The Productive Unease of 21st-century Digital Scholarship

Julia Flanders <Julia_Flanders_at_brown_dot_edu>, Brown University

Abstract

Despite prevailingly progressive narratives surrounding the impact of digital technology on modern academic culture, the field of digital humanities is characterized at a deeper level by a more critical engagement with technology. This engagement, which I characterize as a kind of "productive unease", is focused around issues of representation, medium, and structures of scholarly communication.

Technological Progressivism

The narratives that surround technology tend, understandably, to be progressive. Moore's law, which states that the complexity and hence the processing power of computer chips is doubling every couple of years, and Kryder's law, which says something similar about disk capacity, have visible and in some cases stunning illustrations in the world around us. We see evidence in products such as palmtop devices that have thousands of times the computing power and storage capacity of ENIAC, the first stored-program electronic computer; personal disk storage is now purchasable almost by the terabyte, and processor speed is now measured by the gigahertz; both of these statements will have dated by the time this article is published. We also see the effects of these developments in processes whose increasing speed produces subtle luxuries that creep into our lives, almost without our taking particular notice: for example, color screens for computers, three-dimensional icons, the clever animation behaviors that are as ubiquitous (and as useful) as small plastic children's toys. Or, more substantively: the fact that you can now store and edit digital video footage on your laptop, or view streaming movies on a device you can put in your pocket. These kinds of change produce easy metrics for success and a correspondingly easy sense of progress.

[1]

Digital humanities scholarship to a large degree shares this sense of progress. We see, first of all, simple infrastructural developments that change the social location of computers and bring them into our sphere of activity. The ubiquity of computing resources means that it's no longer remarkable for humanities scholars to work with computers: one doesn't have to get a special account from Central Computing or explain why one needs it; it's not considered quaint or cute or bizarre. Certain efficiencies and conveniences are now commonplace; it has become expected that things scholars want to read or learn will be more or less easily available from anywhere, at any hour, electronically. And there are indirect effects as well: all of these changes produce the conditions for consumer-level products like electronic book readers, hand-held browsing devices, social software like Flickr and YouTube. These products provide extended horizons of usage, and produce a generation of students (and eventually future scholars) for whom computers mean something completely different: for whom they are not a specialized tool but part of the tissue of the world.

[2]

The effects of these developments are all around us in the emerging shape of digital scholarly tools and research materials. At a basic level, the increased power of modern computers is almost literally what makes it possible to use them effectively for humanities research. In early computer systems, scarcity of storage space dictated extremely frugal methods of representing characters: because it only uses 7 bits of information to represent each character, ASCII can represent only 128 characters, of which only 95 are actually printable characters. This limited the effective alphabet to upper- and lower-case roman letters, Arabic numerals, and common punctuation marks, with no accented characters or characters from non-roman alphabets. The advent of Unicode in the 1990s is a direct outcome of the increase in

[3]

storage space, allowing the representation of nearly all human writing systems and freeing digital scholarship on texts from early artificial limitations.

This same comparative abundance of space has also opened up the whole domain of image processing, giving us another information vector to use for research, leading to work in which the graphical meaning of text can be explored alongside its linguistic meaning. and allowing us also to explore the interpenetration of image-based and text-based approaches. To appropriate a term Jerome McGann suggested in his opening keynote to the conference at which this paper was originally presented, there is a dialectical process opening up here as well: the mutual pressure of image and text, of alphabetic and figural modes of representing meaning, is now blossoming into an extremely lively field of study.

