A Smell of Higher Honey

E-Humanities Perspectives

INGER LEEMANS

147

Is e-humanities 'The next big thing' and are we on the verge of a 'Humanities 3.0'? The field of e-humanities, at this moment, seems to be too diverse and scattered to move in sync. One of the challenges that face e-humanities research at this moment lies in the field of uniting data digitisation/management and data analysis/representation. Here, heritage institutions, libraries, archives and universities should cooperate closely. Existing (digital) humanities data corpora should be connected and integrated through on the basis of question-driven analysis possibilities. However the biggest challenge would be to take hermeneutics as a base for e-science, by developing digital analytical strategies for the knowledge fields unique to the humanities – meaning attribution, interpretation and concept formation in text, image, sound, object or space, and combining these with numeric data about production, consumption patterns, networks, et cetera. How can we trace complex and essentially contested concepts? Can digital analysis set us on a trail for new interpretations? How can we accommodate the complexity and ambiguity of the sources with which we work and develop digital methods to automate the way humanities scholars look for patterns, interpret and evaluate them? The evaluative aspect of e-humanities could really alter our research field, since for a long time humanities scholars have tended to work with implicit evaluative schemes. The advantage of e-humanities research therefore would lie not only in the fact that hermeneutics are structurally taken into account in digital analysis, and that new patterns and interpretations might be found, but also that we would gain more insight in the kind of questions we pose, the steps that we take during this process and in the validation of the results.

> O golden letter age! Sagacious spring!
> O fragrant air, after storm's threatening!
> Joost van den Vondel, *Inwying der Doorluchtige Schoole*

A higher honey! Its smell
Drove us from where we dwell.
Martinus Nijhoff, 'The Song of the Foolish Bees'

Is e-humanities 'The next big thing'?[1] If we are to believe Rens Bod, there
might have been a storm, but for the humanities the sky is rapidly clearing.
Digital humanities will not only supply the humanities with new insights, but
will also give us the opportunity to pose new and undreamed-of questions,
to develop critical reflection on the discipline and even to give the computer
a role in gauging divergent (either more or less suitable) interpretations. If
this were not enough, e-humanities will subsequently also shower us with
money, since we can now market our knowledge products, while at the same
time it will help us to explain the use of the humanities to society in general
and business in particular. 'O, golden letter age!' Vondel would exclaim; after
all those decades on the cusp of oblivion, the humanities' air is filled with
promising fragrances. Or are the critics right and will the e-humanities bees,
enticed by the smell of higher honey, eventually reach an air too rarefied, to the
embitterment of the traditional humanities flowers on the ground? Will they
rise up as sparkles, 'distraught, discarded, disembodied'?

A soft buzzing – humanities 2.0

Let us begin with Bod's statement that the use of digital methods is bringing
innovation to the humanities at a fundamental level, to wit, that of the
research question. According to Rens Bod, 'due to the advancing digitisation'
we are now 'able to pose questions which humanities scholars have for a long
time been avoiding'. What Bod has in mind are questions about universal
patterns for which large text or data corpora can be searched, questions such as
'does the course of history show patterns?' It remains to be seen whether this is
precisely the innovative aspect of e-humanities. Curiously, Bod himself already
tells us that, even before the arrival of digital humanities, many humanities
scholars have posed questions about universal patterns and have gathered data
on a large scale in order to answer these. Take for example, the historians who
have tried to explain the unequal distribution of wealth between the West
and other parts of the world and the discussions on the causes and patterns of
subsequent industrialisation and modernisation processes. Take the Braudel
branch of the *Annales* school, take diachronic research into revolutions, revolts,
strikes, but take also literary scholars trying to trace universal story patterns in

---

1   This article is a reply to Rens Bod's inaugural
     address *Het einde van de geesteswetenschappen 1.0*
     [*The End of the Humanities 1.0*] (December 2012).

folk stories, fairy tales or nursery rhymes, take research into visual or musical *topoi* or research into 'the romantic' as a universal motif.

