



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

Digital Historical Research. Context, Concepts and the Need for Reflection

Piersma, H.; Ribbens, C.R.

published in

BMGN - Low Countries Historical Review
2013

document version

Publisher's PDF, also known as Version of record

document license

CC BY

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Piersma, H., & Ribbens, C. R. (2013). Digital Historical Research. Context, Concepts and the Need for Reflection. *BMGN - Low Countries Historical Review*, 128(4), 78-102. <http://www.bmgnlchr.nl/index.php/bmgm/article/view/9352>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl

Digital Historical Research

Context, Concepts and the Need for Reflection¹

HINKE PIERSMA AND KEES RIBBENS

In recent years digital research has been promoted as a promising innovation in humanities research. New answers to old questions and new questions, according to enthusiastic supporters. In this article the authors address the question what this desired methodological innovation means for historical science. Without denying the benefits of large digital sources, they feel the need to make some comments on the current practice and expectations regarding digital historical research.

Introduction

Methodological innovation in the humanities, in which close collaboration between humanities and computer science is aspired, is a priority, in particular at the Royal Netherlands Academy of Arts and Sciences (KNAW). This collaboration, described in 2009 as a 'promising cross-fertilisation', was even referred to in 2010 by Academy Director of Research and declared e-humanities supporter Theo Mulder as a 'great leap forward'. In 2012 historian and media scholar Frank van Vree, Dean and Full Professor at the University of Amsterdam, supported Mulder by declaring the new developments a 'revolutionary movement'.² This article aims to take a critical look at the feasibility of this vision of the future, by linking it to various relevant aspects of research policy as well as substantive historical research.

To outline the broader context, we will first address the intended innovation, using the policy outline memorandum *Contouren van een vernieuwings- en stimuleringsprogramma voor de Geesteswetenschappelijke Instituten van de KNAW* [Contours of an innovation and stimulation programme for the KNAW humanities institutes], that was presented by the KNAW in 2012, as a point of departure.³ Subsequently, we will describe two case studies regarding

the practice of digital historical research and the complexity of text mining (deriving relevant results from large, in this case not-born-digital text files). These are the CLARIN-NL projects *War in Parliament* and *Verrijkt Koninkrijk* [Enriched Kingdom], both with a duration of one year.⁴ *War in Parliament* addresses the full-text search of digitised data using (combinations of) keywords. *Verrijkt Koninkrijk* goes a step further. This project also involves the full-text search of digital data, but with the added challenge of tracking down significant concepts. This article will then present a brief outline of the necessary preconditions for historical research and a conclusion intended to invite further discussion.

The growing volume of digital data and the question how we as historians and scholars should handle this, is outside the scope of this article. This also holds true for the required role of academic education in making digital research into a success, and for the required research environment

- 1 The authors would like to acknowledge the anonymous reviewers, Gerben Zaagsma and the editors of *BMGN - Low Countries Historical Research*, for increasing our awareness of the diversity of views and expectations regarding the subject discussed here. Translation by Maggie Oattes.
- 2 Symposium KNAW June 2009 entitled 'Bridging the Gap between the Humanities and the Computational Sciences'; 'Aanloop voor Grote Sprong in Humaniora' [Build-up to Great Leap in the Humanities], *e-data&research* 5:1 (June 2010), http://www.edata.nl/0501_010610/pdf/Aanloop_voor_Grote_Sprong_in_humaniora.pdf (11 July 2013); Kick-off meeting KNAW, October 2012. Cf. Stanley Fish, who in his blog of 9 January 2012 on digital humanities wrote: 'The Rhetoric of these Statements [Fish lists several supporters here] (which could easily be multiplied) is not one of reform, but of revolution', <http://opinionator.blogs.nytimes.com/2012/01/09/the-digital-humanities-and-the-transcending-of-mortality/> (11 July 2013). Recently, Van Vree was much more critical. He stated that the humanities are in danger because grant providers seems to prefer quantifiable research; the type of research that uses methods similar to the sciences. See: *De Groene Amsterdammer* 137:44 (31 October 2013) 49.
- 3 *Contouren van een vernieuwings- en stimuleringsprogramma voor de Geesteswetenschappelijke Instituten van de KNAW* (2012): <http://www.knaw.nl/Content/knaw/publicaties/pdf/20121019.pdf> (11 July 2013).
- 4 CLARIN stands for Common Language Resources and Technology Infrastructure. CLARIN-NL is the Dutch division of a European organisation that provides subsidies for projects with a duration of six to twelve months. The aim was to convert source material into an international standard and to develop tools to enable and facilitate working with these standards within a European infrastructure.



This parody on the cult movie *Pulp Fiction* has actor Samuel L. Jackson referring to both the *buzzword*-character of digital humanities and the critical reactions to the phenomenon by colleagues.

Probable source: Website *LA Review of Books*, 28 October 2012. As used in Stephen Marche, 'Literature is Not Data: Against the Digital Humanities', <http://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities>.

and infrastructure.⁵ This contribution aims to reflect on current practice and policy of historical research and on the challenges of digital historical research in light of the programmatic desires of the KNAW. We will focus on a specific aspect of digital historical research, namely tracing and analysing significant concepts and the context in which this historical content is located, appropriately described as ‘enhanced IT’. Emphasis is on ‘the special and complex nature of historical data processing in contrast with computer applications in, for example, business and hard sciences’.⁶

In this article we continue the discussion that was initiated in the Netherlands several decades ago by scholars like Onno Boonstra, Leen Breure and Peter Doorn. They are among the pioneers in the field of IT in humanities research. They were active in, among other things, the *Vereniging voor Geschiedenis en Informatica* [Belgian-Dutch Association for History and Computing], founded in 1987 in which the earliest forms of digital history research were institutionalised. When these pioneers introduced the use of computers in historical research and studied its possibilities, no one foresaw that two decades later digital text processing, searching for information via the Internet and a means of communication like email, would be part of the standard tools of historians. Indeed, they are so established today that their utilisation is not really considered part of digital practice of history. On the other hand, when Boonstra, Breure and Doorn took stock of the developments in their field in 2004, they had to conclude that ‘the traditional field of historical science’ was insufficiently aware of the possibilities of ‘history and computing’.⁷

Another disturbing phenomenon these three scholars identified, was the position of ‘history and computing’ vis-à-vis information science in general.⁸ This meant the relatively early detection of the major importance of fruitful communication between historians and digital experts. This was linked to a strong plea for a digital infrastructure to provide overview and structure for the increasing amount of individual projects. The results of the various activities of CLARIN and DARIAH seem to be based on a similar opinion.⁹

5 For an illuminating report on infrastructure and education see: ‘Our Cultural Commonwealth: The Report of the American Council of Learned Societies’ Commission on Cyberinfrastructure for Humanities and Social Sciences’ (2006), <http://www.acls.org/cyberinfrastructure/ourculturalcommonwealth.pdf> (11 July 2013).

6 Onno Boonstra, Leen Breure and Peter Doorn, *Past, Present and Future of Historical Information Science* (Second edition; Amsterdam 2006) 25-83. See <http://www.dans.knaw.nl/sites/default/files/file/publicaties/Past-present.pdf>.

7 *Ibid.*

8 *Ibid.*, 88-89.

