In the age of explorations

Willard McCarty
Centre for Computing in the Humanities
King’s College London
staff.cch.kcl.ac.uk/~wmccarty

Closing keynote for Exploring the Archive in the Digital Age,
King’s College London, 8 May 2010

1. Setting the historical stage

In her keynote address yesterday, Diana Taylor said (here quoting from my notes) that the manner in which the digital changes our ways of knowing have not yet come into focus. She’s largely right, but I want to talk now about a specific coming into focus of what I think will prove a far-reaching change for scholarship. I begin, however, by setting an historical stage on which change is being enacted. Without history our arguments about what is happening and might happen are rootless.

When in the mid 1980s I first began to take a serious interest in humanities computing – the field which all the digital humanities share – one of the commonest ways of deflecting the threat posed by the machine was to say that it was “just a tool”. In my experience this statement was used dismissively by anxious colleagues to relieve themselves of the fear that the computer was or could be anything more than an efficient servant – something which merited no serious consideration intellectually and so would not substantially alter what they did – only, somehow, make it better. Their dismissive stance was, I concluded, a defensive posture against soaring claims from the artificial intelligentsia, reported and often amplified in the press, that (as was actually asserted in the late 1950s) computers already could think, learn and create. Manufacturers at first joined in but, Pamela McCorduck recalls,

sales executives at IBM began to grow nervous lest the very machines they were trying to sell prove so psychologically threatening that customers would refuse to buy them. Thus they made a deliberate decision to defuse the potency of [AI programs] by conducting a hard-sell campaign picturing the computer as nothing more than a quick moron. Countess Lovelace’s dictum, that a machine can do nothing more than we tell it to do... was raised to a universal truth, and parroted by every sales and service person connected with the company. It came to be a popular idea, a sort of slogan of the backlash, and is offered to this day by people who feel threatened by the idea of machine intelligence. (McCorduck 1979: 159f; cf Bowden 1953: 317; MacKay 1991: 171)

It is no wonder, then, as one scholar wrote, [FIGURE 1] that nervous humanists expressed existential “fear and trembling” (Nold 1975), or as others said, took shelter
in the reassurance that scholarship itself was a uniquely human activity (Pegues 1965), sounded a warning against “the more sensational attempts” at venturing beyond concordance-production (Anon: 344) or thundered against what F. R. Leavis called “the professional subhumanities of computerial addiction” (1970: 154). The literature suggests that scholars weren’t (as some are now) in fear of losing their jobs, rather felt their intellectual raison d’être threatened.

Thus while factory workers feared that the new wave of smart machines would deprive them of their livelihood or de-skill their jobs, as in fact widely happened (Zuboff 1988), intellectuals worried that an intelligent machine might, like the discoveries of Copernicus, Darwin and Freud, force a downgrade of the human self-image. Freud had declared in 1917 that psychoanalysis was “the third and most irritating insult... flung at the human mania of greatness” because it showed that we have little idea of who we are, and that what we can see is deeply disturbing, even repugnant. Hence, he noted, the great disturbance caused by his findings, the outrage among colleagues and in the wider society (1922/1917: 246-7). As Bertrand Russell said of Joseph Conrad, Freud seemed to reveal “civilized and morally tolerable life as a dangerous walk on a thin crust of barely cooled lava which at any moment might break away and let the unwary sink into its fiery depths” (1956: 82). During computing’s incunabular period, from the end of World War II to the introduction of the Web in August 1991 (significantly coincident with the Cold War), its threat may have been dressed in different imagery – mostly, given the time, Orwellian – but computing was no less of a threat to human self-conception and so provoked anxious uncertainty.

Calling the computer “just a tool” was also in part the reaction of practical people to philosophical attempts at understanding the computer as a machine to think with. But I suspect that most scholars who used this phrase had little familiarity with the skilled use of tools and so were unaware that no tool in skilful hands is just a tool. The artists and musicians involved with computing at the time – people who would have known better about tool-use – spoke very differently (Brown et al 2008). Artist and computer graphics consultant Frank Dietrich, for example, described the relationship between British painter Harold Cohen and his drawing software AARON as “a functioning harmonic symbiosis between man and machine” (Dietrich 1986: 162) [FIGURE 2]. Douglas Engelbart, inventor of the mouse, thought in much the same way, arguing for the potential of the computer to augment human intelligence dynamically rather than to serve or replace it (1962). More recently, my King’s colleague John Bradley has written software for note-taking that embodies and furthers such thought.

