

Killer Applications in Digital Humanities

Patrick Juola
Duquesne University
Pittsburgh, PA 15282
UNITED STATES OF AMERICA
juola@mathcs.duq.edu

August 31, 2006

Abstract

The emerging discipline of “digital humanities” has been plagued by a perceived neglect on the part of the broader humanities community. The community as a whole tends not to be aware of the tools developed by DH practitioners (as documented by the recent surveys by Siemens et al.), and tends not to take seriously many of the results of scholarship obtained by DH methods and tools. This paper argues for a focus on deliverable results in the form of useful solutions to common problems that humanities scholars share, instead of simply new representations.

The question to address is what needs the humanities community has that can be dealt with using DH tools and techniques, or equivalently what incentive humanists have to take up and to use new methods. This can be treated in some respects like the computational quest for the “killer application” – a need of the user group that can be filled, and by filling it, create an acceptance of that tool and the supporting methods/results. Some definitions and examples are provided both to illustrate the idea and to support why this is necessary. The apparent alternative is the status quo, where digital research tools are brilliantly developed, only to languish in neglect and disuse.

1 Introduction

“The emerging discipline of digital humanities”. . . . Arguably, “digital humanities” has been emerging for decades, without ever having fully emerged. One of the flagship journals of the field, *Computers in the Humanities*, has published nearly forty volumes, without having established the field as a mainstream sub-discipline. The implications of this are profound; tenure-track opportunities for DH specialists are rare, publications are not widely read or valued, and, perhaps most seriously in the long run, the advances made are not used by mainstream scholars.

This paper analyzes some of the patterns of neglect, the ways in which mainstream humanities scholarship fails to value and participate in the digital humanities community. It further suggests one way to increase the profile of this research, by focusing on the identification and development of “killer” applications (apps), computer applications that solve significant problems in the humanities in general.

2 Patterns of Neglect

2.1 Patterns of participation

A major indicator of the neglect of digital humanities as a humanities discipline is the lack of participation, particularly by influential or high-impact scholars. As an example, the flagship (or at least, longest running) journal in the field of “humanities computing” is *Computers and the Humanities*, which has been published since the 1960s. Despite this, the impact of this journal has been minimal. The Journal Citation Reports database suggests that for 2005, the impact factor of this journal (defined as “the number of current citations to articles published in the two previous years divided by the total number of articles published in the two previous years”¹) is a relatively low 0.196. (This is actually a substantial improvement from 2002’s impact factor of 0.078.) In terms of averages from 2002–4, *CHum* was the 6494th most cited journal out of a sample of 8011, scoring in only the 20th percentile. By contrast, the most influential journal in the field of “computer applications,” *Bioinformatics*, scores above 3.00; *Computational Linguistics* scores at 0.65; the *Journal of Forensic Science* at 0.75. Neither *Literary and Linguistic Computing*, *Text Technology*, nor the *Journal of Quantitative Linguistics* even made the sample.

In other words, scholars tend not to read, or at least cite, work published under the heading of humanities computing. Do they even participate? In six years of publication (1999-2004; volumes 33–38), *CHum* published 101 articles, with 205 different authorial affiliations (including duplicates) listed. Who are these authors, and do they represent high-profile and influential scholars? The unfortunate answer is that they do not appear to. Of the 205 affiliations, only 5 are from “Ivy League” universities, the single most prestigious and influential group of US universities. Similarly, of the 205 affiliations, only sixteen are from the universities recognized by *US News and World Report* [USNews, 2006] as one the top 25 departments in in any of the disciplines of English, history, or sociology. Only two affiliations are among the top ten in those disciplines. While it is of course unreasonable to expect any group of American universities to dominate a group of international scholars, the conspicuous and almost total absence of faculty and students from top-notch US schools is still important. Nor is this absence confined to US scholars; only one affiliation from the top 5 Canadian doctoral universities (according to the 2005 MacLean’s ranking) appears. (Geoff Rockwell has pointed out that the MacLean’s rankings are

¹<http://jcrweb.com/www/help/hjcrqls2.htm>, accessed June 15, 2006

School	Papers (2005)	Papers (2006)
USNews Top 10	7	4
Harvard		
Cal-Berkeley	1	1
Yale		
Princeton	1	
Stanford	1	2
Cornell		
Chicago		
Columbia	1	
Johns Hopkins		
UCLA		
Penn		
Michigan-Ann Arbor	2	
Wisconsin-Madison		
UNC-Chapel Hill	1	1
MacLean's top 5	2	3
McGill		
Toronto	1 (3 authors)	1
Western		1
UBC	1	1
Queen's		
Ivies not otherwise listed	4	6
Brown	4 (one paper 2 authors)	6
Dartmouth		

Table 1: Universities included for analysis of 2005 ACH/ALLC and 2006 DH proceedings

not necessarily the “best” research universities in Canada, and that a better list of elite research universities would be the so-called “Group of 10” or G-10 schools. Even with this list, only three papers — two from Alberta, one from McMaster – appear.) Australian elite universities (the Go8) are slightly better represented; three affiliations from Melbourne, one from Sydney. Only in Europe is there broad participation from recognized elite universities such as the LERU. The English-speaking LERU universities (UCL, Cambridge, Oxford, and Edinburgh) are all represented, as are the universities of Amsterdam, Leuven, Paris, and Utrecht despite the language barrier. However, students and faculty from Harvard, Yale, Berkeley, Toronto, McGill, and Adelaide — in many cases, the current and future leaders of the fields — are conspicuously absent.

