentive for involving a team of mid-career and younger scholars whose formulation of problems of linguistic change and diffusion will also be essential. They are demanding an augmented publication format that will allow their work to be published rapidly, with a short peer-review cycle and with various scales of intellectual contribution, links to other existing corpora and repositories, and even an agonistic spirit of gamesmanship.



Methods

To do this project effectively, some crowd-sourcing of the translations, decipherment, and editing of the documents seems like a viable possibility. Statistical methods for doing large corpus analysis and comparison will also be essential. Thus, both close-reading and distant-reading techniques need to be involved in the study of the texts, the artifacts, and the scripts. The first phase of the project will require extensive integration of the imaging and digitization, with all uncertain signs or graphic elements marked so that the guesswork part of the project is conspicuously noted. Using the cultural analytics platform for display of large numbers of artifacts as well as pattern recognition software, similarities in script forms will be used to pinpoint linguistic similarities. The text translations will remain fluid, with variants and disputed elements conspicuously marked. Publishing these bits and pieces on a regular basis will be essential if any crowd-sourced work is going to occur. The senior scholar has asked that his interpretations be given a separate layer for presentation online so that his work can stand alone and be scraped off for later publication in print format.

Work plan

- ✘ Identify imaging techniques and sources of equipment
- ✘ Establish partnerships for shared access
- ✘ Create image files and test integration and comparison techniques
- ✘ Test probabilistic methods of text analysis for fragmentary data sets
- ✘ Identify translations of source texts where appropriate
- ✘ Link to existing repositories and online translations
- ✘ Test annotation and version-control platforms
- ✘ Test the cultural analytics and pattern-recognition software
- ✘ Put peer-review system into place for short-cycle contributions
- ✘ Create a platform of publishing and crowd-sourcing translation, editing, and decipherment
- ✘ Continue iterative process of imaging, translation, and decipherment
- ✘ Do sequential publishing of the findings in the form of an augmented edition that contains links, comparisons with existing corpora, and other versions of the texts

Dissemination and participation

Create a Twitter feed and RSS feed to publicize the project and engage participants; publish a beta version of the project in a print-plus mode online and establish a workflow to repurpose this content for traditional publication; crowd-source the translation and comparisons as well as the decipherment; augment the edition on an ongoing basis as scholars in the field indicate points of connection or comparison with other existing papyri, texts, or fragments of ancient scripts and languages. The project will be linked to major repositories in the Near and Middle East, Europe, and the United States by using an aggregation engine to allow for a larger statistical sample for investigation and comparison. The bridging of traditional and new modes

of scholarly engagement through distributed knowledge production approaches will allow the senior scholar to work effectively with younger scholars and allow for crowd-sourced input without collocation.

Assessment

The technical, intellectual, and cultural/institutional aspects of scholarship are interdependent. Success will be gauged in part by the extent to which the decipherment is completed and legibility for various layers established with credibility through the imaging and textual analysis. Another measure of success will be the number of contributions that enable links to existing digitized fragments and/or translations of ancient scripts.

CASE STUDY 3

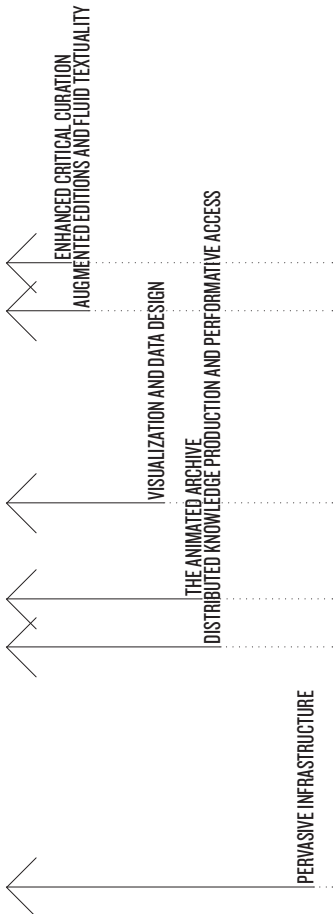
AUGMENTED OBJECTS & SPACES: JEWISH RITUAL OBJECTS IN DIASPORA

This project in critical curation and the augmentation of objects with commentary will resituate religious objects through a multi-modal approach that captures ritual practices across the time and space of diaspora. Spearheaded by a museum, the goal of this project is to produce an animated archive of cultural materials attuned to questions of provenance, use, and scholarly interpretation.

