The field of Humanistic Informatics and its relation to the humanities

Espen Aarseth

Dept. of Humanistic Informatics in the Faculty of Arts,
University of Bergen

Abstract

This essay discusses how humanistic informatics (humanities computing) can be established as an autonomous field, rather than to go on as a supporting discipline in the service of traditional humanistic scolarship. This is important if the field is to go on expanding and its practitioners gain both self respect and the respect of others, something which today often is lacking, especially in the subfield of literary computing.

To do so, the author argues, the field must be able to focus on something which is not already dealt with by other fields, and which is not an obvious object for other fields. The answer lies in focusing on aesthetic and media issues of information technology (computer games, Internet culture, and hyper/cyber/media). This direction opens up a fresh territory of huge potential and importance for humanistic research.

Innehåll

Introduction

To discuss the theme put forward in my title, I must first address a more basic question: What is "Humanistic Informatics"? Unfortunately, I am not quite sure I know the answer to that. However, I shall try to construct an answer which will serve the purposes of this essay. After that, I will examine the relations between the Humanities and "Humanistic Informatics" (for which I find the acronym "HI" a bit too frivolous) as best I can, and from my limited perspective on the subject. Finally, I will suggest some areas and topics where the field (as I see it) is in a special position to develop a body of knowledge, that so far has been marginalized and neglected by other humanistic fields. To precipitate the conclusion, I think that the Humanities is best served by a field that is able to claim an autonomous research program, independent of the needs and opinions of other disciplines. And I don't think we have that quite yet. (If we had, there would be no need for this essay.)

I am trained in the study of certain aesthetic theories and objects (literary texts), but I have also worked for a number of years as a computing consultant for researchers in the humanities, so I speak from an interstitial position, between two different academic cultures or, if not cultures, then at least two very different ways of thinking. Whether this makes my position a priviledged one, akin to Edward Said's "unique double vision" of the migrant, the intellectual who can move between different worlds, or whether I am just a confused "Mr. In-between" whith no place to call my own, and no space within which to develop a consistent professional identity, I am not quite sure. In fact, when asked to respond to questions such as the present one, I am, especially after a few weeks of brooding, inclined towards the latter alternative.

The reason I bring up the question of identity is of course not to bore you with my personal or professional problems, but to illustrate what seems to be a very prominent dilemma for many researchers in Humanistic Informatics: That is, what kind of field are we in, anyway? Is it a field?

What is "Humanistic Informatics"?

The history of "Humanistic Informatics" is a problematic one, since it can be seen from many different perspectives, and not always be seen as the same thing. Sometimes, it is not seen at all. Only a few months ago, I was asked, as acting Chair of our very small Dept., by the Director of the Faculty, to justify the name "Humanistic Informatics" as a label for our new undergraduate course, because there had been objections from other faculties about the legitimacy of that name. And, from within the Humanities, especially from the traditional humanists of a certain "critical" bent, there are many who regard "Humanistic Informatics" simply as an oxymoron, a contradiction in terms. And there are those who regard us as a foreign ideological body, to which not all humanists, unfortunately, are immune.

But today, even if humanities computing (as it used to be called) is still invisible, it is also omnipresent. Everybody's doing it. Computers, just like telephones, are everywhere. And they are being used, in spite of popular mythology to the contrary. But this poses a major problem: If computing is done in every field, then why do we need a separate field? Can there be a separate field?

Some years ago I served on a central committee for reorganising the computing infrastructure at Bergen university. At one point, it was suggested that all Departments and research groups focused on information technology should be gathered together in a new, separate Faculty. This notion was of course quickly rejected, for the task of extracting computer research from the local research communities would not only be impractical, but also counter-productive, since new activities involving IT would develop locally even as the old activities were being centralized. Computing is simply too well integrated into all kinds of research to be isolated in one place.

So, one may well ask, should not the same conclusion be drawn for the Humanistic field(s) of computing? Do we need a separate field for what goes on everywhere anyway? We may need some kind of training centre for our students and research candidates, if only for the economical advantages of scale, but does that justify a specialized research dept.? The answer is, of course, no. Only the need for research can justify research. A computing section with technical support staff and programmers for special projects, perhaps a lecturer for introductory undergraduate courses, fine. But a autonomous department?

Looking at the humanities computing activities at a place like Bergen, we find a very rich diversity of computer-based research: Computational linguistics, historical informatics, corpus-based linguistics, computational art-history, classical philology, digital runology, machine translation, textual criticism by exploratory data analysis, computerized teaching methods, and much more. Most departments at our faculty can boast some sort of computer-based applied or basic research, many at a very sophisticated level, and in addition there are three or four centres also devoted to IT strategies. Most of these activities seem to be doing very well on their own, and seem to be welcomed, rather than frowned upon, by their mother departments. So the last thing a place like Bergen needs is a Department of humanistic computer research as well, or so it would seem.