Perhaps the real heavyweights are simply publishing their DH work elsewhere, but are still a part of the community? A study of the 118 abstracts accepted to the 2005 ACH/ALLC conference (Victoria) shows that only 7 included affiliations from universities in the “top 10” of the USNews ranking. Only two came from universities in the “top 5” of the Maclean ranking, and

only 6 from Ivies (Four of those six were from the well-established specialist DH program at Brown, a program unique among Ivies.) A similar analysis shows low participation among the 151 abstracts at the 2006 DH conference (Paris). The current and future leaders seem not to participate in the community, either.

2.2 Tools and awareness

People who do not participate in a field cannot be expected to be aware of the developments it creates, an expectation sadly supported by recent survey data. In particular, [Siemens et al., 2004, Toms and O'Brien, 2006] reported on a survey of “the current needs of humanists” and announced that, while over 80% of survey respondents use e-text and over half use text analysis tools, they are not even aware of “commonly available tools such as TACT, WordCruncher and Concordancer.” The tools of which they are aware seem to be primarily common Microsoft products such as Word and Access. This lack of awareness is further supported by [Martin, 2005] (emphasis mine):

Some scholars see interface as the primary concern; [electronic] resources are not designed to do the kind of search they want. Others see selection as a problem; the materials that databases choose to select are too narrow to be of use to scholars outside of that field or are too broad and produce too many results. Still others question the legitimacy of the source itself. How can an electronic copy be as good as seeing the original in a library? Other, more electronically oriented scholars, see the great value of accessibility of these resources, but are unaware of the added potential for research and teaching. The most common concern, however, is that scholars believe they would use these resources *if they knew they existed*. Many are unaware that their library subscribes to resources or that universities are sponsoring this kind of research.

Similarly, [Warwick, 2004a] describes the issues involved with the Oxford University Humanities Computing Unit (HCU). Despite its status as an “internationally renowned centre of excellence in humanities computing,”

[P]ersonal experience shows that it was extremely hard to convince traditional scholars in Oxford of the value of humanities computing research. This is partly because so few Oxford academics were involved in any of the work the HCU carried out, and had little knowledge of, or respect for, humanities computing research. Had there been a stronger lobby of interested academics who had a vested interest in keeping the centre going because they had projects associated with it, perhaps the HCU could have become a valued part of the humanities division. That it did not, demonstrates the consequences of a lack of respect for digital scholarship amongst the mainstream.

3 Killer Apps and Great Problems

One possible reason for this apparent neglect is a mismatch of expectations between the expected needs of audience (market) for the tools and the community's actual needs. A recent paper [Gibson, 2005] on the development of an electronic scholarly edition of *Clotel* may illustrate this. The edition itself is a technical masterpiece, offering, among other things, the ability to compare passages among the various editions and even to track word-by-word changes. However, it is not clear who among *Clotel* scholars will be interested in using this capacity or this edition; many scholars are happy with their print copies and the capacities print grants (such as scribbling in the margins or reading on a park bench). Furthermore, the nature of the *Clotel* edition does not lend itself well either to application to other areas or to further extension. The knowledge gained in the process of annotating *Clotel* does not appear to generalize to the annotation of other works (certainly, no general consensus has emerged about "best practices" in the development of a digital edition, and the various proposals appear to be largely incompatible and even incomparable). The *Clotel* edition is essentially a service offered to the broader research community in the hope that it will be used, and runs a great risk of becoming simply yet another tool developed by the DH specialists to be ignored.

Quoting further from [Martin, 2005]:

[Some scholars] feel there is no incentive within the university system for scholars to use these kinds of new resources.

— let alone to create them.

This paper argues that for a certain class of resources, there should be no need for an incentive to get scholars to use them. Digital humanities specialists should be in a unique position both to identify the needs of mainstream humanities scholars and to suggest computational solutions that the mainstream scholars will be glad to accept.

3.1 Definition

The wider question to address, then, is what needs the humanities community has that can be dealt with using DH tools and techniques, or equivalently what incentive humanists have to take up and to use new methods. This can be treated in some respects like the computational quest for the "killer application" – a need of the user group that can be filled, and by filling it, create an acceptance of that tool and the supporting methods/results. Digital Humanities needs a "killer application."