A university museum has an extensive collection of costumes, ritual objects, and recordings of ceremonies from Jewish families in Poland. These were obtained across several generations and do not all have fully documented provenance. Some were the work of early 20th century anthropologists who brought objects, photographs, and artifacts back from the field. Some are materials that were part of a large collection developed by a mid-century alumnus who donated them to the museum on the condition that the materials be used for teaching and public education about Jewish art and culture. Other materials were acquired through a fund established for the promotion of the study of the Jewish diaspora and were bought at auction or through reputable dealers by the museum curator. But some materials are of uncertain origin, and have been the subject of controversy, since they may have been looted, stolen, or smuggled out of the countries of origin, possibly as the result of Nazi appropriation. The museum has started a major initiative to make use of its Jewish ritual object collection and create a series of public programs, research opportunities, and curricular initiatives, as well as produce some permanent exhibits. In order to do this effectively, the director has determined that a digital approach based on augmenting the objects with multifaceted information displays will be the most effective way of addressing the ethical and intellectual dimensions of this cultural legacy.

While much of this material has been cataloged, not all of the descriptive information about the works has been put into digital form. Some of the earliest material was entered into the registrar's ledgers in handwritten form, while the most recent metadata conforms to the Getty's standards for Cataloging Cultural Objects. While the museum staff would like to standardize metadata for the purposes of managing the collection, they do not want to lose the important record of earlier approaches to the classification of artifacts. The idea of displaying different sub-collections within the larger whole also suggests some interesting historical narratives about the development of diasporic anthropology and cultural studies of ritual objects. Finding an effective way to display different interpretive approaches is crucial.

The artifacts in the collections range in size and scale from tiny mezuzot to tefilin, prayer shawls, Torah coverings and a fully rebuilt antique ark and bimah. Photographing these works for digital presentation also poses some challenges. Thinking through the organization of images to show multiple views, to facilitate detailed as well as overall study of the objects, and to allow for research as well as interpretive exhibits will take some serious repository building and design. The curators want to avoid any kitsch or special-effects approaches and also do not wish to create fictional spaces for actual artifacts.



One part of this collection came from a synagogue in Poland that was destroyed in World War II. But the site has been excavated, and there are extensive field notes and site photography to accompany the artifacts. These objects could be resituated through the narrative of the dig and accompanied by a story of the excavation. Other objects are of dubious provenance, and so need to be presented in a manner that allows serious scholarly engagement with their history and forms. A significant number of artifacts are known to be stolen, and finding descendants of survivors who can claim them is important and will be a part of the outreach supported by distributed access to the collections. Descendants of survivors will need to show appropriate credentials and be vetted before they can search and annotate the archives and repositories. Display of these materials may need to be limited, but research on them needs to be supported digitally so that some of the scholars best positioned to do this detective work can access them. The museum will collaborate with restitution groups which provide legal expertise and advocacy on behalf of survivors and their families.

Perhaps the most challenging materials among the collections are recordings of ritual practices that were never meant for public display. Some of these recordings were obtained surreptitiously. Others were obtained under very carefully worked out privacy agreements and intellectual property negotiations. Creating an environment that respects these agreements or goes further in using the museum environment to educate the public about the restraints on viewing seems essential in today's critical frameworks.

The digitization of these collections and creation of critical exhibits for public programming, education, and research is the focus of this project. The goal is to create exhibits that augment the objects and artifacts by exploring these many intellectual, critical, ethical, and political dimensions. Relating the artifacts to the geo-temporal history of diaspora is one component of the exhibit. But community testimony and archival materials that provide demographic data are also crucial contributions from which to generate display. The ultimate goal is to present the history of diaspora told through the movement of things and the rituals around their use. Creating a network analysis and information visualization will be one part of this presentation. Another will be the attempt to situate all of the artifacts within practices. Thus an artifact will never appear as a single thing on display. No artifact will be an autonomous object with a single text label. Instead, all objects will be accompanied by a digital matrix that exposes provenance questions, communities of use, and historical information about each as well as information on its acquisition, transmission, and debates about the ways it should be displayed and interpreted. In other words, the display of augmented objects will refract them along multiple lines of inquiry and interpretation that invite scholarly and critical engagement with an animated archive of materials.

