Controversies around the digital humanities: an agenda

Thaller, Manfred

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Controversies around the Digital Humanities: An Agenda

Manfred Thaller [*]

Abstract: *»Kontroversen um die Digitalen Geisteswissenschaften: Ein Arbeitsplan«*. Observations on the current stage of the Digital Humanities and their environment identify four dangers: (1) The focus on infrastructures for the Digital Humanities may obscure that research ultimately is driven by analytical methods and tools, not just by the provision of data or publishing tools. (2) Information technology can support the Humanities in many forms and national traditions. That textual analysis is much discussed right now, should not hide the view of a broader disciplinary field. (3) The mobile revolution looming may once again lead to a repetition of highly destructive processes observed at the PC and the internet revolutions. (4) The Digital Humanities may have to take a much stronger part in the development, not only the reception, of technology. – A series of concrete and controversial questions, which allow the discussion of some of these trends, is derived.

Keywords: Digital Humanities, research infrastructure, Digital Libraries, Digital Preservation.

1. Background and Motivation for a Discussion of the Digital Humanities

"Computing in the Humanities" has been a field of study, which originally derived much from two different roots: On the one hand the strong belief, that there must be easier ways to handle a couple of hundred thousand small details than write them on 5" x 8" cards; on the other an intellectual climate, which assumed that different methodological and intellectual traditions of the Humanities and hard sciences could and/or should converge, of which C.P. Snow's famous lectures have rather been a symptom, than the cause. The field has been influenced by rather different developments. *Intellectual or methodological fads*, as e.g. the popularity of interdisciplinary reception of the methodology of the (empirical) Social Sciences during the sixties and seventies. *Technological innovation*: When new key technologies arise, as the PC or the WWW, during the first few years there seems always to be a widespread conviction, that with that technology a totally new methodological platform for the

[*] Address all communications to: Manfred Thaller, Historisch-Kulturwissenschaftliche Informationsverarbeitung, Albertus-Magnus-Platz, 50923 Cologne; Germany; e-mail: manfred.thaller@uni-koeln.de.

interdisciplinary cooperation between the Humanities and information technology and its conceptual underpinnings has been created. And, of course, *funding possibilities*: It seems to be doubtful whether the term "Digital Humanities" would exist, if the huge funding programs for Digital Libraries at the turn of the millennium would never have arisen; it seems to be almost beyond doubt, that nobody would have thought of "eHumanities", if the promise of funding for eScience would not have been there.

In 1962 at the castle Wartenstein in Austria a group of scholars met to discuss *The Use of Computers in Anthropology*[1], presumably the first attempt to clarify a methodological position for the interdisciplinary world between the Humanities and Computer Science. Fifty years later, there is consensus that there is a shadowy subject between these disciplinary worlds, to describe it precisely seems only marginally easier than in 1962; or maybe harder, as the number of possibilities has exploded.

The University at Cologne has been the first German University to create a professorship for applied Computer Science in the Humanities, outside Computational Linguistics. It has recently strengthened its commitment to the field by founding the Cologne Centre for the eHumanities as a framework for the cooperation between interdisciplinary projects in the realm indicated by its name. Being strong in the field, the university sees it as part of its mission, to contribute to the intellectual consolidation of Humanities Computing/Computing in the Humanities/Humanities' Computer Science/Digital Humanities/eHumanities.

To do so, it celebrated the fiftieth anniversary of the methodological discussions on "the field" by a high-quality workshop, which hopefully shall initiate a series.

2. How to Discuss the Digital Humanities?

The Digital Humanities have been with us for a long time. Nevertheless, they are less well established as one would assume after half a century. People active in the interdisciplinary field between the Humanities and Computer Science have for a long time tended to feel like underdogs. The experience that an interdisciplinary field which arguably started in '49, but most definitely with the conference wave of the sixties, has frequently still to explain that it exists as a field of research in 2012, has left its traces. Therefore Digital Humanities conferences are frequently concerned to project the most harmonious picture possible, and smaller workshops frequently bring only people together who have been invited because they are known to agree on many terms beforehand.

[1] D. Hymes (ed.): The Use of Computers in Anthropology. Studies in General Anthropology II. The Hague: Mouton & Co., 1965.

And all to frequently "methodological" statements about the Digital Humanities seem primarily to be addressed to funding bodies, hoping to convince them of a specific type of project.