"Killer application" is a term borrowed from the discipline of computer science. In its strictest form, it refers to an application program so useful that users are willing to buy the hardware it runs on, just to have that program. One of the earliest examples of such an application was the spreadsheet, as typified by VisiCalc and Lotus 1-2-3. Having a spreadsheet made business decisionmaking so much easier (and more accurate and profitable) that businesses

were willing to buy the computers (Apple IIs or IBM PCs, respectively) just to run spreadsheets. Gamers by the thousands have bought Xbox gaming consoles just to run *Halo*. A killer application is one that will make you buy, not just the product itself, but also invest in the necessary infrastructure to make the product useful.

For digital humanities, this term should be interpreted in a somewhat broader sense. Any intellectual product — a computer program, an abstract tool a theory, an analytic framework — can and should be evaluated in terms of the “affordances” [Gibson, 2005, Ruecker and Devereux, 2004] it creates. In this framework, an “affordance” is simply “an opportunity for action” [Ruecker and Devereux, 2004]; spreadsheets, for instance, create opportunities to make business decisions quickly on the basis of incomplete or hypothesized data, while *Halo* creates the opportunity for playing a particular game. Ruecker provides a framework for comparing different tools in terms of their “affordance strength,” essentially the value offered by the affordances of a specific tool.

In this broader context, a “killer app” is any intellectual construct that creates sufficient affordance strength to justify the effort and cost of accepting, not just the construct itself, but the supporting intellectual infrastructure. It is a solution sufficiently interesting to, by itself, retrospectively justify looking the problem it solves — a Great Problem that can both empower and inspire.

Three properties appear to characterize such “killer apps”. First, the problem itself must be real, in the sense that other humanists (or the public at large) should be interested in the fruits of its solution. For example, the organizers of a recent NSF summit on “Digital Tools for the Humanities” identified several examples of the kinds of major shifts introduced by information technology in various areas. In their words,

When information technology was first applied [to inventory-based businesses], it was used to track merchandise automatically, rather than manually. At that time, the merchandise was stored in the same warehouses, shipped in the same way, depending upon the same relations among produces and retailers as before[...]. Today, a revolution has taken place. There is a whole new concept of just-in-time inventory delivery. Some companies have eliminated warehouses altogether, and the inventory can be found at any instant in the trucks, planes, trains, and ships delivering sufficient inventory to re-supply the consumer or vendor — just in time. The result of this is a new, tightly interdependent relationship between suppliers and consumers, greatly reduced capital investment in “idle” merchandise, and dramatically more responsive service to the final consumer.

A killer application in scholarship should be capable of effecting similar change in the way that practicing scholars do their work. Only if the problem is real can an application solving it be a killer. The *Clotel* edition described above appears to fail under this property precisely because only specialists in

Clotel (or in 19th-century or African-American literature) are likely to be interested in the results; a specialist in the *Canterbury Tales* will not find her work materially affected.

Second, the problem must get buy-in from the humanities computing community itself, in that humanities computing specialists will be motivated to do the actual work. The easiest and probably cheapest way to do this is for the process of solution itself to be interesting to the participating scholars. For example, the compiling of a detailed and subcategorized bibliography of all references to a given body of work would be of immense interest to most scholars; rather than having to pore through dozens of issues of thousands of journals, they could simply look up their field of interest. (This is, in fact, very close to the service that Thompson Scientific provides with the *Social Science Citation Index*, or that Penn State provides with *CiteSeer*.) The problem is that though the product is valuable, the process of compiling it is dull, dreary, and unrewarding. There is little room for creativity, insight, and personal expression in such a bibliography. Most scholars would not be willing to devote substantial effort — perhaps several years of full-time work — to a project with such minimal reward. (By contrast, the development of a process to automatically create such a bibliography could be interesting and creative work.) The process of solving interesting problems will almost automatically generate papers and publications, draw others into the process of solving it, and create opportunities for discussion and debate. We can again compare this to the publishing opportunities for a bibliography — is “my bibliography is now 50% complete” a publishable result?

Third, the problem itself must be such that even a partial solution or an incremental improvement will be useful and/or interesting. Any problem that meets the two criteria above is unlikely to submit to immediate solution (otherwise someone would probably already have solved it). Similarly, any such problem is likely to be sufficiently difficult that solving it fully would be a major undertaking, beyond the resources that any single individual or group could likely muster. On the other hand, being able to develop, deploy, and use a partial solution will help advance the field in many ways. The partial solution, by assumption, is itself useful. Beyond that, researchers and users have an incentive to develop and deploy improvements. Finally, the possibility of supporting and funding incremental improvements makes it more likely to get funding, and enhances the status of the field as a whole.

