Universiteit
Leiden
The Netherlands

Reflexive inertia : reinventing scholarship through digital practices

Kaltenbrunner, W.

Citation
Kaltenbrunner, W. (2015, May 27). *Reflexive inertia : reinventing scholarship through digital practices*. Retrieved from https://hdl.handle.net/1887/33061

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/33061

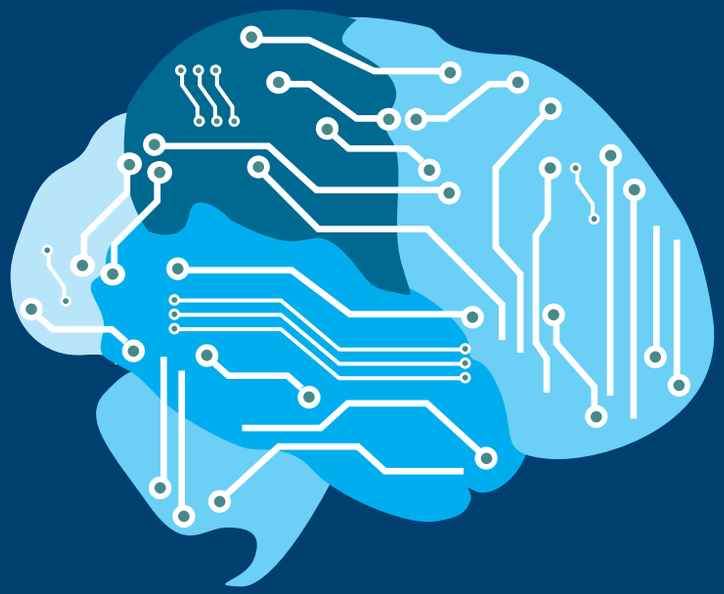
Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/33061 holds various files of this Leiden University dissertation

Author: Kaltenbrunner, Wolfgang
Title: Reflexive inertia : reinventing scholarship through digital practices
Issue Date: 2015-05-27



REFLEXIVE INERTIA

REINVENTING SCHOLARSHIP THROUGH DIGITAL PRACTICES

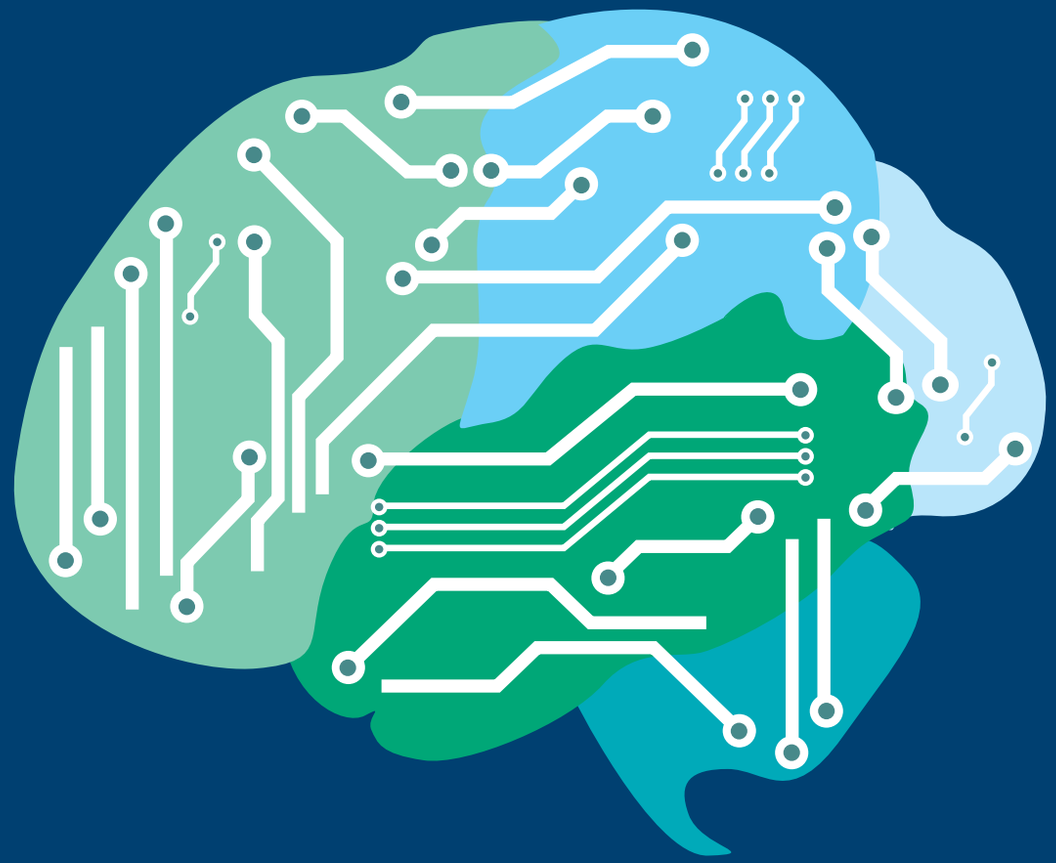
Digital humanities is an emerging field whose practitioners apply digital technology to humanistic research problems. Its manifestations are diverse: from the use of online annotation tools in the collaborative study of empirical sources, the computational analysis of large corpora of textual data, to the use of provocative digital performances for exploring the twists and turns of poststructuralist theory. At the same time, such engagement with novel technologies is often full of tension. In contrast to the single-author, monograph-oriented research that characterizes established forms of scholarship, digital humanities is often practiced in collaborative, interdisciplinary projects that produces digital output rather than traditional publications. The use of digital technology thus creates exciting new possibilities to supplement and extend humanistic knowledge production, but it also entails uncommon requirements regarding the epistemic, social, and material organization of research.

Drawing on a combination of ethnographic work and theories from Science & Technology Studies, this thesis investigates the conflicts that arise as scholars try to incorporate digital approaches into their established practices. Its main argument is that lasting innovations in the scholarly work process will only be possible if they are informed by a reflexive sensibility for the history and organizational specificities of the humanities.

Wolfgang Kaltenbrunner (1981) is a postdoctoral researcher at the Center for Science & Technology Studies at Leiden University. He holds degrees in Comparative Literature as well as in Science & Technology Studies.

REFLEXIVE INERTIA REINVENTING SCHOLARSHIP THROUGH DIGITAL PRACTICES

WOLFGANG KALTENBRUNNER



REFLEXIVE INERTIA

REINVENTING SCHOLARSHIP THROUGH DIGITAL PRACTICES

DOCTORAL DISSERTATION

WOLFGANG KALTENBRUNNER

Reflexive Inertia

Reinventing Scholarship Through Digital Practices

Wolfgang Kaltenbrunner

© 2015 Wolfgang Kaltenbrunner
Cover design: Julia de Jong
Printing: Leiden University. This thesis was printed with financial support by Leiden University and the Netherlands Graduate School of Science, Technology, and Modern Culture (WTMC).
ISBN 978-90-9029014-0

Reflexive Inertia
Reinventing Scholarship Through Digital Practices

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof. Mr. C.J.J.M. Stolker ,
volgens besluit van het College voor Promoties,
te verdedigen op woensdag 27 mei 2015
klokke 11: 15 uur

door

Wolfgang Kaltenbrunner
geboren te Vöcklabruck, Oostenrijk
in 1981

Promotiecommissie

Promotores:
Prof. dr. Paul F. Wouters (Universiteit Leiden)
Prof. dr. Francisca M.G. de Jong (Erasmus Universiteit Rotterdam)

Overige leden:
Dr. Anne Beaulieu (Universiteit Groningen)
Prof. dr. Sally Wyatt (Universiteit Maastricht)
Prof. dr. Willard McCarty (King's College London, UK)

Table of Contents

Acknowledgements **1**

Introduction **3**

The success of digital scholarship - an unexpected turn of events 3

New tools, new knowledge? 9

Infrastructural inertia and reflexive adaptation 14

Summary of the research questions 23

Methodology: the case study and the article-based PhD dissertation 23

Chapter 1 - E-research and methodological innovation in Dutch Studies **35**

E–research and the humanities 35

History of the BNTL 36

The BNTL in different research practices 38

Technological innovation and disciplinary identity 42

Discussion and conclusions 47

Postscript to chapter 1 *51*

Chapter 2 - Scholarly labor and digital collaboration in literary studies **53**

Parenthesis – relation to conceptual framework *53*

Introduction 56

The organization of scholarly labor in literary studies 57

Case study & methods 63

Discussing theory in a pragmatic way 65

Delegating data input 70

What can you expect from a database, and what can it expect from you? 75

Conflicting demands 78

Resolving the tensions in local contexts 80
Discussion 82
Postscript to chapter 2 *86*

Chapter 3 - Decomposition as practice and process: creating boundary objects in computational humanities **89**
Parenthesis – relation to conceptual framework *89*
Introduction 92
Case and methods 93
Hinterlands and boundary objects 94
Different hinterlands in ENS 98
Data as a link between hinterlands? 101
Emerging boundary objects 104
Decomposition and cyberinfrastructure design 108
Postscript to chapter 3 *112*

Chapter 4 - Infrastructural inversion as a generative resource in digital scholarship **115**
Parenthesis – relation conceptual framework *115*
Introduction 117
Infrastructural Inversion as Articulation Work 118
Four Case Studies of Inversion in Digital Scholarship 124
Conclusion 137
Postscript to chapter 4 *140*

Chapter 5 - Digital infrastructure in the humanities: reconfiguring the organization of scholarly tool development **143**
Parenthesis – relation to conceptual framework *143*
Introduction 145
Infrastructure as a discursive interface between policy and scholarship 146

Methods 152

European Union: Digital infrastructure as a catalyst for integration 153

United States: Infrastructure as an emergent property of ongoing digital scholarship 162

Discussion 170

Postscript to chapter 5 *174*

Conclusion **177**

Summary of empirical findings 179

Theoretical and practical implications 191

Directions for further research 197

References **201**

Written sources and digital artifacts cited 201

Interviews cited 227

Summary in Dutch **229**

Curriculum Vitae **239**

Acknowledgements

Firstly, I would like to thank Paul Wouters for being an exceptionally inspiring and patient supervisor, as well as an amazing person all around. A supervisory meeting with him always helped me to shrug off the inevitable everyday frustrations that come with doing research, and to remind myself why I wanted to become an academic in the first place. I also wish to express my gratitude to my co-supervisor Francisca de Jong, who has provided consistently constructive and useful feedback even though her disciplinary background is very different from mine.

Secondly, I would like to thank all the members of the former Virtual Knowledge Studio for the Humanities and Social Sciences (VKS), and its sister institution, the Erasmus Studio Rotterdam (Sally Wyatt, Anne Beaulieu, Andrea Scharnhorst, Charles van den Heuvel, Jeanette Haagsma, Smiljana Antonijevic, Jan Kok, Nick Jankowski, Krzysztof Suchecki, Anja de Haas, Stef Scagliola, in addition to numerous temporary visitors). The VKS – prematurely discontinued in spite of its impressive academic output and international reputation – provided an incredibly inspiring intellectual environment, and it was through my stay at the VKS that my interest in digital scholarship was originally piqued. An important stimulus for my work were also the research meetings at the e-humanities group in Amsterdam, led by the energetic Sally Wyatt. Together with my discussions with Joris van Zundert at the Huygens Institute, as well as valuable feedback and support by Willard McCarty, this made it easier for me to stay in touch with ongoing developments in digital scholarship in the Netherlands and abroad.

Apart from the above mentioned venues, I had the good fortune of participating in WTMC, the national graduate school for Science, Technology and Modern Culture. Attending WTMC workshops and summer schools encouraged me to delve into the various strands of an ever growing Science & Technology Studies literature, as well as breaking up the often lonely routine of thesis writing with welcome social events. I would like to thank Teun Zuiderent-Jerak and Willem Halffman for doing a wonderful job at organizing the WTMC meetings, and of course all the great colleagues and friends I was lucky enough to meet there.

My department, the Center for Science & Technology Studies (CWTS), has provided a very supportive environment for my research. I am grateful to all of my colleagues, and in particular my fellow PhD candidates Cathelijn, Zoreh and Joost, with whom it was a pleasure to share a room. Two good friends, Sarah de Rijcke and Clifford Tatum, came to the CWTS

some time before me. They helped me to find my way around, and they accompanied my thinking process through regular input. Next to that, I benefitted from regular discussions with Thed van Leeuwen and Alex Rushforth.

Much of the research I present in this thesis is based on ethnographic work in the context of two digital humanities projects, the COST Action Women Writers in History and the Elite Networks Shifts. I would not have been able to collect my empirical material without the generous support of their participants, and I would like to express my sincere gratitude to all of them.

Finally I am grateful to my friends and family. My mother and sister have been tremendously supportive, and my regular visits at home not only allowed me to focus, but also to relax and enjoy traditional Austrian cuisine. Research, after all, is only one of the things we do in our lives. Similarly important were the good friends that accompanied me during the past years: Rainer, Carla (looking forward to reading your thesis soon!), Harald, Sarah, Astrid, Lucas, Anil and Irene. Thanks for bearing with me, and for regularly taking my mind off the research.

Introduction

The success of digital scholarship - an unexpected turn of events

Digital scholarship in the humanities has a much longer history than traditionally trained scholars and casual followers of higher education debates might realize. Arguably the earliest instance can be found in the machine-generated, systematic word index of the writings of St. Thomas that the Jesuit Roberto Busa created for his doctoral research in the late 1940s (Busa, 1980; Hockey, 2004). Busa, who is now widely credited as the foundational figure of digital humanities, used this concordance to develop an argument about the composition and structure of the Church Father's language, which in turn served him as a basis for engaging with hermeneutic debates in theology. The first attempts to apply computers in the disciplines of linguistics and history date back to the 1950s (Boonstra, Breure & Doorn, 2006). Linguistics has a long tradition of using large language corpora for statistical analysis (see Swadesh, 1952), and computational analysis provided a useful way of accelerating research as well as broadening its empirical scope.[1] Social and economic historians again showed interest in using digitized demographic information, tax records, and other data collected by public administrators to replace the heroic simplifications of 19th century historiography with more detailed and comprehensive accounts (Boonstra, Breure & Doorn, 2006; Kok & Wouters, 2013). The terms most commonly used to summarize such work - 'history and computers' or the more encompassing 'humanities computing' - drew attention to the novelty of modern information technology as an artifact in everyday scholarly practice. Similar to current discussions, early efforts in digital scholarship were characterized by sometimes bold visions of what computers can do for the humanities. At the same time, these were primarily formulated by practicing researchers and often strongly embedded in the discourse of their 'home discipline'. Scholars of social history for example used computational approaches to extrapolate existing research questions through larger quantities of empirical material than previously possible, and bible scholars considered digital indices verborum primarily as the logical, more powerful extension of print concordances.