We seldom hear the phrase “just a tool” today because, I think, we have so successfully domesticated and house-trained the computer that the statement seems too obvious to make. We’re so busy getting on with business – doing our research
with computing, encoding texts, building helpful resources and the like – that we overlook the Blakean corrosive that computing still is. All these activities are good when done well, but they are not good enough to do justice to our disciplines and ourselves. As the literary critic Louis Milic wrote in 1966, our “satisfaction with such limited objectives denotes a real shortage of imagination.... We are still not thinking of the computer as anything but a myriad of clerks or assistants in one convenient console” (1966: 3-4). In a sense the fear then expressed by those who felt it was healthier. It meant at least a modicum of sensitivity to that Blakean corrosive.

I am not recommending that we learn to be afraid again, rather that we open our eyes to that to which fear was then a reaction. What we need now, I think, is, in Roman Jakobson’s words, to assume the artist’s task of making “the ordinary strange” (Bruner 1991: 13) and so able to awaken what a former teacher of mine used to call a “beginner’s mind”. To understand computing’s potential beyond what we now know – and to see what in spite of our pedestrian focus it is yeast-like doing right now – we have to recover its perceptual freshness.

Freud’s argument for the successive downgrading of human self-regard, to which computing’s reinvigoration of the problem of the human adds a new chapter, strikes more forcefully than you may suspect. Its power derives from a defining cultural pattern traceable to the Hebrew Bible: of history as a sinusoidal alternation of redemptive ups and degenerate downs going on and on through time, each time requiring the people to refigure themselves – from Edenic freedom to Egyptian slavery to Red Sea passage to Wilderness wandering to Moses’ Pisgah-sight [FIGURE 3] and so forth and so on, to “a paradise within thee, happier far” at the end of time, as the Archangel Michael says to Adam in Milton’s Paradise Lost (12.587). In other words, we have a very long tradition of re-envisioning our goals and ourselves. This is how, I suggest, we make sense of what’s happening now.

2. Exploration

My title, “In the age of explorations”, refers deliberately to the so-called Age of Exploration, a period from the 15th through the 17th Centuries when Europeans – the Portuguese and Spanish first, then the French, English and Dutch – explored and mapped the world. I justify the reference on two grounds: first, more obviously, that the actions and imagery of genuine, curiosity-motivated research allude non-trivially to exploration of terra incognita; second, that the digital medium has, since its introduction into the scholarly life, begun to work on us much the same kind of (dare I say it?) revolutionary change which the Age of Exploration brought to biological taxonomy, when the variety of hitherto unknown species provoked the overthrow of European folk-biological by scientific classification (Atran 1990) [FIGURE 4]. This is not at all to suggest that the humanities are being made more scientific by computing, rather that our understanding of the artefacts we study and
the resources we use to study them is profoundly affected by the extent to which we command their variety, and that by bringing greater masses of raw material into view computing becomes midwife to cognitive change.

I accept without dispute that we are in an age that may be characterised by its preoccupation with digital media. Aspects of this preoccupation, specifically its trendiness, will pass. But I think it is essential to realise not only that we are in an age far more of beginnings than of accomplishments but that especially for computing the emergent condition is where we always are.

More than 15 years ago the late and much-missed Paul Evan Peters declared that we had just come out of humankind’s palaeo-electronic period; he welcomed us to the meso-electronic (1994). Technologically we have made some strides since then, but for us scholars the issue is not so much that progress has been slow where it most concerns us, rather that our concerns are different. Comparison of the humanities and the sciences, which establishes the context for our meditations on computing in the humanities, offers several ways of thinking about the problem, from Wilhelm Windelband’s contrast of the particularising humanities and the law-seeking sciences in the late 19th Century (1894) to Jerome Bruner’s modification of it in the late 20th, in which the humanities are not so much focused on the particular as on opening up “the alternativeness of human possibility” (1986: 53). Bruner’s generous formulation identifies what seem to me two enduring ways of knowing, but C. P. Snow’s two-cultured distinction has, since computing entered our scene, become considerably more difficult to maintain. Like it or not, computing as product of the technosciences blurs this distinction, making the difference more difficult to grasp, which is to say, richer, more complex and (again I say it) revolutionary.