3.2 Some historical examples

To more fully understand this idea of a killer app, we should first consider the history of scholarly work, and imagine the life of a scholar c. 1950. He (probably) spends much of his life in the library, reading paper copies of journal articles and primary sources to which he (or his library) has access, taking detailed notes by hand on index cards, and laboriously writing drafts in longhand which he will revise before finally typing (or giving to a secretary to type). His new ideas are sent to conferences and journals, eventually to find their way into the libraries

of other scholars worldwide over a period of months or years. Collaboration outside of his university is nearly unheard-of, in part because the process of exchanging documents is so difficult.

Compare that with the modern scholar, who can use a photocopier or scanner to copy documents of interest and write annotations directly on those copies. She can use a word processor (possibly on a portable computer) both to take research notes and to extend those notes into articles; she has no need to write complete drafts, can easily rearrange or incorporate large blocks of text, and can take advantage of the computer to handle “routine” tasks such as spelling correction, footnote numbering, bibliography formatting, and even pagination. She can directly incorporate the journal’s formatting requirements into her work (so that the publisher can legitimately ask for “camera-ready” manuscripts as a final draft), eliminating or reducing the need both for typists and typesetters. She can access documents from the comfort of her own office or study via an electronic network, and use advanced search technology to find and study documents that her library does not itself hold. She can similarly distribute her own documents through that same network and make them available to be found by other researchers. Her entire work-cycle has been significantly changed (for the better, one hopes) by the availability of these computation resources.

We thus have several historical candidates for what we are calling “killer apps”: xerographic reproduction and scanning, portable computing (both arguably hardware instead of software), word processing and desktop publishing (including subsystems such as bibliographic packages and spelling checkers), networked communication such as Email and the Web, and search technology such as Google. These have all clearly solved significant issues in the way humanities research is generally performed (i.e. met the first criterion). In Ruecker’s terms, they have all created ‘affordances’ of the sort that no modern scholar would choose to forego. The amount of research work — journals, papers, patents, presentations, and books — devoted to these topics suggests that researchers themselves are interested in solving the problems and improving the technologies, in many cases incrementally (e.g., “how can a search engine be tuned to find documents written in Thai?”).

Of course, for many of these applications, the window of opportunity has closed, or at least narrowed. A group of academics are unlikely to be able to have the resources to build/deploy a competing product to Microsoft and/or Google. On the other hand, the very fact that humanities scholars are something of a niche market may open the door to incremental killer apps based upon (or built as extensions to) mainstream software, applications focused specifically on the needs of practicing scholars. The next section presents a partial list of some candidates that may yield killer applications in the foreseeable future. Some of these candidates are taken from my own work, some from the writings of others.

3.3 Potential current killer apps

3.3.1 Back of the Book Index Generation

Almost every nonfiction book author has been faced with the problem of indexing. For many, this will be among the most tedious, most difficult, and least rewarding parts of writing the book. The alternative is to hire a professional indexer (perhaps a member of an organization such as the American Society of Indexers, www.asindexing.org) and pay a substantial fee, which simply shifts the uncomfortable burden to someone else, but does not substantially reduce it.

A good index provides much more than the mere ability to find information in a text. The Clive Pyne book indexing company² lists some aspects of what a good index provides. According to them, “a good index:

- provides immediate access to the important terms, concepts and names scattered throughout the book, quickly and efficiently;
- discriminates between useful information on a subject, and a passing mention;
- has headings which are concise, accurate and unambiguous reflecting the contents and terminology used in the text;
- has sufficient cross-references to connect related terms;
- anticipates how readers will search for information;
- reveals the inter-relationships of topics, concepts and names so that the reader need not read the whole index to find what they are looking for;
- provides terminology which might not be used in the text, but is the reference point that the reader will use for searching through the index;
- can make the difference between a book and a very good book”

A traditional back-of-the-book (BotB) index is a substantial intellectual accomplishment in its own right. In many ways, it is an encapsulated and stylized summary of the intellectual structure of the book itself. “A good index is an objective guide to the text, a link between the author’s ideas and the reader. It should be a road map that leads readers to every relevant idea without frustrating detours and dead ends.”³ And it is specifically not just a concordance or a list of terms appearing in the document.

It is thus surprising that a tedious task of such importance has not yet been computerized. This is especially surprising given the effectiveness of search engines such as Google at “indexing” the unimaginably large volume of information on the Web. However, the tasks are subtly different; a Google search is not expected to show knowledge of the structure of the documents or the relationships

²http://www.cpynebookindexing.com/what_makes_a_good_index.htm, accessed 5/31/2006

³Kim Smith, <http://www.smithindexing.com/whyprof.html>, accessed 5/31/2006.

among the search terms. As a simple example, a phrasal search on Google (May 31, 2006) for “a good index,” found, as expected, several articles on back of the book indexing. It also found several articles on financial indexing and index funds, and a scholarly paper on glycemic control as measured (“indexed”) by plasma glucose concentrations. A good text index would be expected to identify these three subcategories, to group references appropriately, and to offer them to the reader proactively as three separate subheadings. A good text index is not simply a search engine on paper, but an intellectual precis of the structure of the text.