1 A foundational research problem for the related field of language technology was the development of automated techniques for translating scientific literature from Russian into English (Hutchins, 1999).

Throughout the second half of the 20th century, the development of digital scholarship followed a relative steady course. The use of computational approaches in the humanities was generally perceived as rather exotic by the vast majority of traditional scholars, posing not least a risk for disciplinary career development (see Nyhan, 2012). Rather unexpectedly for all, however, digital scholarship has begun to attract significant amounts of funding and public attention over the course of the last fifteen or so years. A range of new appellations have recently been proposed to replace the older term humanities computing: e-research, e-humanities, and, by far the most common, digital humanities. In 2010, the *New York Times* featured a series of articles in which several well-known instances of data-driven scholarship are presented as a model for what academic work in the humanities will look like in the future (Cohen, 2010). The application of computational methods to scholarly research problems is moreover one of the few areas of humanistic inquiry that have actually grown in terms of institutional presence and volume of public, private, and philanthropic funding (Gold, 2012). Yet at the same time as it has gained wider currency, the authority to define the essence and boundaries of digital scholarship is no longer exclusively with the individuals who practice it. In the words of Matthew Kirschenbaum (2012), the term digital humanities has become 'a free-floating signifier', which does not primarily denote a particular methodological approach within a discipline or field of research, but also serves as a label for a wider discussion about the future of the humanities that prominently involves research policy and funding bodies. Perhaps precisely because of its lack of specificity, the 'digital' can connect researchers, funders and policy makers in expectation scenarios, thus providing a refracting lens through which those actors reimagine and gradually change what it means to do scholarship in the humanities.[2] In the language of actor-network theory, we could describe current events as a multitude of simultaneous translation processes (Callon, 1986), in which technologically mediated research practices, institutional arrangements, and relations between different actor groups are reconfigured.[3] However, it would seem naïve to think of this

2 Hine (2008) makes a similar argument for the role of ICT in the disciplinary reorganization of biological systematics.

3 Actor-network theory is one of the most influential approaches to social theory of the last decades. Its main conceptual argument is that reality can be thought of as material-semiotic networks that distribute agency across both human and non-human entities, such as material objects, texts, and institutions (Callon, 1986; Latour, 1999; Law, 2004). Networks emerge through the efforts of entrepreneurial individuals, for example scientists, who

process as a transition towards a singular new model of digital scholarship. Rather we are likely to see competing attempts by different actors to seize the current opportunity and change the organization of academic life in the humanities according to their respective interests.

In terms of intellectual aspirations, practitioners of digital scholarship adopt diverse positions. They include computationally intensive research in an empirically oriented tradition[4], but also initiatives in what could perhaps be termed digital poststructuralism. Intellectual goals are of course often difficult to separate from institutional and professional interests. In fact, the current popularity of digital scholarship, and its status as a recipient of funds to "bring information technology into the humanities" (Liu, 2005: 11) has not least been a result of earlier struggles to reposition humanities computing in the institutional hierarchy of the academic system. This can be gauged by looking at the short history of the term 'digital humanities'. Many of its now famous North American representatives started their careers in IT support centers that had originally been set up as service providers to faculty researchers (Nyhan, 2012; Flanders, 2011). Intent to consolidate these centers as an academic workplace that would allow to combine service functions with more explicitly intellectual ambitions, digital scholars increasingly took issue with the label humanities computing, which in their view had become widely associated with an auxiliary activity that merely 'facilitated' the work of other researchers (Kirschenbaum, 2010; 2012). Instead, they wanted to underline that the dedicated use of digital technology makes for a fundamentally different practice that mediates and shapes the very intellectual substance of scholarship. The term digital humanities, originally used in various pioneering institutions in the US, was gradually adopted as a more desirable alternative that suppressed the connotation of service work.

Apart from dedicated efforts in which novel technology is given a central role in the organization of the research process, there is currently a large majority of scholars who apply digital tools in more basic ways (see Bulger et al., 2011), for example by using digital library services, social media, or networked reference management software. However, digital

---

selectively enroll human and non-human actors and subsequently draw boundaries between elements designated as either 'natural' or 'cultural' (Latour, 1993). While not at the heart of my theoretical framework, I will variously refer to actor-network theory in the following chapters.

4 See for example publications in the journal *Literary & Linguistic Computing*, a longstanding outlet for computationally intensive research in the humanities.

scholarship has stimulated reflection not only in terms of its potential beneficial effects. In fact, its very success in attracting public attention and funding has also generated a certain suspicion in recent years. A number of renowned academics have expressed discomfort about a seeming proximity of the digital humanities to neoliberal approaches to research management (see Pannapacker, 2013; Chun, 2013), for example in terms of the distinctively upbeat, promotional rhetoric by which some digital scholars present their work in the media (Cohen, 2010), the coincidence of policy calls for more collaboration in research and the traditionally project-oriented format of most digital scholarship, or the pronounced pragmatism and aversion to high theory displayed by some influential practitioners (Scheinfeldt, 2008).

An important background for the policy interest in digital scholarship is provided by the many digital infrastructure projects currently underway in Europe, the US, and Asia. These aim to create a pervasive technical basis that would allow researchers from all fields to draw on large amounts of data, get access to sophisticated analytical tools, as well as technologies for facilitating collaboration across disciplines and countries. Digital scholars are often the primary beneficiaries of the grants disseminated by such project frameworks, since they have a longstanding experience in the development and application of research technology. Particularly influential in promoting the idea of digital infrastructure has been a report authored by computer scientist Dan Atkins and his colleagues (2003) for the National Science Foundation (NSF). Introducing the popular term cyberinfrastructure, Atkins et al. argue that investment in digital infrastructure is absolutely indispensable if the US wish to retain their position as the scientifically most productive nation in the 21st century. In 2004, the NSF acted on their recommendations by setting up a specialized division that distributes grants for infrastructure projects in various disciplines of science and engineering. The National Endowment for the Humanities created a complementary unit, named Office for Digital Humanities, in the following year. In Europe, digital infrastructure projects are connected to the idea of integrating the various national research systems into a European Research Area, thus increasing their research performance as well as the translation of basic science into economically viable innovation (EC, 2000; ESFRI, 2006a). Respective initiatives can in several respects be seen to intervene into the organization and practice of humanistic inquiry (Beaulieu & Wouters, 2009; Barjak et al., 2013). The European Strategy Forum for Research Infrastructures (ESFRI), set up in

2002 by the European Commission and Council of Ministers, offers not only new grant opportunities for digital scholarship and the development of digital research tools, but also coordinates these activities by issuing a regularly updated roadmap. Funded projects are typically expected to create reusable digital resources. Specific funding decisions will moreover be informed by rationalizing considerations, for example regarding the relative redundancy or complementarity of a prospective new research tools in relation to already existing facilities (cf. Zorich, 2008). The policy perspective of digital infrastructure thus has certain consequences for research governance: it refracts the organizational structure of the humanities in such a way as to foreground a shared layer of analytical applications and datasets, which can then be applied according to the specific needs of the user. This will create a certain tension with the way many scholars view their activities. Especially if they are not yet initiated to digital research practices, they are likely to think about their work in terms of longstanding disciplinary traditions, theories, and methods, as transmitted through institutionally embedded curricula, rather than in terms of an underlying layer of shared digital data that must be used as efficiently as possible. Other scholars again may agree with research policy and funding bodies on the need to create digital infrastructure, but in so doing may pursue different normative and intellectual priorities.

Particular conceptualizations of digital infrastructure also have epistemic implications. In many cases, research policy associates the development of digital infrastructure with the expectation that it will make research in the humanities more data-driven, and less hermeneutic. The idea here is that the availability of digital data and analytical tools will enable scholars to make firmer claims with respect to their research questions. For example, the prestigious funding framework Digging into Data – a joint initiative pooling resources by four international funding bodies – takes its participating projects as proof that the humanities are no longer adverse to the use of technology, and by now underway to embrace computationally intensive research. The report insists that such research constitutes a singular culture of e-research, in which the bifurcation of the humanities and sciences, as (in)famously posited by C.P. Snow, is no longer existent (Williford & Henry, 2012: 7). Similarly, British attempts to create digital infrastructure for humanities scholarship were originally a byproduct of the UK e-science program. The latter is strongly informed by the needs of data-intensive fields like particle physics, astronomy, or genetics, and its leading promoters usually are computer scientists with distinct computer science

research agendas (Hey & Trefethen, 2002; Hey & Trefethen, 2005; Wouters & Beaulieu, 2006). Although there have in the meantime been dedicated efforts to make the specific features of humanities scholarship a more important consideration in the design process (Anderson, Blanke & Dunn, 2010), the British concept of digital infrastructure as a source of high-performance computing and large-scale datasets is still indicative of its original modeling on particular natural sciences. To give a third example, efforts to promote digital scholarship in the humanities in the Netherlands are funded under the title of computational humanities. Conceptually very similar to the above mentioned efforts, the Dutch approach is moreover particularly explicit about the goal of using digital infrastructure to reduce the organizational fragmentation of the humanities. An underlying assumption is that computational humanities can remedy what research policy perceives to be a lack of internal coherence of humanities scholarship in the Netherlands (Willekens et al., 2010).

If it is analytically useful to think of current developments in digital scholarship as a multitude of simultaneous translation activities, then it is also important to operate with a sufficiently complex idea of the object that is being translated. Drawing on a range of complementary theoretical traditions, I will in the following section develop a conceptual view of the humanities as a knowledge-producing machine[5], which has developed historically into its current configuration. This machine consists of many moving parts – institutional mechanisms by which disciplinary practices and identities are reproduced, established epistemic and methodological conventions that are embedded in material tools, and nationally specific relations between researchers, funders, and policy makers, to name but a few. The complex, distributed nature of this machine also implies that its

5 The literature on infrastructure studies, which provides the main conceptual source for my own theorizing, is characterized by the use of both organic and mechanical metaphors. Star & Ruhleder (1996) for example picture distributed but interrelated practices as an ecological system that develops in an evolutionary fashion. Drawing on Hughes' (1983) foundational work on large technical systems, on the other hand, Edwards (2010) and Edwards et al. (2011) visualize scientific networks as a complex machine or engine that is developed in small, adaptive steps. In this thesis, I use both organic and mechanical metaphors so as to draw out different conceptual aspects. The ecological metaphor has the advantage of stressing the delicate balance of practices, while the notion of evolutionary change has an unfortunate connotation of inevitability. By contrast, while perhaps less effective in conveying a sense of mutually sustaining practices, the metaphor of a machine that undergoes partial modifications usefully emphasizes the need for intentional, reflexive human agency.

functioning is characterized by a certain inertia.[6] Modifying or rebuilding it is not something that can be achieved by an isolated, centrally operating engineer, however skilled he or she might be. This inertia should not be seen as a negative feature. After all, it is what ensures that distributed practices are compatible, and that individual knowledge contributions form more or less coherent intellectual traditions. It also means, however, that incorporating new parts – for example in the shape of new tools and digitally mediated practices – must be done with circumspection for the history and complex design of the machine. Adapting a metaphor from Edwards et al. (2011), we could think of such reflexivity as a lubricant that allows to develop its mechanics without producing a jam due to carelessly fitted new parts.

New tools, new knowledge?

When reimagining scholarship as a digitally mediated activity, actors make implicit or explicit assumptions about the relation between changes in research technology and changes in the research process itself. These assumptions may often have an important effect on practical strategies for implementing digital tools in established practices. For example, the promise of a technology-induced, revolutionary change in the epistemic fabric of scholarship logically suggests an approach of concentrating technical expertise and funding in the framework of a centrally managed digital infrastructure project. I would go as far as to say that the discourse on digital scholarship proper is accompanied by a meta-discourse on innovations in technological instruments and their epistemic implications. Partly promotional, partly academic, this meta-discourse always refracts our perception of digital scholarship, albeit serving potentially different purposes.

A popular rhetorical device that the more enthusiastic advocates of digital infrastructure have regularly mobilized to promote their vision is the comparison between digital technology and the printing press. In his keynote speech at the World Social Science Forum 2013 for example, media theorist Derrick de Kerckhove argued that the advent of the printing press triggered a cognitive and scientific revolution, and he drew a parallel between these historical developments and the ways in which digital

6 An alternative metaphor compatible with my perspective is that of infrastructure as following a historical drift (Ciborra & Hanseth, 2000).

technology currently changes knowledge production in the social sciences and humanities. Such comparisons resonate with a venerable academic tradition (often explicitly referenced) that credits the printing press itself with a causal role in promoting the scientific revolution of the 17th century. The scholarly work that has most fully developed this idea is Elizabeth Eisenstein's *The Printing Press as an Agent of Change* (1979). Eisenstein argues in impressive detail that the printing press was instrumental in the separation between experimental philosophy and deductive traditions of thought. For the first time in (Western) history, the printing press allowed to circulate texts in an authorized, stable shape, thus overcoming the susceptibility of manuscripts to errors from translation and reproduction. The possibility to exchange standardized information in turn enabled scholars to engage in comparative empirical work on a large scale. Eisenstein's work is in many ways shaped by the influential theorizing of Marshall McLuhan, who ascribes media technologies a causal effect on public discourse, the organization of society, and the epistemic principles of science. According to McLuhan's popular bottomline, "the medium is the message".

