The Humanities in the Digital Age

On January 23, 2014, the Academy hosted a meeting at Stanford University on The Humanities in the Digital Age. Richard Saller (Vernon R. and Lysbeth Warren Anderson Dean of the School of Humanities and Sciences and Kleinheinz Family Professor of European Studies at Stanford), Elaine Treharne (Roberta Bowman Denning Professor of the Humanities and Codirector of the Center for Medieval & Early Modern Studies at Stanford), Franco Moretti (Danily C. and Laura Louise Bell Professor and Professor of English and Comparative Literature at Stanford), Joshua Cohen (Marta Sutton Weeks Professor of Ethics in Society; and Professor of Political Science, Philosophy, and Law at Stanford), and Michael A. Keller (Ida M. Green University Librarian; Director of Academic Information Resources; Publisher of the Stanford University Press; and Founder/Publisher of HighWire Press at Stanford) discussed the humanities in the context of rapidly developing new technologies. The program served as the Academy’s 2004th Stated Meeting. The following is an edited transcript of the presentations.

The title of this evening’s program, “The Humanities in the Digital Age,” captures both the sense of possibility for the humanities with the rapid development of new technologies, and also the sense of foreboding that comes from feeling threatened by obsolescence. There is perhaps no better place than Silicon Valley to explore this issue. Our four speakers will present some tantalizing examples of the possibilities for extending humanities research through new technologies. But Josh Cohen, who always surprises me, has told me that he has changed his subject since he last wrote to me, and that he is actually going to take a bit of a contrarian’s view on the limits of digital technology in the humanities – that is, he will look at what it won’t do. And I think that is a very important subject.

Our first speaker is Elaine Treharne, the Roberta Bowman Denning Professor of the Humanities, whose home is the English department here at Stanford. Elaine’s education was in the field of Old English, with a focus on early British manuscripts: their materiality, contents, and context of production and reception. What makes Elaine especially appropriate for this evening’s program is that she places the medieval manuscript in the much broader suite of technologies related to texts.

Next will be Franco Moretti, the Danily C. and Laura Louise Bell Professor of English and Comparative Literature. Franco has done many things at Stanford, as well as writing prodigiously, but relevant here is that he is founder of the Center for the Study of the Novel and founder of the Literary Lab.

Our third speaker will be Josh Cohen, the Marta Sutton Weeks Professor of Ethics in Society; he is also Professor of Political Science, Philosophy, and Law. He is a political theorist trained in philosophy, a principal investigator in the program on liberation technology at the Freeman Spogli Institute for International Studies, and Editor of Boston Review.

Finally we will hear from Michael Keller, the Ida M. Green University Librarian, Director of Academic Information Resources, Publisher of HighWire Press, and Publisher of Stanford University Press. These titles touch on his major professional preoccupations: that is, a commitment to supportive research, teaching, and learning; the effective deployment of information technology hand-in-hand with materials; and active involvement in the evolution and growth of scholarly communication.

Richard Saller
Richard Saller is the Vernon R. and Lysbeth Warren Anderson Dean of the School of Humanities and Sciences and the Kleinheinz Family Professor of European Studies at Stanford University. He was elected a Fellow of the American Academy in 2005.
As a medievalist and book historian, working on text technologies from c. 60,000 BCE to the present day – that is, from inscribed rocks to Flickr and YouTube – I find it both exciting and exhausting to be part of the digital age during its infancy. The opportunities that digital technologies afford are potentially limitless. Currently, though, we are hamstrung by our print-culture perspective. That is, in these decades of transition, we are only slowly moving away from the fixedness of the page, from the ostensibly static nature of word and image on paper, from the lexis of the book, the scroll, the text, the author, the reader.