Work plan

- ✘ Inventory the objects and artifacts
- ✘ Do an assessment of the metadata and cataloging protocols
- ✘ Create a set of crosswalks and schemes for description of the objects
- ✘ Address the multiple representations of the objects in existence and those to be created through the process of photographing or scanning
- ✘ Identify a content management system appropriate for museum management
- ✘ Test various network analysis tools and visualizations to display the movement of objects through time and space
- ✘ Modify the system so that it is customized for appropriate workflow and use
- ✘ Consider the administrative issues of permissions, access, and use of digital materials
- ✘ Consult appropriate scholars and authorities on legal and ethical issues around these materials
- ✘ Create an academic and community advisory board for ongoing review of ethical practices
- ✘ Create proof of concept demonstrations of augmented object displays that contain multiple viewpoints and artifactual histories
- ✘ Develop an appropriate permissions system and demonstrate the ability to create tiered levels of access and use for various audiences
- ✘ Design a method for processing input from professional and amateur scholars
- ✘ Consider the ways to engage stakeholders in the larger questions of cultural ownership

Dissemination and assessment

Create text labels, commentary, and debate that carry an author attribute; design a system of searching and indexing according to author; record and display relevant debate trails generated by objects; create a public forum in which these debates are edited or represented for study; engineer an app for mobile devices that allows input from contemporary sources.



CASE STUDY 4 VIRTUAL RECONSTRUCTION OF AN AFGHAN REFUGEE CAMP AS A SITE FOR CULTURAL MEMORY

This project repurposes the technology of online multi-player games to create a virtual community of testimony, witness, recovery, and social bonding. A spirit of joy and community-building is present as the shared repository of memories—photographs, some video materials and audio tapes, as well as letters, diaries, journals, and other materials—is being used by a younger generation to create a shared history through a series of mash-ups in which nuclear family histories become the common property of extended “families” through database documentaries and remix storytelling. This shared history may promote political activism, but also may become a target of unintended surveillance. Use of avatars and assumed identities is standard practice, and sensitive materials are likely to be part of the repository.

A professor specializing in politics, one in architectural history, and another in performance studies have been gathering material for a collaborative research project that would allow them to create a virtual model of one of the largest refugee camps that came into being after the Soviet invasion of Afghanistan in 1979. Interest in the site and its inhabitants has been spawned by recent events, including the pullout of American troops. Many of the children born in that camp grew up outside of Afghanistan—in Pakistan, Iran, India, and elsewhere throughout the region and beyond. An international organization interested in “virtual” repatriation is looking at patterns of diaspora, assimilation, and cultural memory. The idea of using a virtual reconstruction of the camp as a point of shared experience touches many nerves. The site itself, though still in existence, is in sensitive territory, difficult to access. But photographs taken by a U.S. soldier have been smuggled to the organization, Jalozai International, and offered to the U.S. academic team. A mobile phone application that repurposes these photographs to create an augmented reality experience of the original site has gone viral.

The site needs to be re-created virtually, but should be as accurate as possible with respect to the layout of the original camp. The group is working with a refugee community organization, which is in touch with a worldwide diasporic network of displaced persons and refugees. This community organization is eager to participate, as much as possible, in the creation of a virtual environment that could serve as a theater of reconciliation, testimony, memory, and commemoration. The United Nations High Commissioner for Refugees helped with repatriation, while the Pakistanis issued ID cards to all Afghani people living within their borders. For political reasons, many persons slipped through these official programs.

Much research has been done, and hundreds of hours of interviews have been logged. Descriptions of the camp from firsthand accounts, and from photographic and drawn images, have been gathered. The reproduction of everyday life in the camp would be made in virtual form in a simulation lab. The integration of the stories and eyewitness accounts and the creation of a fully immersive environment represent the next phases of research for this project. The questions surrounding the use of the environment and the quasi-game-like virtual world it suggests are all beginning to raise some concerns in the university and in the community. Creating a way to allow active participation and contributions from the community without trivializing the trauma of those who experienced the camp firsthand is one problem. Keeping fake testimony and malicious content from appearing is another. Protecting sources is yet another. However, all involved are interested in using theater and performance art as a way of engaging with recent history. They want to treat the camp not only as a historical site, but as a living memory that has to be engaged directly through imaginative experience if it is to be fully understood. Members of



the younger generation are engaging in the creation of remix narratives and role-playing games based on materials in the repository.

The goals of this project are to create the immersive simulation that allows for performance of recent trauma in an environment that may or may not be able to be controlled. Some simulations and predictive models are being built into the system, particularly those that use complex adaptive systems modeling techniques, and these will be used in dialogue with live user contributions to monitor emerging trends in the environment. However, any hint of surveillance or control will have disastrous results, and the simulations need to be fully transparent to all participants. The performance studies professor has been working with the interface designers to produce some avatar representations and thumbnail theaters that show possible scenes and probable story lines among live participants. Creating an interface that allows multiple users to participate actively in a multi-person performance while also making use of historical materials and documents will require careful scripting and guidelines. A polymorphous browser that displays materials differently depending on how parameters identifying the user are set is in beta.