The function of comparisons such as de Kerckhove's has been extensively studied by the sociology of expectations (van Lente, 1993; Brown & Michael, 2003; Borup et al., 2006). The latter has shown that the very hyperbole that often surrounds new undertakings in science and technology correlates to the underlying uncertainty of these visions. Precisely because individual actors wish to change the status quo in some consequential way, they need to make bold promises, thus mobilizing resources and creating protected niches for nascent practices. Many advocates of digital scholarship and e-science have pursued this strategy in order to mobilize resources and make possible targeted interventions in the form of digital research projects and tool building initiatives. Inevitably, however, the use of new research tools – even if they are immediately adopted by researchers in the context of a project – implies a reshuffling of the socio-material setup of extant scholarly work processes, thus creating an inevitable amount of tension with the complex and often invisible ties that connect local practices to the institutional and disciplinary history of the humanities. While seasoned practitioners will tend to be very careful about the strategic promises they make to funders and policy makers (see for example Unsworth, 2007), the tremendous current interest in digital scholarship also brings into play actors with little experience in actually developing and implementing new research tools in a humanities context. Hype often does create the economic

and political preconditions for such large-scale undertakings, but by emphasizing the role of inherent technological potential in effectuating change, it also fails to create awareness or even provide a language for addressing the issues that arise when promises turn into requirements (van Lente, 2000). A good illustration is Project Bamboo, a prestigious US project that is now widely considered an example for how *not* to go about the development of digital infrastructure (Dombrowski, 2014). The goal of Project Bamboo was to create a comprehensive set of digital tools for a large bandwidth of scholarly purposes over the course of 7 to 10 years. Renowned digital scholars criticized Project Bamboo from early on for concentrating significant conceptual authority in the hands of a small group of funding officers, computer scientists, and scholars (Ramsay, 2013). According to these observers, lack of sensitivity to actual scholarly needs combined with its exaggerated ambition let to a disconnect between designers and users, as well as bogging down the development process in lengthy negotiations. Project Bamboo eventually failed to deliver a workable proof of concept at the end of the initial funding period, not to mention the longer term goal of a widely used infrastructure (Dombrowski, 2014; Boast, 2009). A recent report commissioned by the British Research Information Network similarly indicates a conceptual disconnect between technological promises and actual disciplinary needs. The authors conclude that while scholars do have a strong interest in tools that they perceive to meaningfully support their respective practices, they are equally disinterested in applications that primarily seem to result from the research agenda of computer scientists (Bulger et al., 2011).

Here it is useful to highlight that Eisenstein's view of the historical trajectory of the printing press – while still serving as a blueprint for the promotional meta-discourse about digital scholarship - is generally considered to be outdated by more recent generations of scholars of science and technology. In his influential critique of Eisenstein's work, Adrian Johns (1998; 2002) summarizes the distinctive features of this conceptualization of technology: firstly, particular tools are ascribed an inherent potential that determines their use. Secondly, this means that once invented, innovative technologies such as the printing press exert a revolutionary force that will bring about potentially fundamental change in a society. Johns dismisses this view as reductive, and he argues that the book as an artifact was constructed in a process of social shaping. Rather than following the unfolding of some revolutionary potential inherent to technology, it was due to the confluence of various actor interests that the book as we know it today

became stabilized, i.e. associated with particular material format that is invariant across time and space. To support this argument, Johns sheds light on the significant efforts actors had to make to achieve such standardization, and in particular also the persistent attempts to subvert it, for example through piracy and illegitimate editions (see also Johns, 2009).

Theoretically, Johns' critique is heavily informed by research in Science and Technology Studies (STS), a field whose intellectual outlook makes it a highly relevant contributor to the current meta-discourse on the relation between technological instruments and the research process.[7] Historically positioning itself as an academic contender against critical rationalism and Mertonian sociology of science, STS has a rich tradition of problematizing the conceptualization of technology as an autonomous agent that characterizes some current expectations towards digital infrastructure. A core theoretical development in STS scholarship in the 80s in fact was the extension of constructivist views from scientific knowledge to technology (Bijker & Pinch, 1984; Bijker, Hughes & Pinch, 1987). According to this view, technology does not in any straightforward sense determine the practices of its users, but is also itself formed through application in specific contexts. Design here is not exclusively seen as taking place in the offices of engineers and technologists, but as a process that extends to the sites in which technology is being put to use (Oudshoorn & Pinch, 2003). This idea is condensed in the notion of mutual shaping (see Williams & Edge, 1996). Moreover, STS has a longstanding tradition of conceptualizing knowledge production as a situated socio-material practice (Knorr Cetina, 1981; Knorr Cetina, 1999, Latour & Woolgar, 1979; Latour, 1987) that is tied to historically developing, institutionally anchored disciplines (Whitley, 2000; Becher & Trowler, 2001). STS typically pictures tools, conceptual frameworks, and social order in a site to form particular configurations, thus making possible the generation of specific forms of knowledge. A change in the technological base of this configuration will affect, and be affected, by its epistemic and social features (Galison, 1997). Given the complex intertwining of these elements, and their specificity to historical and geographical context, no such change is likely to follow the pattern of sweeping revolution.

A range of different actors have applied the constructivist lens to look at

7 Johns' monograph is based on his PhD research, which was supervised by the influential STS scholar Simon Schaffer. Moreover, Johns explicitly points out his intellectual debt to the equally eminent Steven Shapin, in particular the book *A Social History of Truth* (Shapin, 1994).

current developments in digital scholarship. On the one hand, a relatively small number of STS scholars (Wouters & Beaulieu, 2006; Beaulieu & Wouters, 2009; Dutton & Jeffreys, 2010; Fry & Schroeder, 2010; Meyer & Schroeder, 2010; Schroeder, 2008; Schroeder & Meyer, 2013; The Virtual Knowledge Studio, 2008; Wouters et al., 2013). One explanation for this valuable, though overall limited interest could be that STS has historically had a focus on the natural sciences, in order to show that constructivist accounts can be applied to 'hard' forms of knowledge production like particle physics or molecular biology. On the other hand, theoretical concepts from STS have recently been taken up by actors who have a more immediate stake in current translation activities. Digital scholars in the US have used STS knowledge to stress the grounding of technology use in local practices (ACLS, 2006; Unsworth, 2007), thus promoting a vision of infrastructure as an emergent property of ongoing digital scholarship. Many European policy makers and infrastructure project leaders have in turn switched to frame the goals of respective funding frameworks in a terminology of mutual shaping. A strategic document published by the European Science Foundation for example argues that the development of digital infrastructure must involve prospective scholarly users from early on, if the technology is actually to be adopted on a wide basis (ESFRI, 2011). Similarly, Anderson, Blanke & Dunn (2010) outline how the European infrastructure initiative DARIAH is informed by the concepts of mutual shaping and trading zones (Galison, 1997). Anderson, Blanke & Dunn suggest that both a techno-deterministic view in which new tools inevitably bring about new practices, as well as a radical constructivist view in which research practices are largely immune to technological stimuli, fall short. Instead, digital infrastructure should be conceptualized as a 'marketplace of services', from which scholars can adopt those tools that suit their needs, and in ways that individually make sense for them.

Naturally, such use of analytical concepts from STS should be seen with a critical sensibility as to its political function. The mere adoption of a terminology of mutual shaping does not tell us anything about whether and how the concept will inform the design process of digital research technology, or the way conflicts resulting in the implementation phase are resolved. The use of trading zones in infrastructure projects similarly requires close attention to the organizational and political context of respective initiatives. Trading zones can emerge from a relatively 'peaceful' encounter between different disciplinary cultures who work towards viable arrangements for all parties over time, or they may be the result of coercion

in which one set of practices simply supplants another one (Collins, Evans & Gorman, 2010). The very choice that something should become a trading zone, taken by one group of actors for another one, is itself an important expression of power. Framing infrastructure projects in the terminology of trading zones, then, does not necessarily guarantee a more democratic or context-sensitive way of creating technology, but may well serve a similar rhetorical purpose as older techno-deterministic accounts, namely that of creating political legitimacy for technological choices.

We are thus confronted with a somewhat paradoxical situation: STS terminology and ideas have diffused into policy contexts, where they are used for various political and intellectual purposes, while STS itself has devoted relatively little, proper attention to digital scholarship. What is therefore needed, I suggest, is empirical research and further theoretical reflection on the actual development and use of digital research technology by humanities scholars. This analysis should be reflexive with respect to the inevitable partiality of its underlying conceptual choices, and to the way social scientific insight may itself become a resource for actors. The following thesis is the outcome of my attempts to engage in exactly such research. The chapters are conceived as separate journal publications. They are held together, however, by the overarching theme of scholarship as a historically grown, institutionally embedded, and therefore inert phenomenon, which is currently being reinvented through the lens of the digital. My argument is that only those translations that are sensitive to the local manifestations of this inertia will lead to longer lasting innovations in the scholarly work process.

Infrastructural inertia and reflexive adaptation

A first assumption of this thesis is that tools have no essence that inherently determines their use. Instead, the function of individual research tools must be seen in the context of scholarship as a larger, historically developing infrastructure. Drawing on the work of Suzan Leigh Star, Karen Ruhleder, Geoffrey Bowker, and others (Star & Ruhleder, 1996; Bowker & Star, 2000; Hughes, 1983; Edwards, 2010), I here use the term infrastructure in a specific theoretical sense: it denotes the relational state that obtains when heterogeneous, but cooperative practices, for example in the academic labor ecology, achieve a state of smooth coordination.[8] The conduct of scholarly

8 In the present chapter as well as in the remainder of the thesis, I have tried to distinguish

work in fact depends on the accomplishment of a large number of other tasks, ranging from administrative work and the maintenance of physical workspace to the provision of library and information services, to name but a few. Each of these tasks constitutes a practice in its own right, complete with specific conventions that must be mastered by new entrants, and which therefore also inform the larger whole in one way or another. An important interface between different task areas is provided by all sorts of standards and classifications, which are simultaneously material and cognitive. Think of the countless norms that are embodied in a tool such as a computer keyboard, and the way it shapes the act of writing a scholarly paper both as a physical act and as an intellectual operation. In the absence of a singular management structure, standards exert a subtle normative force that makes heterogeneous practices compatible. Infrastructure thus invisibly supports tasks, rather than having to be assembled or reinvented every time anew. As a side effect, infrastructure tends to be transparent to its acculturated users, precisely because it is reproduced through the work routines those actors engage in on a daily basis.

Infrastructure studies is usefully complemented by the concept of hinterlands as proposed by John Law (2004). This perspective similarly emphasizes the invisible constraints of historically grown research practices, albeit with more explicit attention to their implications for the specific forms of knowledge created. Hinterlands denote particular assemblages of material and symbolic elements that temporarily cohere in the methodological apparatus underlying disciplinary knowledge production. While this concept is in many ways compatible with Kuhn's paradigms, it puts more focus on the notion that established research practices also enact reality, rather than merely providing historically shifting possibilities for describing it. Law argues that instead of discovering facts 'out there', methods amplify certain realities out of a wide range of possibilities, thus foregrounding some elements while pushing others into invisibility. The economic metaphor of the hinterland is meant to express that some realities are easier to establish because they draw on standardized packages of material and semiotic relations (see Fujimura, 1987; 1992). Such packages

---

this specific theoretical meaning from the unrelated terms 'digital infrastructure' and 'cyberinfrastructure' through appropriate contextualization. For example, when discussing the development of digital infrastructure in Europe and the US, I provide enough description to make clear that I am referring to particular empirical phenomena, rather than the theoretical framework of Star & Ruhleder.

gradually develop through knowledge and practices that become widely acknowledged in a field, thus crystallizing into a relatively stable, taken-for-granted epistemic and praxeological framework. Examples of packages include established data formats in a given field, stylistic and argumentative conventions for communicating research findings, and implicit social protocols that order the interaction among disciplinary peers. For example, in the monograph-based scholarship of history and literary studies, narrative forms of knowledge are traditionally privileged. Historical events are primarily narrated, even when the author draws on quantitative empirical information. A piece of scholarship that is unconventional in its material or literary format - for example by presenting an algorithmic analysis of historical events, or by publishing findings in the format of a digital database – may be difficult to recognize as a valid disciplinary contribution by peers. Insofar as it can only partially build on the work of predecessors, it may require 'redoing' some foundational empirical or theoretical work, and thus end up being dangerously demanding in terms of resource and time investment. This aspect also highlights an important political dimension of the hinterlands perspective. Specific forms of humanistic knowledge here are seen not as the natural ideographic expression of underlying historical or social reality, but as the result of specific methodological choices constrained by various external factors.[9]

An important feature that distinguishes modern academic research from historical forms of knowledge production in fact is the formation of a disciplinary structure as well as the institutionalization of science and scholarship in the 19th century (Stichweh, 1984; Whitley, 2000). According to the traditional model of disciplinary organization, the dynamics of research are determined by mechanisms such as peer review, the circulation of influential publications, and the accumulation of intellectual reputation among colleagues. This disciplinary organization is entwined with institutional power, most importantly in terms of the control over resources

9 While strongly informed by Law's interest in actor-network theory, I would argue that the concept of hinterlands is not least meant as a reaction to the longstanding critique that actor-network theory is politically conservative in the sense of being too descriptive and malleable, thus allowing to see seamless networks where one actually is confronted with complex, yet difficult to discern structural conflicts (see for example Whittle & Spicer 2008). The focus on the unseen and structure-like that is implied by metaphor of the hinterlands has a tendency to highlight rigidity and tensions, rather than malleability. Such an analytical perspective, I suggest, creates a useful contrast to the often hyperbolic policy expectations towards the revolutionary potential of digital research tools.

and employment opportunities exercised by universities. A small but influential body of literature has provided detailed comparative accounts of the various disciplinary cultures that have emerged across the humanities and natural sciences over the last 200 years (Becher & Trowler, 2001; Whitley, 2000; Knorr Cetina, 1999). Particularly useful for my own research is the work of Richard Whitley (2000), who has introduced the analytical concepts of mutual dependence and task uncertainty to distinguish different fields of research on the basis of their relative degree of social and intellectual integration. The field of literary studies for example is characterized by a low degree of mutual dependence and a high degree of task certainty when compared to most natural sciences. Knowledge is circulated in monographs, a format that grants individual academics considerable analytical and stylistic freedom. Populated by proverbially 'lone scholars', there is a large variety of research goals and coexisting theoretical views. Knowledge production in fact is partly driven by productive disagreement between representatives of different theories, rather than by integration of individual knowledge contributions in a single dominant framework, as is the case in more highly integrated domains in the natural and quantitative social sciences. Different organizational features also go along with specific types of research instrumentation. Fields characterized by strong consensus on theoretical frameworks and data formats across sites tend towards uniform, often large-scale research instruments. Research here is typically organized as collaborative work process with clearly divided tasks, proceeding from the generation to the analysis of large amounts of data (Galison, 1997). Loosely integrated fields on the other hand are likely to operate with instruments that are more specific to local research contexts (Shrum, Genuth & Chompalov, 2007). At the same time, there are indications that the institutionalized mechanisms of disciplinary research organization have begun to undergo change (Whitley, Gläser & Engwall, 2010). A relatively recent development is the practice of evaluating research on the basis of formalized indicators. These indicators are not only used by research policy, but also by individual researchers and university administrators, with potentially significant effects on the inner workings of the disciplinary reputation economy (Espeland & Sauder, 2007; Martin & Whitley, 2010). The most important development for the present thesis certainly is the emergence of often large-scale, transnational funding frameworks for digital infrastructure, as well as new managerial structures that aim to coordinate research and tool development internationally. Such grants partially loosen the economic and disciplinary constraints on the epistemic choices that

researchers make in preparing knowledge contributions. Yet the managerial imperatives connected to these grants – for example the call to avoid redundant investment and to aim for utmost reusability of data and tools - also create new constraints that may not be congruent with established conventions, thus resulting in a tension between infrastructure funding frameworks and nationally based, disciplinary scholarship.