This is therefore an obvious candidate for a killer application. Every humanities scholar needs such a tool. Indeed, since chemistry texts need indexing as badly as history texts do, scholars outside of the humanities also need it. Unfortunately, not only does it not (yet) exist, but it isn’t even clear at this writing what properties such a tool would have. Thus there is room for fundamental research into the attributes of indices *as a genre of text*, as well as into the fundamental processes of compiling and evaluating indices and their expression in terms of algorithms and computation.

I have presented elsewhere [Juola, 2005, Lukon and Juola, 2006] a possible framework to build a tool for the automatic generation of such indices. Without going into technical detail, the framework identifies several important (and interesting) cognitive/intellectual tasks that can be independently solved in an incremental fashion. Furthermore, this entire problem clearly admits of an incremental solution, because a less-than-perfect index, while clearly improvable, is still better than no index at all, and any time saved by automating the more tedious parts of indexing will still be a net gain to the indexer. Thus all three components of the definition of killer app given above are present, suggesting that the development of such an indexing tool would be beneficial both inside and outside the digital humanities community.

3.3.2 Annotation tools

As discussed above, one barrier to the use of E-texts and digital editions is the current practices of scholars with regard to annotation. Even when documents are available electronically, many researchers (myself include) will often choose to print them and study them on paper. Paper permits one not only to mark text up and to make changes, but also to make free-form annotations in the margins, to attach PostIt notes in a rainbow of colors, and to share commentary with a group of colleagues. Annotation is a crucial step in recording a reader’s encounter with a text, in developing an interpretation, and in sharing that interpretation with others.

The recent IATH Summit on Digital Tools for the Humanities [IATH Summit, 2006] identified this process of annotation and interpretation as a key process underlying humanistic scholarship, and specifically discussed the possible development of a tool for digital annotation, a “highlighter’s tool,” that would provide the same capacities of annotation of digital documents, including multimedia documents, that print provides. The flexibility of digital media means, in fact, that

one should be able to go beyond the capacities of print — for example, instead of doodling a simple drawing in the margin of a paper, one might be able to “doodle” a Flash animation or a .wav sound file.

Discussants identified at least nine separate research projects and communities that would benefit from such a tool. Examples include “a scholar currently writing a book on Anglo-American relations, who is studying propaganda films produced by the US and UK governments and needs to compare these with text documents from on-line archives, coordinate different film clips, etc.”; “an add-on tool for readers (or reviewers) of journal articles,” especially of electronic journal systems (The current system of identifying comments by page and line number, for example, is cumbersome for both reviewers and authors.); and “an endangered language documentation project that deals with language variation and language contact,” where multilingual, multialphabet, and multimedia resources must be coordinated among a broad base of scholars. Such a tool has the potential to change the annotation process as much as the word processor has changed the writing and publication process.

Can community buy-in be achieved? There is certainly room for research and for incremental improvements, both in defining the standards and capacities of the annotations and in expanding those capacities to meet new requirements as they evolve. For example, early versions of such a project would probably not be capable handling all forms of multimedia data; a research-quality prototype might simply handle PDF files and sound, but not video. It’s not clear that the community support is available for building early, simple versions – although “a straw poll showed that half of [the discussants] wanted to build this kind of tool, and all wanted to use it.” [IATH Summit, 2006], responding to a straw poll is one thing and devoting time and resources is another altogether; it is not clear that any software development on this project has yet happened. However, given the long-term potential uses and research outcomes from this kind of project, it clearly has the potential to be a killer application.

3.3.3 Resource exploration

Another issue raised at the summit is that of resource discovery and exploration. The huge amount of information on the Web is, of course, a tremendous resource for all of scholarship, and companies such as Google (especially with new projects such as Google Images and Google Scholar) are excellent at finding and providing access. On the other hand, “such commercial tools are shaped and defined by the dictates of the commercial market, rather than the more complex needs of scholars.” [IATH Summit, 2006] This raises issues about access to more complex data, such as textual markup, metadata, and data hidden behind gateways and search interfaces. Even where such data is available, it is rarely compatible from one database to another, and it’s hard to pose questions to take advantage of the markup.

In the words of the summit report,

What kinds of tools would foster the discovery and exploration

of digital resources in the humanities? More specifically, how can we easily locate documents (in multiple formats and multiple media), find specific information and patterns in across [sic] large numbers of scholarly disciplines and social networks? These tasks are made more difficult by the current state of resources and tools in the humanities. For example, many materials are not freely available to be crawled through or discovered because they are in databases that are not indexed by conventional search engines or because they are behind subscription-based gates. In addition, the most commonly used interfaces for search and discovery are difficult to build upon. And, the current pattern of saving search results (e.g., bookmarks) and annotations (e.g., local databases such as EndNote) on local hard drives inhibits a shared scholarly infrastructure of exploration, discovery, and collaboration.