We think that this emphasis on presenting a harmonious picture to the outside world deviates in an important way from what research and scholarship should be: Productive, exciting and challenging discussions require a difference in opinion about a topic, which can be discussed. Research is not driven by harmony, but by controversial discussion. To support this, we have brought together highly visible researchers which could be expected to hold conflicting opinions on central issues. The following collection of papers documents these differences in opinion. To clarify for the reader, what the underlying controversies have been assumed to be, we start describing these controversies as we see them. We try to do so in a neutral and impartial way. The simple act of diagnosing a controversy is unavoidably an act of partiality already, however: Inviting to discuss the question whether the Earth is round or flat implies that it *could* be both. In cases where the number of followers of one possible position is much larger, than the number of followers of the other one, the majority tends to find it irritating, that one considers the possibility that it *might* be wrong. We hope we have avoided considering Flat Earth models of the Digital Humanities – but we also have to respectfully note, that there has been a stage in history, where the *majority* considered the world to be flat. Being a majority opinion did not make the idea right.

3. Observations

The individual controversies that have been selected for this workshop try to pinpoint individual questions, which are sufficiently precise, that concrete answers are possible. I would like however, to start with four personal observations on the development of the Digital Humanities, as I experienced it. This may make the reasons for the selection of *these* controversies selected more transparent.

3.1. A Loss of Focus?

My personal involvement with what today is called the Digital Humanities started in the autumn of 1976 when an interest in the application of Social Science methods to historical questions lead a young PhD in History into a two-year post-doc training in empirical sociology at Vienna's Institute of Advanced Studies. From there the idea, that software supporting historical methodology was needed which operated on the same level of complexity as SPSS did for empirical sociology, lead into a bundle of research projects first and then for roughly twenty years to the Max-Planck-Institute for History at Göttingen. Already at that time it was clear to everybody who watched the field

closely, that two conflicting goals existed in the application of computers to historical, or any other Humanities discipline, research: Easing the drudgery of routine tasks on the one hand and trying to work towards a "methodologically better" type of history at the other. Both goals could blur into each other, when types of source material could be used with the help of databases and statistical techniques, which otherwise eluded historical interpretation – like the extensive series of administrative records.

It was rather clear, however, that historical *analysis* improved by methodologically designed tools was the *goal*, and what had to be done to achieve that, like preparing analog sources digitally in such a way, that afterwards they would be accessible for the most diverse type of analysis possible, was (only) the *way*. That this preparation of digital material was so labor intensive, that at the end of the day many projects funded for short periods of time discovered at the end of that funding period, that no more time was left for analysis, seemed to be the most vexing and occasionally almost tragic problem of the field.

Having worked in that context with image processing since the late eighties, in 1996 we started a project to systematically digitize as images a collection of 80.000 pages of administrative material of the city archive Duderstadt.[2] Such a platform of archival material would be the ideal starting point, we assumed, of integrating analytical tools on several levels into the digital collection itself, working towards a solution of the vexing problem mentioned above, by separating the preparation and the analysis of the material. To propagate that idea we organized a series of three annual workshops, where we assumed, that reactions to the image processing tools presented plus the analytical and editorial possibilities described would allow us to optimize the platform for historical research. In hindsight the idea to let the feedback direct the later stages of the project may have been a mistake: While the interest in analytical possibilities was moderate, the interest in the proven possibility of having tens of thousands of pages of archival documents instantly available in the early WWW was overwhelming.

Still personally shaped by twenty years of watching how the drudgery of preparing Humanities' data for analysis was so exhausting, that not enough stamina remained for analysis, these developments moved my own research into the creation of digital libraries for cultural heritage and later on to the question how these could be kept stable over seriously long periods.

What became truly and increasingly irritating, particularly when out of the Digital Library movement notions of "Humanities research infrastructures", "Virtual Research Environments", "eHumanities" have been derived, is a simple observation, however. As I am not aware of any generally accepted definition of eScience, let me start with my own:

---

[2] S. Aumann et al.: "From Digital Archive to Digital Edition", in: *Historical Social Research* 24.1 (1999), p. 101-44.

> eScience describes the concept of performing scientific research in a distributed digital working environment, which supports equally well: (1) access to the information needed to tackle a research question, (2) the analysis of that information by tools reflecting the methodological requirements of the specific discipline and research problem and (3) the publication of the new information gained by the analytical process.

Translating that into a definition of eHumanities is clerically simple.