Indeed, it was twenty-one years ago, in 1993, after substantial work on the early eleventh-century *Beowulf* manuscript held at the British Library, that Anglo-Saxonist Kevin Kiernan recognized the potential of digital photography for reclaiming lost portions of the codex. As Andrew Prescott reveals in his account of events, with a digital camera, Professor Kiernan produced a high-resolution, 21-megabyte image of the badly burned manuscript, originally part of the library of Sir Robert Cotton (the Cottonian library), which formed the basis of the British Library. Having saved the image to an external hard drive, which was subsequently erased by security equipment at Gatwick Airport, Kiernan was glad that he had also transmitted the image via an expensive, hours-long telephone call from London to his home in Kentucky. This transatlantic transfer of a digital image was quite likely the first of its kind – a historic moment indeed, and especially for *Beowulf* studies, because it opened the window for the reparation of many obscured and partial readings damaged by fire and water.

Now, if I want to examine the *Beowulf* manuscript, it is available by way of open access, courtesy of the British Library. I should, however, add a “what you see is what you get” warning: none of the detailed work recaptures lost graphemes, for example; none of it recaptures the high-quality imaging that Kiernan was able to do (and has subsequently gone on to do). Open access, then, is a major advance for the community of scholars and other interested parties who want to see, study, and teach from digital primary sources held at thousands of repositories globally. Rare-books scholar William Noel, of Archimedes Palimpsest fame, has overseen the publication of the holdings of the Walters Art Gallery, the images of which are hosted here and at the Schoenberg Institute for Manuscript Studies at the University of Pennsylvania. The British Museum has long been moving toward open access, with eight hundred thousand digital objects, including eleven thousand manuscripts now available. The Wellcome Institute in London and Trinity College, Cambridge, have just joined the ranks of the open access crowd. When all primary materials are open access, free to reproduce and use in scholarship, then the goalposts will have moved substantially. A real danger, though, is selective open access or digitization: the privileging of the pretty and the re-canonization of already well-known corpora. This has serious knock-on effects for all research and funded projects, when all the project team has access to work with and display freely is a fraction of what actually exists in libraries, archives, and museums worldwide.

Still, there is the thrill: like me this morning in my study at home, all-comers can travel to London to see the Lindisfarne Gospels in a digital display, where the detail of the manuscript is available in a way that simply is not possible otherwise, at a level of decorative amplification equaled only by the experience of an actual user leaning over the codex with magnifying glass in hand. My students, or high school students, or students at the smallest college with no manuscripts of its own can travel with me to Philadelphia to view illustrations of jousting snails, or to Salisbury (whenever the Salisbury material might be made available) to pity the poor mouse squashed inside an eighteenth-century printed Suetonius. Obviously, not all manuscripts and printed books are accessible, often for reasons of preservation. Even
New approaches to book history that have been facilitated by digital tools mean that we can virtually re-create medieval and early-modern libraries.

but in order to do this, theory and practice are essential—and perhaps more important, so is training. Large numbers of amateur and academic crowds can collaborate in inventive and fruitful ways in these initiatives. And at the undergraduate and graduate levels, it is critical that academics involved in digital humanities scholarship provide training to maximize the potential of new and probably game-changing information as well as new mechanisms of analysis. Numerous projects, like DigiPal at King’s College, London, and Inscribe at the IAS in London, are ensuring that carefully explained principles of description and interpretation are available, but only—and herein lies the rub—to those with reliable connectivity. It is arguably the most essential desideratum of digital humanities across the board that scholars equip the next generations with fundamental skill sets to make the most of the mountains of data now available, and the whole ranges yet to come.

Generous training and collaboration in the interstices of science, technology, education, the arts, and humanities will inevitably lead to “creative chances,” to revisit C. P. Snow’s famous 1959 Cambridge lecture on the “two cultures” of arts and sciences. These creative chances are where much true innovation happens, often serendipitously. Even then, as my collaborator, Professor Andrew Prescott, has said to me, it’s not just about the digital. As amazing as it was when Kevin Kiernan sent himself that 21-megabyte image, heralding a small part of this new age, we should understand that something amazing is happening now, too. We need to develop a capacious recognition that developments in nanotechnology, for example, will help us preserve and conserve our immense cultural heritage, or that the Internet of Things will encourage and permit the creation of augmented forms of already-existing books, made richer with added digital data. There will be books and print with conductive ink or inbuilt hyperlinks and multimedia; they will be accretive and connected.