Through the combination of theoretical resources from infrastructure studies, the perspective of hinterlands and sociological theory, we can begin to outline a number of areas for research. For one, infrastructure as well as hinterlands picture knowledge production as sedimented socio-material practices, which in turn reproduce (or challenge) institutionally recognized definitions of 'proper' scholarship. The use of research tools is often part of a disciplinary curriculum, and the tools are conceptually bound up with the theoretical and methodological base of that field. Insofar as it is widely used in accordance with received notions of its meaning, established technology therefore has a tendency to become transparent to its users. Put differently, specific ways of using technology are encouraged by the institutional, conceptual, and organizational features of infrastructure. Hypothetical affordances of new instruments – perhaps built into them by computer scientists or software engineers operating under a very different set of constraints – in turn may be discouraged and perhaps not even recognizable to acculturated members of a field in the first place. At the same time, scholars may creatively adapt the tools in ways not anticipated by their designers. An important task therefore is to interrogate how the inertia of infrastructure manifests itself in the move to digital scholarship, and to analyze how this inertia shapes the embedding of new research tools in scholarly practice (Chapters 1 and 2).

Moreover, the perspective of scholarship as embedded in a larger infrastructure draws attention not only to the interaction of scholars with technology, but also to the many other task areas that enable this 'primary' research activity in the first place (Strauss & Star, 1999). In traditional modes of humanistic inquiry, the work of scholars has been enabled and constrained by the often invisible work of archivists, librarians, and bibliographers. Historians and literary scholars after all depend on another group of professionals who make sure that textual sources in archives and special collections are accessible and well curated, so that they can be referenced in academic monographs and papers. However, in many collaborative digital projects, the traditional boundaries between scholarly activity and support work are blurred – the various tasks necessary to create,

maintain, and expand a scholarly database for example cannot straightforwardly be divided into curation and research. Insofar as different forms of labor are subject to different reward systems, such shifting of boundaries seems to pose an interesting opportunity for research. At the same time, STS does not have a strong tradition of taking into account the economic dimensions of science. Some scholars (Vann & Bowker, 2001; The Virtual Knowledge Studio, 2008) have attributed this to the common STS approach of conceptualizing scientific work as practice, a perspective that – while meant to broaden Mertonian and Popperian views of science as a purely cognitive operation - also tends to downplay the material cost of research. The move to digital scholarship in the humanities, however, and the transformation of the academic labor ecology it has occasioned, provides a very good reason for studying how the exchange value of scholarly labor mediates its intellectual substance (The Virtual Knowledge Studio, 2008).

The current discourse on digital scholarship in fact strongly tends to emphasize the consequences of digital technology for intellectual work rather than for curation and data work. For example, it has become something of an informal requirement in funding proposals to promise that digital tools will scale up the empirical scope of research. This resonates well with the policy expectation towards a more data-driven, algorithmic form of humanities scholarship, but it tends to downplay the huge amounts of labor that will be necessary for digitizing sources, entering and harmonizing digital data and metadata, and providing sustained support to keep digital materials and applications usable in the long run. There is even a certain tendency in the policy discourse to associate 'digitization' initiatives with cost-cutting. Transforming collections and bibliographies into digital artifacts here is often connected to laying off staff for the maintenance of physical facilities (Baars et al., 2005; PLG GTA, 2013). Projects that primarily aim to hire new staff for large-scale digitization of print sources generally do not have very good chances of acquiring funds (personal communication). Over the past decade, we have witnessed a number of curious attempts to tackle the problem of data work. For example, we have seen recent collaborations between venerable academic institutions and Google. The latter has an interest in improving the attraction of its Google Books service, and therefore has invested significant amounts of money into the digitization of major university collections. However, the mark-up and metadata applied in these digitization efforts has been criticized for not living up to the specialized needs of scholarly inquiry (Duguid, 2007). Moreover, there has been a number of ambitious crowdsourcing projects, the

most publicized of which is probably Transcribe Bentham. The idea behind this initiative was to mobilize interested laypeople to digitize and mark up the complete writings of Jeremy Bentham, given that funding bodies are usually very reluctant to finance such work. Project staff have recently evaluated the success of the undertaking somewhat critically, however (Causer, Tonra & Wallace, 2012). Perhaps unsurprisingly, it has proved very difficult to mobilize a large enough number of volunteers and organize their transcription work in such a way as to attain a satisfying (and economically viable) data quality. Large-scale algorithmic knowledge claims about the works of Jeremy Bentham will therefore remain impossible in the foreseeable future. Under the radar of most casual observers, it seems, digital scholarship is importantly shaped by allocation of funds for what is often seen as mundane data work. An important question therefore is: How do actors construct forms of labor as 'scholarly', 'technical', or 'support activities', and how does the distribution of such labor make possible certain forms of knowledge, but not others (Chapters 1 and 2)?

Another area for research is the way that disciplinarity and situatedness of specific research practices relate to the construction and use of digital research technology. The concept of digital infrastructure, as well as most tools for data-intensive, algorithmic research, are often claimed to be of universal benefit to all fields of research. In reality, these technologies are often informed by the particular requirements of certain natural sciences, thus confronting its prospective scholarly users with analytical possibilities that they do not necessarily have an existing disciplinary need for (Wouters, & Beaulieu, 2006). Previous research has already stressed how the adoption of given tools depends on disciplinary traditions (Borgman, 2007; Collins, Bulger & Meyer, 2012; Fry & Talja, 2007; Fry & Schroeder, 2010). An observed pattern has been that disciplines such as linguistics, which has a strong tradition of computational empirical research and a relatively high degree of internal integration, are quick to take up certain tools, say, for the analysis or large-scale corpora of textual data. On the other hand, disciplines with more internally divergent research priorities and less mutual dependence among individual scholars have found to be reluctant in their uptake of the same tools. There is a number of possible ways in which this could play out in the future. For one, tools could be adopted one-sidedly by fields with established data-intensive traditions. Alternatively, a process of mutual shaping could transform the more hermeneutic fields in such a way as to create a need for such tools. Another hypothesis is that we will see more efforts by scholars to tailor technologies to their specific requirements

and intellectual preferences, rather than adopting applications modeled on particle physics or computer science. An important question therefore continues to be: how does disciplinarity shape the use of digital research tools, and how does the use of these tools shape disciplinarity (Chapters 1, 2 and 3)? Directly related to this is the question of interdisciplinarity. A characteristic expectation towards digitally mediated research – both in the sciences and the humanities – is that it will create new possibilities for collaboration. Digital infrastructure projects in particular promise to facilitate the sharing of analytical tools and data across disciplines and geographical distances (Atkins et al., 2003; ACLS, 2006; Hey & Trefethen, 2005). However, Edwards et al. (2011) have cautioned that greater interaction among researchers will inevitably create 'science friction', i.e. difficulties in communication that arise due to diverging ways of framing research questions, appropriate methods for answering them, as well as differences in handling data. We can reasonably assume that the amount of friction – and thus the additional work involved in bringing collaborative research to closure – will at least partly depend on how strongly the interacting disciplines differ in terms of their characteristic epistemic and organizational features. What does this mean for interdisciplinary collaboration in digital scholarship and the development of digital research tools for the humanities? How do scholars, typically acculturated in hermeneutic traditions such as close reading and thick description, work out a collaborative arrangement with computer scientists and software developers? How do they resolve tensions between very different epistemic frameworks, yet without giving up their commitment to their respective hinterlands (Chapter 3)?

Several aspects of my conceptual framework highlight how the dynamics of digital scholarship are shaped by various forms of constraints. This should not lead us to neglect the inherent underdetermination of new scholarly tools as an important area for investigation, however. STS has rightfully pointed out that while technology often comes with specific scripts built into it (Akrich, 1992; Oudshoorn & Pinch, 2003), i.e. implicitly or explicitly formulated aids for how to 'read' and use a technological artifact, there is nothing in a tool that would fully determine how it is eventually deployed in practice. Similarly, while infrastructural inertia and disciplinary logic in many ways constrain digital scholarship, the development of a field of research is not fully determined by its history. How, then, do actors come to select a specific use of new technology over another one, and how are hinterlands restructured as new tools are embedded in scholarly practices? I

suggest that the development of infrastructure can at least partly be understood as driven by an intentional reflexive agency of scholars. Individual practitioners of digital scholarship typically also try to shape their academic work environment according to their specific visions and normative interests. This often entails going against the grain of established infrastructural routines, thus requiring an investigation of the process by which scholars 'untie' the standardized material and symbolic packages that constitute a given hinterland (Law, 2004; Fujimura, 1987; 1992). Reflexivity in discourse and practice allows to reimagine scholarship in the light of potentialities, and thus fulfills an evolutionary functionary in the development of scholarly infrastructure. Different forms of reflexivity likely will open up different paths for development, however, thus making it also a site of controversy in which different actor interests clash (Chapters 3 and 4).

A last aspect worth addressing in more detail is the relation between digital scholarship as practiced in individual projects and the various, overarching infrastructure initiatives that are concurrently undertaken (Atkins et al., 2003; ACLS, 2006; ESFRI, 2010). In most Western postwar science systems, choices about the development of particular research tools and facilities used to be reserved for disciplinary elites, i.e. researchers who had acquired significant reputation among their peers (Whitley, 2000). Relations between researchers on the one hand, and funding bodies and policy makers on the other, were based on the relatively stable agreement that science would produce a continuous stream of useful knowledge and technology if given a certain discretion in organizing its activities (Mirowski & Sent, 2008). Current digital infrastructure initiatives, however, take place against the background of changing relations between those actor groups (Nowotny, Scott & Gibbons, 2001; Etzkowitz & Leydesdorff, 2000). Policy and funders now often take a proactive role in scientific agenda setting, while university block funding, an important economic basis for disciplinary self-governance, is being reduced (Mirowski & Sent, 2008). Current infrastructure initiatives play a particular role in this reconfiguration process. Usually framed as a foundation for future economic and scientific success of a country or region, they connect researchers, funders, and policies in strategic alliances (Barjak et al., 2013). Not much published research has addressed how intellectual and political interests of different actor groups are bound up in the development of digital infrastructure, and in what ways such initiatives differ between countries. However, this is important empirical knowledge for an STS analysis of the sort I propose, because

specific policies of funding and coordinating tool development will also create distinct conditions for individual projects in digital scholarship. My final research question proposes a comparative analysis of this interaction: What strategic considerations underlie current infrastructure initiatives in Europe and the US, and how do they affect the organization of tool development for digitally mediated scholarship (Chapter 5)?

Summary of the research questions

1. What does it mean to think of infrastructure as inert, and how does this inertia shape the embedding of new research tools in scholarly practice?
2. How do actors construct forms of labor as 'scholarly', 'technical', or 'support activities', and how does the distribution of labor make possible certain forms of knowledge, but not others?
3. How does disciplinarity shape the use of digital research tools, and how does the use of these tools shape disciplinarity?
4. What is the reflexive agency of scholars in the embedding of new tools into their infrastructural work setting?
5. How is infrastructure conceptualized differently across countries, and what role do such conceptualizations play in organizing infrastructure development 'on the ground'?

Methodology: the case study and the article-based PhD dissertation

While each of the following chapters contains a separate methods section, a few words are in order to outline and reflect on the common methodological principles that underlie this thesis. The specific form of knowledge I have produced can perhaps best be described in a reflection on the intertwinement of two changing organizational aspects of academic scholarship: the move from the monograph-based dissertation to an article-based model, and the changing socio-material conditions that accompany the adoption of digital research technology in the humanities.

This thesis follows the model of an article-based thesis, which is becoming an increasingly popular modality of gaining a doctorate in the Netherlands and a number of other European countries, as well as Australia and Canada (Park, 2005; Powel & Green, 2007; Kamler, 2008). An important context for this development is a loosely concerted international effort by academic institutions, policy makers, and researchers to make postgraduate work more 'accountable' and 'transparent' (Park, 2005; Bartelse, Oost &

Sonneveld, 2007). The traditional monograph model dominant in the humanities and qualitative social science entails that the PhD candidate is largely invisible to administrators for a number of years – to put it casually, students would disappear in field sites and libraries for several years, and then hopefully emerge with a scholarly book in their hands. The switch to the article-based thesis on the other hand is many countries related to a contractual stipulation of supervisory obligations (Robins & Kanowski, 2008), as well as the spread of graduate schools that add an educational component to the research process (Bartelse, Oost & Sonneveld, 2007; Sonnveld, 2010). At the same time, there is still a degree of legal uncertainty surrounding the article-based thesis. In the Netherlands for example, no single university offers binding rules for the required amount of papers, for how many papers must be published at the time of submission, or in what type of journals (see Leiden University, 2008; University of Amsterdam, 2014). There is merely the informal recommendation that the thesis should consist of four papers in total with at least one accepted for publication (Vrije Universiteit Amsterdam, n.d.). Of course, the new PhD model also resonates with a parallel process of accelerating the 'throughput' of academic knowledge production, accompanied by the increasing importance of publication-based methods of evaluation. The intellectual and stylistic freedom implied by the monograph makes it particularly difficult for non-disciplinary evaluators to assess its intellectual value, while journal impact factors facilitate relatively simple (though frequently dubious) quality judgments. Critics in fact have portrayed the article-based thesis as a concession to managerial sensibilities that produces 'audit-ready PhDs' as well as a form of 'precocious professionalism' (Park, 2005), thus suggesting a certain leveling of the intellectual quality of doctoral work. As I know from conversations with fellow graduate students from the national graduate school WTMC, the article-based thesis is frowned upon by some faculty in some universities, thus putting graduate students in the uncomfortable situation of having to side with either the defenders of the monograph-model, or those that embrace the article-based dissertation.