On the one hand we tend to regard the humanities as subject to major and minor stylistic changes but fundamentally unaffected by progress, implying that our disciplines stand in isolation from the great engine of 21st-century change. (Hear here an echo of the just-a-tool argument.) On the other hand, it is now quite clear that computing has affected nearly all of the ways in which research is done. This is a problem for isolationists because technological progress is undoubtedly real: the greater capacity and faster speed of computing machinery without question make it better. Whatever may be said about oldsters’ age-old complaints of a world speeding up out of control, acceleration of technological change has become a fact of life and can hardly be controlled. It has become an economic necessity. Isolationists cannot claim that the humanities nevertheless remain unaffected because the ever increasing speed of the machine’s response over the greater volume of data it can hold changes how we (limited and mortal as we are) think with it. Indeed, close observation reveals that the development of computing is demand-driven as much or more than it drives demand. We want the better equipment. Furthermore, computing is not one thing. Alan Turing’s scheme gives us not a single design but a
way of inventing an indefinite number of computings, limited only by the human imagination. So computing not only gets faster and more capacious but also more talented. It is a luxurious jungle growing at speed. It is desire industrialised. Our desire.

At the same time computing is by Turing’s rigorous definition severely constrained in what it can do for us. Put as simply as possible, it can compute what is computable, not what is not, or in other words, computability both defines what it means to state something exactly and gives us a universal process by which anything that is exactly stated can be processed. The surprise is that within such a Spiegelkammer so much can be done. Some of its forms are now so pervasive and seamlessly efficient – in wristwatches, pacemakers, refrigerators, greeting cards, cash machines and automobiles, for example – that we might be tempted to think computing has reached a proliferating maturity, that one day soon the adjective “digital” will no longer be needed. But, again, computing’s closed world never reaches, rather it is always reaching toward equilibrium: “digital” will always be taking on new meanings in new devices, and it will always prove incapable of doing something or other. Furthermore, it makes no sense to draw a principled line between what computing can and cannot accomplish (in fear still of how far it may go). All arguments of the kind are fatally weak. The point, rather, is to use it in whatever up-to-the-minute form it might take, for asking ever better questions that will themselves suggest new forms of computing to be devised and applied. The computer is a forward-looking, experimental machine. Let’s see how far we can get.

Do scholars therefore get better at what they do? Attempts to answer this question usually fall back on the old argument that computing does the donkey-work so that we can be insightful, creative and so forth – i.e. uniquely human and thus safe from the Copernican-Darwinian-Freudian-computational downgrade. Apart from the fact that computing leads to considerable donkey-work of the technical kind (do we actually ever save time? I don’t), we know from work in archives and work elsewhere how dangerous it is to assign any aspect of what we do to drudgery, how isolation from the data sets thought adrift. I for one strongly prefer to let ideas bubble up from raw material rather than to enter the archive with mental template ready to hand. Who knows in such circumstances what supposed drudgery will yield, what opportunities for meaningful surprise? Who would hand over such opportunities to an obedient assistant except as a precious gift, in the hope and full expectation that he or she will strike out for “pastures new”?

Viewing the computer as servant or slave, fit in principle only for obedience and drudgery, thus buries its intellectual potential. And it traps us because, Frederic Jameson notes in Valences of the Dialectic, “the slave is not the opposite of the master, but rather, along with him, an equally integral component of the larger system called slavery or domination” (2009: 20). This is why I argue that archival engagement with
data is the way ahead for computing in the humanities. We don’t want things delivered however conveniently, we want to be *there*, messing with stuff, finding things out for ourselves as directly as possible.