9 DARIAH stands for Digital Research Infrastructure for the Arts and Humanities and is a result of the Roadmap for the humanities of ASFRI (European Strategy Forum for Research Infrastructures).

Whether this also automatically intensifies and improves the relationship between a varied group of historians and information experts remains to be seen, especially as the field of digital historical research in recent years has expanded considerably, which does not automatically improve transparency for outsiders. In our opinion the relationship will benefit from a gradual process of increasing mutual familiarity with the variety of research questions and methods, which must be continuously encouraged – including the possibility of debate, – rather than from the perhaps overly optimistic desire to elicit programmatic revolutions. This is the context that is at the basis of the somewhat polemic nature of this article.

Silent ideology

In the abovementioned plan (*Contourennota*) presented by the KNAW in 2012 five humanities institutes (Huygens ING, International Institute of Social History (IISG), Royal Netherlands Institute of Southeast Asian and Caribbean Studies (KITLV), Meertens Institute, and NIOD Institute for War, Holocaust and Genocide Studies) that come under the Academy, took central stage. Shortly before the publication of the *Contourennota* these institutes were evaluated. Although the Academy Board took ‘the results of the evaluation seriously’, they ‘did not agree with the commission’s opinion that information technology, while being of importance, is only *one* of the available methods of conducting research’. The KNAW views ‘the technological development as an irreversible process that will profoundly affect the methods applied in the humanities’.¹⁰

Few people will deny that technological development progresses, or that this has an impact on the humanities. It remains to be seen, however, whether the methodological innovation sought after by the KNAW can actually be realised without a critical discussion about the applicability of the digital tools in historical research. Although there are many digital historical projects, a thorough evaluation of their results and added value by a wide body of historians and other scholars is lacking. Evaluation is even more urgent in view of the frequently implicit claims creating the impression that technological progress also implies a new historical-scientific paradigm. Such a paradigm would be based on various assumptions, namely that this type of research is quantifiable to a high degree and that large amounts of diverse sources are suitable to be used in this approach, that (un)suspected interrelations between

10 *Contouren van een vernieuwings- en stimulerings-programma voor de Geesteswetenschappelijke Instituten van de KNAW* (2012). For the evaluation report see: <http://www.knaw.nl/Pages/DEF/33/501>.

[bGFuZz10TA.html](http://www.knaw.nl/Pages/DEF/33/501). The involvement of the KNAW goes back longer and has been visible from 2006 in, among other projects, the establishment of the former Virtual Knowledge Studio and the current e-Humanities Group.

a variety of data can be demonstrated on a large (and perhaps international and longitudinal) level, and that the answers to existing as well as new questions obtained in this way will have a more solid basis than currently used forms of interpretation that are considered more subjective. One can sometimes even detect the ambition that in time digital historical study can replace the analog variant(s).

The appeal of the large-scale variant of digital historical research rests strongly on two assumptions. The first is that it would be relatively simple to conduct research much quicker and in principle also on a larger scale, because computers are able to process large quantities of data and they can do it rapidly. The second assumption is that complex historical questions can also be answered with the aid of computer tools. By means of machine-learning techniques computers could *learn* to trace the desired information.¹¹ Using sentiment mining techniques to trace and interpret subjective concepts, the emotional meaning of the content could also be retrieved.¹²

As yet this proves to be quite a challenge in practice, partly because the way in which computers work is not automatically compatible with the way historians work. Furthermore, there is the question of the degree of uniformity in ‘the’ historical method. As a consequence of the aim of historians to basically cover all of human history, and knowledge being extracted from a wide array of sources, there is a large variety of methods that – depending on the research question and the nature of the source material – are used to gain insight into very different phenomena. Oral history is a very different way of gathering and interpreting sources than analysing medieval charters or interpreting ancient Greek potshards. Each source and each method of study inspires new questions, as is the case with digital historical research. But sources and methods also have their limitations.

The similarity between the different ways of studying sources is not, however, only found in their limitations. Despite the methodological diversity a commonality can be observed, for example in the way in which knowledge

11 In *War in Parliament* we intended to create a classifier capable of recognising references to World War II using memory based learning, in order to generate all relevant World War II-references in the period under research. In this way we wanted to avoid depending (solely) on keywords. To create this classifier we needed training materials, which were supplied by three researchers from NIOD. The next step was to have the computer scientist translate these results into the aimed classifier. He concluded that the classification of World War II-references was

more difficult than initially thought. One of the difficulties was that the Second World War is not a ‘natural’ topic in Parliament.

12 More information on the background and complexity of sentiment analysis and possible solutions, see: Isa Maks and Piek Vossen, ‘Modeling Attitude, Polarity and Subjectivity in Wordnet’, in: P. Bhattacharya, C. Fellbaum and P. Vossen (eds.), *Proceedings of the 5th Global WordNet Conference (GWC 2010), Mumbai, India* (Mumbai 2010).

about the past is collected, organised and interpreted. Commonality can also be found regarding the importance of context, the way different sources and perspectives are combined and in the fact that different layers are distinguished in historical processes. These aspects are all complex characteristics of ‘the’ historical method, based partly on interpretation, which a digital approach cannot match in the present situation. Before examining the practice of digital historical research in the case studies, we will first address the methodological innovation that, inspired by the possibilities of advanced technology, is considered necessary in the humanities.

Technology is generally presented as neutral, and only its applications, not the underlying assumptions, are subject of debate, as illustrated by the 2012 volume *Stille ideologie* [Silent ideology]. Silent ideology is defined in the contribution of Meike Bokhorst (affiliated with the Scientific Council for Government Policy (WRR) and Tilburg University) as an ideology that is not designated as such because it ‘is (too much) presumed to be matter of course’. The protagonists deny there is an ideological nature by ‘presenting the ideas as factual, objective, value-neutral, scientific, or universally accepted’.¹³

A comparison with the above quoted KNAW statement that the technological development is irreversible and will affect humanities research, springs to mind. Nobody will deny this statement as such, but the implied normative message is problematic. This message is that technological development makes methodological innovation unavoidable if humanities scholars do not want to miss the boat. The (apparent) contrast between the traditional scholars and those who (dare to) master the new techniques also fits in this discourse. The policy plans generally do not discuss these assumptions, nor the desire to make the humanities as ‘hard’ as the sciences. However, one can see it in the importance supporters of the e-humanities attach to *evidence* – which is assumed to be checked more easily in the case of digital source material and digital techniques – unlike the *interpretation* of historical phenomena crucial in historical publications. The similarity with debates in earlier decades is striking.