Again, this has the potential to effect significant change in the day-to-day working life of a scholar, by making collaborative exploration and discovery much more practical and rewarding, possibly changing the culture by creating a new “scholarly gift economy in which no one is a spectator and everyone can readily share the fruits of their discovery efforts.” “Research in the sciences has long recognized team efforts. . . . A similar emphasis on collaborative research and writing has not yet made its way into the thinking of humanists.”

But, of course, what kind of discovery tools would be needed? What kind of search questions should be supported? How can existing resources such as lexicons and ontologies be incorporated into the framework? How can it take advantage of (instead of competing with) existing commercial search utilities? These questions illustrate many of the possible research avenues that could be explored in the development of such an application. Jockers’ idea of “macro lit-o-nomics (macro-economics for literature)” [Jockers, 2005] is one approach that has been suggested to developing useful analysis from large datasets; Ruecker and Devereaux [Ruecker and Devereux, 2004] and their “Just-in-Time” text analysis is another. In both projects, the researchers showed that interesting conclusions could be drawn by analyzing the large-scale results of automatically-discovered resources and looking at macro-scale patterns of language and thought.

3.3.4 Automatic essay grading

The image of a bleary-eyed teacher, bent over a collection of essays at far past her bedtime is a traditional one. Writing is a traditional and important part of the educational one, but most instructors find the grading of essays to be time-consuming, tedious, and unrewarding. This applies regardless of the subject; essays on Shakespeare are not significantly more fun to grade than essays on the history of colonialism. The essay grading problem is one reason that multiple choice tests are so popular in large classes. We thus have another potential “killer app,” an application to handle the chore of grading essays without interfering with the educational process.

Several approaches to automatic essay grading have been tried, with reasonable but not overwhelming success. At a low enough level, essay grading can be done successfully just by looking at aspects of spelling, grammar, and punctuation, or at stylistic continuity [Page, 1994]. Foltz [Foltz et al., 1999] has also shown good results by comparing semantic coherence (as measured, via Latent Semantic Analysis, from word cooccurrences) with that of essays of known quality:

LSA's performance produced reliabilities within the range of their comparable inter-rater reliabilities and within the generally accepted guidelines for minimum reliability coefficients. For example, in a set of 188 essays written on the functioning of the human heart, the average correlation between two graders was 0.83, while the correlation of LSA's scores with the graders was 0.80. . . .

In a more recent study, the holistic method was used to grade two additional questions from the GMAT standardized test. The performance was compared against two trained ETS graders. For one question, a set of 695 opinion essays, the correlation between the two graders was 0.86, while LSA's correlation with the ETS grades was also 0.86. For the second question, a set of 668 analysis of argument essays, the correlation between the two graders was 0.87, while LSA's correlation to the ETS grades was 0.86. Thus, LSA was able to perform near the same reliability levels as the trained ETS graders.

Beyond simply reducing the workload of the teacher, this tool has many other uses. It can be used, for example, as a method of evaluating a teacher for consistency in grading, or for ensuring that several different graders for the same class use the same standards. More usefully, perhaps, it can be used as a teaching adjunct, by allowing students to submit rough drafts of their essays to the computer and re-write until they (and the computer) are satisfied. This will also encourage the introduction of writing into the curriculum in areas outside of traditional literature classes, and especially into areas where the faculty themselves may not be comfortable with the mechanics of teaching composition. Research into automatic essay grading is a active area among text categorization scholars and computer scientists for the reasons cited above. [Valenti et al., 2003]

From a philosophical point of view, though, it's not clear that this approach to essay grading should be acceptable. A general-purpose essay grader can do a good job of evaluating syntax and spelling, and even (presumably) grade "semantic coherence" by counting if an acceptable percentage of the words are close enough together in the abstract space of ideas. What such a grader cannot do is evaluate factual accuracy or provide discipline-specific information. Furthermore, the assumption that there is a single grade that can be assigned to an essay, irrespective of context and course focus, is questionable. Here is an area where a problem has already been identified, applications have been and continue to be developed, uptake by a larger community is more or less guaranteed,

but the input of humanities specialists is crucially needed to improve the service quality provided.

4 Discussion

The list of problems in the preceding section is not meant to be either exclusive or exhaustive, but merely to illustrate the sort of problems for which killer apps can be designed and deployed. Similarly, the role for humanities specialists to play will vary from project to project – in some cases, humanists will need to play an advisory role to keep a juggernaut from going out of control (as might be needed with the automatic grading), while in others, they will need to create and nurture a software project from scratch. The list, however, shares enough to illustrate both the underlying concept and its significance. In other words, we have an answer to the question “what?” — what do I mean by a “killer application,” what does it mean for the field of digital humanities, and, as I hope I have argued, what can we do to address the perennial problem of neglect by the mainstream.