> eHumanities describes the concept of performing Humanities research in a distributed digital working environment, which supports equally well: (1) access to the information needed to tackle a research question, (2) the analysis of that information by tools reflecting the methodological requirements of the specific discipline and research problem and (3) the publication of the new information gained by the analytical process.

Looking at the reality of digital infrastructures for the Humanities I find it, however, extremely, and increasingly so, difficult to recognize many precautions "for the analysis of that information by tools reflecting the methodological requirements of the specific discipline and research problem". Or more bluntly: I can in no way recognize, that the abundance of digital material made available during the last decade has been augmented by a similar increase in the ambitiousness or power of the analytical tools applied to them.

It is extremely welcome, that librarians nowadays take an active interest in providing access to digital information. But does the way, in which this digital information is handled, not have to come out of an understanding of the analytical requirements, which can only be derived from an understanding what on the level of analysis – not skills – is defined by and defines the Digital Humanities? An understanding, that has to come out of the Humanities themselves, not the libraries – nor any other type of repository.

Have the Digital Humanities lost their focus on striving for "better" research?

3.2. A Narrowing of Scope?

Another observation. In the opening lecture of my introductory class of the BA for *Humanities Computer Science* at Cologne I define the scope of the Digital Humanities as follows:

The field of Digital Humanities can be divided into four broad subfields or paradigms.

a. There exists a subfield oriented towards the analysis of text as text.
   At least three directions of research are recognizable in this subfield. Within *literary studies* questions can be answered by computational tools: These can be as trivial as finding indications for the specific vocabulary of an author or a group of texts, or formally more ambitious as the computational definition of the style of an author or "school". *Editorial philology* focuses

on the reconstruction of "proper texts" which may have become lost during transmission over time, the identification of different layers within a text originating from a creative process and the presentation of the results. Digital support of it concentrates on forms of presentation not possible in the printed medium. *Computational linguistics* tries either to model formally the rules according to which language is used and utterances are constructed or it focuses on the empirical analysis of linguistic phenomena within *corpus linguistics*. As it is the only subfield in the Digital Humanities, which has acquired status as a well defined separate discipline, it is doubtful whether it should be seen as part of the Digital Humanities.

b. Another subfield focuses on chunks of information extracted from many sources: Texts, descriptions of images, spatial relationships. It is not concerned with the texts as texts or the images as images, but uses them to derive such chunks of information – sloppily called "facts" frequently – for further analysis between these extracted "facts". These approaches are most frequently used in disciplines which are related to the study of societal phenomena – like history, anthropology – or where tangible objects are examined – archaeology, history of art.

   The tools most frequently used today for these approaches are *data bases*, which can be used within a conceptual framework which can be very loose or rather rigid. The more rigid such a conceptual framework becomes, the more important techniques of statistical analysis become, which can lead to approaches described as *quantitative studies* of the Humanities.

   Within these approaches towards the analysis of "facts" two specialized groups can be identified: *Geographical Information Systems* connect data bases, where at least some components represent statistical information, with a spatial representation which usually leads to the visualization of phenomena in the form of maps. *Simulation* studies, more important in the eighties of the last century than right now, try to compare the result of developments of a segment of society or another system that can be modeled formally predicted by an algorithmic simulation with the outcomes that can be observed in reality.

c. There is, furthermore, a large subfield of the Digital Humanities, which deals with non-textual resources. This covers on the one hand the administration of large collections of images in fields like archaeology or the history of art, on the other it involves the usage of three-dimensional models of artifacts in these – and other – "visual" disciplines. Visualization, however, extends more and more also into the other subfields, as there exist many results of literary studies, linguistics, data base queries or statistical analysis, which are easier to understand by a graphic representation than by a tabular listing.

   From these approaches there are very close connections to the development of computer games. On the one hand as "serious games" which are used to

allow the playful acquisition of knowledge in various Humanities disciplines; on the other on the level of tools, as for many applications which require 3D models the quality which can be reached by game related software is quite sufficient.

This subfield comprises also the *cultural heritage* domain, where in the so called *memory institutions* (libraries, archives, museums) huge collections of material, which is available in the form of images exist, which require for their handling a mixture of technical knowledge with a Humanities background.

d. Finally there exists a subfield with two branches, which we call Humanities Computer *Science*.