In 1959 Z. R. Dittrich and A. M. van der Woude published an article entitled ‘De geschiedenis op de tweesprong’ [History at the crossroads]. They felt that historical science was in a crisis and that historians were operating on the scientific fringes; all they occupied themselves with were insignificant anecdotes in the past. Dittrich and Van der Woude called for a scientification of historical science that was to be realised in collaboration with the social

13 Meike Bokhorst, ‘Europese integratie. Laveren tussen stille ideologie en politiek pragmatisme’ [European integration: Navigating between silent ideology and political pragmatism], in: Cor van

Montfort, Ank Michels and Wouter van Doorn (eds.), *Stille ideologie. Onderstromen in beleid en bestuur* [Silent ideology: Undercurrents in policy and administration] (The Hague 2012) 197-203, 200-201.

sciences – their article was therefore published in the sociology journal *Mens en Maatschappij* [Man and Society]. Not the uniqueness of historical events was to be the central theme, but recognition of patterns.¹⁴ In the elaborated KNAW plans pattern recognition and the discovery of regularities are also important spearheads.¹⁵

More than a decade after the article by Dittrich and Van der Woude, prominent representative of the French *Annales* school E. Le Roy Ladurie stated that: ‘history that is not quantifiable cannot claim to be scientific’. He was furthermore of the opinion that the future historian ‘will be a programmer or he will be nothing’.¹⁶ However, at the peak of Le Roy Ladurie’s popularity – who as a microhistorian may have unintentionally stressed how much the study of small-scale textual sources was a part of the practice of history – narrative history was on the rise (again). This advance continues to this day, witness the many historical bestsellers. There seems to be a wave-like motion in which the scientific character of historical science is measured and found wanting from time to time. That this is happening again in an age in which ‘market forces’ prevail, scientific output is increasingly translated into economic gain, and research policy is cut up into top sectors, should not come as a surprise. This political agenda should not, however, be an obstacle to further reflection on the significance of digital historical research or to exploring its possibilities and impossibilities. Based on this consideration we will now describe the two projects carried out within NIOD: *War in Parliament* and *Verrijkt Koninkrijk*.

14 Z.R. Dittrich and A.M. van der Woude, ‘De geschiedenis op de tweesprong’ [History at the crossroads], *Mens en Maatschappij* 34 (1959) 361-380. Quoted in: Marjolein ‘t Hart, ‘Verzin een list, jonge vriend. Over generaties en de toekomst van het vak’ [Find a way, young friend: On generations and the future of the profession], *Tijdschrift voor Sociale en Economische Geschiedenis* 5:2 (2008), 138-150, 139-140. Cf.: Jo Tollebeek, *De toga van Fruin. Denken over geschiedenis in Nederland sinds 1860* [Fruin’s gown: Thoughts on history in the Netherlands since 1860] (Second edition; Amsterdam 1996) 388.

15 The KNAW Board in its reply of 7 May 2013 to the directors of Huygens ING, IISG, KITLV, Meertens Institute, NIAS, NIOD and DANS, regarding Humanities Centre, expresses the hope that ‘patterns and regularities may be found that show

new links between economic, political, cultural and religious changes and that (alternative) explanations may be provided for historical developments in culture, politics and economy’, which corresponds with the KNAW’s *Strategische Agenda 2010-2015* [Strategy Agenda 2010-2015], 24.

16 Emmanuel Le Roy Ladurie, ‘L’historien et l’ordinateur’, in: *Le territoire de l’historien* (Paris 1973) 11-14, 11. Quoted in: Lawrence Stone, ‘Revival of the Narrative: Reflections on a New Old History’, *Past & Present* 85 (November 1979) 3-24, 5 and 13. Le Roy Ladurie was spokesman of the French *Annales* school, but ironically became famous for his ‘narrative’ work *Montaillou*, published in 1975, about life and death, work and sex, religion and customs in the early fourteenth century.



Hendrik Koekoek, leader of the *Boerenpartij*, and chairman Frans-Jozef van Thiel during the discussion of the problems within the *Boerenpartij* in the Dutch House of Representatives, 25 June 1968.

ANP Photo.

Historians are traditionally explorers of sources, digging around in archives for days to look for relevant material. During the research period, as they become familiar with the material and the historical process, the research question they formulate can be modified, adapted, or even changed completely, depending on where the material leads them. However, a digital approach of the research material requires the early formalisation of the research question. It cannot be denied that this is challenging to historians, a phenomenon that is described by Joris van Zundert, who has a background in Dutch literature and linguistics, as an ‘almost hostile act’.¹⁷ Van Zundert has a point here. First of all because the nature of the material historians work with is often grim and multifaceted, which frequently means the research question is gradually specified; and secondly, because in historical practice, research questions and hypotheses formulated by historians are often not easy to quantify and to accommodate in models, a fact recognised by Van Zundert.

Some studies, however, seem to be very suitable for a digital approach because formalisation of the research questions should be relatively simple. This was true, for example, for the *War in Parliament* project, in which we investigated how World War II was used in the postwar political debate in the Netherlands.¹⁸

To this day references to fascism and national-socialism are considered a tested means of stigmatizing fellow-politicians, and they could even result in the effective exclusion of political parties.¹⁹ By referring to its resemblance to fascism an extreme right-wing party like the Dutch *Centrumdemocraten*, was marginalized and effectively isolated in the 1980s. In the more recent past, in 2002 – and outside the parliament – Thom de Graaf of the social-liberal Democrats 66 (D66) quoted Anne Frank to defame Pim Fortuyn.²⁰

War in Parliament, for which we utilised the digitised *Handelingen der Staten-Generaal* (Dutch Hansard) that are searchable as full-text, was intended to systematically map all references to World War II, to find an answer to our hypothesis that the utilisation of the Second World War as a political argument increases in times of political transition that are characterized by

17 Joris van Zundert et al., ‘Cultures of Formalization: Towards an Encounter between Humanities and Computing’ (March 2010), <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxkbWJlcnJ5fGd4OmUxNT3NGU5MDdmMzBk> (11 July 2013).

18 The project *War in Parliament* is funded by CLARIN-NL and is conducted in cooperation with the University of Amsterdam (UvA) (Lars

Buitinck, Johan van Doornik, Maarten Marx) and DANS (Marjan Grootveld).

19 References to World War II can also be used to legitimize a particular political position, for example to condemn the policy of *Apartheid* in South Africa.

20 See: <http://www.pimfortuyn.com/asp/default.asp?t=show&id=1392> (11 July 2013).

the emergence of new parties, threat of war, et cetera. In order to harvest the relevant results we used the Boolean search operators AND, OR and NOT. We then formulated our queries thinking we could use the NIOD subject index. Although this index is definitely suitable to open up the collection of books and archives of NIOD, we could not use it for the *Handelingen der Staten-Generaal*, as the example of the keyword ‘politie’ [police] can clarify. Naturally we were not interested in reorganisations of the police force, more uniforms on the streets, or any other of the post war police-related issues.

An additional problem turned out to be the disappearance of some words from the vocabulary, such as ‘politieke delinquenten’ [political delinquents] to refer to former NSB members, or the emergence of new terms like Holocaust after 1979. In practice it proved to be quite a challenge to adequately generate references to World War II with a simple keyword search. We already mentioned the example of ‘police’, but domain-specific keywords were also problematical, for example ‘bezetting’ [occupation]. We were looking for ‘occupation’ as in the German occupation of the Netherlands from 1940 to 1945, but also generated the occupation of hospital beds, ministries and the student protest at the *Maagdenhuis*. The keyword ‘World War II’ created another type of noise, as it was also used as a caesura in time, in combination with ‘since’ ‘after’ and ‘before’, while ‘Hitler’ was often mentioned in a list of detested statesmen. In itself an interesting phenomenon, but not relevant to our research question. And then there were the more or less hidden references that were found by accident, for example when the Second World War was referred to as the ‘darkest period’.