## 3. Machine and mind

Where this desire takes us is ultimately to our own ignorance of how we might render our core activity, interpretation, into a digital medium. What we now do with computing, and have done quite well in many cases, is to serve the interpretative act by delivering nuanced information, i.e. by embedding interpretations in metadata or in data structures. This is fine though inadequate, as I noted earlier. What we have not done, and don’t know how to do, is to build the computational equivalent of a musical instrument, i.e. an instrument which allows the interpreter’s mind to engage at the point and in the moment of interpretation. In the early days it seemed to many that the digital computer was “faster than thought” because it could perform numerical and logical operations at dazzling speed (Bowden 1953) [FIGURE 5]. But, amidst all the excitement which computing machinery then caused, some realised that the basic question was not so much how to implement thinking as to consider what thinking is. As it shifts in conception from doing sums toward interpreting cultural artefacts, the computing we know (note: not the only one possible) seems not just progressively slower but more and more the wrong kind of thing. We begin to notice that it is based on a concept of mind which more than meets it halfway, that it mirrors exactly the kind of mind which it is designed to serve. The philosopher Vernon Pratt suggested about 25 years ago that we take “a fresh look at the picture, partly in light of those things that the machine becomes able to do for us, and then [change] our ideas about which aspects of thinking or mentality are truly important – that is, about what thinking is” (1987: 245).

This is not just a philosophical question, or rather, it is a philosophical question best pursued nowadays from the multiple perspectives of the cognitive sciences. Christopher Longuet-Higgins, coining the term, defined this plural field in his 1973 commentary on the state of artificial intelligence in the U.K.: “All those sciences which are directly relevant to human thought and perception” (1973: 37). More recently Margaret Boden has defined it more narrowly as “the study of *mind as machine*” (2006: 9). For my purposes here it doesn’t really matter whether you believe that your or anyone else’s mind is a machine, even a very special kind, or are willing to entertain this belief to see what comes of it. Actually, as Boden’s two-volume history of cognitive science demonstrates, quite a bit does come from this belief, but again you don’t need to embrace it. The meeting of mind and machine that I identified as our goal only requires of you that you believe the mind can simulate the machine a human mind designed, and in fact does so in moments of perfect fusion, when you forget that you are using the computer and pay attention only to what you are doing with it. The philosopher Michael Polanyi describes this more
accurately as attending from something in order to attend to something else better or at least differently (1983/1966: 9-10).

The contribution of the humanities in all of this is to take Vernon Pratt’s “fresh look... partly in light of those things that the machine [is] able to do for us” and partly in light of those other things we can imagine on the basis of what, I’d think, is the most demanding and complex material human beings have made. The contribution of the humanities is to give the cognitive sciences the most challenging of all problems to work on.

4. Implications of the online archive

In the remainder of this lecture, I am going to turn from those genuinely interpretative uses of the machine which define our imaginative horizon and in all likelihood will not be realised for some time but give us an idea of our trajectory. Instead I am going to ask what the machine primarily does for us scholars now, and what long-term implications arise from our primary machine-use.

Clearly the primary use for research is enabling access to sources and resources online. But what is it actually like to do research online?

You will all be familiar with the mixture of problems, worries and garden-variety successes common to everyone’s online practices. No doubt there is much to learn from studying these practices, but here I am going to concentrate on the seemingly less problematic set of activities in which we engage when we use what is loosely called a digital library. For my purposes a digital library is a large collection of scholarly literature organized in some obvious way and supplied with a standard search mechanism. For my example I will use JSTOR, with which I assume everyone here is familiar [FIGURE 6]. First, however, a brief account of how search-mechanisms are designed and used.

In its classical form the field of research in computer science known as “information retrieval” attempts to achieve results along two dimensions: precision, the measure of fidelity to the user’s expressed intentions; and recall, the measure of completeness, or percentage of relevant items in a collection that a query actually finds. To serve the maximisation of both – greatest precision with highest recall – query languages have been designed with operators which the user combines in an exacting syntax to recall as close to all qualifying items as possible. Google surprised many people educated in classical information retrieval by shifting the burden from precise and comprehensive statement in a query language to reliance on recorded choices people had already made, i.e. from specifying as exactly as possible what one wants on the one hand to accepting what the majority have shown that they want on the other.
Nevertheless on collections much smaller and more coherent than the whole Web, query languages remain the dominant means of access.