An *algorithmic* view of Humanities Computer Science assumes that there exist some fundamental differences in the kind of information that is handled by the Humanities disciplines from that which is used in engineering, business or other knowledge domains. Differences which are sufficiently fundamental, that they require an adaptation of the basic building blocks of software systems. While this is the strongest and most fundamental claim of this view, in reality it is usually expressed quite practically: If researchers' interests are primarily in one of the first three major subfields, they will usually try to apply existing software to solve a question which is justified by the requirements of a specific discipline in the Humanities. Humanities Computer Science holds the knowledge and skills to create solutions of problems, where existing software is insufficient or simply not there.

A related, *epistemological*, view explores the question, how the study of the Humanities is changed by methods which become available only, if they can be supported by computational tools. What, on an abstract conceptual level, makes Humanities in a world where digital information and tools for its handling are available different from previous stages of the development of the Humanities?

I apologize for quoting so extensively from an introductory lecture of mine. I feel justified in doing so, as we are all aware, that our mutual definitions of the Digital Humanities and their scope are by no means clear. Using the above definition as my personal reference, I am concerned, that it seems to be unusual in being so broad. We pay a price for this narrowness of many definitions on a very practical level, as it seems to be next to impossible to transfer solutions from one project defining itself within a narrow context to another one within a context defined equally narrowly. As an example: In many projects of the Humanities, which are spread over all the subfields, views and paradigms referred to above, people encounter again and again the problem that time, as a data type supported by current software, does not cover the kind of temporal references we find in sources of all types, be it because the temporal notations refer to intervals explicitly – "from March 18th 1723 until August 23rd 1724" – or implicitly – "2nd quarter of the 19th century". I never collected the refer-

ences systematically, but I am sure that various projects have been proposing software solutions to that problem in at least two hundred conference papers and other contributions, ranging from highly simplistic uses of regular data bases to quite sophisticated designs of fuzzy data types encapsulated in component libraries, presented by computer scientists. That we can be absolutely sure that right at this moment somewhere a Digital Humanities project exists which strives heroically to solve that problem for the two hundred and first time is in my opinion directly traceable to an understanding of the field, where the community of reference is extremely small, so the generalizability of solutions is not really of great concern.

What has been exemplified by a small detail could also be shown by much broader concerns. Is a controlled vocabulary the only way to sensible recall and precision, or is it an inherent sin against a fundamental property of Humanities research, where categories are not fixed, but their change and development is the very substance of the research in the field?

Not every historian has to be knowledgeable about the history of Mr. Madison's war; not every chemist has to embrace the study of Fluoropolymers. A historian who denies that the events of 1812 are a topic of the larger field of history, denies to be a historian. A chemist, who denies, that Teflon is a topic of chemistry, isn't one. A Digital Humanist who denies, that the broader definitions define a frame of reference for his or her own work…

It is extremely welcome that by the increasing number of philology related projects under the Digital Humanities label the reference community for that angle of research is increasing. But does the way, in which this results in a concentration on rather plain and basic questions of handling texts, not lead to a serious loss of vision what computational methods could do for the Humanities as a whole?

Have the recent implicit definitions of "Digital Humanities" as a field of handling texts narrowed the scope of methodological interest so much, that the wider community suffers?

3.3. Another Iteration?

Another observation, on the relationship between technological development and stages in the development of the Digital Humanities.

I would propose to differentiate between the following main periods of the Digital Humanities, pointing to some characteristics of each of them.

a. 1949 - ca. 1970 in my periodization is the nascent period of computer usage in the Humanities. Being nascent does not mean, that it was not taken serious by its followers: Le Roy Ladurie's famous ... *dans ce demain au*

moins, l'historien de demain sera programmeur ou il ne sera plus[3] has been written in 1968. A statement which also describes the crucial technical terms of reference of that period: Applying computers meant, that programs in higher programming languages had to be written in each individual case, which restricted interdisciplinary projects mostly to rather large, separately funded ones – or to cases, where individuals immersed themselves very deeply into the software engineering part of the task.