4

The rhetoric of abundance which has characterized descriptions of digital resource development for the past decade or more has suggested several significant shifts of emphasis in how we think about the creation of collections and of canons. It is now easier, in some contexts, to digitize an entire library collection than to pick through and choose what should be included and what should not: in other words, storage is cheaper than decision-making. The result is that the rare, the lesser-known, the overlooked, the neglected, and the downright excluded are now likely to make their way into digital library collections, even if only by accident. In addition, the design of digital collections now frequently emphasizes precisely the recovery of what has been lost, the exposure of what has been inaccessible. Projects like the Women Writers Project, or Early English Books Online, or any one of countless digital projects now under way at universities across the country, focus on providing access to materials that would otherwise be invisible to researchers. This access proceeds on two fronts: first, by digitizing them so that they can be read without visiting the specific archive where they are held, but also, more importantly, by aggregating them and making them discoverable, by heaping them up into noticeable piles. The result is that minority literatures, non-canonical literary works, and the records of what goes on in (what appeared earlier to be) the odd corners of the universe are all given a new kind of prominence and parity with their more illustrious and familiar cousins.

5

Invisibly, under the hood (so to speak), increased speed and computing power has also given us tools that finally propel us over the threshold of possibility: humanities novices are becoming able to participate meaningfully in what would formerly have appeared to be impossibly technical projects. Examples include tools for XML text encoding that are good enough, and fast enough, that anyone can learn to use them within ten minutes; or, similarly, tools for image manipulation that put real power in a novice's hands. Even improvements in things like compression algorithms, as Morris Eaves observes in his contribution to this issue, have a huge impact on the accuracy and effectiveness of digital image representation.

6

But despite the fact that these are tangible improvements, there is also an important sense in which their progressive momentum is not, ultimately, what is characteristic of the digital humanities as a field. John Unsworth, in an article entitled "What is Humanities Computing and What is Not?" makes a point of noting the difference between using a computer for any of its many practical purposes, and using the computer as a scholarly tool:

7

> ...one of the many things you can do with computers is something that I would call humanities computing, in which the computer is used as tool for modeling humanities data and our understanding of it, and that activity is entirely distinct from using the computer when it models the typewriter, or the telephone, or the phonograph, or any of the many other things it can be. [Unsworth 2002]

Unlike its comparatively recent ability to model the telephone or the phonograph, the computer's role as a tool for modeling humanities data is of long standing — arguably extending back to Father Roberto Busa's 1945 *Index Thomisticus* and certainly including early tools and methods including concordancing, text analysis, and text markup languages. Although our ability to work with these models has without doubt been made easier by the advent of faster, more seamless tools, the complexity and interest of the models themselves has been affected little if at all. We have only to consider as an example Willard McCarty's remarkable project of modeling mutability in his *Analytical Onomasticon to the* Metamorphoses *of Ovid*, a project of great complexity and nuance which was undertaken almost entirely through markup and without the aid of any specialized tools for model construction, visualization, or data

manipulation. The nature of the models being created in the digital humanities may be changing with time, but not as a function of speed or power, but rather as a result of changes in emphasis or theoretical concern.

In this respect, the digital humanities domain reflects the non-progressiveness of the humanities disciplines more generally, and also reveals what may be a fundamental tension at its heart. If the rhetoric at the heart of the "digital" side of "digital humanities" is strongly informed by a narrative of technological progress, the "humanities" side has equally strong roots in a humanities sensibility which both resists a cumulative idea of progress (one new thing building on another) and yearns for a progressive agenda (doing better all the time). The theoretical and methodological shifts that constitute disciplinary change in the humanities, when viewed in retrospect, do not appear clearly progressive in the way that sequences of scientific discoveries do, though they do appear developmental: they are an ongoing attempt to understand human culture, from the changing perspective of the culture itself. But the resilience of fundamental habits and assumptions concerning literary value, scholarly method, and academic standards suggests that the humanities are in fact governed by a self-healing ideology that persists comparatively unchanged.