Query languages can be useful, but beyond the simplest concatenations of AND, OR and NOT they take too much time in the devising to be practical, in part because they implicitly assume that the collected material reflects the same kind of logical syntax, which in our case it does not. Even in the simplest of concatenations a fair amount of conscious sophistication is needed to come up with a reasonable expression with any chance of success. My point – the major point I wish to make in this lecture – has to do with the consequences of the tendency to abandon the explicit complexity of a query language within a digital collection and go for a simple, straightforward string-search, as I suspect almost all of us do most of the time.

I once knew a Sinologist at Toronto who told me a story about recall that illustrates my point. For years he had been using a well-known classical Chinese encyclopaedia that had no index, for the very good reason that Chinese is actually quite difficult to organize indexically. (Without an alphabet, one must rely on common elements or “radicals”, the total number of strokes in a character and a fair amount of searching through lists.) One day my friend saw an advertisement for an index that a modern scholar had at last devised. He reached for his pen to fill out the order form but stopped, and in fact never ordered the index. He realised that in all the years he had used the encyclopaedia he had on many occasions found something far more valuable and interesting to him in his laborious searches than that for which he had started to look. He knew that serendipitous discovery would be taken from him under pressures of time if he had the index. He was a wise man.

We often somehow know more than we can say, or the subject we are investigating requires us to know more than we can specify or know that we might attempt to specify. For this reason the kind of searching that computer scientists call “dumb” – searching based merely on a single character-string with no qualification, yielding high recall at the cost of poor precision – turns out to be much more useful, smarter and far more demanding of our intelligence than we are likely to suspect. Let me illustrate with a brief example. I will depend on you to change whatever needs to be changed so that you can imagine yourself in the situation I depict.

Let us suppose a question arising from an investigation of structures or qualities that seem to emerge during the reading of a novel, in some sense implied by but, strictly speaking, never set down in the novel. Let us suppose that you get interested in this phenomenon, for which the term emergence seems apt. You are stuck, however; you lack the vocabulary with which to discuss what seems a crucial aspect of this novel but cannot find assistance in any of the obvious literary-critical sources. Let us suppose further than you turn to JSTOR, which as a collection of scholarly articles across a large number of disciplines is likely to give you a grip on the idea, if it
exists, and a handle on where to look for its development. As happens full-text searching for the adjective “emergent” yields in the first 25 of the 53,991 hits articles in at least 10 disciplines, namely, design theory, operational research, ecology, sociology, zoology, economics, philosophy, developmental psychology, anthropology and computer science. Glancing quickly at some of these you discover stimulating lines of thought. Investigation of the topic by other means, such as Wikipedia and Amazon, leads you to other areas and gives you additional things to read. Sampling is, of course, necessary, but with reasonable rapidity you close on areas of enquiry in the major fields and so draw up a reading list – all from your computer at home. I have done exactly this sort of thing on numerous occasions.

If you yield to the temptation to go further you subject yourself to two strong constraints apt to shape the work you subsequently do.

The first is necessary redeployment of effort. Any time spent looking elsewhere is necessarily taken away from the time you have to burrow into scholarship in your specialist area. In other words, going wide means being less deep but over a greater number of sources. We can choose to call those who do so intellectual butterflies and rail against the unavoidable superficiality of their work, countering with the conventional phallocentric metaphors of depth, profundity and the like. Or, more wisely I think, we can observe that going wide is what is happening and choose to ask how it can be done well. We can ask, what difference is made by paying attention laterally rather than hierarchically? What kind of truth to the subject is entailed if we replace metaphors of depth with metaphors of breadth in our thinking? Richard Rorty, citing Hans-Georg Gadamer, argues for the nominalist view that there are no privileged descriptions, therefore none “more true to the nature of [an] object than any other”. Hence, he says, “to understand something better is to have more to say about it.... the more descriptions that are available, and the more integration between these descriptions, the better is our understanding of the object identified by any of those descriptions” (2000: 22-3). I am not going to take Rorty’s argument further – I haven’t the time nor the philosophical skills. Nor I am going to argue that you must sign up to nominalism, turn against essentialism or proclaim a new chapter in the sociology of knowledge. My point, rather, is that we need to recognise in our own and in other people’s information-seeking behaviours the strong tendency to go wide rather than deep online, understand its potential, work out how to do it well and think about its implications. It seems to me that these implications are, to put the matter mildly, non-trivial.

