

CRITICAL EXPLORATION OF MEDIEVAL PHILOSOPHY WITH DIGITAL HUMANITIES TOOLS: CHALLENGES AND SOLUTIONS

MIHAI MAGA*

Abstract In the last few years, the Digital Humanities have gained more and more attention from the medievalists because of the perceived advantages in encoding and analyzing digital critical editions. But there are many challenges for the scholars working on medieval philosophical texts. We shall try to identify some of these challenges and to discuss some solutions on three subjects: critical editions in digital format, organizing a digital corpus, and computer analysis. A general conclusion emerges from the principles and illustrations discussed here: the issues must be addressed through critical thinking in order to achieve relevant and accurate results.

Keywords Digital Humanities, critical edition, medieval philosophy, semantic encoding, digital text corpus, semantic analysis

Every medievalist nowadays uses a computer. And, at least once, everyone has done a search on the internet for some medieval text, or manuscript, or medieval author. This practice has brought us into a new age of studies, but it also unfolded new challenges, not only regarding the technical difficulties in using a computer, but questions of principle as well. A new research domain was born from the attempt to put the computing technology at work for the humanists, and from the 2000s it is

* *Babeş-Bolyai University, Cluj-Napoca.* mihai.maga@ubbcluj.ro

This article benefits from the support of the Babeş-Bolyai University Young Researchers Grant no. 31797/2016 “The Impact of the Political Philosophical Sources on the University Discourses from the 14th–15th Centuries in Central Europe,” project leader: Dr. Mihai Maga. Some of the research on digital tools was previously carried out within the ERC Starting Grant no. 313339 “THESIS. Theology, Education, School Institution and Scholars-network: dialogues between the University of Paris and the new Universities from Central Europe during the Late Middle Ages”, project leader: Dr. Monica Brînzei.

called Digital Humanities.¹ Its ambitions are high, but it is still in an early phase and it still needs a lot of effort to reach its potential. It seems that even its definition is a point of controversy.² Nevertheless, it gained popularity due to its promises and the quick pace of its development.³

There are some general controversies concerning the Digital Humanities, and some of them come from reducing the intellectual work in humanities to technicalities and computer-processing standards. Adam Kirsch, in his article “Technology Is Taking Over English Departments. The false promise of the digital humanities”⁴, warns about “anti-humanistic manifestations of digital humanities” in the sense that “digital humanities has less to do with ways of thinking than with problems of university administration”. Indeed, the initial intentions of the pioneers of this area, that is to use the computer as an aid for text study, were gradually subordinated by the emergence of a domain by itself established and judged through different principles. Firstly, the focus on the technology more than on meaning, with the hidden belief that the computer can produce new knowledge which can be quickly appropriated by an author who does not need to read the texts him/herself. Secondly, the promise that this technical revolution will radically change the humanities, bringing them into the most modern industry standards in order to accommodate the current paradigm of corporate project-based research which can be economically quantifiable and bureaucratically manageable. Thirdly, the dream of a universal device which can unravel all the major issues of the human intellect: historical incompleteness, plurality of meanings, objective knowledge, and so on. If we try to summarise these arguments, they all claim the ignorance of critical thinking, which is the foundation of humanistic studies. Therefore, I shall try to approach this subject from the point of view of the modern medievalist, who is, first of all, an intellectual, a critical thinker. It is her or him who (still) has the lead in using the new technology to produce meaningful knowledge.

In order to be more explicit, I shall illustrate the generic approach on these issues with some of the challenges from my current project. The project intends to explore the political content of the late medieval university discourses from the Central European universities. These speeches (*recomendationes, sermones*) often

¹ *A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens, John Unsworth (Oxford: Blackwell, 2004).

² Fred Gibbs, “Digital Humanities Definitions by Type,” in *Defining Digital Humanities: a reader*, eds. M. Terras, J. Nyhan, E. Vanhoutte (Farnham–Burlington: Ashgate 2014), 289–297.

³ There are at least 191 Digital Humanities centers in the world today, according to CenterNet, <https://dhcenternet.org/centers> (accessed on 10.01.2017).