In charting the intellectual aspirations of the digital humanities, it is tempting to elide the difference between this sense of ongoing debate and the gains in size and speed that come from the technological domain. But the intervention made by digital technology when it truly engages with humanities disciplines is something apart from both the simple progressivism of technology and the canonical resilience of the traditional humanities. In the same article I quoted from earlier, John Unsworth characterized humanities computing as follows:

> [h]umanities computing is a practice of representation, a form of modeling or [...] mimicry. It is[...] a way of reasoning and a set of ontological commitments, and its representational practice is shaped by the need for efficient computation on the one hand, and for human communication on the other. [Unsworth 2002]

In other words, it is neither about discovery of new knowledge nor about the solidity of what is already known: it is rather about modeling that knowledge and even in some cases about modeling the modeling process. It is an inquiry into how we know things and how we present them to ourselves for study, realized through a variety of tools which make the consequences of that inquiry palpable. This is why, when humanities practitioners learn a technology like text encoding, they feel both a frisson of recognition — of a process that is familiar, that expresses familiar ideas — and also the shock of the new: the requirement that one distance oneself from one's own representational strategies and turn them about in one's hands like a complex and alien bauble. As Unsworth puts it further along,

> Humanities computing, as a practice of knowledge representation, grapples with this realization that its representations are surrogates in a very self-conscious way, more self-conscious, I would say, than we generally are in the humanities when we "represent" the objects of our attention in essays, books, and lectures. [Unsworth 2002]

Representational technologies like XML, or databases, or digital visualization tools appear to stand apart from the humanities research activities they support, even while they encapsulate and seek to do justice to the assumptions and methods of those activities. Humanities scholarship has historically understood this separateness as indicating an ancillary role — that of the handmaiden, the good servant/poor master — in which humanities insight masters and subsumes what these technologies can offer. Technology implements what humanities insight projects as a research trajectory. But in fact the relationship is potentially more complex: by expressing "human communication" in the formal language needed for what Unsworth calls "efficient computation," these representational technologies attempt to restate those methods in terms which are not identical to, not embedded in the humanities discourse. They effect a distancing, a translation which, like any translation or transmediation, provides a view into (and requires an understanding of) the deep discursive structures of the original expression.

Unsworth is careful to observe that not all digital humanities activities — in fact, very few — really constitute this kind of intervention, or count as "humanities computing" according to his strict definition. The act of publishing digital content, of making an uncritical digital facsimile of a physical artifact, does not produce this effect of translation or the resulting potential for insight. I would argue that we can recognize humanities computing in his sense of the term, precisely by a kind of productive unease that results from the encounter and from its product. This unease registers for the humanities

scholar as a sense of friction between familiar mental habits and the affordances of the tool, but it is ideally a provocative friction, an irritation that prompts further thought and engagement. In the nature of things — systems and people being imperfect — it might produce a suspicion that the tool in question is maladapted for use in humanities research. In some cases that may be true, and in some cases that may be a self-defensive response which deserves further probing. But where that sense of friction is absent — where a digital object sits blandly and unobjectionably before us, putting up no resistance and posing no questions for us — humanities computing, in the meaningful sense, is also absent. Humanists may learn from the content of such objects, treated as research materials, as they always have. These objects will serve as more or less effective surrogates for their physical originals and may produce efficiencies of access and other practical benefits of one sort or another. But they have no contribution to make to humanities scholarship: they make no intervention, they leave no intellectual mark.

Productive Unease

Where, then, is this unease manifesting itself? and what useful insights and intellectual traction does digital humanities scholarship provide on the central problems of the humanities? Here are three areas where I would argue that interesting critical friction is being produced by work in digital humanities. | 13

1. Digital scholarship is uneasy about the significance of medium.

One almost immediate effect of the emergence of digital texts was to instigate a discussion of medium: a discussion which raised the stakes and broadened the scope of the discussion, which had previously been of concern primarily in scholarly editing, in the tradition of D. F. McKenzie. The initial manifestations of this discussion were expressed as anxiety about the unreliability of digital texts, linking this quality to the medium itself rather than to social practices such as peer review. As the Women Writers Project reported in summarizing its 1995 survey of scholars, "anxiety about the accuracy of electronic texts was so acute that some respondents discussed it even in answer to questions on other subjects, and it clearly represented the single largest obstacle to general scholarly use of electronic texts." Early threads in electronic discussion forums such as SEDIT-L also foregrounded this problem of inaccuracy as a kind of worrisome dark side to the "polymorphic, polysemic, protean" qualities attributed to digital texts in more optimistic analyses. The theme attests to an odd sense of self-consciousness about how to make digital texts reliable — in other words, how to transplant a familiar set of social practices into unfamiliar territory, as if this might involve profoundly different processes from those which had been used to produce reliable print texts. | 14