It should be obvious from my example and stand to reason that going wide rather than deep involves interdisciplinary research skills. Developing and deploying these skills constitute quite a challenge. There are, first, the formidable difficulties of attempting to think your way into disciplines other than your own. I will return to these in a moment. But before you even get started, you are likely to be hopelessly
confused by the discourse of interdisciplinarity, which is severely muddled by the prevalent tendency to treat it as what Peter Galison calls a transcendental virtue (2004: 380), i.e. something that is good without qualification – a term you must use, for example in grant applications, without stopping to understand what might be entailed. My familiarity with reports from actual interdisciplinary research projects, in the few instances where these have been closely observed and their progress described, suggests how much thought and care is required, how ill prepared researchers tend to be. If you choose to look into the literature on the topic, you are also likely to be confused, indeed perhaps put off, by the seemingly endless, neo-scholastic discussions of what interdisciplinarity is, how it differs from trans-, multi- and other disciplinarities. The mistake here (made repeatedly, for example in the forthcoming Oxford Handbook of Interdisciplinarity) is, I think, to reify a way of acting by using an abstract noun to name it. Once you see it not as a thing to be defined but as acting in a particular way or set of ways, then Karin Knorr Cetina’s notion of disciplines as “epistemic cultures” (1999) – or, as Clifford Geertz seems once to have remarked, “savage tribes” (Apter 2007: 112) – comes almost without bidding. It is then relatively straightforward to see that the literature on ethnography is where one must start and that it is richly rewarding.

The question to ask is how to do interdisciplinary research, and do it well, as you attempt to assimilate the articles turned up in a search such as I have described. Stanley Fish has argued quite sensibly that achieving a fully interdisciplinary stance is impossible because there is no neutral ground (1989). No large undifferentiated field of knowing is possible, nor desirable, because as Northrop Frye has argued, “every field of knowledge is the centre of all knowledge” (1988: 10), the centre from which you begin to expand into other worlds of knowledge. But Fish then goes on to argue that since perfect interdisciplinarity is impossible it should not be tried. This is remarkably like the theological argument that since one cannot be perfectly good, the attempt shouldn’t be undertaken. Nonsense. As in the moral life, so in the scholarly: the whole point is to try, to do the best you can, as aware as possible of the perils, as cautious as you need to be, keeping in mind the old English proverb Edmund Spenser used in The Faerie Queene: “Be bold, Be not too bold”!

Apart from doing interdisciplinary research for the last 25 years to help interconnect humanities computing with the disciplines it serves and from which it must learn, I designed late last year a course for the King’s Graduate School which I’ve called “Exploring Disciplines”. It was offered for the first time last Spring and was, I am happy to say, well received. For more on the subject see that course, or better yet, if you are in a position to take it, do so.
5. Conclusion

The digital archive, or the digital library viewed archivally as a set of primary sources to be trawled, lies all before us. It is hard not to conclude from even such brief experience we have had – brief in the cultural time-scale which matters – that new ways of knowing the knowledge we have inherited are on offer or are even unavoidable. But, speak such promises of the new in this digitally preoccupied age and you put yourself in immediate danger of resonating with the hyperbolic cant which so vexes everything having to do with our beloved machine. [FIGURE 7] Raymond Williams is possibly the best guide to the technological determinism that, because of this cant, seems the next step in my argument and must at all costs be avoided – the assertion that because of computing, because of the Web, thus and such is happening to scholarship. This assertion, he would argue, is a one-sided oversimplification of limits set and pressures exerted, by people and their inventions, “within which variable social practices are profoundly affected but never necessarily controlled” (2005/1974: 133). We could, of course, act as if we were controlled, making it so in the only way it could be. But I greatly prefer Williams’ argument because it preserves the unknown, leaves open the possibility of freedom without denying constraints to it. The world of scholarship I recommend to you is so much less known, so much richer and more interesting than the effects of technological causes, whatever the thrill of thinking oneself a part of invincible progress and whatever the comforts of avoiding uncertainty. Let us not be controlled. Let us do something new. [FIGURE 8]
Works cited.


McCarty, In the age of explorations / 13