⁴ Adam Kirsch, “Technology Is Taking Over English Departments. The false promise of the digital humanities,” in *New Republic*, May 2, 2014, <https://newrepublic.com/article/117428/limits-digital-humanities-adam-kirsch> (accessed on 10.01.2017).

contain references to authoritative texts in order to argue points of view on current events or disputes on ideas. They have the role of reception and reinterpretation of the political thought tradition, and of propagation of these ideas to a broad audience. Because they have been less studied and because most of them survive only in manuscripts, the main challenge is to go from the discovery of manuscripts, through procurement, critical editing and analysis, towards building a general representation of this phenomenon. Given the fact that the number of the preserved discourses is quite large (in terms of hundreds), the Digital Humanities approach seems the most suitable. But, as we shall see, this too raises problems, and demands a critical treatment.

1. Critical editions in digital format

For some, producing a digital edition represents the end of Digital Humanities. There are a lot of projects on the internet whose only goal is to reproduce works in digital format in order to reconstitute, preserve and disseminate certain texts considered worthy of this action. But the critical thinker may simply ask: why? Of course, it is a meritorious effort to make available such texts which may be of crucial importance for the history of thought, or which may be endangered by their material degradation. However, the fact that they exist does not directly enrich our knowledge, unless they are read, analysed and interpreted. At this point, there seems to be an unequal race between the production and the interpretation of these texts: more and more become available (some through careful editing efforts, some just through machine recognition), but some are understudied, even ignored by those who should integrate them in a renewed view of the subject (and sometimes the histories taught in schools or popularised elsewhere only refer to the “classical” few).

In terms of new media, we should ask ourselves: how do we connect the resources with the consumers? How do we bring these texts to the attention of the specialists? For my project, the question took the form: how do I know which texts are more relevant for my research?

It is primarily for this reason that the semantic encoding was invented. Basically, it explicitly exposes the meta-textual content in a way that machines and humans can unequivocally identify certain meaningful elements in a standardised way. The semantic encoding makes the difference between a simple text document and a digital edition. Therefore, we must clearly differentiate between digital edition and computer typesetting, because computer typesetting has the goal of producing visual content, namely printed pages, while a digital edition is an electronic document semantically encoded.

The signification of a text is conveyed through the text itself (characters, words) and meta-textual elements (formatting, position etc.). In the visual (physical) format, the meta-textual signification is inferred from general or local rules. These rules are sometimes explicitly indicated: by general known conventions (e.g. the title

is larger, centred, separated from the text), by an explicit label (e.g. *Notes*, *Index*, *Contents*), or by an explicit declaration (e.g. typographic conventions described in a *Preamble*). But other times the conventions are unclear and they are inferred from the context (e.g. italics may be used for foreign words, short quotes or expressions, referenced titles, emphasised words etc.) Sometimes the conventions vary so much that they can induce ambiguities in reading. And this does not happen only to us, but also to medieval scholars (there are a lot of examples of foreign marginal notes integrated in the texts by copyists who mistakenly took them for the author's text, or notes written by authors but left out by copyists). Therefore, in the semantic encoding, every meta-textual element should be clearly and unambiguously labelled.

From the beginnings of the Digital Humanities in late 40s until late 80s there were many efforts to establish a standard for digital critical editions. From the elaboration of the first draft of the TEI Guidelines⁵, a *de facto* standard emerged and reached general acceptance⁶. The Text Encoding Initiative is a collaborative project maintained by TEI Consortium which aims to develop guidelines for the digital encoding of literary and linguistic texts. It is merely a structured documentation for standardised annotation of texts in XML (Extensible Markup Language) format which tries to address every possible usage of digital semantic encoding: text structure, manuscript description, critical apparatus, bibliography, graphics etc. For example, instead of implying that italics indicate a title and small caps indicate a name in a fragment like "quod scribit KATHO in libello suo *De doctrina morali*", we note this in TEI XML as: "quod scribit <name ref="#Cato">Katho</name> in libello suo <title ref="#Cato-Disticha">De doctrina morali</title>" and, besides clearly and abstractly indicating the meaning of these words, we can also add more meta information like, for example, links to standardised reference lists where we can further enrich the data with bibliographic details, perhaps even references to full text electronic sources. This allows developing an arborescence of metadata (or, in information science terms, an ontology) which is unattainable in printed format, due to its limitations. In its abstractness, this semantic markup does not have a visual representation, but documents can be easily transformed into visual formats (standard printed editions or web page) by applying a certain styling to the elements globally.