I would argue that the new modalities do not produce inherently less valuable knowledge, but rather entail a different way of structuring doctoral work intellectually and in terms of everyday work routines. I personally found the prospect of organizing my work around individually published milestones more appealing than having to work on one single big chunk, with less manifest possibilities for assessing progress. The other reason for opting for individual articles was that I did not see the point of

sticking to the monograph format when research evaluation increasingly values peer-reviewed articles. This also constitutes a specific choice in terms of the prospective audiences, however. A (published) monograph might potentially have reached a wider, non-specialist group of readers interested in digital scholarship, while articles are more likely to be read by social scientists and scholars with access to a digital university library. On the other hand, the article format also means that my arguments feed back to the actors I study in a formal, but quicker way than if I had opted for a monograph. Before discussing this latter implication in more detail, it is necessary to address the specific epistemic constraints that come with an article-based thesis. Rather than a single unified narrative, an article-based thesis implies a collection of several independent publications, which nevertheless amount to a larger argument in their totality. Given the often excruciating duration of peer review, at least two articles should be finished well ahead of the last year before graduation. This obviates the possibility to modify an argument after it has been published, while a monograph in principle allows to continuously develop the argument in light of the most recent insights one develops about the subject. At the same time, when working on individual articles, one often has to respect the wishes of the journal editors and reviewers, thus bringing into play the judgment of a number of additional scholars who may not always be from the same discipline as the PhD candidate, the supervisors, or the PhD committee. An article-based thesis thus trades off speed of circulation for internal coherence of a monograph. Perhaps the most important constraint of the article-bases thesis in STS, however, is that it will tend to further consolidate the case study approach that is dominant in the field. In line with an established tradition in STS, research must be simultaneously based on theoretical reflection and empirical field work, most often in the form of ethnography, participant observation, and interviews. Analyses of individual 'cases' usually provide just the right amount of empirical material for an article, which in turn is becoming a more and more important complement to the monograph as the dominant form of scholarly output in STS.

Of course, rather than picturing the case study as a lucky fit for the genre conventions of the scholarly article, one could also consider it as an artificial way of segmenting reality that actually results from the constraints of academic social science. Much recent writing on STS methodology in fact has critically reflected on the limitations of the case study approach (Shrum, Genuth & Chompalov, 2007; Wyatt & Balmer, 2007; Law, 2004; Hine, 2007; Beaulieu, Scharnhorst & Wouters, 2007). As Shrum, Genuth & Chompalov

(2007) argue, the notion of a case study implies that what is under investigation is a meaningful unit of analysis that simultaneously captures a representative empirical element of a larger phenomenon. This assumption has historically fulfilled the function of legitimizing the knowledge produced by the case-studying researcher, who could thereby claim empirical representativity and authoritative insight about the phenomena he or she studied. At the same time, STS has a long tradition of methodological reflexivity. The explicitly relativist orientation of much foundational STS research with respect to the knowledge claims of natural scientists (Bloor, 1976; Collins, 1985) has early on raised the question as to the status of STS knowledge itself. On what grounds can social scientists claim a form of epistemic validity when their main argument is that knowledge is always socially constructed (Ashmore, 1989)? Since then, method in STS has widely come to be seen as generative of reality, rather than providing a neutral way of describing it (Law, 2004). From this perspective, uncritical use of the case study threatens to reify both the level of an individual case as the natural unit of analysis, and that of a larger culture or system of which the individual case allegedly constitutes a part, and which in turn must be specified by a number of shared features (e.g., organizational characteristics, tools used, conceptual elements) (Beaulieu, Scharnhorst & Wouters, 2007). Moreover, some scholars (Hine, 2007; Mol, 2002) have criticized the traditional understanding of the case study for its tendency to associate the individual case with the local, and the level of a discipline or field with the global, supra-local. The latter metaphors imply a potentially misleading spatiality that will tend to undermine how, say, what happens in the context of an individual project is simultaneously affecting and being affected by an academic field as such (Beaulieu, Scharnhorst & Wouters, 2007; Beaulieu & Simakova, 2006; Beaulieu, 2010; Jensen, 2007). An important influence for this type of critique has been non-dualistic theorizing by authors at the intersection of STS and anthropology (Strathern, 1991; Verran, 2001; Tsing, 2005; Mol, 2002; Law, 2004). The central argument of this literature is that dichotomies such as local/global, micro/macro, and field/laboratory cannot be used as stable concepts on which to base methods, but must themselves be analyzed as achievements of actors.

This methodological reflection has been additionally stimulated by the proliferation of ICT and the need to adapt existing theories to the specificities of digitally mediated research. While not inherently more complex than older modes of knowledge production, the development of digital infrastructure does go along with an epistemic, material, and

economic reorganization of research. As I have variously mentioned, a widely shared expectation is that the spread of digital instrumentation will facilitate collaborative relations between researchers across individual sites and disciplines, for example through the reuse of large datasets and networked tools for analysis. This suggests that individual epistemic artifacts will travel a lot more than was common in the past. Several scholars in STS and related fields such as Computer Supported Cooperative Work (CSCW) (Beaulieu, Scharnhorst & Wouters, 2007; Hine, 2007; Williams & Pollock, 2012) have therefore called for replacing a methodological focus on the interaction between actors and technology in isolated sites, and under the assumption of stable dichotomies of micro/macro, with a multi-sited ethnography (Marcus, 1995). The goal here is to study for example the mutual shaping of practices and technologies in use (Pipek & Wulf, 2009), as well as how seemingly foundational dichotomies are constructed and performed by actors. This also includes the ambition to trace the historical development of infrastructure over extended periods and from different viewpoints. Infrastructure may involve multi-dimensional temporalities specific to individual elements (e.g. the temporal logic of scientific careers vs. the lifetime of a given software package), and thus give rise to complex dynamics that become visible only through simultaneous attention to short and long-term events (Ribes & Lee, 2010; Ribes & Finholt, 2009). Beaulieu & Simakova (2006) moreover have proposed to take seriously the topology of the network, for example by studying how hyperlinks hook up different contexts in ways that transcend physical and intangible field boundaries. While few scholars would deny that it is desirable to extend the focus of ethnographic work in both time and space, Jensen & Winthereik (2013) have again cautioned against an empiricist tendency among practitioners of multi-sited ethnography themselves. Especially in the field of CSCW (e.g., Pollock & Williams, 2010), they argue, there is often an assumption that infrastructure can be strategically charted beforehand, with the expectation that carefully aggregated case studies can reach a degree of saturation over time and will add up to a comprehensive empirical picture if only enough empirical work is conducted.

By and large, however, STS advocates of multi-sited ethnography seek to establish methodological legitimacy not through expansive empirical coverage (i.e., through distributed and longer fieldwork sessions), but rather through intensity of their ethnographic interaction with infrastructure. The underlying theoretical assumption is that an essential distinction between the activities of actors and analysts cannot be drawn, since both engage in

mutual translation activities (Zuiderent-Jerak & Jensen, 2007; Zuiderent-Jerak, 2007; Vikkelsø, 2007). Thus, actors enroll other people and objects into networks, but analysts do the same when they collect empirical material and distill it into papers. In this perspective, it is no longer possible to claim that social scientists – after having 'covered' a preexisting fields through enough representative case studies – have straightforwardly superior, objective insight that they can for example use to advise policy. Rather, social scientific interventions are seen as implicated in partly unpredictable dynamics of alliance, betrayal, and negotiation with actors, with the latter pursuing often incommensurable normative interests as well as incommensurable ways of framing relevant issues. Zuiderent-Jerak & Jensen (2007) therefore argue for a social scientific 'ethics of specificity', i.e. the attempt to contribute to the explicit framing and resolution of carefully studied, situated conflicts. Hine (2007) suggests that such an approach can also provide a new means of making social scientific research relevant. Rather than framing the timeliness or adequacy of research in absolute terms, she suggests that social science is relevant if it creates meaningful resonances with actors in specific contexts, e.g. by meaningfully framing problems that in turn feed back into the translation work of actors. A number of researchers has moreover proposed to use the isomorphic relation between the translation activities of STS researchers and the actors they study as a methodological device. Jensen & Winthereik (2013) for example use their own experiences in doing fieldwork as an empirical instance of how infrastructure develops – here, the need to make certain connections with actors, to enter specific kinds of partnerships in order to get access to information, is not considered a practical nuisance, but as an opportunity for learning about the nature of infrastructure.[10]

In approaching my own empirical work, I have tried to reconcile a case study approach with several key aspects of the methodological discussions rehearsed in the above. For one, I have tried to engage in multi-sited ethnographic work that covers a number of different empirical

10 There are two reasons why such an approach should be applied with caution. First, its usefulness will depend very much on how detailed the resulting empirical accounts are. Given its rather descriptive character, uncritical use of actor-network theory could lead to a superficial portrayal of translation processes, without providing a sufficiently complex account of underlying structural constraints (e.g., of historically grown disciplines, technologies, institutions). Second, the method of emulating the translation behavior of actors should not be reified in its own right, perhaps under the assumption that it captures the *actual* reality of infrastructure development.

phenomena, yet without aiming for data saturation in an empiricist fashion. Moreover I have sought to be attentive to how categories such as 'discipline', 'field' and 'project' are not stable entities, but are themselves performed and subject to ongoing translation activities by various actors (including myself). The first three chapters can be grouped together in that they are based on the interrogation of a specific project (Chapters 2 and 3) or controversy (Chapter 1) within the ambit of digital humanities. In the last two chapters, I define more expansive research objects, namely the role of reflexivity in the discourse of digital humanists (Chapter 4), as well as the development of digital infrastructure in the US and Europe (Chapter 5). A useful methodological strategy in analyzing these different cases was to focus on the very tensions that arise as actors try to reconfigure the scholarly knowledge machine. For example, a characteristic of digital scholarship is that it is often practiced in collaborative projects, whereas knowledge production in the humanities is traditionally organized around the single-author, scholarly monograph as the predominant format. Moreover, different actor groups frequently try to combine digital scholarship with a strategic agenda, but often in ways that are informed by rather specific normative interests. In most of the following chapters, the organizational, epistemic, and institutional tensions that follow from these competing translation activities process appear center stage in the analysis.

In Chapter 1, I study the controversy around the decision of the Royal Netherlands Academy of Arts and Sciences to cease publishing the national bibliography for Dutch Studies in print, and to instead transform it into a digital database with a new set of functionalities and a reduced budget. Many practitioners of Dutch Studies perceived this as a threat to their field, thus prompting them to publicly insist on the importance of keeping the bibliography in its original form. This debate is not of the same type as investigated in older STS research in controversy studies, which focused on scientific arguments about whether or not a particular experiment has been successfully replicated, or what constitutes definitive empirical evidence for a given knowledge claim (Collins, 1985; Nelkin, 1995). Rather, the controversy around the bibliography is a clash between different ways of defining an area of scholarly inquiry, occasioned by an overarching discourse of the 'encounter' between the humanities and digital technology. My account of these discussions balances the presupposition of certain units of analysis (for example that of research practice and discipline) with a sensibility towards the way actors themselves leverage and contest such boundary-drawing.

In Chapter 2 I study how a group of 120 literary scholars from across Europe struggled to combine their university-based, disciplinary careers with participation in an international project in digital literary history. The goal of this initiative was to create a comprehensive empirical picture of 'forgotten' women writers through the collaborative use of an online database. Of particular analytical interest for the present thesis were so-called training school events in which the project participants learned how to use and ingest information into the database. These meetings were an occasion in which the established social and intellectual organization of literary history was challenged, for example insofar as theoretical concepts that are usually left to the discretion of individual scholars in the monograph-oriented model of research had to be operationalized in a consensual manner, so that the project as a whole could define shared analytical categories. Instead of simply enhancing established practices through the use of a database, the ensuing discussions brought to light numerous conflicts between the current infrastructural configuration of scholarly practice and the envisioned model of digital collaboration.

In Chapter 3 I study a project aiming to investigate regime shifts in contemporary Indonesian history through the algorithmic analysis of comprehensive corpora of digitized newspapers. A joint undertaking by scholars of Indonesian Studies, networks researchers, and computer scientists, it provided a good opportunity to study the relation between tool development and disciplinarity. More specifically, my analytical focus is on how the participants tried to devise a shared project workflow that allows for interdisciplinary collaboration, yet without overriding domain-specific epistemic conventions. This illustrates how 'the digital humanities project' is not some readily defined organizational entity that scholars simply 'join'. Instead, projects constitute a distinct new format of knowledge production that emerges from the reflexive attempt to balance the participants' diverging disciplinary commitments and career interests.

In the last two chapters, the methodological insight that the activities of STS researchers and the actors they study are essentially isomorphic has actually provided a foundational conceptual inspiration. In Chapter 4 I analyze how digital scholars engage in a reflexive discourse on the conditions of their research, with the aim of changing how contributions to scholarly knowledge are produced and associated with each other. Here I make the point that the circulation of reflexive arguments by digital scholars is not essentially different from the 'deconstruction' or 'inversion' of scientific practice as performed by STS scholars (e.g., Bowker & Star, 2000) – reflexive

representations are attempts to untie existing 'standardized packages' of socio-material activity (Fujimura, 1987; 1992), thus potentially changing the very structure of 'the field' over time.

Chapter 5 compares European and US approaches to developing digital infrastructure. Here I take an empirical look at various infrastructure projects and the policy discourse surrounding them. Different visions of infrastructure, I argue, can also help establish paradigmatic kinds of logic in how actors think about 'the field'. Once instantiated in funding frameworks and managerial structures, they tend to facilitate for example certain judgments about the relative similarity/difference of individual digital research tools.

In the two case studies that involved ethnographic work, I have tried to use my presence as an ethnographer as a methodological device, rather than treating it as an epistemic contamination of the field. An important aspect in getting access to both projects was the intellectual reputation and visibility of an institution I was formerly affiliated with, namely the Virtual Knowledge Studio for the Humanities and Social Sciences (VKS), which until its dissolution in 2011 was a very active player in digital scholarship in the Netherlands (The Virtual Knowledge Studio, 2008). Using my VKS relations to make contact with the project leaders, I introduced myself as an aspiring STS researcher, thus emphasizing my disciplinary expertise in theorizing and practically moderating the implementation of digital research tools in the humanities. In this sense, I drew advantage from the intellectual capital accumulated by previous generations of STS researchers when positioning myself as a participant observer. This does not mean that I had full control over how I was perceived by different actors within the projects. Sometimes, my work was perceived as a sort of requirements engineering research that would make implicit aspects of scholarly practice visible. At other times, I was seen more as a project therapist who drags suppressed group issues out into the open. Not least, project participants used my interest as a means of promoting their work to funding bodies and peers, according to the logic: 'we even managed to attract the attention of an STS researcher!' These experiences have informed my analysis insofar as they sharpened my awareness for the manifold forms of articulation work necessary to bring unconventional forms of scholarship to closure, both among the participants of a project and between the project and external actors such as funding bodies and evaluators (Fujimura, 1987; 1992).

Here it is worth noting that the article-based PhD format conditioned the interaction between my own knowledge and that produced

by the actors I studied in a rather specific way. For example, during my fieldwork in the above mentioned project in literary history (Chapter 2), I circulated a well-developed draft of my argument on the simultaneously technical and intellectual dimensions of datawork about two years into the project. The relatively formal character of the document - I distributed it with an explicit announcement that I was planning to submit to a peer-reviewed STS journal – arguably contributed to the attention that the participants paid to my findings. Had I been working on a monograph, they might have received my findings only in the more ephemeral and less 'weighty' form of a presentation at a project conference, or at best through a very early draft of a prospective book chapter that would be submitted only a few years later. The specific rhythm and format of an article-based thesis thus also affected how my own research acquired agency within the projects I studied. Not least, my peer-reviewed articles have become part of the way both projects present themselves to their funders, in the sense that they are listed as 'output' in the concluding evaluation reports.