It is only through connections, networks, and collaborations that creative chances can happen—and not just between technologists, scientists, and humanists, but between artists, librarians, archivists, literary specialists, historians, and engineers. We all are cultural heritage practitioners, all digital humanists. For while information might expand, almost as if automatically, knowledge is learned; and while data increase exponentially, wisdom must be acquired. The digital is all very well, but the human outdoes it all.

so, digital processing increasingly permits arts and humanities scholars the ability to gather sources for original purposes, utilizing new methods and new tools.

New tools and developments by project teams include those that permit graphemes to be repositioned in manuscripts in order to reconstruct areas that have been damaged over time; this was recently done with the later tenth-century anthology of Old English poetry, The Exeter Book, as scholars at the University of Glasgow managed to provide hypothetical readings of lacunae caused by a burning firebrand, probably in the later Middle Ages. Multiple forms of librocasm, or the destruction of books, could in fact be silently and virtually mended by new digital tools. As Professor Monica Green at Arizona State University has suggested, new connecting scholars with browser interfaces such as Mirador, and data-modeling services like Shared Canvas. These tools allow virtual textual communities to flourish in interoperable environments, where annotations, explanations, and transcriptions can be used to push scholarship forward, and to build valuable collaborative endeavors in which the latest knowledge can flourish and be openly shared. This will be the outcome, I think, of a new project at Stanford, “Global Currents, 1050 to 1900,” which received support from the National Endowment for the Humanities. Along with colleagues from McGill and Groningen Universities, we will use visual language processing and social network modeling to analyze how literary communication functioned across space and time.

Critical Arts & Sciences
So what are they like, these objects we study in the Literary Lab? They are things like Figure 1: The Correlation between Sentence Types (black vectors) and Common Nouns (in blue and gray).

This image comes from our recent collective pamphlet, “Style at the Scale of the Sentence,” and the full argument can be found at http://litlab.stanford.edu/LiteraryLabPamphlet5.pdf. Here, let me just say that the chart correlates a certain number of words, in blue and gray, with four types of clauses, indicated by the black lines, that are particularly significant in nineteenth-century novels; we spent quite a few hours trying to understand the logic behind this distribution, and others like it. These are the objects we study. Or this – Figure 2: Types of Speaking Verbs by Decade.

The black segments at the bottom of the figure express the declining presence of loud speaking verbs, hence the “silencing” of the English novel that twenty-one-year-old Holst Katsma discovered in our database. This is what our objects are like. And no one had ever seen them because they exist on a different scale from that at which we typically experience literature: one that is simul-

Franco Moretti

Franco Moretti is the Danily C. and Laura Louise Bell Professor and Professor of English and Comparative Literature at Stanford University. He was elected a Fellow of the American Academy in 2006.

The first thing that happens when a literary historian starts using computers to think about literature is that the object of study changes. Not just the tool; the object itself. “The objects studied by contemporary historians” have this peculiarity, Krzysztof Pomian observed some time ago, that “no one has ever seen them, and no one could ever have seen them [. . .] because they have no equivalent within lived experience.” He was thinking of things like demographic evolution and literacy rates, and it’s true, no one can have a “lived experience” of these “invisible objects,” as he also calls them; our objects are different, of course; they are literary ones, but they too have no equivalent within the usual experience of literature.
What is at stake is not reading, it’s the continuity between the experience of reading a book and the production of knowledge.

The eleven different literary genres are represented by the colored word strips, and the twenty-odd verb forms are indicated by the black vectors. No one experiences literature as a scatter plot of verb forms and genres. Reading a novel; watching a play; listening to a ballad: this is the lived experience of literature. And instead, here literature is decomposed into its extremes; but this radical reduction also allows us to see a relationship between the very small and the very large that would otherwise remain hidden: how crucial the passive past simple is for the rhetoric of Gothic novels, for instance, or progressive tenses for the Bildungsroman. And it’s not just a matter of “seeing” the relationship; you can work on it: change the variables, use adjectives instead of verbs to test if they differentiate genres better; exclude function words or include them; you can conduct small experiments with historical evidence. This says something important about the new object of study: it is not something we have found somewhere (in an archive, say); it’s something we have constructed for a specific purpose; it’s not a given, it’s the result of a new practice. A new type of work that, before the advent of digital corpora and tools, was simply unimaginable.