All of these are merely technical aspects. The challenge for a scholar here is to decide what types of metadata are relevant for annotation and how they should be annotated, which is the critical usage of digital instruments. At one end, there is the desire to include every possible detail of the text, from philological and material details to complete external references for everything connected to the text: citations, geographic data, biographies etc. At the other end, there are the resource

⁵ *Text Encoding Initiative*, <http://www.tei-c.org/About/> (accessed on 10.01.2017).

⁶ *A Companion to Digital Humanities*, eds. Susan Schreibman, Ray Siemens, John Unsworth (Oxford: Blackwell, 2004).

limitations: working time, deadlines, computing limits, and limited purpose of the project. Therefore, one should decide what should be annotated and how. Furthermore, these decisions must be clearly stated so that the scientific community may judge the quality and usability of the work.

Coming back to my project, I am limited by two goals: to produce some critical editions for a selected number of unedited texts and to expose the textual sources referenced by the authors. Consequently, I decided to encode the critical apparatus and the references to other works. But one should not forget that other users of these digital editions may have different goals. As a result, I set some principles for myself which may be useful for other scholars working in this field. I do take into account and adhere to other general principles proposed by other scholars, especially the four put forward by J.C. Witt in *Mediaeval Commentaries on the Sentences of Peter Lombard*, 3rd volume⁷. Here are my own principles for developing an editing application for medievalists:

1. Parallel aligned segmentation method. Each word should be considered a text segment and this may correspond to different readings in the manuscript copies and there are at least two possible readings: the presence of the word and its omission. This method poses problems for those texts which differ substantially, but may be treated as block omissions and block additions. With the aid of external scripts, the similar events in consequent segments may be collapsed into a single critical note.

2. Headless edition. The advantage of digital editing is that the variants of the text can be encoded without the need to establish the 'right' version of the text. This speeds up the editing and avoids long suspensions of work awaiting the consultation of missing manuscripts. It may seem peculiar for the classical humanist, but it is implied by a postmodern view of the text in which every instance of this text is merely a witness, a version. Therefore, even modern prints and what is known as 'definitive' editions may be regarded as just manifestations of the same source and may be studied with the same principles, identifying differences, mistakes, additions, grouping of variants, sources and so on. And, in this sense, our work in progress on new digital editions produces only a new manifestation of the text. What we think is the 'original text' is most of the time lost and what we reconstitute is based only on surviving witnesses, so the 'prototype' is a reconstruction which is the result of our sources, tools and conceptions. That is why, instead of concentrating on early restoration of the sense with many possible afterthoughts and revisions, we can represent the sources in an adequate manner and, when doubts are clarified, we can add a new, virtual witness or manifestation, which is our critically established text. For us, this would be the 'head edition', but for further work this may be considered

⁷ Jeffrey C. Witt, "Texts, Media, and Re-Meditation: The Digital Future of the *Sentences* Commentary Tradition," in *Mediaeval Commentaries on the Sentences of Peter Lombard*, vol. 3, ed. Philipp W. Rosemann (Leiden–Boston: Brill, 2015), 511–514.

to be just a recent manifestation, and, if new sources and principles occur, a new head edition may be elaborated with minimal effort.

3. Reusable semantic encoding. As I have said before, the goals of the editor may differ from other intentions further developed regarding the same text. Therefore, in order to avoid repeating the electronic effort for further usage, the resulting edition should be reusable. This is done firstly through the careful observation of the standards, for example the TEI Guidelines. The result is an interchangeable file format which can be integrated in other tool chains and procedures. For example, a critical edition may be used for linguistic analysis, or for generating structured diagrams. This may complicate the decisions of the editor who is limited by the assumed goals, but it ensures the compatibility with other software. It may happen that at a certain point the researcher discovers that there is a new tool suitable for his work; having a compatible edition means that he or she can immediately use the new tool.

4. Nonlinear workflow. For the classical procedure of critical edition, there is a typical workflow in which each step must be completed in order to pass to the next one. In short, one should gain access to all the manuscripts, transcribe one which is considered better (but this is of course a subjective decision), collate the others, elaborate the *stemma codicum*, establish the edition based on stemmatic decisions, identify the sources, prepare the apparatus, write the introduction. In a digital edition, these steps may be attained in almost any order. An example is the headless edition principle, by which a critical apparatus may be prepared before establishing the head edition, but this can go further. One can work on a short sample of text in order, for example, to establish the *stemma codicum*, or one may do the collation of available manuscripts without waiting to obtain others difficult to reach. At any point, the work done may be continued towards further steps, because what is done in a standard way may be easily integrated with the rest. Furthermore, the nonlinear workflow allows early access to the text through other instruments, e.g. partial online publication, statistical analysis, computer aided searching. This workflow is also convenient for collaborative projects where different persons may work on different editorial aspects and their contributions are continuously integrated in the project, no matter their state of progress.