The search for recurring or universal patterns is nothing new in the humanities – and this kind of questions is not reserved to language scholars either: they have a long tradition in many other fields of the humanities. That the use of technology in this kind of research can be of help seems evident to me. That is one of the reasons why new database projects, collecting numeric data or text (fragments), originated in these fields. It is therefore debatable whether the question did not drive the technology, instead of vice versa.[2]

It is also debatable whether humanities 2.0 bridges two different research paradigms operating within the humanities, between hermeneutic research, traditionally aimed at the understanding of the unique and context-determined, and the explanatory humanities that has concentrated its effort on pattern recognition and the quest for laws. The humanities 2.0 projects mentioned by Rens Bod, especially those projects undertaken within the framework of the Amsterdam Centre for Digital Humanities, appear to fall largely within the second tradition. They compare data on art producers with data on locations, thus enabling us to discover where which painters lived in the Amsterdam Golden Age; they count quotes and references from and in academic texts, or compare word patterns or stylistic data in different (literary) works.[3] Is this hermeneutical research as envisioned by Dilthey?

Moreover, does this kind of research supply us with a unique voice within the broad e-science field? Does it help us in making an essential, new contribution? The prime advantage of the humanities within the e-science field would be the fact that we can work with 'complex, fuzzy, incomplete data'. It does seem that our present projects propose to neutralise this complexity at the earliest possible stage. The challenge would seem to lie precisely in accommodating complexity and ambiguity and to develop digital methods to automate the way we look for patterns, interpret and evaluate them.

2   That is not to say that the supply and incorporation of large data files cannot bring innovation to these research disciplines; they can facilitate the analysis of research questions or challenge existing assumptions.

3   The projects look for recurring patterns in the assembled data, partly checking assumptions ensuing from earlier accumulated humanities research. For instance, the Huygens ING project *The Riddle of Literary Quality* investigates whether literary works which have been awarded a place in the canon (among other things on the basis of humanities arguments) can be stylistically distinguished from non-canonical works.

Hermeneutic computing

The greatest advantage seems to lie in the field of interpretation of meaning attribution and concept formation in text, image, sound, object or space. These are the knowledge fields unique to the humanities. Digitisation of such questions and analyses would be truly innovative. For this, I have in mind digital forms of conceptual history: how can we trace complex and essentially contested concepts such as 'freedom', 'equality', 'citizenship' and related connotations in texts? How can we show that within a certain text or context, different, conflicting perspectives on a single concept can operate? What is the genesis of conflicting historical narratives? What models can we develop to make visible and understandable that in different eras, different groups have used different interpretations of specific concepts?

Concept formation is only one possible example, but in my opinion it can illustrate where our strength might lie – in making interpretative processes digitally accessible, measurable and visualisable. This applies to both historic processes – the ways in which people attribute meaning to the world around them, make classifications, comments, et cetera and the way we analyse and interpret these as humanities scholars. The next step should be to combine these insights with numeric data, for instance about producers and production locations of texts, in order to grasp the cultural and geographic infrastructure of texts and concepts. Thus, *erklären* and *verstehen* can be combined.

Have we, with this hermeneutically driven, but combined approach to e-humanities, arrived at humanities 3.0 the next phase of digital humanities Rens Bod proposes? This is hard to say. What he understands under humanities 3.0 sadly is rather unclear. All we have is the specification 'critical reflection, or the connection of patterns with interpretation'. Although I do not know whether we mean the same, I am willing to make a proposal for the further elaboration of these terms. I hope I have made it clear above what I have in mind for 'the connection of patterns with interpretation', to wit the tracing of patterns in (conflicting) interpretations. Before going into speculation on the 'critical reflection' possibly valuable for humanities 3.0, I will first consider the concept of e-humanities, which is itself essentially contested.

E-Humanities – attuning to data accessibility & data analysis

Digital humanities and e-humanities are umbrella terms. Nearly any form of humanities research employing ICT can be bracketed under e-humanities, from data management and data digitisation to data analysis and representation or visualisation. In this broad field many people, groups and institutions are at work and I think it is somewhat reckless to put them all in one paradigm such as humanities 2.0, and furthermore to expect them all simultaneously to make the turn to 3.0. This is impossible, because the fields concerned are operating

largely independently. The developments in the field of digitisation of texts, or the making accessible of objects in the museum world do not walk apace with the developments of geographical information systems, or of text and concept mining, if only because data digitisation and management usually involve more expenses, work hours and legal problems. Data analysis and representation are less dependent upon institutes and management technical problems, but the progress in these fields is dependent upon the available software or of the fragmentation of research data and tools.