Historically, the tradition of humanities computing began soon after the first computers were built. The first project was Father Roberto Busa's concordance of the texts of Thomas Aquinas, begun as early as 1949. A concordance is an index of every occurrence of a word in a body of texts, and immediately useful to anyone who needs to study longer texts closely. Indeed, concordances are much older humanistic tools than computers (they have been around since the middle ages), and are suitable to a number of different purposes, not only philological ones. The number of humanists who might need a concordance in their research probably far outnumber those who don't. But how many of them know how to build one?

But the promise associated with the new computer technologies gave rise to research far beyond simple indexing and digitalisation of texts. Coincidental with the rise of such futuristic and initially optimistic disciplines as Artificial intelligence and computerized simulations ("Virtual reality"), the computing humanists steadily grew in number,

organized themselves in international societies, established journals, and so on. In some fields (particularly linguistics) they were more or less quickly accepted and welcomed, in others their methods and perspectives were regarded with suspicion and ridicule, if not simply ignored.

It is easy to see how the lack of successful integration into existing disciplines might motivate the forming of a separate field of Humanistic Informatics, but this is of course not a very good justification for any field, and conjures up the image of a ghetto or a reservation. A far better reason for such a field is the computing humanists' methodological community, which cross the traditional disciplinary boundaries, and which can stimulate interdisciplinary research and the exchange of ideas even far beyond the Faculty boundaries. This would be reason enough for a research centre, where researchers could receive special training and eventually return to their home fields with new methods and ideas. But for a permanent, independent, tantamount humanistic field to exist, one must be able to establish that 1) its research is a worthwhile addition to existing fields, and that 2) it could not be better cared for within existing fields. Therefore, humanistic informatics, in order to exist as an independent field, must display a core research activity that does not naturally belong to the established fields.

It may of course still exist as a department, since the boundaries between departments seldom observe other principles of delimitation than political ones. Therefore, a Department of Humanistic Informatics may well resign itself to the role of an area study; in this case the study of how computers are used in humanistic research. If we look at the actual state of the World, we find very few Departments of Humanistic Informatics, but quite a few research centres for humanities computing. There are permanent departments in Denmark, Netherlands (called "alfa-informatica"), and Norway, but not, as far as I know, anywhere else.

A field based on the premise that it exists primarily to assist and "contribute to" other fields, will probably never reach a healthy, self-respecting identity as a platform for scientific or scholarly enterprise. Contributions to other fields should not be offered, they should be obvious. In that sense, I find the question "What can Humanistic informatics contribute to the Humanities?" a little offensive, as it implies that the field of "Humanistic Informatics" should make itself useful, like a newly freed slave, or else full citizenship might be withheld. No one dares ask, say, Media studies, how it "contributes to the Humanities" —whatever that means. The field of Humanistic Informatics can only justify itself through a unique research profile, and it should let the usefulness of that profile be decided by the rest of the Humanities. I will return to the question of what such a research profile may be in the final part of the essay.

The relationship between Humanistic Informatics and the Humanities

The notion that computing in general is irrelevant to humanistic research is of course a false one. Computers provide useful and elegant tools for doing what we have always been doing. We need them just like we needed paper and libraries in the past. The

hostility towards computers in the Humanities, when separated from the general technophobia of traditional humanists, is usually concerned with methodological issues, and, through those, with far-reaching questions about the nature of humanistic research.

The "nature of humanistic research" is in itself a very questionable phrase. Does humanistic research have a consistent nature, e.g. one that would distinguish it from the social or natural sciences? C. P. Snow, in his famous 1959 essay on the cultural conflict between the Humanities and Science, perceived a gap between the two traditions. But from my own, much more recent, experience, I would say that there seems to be a larger gap within the Humanities itself, one that has perhaps become visible only some time after Snow's essay, although divisions like the one between "lang." and "lit." are not new at all. This gap is especially obvious in those parts of the Humanities that were affected by the turns of post-structuralism and critical theory, but it is also a gap with much older roots in intellectual history than those recent developments. In this present situation one type of humanist may find much more common ground with scientists and mathematicians than with humanists belonging to another theoretical school. Insofar as this fault line follows the humanistic disciplines (broadly speaking: history, aesthetics, linguistics, philology, philosophy) then my own discipline, (literary) aesthetics, along with philosophy, most clearly belong to the anti-scientific side of the fence (yes, I know that this is a *very* simplified picture).