Which brings me to a question I have often been asked, and rightly so: Will the humanities of the digital age lose what has so powerfully characterized them—the experience of reading a book from beginning to end? And, I don’t want to answer for the humanities in general, but for those of us in digital literary studies the answer has to be, Yes: reading a book from beginning to end loses its centrality, because it no longer constitutes the foundation of knowledge. Our objects are much bigger than a book, or much smaller than a book, and in fact usually both things at once; but they’re almost never a book. The pact with the digital has a price, which is this drastic loss of “measure.” Books are so human-sized; now that right size is gone. We’re not happy about the loss; but it seems to be a necessary consequence of the new approach.

Now, let me be clear about this: this does not mean that literary critics, let alone readers in general, shouldn’t read books any more. Reading is one of the greatest pleasures of life; it would be insane to give it up. What is at stake is not reading, it’s the conti-
nty between the experience of reading a book and the production of knowledge. That’s the point. I read a lot of books, but when I work in the Literary Lab they’re not the basis of my work. The “lived experience” of literature no longer morphs into knowledge, as in Ricoeur’s great formula of the “hermeneutic of listening,” where understanding consists in hearing what the text has to say. In our work we don’t listen, we ask questions; and we ask them of large corpora, not of individual texts. It’s a completely different epistemology.

Do we not read at all, then? Well, not exactly. You may have noticed a crazy outlier at the top of Figure 3: each of the strips indicates a set of two hundred narrative sentences from various novels, and that one, from the early chapters of Middlemarch, was so extreme, we of course took those two hundred sentences and read them very, very carefully. The question is, were we thereby reading Middlemarch? I don’t think so. The sentences came from Middlemarch, yes, but they couldn’t be “read” like one reads a novel because they were not continuous with each other; rather, they formed a series only on the basis of a grammatical peculiarity we wanted to investigate. No one could have ever “seen” them together while reading Middlemarch. We were studying Middlemarch, then, but not by reading it.

The objects have changed, and the scale has changed, and the type of work, and of knowledge, and the relationship to reading. And this of course raises all sorts of other questions: are the old and the new type of knowledge – in conflict? Complementary? Independent of each other? And the study of these new objects – what exactly has it achieved? Has it achieved anything? But, for today, this is enough.

**Figure 3. Correlation between Verb Forms (black vectors) and Nineteenth-Century Novelistic Genres**
Joshua Cohen
Joshua Cohen is the Marta Sutton Weeks Professor of Ethics in Society and Professor of Political Science, Philosophy, and Law at Stanford University. He was elected a Fellow of the American Academy in 2002.

My presentation is entitled “If I Had a Hammer.”

I am a philosopher by training and sensibility. A political philosopher. That means I bring a philosopher’s sensibility and training to social and political issues. I have written about equality, liberty, democracy, and global justice.

I think that qualifies me as a kind of humanist. But with apologies to Elaine Treharne, I am not a digital humanist. I am not saying that as a confession; it is not an invitation for congratulation; and it is certainly not a criticism of others who are. I am just giving a self-description. When I am working as a philosopher, I use digital tools as devices of communication and presentation; I don’t use them as tools for analysis. Their availability has not changed the way that I approach questions.

Moving now from self-description to a normative statement: nor do I think they should. I want to use this occasion to reflect on why. As I see the issue, it is all about the questions you ask. Inquiry, including philosophical inquiry, aims to answer questions, and philosophical questions, generally speaking, are not nails awaiting the power of a digital hammer. To put the point less metaphorically, they are not best answered using digital analytical tools.

But – a little more windup before I get to the pitch – who cares that I am not a digital humanist?