Sometimes, my arguments were also read as a critique of managerial decisions or positions adopted by individual actors in the two projects. My strategy in maneuvering such situations was to be explicit about the partiality implied by my theoretical framework. Against the conceptual backdrop of infrastructure studies, the various kinds of friction that occur in the current institutional and praxeological re-organization of scholarship appear as a result of structural conflicts – conflicting career demands (do what is necessary to run a database project vs. do what is necessary to become a full professor in a university department), conflicting ways of making knowledge claims in different fields (algorithmic claims in network research vs. narrated, hermeneutic arguments in Indonesian Studies), and different disciplinary ways of performing epistemic concepts such as data (data as a highly standardized entity vs. data as highly specific to individual practices). Had I chosen a different conceptual approach, say, theories drawn from usability engineering or management-oriented anthropology, I might have explained these tensions as a result of poorly designed interfaces (Santos & Frankenberg-Garcia, 2007) or failure to create shared repertoires of communication between project participants (see Contu & Wilmott, 2003). In this sense, my research also co-produces the overarching theme that it professes to elaborate, namely the notion that digital scholarship will only be successful it is developed with a reflexive sensibility to the evolution of the socio-material infrastructure that constitute the humanities. I would maintain, however, that this approach is a useful second step after the more

hyperbolic claims that are initially necessary to mobilize funding and political will for digital scholarship, as well as a useful complement to perspectives such as usability engineering. My analysis provides a context-sensitive means of framing the inevitable conflicts that arise when researchers begin to experiment with new forms of knowledge production. This makes the conflicts amenable to discussion, rather than preemptively dismissing them as expressions of intellectual conservatism or lack of collaborative spirit.

Table 1: Chapter Overview

	Pub-lished	Title	Co-author	Journal	Methods
1	2010	E-research and methodological innovation in Dutch Studies	Paul Wouters	*First Monday* 15(9)	Semi-structured interviews, document analysis
2	2014	Scholarly labor and digital collaboration in literary studies	-	*Social Epistemology* 29(2), pp. 207-233.	Participant observation, semi-structured interviews, document analysis
3	2014	Decomposition as practice and process: creating boundary objects in computational humanities	-	*Interdisciplinary Science Reviews* 39(2), pp. 143-161.	Participant observation, semi-structured interviews, document analysis
4	2014	Infrastructural Inversion as a generative resource in digital scholarship	-	*Science as Culture* 24(1), pp. 1-23.	Document analysis
5	-	Digital infrastructure in the humanities: reconfiguring the organization of scholarly tool development.	-	Under review at *Computer Supported Cooperative Work*	Semi-structured interviews, document analysis

Chapter 1

E-research and methodological innovation in Dutch Studies[11]

E–research and the humanities

It is well known that innovations in data collection and analytical instruments have regularly spawned new scientific and scholarly fields (Beaulieu, 2001; Lemaine et al., 1976; Shinn & Joerges, 2002), e.g., imaging technologies have led to radical innovations in medical, cognitive and neurosciences. Techno–opitimistic stories about the revolutionary potential of e–research applications (Atkins et al., 2003; Hey et al., 2009) seem to fit the picture of an innovative research technology with far–reaching consequences for the cognitive, social and material aspects of the sciences (Joerges & Shinn, 2001). E–research promises to enhance and innovate research in a number of regards: by facilitating cost-efficient, distributed access to large datasets, by providing the computing power necessary to process these data (e.g., through grid computing), and by facilitating collaboration across disciplinary and geographical boundaries (Jankowski, 2007; Wouters, 2006). The concept of e–research emerged in natural and biological sciences such as particle physics, astronomy, meteorology, and DNA research, and its characteristic features are tailored to the needs of quantitatively oriented, collaborative fields of research (Jankowski, 2007).

But what does e–research mean for interpretative social sciences and humanities? How are the dynamics in these fields influenced by technological and managerial innovations in research instrumentation and infrastructure? And how does this impact the identity of the field and its practitioners? To shed light on these questions we study the controversy around the recent digital innovation of the *Bibliografie van de Nederlandse Taal– en Literatuurwetenschap* (BNTL), a well–established bibliographical tool for Dutch Studies, i.e. the academic field concerned with Dutch language and literature. As we will show, the digitization of the BNTL is representative of many implications of e–research for the humanities.

The history of the BNTL is intimately connected to the disciplinary

11 This chapter has been published as: Kaltenbrunner, W., and Wouters, P. (2010) E-research and methodological innovation in Dutch literary studies, *First Monday*, 15(9). The present version is slightly abridged and contains minor stylistic changes in comparison to the published article.

history of Dutch Studies, and many practitioners used to regard the bibliography as an important tool for research. In 2004, the Royal Netherlands Academy of Arts & Sciences announced that funding for the bibliography would be decreased by more than 50 percent. The Royal Academy also decided that the BNTL should be no longer published in print, but in the format of an online database. As soon as this plan became public, a number of practitioners voiced their concern about the impact of this decision on everyday scholarly work routines and the future of Dutch Studies as a discipline.

Our paper will try to understand the innovation and the discussions accompanying it on two analytical levels. First, we will analyze how the innovation affects research practices in Dutch Studies. Second, we will investigate the implications of the digitization for the way practitioners think about themselves as scholars. Analyzing the transformation of a key research instrument on these two levels provides us with a first impression of the co–construction of scholarly knowledge, practices and identities through the implementation of technological and managerial innovation. We derived the most important sources for our study from written documents and qualitative interviews with members of the BNTL editorial team, scholars of Dutch Studies, and policy makers, all of which were conducted between September and December 2008.

History of the BNTL

The BNTL was first published in 1970, following a grassroots initiative to identify and make accessible a canonic body of scholarly works in Dutch and Flemish literary studies and linguistics.[12] The composition of its editorial staff fluctuated over the years, but usually consisted of five editors with a degree in Dutch Studies, and two university–trained documentalists (Baars et al., 2004). The BNTL was a retrospective disciplinary bibliography in the traditional sense. Individual cumulative additions were published on an annual basis, ordering relevant scholarly sources according to an elaborate decimal categorization. The editorial team simultaneously extended coverage backward and forward in time, ultimately encompassing the period from 1940 to 2004. From 1993 on, the BNTL database could be accessed online via university library portals and as an MS DOS or Windows version, with the print edition being published in parallel (Doorenbosch,

12 Personal interview with Elly Kamp, 26 November 2008, The Hague.

1993). Originally an independent organizational unit within the Academy, the BNTL was in 2005 taken over by the Huygens Institute, an institute specialized in high–quality editions of historical texts in science, philosophy, and literature. Funding for the BNTL was subsequently reduced from 5,7 FTE to 2 FTE (Baars et al., 2004). The Royal Academy furthermore decided that the BNTL should no longer be published in print at all, but exclusively as an online database. While the editorial team previously guaranteed comprehensive coverage of relevant sources, the bibliographical dataset of the digital BNTL is now limited to a list of core journals. Articles appearing in these journals are automatically added, thus making users independent from the publication rhythm of the old print bibliography. Monographs, however, which still constitute a very important publication format in Dutch Studies, are no longer indexed in a comprehensive fashion. To make up for this, registered users now have the possibility to add publications themselves, which are then double-checked by the editors on a weekly basis. Another change is that the decimal categorization system of the print BNTL has been replaced by a new online query form, as for example used by digital library catalogues. Moreover, users have the possibility to inspect abstracts and access full texts of publications if available (Huygens Instituut KNAW, 2004a; 2004b).

The announcement of changes to the BNTL led to a controversy in which many practitioners of Dutch Studies as well as members of the editorial staff expressed their strong disapproval. One of the critics even called for a collective publication strike (Verkruijsse, 2005), and the Dutch Minister of Science and Education attempted to directly intervene at the Royal Academy by an open letter (Verkruijsse, 2004). Major points of critique raised against the innovation concerned the reduction and automation of bibliographical coverage. Another controversial aspect was the original plan to completely exclude publications in modern Dutch linguistics from the bibliographical dataset. The Huygens Institute reacted by setting up an advisory board of external users who were invited to participate in the digitization project. Among them were also some of the most outspoken critics of the changes. In response to the fierce criticism, the plan to exclude modern Dutch linguistics was finally abandoned. While one of the original critics who had joined the board reaffirmed his objections in our interview, another one indicated that the advisory board meetings had given him a better idea of the changes introduced, thus mitigating his original concerns. The new BNTL Web site was officially launched on 24 April 2008.

The BNTL in different research practices

The implementation of e–research tools in the Netherlands is linked to attempts by policy makers and individual academics to stimulate a methodological innovation in how science and scholarship is practiced. On the one hand, e–research is about enhancing knowledge production by bringing together and facilitating access to existing datasets in a centralized virtual environment, thereby enabling researchers to pursue wholly new lines of inquiry (KNAW/NWO, 2004). Another expected benefit is that the use of ICT will make research more cost-efficient, in that it will allow to automate many tasks previously carried out by humans. With respect to textual scholarship specifically, policy makers and e–research advocates often express the hope that the use of digital tools will encourage scholars to move from narrowly circumscribed research topics (e.g., the production circumstances of a single literary work, or the way a classic literary leitmotif is treated by a single writer) to larger scale comparative research (e.g., a comparison of production circumstances of many literary works across different countries, or a comparative international history of a given leitmotif) based on a strong basis of hard empirical data. In the following, Henk Wals, the Director of the Huygens Institute, exemplifies the characteristic advantages he expects of e-research on the basis of a recently developed tool for collaborative annotation:

> We have recently developed a tool called eLaborate. On the one side of the screen you have a digital facsimile of a medieval manuscript, on the other side you can insert a transcription and annotations. That's a Web–based tool, meaning that whole teams of researchers can simultaneously transcribe and annotate a text, and share their annotations. This allows not only to translate a text into machine–readable form quickly and efficiently, but also to create a research tool, a text which is constantly enriched, to which data are constantly added. (…) if you are a literary scholar dealing with a specific question in a project, which usually run for three or four years nowadays, then you can only do so much work on your own, only a limited number of texts at one time. In other words, it is always a sort of sample that you take. On the basis of a relatively small number of sources you try to draw a more generally valid conclusion. (…) But if it becomes easier to pose the same question to a larger corpus of texts then your research becomes much better grounded. If you then also take advantage of quantitative methods,

measure word frequencies, etc., you take another step towards more objectivity.[13]

The use of databases and other digital tools in various scientific and scholarly domains has in recent years become a topic of study for researchers in Science and Technology Studies (STS), information science, and Computer Supported Cooperative Work (Beaulieu, 2004; Borgman, 2007; Bos et al., 2007; Bowker, 2000; Hilgartner, 1995; Hine, 2006). Research foci and analytical approaches vary significantly, however. Publications in information science for example often provide descriptive accounts of the proliferation of ICT across the sciences, thus implicitly suggesting an inevitable epistemological development towards ever more data-intensive, ever more networked modes of research (Nentwich, 2003; Borgman, 2007). In such a view, digital databases and other tools tend to be treated as readily black–boxed instruments that transform scholarly practice by virtue of inherent technological potential.

In the perspective of the more ethnographically and anthropologically oriented approaches to e–research, the unit of analysis normally is the interaction of disciplinary culture, users, and technology (Beaulieu, 2004; Bos et al., 2007; Davenport, 2001). This line of inquiry stresses the embedding of tools in individual research practices, implying that the shaping of e–research technology follows a logic of social construction (Hine, 2006; Bijker et al., 1987). In such a perspective, the question as to whether tools like the digital BNTL will indeed lead to a more efficient organization and methodological enhancement of scholarship depends not on inherent technological features, but on how well practitioners manage to integrate them with the specific cognitive and praxeological needs of their research.

In investigating the role of the BNTL in the work routines of scholars, we take theoretical inspiration from Karin Knorr Cetina's (1999) concept of epistemic cultures. Knorr Cetina's theory was originally developed to study knowledge production in laboratory sites in the natural sciences, but can also be applied to textual scholarship. It allows us to relate the use of technologies in everyday research practice to issues of heuristic interest and epistemology. The concept of epistemic cultures describes research practice in terms of three characteristics: the way researchers construct their objects of study; the way they experimentally validate knowledge; and the way

13 Personal interview with Henk Wals, 20 October 2008, The Hague (my translation).

epistemic units in a research site are related to each other (Knorr Cetina, 1999). A particular category of factors can only be analyzed with respect to the configuration as a whole. Conceptual frameworks for example shape technological instruments for research, which are used in turn to validate knowledge and thus reproduce the overarching conceptual structure. Symbolical, material, and social aspects of an epistemic culture are seen as interrelated in a specific configuration. Changing one constitutive aspect, such as a specific research tool like the BNTL, may result in a reconfiguration of the epistemic culture, but perhaps in ways not originally anticipated.

We adapt Knorr Cetina's concept to our own case in the following way. Under symbolical aspects, we subsume characteristic research questions (e.g., "when, where, and by whom was this particular literary manuscript written?"), underlying theories and theoretical assumptions (e.g., "linguistic analysis of texts allows to infer statements on its production process"), and methods (e.g., the comparison of different sets of empirical material) in Dutch Studies. Material aspects comprise tools and empirical sources for research, i.e., libraries, textual corpora, and specific instruments like the BNTL. As regards the social aspects, research and writing in literary studies has traditionally been organized as a solitary endeavor, although one of the expectations towards e–research is that it will bring about a more collaborative form of scholarship.

Dutch Studies is a continuum of very different research practices, rather than a methodologically and theoretically homogeneous field. Traditional ways of ordering these practices are to group them either according to the object of study (e.g., the writer investigated; the literature of a given historical period) or according to the methodological approach taken (e.g., quantitative reception studies). We decided that it is most insightful for the purpose of this paper to focus on a particular object of study, Dutch literature of the late medieval and early modern period. More specifically, we will discuss three distinct approaches to older Dutch literature as professed by three individual researchers. This allows us to give an overview of the bandwidth of techniques deployed to study a single topic, and of the different functions of the BNTL in these research practices.