Here, instead of looking at all the humanistic disciplines and their relation to humanistic informatics one by one (a task for which I am not really qualified, even if there had been space to do it), or, alternatively, engage in some vague and overgeneralized speculation about a Humanities that does not exist except in banquet speeches, I will limit this part of my discussion by using the relationship between literary aesthetics and Humanistic Informatics as a representative example. Not to make things easier for myself by using the field I happen to know best, but because I find this particular relationship the most interesting and problematic, almost traumatic; much more so than the relationship between, say, history and historical informatics, or between linguistics and linguistic computing.

For at least three decades, statistical methods have been used in the study of literature. Using statistical analysis, the "stylistic features of texts, writers, periods and/or genres" (Potter 1991: 413) are identified. Despite huge and admirable efforts, this type of research has met with very little success, and almost no recognition at all from mainstream literary critics. Its reception can probably better be described as generally scornful. A clear exception is the sub-field of authorship attribution, where the personal stylistic fingerprint of an author is used to determine whether a particular writer is the author of a particular work. In such cases, the results have an obvious historical and text-critical value. But in general, "The statistical analysis of Literature" has been very poorly received as a new method of interpretation.

Some of you may have read Stanley Fish's 1973 essay, "What Is Stylistics And Why Are They Saying Such Terrible Things About It?" where he, in his usually thorough manner, dismisses the practice of computational stylistics as a viable alternative to what he calls "impressionistic criticism", that is, the tried and true method that most critics use. Fish holds that, because the statistical approach cannot be used to identify elements not already specified by the analysis, it will be used to find exactly what the

analyst was looking for.

It would be easy to join the polemic against computational stylistics, but also quite unnecessary, because the practitioners themselves are among the most candid and critical in their recognitions of the failings of their field. In a 25th anniversary issue of *Computers and the Humanities*, the leading journal for humanities computing, Rosanne Potter surveys the journal's articles on literary computing and concludes that "Too often we have counted because the computer can and run tests because statistical packages exist" (1991: 428). In a similar article in the same issue, Louis Milic admits that "measured against the expectations that we had twenty five years ago and more, I must say the net is disappointing" (1991: 399).

But this admirable self-criticism did not stop there. A few years later, a full double issue of *Computers and the Humanities* were devoted to this crisis in literary computing, with a very critical position paper by Mark Olsen from 1991 as a starting point. Olsen claims that the discipline of computer-aided literature have failed to have a significant impact on the field as a whole, and goes on to discuss various possible reasons for this failure, the most prominent of which seems to be that the theoretical basis for the computer critics was outdated; they simply did not address the current problems of literary theory, but were still examining texts from older, outdated perspectives such as new criticism or thematic analysis. Olsen suggests a "new direction" oriented towards (Barthesian) semiotics and post-structuralism, because, he claims these approaches contain textual models that lend themselves better to a computer-based approach than the more analytical theories.

The reactions to Olsen's proposal have been varied and interesting. Some of his colleagues, notably Paul Fortier, found no reason to exchange the old paradigm for the new ones suggested by Olsen. Later, in a paper given at the ALLC-ACH conference in Santa Barbara, 1995, Fortier suggested that literary computing is better off without any connections to current literary theory:

In a recent article Jonathan Culler declares a complete break between literary theory and methodologies for studying literary texts. Given the current state of much theoretical speculation, one can only applaud this distinction. It then becomes a question of generating new theoretical underpinnings for the study of literature particularly when computers are used.

In other words, if literary theory won't come to the computational critic, the computational critic must build a new theory of literature, suited to the methods of computer-assisted study. One can easily understand the frustrations of a scholar who is faced with theories (such as deconstruction) for which his methods must seem completely irrelevant. But this is clearly not the way to build a discipline with any hope of making an impact on the field of literary studies.

However, although the discipline of literary computing is without much consequence for the field of literature, the relationship between information technology and literature in general is far from non-existent. In fact, as Allen Renear recently noted (in *Computers and the Humanities*, no 29, 1995), the rise of new communication technologies has made the field of textual computing more relevant to humanistic research than ever before. Renear is of course thinking of technologies of writing and reading such as

hypertext and hypermedia, and their recent easy integration in the humanistic tradition, particularly in the study of literature. Where literary computing has failed, it seems that hypertext scholarship has succeeded beyond anybody's wildest dreams. Not only does hypertext promise a tool for critical annotation and the representation of intertextuality, as well as a useful method for representing complex editions of variorum texts, it also has become, for many, an incarnation of the post-structural concept of text. Where the computational critic perceives an unbridgable gap between himself and the deconstructors, the hyper-critic simply constructs a link from the hypertext to the French semiotics, and Derrida, Barthes et al. are caught in the Net. As I have argued in *Cybertext: Perspectives on Ergodic Literature* (Johns Hopkins UP 1997), the hypercritics' "French connection" is dangerously superficial, but it is still very revealing how successful this idea has become.