David C. Blair, specialist in computer and information systems, pointed out the complications occurring in ‘full-text document-retrieval’ as early as 1985, and he spoke of ‘output overload’ on the one hand, and problems related to ‘meaning’ on the other. In his article he also addressed the formulation of queries (in particular the (im)possibility of predicting every relevant keyword that fits in such a query), and the pros and cons of combining keywords. With regard to the latter: in *War in Parliament* not combining keywords resulted in the abovementioned output overload (‘World War II’ generates a huge amount of non-relevant hits), but narrowing the number of hits by using combinations with other keywords (World War II AND ‘persecution’ AND et cetera) learned that we missed relevant passages.²¹

Developments in the field of computer technology did not stop after 1985, a fact which did not escape Blair; he analysed successful and less successful projects regularly. However, the difficulties he observed in

21 David C. Blair and M.E. Maron, ‘An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System’, *Computing Practices* 28:3

(March 1985) 289-299; See also: Don R. Swanson, ‘Historical Note: Information Retrieval and the Future of an Illusion’, *Journal of the American Society for Information Science* 39:2 (1988) 92-98.

1985 have not been solved, despite the progress in technology. Today there is a considerable amount of literature on the question how to improve the harvesting of relevant hits and where the problems are located. An illustration can be found in the article by linguist Robert Krovetz and computer scientist W. Bruce Croft on lexical ambiguity, in which the authors distinguish syntactic and semantic ambiguity.²² Syntactic ambiguity means that a word can be a verb as well as a noun (for example, the Dutch word ‘kampen’ can refer to ‘concentriekampen’ [concentration camps] as well as ‘te kampen hebben met’ [to struggle with]). Semantically ambiguous words have a different meaning or, as in the case of ‘occupation’, a different connotation. Although efforts are made to find solutions – see for example the work of Dutch linguist Piek Vossen²³ – the ambiguous nature of the information we searched for in the *War in Parliament* project was still a real problem.

To get a result in the available time, we therefore had to modify our initial ambitions in the project (finding all references to World War II). We opted for a case study in which limitation was not sought in the war-related keywords (‘World War II AND fascism AND et cetera), but by focusing our search on one political party. We chose the *Boerenpartij* [Farmer’s Party], represented in the States General in the period 1963–1981, because it was the first right-extremist party to enter the political stage after World War II. Because of the party’s right-wing nature we expected to find many references to the war years, an expectation also suggested by the fact that there was some turmoil with regard to the *Boerenpartij* after Hendrik Adams was installed in the Senate [*Eerste Kamer*] in 1966. Adams had collaborated with the Nazis during World War II and had been tried and punished after the war. In 1966 he obtained a seat in the Senate, which resulted in much commotion and forced Adams to withdraw (the Adams affair).²⁴

Our research question was: How often and in what way was the *Boerenpartij* linked with fascism and national socialism in parliamentary debate (and therefore associated with the label ‘fout’ (‘wrong’)? We defined ‘fout’

22 Robert Krovetz and W. Bruce Croft, ‘Lexical Ambiguity and information Retrieval’, *ACM Trans. Inf. Syst.* 10:2 (1992) 115–141, <http://www.lexicalresearch.com/tois-lex-ambiguity.pdf> (11 July 2013).

23 Attila Göring and Piek Vossen, ‘Computer Assisted Semantic Annotation in the DutchSemCor Project’, http://hmk.ffzg.hr/bibl/lrec2010/pdf/269_Paper.pdf (10 November 2013).

24 For the Adams affair see: Koen Vossen, *De andere jaren zestig. De opkomst van de Boerenpartij (1963–1967)* [The other sixties: Emergence of the Boerenpartij (1963–1967)], <http://dnpp.eldoc.ub.rug.nl/FILES/root/jb-dnpp/jbo4/vossen.pdf> (11 July 2013). See also: Hinke Piersma, Ismee Tames, Lars Buitinck, Johan van Doornik, Maarten Marx, ‘War in Parliament: What a Digital Approach can add to the Study of Parliamentary History’, *Digital Humanities Quarterly* (intended publishing date December 2013).

based on seventeen keywords and we defined the *Boerenpartij* (the party and its representatives). Our query looked like this:

(fascis* OR nsb OR "politiek delinquent" OR "politieke delinquenten" or collaborat* OR "nationaal socialistisch" OR "nationaal socialistische" OR antisemitis* or oorlo** OR Hitler or Mussert or Roskam or Boerenleider OR Jeugdstor* OR NSK* OR Waffen-S* OR Landstand) AND (Boerenpartij or Adams OR Koekoek OR Voogd or Brake or Harmsen or Harselaar or Bossche or Koning or Kronenburg OR Leffertstra OR Nuijens OR Verlaan)

This combined query resulted in 179 hits (hits were harvested at the paragraph level). After removing all non-relevant hits we were eventually left with twenty relevant hits, with a peak of eight hits in 1966.

Based on the analysis of these hits we were able to conclude that the political opponents of the *Boerenpartij* did repeatedly make a connection between the *Boerenpartij* and ‘fout’. However, until the Adams affair the references were related to a ‘wrong’ mentality, not a ‘wrong’ ideology; the *Boerenpartij* was not ‘wrong’, but acted ‘wrong’. Ideology was an element in the debate about Hendrik Adams, but, again, not very clearly; more than whether the *Boerenpartij* was fascist in nature, the question was whether a former supporter of fascism, even one who had undergone his postwar punishment, could bear political responsibility. After 1966 there were very few links between the *Boerenpartij* and ‘fout’. This may perhaps be explained by the fact that after the Adams affair the *Boerenpartij* was pushed into the political fringe as a result of rows and rifts, and the reference to ‘wrong’ had done its job: the *Boerenpartij* had been rendered politically harmless.

What is the significance of this case study for digital historical research? First of all we must stress the importance of the conclusion that after 1966 almost no links were made between the *Boerenpartij* and ‘wrong’. Without full-text, digitally searchable data it would not have been possible to make such a statement. For unless someone makes it his life’s work, it would require far too much time to examine by hand every debate between 1966 and 1981 in which the *Boerenpartij* played a role. At the same time we must conclude that, in order to get beyond the observation that there was indeed a link between a specific party and ‘wrong’, qualitative research remains necessary to 1) eliminate the noise, and 2) be able to make statements about significance and context. The required qualitative research implies that it is important to keep the number of hits limited and practicable – a problem already discussed earlier. In this case study linking ‘fout’ to one particular political party proved successful. However, without this link (‘fout’ AND ‘geen selectie op partij’) [‘wrong’ AND ‘no selection on party’], we ended up with the output overload described by Blair.²⁵

Verrijkt Koninkrijk (Enriched Kingdom)

Our second example is the research project *Verrijkt Koninkrijk* related to the famous work by dr. L. de Jong entitled *Het Koninkrijk der Nederlanden in de Tweede Wereldoorlog* [The Kingdom of the Netherlands during World War II].²⁶ NIOD prepared a digital version of the scientific edition of this series in 2011, which was further unlocked with the financial support of CLARIN-NL.²⁷ The goal was to make the complete text of De Jong more accessible. Apart from the systematic analysis of the structure of the publication, this was done especially by setting up a user-friendly and quick search engine and creating links to other databases (in particular Wikipedia), thus contextualising the information in the current knowledge landscape. In addition the project also included a substantive, historical scientific question aimed at retrieving how De Jong dealt with the concept of ‘pillarisation’ in his work.