b. ca. 1970 - ca. 1985 changed this radically by the advent of program packages. Still having to rely on their university's computing centre this meant, that Humanists who spent relatively little time to learn the command language of a program package, could realize a project within a much smaller project, or indeed, relatively easily, without support staff at all. In hindsight one wonders, whether this period also being the heyday of quantification, may not be more related to the fact, that this methodology supported by packages like SPSS was the only one, where such packages were so easily available, than any inherent interest in quantification.
c. ca. 1985 - ca. 1997 saw the effect of the *PC revolution*. The major change was obviously the possibility to do all of the work at your own desk. Another major change – if we use the broad definition of Digital Humanities proposed above – has been, that quantification lost much of its previous importance. Leaving word processing aside, the technical index fossil of the period is certainly the database, not necessarily easier to handle than a statistical package, but allowing the user to ignore the formalization of the problem. And this was certainly the first period, when computer supported projects in the Humanities became predominantly one person projects.
d. The *Internet revolution* formed the next period than, ca. 1997 - ca. 2010 (?). As this is, what we still experience daily, I'll be brief: It seems to be significant to me, that during this period the assumption what a "typical use" of information technology was changed slightly – towards the usage of computers as presentation medium and for the access to huge corpora of material, leaving analysis often as a human's task after information technology collected the items to be analyzed.

As in the previous observation, I apologize on spending so much space on the seemingly obvious. I wanted to specify these periods, as at least the 2nd, 3rd and 4th in my opinion show a common structure, which to the best of my knowledge has not been examined in depth so far. Simplifying things very much, we can observe:

Stage 1 (ca. the first 5 years): A new technology arises. It is accessible with much less technical knowledge. On the other hand, its capacity is insufficient to

[3] Originally in *Le Nouvel Observateur*, 8 mai 1968; later reprinted in: E. Le Roy Ladurie, *Le territoire de l'historien*, Paris 1973, pp. 11-14. Here with the title: *L'historien et l'ordinateur*.

handle the problems the established community of Digital Humanities is interested in. The old specialists therefore belittle the new technology or ignore it; the young enthusiasts consider the experience of the established professionals of the previous stage irrelevant as they do so.

As things are so incredibly easy now, the number of low level university courses explodes. (My personal all time favorite is *Informatik für Germanisten: WordStar 2000*. I.e.: Computer Science for German Studies: WordStar 2000, ca. 1987.) The methodical level of the discussion plummets vertically. (My personal all time favorite is a presentation at a historian's conference, where a young hopeful explained that PCs were much better suited methodologically to history, than main frames, as word-processing documents on dot-matrix printers were much more esthetically pleasing than statistical tables on high-speed printers, 1984.)

As the old professional organizations and conferences do not allow sufficient room for the new technology new organizations and conference series are established. (It may be doubtful to point to the start of the *Digital Resources for the Humanities* (1996) here; it very soon focused much more on the Internet, however, than the established series.)

Stage 2 (ca. the second 5 years): The number of "computer using" Humanities' scholars has been rising by an order of magnitude, compared to the situation before stage 1. Nevertheless, the number of well defined projects drops considerably against stage 1. As it turns out, what promised to be a very easily applicable technology, when done experimentally with small amounts of data, needs serious planning, when applied to significant and rewarding problems. It also turns out, that while the new technology *is* very simple to use, there are a couple of peculiarities of Humanities' data, which make the application of that technology not quite as straightforward as originally envisaged.

A considerable part of the Humanists prominent in the time of the previous reference technology have withdrawn into the Humanities' discipline from which they came (or into the industry) and are not interested in the interdisciplinary field any more.

These problems notwithstanding, however, a significant number of projects uses that technology and produces very promising results. Some of the pioneers disappear from the conference circus, however. It turns out, that used at a low level, the new technology can be integrated into a traditional Humanists job without the need of contact to a dedicated interdisciplinary community.

Low level courses at the universities decline. Talented students think that the skills taught are trivial.

Stage 3 (last five years, roughly): The proponents of the new technology are linked closely with the remnants of the interdisciplinary community of the previous stage. It has turned out, that the possibilities of the extremely simple technology are so far reaching, that dedicated institutions and infrastructures for their support are needed if the potential shall be used fully. Some of these

are connected to those of the previous stage; some of the institutions of the previous stage have not adapted sufficiently rapidly, so they have been closed and are replaced by new ones. Quite a few of the forgotten methodological discoveries of the previous stage are made again.

There is an intensive discussion about the need for a professionalization of the field. Almost all of the low-level courses at the universities have died. The broader community occasionally wonders, whether the few people still focusing on the application of these technologies are still real Humanists.

Stage 4: A significant new technology arrives and changes the basic parameters of the accessibility of information technology for Humanists radically. The cycle repeats.

It is extremely welcome, that changing habits of communication and the advent of mobile devices intensify the access to information about the Humanities. It is worrying, however, that the infrastructures currently being established seem to have few concepts for using particularly the mobile devices, though some of their usage characteristics are radically different from the PC-bound web. It is also worrying, that first indications of separate *fora* for these technologies appear.