There is very little friction, and much productive co-operation, between Hypertext and the Humanities, and this gives ground for optimism. However, one cannot sometimes but wonder that things are going a bit too smoothly; it is, after all, the Humanists' task to be critical, especially where their own tools are concerned. What is needed now is something that the Literature professors in love with hypertext can't give us, and that is a critical perspective on technologies of communication.

A research profile for Humanistic Informatics

The Humanities has always been concerned with human expression, whether Literature, Drama, Visual Art, or cultural discourse in general. In fact, we have organized a large part of our departmental structure to follow the media genres rather slavishly. When the electronic mass media arrived, they gradually became worthy of our attention, and eventually got their own department, as the departmental logic dictates.

Today we are faced with a new addition; a new type of technology of expression has arrived. We may then ask the questions to determine whether we have cause to set up a new field:

Is it an important addition that entails significant new structures of communication?

—So it seems. Today 50-60 million people are connected to the Internet. It is estimated that by the current growth rate, the figure will be 200 million in three years time. Computer games threaten the cultural hegemony of movies, and have passed the movie industry in terms of annual income. And perhaps most importantly, the digital media entail a shift in the way we organize our stored experience, from narratives to the dynamic models of games and simulations.

Can the same phenomenon be studied sufficiently by an existing discipline?

—In my opinion, it cannot. The opaque nature of digital information technology, the programmed mechanisms beneath the sign surfaces, makes special knowledge of computing necessary for the study of these media. Criticism, as well as exploration, must be informed, or it will be worthless.

To study the effects and consequences of digital technology on our culture, and how we are shaping these technologies according to our cultural needs, we can now begin to see the contours of a separate, autonomous field, where the historical, aesthetic, cultural and discursive aspects of the digitalisation of our society may be examined. That way, the field of Humanistic Informatics may contribute to the goal of the Humanities, which is the advancement of the understanding of human patterns of expression. We cannot leave this new development to existing fields, because they will always privilege their traditional methods, which are based on their own empirical objects.

But what about the needs of the Humanities, in terms of better computer methods, more useful software, and project assistance? If the other fields need specific support, then they must allocate the resources from their own budgets, or from central funds. This is just common sense. But I don't think there is a real conflict of interest here, where Humanistic Informatics changes from a useful servant to an unpaying house guest. What a department of Humanistic Informatics can and should do is to focus its research on the processing and exchange of information, particularly as it is being conducted in the Humanities. Thus, the problems of digital document representation, the rhetoric of pedagogical software, the uses of hypertext and multimedia, the limits of formal representation of aesthetic objects, the cognitive and political aspects of dynamic models, and many other activities are obvious objects for study. In particular, text mark-up systems such as SGML, and the potential and limits of exploratory data analysis, can and should provide extremely interesting subjects for the field.

A final illustration is the global hypertext system known as the World Wide Web. This is rapidly becoming the largest source for the world's textual knowledge. Like any successful information systems, its growth is exponential: every fifty days the information awailable of the WWW doubles in size. If this were a library, we would have to build a new building, twice as spacious as the current buildings, every other month or so. Unless, of course, we could come up with an alternative way to store information. Which is exactly what we have done with the WWW, which may be described as part library, part encyclopaedia, part scrap heap, and actually very useful, if you only know how to use it.

In the past, humanists led the development of the new technologies of reading and writing. Today, it is only logical that this tradition should continue. We must also develop new means of accessing this information, and teach our students and colleagues how to use it. Perhaps even more important is the fact that the Internet is different from the old media, in that it allows its users (those that are given the privilege of writing as well as reading) to participate as individuals. This shifts the rules and rituals of academic discourse in subtle and yet poorly understood ways.

In terms of theory, the field must be eclectic, seek inspiration from a wide variety of sources, and synthesize these as needed. Here we must be trusted to come up with the theories and methods that will further the research of the field, and that these will not be detrimental to the spirit and goals of the Humanities.

In short, the rapidly changing role of technology in the Humanities, as in society in general, is a very fascinating and important topic, which can and should be studied from a position within, but also only from a position of equality and autonomy.

The author

Espen Aarseth is Associate Professor in the Department of Humanistic Informatics in the Faculty of Arts, University of Bergen. He teaches undergraduate courses in the theory and design of digital media and the uses of IT in the Humanities, focusing on aesthetic and critical issues of hypermedia, computer games and cyberculture in general. He advises several doctoral projects on these topics, and is Director of the CALLMOO project . He has written numerous articles for Norwegian and international books and publications, and his doctoral dissertation, Cybertext: Perspectives on Ergodic Literature is published by the John Hopkins University Press. For more information please visit his Web site.

Åter till Human IT 4/1997