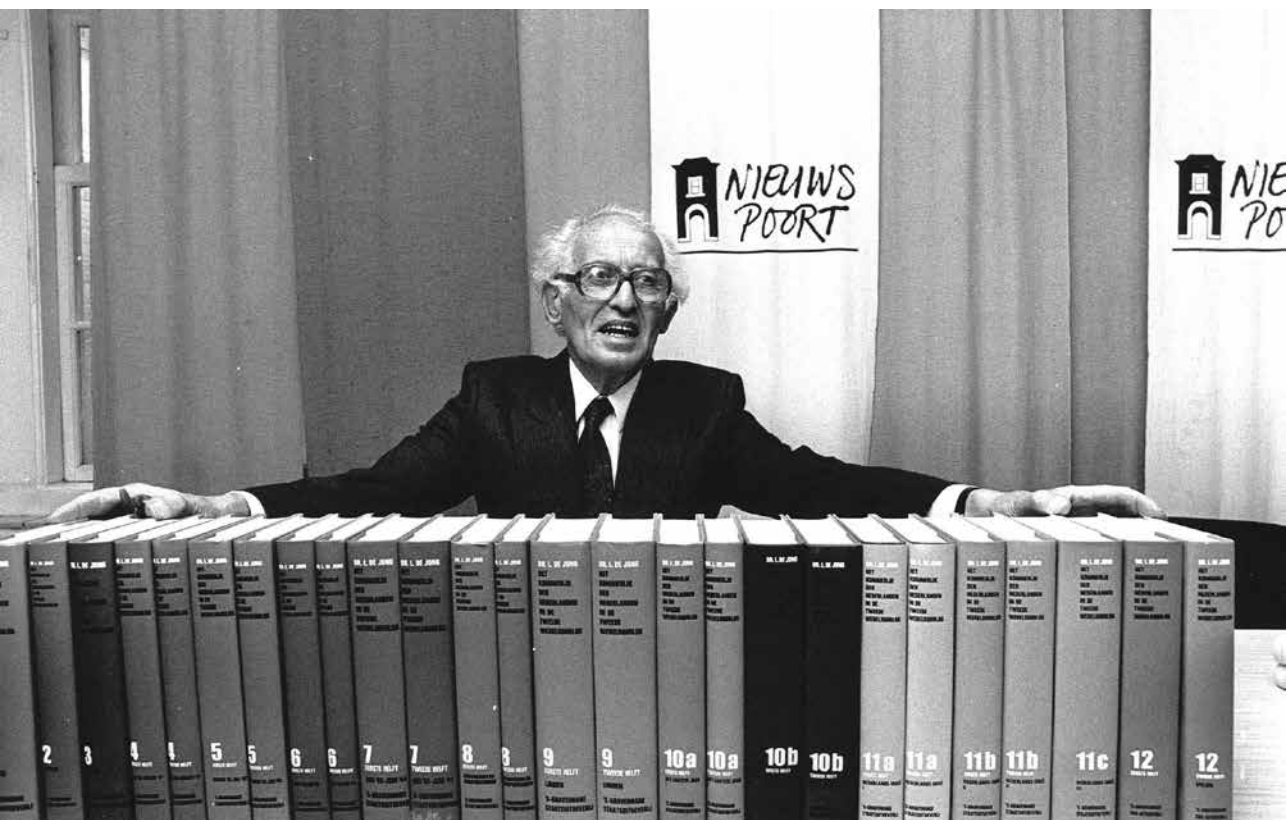
The answer to this question proved complex. The full-text search for the keywords ‘zuil’ [pillar] and ‘verzuiling’ [pillarisation] or plural forms of these words, generated few hits. It quickly became clear that De Jong spoke more frequently about one or more sections of the national community (‘volksdelen’). This emphasised the importance of incorporating related concepts and other terms that could be associated with the research subject in the search strategy. Several of these concepts could be formulated in advance based on general historical knowledge of both the period described and the period in which De Jong’s work was created. Other alternative terms were found in the proximity of the predefined terms. The fact that not all

25 W.C. Lee and E.A. Fox published an article in 1988 ‘Experimental Comparison of Schemes for Interpreting Boolean Queries’, *Technical Report TR-88-27, Computer Science, Virginia Polytechnic Institute and State University*. In this article they prove that the use of AND-queries results in high precision (relevant hits) at the expense of recall (large number of non-relevant hits), and OR-queries do the opposite: high recall at the expense of precision. Perhaps ‘statistical documentation retrieval’-methods (or: extended Boolean information retrieval) can be useful by focusing on assigning more weight to relevant keywords. See for example: David D. Lewis and Karen Sparck Jones, ‘Natural Language Processing for Information Retrieval’ (July 1993). See also: Edgar Meij, Dolf Trieschnigg, Maarten de Rijke and Wessel Kraaij, ‘Conceptual Language Models

for Domain-specific Retrieval’, *Information Processing and Management* 46:4 (July 2010) 448-469; C.D. Manning, P. Raghavan and H. Schütze, *Introduction to Information Retrieval* (Cambridge 2008). For an informative case studie see: Martha van den Hoven, ‘Strikes that never happened: Text mining in Historical Data’, *ILK Technical Report Series 10-05* (August 2010).

26 L. de Jong, *Het Koninkrijk der Nederlanden in de Tweede Wereldoorlog* (14 volumes; Leiden, The Hague 1969-1994) [The Kingdom of the Netherlands during World War II].

27 Opening up and making the volumes available occurred in collaboration with information specialists from NIOD (Tim Veken), UvA (Johan van Doornik, Lars Buitinck, Maarten Marx), VU (Victor de Boer), DANS (Marjan Grootveld) and the Meertens Institute (Marc Kemps Snijders).



▲
Historian Loe de Jong (1914-2005) at the presentation in 1988 of volume 12 of his life's work *The Kingdom of the Netherlands during World War II* – a time in which the growth of the use of computers in historical research was scarcely conceivable.
ANP Photo.

terms have exactly the same meaning and they sometimes reflect a different, somewhat instinctive evaluation or otherwise additional connotation, had to be taken into account in the further examination (by hand) of the hits and their context.

Given that De Jong (based in all likelihood on ambivalent thoughts about his own cross-pillarised position) proved to be even less taken with a concept like pillarisation than expected – although this was undeniably a strong phenomenon in the society he describes – it was necessary to further specify the concept of pillarisation. The first step was the search for mention of the most recognisable and high-profile pillars, the Catholics and Protestants, and the relatively less strongly organised pillars of Social-Democrats and Liberals. Adjectives and nouns based on these four terms were relatively easy to find. One thing that became clear was that the quality of the optical character recognition (OCR), that converts the printed text of the books into machine-readable text, frequently leaves much to be desired; the recognition rate was suboptimal.²⁸ Despite the apparently high print quality of the relatively recent books, a particular letter or letter combination was frequently mistaken for another. Especially when the mistakes seem to occur inconsistently, this means an impoverishment of the available information for the researcher. Incidentally, the desires of the historians in this respect seem to surpass the expectations of experienced computer scientists.