This anxiety looks dated in retrospect, but it has had a salutary effect: it has produced an interest in understanding medium and its role in anchoring our textual perceptions. Digital humanities scholarship now includes an awareness of the representational significance of medium as a fundamental premise. This is not only because the digital medium is seen as a kind of meta-medium in which other media can be modeled or represented (which requires us to think about the special characteristics of those other media, for modeling purposes), but also because the digital medium itself is not representationally uniform. The kinds of sampling and decomposition that seem at first blush like typically "digital" effects are very different from the formalizing properties of text encoding or vector graphics. | 15

The digital humanities world is in fact full of intensive and fruitful debate about representation and medium. Jerome McGann's sustained engagement with the question of how structured text markup may fail or succeed at representing literary texts — his account of the dialectical influence of different representational modes and what we can learn by their insufficiencies — and the responses and research this work has elicited from markup theorists, taken together have provided a great deal of insight into how digital formats represent textual and figural information. And this insight has in turn shed light backwards (as it were) upon the traditional printed scholarly edition. | 16

2. Digital scholarship is uneasy about the institutional structures of scholarly communication.

By its emergence through innately cross-disciplinary and cross-organizational (and widely differing) formations, the digital humanities domain has helped to create a critical self-consciousness about the role institutions play in establishing and maintaining cultural habits that affect how humanities research is done. Alan Liu, in a 2003 MLA | 17

presentation, asserted that "The humanities should embrace the poiesis of IT for alternative ends — first of all at the level of organizational imagination" to "reimagine the protocols of the work of education" [Liu 2003, 6]:

> Here I come to what I perceive to be one of the frontiers of IT in the humanities. That is the far territory on which the many, scattered humanities computing programs, centers, projects, and so on that have used IT as a catalyst to reorganize the normal disciplinary work of the humanities evolve from ad hoc organizational experiments into strategic paradigms of interest to the profession as a whole. In general, we must acknowledge, the profession of the humanities has been appallingly unimaginative in regard to the organization of its own labor, simply taking it for granted that its restructuring impulse toward "interdisciplinarity" and "collaboration" can be managed within the same old divisional, college, departmental, committee, and classroom arrangements supplemented by ad hoc interdisciplinary arrangements. [Liu 2003, 7]

Digital humanities projects, practices, and practitioners typically emerge out of working relationships which by their nature raise questions about the politics of work, and occupy a space that is naturally and productively critical of current tenure and reward systems. These systems are still struggling to understand the fundamentally collaborative and interdisciplinary work of digital humanities or the new modes of scholarly communication it is engendering. [18]

Following almost inevitably from this unease about institutional and organizational containers for professional identity is a related concern with published expressions of professional identity and the question of how we evaluate new forms of communication and scholarly work. In effect, digital scholarship reveals a conundrum that has lain at the heart of humanities scholarship for decades: how can we simultaneously encourage paradigm shifts and radical revisions of our modes of analysis, and also know how to evaluate them once we have them before us? Digital scholarship proceeds through collaborations and hybridizations that challenge our notions of discipline — indeed, often that is the desired goal — but evaluation and professional acknowledgement are typically provided through conduits that are slower to adapt and may not necessarily view such a challenge as *ipso facto* valuable. The MLA's "Guidelines for Evaluating Work with Digital Media in the Modern Languages" acknowledge this difficulty, marking the disciplinary changes that are taking place and the uncertain position that "traditional notions of scholarship" occupy in relation to emerging forms of academic work: [19]