Let me try to provoke a bit of interest. Though I am not a digital humanist, I teach a course at Stanford’s design school (the d.school) called Designing Liberation Technologies, in which students try to develop innovative uses of mobile technology to address human development issues in Nairobi’s informal settlements. I am also principal investigator for a few of the projects that have grown out of the course. In addition, I recently coauthored, with eighteen others, an article in *Science* that is partly about uses of digital tools to ensure research transparency in the social sciences. And I work half time at Apple. So I don’t embrace a generic digital Luddism. My reluctance to use digital tools when I am working as a political philosopher comes from a sense of the specifics of the intellectual terrain – the questions at issue – not from a general predisposition.

To explain what I mean, I will tell a couple of stories.

When I am working as a philosopher, I use digital tools as devices of communication and presentation; I don’t use them as tools for analysis. Their availability has not changed the way that I approach questions. . . . nor do I think they should.

In 1983, I was thinking about writing a book on Rousseau. The book – *Rousseau: A Free Community of Equals* – did eventually appear in 2010. (My mind was on other things for some time in between.) But the year is relevant, because I began thinking about the book before I had a computer. As some of you may have observed, Apple’s Macintosh computer was released thirty years ago, certainly long before I could have even imagined doing interesting things with digitized texts. I finished the book in 2010, when those tools were available.

When you start to write a book on Rousseau, one of the things you have to do is figure out his idea of the general will. I was at MIT when I started the book, and I asked my research assistant to go through *The Social Contract* and a couple of other Rousseauan texts to find all uses of the phrase “general will” and its cognates. I cannot recall how long it took her, but I am pretty confident that it took more time than it would have taken when I was finishing up the book several years later. But in the final year of working on the book (2009), I didn’t see any need to check or redo her analysis. I could have quickly redone it to check the results. I could have checked a larger range of texts, including texts by others. And as Franco Moretti’s comments indicated, I could have done some more interesting things than sheer enumeration. I did think about it. But I decided not to do it. Why? Well, for one, I was anxious to be done. But more to the point, it didn’t seem important to me. I had a hammer, but
I wasn’t inclined to see Rousseau as a nail. Why not?

When I was writing this book on Rousseau, I had two connected purposes— one smaller, one larger. The smaller purpose was to provide an interpretation of Rousseau’s political philosophy that came closer than existing interpretations to meeting the standards of analytical philosophy: the theses needed to be more crisply stated than in other interpretations, maybe even more crisply than Rousseau himself. I wanted the book to make the arguments more explicit, assess the arguments, reformulate theses and arguments, and get them to a point where they might be able to withstand critical scrutiny. (I am well aware that Rousseau’s writing is filled with apparent paradoxes defying crisp statement and argument, and aware, too, that some people celebrate those paradoxes as the true measure of his genius. I demur.)

The second, larger purpose served by the first was to present a distinctive political-philosophical outlook, arguably occupied by Rousseau: a position that presents social cooperation as a way to live autonomously, and political participation as the key to sustaining that cooperation, rather than thinking of society as a bargain that requires a sacrifice of autonomy for the safety assured by a commanding authority. Really? Live with other people in a community in which we give the law to ourselves? What a wild idea. Wild and profoundly important. By presenting that outlook, I thought of the book as contributing to a debate in political philosophy. In that debate, the central question is: what should we think about justice, autonomy, obligation, equality, and related ideas? What views on these issues are most reasonable or true?

Text analysis of the kind that I had my assistant do was helpful, but in a very secondary, supportive role. It wasn’t the heart of the matter; because in writing the book, I thought of Rousseau as a partner in the discussion, not an object of theoretical scrutiny. The largest purpose in examining and reconstructing Rousseau’s view was not to explain why he thought what he thought, but to figure out what to think. The assumption behind my writing on Rousseau was that he was a remarkable thinker who offered a distinctive way of thinking about social and political life. So in deciding what to think, it was important to come to terms with other views: not only Rousseau’s, of course, but his, too.