Methods

Identify the available software for creation of a virtual site with social media and participatory capabilities, or consider making this site in Second Life or another virtual world application. Get information on issues of security and privacy if a third-party platform is used. Make sure the site can be used with mobile apps as many of the participants will not have cable connections, but will access the materials through their cellphones.

Work plan

- ✘ Establish communication with appropriate international organizations
- ✘ Develop collaboration with the refugee groups and their leadership
- ✘ Create a beta version of the virtual site
- ✘ Invite a small group for user testing
- ✘ Perform iterative user testing based on initial results
- ✘ Build out the virtual environment
- ✘ Create a public forum for input
- ✘ Create avatars to protect privacy and identity as appropriate
- ✘ Document use and participation
- ✘ Create cross-referencing tools for tracking shared information and memory

Dissemination and participation

Engagement of the UNHCR and dialogue with the Jalozai International leadership is crucial, as is ongoing support of the university where the project is housed; plans for a small working group to meet in Iraq and another in Pakistan will facilitate direct contact with academic team members; YouTube presentations and

virtual encounters are also planned, as is a series of performances in the virtual environment. These will be publicized as real-time events in the virtual space. Scholarly publication of findings will take various forms, including but not limited to traditional conference presentations and publications supplemented by digital collections and archives of the project and its materials deposited in the university library.

Assessment

Assessment will be ongoing; monitoring the participation and reaction of participants will be crucial to safeguarding privacy and gauging comfort levels as well as the effectiveness of debate, dialogue, and documentation; getting solid documentation of the contributions and testimonies is essential.

CASE STUDY 5 MULTI-AUTHORED LOCATIVE INVESTIGATION OF THE ZENON HEADQUARTERS AND CORPORATE ARCHIVE

This project aims to design a prototype tablet application that interweaves three components: a collaboratively developed body of interpretive research, an archive, and an architectural heritage site. The team will build a content management system for the online generation of a multi-authored essay and a digital archive that will be keyed to the physical structure of the heritage site. User-visitors will be provided with a navigation system that allows for multiple entry points and pathways for on-site and off-site access as well as for “reading” this “published” work. The prototype will be employed to beta test a new way of accessing information, interacting with knowledge, and experiencing data and research in physical and virtual space.

An archive of documents from the Zenon Corporation, whose headquarters occupied a historical building designed by Louis Sullivan, was discovered during recent restoration work. Stored in a safe in the basement of the building, the documents have now been moved to the University of Illinois, Chicago, where they have become the focus of interdisciplinary study. The archive contains copies of business correspondence, internal memos, meeting notes, drafts, minutes, calendars, and personal notes. Some are handwritten, some are in shorthand, and some are hand-edited drafts. Others are typewritten documents, many on letterhead featuring the building with its original facade elements, including downspouts and decorative features that were removed in the 1970s. The documents are from the 15-year period (1945-1960) during which Zenon, contracting with UIC, was developing a secret project for the U.S. military to build a distributed computing system for defense that many cultural historians believe was an early version of the networked computer.

The cast of characters is colorful. Zenon's president was a Harvard-educated scion of an old Chicago family, and the Sullivan building had originally been built for the home offices of his grandfather's insurance business. Plans for the project were leaked by his secretary and mistress, whose connections to gangland South Side mobsters seemed to hark back to Depression-era speak-easies. Scandal and intrigue dogged the Zenon project, and the secretary's body was found at the bottom of an elevator shaft, a crime concealed for a half decade. The documents also open up a window into the culture of the corporation and the research university during those years as well as the relationships among employees as mediated by the writing technologies of the time (pen and paper, stenographic machines, typewriters, shorthand).