5. Computer-aided editing. The benefits of computer-aided editing are generally visible from the other principles above, but there are certain points which must be emphasised. First, automatic data validation may be carried out in the background or when a step is finished; this ensures the integrity of the data, the conformance to the language specifications, the completeness of the structures and the overall compatibility. Thus, typing mistakes can be signalled early, references are verified and nonstandard annotations are normalised. Secondly, computation statistics may be performed on the differences between manuscripts; this may be used to highlight the rapport between manuscripts, elaborate a provisional *stemma codicum* (I was even able to develop some algorithms to calculate an approximated

stemma by counting manuscript differences), and to single out some peculiar situations which are due to editing mistakes. Thirdly, automatic comparisons may be performed both inside the text and with external sources, for example to identify textual references, copied fragments or recurring expressions. The advantage is that, since an electronic annotated text already exists, these instruments and procedures are directly available and require no effort to access.

6. Intuitive graphical interface. For those ones who worked in TEI, the first difficulty is to learn and read computer encoding. Even if the XML subset used by TEI is intended to be both human friendly and computer compatible, writing XML code is complicated, redundant, prone to human errors and visually awkward. That is why there is a need for visual interactive editing in a simple graphical interface with maximum control and minimal interaction. There are some complaints that the TEI Initiative was too focused on document encoding and it left out the human readable representation of the data. There are several projects aiming to surpass this difficulty⁸, but every one of them lacks some aspects, because of one of the two following reasons: by aiming to be a universal tool, they ignore specific usage; by aiming at specific usage, they miss all the other practices. From my point of view, universal tools, as desirable as they are, may not be easily imposed and the reason comes from what I discussed earlier: the critical usage of electronic instruments. There are indeed encoding standards, but there are no standard goals. The originality of a scholar should not only come from the casual discovery of interesting texts, but also from new ways of questioning the texts. Therefore the interface should be modelled after the principles of reading the texts. One of the projects created by me for this purpose, provisionally called Collexy⁹, was developed step by step from the needs of collating manuscripts. Therefore, the most commonly used functions have simple access by a single click or a keyboard shortcut; those used less often are accessible through menus or attached to other tools; finally, there is an integrated XML editor for any other encoding which does not have an interactive interface. The incompleteness of this application (and consequently its only beta version) is the result of rising and solving punctual problems and a balance between automating functionalities and manual editing complex structures. But it serves the mentioned principles by providing the functionalities for segmented, headless edition, TEI compatibility, nonlinear workflow, aiding tools and a graphical interface.

As we have seen, the critical thinking is not only necessary for establishing the edition of a medieval text, but also for choosing the standards and the instruments for editing. The general aim is not only to produce the edition, but also to make this process transparent. This way the editor's work may be criticised by

⁸ For example, the "Author" mode of the oXygen XML Editor, in <https://www.oxygenxml.com/>. Other editors are listed on TEIWiki, <http://wiki.tei-c.org/index.php/Editors> (accessed on 10.01.2017).

⁹ Collexy, <http://collexy.mihaimaga.ro/> (accessed on 10.01.2017).

others, errors may be corrected, further discoveries may be integrated and new questions may be answered.

2. Organising a digital corpus and additional information

The work of medievalists focuses not only on isolated texts, but on subjects, periods, authors, places of intellectual development and so on. Therefore, there is a desire to be able to access all the texts from a particular set. Furthermore, most of the institutional projects target a number of sources composing a defined set. By organising a set of digital sources, a digital corpus is created.

The usefulness of text corpora in general is obvious: anyone can have access to a standardised version of every text in a corpus and the entire collection is available. The usefulness of digital text corpora goes beyond that: text searches can be quickly performed on the whole collection, statistical analyses may be achieved fast, it may be used in hypotheses testing or in automatic linguistic analyses. The labour of many years and many people working on paper some time ago in order to produce tables of indices and concordances can be carried out in milliseconds by computers and is less prone to human error.