The fact that these two sides of e-humanities can only be brought together with great effort appears for instance from the manner in which the Dutch university libraries digitise their collections. Once there was the plan to create one 'Libratory', the Dutch version of the DPLA – the Digital Public Library of America. As a united force the Dutch university libraries would make their old collections digitally available. In so doing, they would develop new forms of digitising material to facilitate digital scholarly research. Sadly, this project died a silent death. The KB, the National Library of the Netherlands, decided to put out a large part of its collection to Google Books, for mass digitisation. In the words of Sander Dekker, state secretary for Education, Culture and Science: 'The great winner is the public, especially scientists, who now get better access to the collection of our national library'.[4]

The first 80,000 books are already online, both via Google Books and via boeken.kb.nl, but the question remains whether this milestone will create any momentum for digital scholarly research. The scans can be read, but the OCR quality is still very limited, the obvious link with the metadata of the KB catalogue, Picarta or the *Short Title Catalogue of the Netherlands* (STCN) has not (yet?) been established and to be actually able to mine a text, you have to download a pdf via Google Books and then store it as a text file. Making a selection on the basis of year of publication, genre, publisher et cetera is not possible and bulk downloads are not supported. It is therefore impossible (as it is with that other beautiful site of the KB: Early Dutch Books Online) to select a group of texts, to download these and to process them for digital text analysis.[5]

The problem is clear: too often digital humanities research still operates apart from the data. In digitising the data, too little attention is paid to the kinds of questions researchers ask and the facilities they need to be able to innovate. Add to this the problem that many data suppliers can shield their data out of legal or commercial concerns. A fundamentally open data environment, in which humanities scholars can adjust links according to their preferences will remain 'a higher honey' for a long time.

---

4   Sander Dekker, citation via National Library: http://www.kb.nl/nieuws/nieuwsarchief-2013/80000-boeken-tot-1870-online-beschikbaar-in-nieuwe-webdienst-kb.

5   I should here remark that the staff members of the National Library of the Netherlands (KB) have always been very willing to accommodate the requests of individual researchers.

That is why e-humanities' greatest challenges at this moment lie in the field of uniting on the one hand data digitisation and management and on the other hand data analysis and representation. Existing digital humanities data files should be connected and integrated through complex question-driven analysis possibilities. If this were possible, it would mean an important shift in the organisation of humanities research. It would mean a great advance for e-humanities, but there are also all kinds of drawbacks connected to it. During the last few decades humanities research has developed from individual to project-driven research. These projects are usually still relatively small and short-running. Thus the dynamic so characteristic for humanities research could still be guaranteed: but what will happen if we were to build ever larger integrated data collections and search engines on the basis of humanities research questions? Not only is there a danger that such a 'machinery' would suck a lot of power from the field, thus advancing a very limited number of research fields or questions, but the 'machinery' would eventually (if it were successful) determine the research agenda for a long period. This seems only desirable when there is a consensus within a certain part of the scholarly field about a number of core problems worthy of dominating all others. Such an approach was tried out with the top sectors, but for the humanities, this has yielded a limited field. What themes or questions offer enough mass and weight to drive the (re)organisation and linking of data files to be dynamically questioned?

Humanities 3.0 – 'critical reflection'

Luckily, we still have other paths to pursue. Not all data collections are anxiously kept under lock and key. Heritage institutions such as the *Rijksmuseum* and the *Tropenmuseum* also collaborate on research projects aimed at dynamic ways of making their collection accessible and searchable. The KB has made the data from the *Short Title Catalogue of the Netherlands* available to give historians and computer scientists the opportunity to explore to what extent this data file could facilitate research into the creative industries of the Dutch Republic. In my opinion, the projects organised with this aim constitute a good example of what I have in mind for e-humanities – the combination of several (disciplinary) approaches surrounding an integrated way of data digitisation and analysis. In this sort of 'smart or meaningful search' projects, the final component of e-humanities also plays its part, critical reflection on the methods used in the humanities.