I am emphasizing that the question I wanted to answer was not why Rousseau held the views that he held or how those views were shaped by a mix of personal experience, political context, and regnant ideas. Those are great questions on which much of interest has been written. But I think of those as questions more for historians or social scientists. Writing the book as a philosopher, I wanted to know what to think about the normative questions that Rousseau addressed, not to explain why he thought what he thought (except insofar as figuring that out helped to address the more fundamental question). Had I been addressing the historical question of why he thought, for example, that a social compact is the right way to think about political legitimacy, then I would have wanted tools for handling large data, and not just textual data. But in thinking about the very same normative questions that Rousseau was addressing—in imagining myself in a kind of conversation with him about justice, freedom, and equality—those tools are, at most, of indirect use.

Now you may be thinking that this is all a matter of idiosyncratic personal choice, or something specific to political philosophy. I do not think so. And to explain why, I will switch to a second example.

For the past few years, I have been working with three wonderful philosophers—Alex Byrne from MIT, Gideon Rosen from Princeton, and Seana Shiffrin from UCLA—on the first ever Norton Introduction to Philosophy. The book, out later this year, will cover a broad range of issues: God’s existence, self-knowledge, the mind/body problem, the nature of color, the existence (or not) of numbers, the metaphysics of morals, how to reason about what’s right, and whether equality is essential to justice, among many other things. Our work has been made vastly easier by Dropbox and costless search (and I don’t mind somebody is monetizing all my information). Still, nothing in the substance of the book is significantly different from what it would have been in 1983: that is, nothing has changed as a consequence of the existence of tools for analysis of vast amounts of data. We had a hammer, but did not find nails. In putting the book together, we did not see how thinking about or answering the questions of the book would have been aided by the tools of the digital revolution.

One area where you might think I am off-base is the philosophy of mind. Consider what the philosopher David Chalmers has called “the hard problem of consciousness,” which he describes this way:

It is undeniable that some organisms are subjects of experience. But the question...
of how it is that these systems are subjects of experience is perplexing. Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion? It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does.¹

For about fifty years now, philosophers of mind have been thinking about computational models of mind: whether those models are correct, and how, if at all, they throw light on this hard problem of consciousness – about the relationship between physical processing and an inner life. But that line of thinking, as important and rich as it is, strikes me as completely different from anything in the enterprise of the digital humanities, which is about using digital tools to bring large amounts of data to bear on answering humanistic questions. The philosophical questions that Rousseau was addressing, or the question that Chalmers describes as the hard problem of consciousness – about what to think about fundamental normative and conceptual questions – are addressed by clear thinking, close attention to argument (what follows from what), reflection on the distinctions between cases and among concepts, and an exploration of imaginative hypotheticals and possibilities.

Given the questions, I don’t see how having more data helps.


The French Revolution Digital Archive is an example of an anthology, in the taxonomy of types of research in the digital humanities. Another classic example is the Valley of the Shadow (http://valley.lib.virginia.edu), organized by Ed Ayres and others at the University of Virginia. These are in essence modern digital versions of anthologies that we all have used, but they are much more extensive, they allow more searching, and they allow deeper understanding of texts. Often they involve images as well, as in the case of the French Revolution Digital Archive. They allow scholars to do their own work from a great distance, to collaborate with one another at great distances, and to comment in a very quick, modern way.

The second type is what I would call an interactive scholarly work, with three subtypes. Almost all of these interactive works involve novel visualization techniques, and a great many of them use geographic information systems to present the results of their work. The first subtype is what I might call a simulation model. One example is ORBIS: The Stanford Geospatial Network Model of the Roman World (http://orbis.stanford.edu). It allows one to ask questions about trade routes and communication in the Roman Empire at different times of the year and among the more than seven hundred cities that made up the Roman Empire at its height. It is a focused look at the Roman Empire, but it also allows you to interrogate the data, which have been aggregated from many different sources, in order to form a new hypothesis or create a new research direction. As a model for other projects, ORBIS is marvelous because it doesn’t have to be just about trade routes in the Roman Empire. The ORBIS approach could be employed on other social and human behaviors, using various approaches to aggregating information, then displaying findings in order to allow humans to interrogate it.

Michael A. Keller

Michael A. Keller is the Ida M. Green University Librarian, Director of Academic Information Resources, Publisher of Stanford University Press, and Founder and Publisher of HighWire Press at Stanford University. He was elected a Fellow of the American Academy in 2010.