But how do we organise a digital text corpus? It must be structured in such a way that it is simultaneously adapted to the specificity of each text and to the demands of the users. At the same time, it must start with something and end by comprising all the available texts and details. In information science terms, it must be organised as a dataset in a tree-like structure. But, for predictability and compatibility, the structure must be homogenous, which means that every division must follow the same principles. This may seem redundant, but when we try to apply these principles strictly, we remark how often we admit exceptions and heterogeneity.

For example, there is still a debate about how the corpus of the *Sentences* commentaries and other scholastic corpora should be organised. A workshop on this subject was held at Basel¹⁰ and a proposal was made. The point was that there is no consensus about how to interchange the data representing digital editions and digital information on scholasticism. Every project has its own principles in organising the data and when someone wants to reuse data from another project (e.g. to integrate an external digital text in its own functionality), he or she must deal with its different principles and work on data conversion, which is futile. For example, someone may build the tree on the assumption that its levels are: author, work, chapter, paragraph. But this structure fails on several instances: the identification of the author may be problematic (is 'Anonymous' a single author, or a different author for each work? what about works proven or suspected to be written by the same anonymous author?); the work may have several different versions (is it

¹⁰ *Linked Data and the Medieval Scholastic Tradition*, University of Basel, August 17–19, 2016, organized by Jeffrey C. Witt and Ueli Zahnd.

still the same work if it differs a lot? is an abbreviation by some other author considered the same work or should it be assigned to the abbreviator? what about a heavily altered copy?); the work may be divided differently than in chapters (how to deal with multi-level divisions and how much of them is to be stored?) These issues lead to some quite philosophical questions: what is an author? what is a work? There are proposals to abstract this,¹¹ but for some projects it may seem redundant. In addition, the digital humanist must avoid inventing information for the sake of the dataset structure (e.g. forcing items into categories where there is no scientific motivation for their inclusion).

Concerning my project, I faced the same questions on data modelling. Not only that I had to put all the discourses and information about them together, but I had to put them in such a way that they may be easily interrogated. I had no intention in solving the tree problem, so I took a somehow different approach: different types of data are stored separately, but they are put together by means of pointers and a tree is dynamically generated by the computer. There is a testing website called RecommDB¹² for illustration. There are several TEI documents containing data tables: manuscripts, authors, discourses, sources etc. Each entry points to other entries, for example a discourse points to one or several manuscripts, to one author, to several sources. By accessing any table we may consider that we are on a first level of a tree, having the other linked data as descendants. Accordingly, when we access the authors list, we may traverse the tree by going to their discourses, then from discourses to manuscripts and so on. For the purpose of analysis, I also encoded geographical places and chronological data (for manuscripts, for discourse authors, for ancient source authors) so that I can generate maps and timelines. The advantage is that the corpus can be easily extended through adding new datasets and pointers. When a consensus about the structure of datasets is reached, it should hopefully be just a particular case of traversing the corpus through a specific tree model. Therefore, we could also abstract the model design.

From what we have seen, the building of a digital corpus needs important critical decisions. It does not only pertain to the specificity of the data, but to the interrogation possibilities as well. A corpus which is inadequate to certain interrogations may produce false results when these interrogations are still performed. But the work will not, or should not rely only on the computer¹³. We already have the historical experience of organising text corpora: the libraries. Every

¹¹ Jeffrey C. Witt, *A Distributed Text Service Modeling Proposal – case study of the scta ontology and api*, <http://lombardpress.org/2016/06/12/DTS-modeling-proposal/> (accessed on 10.01.2017).

¹² Mihai Maga, *Recommendaciones Database*, <http://www.mihaimaga.ro/recommdb/> (accessed on 10.01.2017).

¹³ See Jerome McGann, “Information Technology and the Troubled Humanities,” in *Defining Digital Humanities: a reader*, 49–65.

time when someone is in doubt about a digital edition, she or he must be able to check the edition, the manuscript, the catalogue in the library.