Do we have to go through the cycle all over again?

3.4. Is Technology Designed by God?

My fourth observation is much wider in scope than the previous ones. In a sense it is not so much an observation about the Digital Humanities, but about some aspects of the way in which information technology is frequently reacted to in the society of the early twenty first century.

It also allows me to change an impression, which some of my earlier remarks have possibly created. On the methodological level I have observed with considerable unease, that current considerations of digital infrastructures for the Humanities seem to be overwhelmed by the notion of publishing information, so a infrastructure for the Digital Humanities in some of the current discussions can be so devoid of analytical considerations, that it becomes almost indistinguishable from a Digital Library (and not even a very sophisticated one, at that).

While analytical methodologies for the Humanities must come from the Humanities, it is clear that the development of Digital Libraries changes among other things the relative roles of scholars and librarians. Particularly the discovery, that information can only be understood within a context, so that context has to be preserved together with the information, if the later shall remain available, has created new and fruitful connections between Humanities and the curatorial disciplines, particularly in the context of the emerging concept of digital curation. It is quite a significant example, that Glasgow's *Humanities*

Advanced Technology and Information Institute (HATII)[4], starting in 1997 as a classical "Computing and the Humanities" unit, has then moved completely towards support for the memory institutions and runs MSc courses on "Information Management and Preservation", "Computer Forensics and E-Discovery" and "Museum Studies" rather than an MA that prepares for Humanities research within the Digital Humanities field.

Indeed, it seems to me that problems of Digital Libraries and keeping their content available over long periods of time are much closer to the Digital Humanities than pre-digital library studies and the pre-digital Humanities have been. (The long emotional relationship of Humanities' scholars and libraries notwithstanding).

Precisely because this relationship is close, I observe with considerable alarm, however, that the highly contradictory attitude of many Humanities' scholars to information technology seems to run particularly strong in the area where digital library studies and digitally supported approaches towards the Humanities overlap. It is contradictory, as on the one hand Humanities scholars and librarians consider information technology as deeply below the lofty intellectual domain on which they reside. Information technology, in that part of the world view, is a coolie's or peasant's job and it is quite obvious that these unthinking technicians should be directed by the wise intellectual requirements formulated by Humanists and Librarians who are in no way distracted by technical trivia. On the other hand, information technology is a given: If "the technology" provides data structures, software restrictions or constraints, these have divine authority and have to be accepted by Humanists and Librarians who will never be able to change them. Creating software is an arcane mystery; the further development of the technological frameworks can definitely not be predicted and has to be accepted as the inscrutable working of a superior entity completely removed from human influence.

As mentioned, this is an attitude, which I find detrimental and counterproductive in the development of the Digital Humanities, worrying and disturbing, as I seem to observe them in society at large as well.

To be more specific: Being rather active within research on digital long term preservation since 2004, I have watched with great interest the blossoming of many procedural standards and conceptual models. Unfortunately computer science – or software engineering, more appropriate in the short run – simply has no concept of long term persistency. Data are considered to be stored persistently, if they are still there, after a computer system has been powered down and rebooted. A formal concept of persistency which describes conditions of persistence across radical changes in system design simply does not exist as far as I know; nor does an engineering solution which would provide such a possi-

[4] <http://www.gla.ac.uk/departments/hatii>.

bility. Why we have spent uncounted millions in the last decade, to preserve conceptually digital objects, which have technically been designed without longevity in mind, instead of investing into the possibility to engineeringly redesign systems to make their content fit for long term survival, has always been a mystery to me.

An amusing example for the effects of this strange position towards technology: At a recent conference on long term preservation I had the privilege to listen to a brilliant presentation on the legal conditions for long term preservation. It discussed and proved in great detail, that in Germany it was not allowed to keep redundant copies in digital preservation systems as legally a customer was only allowed to keep "one copy" of a digital object as backup. As far as I am a judge of it, it was a brilliantly argued legal opinion. It set my mind wandering, though. As this audience certainly is aware of, a RAID system – virtually all file servers with any kind of claim of professionalism are using this or a later technology to the same effect – always keeps two copies of everything it holds, which is, why the data survive if one of the hard disks break. Of course the same is also true of any professional archiving system. This is not usually apparent, as on the level a user, say a lawyer, interfaces with the system, a file appears as one item, even if it is stored twice.