The second step in our search for pillarisation was to not look exclusively at the explicit occurrence of the names of the pillars²⁹ but also at the organisations and individuals associated with these pillars. To this end we drew up a list for each pillar of some ten names of persons and institutional designations (political organisations, trade unions, resistance groups and newspapers of a particular religious or political affiliation, et cetera): the so-called ‘named entities’. Where these names occurred in the text, whether in combination with an explicit mention of the pillar or not, we spoke of a pillarisation passage. Computer scientists tend to link these passages to one specific pillar as much as possible, whereas the historian is more used to recognising the frequently more ambiguous nature of such passages that regularly refer to more than one pillar (representative). This illustrates the need for further coordination and further familiarization with each other’s basic principles.

28 Victor de Boer, Johan van Doornik, Lars Buitinck, Maarten Marx, Tim Veken and Kees Ribbens, ‘Verrijkt Koninkrijk: Linking a Historiographical Text to the Web of Data’ (March 2013), http://cls.ru.nl/~ihendrickx/Posters_ehum/14_deboerehuman.pdf (10 November 2013).

29 In addition to mentions of the four pillars, the project also examines the position and representation of three other communities with a prominent role during the German occupation, namely Jews, national-socialists, and communists.

Terms like collaboration and interdisciplinarity are important in discussions about the digital humanities. Prominent Swedish linguist Patrik Svensson writes that ‘the digital humanities can be seen as a fractioned (not homogeneous), collaborative (not coerced) trading zone with a continuous development of interlanguage’.³⁰ Svensson does not, however, offer an opinion on what exactly the scientific outcome of all this should be. In the execution of *War in Parliament* and *Verrijkt Koninkrijk* there proved to be tension between the goals both disciplines set and the interests they had. Earlier we presented the example of the interpretation of certain passages where the historian sees a more complex reality. In addition computer specialists want to make *new* tools that *work*. It is not deemed absolutely necessary that the results generated by these tools – and that are definitely improved during the course of the research project thanks to the further specification of the algorithms used – are the most useful results or provide relevant answers to the research question, not least because development and experimentation are important goals in themselves. One specific problem occurs when it becomes clear that – similar to the problems in *War in Parliament* with terms like ‘occupation’ – particular words and names have more than one meaning, and are therefore not always relevant from the perspective of a specific question. The human eye can usually quickly determine relevance from the context. However, this is not always easily translated into instructions to improve the search engine.

Due to De Jong’s suspected ambivalence regarding the concept of pillarisation – which implies that from his own hybrid position as a social-democratic Jewish Dutchman he must have experienced more than once that such an absolute and seemingly straightforward societal categorisation was inadequate and unsatisfactory – it was deemed important to also look at the how the author *evaluated* the general phenomenon of pillarisation as well as specific pillars or sections of the community. We hoped to be able to accomplish this by using the abovementioned sentiment mining techniques. Unfortunately the available Dutch language corpora with which emotionally charged terms could be identified and evaluated, proved to contain mostly extreme terms from modern colloquial speech, including terms of abuse that were (almost) absent from the accessible but academically formulated works of De Jong. There was, however, a test version of a programme, the so-called sentiment analyser developed by Professor dr. Piek Vossen of the vU University Amsterdam (vU), which could also trace less obvious terms. We consulted this version under development in October 2012, but unfortunately it was not yet refined enough: a word like ‘very’ [*erg*] was classified as negative, a

30 P. Svensson, ‘The Digital Humanities as a Humanities Project’, *Digital Humanities Quarterly* 3:3 (2009) 42-60, 53.

word like ‘good’ [*goed*] as positive; they then cancelled each other out and the combination ‘very good’ was evaluated as neutral.³¹

The above examples illustrate that text mining in historical research using these new tools – that are still under development – requires a willingness on the part of the historian to do his best to make his reasoning explicit, when possible dividing it into manageable elements or building blocks and always translating it for the more technically oriented information experts. The developments both parties can realise in co-operation generally consist of small steps.

This is an important reason to assess the feasibility of the ambitions in interdisciplinary projects in which historians and information experts join forces. These ‘lessons learned’ have been taken to heart in a proposal involving a follow-up study on *War in Parliament* that was recently submitted to the Netherlands Organisation for Scientific Research (NWO), and builds on the experiences gained in both case studies. An extensive discussion of the complexity of the research question eventually led to a decision to further improve and expand the (traditional) Boolean search method. In today’s state of the art this method offers the best ‘guarantee’ for results, as it does in patent research.³² Although we utilise methods (full-text keyword search) that are already considered ‘traditional’ in computer science, we deliberately choose to avoid high-tech experiments with sentiment mining or sophisticated automatic machine learning techniques. Instead we apply a controlled digital approach without denying the historical method.³³

Mass digitisation

The examples of digital historical research presented here, were conducted on the basis of the existing digital version of the *Handelingen der Staten-Generaal* and the works of Loe de Jong that were digitised just before the start of the

31 The initial version was found on: <http://debvisdic.let.vu.nl:8080/SentimentWebClient/pages/textarea.jsp>. For an overview from 2011 see Isa Maks and Piek Vossen, ‘A Verb Lexicon Model for Deep Sentiment Analysis and Opinion Mining Applications’, in: A. Balahur, E. Boldrini, A. Montoyo, P. Martinez-Barco (eds.), *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA 2011)* (Portland, OR 2011) 10-18. Progress is expected from the Opener project: <http://www.opener-project.org/project.php?mod=1> (11 July 2013).

32 R. Bache and L. Azzopardi, ‘Improving Access to Large Patent Corpora’, in: *Transactions on Large-Scale Data and Knowledge-Centered Systems II* (Berlin, Heidelberg 2010) 103-121.

33 ‘Mobilising emotions in parliament: The Second World War as a benchmark of political morality in post-war political discourse in the Netherlands, the United Kingdom, and the Federal Republic of Germany’, research proposal submitted to NWO Free Competition, May 2013.

project *Verrijkt Koninkrijk*. Digital source material is a prerequisite for digital historical research. This may sound like we are stating the obvious, but we are not; in practice much of the material is still in analog form, and can therefore not be searched using text mining techniques. If the large-scale version of the methodological innovation is to be successful, mass digitisation of paper collections is inevitable. This is a step that is in danger of being skipped over because there is so much focus on the development of tools – which is considered more innovative – and on visions of a more sophisticated historical practice. The *Contourennota* of the KNAW mentioned earlier, speaks optimistically of a ‘breakthrough of mass digitisation’ in the collection-holding KNAW Institutes, but this breakthrough is not as clear in practice. At NIOD, for example, less than three per cent of the total archival materials has been digitised. In other collection holding KNAW institutes, such as IISG and KITLV, only a fraction of the material is available in digital form. Although great strides have been made in the digital opening up of newspapers and magazines, it is obvious that this source concerns only one aspect of the historical source material, namely material relating to the public domain.