These remarks will address the taxonomy of types of research output in the digital humanities, about why people undertake research in the digital humanities, and some of the associated issues. I will conclude with a proposition, but let me begin with an announcement. Just today, Stanford’s Libraries, in collaboration with the Bibliothèque nationale de France, released the French Revolution Digital Archive (http://frda.stanford.edu), a searchable, readable collection of 101 volumes of the Archives parlementaires and a collection of many thousands of images from contemporaneous sources. It took seven years for our large transatlantic team to put together the archive, four or five of those years being taken up with determining who owned the copyright on the most recent volumes.
The second subtype is what one might call a reference tool. *Kindred Britain* (http://kindred.stanford.edu) is an example of a reference tool. It shows family relationships among the great and good in Britain. It is a project that started by coding biographical information from about eight sources, and it has many more to add. *Kindred Britain* allows one to see such connections as how Charles Darwin is related to Henry VIII. It has fascinating possibilities for understanding relationships in any society.

Digital humanities allow more people to interact with the humanities. . . . These projects allow much larger audiences to participate in our work, and that is very important.

The third subtype is the Spatial History Lab work (http://spatialhistory.stanford.edu). The example that I like best is The Broken Paths of Freedom project, which focuses on the trading of people as slaves in Brazil. It aggregates data and information across time, and then plots it onto maps. It looks at texts, revealing new knowledge and new understandings at various points in time. The kind of work that Franco Moretti described earlier, researching texts as data by using quantitative methods, would also fall within this subtype. A Stanford graduate student is carrying out another project that seeks to understand the evolution of Portuguese language in mainland Europe as well as Portuguese in Brazil. Around fifty thousand texts have been analyzed so far, some from Brazil and some from Portugal. It is a fascinating linguistic study that in prior years would have been impossible to do. It is also a truly interdisciplinary project, bringing language and literary scholars together with an evolutionary biologist who hopes to understand what is going on in these two lines of development – one in Portugal, one in Brazil – from an evolutionary-biological or mathematical-biological perspective.

Another type that Elaine Treharne mentioned, which she called augmented books, I would call the new narrative. It is a ribbon of text, or an oral narration, interspersed with media objects. It enlivens the repertory of communication. One good example is something called Composing Southern (http://www.jacquelinehettel.com/composing-southern/), which is about the language and culture of the Southern United States. Another is A Game of Shark and Minnow (http://www.nytimes.com/newsgraphics/2013/10/27/south-china-sea/), a project about the seaside culture off the Philippines published by The New York Times with high production values.

So why the digital humanities? Well, these methods of research, of presentation and aggregation, allow new questions to be asked. The old questions are still there, and they can still be asked, but there are new, often interdisciplinary questions, with interdisciplinary answers made possible. Knowledge can be derived that otherwise would not be derived at all. Franco Moretti has called this “the macroscopic study of cultural history.” Another reason that people engage in the digital humanities is because it allows them to engage a lot more data and metadata than ever before. The data could be of a single form, such as text, or it could be of multiple forms, such as text, images, statistics, or maps. This often involves digitization, and encoding is usually necessary. It is the relentless, stupid consistency of computers that makes these kinds of exercises possible, but it is the creative minds of scholars who put the relentlessly stupid computers to work. A fantastic project on Romanesque and Gothic structures in France is under way by Stephen Murray at Columbia University (http://learn.columbia.edu/bourb/). He measures, analyzes, and then visualizes Romanesque and Gothic structures, enabling him to draw conclusions about the types of models that the master builders carried in their heads in order to build everything from parish churches to cathedrals. Using lasers to take dozens of measurements, he can recreate, with great exactitude, the proportions of these structures, and then compare them.