The socially agreeable private part of my mind drew the conclusion from that that obviously all German copyright lawyers are in default of copyright, by the simple act of using a professional server. The professional part of my mind was worried, however, that the conclusion drawn by the long term preservation community from the legal problem posed would probably not be to point out that the argument was mute, as it used a technically hopelessly obsolete concept of "copy", but an application to one of the funding agencies to find a legal solution to the problem.

I described this observation within the field of Digital Libraries and Digital Preservation. It is more obvious there, as this field is dominated by large scale, long term *fora* and projects, where the discussion can easily remove itself from the underlying technical problems. I consider this a serious problem for the Digital Humanities as a whole, too, however. Are the Digital Humanities consumers, who consume the work of Computer Scientists? Or are they able to influence the further development of the technology themselves, by including projects and people, who overlap with the community of Computer Scientists?

It is extremely welcome, that conceptual designs how to use the emerging technologies in the Humanities should be directed by the concepts of the Humanities. Should these concepts not also try to influence, which technologies shall emerge into the future?

Shall we be driven by the technological development, or shall we participate in driving it along?

4. Controversies: An Agenda

The observations described lead to the concrete controversial questions that have been submitted to the authors of the papers which follow. Please notice, that the presentations and videos of the lectures of the workshop at which these papers have been presented are available at <http://www.cceh.uni-koeln.de/events/CologneDialogue>. Please note also, that these pages are connected to a mechanism for comments, so more information than printed here will be available there.

These controversies form an agenda in so far as we hope to continue the discussions on each – or some – of them at separate workshops later. Proposals for these are welcome, and information may be available, at <WahnDialogue-info@uni-koeln.de>.

Do the Digital Humanities have an intellectual agenda or do they constitute an infrastructure?

This controversy is the one most directly derived from the preceding observations. To my great regret only one of the two opponents was present. *Willard McCarty* of *King's College London/University of Western Sydney* gives an intriguing view of the Digital Humanities as being deeply rooted within the larger concept of the Humanities in general (pp. 24-45). His opponent, *Patrik Svensson* of *Umeå universitet, Sverige*, had unfortunately to cancel at the last moment for reasons of health.

Are all approaches towards interdisciplinary research between the Humanities and Computer Science meaningfully represented by the current concept of Digital Humanities?

This controversy was derived indirectly from the observation, that while Digital Humanities *can* have a very broad definition, the term is used most of the time as a much narrower label. This can be observed by the disciplines and approaches covered: It can also be observed, however, when we notice that strong national traditions of Digital Humanities which go back to the very beginnings of the field, are clearly less well known than others. *Susan Schreibman* of *Trinity College Dublin, Ireland*, provides a thorough introduction into the way in which the understanding of "Digital Humanities" that shaped the *Companion to Digital Humanities* was conceived and how it has developed beyond 2004 (pp. 46-58). *Domenico Fiormonte* of *Università Roma Tre, Italia* describes, why he sees the currently prevailing concept of Digital Humanities to discriminate against some of the national traditions of the field (pp. 59-76).

What is the scope of the Digital Humanities? What is the relationship between individual disciplines served by them?

This controversy is closely connected to the preceding one, focusing not on the cultural, but on the disciplinary traditions. *Jan Christoph Meister* of the *Universität Hamburg* starts by a paper trying to locate the Digital Humanities within the disciplinary arena of the Humanities as a whole and the various levels of reasoning covered by them (pp. 77-85). *Jeremy Huggett* of the *University of Glasgow, United Kingdom* presents an overview of digital techniques and methodologies within archaeology, a field with a very strong tradition, in its conferences and publications with little overlap with Digital Humanities as the term is used in general, and discusses the connections of the fields (pp. 86-105).

What is the appropriate role of markup?

This controversy is not so directly related to the initial observations. It is derived from the fact that many followers of the Digital Humanities have a tendency to understand markup as the core of the field; or more specifically one logic and model of markup, as incarnated in XML/the TEI. Being such a basic majority assumption, it connects to many of the observations made – be it the scope of the services, an infrastructure should expect to provide; be it the relationship to the underlying technological assumptions. *Espen S. Ore* of the *Universitetet i Oslo, Norge* describes the general methodological depth and potential of markup as defined above (pp. 106-24). *Desmond Schmidt* of the *Queensland University of Technology, Australia* follows with an analysis of the limits of that approach and proposes alternatives to it (pp. 125-46).

Big structures or lightweight webs. What is the most sensible technical template for research infrastructures for the Digital Humanities?