3. Computer analysis and interpretation

A computer does nothing with the data unless instructed so. Thus it needs software and commands in order to process the data, but this is also a complex situation. While the file format standards are developed for general compatibility, hence a TEI well-formed file should be processed with any TEI aware software, the reality is different. Firstly, because of the decisions made at the encoding level, which we discussed earlier, and by which only a selected subset of instructions is used. Secondly, because each software application assumes that the file contains certain data and structure and only relies on processing it. For example, there are at least three very different ways of linking the critical apparatus to a text in TEI,¹⁴ but the applications usually only deal with one of them (the parallel segmentation method seems to be the most used). The software available may be divided into two categories: generic software which it happens to also process TEI (most of the advanced XML processors can validate a TEI schema), and specific software, built for a certain project, corpus or analysis. The first one may be difficult to use in specific semantic interrogation; the second one may be incompatible with other encoding strategies.

For a beginner in Digital Humanities, the practice may be quite discouraging: most of the scholars who work in this domain have at least medium level knowledge and training in programming. The situation may be explained intuitively by mentioning the early development of this field, the scarcity of easy-to-use tools, or the conflicting approaches of encoding and processing data. But there is a more profound reason: after putting the texts in computer files, any question on these must be formulated in computer language. Consequently, for a simple text search, the retrieval of results must be aware of the semantics embedded in the file and of the document structure (for example, if the complete titles mentioned by a work are not stored within the work, but pointed to a different location, one will not find them by searching in the work). Therefore, a medievalist who wishes to use the electronic texts to their maximum potential must learn some computer language. The alternative is to research the available tools, to wait for new ones to be developed, or to support and advise the colleagues with proper training in order to achieve his own goals. Otherwise, he should be familiar with the fact that a question like “Which works belong to the author X?” should be reformulated as “SELECT `title` FROM `bibl` WHERE `author`="X"” (in MySQL) or “//bibl/author[.="X"/]/./title” (in xQuery). By reformulating the questions in computer language, they seem more

¹⁴ *The TEI Guidelines*, “12.2. Linking the Apparatus to the Text,”: <http://www.tei.org/release/doc/tei-p5-doc/en/html/TC.html#TCAPLK> (accessed on 10.01.2017).

objective, but we should keep in mind that this translation does not surpass the subjective thinking¹⁵.

Nevertheless, for most users of a digital corpus a small number of usage scenarios are the most prevalent. The text search functionality is probably the most commonly used. It may seem trivial for users, but a lot of effort must be put into a good search functionality, not only through programming, but also through linguistics and mathematics. Because the user will be happy to find the most relevant results even if they have different spellings, or different grammar inflections, or the word order differs, the developer must invest in language processing and algorithmic computing. Thus, the text must be lemmatised and the application must compute the relevance of results. Yet, the application should make its computed decisions visible to the knowledgeable user while keeping a clear interface for the unaware one.

From my experience, there is quite a thin line between the needs and the possibilities. For example, by implementing full lemmatization and complex search algorithms based on score computing and semantic aware, the resulting data storage is huge and the requests cause a high usage of computing power to the extent that a medium sized server may take minutes to return the results for certain complicated requests.

This is why the general usage in computer programming is to separate and to abstract the processing applications from the data model. This way, the integrity of each part can be ascertained, the same data may be consumed by different applications, and the same application may be fed with data from any other project. At a certain point in the future we may envisage a complex network where a constellation of different applications can perform specific tasks on big datasets gathered together from different sources. Consequently, if an application is unsatisfactory, one can access the same data through another one; at the same time, each developer can focus on what he does best, leaving other tasks to other projects connected to the network. But all these tools and resources should not be left unattended: there are many caveats in using the computer for scientific research in the humanities.¹⁶

Conclusion

I tried to illustrate here some of the challenges of the Digital Humanities for a medievalist and to identify some possible solutions. Some of them may be common to other humanistic fields, some are specific for medieval philosophy. There are

¹⁵ See Leighton Evans and Sian Rees, "An Interpretation of Digital Humanities," in *Understanding Digital Humanities*, ed. David M. Berry (Houndmills–New York: Palgrave Macmillan, 2012), 21–41.

¹⁶ See Bernhard Rieder and Theo Röhle, "Digital Methods: Five Challenges," in *Understanding Digital Humanities*, 67–84.

some fundamental questions behind the technical aspects which demand an investigation of our mind and of the human-computer interaction: what is a text? what is semantics? do computer results have any creative value? can programming be an expression of humanistic research? is there a new philosophy behind Digital Humanities? These are open questions and they should be left open while this scholarly area is still under development. Yet the developer must have some answers for themselves, and any user of their products must know about its fundamental assumptions, else she or he may be led into the temptation of universality and objectivity.