An equally important question, of course, is “how?” Fortunately, there appears to be a window opening, a window of increased attention and available research opportunities in the digital humanities. The IATH summit cited above [IATH Summit, 2006] is one example, but there are many others. Recent conferences such as the first Text Analysis Developers Alliance (TADA), in Hamilton (2005), the Digital Tools Summit for Linguistics in East Lansing (2006), the E-MELD Workshops (various locations, 2000–6), the Cyberinfrastructure for Humanities, Arts, and Social Sciences workshop at UCSD (2006), and the recent establishment of the Working Group on Community Resources for Authorship Attribution (New Brunswick, NJ; 2006) illustrate that digital scholarship is being taken more seriously. The establishment of Ray Siemens in 2004 as the Canada Research Chair in Humanities Computing is another important milestone, marking perhaps the first recognition by a national government of the significance of Humanities Computing as an acknowledged discipline.

Perhaps most important in the long run is the availability of funding to support DH initiatives. Many of the workshops and conferences described above were partially funded by competitively awarded research grants from national agencies such as the National Science Foundation. The Canadian Foundation for Innovation has been another major source of funding for DH initiatives. But perhaps the most significant development is the new (2006) Digital Humanities Initiative at the (United States) National Endowment for the Humanities. From the website⁴:

NEH has launched a new digital humanities initiative aimed at supporting projects that utilize or study the impact of digital technology. Digital technologies offer humanists new methods of conducting research, conceptualizing relationships, and presenting

⁴<http://www.neh.gov/grants/digitalhumanities.html>, accessed 6/18/2006

scholarship. NEH is interested in fostering the growth of digital humanities and lending support to a wide variety of projects, including those that deploy digital technologies and methods to enhance our understanding of a topic or issue; those that study the impact of digital technology on the humanities—exploring the ways in which it changes how we read, write, think, and learn; and those that digitize important materials thereby increasing the public’s ability to search and access humanities information.

The list of potentially supported projects is large:

- apply for a digital humanities fellowship (coming soon!)
- create digital humanities tools for analyzing and manipulating humanities data (Reference Materials Grants, Research and Development Grants)
- develop standards and best practices for digital humanities (Research and Development Grants)
- create, search, and maintain digital archives (Reference Materials Grants)
- create a digital or online version of a scholarly edition (Scholarly Editions Grants)
- work with a colleague on a digital humanities project (Collaborative Research Grants)
- enhance my institution’s ability to use new technologies in research, education, preservation, and public programming in the humanities (Challenge Grant)
- study the history and impact of digital technology (Fellowships, Faculty Research Awards, Summer Stipends)
- develop digitized resources for teaching the humanities (Grants for Teaching and Learning Resources)

Most importantly, this represents an agency-wide initiative, and thus illustrates the changing relationship between the traditional humanities and digital scholarship at the very highest levels.

Of course, just as windows can open, they can close. To ensure continued access to this kind of support, the supported research needs to be successful. This paper has deliberately set the bar high for “success,” arguing that digital products can and should result in substantial uptake and effect significant changes in the way that, as NEH put it, “how we read, write, think, and learn.” The possible problems discussed earlier are an attempt to show that we can effect such changes. But the most important question, of course, is “should we?”

“Why?” Why should scholars in the digital humanities try to develop this software and make these changes? The first obvious answer is simply one of self-interest as a discipline. Solving high-profile problems is one way of attracting the attention of mainstream scholars and thereby getting professional advancement. Warwick [Warwick, 2004b] illustrates this in her analysis of the citations of computational methods, and the impact of a single high-profile example. Of all articles studied, the only ones that cited computation methods did so in the context of Don Foster’s controversial analysis of “A Funeral Elegy” to Shakespeare.

The Funeral Elegy controversy provides a case study of circumstances in which the use of computational techniques was noticed and adopted by mainstream scholars. The paper argues that a complex mixture of a canonical author (Shakespeare) and a star scholar (Foster) brought the issue to prominence. . . .

The Funeral Elegy debate shows that if the right tools for textual analysis are available, and the need for, and use of, them is explained, some mainstream scholars may adopt them. Despite the current emphasis on historical and cultural criticism, scholars will surely return in time to detailed analysis of the literary text. Therefore researchers who use computational methods must publish their results in literary journals as well as those for humanities computing specialists. We must also realize that the culture of academic disciplines is relatively slow to change, and must engage with those who use traditional methods. Only when all these factors are understood and are working in concert, may computational analysis techniques truly be more widely adopted.