On one side of the continuum of research practices in medieval/early modern Dutch literature is analytical bibliography, as practiced by Professor Piet Verkruijsse. Analytical bibliography studies the genealogy of texts as material artifacts. By collating variants, i.e. unauthorized or corrupted editions of early modern printed texts, analytical bibliography aims to

establish the original textual shape as intended by the author. Bibliographical tools, especially old library catalogues, potentially index forgotten copies and can thus help to lead the way back to the original version. For the researcher to stay on top of things, relevant bibliographical databases need to be timely updated and as comprehensive as possible. This goes also for bibliographies of academic publications like the BNTL, insofar as they trace the scholarly progress towards the original textual shape.[14] Verkruijsse welcomes the perspective of facilitated, always up-to-date access to academic publications through online databases like the BNTL. At the same time, he expresses strong concern about the fact that coverage of scholarly publications in the digital BNTL will be limited to a list of core journals, and that it will no longer be provided by a human editorial team. The BNTL does in his perception no longer fulfill the function of delineating and identifying a body of relevant knowledge. Verkruijsse recurrently drew a comparison between the innovated digital BNTL and Google to summarize the combination of facilitated access to sources on the one hand, and of less rigid structuring and quality control on the other.

Another approach to studying old Dutch literature is to look at its reception. The research of Professor Paul Wackers aims to reconstruct the reception of late medieval/early modern texts by historically contextualizing them in contemporary social and aesthetic norms. Texts as material artifacts constitute an essential part of this research practice, insofar as individual copies and editions may give hints about the social status of readers, their reading habits, or the way they received a particular piece of literary writing, e.g., through hand-written annotations in the margin. Wackers stresses the difference between his own research and more normative 19th century approaches to reception studies, which were based on the idea of an allegedly ideal way of interpreting a specific text.

> 19th century philologists thought that there was a general human quality expressed in cultural artifacts that could be discovered by a good researcher. All medieval things were valued according to the standards of 19th century aesthetic ideals, because those were held to be a general standard. That has changed. We have abandoned the idea that there is one standard for literature and culture and we are now trying to investigate the mindsets of medieval people in a more

14 Personal interview with Piet Verkruijsse, 10 December 2008, Amsterdam.

unbiased way.[15]

This hermeneutically oriented approach implies different ways of going about empirical work and validating research findings when compared to analytical bibliography. The latter depends on constant updating and comprehensiveness of bibliographical datasets for identifying a touchstone of relevant knowledge, and it implies a strong concern with the quality and depth of bibliographical source criticism. In hermeneutic reception studies, by contrast, the BNTL is considered one way among others to collect scholarly sources. Wackers regularly uses the bibliography for browsing topical publications, but complete coverage is not an epistemological sine qua non. Since Wackers is not interested in giving exact answers to highly specific research questions, validating findings for him is more a matter of creating intersubjectivity, in the sense of being explicit about the sources and research methods used. While Wackers was initially opposed to the digitization, he has since tended to accept the conceptual changes: "I've seen a list of journals they wanted to cover and I would say that 95 percent of what is important is automatically covered. I can live with that."[16]

The research of Karina van Dalen–Oskam, who is also the leader of the BNTL innovation project, consists in linguistic analysis of old Dutch texts. Of particular importance to this research practice are digital tools for the analysis of rhyme patterns, word frequencies, and syntactical structures. While linguistic analysis can also be conducted manually, the adoption of digital tools in recent decades has significantly expanded the empirical scope of this line of research. Findings here are validated through sophisticated quantitative methods, based on large textual corpora. The BNTL itself does not fulfill a particularly important role in this research practice, since most relevant journals are well covered in other databases, such as the Web of Science. The linguistic research community is generally more internationally oriented than other sub-areas of Dutch Studies, thus making practitioners less dependent on a body of canonic national knowledge as provided by the BNTL.

Technological innovation and disciplinary identity

In her influential 2008 study, Hine argues that the broad adoption of ICT in

15 Personal interview with Paul Wackers, 15 October 2008, Utrecht.

16 Ibid.

systematic biology over the last years has been linked to a reflexive repositioning of the discipline. A field concerned with classifying organisms and exploring their evolutionary relationships, systematic biology has attempted to get rid of its image as an archaic taxonomizing endeavor, and thereby save itself from neglect and underfunding. Practitioners instead have strived to re–imagine systematics as a technologically sophisticated and competitive modern science, a process that is in turn linked to the discourse on biodiversity. Institutions in systematic biology have recently presented themselves as providers of crucial information for the preservation of botanical and zoological species, with the spread of digital networks providing an ideal means to make this information widely accessible. Instead of seeking to capitalize on its robust taxonomic methodology, as in the past, systematics is now eager to prove its relevance as a discipline by catering to enlarged lay and professional audiences (e.g., interested amateurs, other biological sub–fields, museums, biodiversity–rich developing countries). Hine (2008) emphasizes in her analysis of these developments that e–research is not a rigid concept whose implementation straightforwardly transforms a scientific field according to a singular underlying model of data-intensive research. E–research rather figures as a sort of prism through which policy makers and individual researchers re–imagine the goals, methods, and also the history of their discipline. It seems that the adoption of e-research tools in Dutch Studies is related to a similar reflexive discourse about the identity and function of the field in an era in which the relevance of humanities scholarship is regularly questioned. The controversy around the digitization of the BNTL in particular has sparked an emotional debate in which different groups of actors express hopes and anxieties regarding the development of Dutch Studies in the near future.

In this section, we attempt to interrogate the elusive notion of 'disciplinary identity' by looking at how academics speak and think about themselves in terms of the following aspects: research methodology, embodied skills, and the cultural and geographical situatedness of research. The BNTL is bound up with the performance of disciplinary identity in that it represents and enables certain research methods, in that requires certain skills on the part of the user, and insofar as it delineates the cultural and geographical space in which research is conducted. As we will show, the digitization of the BNTL has affected all three of these aspects.

Dutch e–research initiatives envision future scholarly practices as characterized by data–intensive approaches and increased international and interdisciplinary collaboration. But while the digitization of the BNTL is part

of the attempt to induce a methodological innovation along such lines, a strong motive for resistance was precisely the function of the print BNTL in representing the methodological traditions of the field. Originally, the digitization plan foresaw to exclude modern Dutch linguistics from the dataset (Baars et al., 2004). This raised the controversial issue of the unity of Dutch Studies. In the 19th century, language and literature were thought to spring from the essence of national character, thus providing a powerful reason to subsume the study of both under one discipline. Since then, however, linguistics and literary studies have differentiated into methodologically and theoretically neatly distinct fields. The original plan for the BNTL digitization had meant to acknowledge this separation by excluding modern linguistics from the bibliographical dataset, not least because practitioners of the latter field had been found to rely mostly on other bibliographical databases anyway (Voorbij, 1999). This announcement caused fierce protests on the part of many Dutch scholars, however, who considered it absolutely vital that the BNTL guarantees at least formally the traditional methodological unity of Dutch literary studies and linguistics. Ultimately, this led to an agreement that the revised BNTL would continue to cover also the most important journals in modern Dutch linguistics (Huygens Instituut KNAW, 2006).

The strong symbolic value that many researchers still attach to the BNTL can partly be explained by the important role it occupied in disciplinary education. Training in the use of the print BNTL traditionally formed part of the undergraduate curriculum in Dutch Studies. Knowing how to use the print BNTL was part of being a scholar, and it distinguished members of the disciplinary community from other researchers. One of our interviewees, Paul Wackers, indicated that especially older generations of scholars have internalized the decimal categorization system of the BNTL, and that these categories influence the way they intuitively order and combine information.

> The old BNTL was created by people who indexed titles with keywords. The new BNTL does not do that. It searches full–text everything that can be found in abstracts and titles and so on. And I think this is one of the major differences between older and younger scholars. I have been trained in working in the system of the old BNTL. I have a grid of knowledge in my head and I know that for this I have to use this bibliography, and for that I need to use another bibliography. I think my way of researching and writing is

informed by these man–made criteria.[17]

The relation between the use of the print BNTL and disciplinary identity in Dutch Studies however, began to change when the bibliography became accessible online in 1993. Within the following few years, many users switched to consulting the BNTL through their university library portals (Voorbij, 1999). The recent implementation of a new online query form have made training in the proper use of the decimal categorization system of the print version principally unnecessary. The ability to use the BNTL is no longer a skill by which members of the scholarly community of Dutch Studies can distinguish themselves from 'outsiders'.

A major topic of inquiry in STS have been the implications of e–science for the spatial organization of research (Bos et al., 2007). Lenoir (1998) has for example argued that the use of global digital databases may replace the laboratory as the main site of knowledge production in biology. Hine (2006) in contrast has concluded that biological laboratories and digital databases co–exist as different frameworks for organizing particular aspects of research, complementing rather than replacing each other. The case of the BNTL shows that the displacement of research tools into virtual space potentially creates problems specific to scholarship in the humanities. Bibliographies for a national philology delineate the geographical and cultural context in which research is conducted, and this context in turn is an important factor in determining what counts as valid methods and objects of study. The digitization of a bibliographical tool, and the creation of e–research applications in virtual space, seems to be related to a change in the established distribution and hierarchy of research goals in Dutch Studies.

In his sociological analysis of the French 'academic field', Pierre Bourdieu (1988) argues that a discipline such as the national philology is characterized by an inherent methodological tension between 'softer' and 'harder' conceptions of research, which are related to different societal functions. On the one hand, the national philology is expected to produce original knowledge according to disinterested 'scientific' standards. On the other hand, it has the function to conserve and transmit knowledge about national language and literature. This conservatory function implies a more panegyric attitude of scholars towards national writers and literary texts which potentially contrasts with the 'scientific' function of the discipline (Bourdieu, 1988). While national philology as an agent in the conservation

17 Personal interview with Paul Wackers, October 2008, Utrecht.

and reproduction of national culture is geographically situated, national philology as the scholarly pursuit of new knowledge about language and literature is a more international endeavor.

In the case of the Netherlands, Dutch national philology has witnessed an overall internationalization over the last years. The need to publish at least partly in international journals and to participate in international conferences and events has become an imperative. Scholarship is increasingly evaluated in comparison to the international academic context. Also, funding is more often provided by bodies of the European Union. E–research is by many practitioners perceived to promote the internationalization of Dutch Studies by strengthening the 'scientific' function of the field.

The project leader of the BNTL digitization project for example, Karina van Dalen–Oskam, points out a relation between the geographical/cultural context in which research is conducted, and the epistemic goals and methods that are considered appropriate. Scholars of Dutch literature addressing a national audience may reasonably presuppose readers to be familiar with Dutch literary history, and can hope to attract attention by interpreting the content of the works investigated. The cultural value of Dutch literature for a national audience here legitimizes a rather interpretive and hermeneutic approach. Scholars addressing an international audience on the contrary will not be able to legitimize their work simply by virtue of the cultural value attached to their objects of study. In comparison to writers of 'world literature', Dutch literature and language are relatively little known abroad. The work of W.F. Hermans for example, one of the most important Dutch writers of the 20th century, and a particular personal interest of van Dalen, has for the most part not been translated into English. In van Dalen–Oskam's view, Dutch Studies as a field should make up for the lack of cultural capital of its research objects in an international context by capitalizing on 'scientific' virtues of empirical exactitude and objectivity, and through the use of sophisticated technology. As a model for a more internationally relevant scholarship, van Dalen-Oskam points to the type of research she engages in herself, i.e. linguistic studies of early modern Dutch texts that leverage large amounts of data, and that are consistently published in English-language journals.

The digitization of the BNTL was on the other hand perceived as a potential threat for the conservatory function of discipline. Apart from the possibility to implement new functionalities such as full–text search and a collaborative component, an important reason for transforming the BNTL

into an online database were of course budgetary considerations. Replacing manual bibliographic work by an automatic coverage system allows for the database to be maintained by fewer and less–skilled personnel.[18] Perhaps unsurprisingly, the combination of digitization with cost-cutting has had a powerful psychological effect on some practitioners of Dutch Studies. Critics perceive it as proof that the disciplinary function of cultivating national literary heritage is no longer valued by the Royal Academy. Book historian Piet Verkruijsse refers to the changes in the BNTL in terms of a metaphor of globalization — an established, national quality product is replaced by a cheap replica manufactured in low–wage countries (Verkruijsse, 2005). Two of the practitioners we interviewed made clear that they do not consider it part of their job to add their publications to the digital BNTL, if those publications are not automatically covered. Guaranteeing a comprehensive national bibliography in their view is something that the Dutch state should fully support through public funds, since it falls under its responsibility for national cultural heritage more generally. Interestingly, a BNTL documentalist we interviewed indicated that many lay users, for example amateurs interested in contemporary and historical Dutch literature, already make use of the possibility to add publications to the BNTL dataset. This is in stark contrast to professional academic users, who rarely upload any bibliographical information or full scholarly sources.[19]

Discussion and conclusions

Popular accounts of e-research suggest that the adoption of data-intensive, networked research tools will bring about a simultaneously more cost-efficient and more powerful way of producing scientific knowledge (e.g., Atkinson, 2006; Nentwich, 2003; see also Hine, 2008). But while the perspective of collaborative work and the use of larger amounts of quantitative data merely extrapolates the methodological precepts of many natural sciences, it implies a tension with the strong grounding of most scholarly disciplines on qualitative approaches (Wouters, 2006; Wouters & Beaulieu, 2006). To better understand the implications of e–research for the humanities, we have analyzed the recent digitization of the BNTL, a long–standing bibliographical tool for Dutch literary studies and linguistics.

Our first point of interest was the question as to how the digitization

18 Anonymous personal interview, 10 December 2008, The Hague.

19 Ibid.

affects everyday research practices, and also what possible inferences can be made regarding the adoption of digital tools by scholars more generally. Adapting Knorr Cetina's (1999) concept of epistemic cultures, we investigated knowledge production in Dutch Studies as an interplay of research questions, theoretical frameworks, and epistemological assumptions, mediated by material tools. Our comparison of research practices in the area of old Dutch literature has revealed a plurality of ways in which the bibliography is used. These ranged from providing a way to identify relevant sources to an epistemological function in validating findings. The degree to which research practices depend on specific tools more generally seemed to correlate to the degree of epistemological exactitude researchers aim for in the results they produce. Research practices aiming to provide very exact answers to research questions (e.g., 'Which one of a range of surviving copies of an early modern printed text is the oldest one?', or, 'What linguistic patterns can be deduced from this corpora of early modern Dutch poetry?') use bibliographies and tools for linguistic tools in an experimental way, i.e., to corroborate or refute hypotheses. Epistemologically softer practices such as hermeneutic reception studies pose questions that cannot be answered with the same claim to exactitude, and bibliograhical instruments such as the BNTL provide one way among others to collect scholarly sources.