While the relative importance of methodological v. infrastructural elements of the Digital Humanities has to be discussed, it is beyond doubt, that infrastructures are needed. Even so, the question remains what, in the light of recent technical advances in general, those of distributed systems in particular and even more broadly the change of paradigms of system development, is the best way to conceptualize and implement "an infrastructure". *Sheila Anderson* of *King's College London* presents the logic behind the current planning for European infrastructures (pp. 147-64). *Joris J. van Zundert* of the *Huygens ING, Koninklijke Nederlandse Academie vanWetenschappen, Den Haag, Nederlands* follows with a description of ways to understand infrastructures as systems with a smaller footprint (pp. 165-86).

"Digital curation" or "digital preservation" is a topic, which has originated within the world of digital libraries; recently it has been drawn closer and closer to the Digital Humanities. Using it as example: What is the proper balance between conceptual work and technology?

As the first controversy, this is almost one to one derived from one of the initial observations: On the somewhat ambiguous or contradictory relationship between the broader definition of the Digital Humanities and technology. At the same time it is the one, were the points of view presented by the opponents are furthest from each other. *Helen R. Tibbo* of the *University of North Carolina at Chapel Hill* (and former president of the Society of American Archivists) presents an uncompromising argument for the predominance of conceptual considerations in digital preservation (pp. 187-200). *Henry M. Gladney* of *Saratoga, California* is equally pronounced in his opinion, that at the technological level digital preservation has been solved by a concrete model he describes (pp. 201-17).

"Digital Libraries" have started their life as an answer to opportunities created by a specific stage of technical development. Where are they now, between Computer Science and the Digital Humanities?

The final controversy proposed generalizes this question of the relative weight of conceptual and technical knowledge to the very broad concept of *Digital Libraries*. As in the first controversy one of the intended opponents, *Edward A. Fox* of *Virginia Tech, Blacksburg, Virginia* has unfortunately been prevented by last minute health problems to attend. *Hans-Christoph Hobohm* of the *Fachhochschule Potsdam, Deutschland* nevertheless addresses him, or rather: the well known formal model Ed Fox proposes[5], directly in his contribution, which gives a fine picture of the substance of non-Computer Science library concepts behind Digital Libraries as part of the changing world of information infrastructures (pp. 218-29).

[5] <http://www.dlib.vt.edu/projects/5S-Model/> (Accessed June 16th, 2012).

Special References

Contributions within this Special Issue HSR 37.3:
Controversies around the Digital Humanities

Anderson, Sheila, and Tobias Blanke. 2012. Taking the Long View: From e-Science Humanities to Humanities Digital Ecosystems. *Historical Social Research* 37 (3): 147-64.

Fiormonte, Domenico. 2012. Towards a Cultural Critique of the Digital Humanities. *Historical Social Research* 37 (3): 59-76.

Gladney, Henry M. 2012. Long-Term Digital Preservation: A Digital Humanities Topic? *Historical Social Research* 37 (3): 201-17.

Hobohm, Hans-Christoph. 2012. Can Digital Libraries Generate Knowledge? *Historical Social Research* 37 (3): 218-29.

Huggett, Jeremy. 2012. Core or Periphery? Digital Humanities from an Archaeological Perspective. *Historical Social Research* 37 (3): 86-105.

McCarty, Willard. 2012. The Residue of Uniqueness. *Historical Social Research* 37 (3): 24-45.

Meister, Jan Christoph. 2012. *DH is Us* or on the Unbearable Lightness of a Shared Methodology. *Historical Social Research* 37 (3): 77-85.

Ore, Espen S. 2012. Document Markup – Why? How? *Historical Social Research* 37 (3): 106-24.

Schmidt, Desmond. 2012. The Role of Markup in the Digital Humanities. *Historical Social Research* 37 (3): 125-46.

Schreibman, Susan. 2012. Digital Humanities: Centres and Peripheries. *Historical Social Research* 37 (3): 46-58.

Thaller, Manfred. 2012. Controversies around the Digital Humanities: An Agenda. *Historical Social Research* 37 (3): 7-23.

Tibbo, Helen R. 2012. Placing the Horse before the Cart: Conceptual and Technical Dimensions of Digital Curation. *Historical Social Research* 37 (3): 187-200.

Van Zundert, Joris. 2012. If You Build It, Will We Come? Large Scale Digital Infrastructures as a Dead End for Digital Humanities. *Historical Social Research* 37 (3): 165-86.