‘Mass digitisation: not sexy, but very useful’, Ewoud Sanders tweeted following a presentation by Edwin Klijn from NIOD after a final presentation of the (not quite completed) newspaper digitisation project of the *Koninklijke Bibliotheek* [National Library of the Netherlands] on 22 November 2012.³⁴ Apart from its lack of appeal, mass digitisation is also costly; on average the quality digitisation of one page would cost one Euro, as humanists René van Stipriaan and Wijnand Mijnhardt stated in 2011 in their renowned article ‘Het digitale drama’ [The digital drama]. According to these authors one immediate consequence of these costs is that the digitisation projects carried out so far are ‘sloppy’ and therefore difficult to search.³⁵

Besides their critical remarks, Van Stipriaan and Mijnhardt are strong supporters of digital (historical) research and they blame historians for their lack of ‘to-measure-is-to-know’ mentality. Mijnhardt is of the opinion that the humanities have fallen behind the sciences because they came under the influence of German philosophers, who propagated observation by *Verstehen* around the start of the twentieth century. This then, supposedly, resulted in missed opportunities.³⁶ Again, as Dittrich and Van der Woude did in 1959, an accusatory finger is pointed at the historians. But do they deserve this?

Without making archival material digitally available on a large scale there will only be limited adoption of the tools developed by computer technicians, simply because they cannot be applied on much of the (analog)

34 <https://twitter.com/ewoudsanders/status/271629868091965440> (11 July 2013).

35 NRC *Handelsblad* 10 September 2011.

36 <http://tekst123.nl/wp-content/uploads/2011/09/02-20110910->

NTC_20110910_1_048_article1.pdf (11 July 2013).

material that is relevant to historians.³⁷ As long as the preconditions for digital historical research are not met, there will be no methodological innovation. It is striking how little is written about the preconditions for conducting digital historical research in the literature.³⁸ The discussions mostly focus on the quality of the tools and evaluations of algorithms, but one rarely hears about the necessity of digitising specific sources to enable digital historical research.

The diversity and scope of the NIOD archives provide a wealth of information. Improved accessibility and searchability through digitisation would definitely stimulate new research. The digitisation of the approximately 1,300 diaries alone would enrich our knowledge of daily life under German occupation, and also make network analysis possible – another exciting and promising subject in the digital humanities. Without these preconditions however, researchers will continue to fish in the same pond of digital data, with newspaper-research in an undisputed first position.

The increasing number of digitised newspapers unmistakably stimulates new research initiatives.³⁹ This type of large-scale digitisation projects is meant not only for professional researchers but aims to reach a much wider audience of people with an interest in history. In this way digitisation should also further the democratisation of heritage. This noble objective obviously affects the selection of materials to be digitised – digitisation of appealing images (Europeana) versus specific collections that are only of interest to scientific research – but it is as yet unknown to what extent both groups experience wealth or limitations in dealing with these sources, to what extent they can learn from each other in this respect, and to what extent it results in new questions and insights.⁴⁰

37 <http://niodbibliotheek.blogspot.nl/2011/09/ingezonden-het-digitale-drama.html> (11 July 2013).

38 A rare exception is William J. Turkel, Kevin Kee and Spencer Roberts, 'A Method for Navigating the Infinite Archive', in: Toni Weller (ed.), *History in the Digital Age* (Abingdon, New York 2012) 61-76.

39 For example the WASPH project that was concluded in 2012 (see: http://www.biland.nl/wahspcms/media/cms_page_media/4/Leken_samenvatting_WAHSP_projectmetplaatjes04112011_3.pdf (11 July 2013) and follow-up project BILAND. See also the study funded by NWO: 'E-Humanity Approaches to Reference Cultures: The Emergence of the United

States in Public Discourse in the Netherlands, 1890-1990', <http://www.historici.nl/Nieuws/Actueel/horizontaanvraag> (11 July 2013).

40 For example the *Netwerk Democratie. Platform voor democratische innovatie* [Network Democracy. Platform for democratic innovation] that has already looked at digitisation, <http://netdem.nl/category/nieuws/> (10 November 2013). See also Robert Darnton's lecture 'Digitalisering, democratiseren' [Digitisation, democratisation], in which he states that digitisations brings a long-cherished Enlightenment ideal within reach: availability of everything for all people and at all times, <http://www.de-gids.nl/artikel/143570> (11 July 2013).

Humanities and sciences

‘The Humanities are doing better than ever’, Rens Bod writes in his book *De vergeten wetenschappen* [The forgotten sciences]. His conclusion is that ‘in the past as well as the present – and in all likelihood also in the future – the humanities comprise the full spectrum of methods: from the most deconstructivist and relativistic approach to the most mathematical and universalist approach’. However, Bod adds, this multiplicity is not universally appreciated: ‘there is increasing pressure to streamline research, to publish in the same top-level journals, using the same methods’.⁴¹

Bod does not comment on why this pressure is increasing. Could it have something to do with the current neoliberal climate in which, as we argued earlier, human activity is increasingly translated into economic gain (or loss), scientific research is tied to economically defined top sectors and a close working relationship between scientific research and business is pursued? In his recent book *Debates in the Digital Humanities* American humanist Alan Liu writes that debates about the e-humanities rarely address the question ‘how the digital humanities advances, channels, or resists today’s great postindustrial, neoliberal, corporate, and global flows of information-cum-capital’. And he continues:

It is as if, when the order comes down from funding agencies, university administrations, and other bodies mediating today’s dominant socioeconomic and political beliefs, digital humanists just concentrate on pushing the ‘execute’ button on projects that amass the most data for the greatest number, process that data most efficiently and flexibly (flexible efficiency being the hallmark of post industrialism), and manage the whole through ever ‘smarter’ standards, protocols, schema, templates and databases [...].

According to Liu, they do this without any reflection on the ‘whole digital juggernaut to the new world order’, that is to say without cultural criticism. Liu sees a role here for what he calls digital humanists (read: interdisciplinary scientists) who should take part in the societal debate.⁴²

In our opinion the pioneering role that Liu wants digital humanists to play does not alter the fact that, in the current climate of technology-driven research ambitions, ‘hard’ science is doing better in terms of possibilities for public funding than the science of history, because the scientific value and also societal relevance of the latter is relatively difficult to measure and its

41 Rens Bod, *De vergeten wetenschappen. Een geschiedenis van de humaniora* (Amsterdam 2010) 442-443 [The forgotten sciences: A history of the humanities].

42 Alan Liu, ‘Where is Cultural Criticism in the Digital Humanities’, in: Matthew K. Gold (ed.), *Debates in the Digital Humanities* (Minneapolis 2012) 490-509.

economic significance is considered limited. Although this is a fundamental misconception, as practitioners of the humanities have contributed to cultural reflection, raising public opinion and processes of democratisation since time immemorial, it is difficult to quantify the results of historical research in particular.⁴³

The dominance of the market forces places the call for methodological innovation in a new light. For what is the issue here and which discussion should be held? Progress in research methods – fine if supported by the field and if there is open debate among colleagues – or political agenda setting? If the latter is the case, we as historians are in danger, as we are warned both in the Netherlands and abroad. Stefan Collini, for example, Professor of Intellectual History and English Literature in Cambridge, is concerned that ‘quantitative criteria, following the example set by the sciences and business, increasingly squeeze out the qualitative (and therefore less measurable) ones’. In his recent book *What are Universities for?* he argues that universities are not businesses and economic logic therefore does not apply.⁴⁴ This logic should not be applied to historical science either.