An interdisciplinary team of scholars is working on the reconstruction of the Zenon project within a “history of the future” framework. It seeks to publish a “multidimensional essay” built around a core set of archival documents—the correspondence between an executive and his secretary—in order to explore a novel publishing model that allows for the building of connections across media, as well as across digital/physical boundary lines. One team member is a cultural historian of organizations; another is an expert on the history of network architectures. A third is interested in the “gendering” of spaces and forms in corporate architecture and, working with the team, seeks to develop an augmented reality app to reconstruct the original Zenon building at its site. Her work is funded by the Chicago Architectural Association and the Chamber of Commerce, both of which wish to integrate scholarly content into tours of the city's sites available on mobile devices. She proposes structuring the interpretive work being carried out, as well as the Zenon archive itself, as a function of the building's layout and spatial organization such that the site visitors are able to “see” events unfold at set points in the entry, lobby, elevators, and hallways, thereby fulfilling the needs of both funders. But the project must also

be designed so that off-site visitors can experience the full interplay of archival materials, interpretive research, and architectural features.

Working alongside the team of scholars is a technical team comprised of an interface designer who specializes in digital corpora, a computer science doctoral student working on information structures and knowledge models, a designer who is an expert in CAD systems, modeling, and architectural rendering, and an adjunct professor who spends half his time as a technology developer creating tablet apps. They will play the lead role in creating an application that allows: a) on-site users to move and rotate their tablet devices to navigate 3-D models of Sullivan's building framed by and animated by texts, annotations, commentary, and archival documents; and b) off-site users to replicate this experience on their own tablet devices in clear and meaningful ways. This prototype will allow scholars from across the humanities to test the viability of multidimensional formats as publishing platforms.

The software challenge is three-fold: First, to create a content management system that allows the three scholars to collaboratively generate texts, metadata, annotations, and images; interweave them with the Zenon archive; and embed them in the architectural models. Second, to create an intuitive, user-controlled navigational interface that supports both on- and off-site movement through the combined content governed by the spatial structure of the building, semantic features of the content, the user's prior navigational choices, and user settings regarding the density of data layers and frames. Third, to build richly textured digital representations of the interiors and exterior of the Sullivan building that are appropriately keyed to the various data layers that make up the research project.

Work plan

- ✘ The entire team works together to develop the information architecture and content schema
- ✘ The developer customizes or creates a bespoke content management system to be used in the writing of the three intertwined "essays"
- ✘ The scholars generate the writing, annotations, diagrams, images, and other content, along with metadata and links
- ✘ The technical team devises a database architecture for the selected portion of the Zenon archive that will work seamlessly with the content management system and the augmented reality app and its architectural models
- ✘ The designer and developer iterate and test interface designs, user interaction, and navigation for both on-site and off-site access
- ✘ The team goes back and forth between scholars generating material and the designer and developer demonstrating the application until an alpha version is ready for user testing
- ✘ Alpha testing with users both for the performance of the application but also for the reading and interpretation of the content
- ✘ Develop a website for feedback and user participation
- ✘ Iterate a beta version based on feedback and refine for public release

Dissemination and participation

The application will be made available for download from online stores at no charge. Its availability will be broadcast through social media and will be promoted by the supporting institutions in a version that is "tuned" for nonspecialist audiences. The project website will continue to be a site of conversation in the hope that other teams of scholars might contract with the technical team to undertake new and extended versions of the project. The results of the project, including user feedback, will be documented, analyzed, and shared through conventional venues, such as conferences and journals, from each of the team member's fields.

Assessment

Assessment will be built upon three points of engagement. On-site and off-site user testing from the standpoint of human-computer interaction will be a part of the development process. User metrics will be generated through social media rankings, online store rankings, number of downloads, and direct feedback. Critics and scholars will be asked to review the application from a variety of vantage points: as a model for scholarly publishing and for the merits of the scholarship of the three contributors.

This is a section of [doi:10.7551/mitpress/9248.001.0001](https://doi.org/10.7551/mitpress/9248.001.0001)

Digital_Humanities

By: Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, Jeffrey Schnapp

Citation:

Digital_Humanities

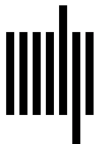
By: Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, Jeffrey Schnapp

DOI: 10.7551/mitpress/9248.001.0001

ISBN (electronic): 9780262312103

Publisher: The MIT Press

Published: 2016



The MIT Press

© 2012 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was typeset in Bembo, Benton Gothic, and Knockout by the authors and was printed and bound in the United States of America.

Cover photo: © 2012 Jeremy Eichenbaum

Library of Congress Cataloging-in-Publication Data

Digital_humanities / Anne Burdick ... [et al.],

p. cm.

ISBN 978-0-262-01847-0 (hardcover : alk. paper)

1. Humanities—Electronic information resources.

2. Humanities—Computer network resources.

I. Burdick, Anne.

AZ195.D54 2012

001.30285—dc23

2012026514

10 9 8 7 6 5 4 3 2 1