Implicit in this, of course, is the need for scholars to find results that are publishable in mainstream literary journals as well as to do the work resulting in publication, the two main criteria of killer apps.

On a less selfish note, the development of killer applications will improve the overall state of scholarship as a whole, without regard to disciplinary boundaries. While change for its own sake may not necessarily be good, solutions to genuine problems usually are. Creating the index to a large document is not fun — it requires days or weeks of painstaking, detailed labor that few enjoy. The inability to find or access needed resources is not a good thing. By eliminating artificial or unnecessary restrictions on scholarly activity, scholars are freed to do what they really want to do — to read, to write, to analyze, to produce knowledge, and to distribute it.

Furthermore, the development of such tools will in and of itself generate knowledge, knowledge that can be used not only to generate and enhance new tools but to help understand and interpret the humanities more generally. Software developers must be long-term partners with the scholars they serve, but digital scholars must also be long-term partners, not only with the software developers, but with the rest of the discipline and its emerging needs. In many

case, the digital scholars are uniquely placed to identify and to describe the emerging needs of the discipline as a whole. With a foot in two camps, the digital scholars will be able to speak to the developers about what is needed, and to the traditional scholars about what is available as well as what is under development.

5 Conclusion

Predicting the future is always difficult, and predicting the effects of a newly-opened window is even more so. But recent developments suggest that digital humanities, as a field, may be at the threshold of new series of significant developments that can change the face of humanities scholarship and allow the “emerging discipline of humanities computing” finally to emerge.

For the past forty years, humanities computing has more or less languished in the background of traditional scholarship. Scholars lack incentive to participate (or even to learn about) the results of humanities computing. This paper argues that DH specialists are placed to create their own incentives by developing applications with sufficient scope to materially change the way humanities scholarship is done. I have suggested four possible examples of such applications, knowing well that many more are out there. I believe that by actively seeking out and solving such Great Problems – by developing such killer apps, scholarship in general and digital humanities in particular, will be well-served.

References

- [Foltz et al., 1999] Foltz, P. W., Laham, D., and Landauer, T. K. (1999). Automated essay scoring: Applications to educational technology. In *Proceedings of EdMedia '99*.
- [Gibson, 2005] Gibson, M. (2005). *Clotel*: An electronic scholarly edition. In *Proceedings of ACH/ALLC 2005*, Victoria, BC CA. University of Victoria.
- [IATH Summit, 2006] IATH Summit (2006). Summit on digital tools for the humanities : Report on summit accomplishments.
- [Jockers, 2005] Jockers, M. (2005). Xml aware tools — catools. In *Presentation at Text Analysis Developers Alliance*, McMaster University, Hamilton, ON.
- [Juola, 2005] Juola, P. (2005). Towards an automatic index generation tool. In *Proceedings of ACH/ALLC 2005*, Victoria, BC CA. University of Victoria.
- [Lukon and Juola, 2006] Lukon, S. and Juola, P. (2006). A context-sensitive computer-aided index generator. In *Proceedings of DH 2006*, Paris. Sorbonne.
- [Martin, 2005] Martin, S. (2005). Reaching out: What do scholars want from electronic resources? In *Proceedings of ACH/ALLC 2005*, Victoria, BC CA. University of Victoria.

- [Page, 1994] Page, E. B. (1994). Computer grading of student prose using modern concepts and software. *Journal of Experimental Education*, 62:127–142.
- [Ruecker and Devereux, 2004] Ruecker, S. and Devereux, Z. (2004). Scraping Google and Blogstreet for Just-in-Time text analysis. In *Presented at CaSTA-04, The Face of Text*, McMaster University, Hamilton, ON.
- [Siemens et al., 2004] Siemens, R., Toms, E., Sinclair, S., Rockwell, G., and Siemens, L. (2004). The humanities scholar in the twenty-first century: How research is done and what support is needed. In *Proceedings of ALLC/ACH 2004*, Gothenberg. U. Gothenberg.
- [Toms and O’Brien, 2006] Toms, E. G. and O’Brien, H. L. (2006). Understanding the information and communication technology needs of the e-humanist. *Journal of Documentation*, (accepted/forthcoming).
- [USNews, 2006] USNews (2006). U.S. News and World Report : America’s best graduate schools (social sciences and humanities).
- [Valenti et al., 2003] Valenti, S., Neri, F., and Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education*, 2:319–330.
- [Warwick, 2004a] Warwick, C. (2004a). No such thing as humanities computing? an analytical history of digital resource creation and computing in the humanities. In *Proceedings of ALLC/ACH 2004*, Gothenberg. U. Gothenberg.
- [Warwick, 2004b] Warwick, C. (2004b). Whose funeral? a case study of computational methods and reasons for their use or neglect in English studies. In *Presented at CaSTA-04, The Face of Text*, McMaster University, Hamilton, ON.