There are more warnings that deserve serious consideration. In the *Digital Humanities Quarterly* Svensson, who cannot be accused of being sceptical regarding the digital humanities, writes:

While the ideas of grand challenges and big humanities certainly have attraction and require forward thinking in order to identify complex problems and large-scale visions, we should be careful not to uncritically accept the frame of big humanities, which, for instance, has a tendency to be coupled with a positivist agenda and a homogenization of the humanities.⁴⁵

43 In this context also see the critical publication by Martha Nussbaum, *Niet voor de winst. Waarom de democratie de geesteswetenschappen nodig heeft* (Amsterdam 2010) [Not for profit: Why democracy needs the humanities]. See also: Geeske Langelans and Joanita Vroom, ‘Wetenschap van woekerwinsten. Het rendement van geesteswetenschappen doorgemeten’ [Science of exorbitant profits: Measuring the output of the humanities], *Mare, Leids Universitair Weekblad* 36:15 (24 January 2013), <http://www.mareonline.nl/archive/2013/01/24/opinion-wetenschap-van-woekerwinsten> (11 July 2013).

44 Quoted by Jacqueline Klooster in de *Academische Boekengids* (November 2012). See also: Chris Lorenz (ed.), *If You’re So Smart, Why aren’t You Rich?* (Amsterdam 2008).

45 Patrik Svensson, ‘Envisioning the Digital Humanities’, *Digital Humanities Quarterly* 1:1 (2012).

And Federica Frabetti, researcher in communication, media and culture studies, argued in 2011: 'The model(s) of rationality on which digital technologies are based cannot be 'imported' unquestioningly into the humanities'.⁴⁶ Frabetti observed a discomfort perceived as fundamental in humanities scholars who have the impression that the often very confidently presented digital possibilities should be viewed as more than an addition to the current repertoire of research possibilities.

The (alleged) presumption raises questions that were recently articulated by American literary historian and critic Stanley Fish:

Does the digital humanities offer new and better ways to realise traditional humanities goals? Or does the digital humanities completely change our understanding of what a humanities goal (and work in the humanities) might be?⁴⁷

Fish has doubts, not least because he fears the decline of the role of interpretive close reading. Expression of these doubts is a desirable intellectual counterbalance to the visionaries who at present want to conquer the humanities landscape and determine its agenda. The techno-utopian vistas they offer to put the science of history on the map bear witness to grand and compelling beliefs. These are definitely not shared by all present practitioners of digital historical research, but they nevertheless leave a powerful mark. This does not exactly help to convince more traditional historians of the need for such a change in direction, all the more since the visionary supporters do not always seem sufficiently aware of the need to reflect on the shortcomings and flaws that, by contrast, demonstrate the value of various more common research methods in the humanities.

The research projects *War in Parliament* and *Verrijkt Koninkrijk* clearly illustrate that digital and 'traditional' historical research complement each other. New tools make new types of source analysis possible but also have their limitations, in our case related to what we may refer to as the two C's: context and concepts. Analysis of the context by means of careful examination of the harvested hits is essential to be able to accurately gauge its significance,

46 Federica Frabetti, 'Rethinking the Digital Humanities in the Context of Orinary Technicity', *Culture Machine* 12 (2011) 1-22, 16-17, <http://www.culturemachine.net/index.php/cm/article/viewDownloadInterstitial/431/461> (11 July 2013).

47 Stanly Fish, blog post, 'The Digital Humanities and the Transcending of Mortality', 9 January 2012, <http://opinionator.blogs.nytimes.com/2012/01/09/the-digital-humanities-and-the-transcending-of-mortality/> (11 July 2013). See also follow-up blog: 'Mind Your P's and B's: The Digital Humanities and Interpretation' (23 January 2013), <http://opinionator.blogs.nytimes.com/2012/01/23/mind-your-ps-and-bs-the-digital-humanities-and-interpretation/>. A further google-search on author's name and title of the blog post shows a flood of reactions.

com/2012/01/09/the-digital-humanities-and-the-transcending-of-mortality/ (11 July 2013). See also follow-up blog: 'Mind Your P's and B's: The Digital Humanities and Interpretation' (23 January 2013), <http://opinionator.blogs.nytimes.com/2012/01/23/mind-your-ps-and-bs-the-digital-humanities-and-interpretation/>. A further google-search on author's name and title of the blog post shows a flood of reactions.

meaning there is a real practical problem when datasets are extensive, hits run into the thousands and relatively expensive man-hours are limited. Concepts in historical research like ‘wrong’ or ‘pillarisation’ are frequently fuzzy phenomena that are not easily translated into adequate queries. The future will show if these problems can be solved. It is clear, however, that traditional, analog, and new digital research methods must always be evaluated critically and be considered and discussed together, including the – not quite as sexy – necessary preconditions, such as digitisation of paper archives.

This does not alter the fact that historians must be careful not to be put on the defensive. The increasing availability of new digital sources and digital research options continually demands a critical approach, but also the willingness to actually explore those new options and, in the closest possible consultation with information scientists, actively work to improve them. Both the historians and their (potential) audience are confronted with amounts of digital data that demand to be utilised in the most illuminating way in the presentation and the analysis of history. In order to strengthen and anchor the role of historians in this process, it is advisable to integrate the diversity of digital historical research possibilities in academic history education in a realistic and reflective manner, but also to persuade historians of the need to – without further visionary distortion – make room for a digital approach in their research practices. ◀

Hinke Piersma (1963) is a political historian and studied history at the University of Amsterdam. Since 1998 she has worked at the research department of NIOD Institute for War, Holocaust and Genocide Studies in Amsterdam. In 2005 she completed her dissertation about four German war criminals who, from their trials in 1948-1949 until the release of the last two in 1989, were the subject of an intense public and political debate in Dutch society (*De drie van Breda. Duitse oorlogsmisdadigers in Nederlandse gevangenschap 1945-1989* (Amsterdam 2005)). Piersma is coordinator of the e-humanities programme at NIOD. Her other interest concerns the history of European integration, looking into national parliamentary debates on the European integration (later the European Union) and World War II as its founding myth. Recent publications: Hinke Piersma, ‘Educatori pubblici: Gli storici olandesi influenzano la politica contemporanea’, in: *Memoria e Ricerca, Anno XVIII, Nuova Serie 37* (2011); Hinke Piersma, *Bevochten recht. Politieke besluitvorming rond de wetten voor oorlogsslachtoffers* (Amsterdam 2010) and Hinke Piersma, *Op oorlogspad. Jaap le Poole, verzetsman voor het leven* (Amsterdam 2006). Email: h.piersma@niod.knaw.nl.

Kees Ribbens (1967) is Senior Researcher at NIOD Institute for War, Holocaust and Genocide Studies. He is also endowed Professor of Popular Historical Culture and War at Erasmus University Rotterdam. Recent publications: Kees Ribbens, *Strijdtoneelen. De Tweede Wereldoorlog in de populaire historische cultuur* (Amsterdam, Rotterdam 2013); Kees Ribbens and Esther Captain, *Exhibiting the War: The Future of World War II Museums in the Netherlands* (Amsterdam 2011) (Available as e-publication: <http://www.niod.knaw.nl/sites/niod.nl/files/exhibitingthewar.pdf>); and Kees Ribbens, 'Global dimensions van de Tweede Wereldoorlog. Een totale oorlog in een vernieuwend historisch kader', *Tijdschrift voor Geschiedenis* 122 (2009) 450-463. Email: k.ribbens@niod.knaw.nl; ribbens@eshcc.eur.nl.