Elaine Treharne also mentioned the possibilities for consultation so that communities of scholars, distant from one another, can work on the same cultural phenomenon or the same collection of objects. More important, in a way, are the new opportunities for scholars to engage students and others, perhaps through crowdsourcing, in research projects. It is a way of bringing novices into the discipline and creating enthusiasm in others by way of one’s own enthusiasm for the possibilities. These types of projects must tackle questions of evidence, attribution, media awareness, logical thinking, organization, and complex arrays of data. Presentation skills in multimedia, digital, and network environments are necessary in these projects. They are great labs, a place for people to get their hands dirty, and a place to learn. Almost all of these projects enable the reuse and remixing of data, whether it is text or images or anything else. We are going to have an expanding collection of information that can be used for many different projects.

Digital humanities, and especially their output, allow more people to interact with the humanities. What a great thing is that!
Our problem in the humanities is in part a problem of our own making. How do we explain to the general intellectual public why it is we do what we do, and why what we do is so interesting? How wonderful many of our projects are, and yet how difficult to convey that enthusiasm to the common man, the so-called man on the street. These projects allow much larger audiences to participate in our work, and that is very important.

There are a thousand flowers blooming all over in the digital humanities right now. There are applications being written or adapted, data being assembled or coded in numerous ways. There are projects that have proceeded over reasonably long periods of time, many with editorial and curation efforts connected to them. The real problem is understanding the spread and the scope of all this work. We don’t have a good way of understanding who is doing what, and how we might intersect with one another. It is important that we do so. On the other hand, it is really important that the scholars who are engaged in this work not have to connect all the dots and all the locations where this work is going on. It is much more important for a thousand flowers to be blooming than to have a single agency harvesting the flowers.

One serious problem is that there is almost no peer review on these projects. Peer review on the importance of the contribution to the commonwealth of knowledge rarely happens with these projects. That is a great shame, because these projects then cannot easily be used in questions of appointment, promotion, and tenure, especially for junior colleagues. Will these interactive projects, which are possibly capstone projects for people receiving an M.A. or Ph.D., be accepted? Not right now (at least not that I know of). We need to separate our judgment of the nature of the contribution from the “gee whiz” aspects of what is being done. They are integrally linked, and yet the nature of the question posed and the answer is quite important. We have in this field a kind of echo chamber. Some of us go to the same conferences. We need to find ways to expand our understanding of what is going on out there so that we might adopt and adapt methods that others have developed.

There is the question of sustainability. Many of the best projects are undertaken by excellent principal investigators with a team of students and staff, but how do these projects look once the PI has stepped away and moved on to another study? How do we sustain these projects so that audiences can come back not just for a decade, but for a century, or five centuries? How do we make it possible for this knowledge to live on, as knowledge has lived on in book form, on paper, for many centuries? How does a library get engaged with this to host and preserve the fundamentals of these research products, so that they may be useful to others in another place and another time? Enabling and supporting the reuse and remixing of the data are essential.

The suitability of the technical architecture to support sharing of data, the quality and extent of metadata, is important. How do we understand the data and the coding of the data? Are APIs (application programming interfaces) available to make the data accessible and to interact across projects? On the projects themselves, is there help to make the novice easily acquainted with the possibilities? Are there explicit models of interaction? Are there user-friendly interfaces? Are there models or samples that a teacher in a K-12 situation might employ to take advantage of the work that has been done by these digital projects and these digital humanists?

How does a library host and preserve the fundamentals of these research products, so that they may be useful to others in another place and another time?

I mentioned copyright when I talked about the French Revolution Digital Archive. These are serious issues, and I didn’t mean to make light of them, but there is also the question of authorship. Who is the author? Who is the principal investigator? Is the principal investigator or the principal developer the author? And what about the scads of students? In the sciences, we see articles with anywhere from ten to a hundred coauthors. What does it mean to be an author in these cases? How do we recognize the vital contributions made by multiple contributors?

Are there licensing issues with the applications and with the data? Should we know about that if we want to try to reuse the data? If they are open source applications, how can we be sure that they will be viable and usable, not ten years hence, which is already a problem, but fifty years? Our technology is not ripe enough to allow that kind of long, sustained access to these projects. So with special regard to one of the big issues, I propose that we think about engaging publishers and libraries in these projects. The issue of peer review, of technical review, and of marketing and distribution are serious issues, and many of them are issues that publishers deal with all the time.