> Digital media have created new opportunities for scholarship, teaching, and service, as well as new venues for research, communication, and academic community. Information technology is an integral part of the intellectual environment for a growing number of humanities faculty members. Moreover, digital media have expanded the scope of textual representation and analysis to include, for example, image and sound. These innovations have considerably broadened the notion of "text" and "textual studies," the traditional purview of modern language departments.
>
> While the use of computers in the modern languages is not a new phenomenon, the popular success of information networks like the World Wide Web, coupled with the proliferation of advanced multimedia tools, has resulted in an outpouring of critical publications, applied scholarship, and curricular innovation. Humanists are not only adopting new technologies but are also actively collaborating with technical experts in fields such as image processing, document encoding, and information science. Academic work in digital media should be evaluated in the light of these rapidly changing institutional and professional contexts, and departments should recognize that some traditional notions of scholarship, teaching, and service are being redefined. [MLA 2002]

At the same time, the Guidelines suggest that there may be a necessary — and fairly durable — interdisciplinarity in play here, which will always place work of this kind in a procedurally awkward interdepartmental space. In their recommendations they advise tenure and promotion committees to: [20]

> **Seek Interdisciplinary Advice.** If faculty members have used technology to collaborate with colleagues from other disciplines on the same campus or on different campuses, departments and institutions should seek the assistance of experts in those other disciplines to assess and evaluate such interdisciplinary

work. [MLA 2002]

If the future of digital scholarship may thus be (for some time at least) to remain "other" to the standard disciplinary structures of the academy, however, this should not be taken as a misfortune. The unease that is the theme of this essay is productive in this case precisely because it makes us aware of discipline as both a formative intellectual constraint, and a somewhat arbitrary institutional reality. Acknowledging the arbitrariness is crucial because it reminds us that change is possible and may be necessary. But acknowledging the formative qualities of the constraint is equally crucial, because it reminds us that we cannot simply posit a return to some pre-lapsarian, pre-disciplinary state of unfettered intellectual free play. Digital scholarship works in relation to established disciplines, even as it stands in some degree usefully apart from them.

21

3. Digital scholarship is uneasy about the significance of representation in forming models of the world.

Jerome McGann observed, at the start of the conference for which this cluster of essays was written, that humanistic study is all about representation: it is about decoding, understanding, historicizing, and critiquing the representational modes and artifacts of the past and present, and reflecting on what they tell us about human culture. But while we are good at distancing ourselves critically from the representational forms we encounter in the materials we study, we're surprisingly less so when it comes to the modes we use ourselves. One of the most significant contributions of the digital humanities on modern scholarship is precisely to foreground issues of how we model the sources we study, in such a way that they cannot be sidestepped. Where printed editions allowed us to treat their contents as if no change in medium had taken place, digital editions force us to confront the very same set of issues with far more rigor and clarity. As John Unsworth observes,

22

> once we begin to express our understanding of, say, a literary text in a language such as XML, a formal grammar that requires us to state the rules according to which we will deploy that grammar in a text or texts, then we find that our representation of the text is subject to verification – for internal consistency, and especially for consistency with the rules we have stated. [Unsworth 2002]

The word *verification* stands out here, sounding very cut and dried, threateningly technical, a mental straitjacket, but in fact the key phrase there is "the rules we have stated": it is the act of stating rules that requires the discipline of methodological self-scrutiny. It is in fact precisely the distance, the discomfort even, that digital representations carry vis-à-vis their print counterparts that reminds us that they are models. At first, this distance registers as a loss: digital representations are models "rather than" the real thing, taking the place it should occupy. But as our tools for manipulating digital models improve, the model stops marking loss and takes on a clearer role as a strategic representation, one which deliberately omits and exaggerates and distorts the scale so that we can work with the parts that matter to us.

23

In effect, digital scholarship embodies an unresolved conflict about scale, human effort, and the nature of digital work. The great bulk of digital research material now available does not look very "scholarly"; with the institutional focus on digital library development and the funnelling of digitization money through efforts of this type, there has been a great deal of emphasis on large-scale activities with light informational yield and strong tradeoff of scale against precision, such as Google Books. Traditionally, humanistic scholarship has been focused on high-labor "craft" activities where care and precision matter, and despite the importance of digital libraries, there is thus a kind of mismatch between current digital library approaches and scholarly expectations. Typically, scholars are not involved closely in the development of these resources: they are alienated, in a way, from the technology because they see it as intrinsically not about their craft, intrinsically maladapted to the kinds of thoughts they are accustomed to think.