What is meaningful search? Via *Open linked data* we have the opportunity to search 'semantically'. Static indicators can be combined with hermeneutic methods based on the content of the data, or with 'social media' methods for

which recommendations of third parties determine the search trajectory.[6] The results are not (as with for instance Google Books) dumped down in front of you, but classified according to content and the various contexts that can be linked to the data. The data can be arranged in ever new ways and therefore can be classified in different (historical) narratives.[7] In this way critical reflection on humanities methods and concepts is developed.

Increasingly, this last step is added to e-humanities projects. The employment of new methods forces an exact reporting and phasing of and a critical reflection on those methods. While new metadata are added to the collections through crowd tagging, simultaneously, research is done into the extent to which this data adds any value and what is for instance the difference between the knowledge of experts and that of the public at large. On the basis of e-humanities analyses, new ontologies of certain disciplines can be designed, and reflection can be developed on the concepts and methods used within these discipline. For instance, projects such as *Semantics of History* and *BiographyNet* contemplate questions such as 'What is an event?' In this, they try to honour the complexity and hybridity of the source material of the humanities. How can we constitute digital analysis environments in which different interpretations can exist next to each other?[8] Can we develop an evaluative system, to make comparisons between the amounts of sources used to underpin a certain interpretation, the number of perspectives taken into account, or those perspectives that were ignored (but used in other cases)? In my opinion, the great promise of e-humanities lies in the design of data files that can be searched in ever new ways, that yield insights that can be represented in relation to each other and subsequently can be introduced in the critical reflection on one's own method.

The advantage lies not only in the fact that hermeneutics are taken into account structurally in data digitisation, but also that such an instrument would give us insight in the kind of questions we pose. This could be the 'critical' aspect, which Rens Bod might consider for his humanities 3.0. Since research questions and results are supplied in the same system, we might also analyse what kind of questions yield what kind of results, and where new forms of questions develop. In this way, we might indeed usher in the 'Golden Letter Age'. ◀

6   See e.g. the CHIP-project of the *Rijksmuseum*: http://chip.win.tue.nl/home.html; and the reporting on this in L. Aroyo et al., 'Personalized Museum Experience: The Rijksmuseum Use Case, in: J. Trant and D. Bearman (eds.), *Museums and the Web 2007: Proceedings* (Toronto 2007): http://www.archimuse.com/mw2007/papers/aroyo/aroyo.html (1 July 2013). See also: M. Van Gendt, A. Isaac, L. van der Meij and S. Schloback,

'Semantic Web Techniques for Multiple Views on Heterogeneous Collections: A Case Study', in: *Proceeding of the ECDL* (Heidelberg 2006) volume 4172, 426-437.

7   This for instance is the point of departure for the Agora project: http://agora.cs.vu.nl/.

8   Semantics of History: http://www2.let.vu.nl/oz/cltl/semhis/. BiographyNet: http://www.biographynet.nl/.

Inger Leemans (1971) is Professor of Cultural History at VU University, where she helps to coordinate and develop the e-humanities educational and research programme. She has published on the history of early modern Dutch literature, about pornography, the (radical) Enlightenment and emotion history. Her interest in cultural economics has resulted in research into censorship, journalism, literary criticism, and the literary 'bubble' that accompanied the financial bubble of 1720. Inger Leemans is one of the founders and directors of ACCESS - The Amsterdam Centre for Cross-Disciplinary Emotion and Sensory Studies. Recent publications include: I. Leemans and G.J. Johannes, *Worm en Donder. Geschiedenis van de Nederlandse literatuur 1700-1800. De Republiek* (Amsterdam 2013); I. Leemans, 'Verse Weavers and Paper Traders: Speculation in the Theater', in: William N. Goetzmann, Catherine Labio, Geert Rouwenhorst and Timothy Young (eds.), *Great Mirror of Folly: Finance, Culture, and the Crash of 1720* (New Haven, CT 2013). Email: i.b.leemans@vu.nl.