24

Alan Liu has observed this shift in the way digital resources are crafted, and noted the politics of the change. In an essay called "Transcendental Data," he charts the emergence of a new aesthetic of the "data pour" in which information is in its most characteristic and powerful state when separated from specific form. The information design of 21st-century digital resources draws on precisely this approach: on presentational models that can scale up by orders of magnitude to accommodate the vast and increasing quantities of material. But they do so by decreasing our ability to

25

apprehend the details of individual objects. The challenge information designers now face is how to span that distance, and how to represent the macrocosm so that we don't lose sight of its parts. This is true not only literally, but also intellectually: the question is how scholarly methods can adapt to this shift in scale without losing their grasp on the concrete and beloved quiddity of texts and words and books and artifacts.

The world of social software is way ahead of us, in some ways, in addressing these problems. Although without our critical sensitivity and unease, casual users are experimenting with tools like Flickr and YouTube and del.icio.us, which attempt to represent the texture of the relevant landscape, as imagined by the people living in it: the photographs that matter to people, the web sites they read, the topics they think these things are about. But the scholarly tribe are not so very far behind, or at least they are in the race: efforts like TAPoR, the Text Analysis Developers Alliance, NINES, and MONK are setting their sights on this same problem, trying to see how far the human perceptual mechanisms can be stretched as they try to grasp both the macrocosm and the microcosm and the informational strands that connect the one to the other.

26

Looking to the Future

All of this unease, as my title has already asserted, is productive: not of forward motion but of that same oscillating, dialectical pulsation that is the scholarly mind at work. Digital tools add a challenge and give us a new set of terms — like a new planet in the system, they change the vectors of all the other things we have in our universe. They will probably change the way humanities research is done. When writing the grant proposals that so often fund digital humanities work, all of the natural rhetoric is progressive — there will be more, and it will be better, and it will open up new ways of thinking. But it is healthy to remember that the most interesting papers and books we read, in any genre, are those that neither foretell doom nor glory, but give us instead an interesting idea about the world to play with. Methods and tools that combine what has been gained in power and scale with a real measure of scholarly effort and engagement can give us such an idea. But the intellectual outcomes will not be judged by their power or speed, but by the same criteria used in humanities scholarship all along: does it make us think? does it make us keep thinking?

27

Works Cited

Flanders 1999 Flanders, Julia. *Scholarly Habits and Digital Resources: Observations from a User Survey*. *The Women Writers Project*. 1999. http://www.wwp.brown.edu/research/publications/reports/rwo/rwo_initial_report.html.

Liu 2003 Liu, Alan. "The Humanities: A Technical Profession". Presented at *MLA 2003*. *Proceedings of the Modern Language Association Convention 2003* (2003). http://www.english.ucsb.edu/faculty/ayliu/research/talks/2003MLA/Liu_Talk.pdf.

Liu 2004 Liu, Alan. "Transcendental Data: Toward a Cultural History and Aesthetics of the New Encoded Discourse". *Critical Inquiry* 31 (2004), pp. 49-84.

MLA 2002 Modern Language Association. *Guidelines for Evaluating Work with Digital Media in the Modern Languages*. *Modern Language Association*. 2002. http://www.mla.org/resources/documents/rep_it/guidelines_evaluation_digital.

Shillingsburg 1993 Shillingsburg, Peter. "Polymorphic, Polysemic, Protean, Reliable Electronic Texts". In George Bornstein and Ralph G. Williams, *Palimpsest: Editorial Theory in the Humanities*. Ann Arbor: University of Michigan Press, 1993.

Unsworth 2002 Unsworth, John. "What is Humanities Computing, and What is Not?". *Jahrbuch für Computerphilologie* 4 (2002). http://computerphilologie.uni-muenchen.de/jg02/unsworth.html.