These exploratory observations suggest that the implementation of e–research tools will unevenly affect the different scholarly approaches in Dutch Studies. The rather exact, technologically dependent practices are more likely to be affected by e–research than the ones leaving larger leeway for interpretation of results. But also in the case of the more technology–dependent approaches, specific predictions about the effects of proliferating digital tools are difficult. For example, the digitization of the BNTL has replaced extensive manual data curation through a human editorial team by a system automatically covering a list of core journals. While fast, continuous updating and the possibility to conduct full-text search on parts of the dataset constitute an undeniable benefit for all users, the reduction in overall coverage is very detrimental for some areas of study, such as analytical bibliography. Further empirical and conceptual work is necessary to unpack the implications of digital approaches in specific research contexts.

A second point of interest was the question as to whether and how the spread of digital tools in the humanities is related to changes in the performance of disciplinary scholarly identity. Much like the case of systematic biology presented by Hine (2008), the implementation of e–

research tools in Dutch Studies does not take the shape of centrally controlled process with a predetermined outcome, but rather of an emotional argument about the very essence and function of the field in the early 21st century. More specifically, the controversy around the digitization of the BNTL touched upon three aspects of disciplinary identity: research methodology, skills/tacit knowledge, and the geographical/cultural space in which research is conducted.

On the one hand, the implementation of e–research in the Netherlands is shaped by the vision that the spread of digital tools will promote more collaborative, data-intensive approaches also in the humanities. Well–established research instruments, however, may represent methodological traditions of a discipline in ways that clash with the intended innovation. The initial plan for the digitization of the BNTL acknowledged the de facto differentiation of Dutch literary studies and linguistics over the past 150 years by excluding publications in modern linguistics from the dataset. This prompted fierce resistance of many practitioners, who considered it crucial that the bibliography of national philology continues to formally represent the historical unity of the two fields.

The digitization of a tool like the BNTL also entails a change in the skills required on the part of the users. In the past, aspiring scholars of Dutch Studies were trained in the use of the print version of the bibliography, in particular its elaborate decimal categorization system. The ability to navigate this system was distinctive of disciplinary culture. By contrast, anybody familiar with digital library catalogues and online search engines can use and contribute to the digital BNTL. The BNTL has thus become a site of collective knowledge production that weakens the boundary between specialists and laymen. The fact that lay users have so far taken much more advantage of the collaborative element than academic researchers would imply that the former are more enthusiastic about this 'opening' of knowledge production than the latter.

The digitization of the BNTL was also perceived as an element in the process of internationalization of Dutch Studies, which is in turn related to a change in the hierarchy and distribution of research goals. Some practitioners associate the methodological innovation expected of e-research with internationally valid 'scientific' virtues, thus suggesting that digital approaches may be useful to promote Dutch literary scholarship among academic audiences abroad. Critics associated the digitization with a demotion of the disciplinary function to conserve and mediate knowledge

about Dutch literary heritage. The case of the BNTL thus illustrates a tension specific to the implementation of digital tools for the humanities in countries like the Netherlands. Scholars understand that they increasingly need to participate in an international academic community, for which the adoption of digital approaches seems to be ideal. The displacement of research tools into virtual space, and the increased focus on research per se, however, may in turn conflict with the pronounced need to cultivate the cultural heritage of an otherwise little studied, small language community.

Postscript to chapter 1

In this first chapter I have argued that the controversy around the digitization of the disciplinary bibliography of Dutch studies has not just been a technical discussion about desirable features and practical design choices. Rather, it has been one about how scholarly work itself should be organized in terms of methods, research goals, and relevant audiences, and how the humanities may be best served through the adoption of new tools. Illustrating the subtitle of this thesis, digital technology has served as a refracting lens through which practicing academics, as well as a variety of other actors, such as administrators and policy makers, began to reimagine what it means to do scholarship in Dutch studies.

The chapter does not portray a transition towards a singular new model of scholarly work, however. Instead, some scholars were extremely critical of the newly introduced features of the bibliography, while others tended to embrace the changes rather quickly. These heterogeneous reactions can be explained by combining the perspective of infrastructure studies (Star & Ruhleder, 1996; Edwards, 2010) with a sensibility for the epistemic and organizational differences between individual scholarly specialties (Beaulieu & Wouters, 2006; Knorr Cetina, 1999). The field of Dutch Studies can then be seen as an ecology of disciplinary subcultures, each characterized by a unique set of properties. These subcultures are interrelated through their shared history, material tools and embedding in academic institutions, but rather loosely integrated in terms of research practices and conceptual frameworks. From such a vantage point, particular technological affordances of a digital bibliography, such as participatory features or immediate updating, are not inherently useful. Instead, they acquire their meaning in relation to the specific research goals and methods of their users. If we take into account the intellectual and methodological diversity of Dutch Studies, it is not surprising that different practices are affected very unevenly by the digitization. Individual opinions differ as strongly as the variety of approaches in the field – this is arguably different from the adoption of tools in comparatively more integrated fields in the natural or quantitative social sciences.

The case also introduces a related aspect of the infrastructure perspective, namely issues surrounding the conceptualization and valuation of particular forms of work in a delicate balance of mutually sustaining task areas (Star & Ruhleder, 1996; Strauss & Star, 1999). At one level, the controversy around the digitization can in fact be read as an altercation about what type of activity the work of bibliographical data ingestion and

quality control actually is. By reducing expenditure and turning that work at least partly into a crowd-sourced responsibility of scholars and interested lay users, the Huygens Institute has effectively redefined a publicly subsidized infrastructural service as an activity that overlaps with the core tasks of university-employed scholars. It would be wrong to read the subsequent protests of academics only as a reaction to the perceived loss of disciplinary prestige. Instead, re-drawing the boundary between technical and scholarly responsibilities also had tangible negative consequences for the everyday conduct of scholarly work in some specialties. Analytical bibliography for example (and possibly other areas of study not covered in the chapter) constitutes an epistemic subculture that is particularly reliant on well-curated and extensive bibliographical information. Continuing to work according to the conventions of this specialty requires practitioners to make up for the reduction in editorially warranted coverage through their individual effort. The digitization thus affects how easily scholars can produce certain forms of knowledge, potentially leading them focus on different sorts of research questions in the future.

In the next chapter, I will refine these first impressions by probing a very different empirical case, namely a grass-roots initiative in digital literary history. This provides me with an opportunity to study the adoption of digital approaches in a context where the intended innovation of scholarly methods and practices is not driven by managerial intervention, but emerges directly from within the intellectual dynamics of a field. Moreover, the case study will allow me to draw out the organizational implications of a core promise often associated with digital research technology – that of harnessing data-intensive research methods in hermeneutic fields of inquiry.

Chapter 2

Scholarly labor and digital collaboration in literary studies[20]

Parenthesis – relation to conceptual framework

Readers of the following chapter may initially be struck by a specific terminological choice, namely my consistent reference to scholarship as labor. This term emphasizes the expenditure of mental and physical resources, and is therefore at odds with the popular idea of research as a disembodied, purely cognitive activity. My use of the word labor is an intentional attempt to radicalize a common theoretical abstraction applied by STS scholars since the late 1980s - that of research as practice (cf. Pickering, 1992). Posited as a challenge to influential mid-20th century accounts of science by philosophers (Karl Popper) and sociologists (Robert K. Merton), the notion of practice cuts through research in such a way as to avoid foregrounding idealized epistemological concepts, as well as reducing the sociology of science to a sociology of individual scientific careers. STS research on science as practice tends to draw attention to recognizably different issues, for example the material mediation of epistemological concepts in lab work (Galison, 1997; Rheinberger, 1997), and the sociotechnical translation processes involved in turning instrument readings into authoritative statements (Latour, 1987; Fujimura, 1992). The concept introduces new blind spots in its own right, however. For one, it tends to deemphasize the function of social, political, and institutional macrostructures, instead picturing knowledge production as a matter of situated, emergent cognition (Vann & Bowker, 2001). Secondly, and as a consequence of this, it neglects questions about how the economic cost of doing research mediates its intellectual substance (The Virtual Knowledge Studio, 2008). Much STS scholarship operating with the notion of practice is in fact based on the methodological assumption that everything relevant about scientific work can be grasped through ethnographic descriptions of the culture, social interaction and concrete physical acts of individual researchers in particular sites.

The notion of research as labor as I use it here is geared to draw attention to the mutual dependence of different forms of work in a larger

20 This chapter has been published as: Kaltenbrunner, W. (2014) Scholarly Labour and Digital Collaboration in Literary Studies, *Social Epistemology* 29(2), pp. 207-233.

infrastructural production process, and especially the way this production is constrained by the economic valuation of individual task areas. The preceding chapter has already shown how a managerial strategy to effectuate budget cuts through the reconceptualization of certain task areas can affect knowledge production. Specifically, I have shown how the digitization of the BNTL has made it more difficult to do research according to the disciplinary conventions of analytical bibliography. The following chapter further pursues this line of inquiry. It investigates in detail the implications of digital instruments for the conceptualization and distribution of different forms of labor that together enable the 'primary' process of scholarship in literary history, and the way that changes in this distribution affect the particular forms of knowledge that can be generated.

The chapter also extends the analysis in another way. The controversy around the digital bibliography of Dutch Studies presented us with a confrontation between research managers and policy makers on the one hand, and practicing scholars in various subdisciplines of this field on the other. If read as an isolated study, the preceding chapter therefore might be taken to imply that boundaries between task areas are stable and undisputed within a given specialty (thus suggesting that scholars of analytical bibliography, reception studies, or historical linguistics form homogeneous communities free of internal tensions). The following chapter instead follows the contentious process of reimagining scholarship on the level of a *single* specialty. Such an empirical focus allows for more refined observations about digital methods and disciplinarity than are commonly provided by social scientists. In fact, a growing body of literature on the implications of digital tools for research adopts a comparative perspective, often suggesting that current disciplinary features of a field will strongly influence the speed and intensity at which its practitioners will engage with novel technology (Collins, Bulger & Meyer, 2012; Fry & Talja, 2007; Fry & Schroeder, 2010). This resonates with my own analysis of the diverse reactions to the digitization of the BNTL across different specialties. At the same time, a downside of the comparative perspective is that it tends to highlight and perhaps overemphasize current organizational features of disciplines, thereby introducing a certain circularity into the argument. A commonly drawn conclusion is for example that the more empirically oriented branches of the humanities and social sciences will quickly adopt data-intensive approaches (Fry & Talja, 2007; Fry & Schroeder, 2010). While this may be true, such a prediction remains tautological if it is not complemented by an investigation of the conditions under which the very

epistemic features of a given area of study may change as they enter a mutual shaping process with digital research tools.

Introduction

The birth of literary history in the 19th century is intimately connected to the formation of the European nation states, and it has long been recognized that this context continues to shape current historiographical narratives (see for example Robinson, 1983). Literary history since the 19th century often has been an account of the heroic literary deeds of male author-geniuses, portrayed as the finest representatives of a Romantic national spirit. Canonical views of literary history typically exclude whole groups of potentially relevant actors, such as women as writers, translators, and mediators in the literary scene (Whittle, 2013). Although various currents of feminist and critical theory have drawn attention to such bias (Warhol & Werndl, 1997), they have not managed to actually replace longstanding canonical traditions. More recently, literary scholars and developers of technology have seen the application of ICT as a way of stimulating attempts to revise the literary canon (Moretti, 2005; Wilkens, 2012) – after all, digital technology often is promised to facilitate collaboration among otherwise scattered, 'lone scholars', and as potentially allowing researchers to take advantage of large amounts of empirical material in ways that combine hermeneutic methods with computational approaches (ACLS, 2006; Babeu et al., 2009).

In this paper I study a collaborative project of literary scholars from 26 European countries, who set out to rewrite literary history from a transnational gender perspective. The goal of the project, funded by the European Science Foundation in the framework of COST (Cooperation in Science and Technology), is to foster collaboration among like-minded scholars and to create empirical knowledge about the reception of marginalized women writers in Europe 1700 to 1900. The project – or COST Action, the official term – aims to integrate the individual research efforts of the participants in a shared conceptual framework. Collaboration is organized around the use of a digital database, which the participants hope can help them remedy some of the many omissions in the literary canon. However, grant-funded collaboration in the humanities is a relatively new phenomenon. Literary studies is organized in a somewhat fragmentary way when compared to the natural sciences, i.e. local contexts such as national disciplinary cultures, and even individual university departments, play an important role in the organization of research. Also, scholarly knowledge is predominantly circulated in monographs, a form of expression that allows for a lot of individual freedom in terms of chosen research goals and analytical approach.

The goal of this paper is to interrogate how specific ways of organizing scholarly labor make possible certain forms of knowledge, and to study the challenges scholars face when trying to adapt established organizational models. What does it mean for university-employed literary scholars, often acculturated in close-reading-based research practices, to work together in the format of a collaborative project? What kind of changes does the shared use of a digital database require in the way they usually organize their labor? How is labor within the project eventually divided among the participants, and how does this division shape the collaboratively produced knowledge?

In order to answer these questions, I make use of empirical materials I have collected through semi-structured interviews with relevant actors, participant observation, and by studying various documents authored in the context of the COST Action. Theoretically, this paper is firstly informed by the work of Richard Whitley (2000), who has compared the organization of research labor across various fields. Secondly, I draw on insights from infrastructure studies (Edwards, 2010; Star & Ruhleder, 1996), which offers a sensibility as to how scholarly work is both enabled and constrained by existing institutional requirements, disciplinary cultures, and technological instruments. The combination of those two perspectives allows me to analyze the move to collaborative digital scholarship in terms of its far-reaching implications for how particular tasks in the academic labor ecology are conceptualized and distributed, as well as providing a framework to describe the inertia of established infrastructural arrangements. Digital collaboration, I will argue, can potentially produce new forms of knowledge in literary history, but especially when undertaken at a large scale, it will also tend to create significant tensions with the way scholarly labor is normally organized.

The structure of the paper is as follows. I will first discuss my theoretical framework in greater detail. Then I will introduce my case study and methods. Subsequently I will present my empirical findings, which are again subdivided in a number of sections that chronologically follow key events in the course of the COST Action, and the debates these events have spawned in the project.

The organization of scholarly labor in literary studies

In order to grasp the implications of collaborative, digital scholarship for literary studies, it is important to understand how scholarly labor in this field has traditionally been organized. Richard Whitley's (2000) comparative

analysis of how scientific fields differ in their organizational characteristics here provides a useful starting point. Whitley introduces the analytical dimensions of task uncertainty and mutual dependence among researchers to distinguish between fields. Task uncertainty describes the degree to which researchers share an understanding of their research object and theoretical priorities, as well as the relative agreement on how technical procedures should be applied. Mutual dependence describes the degree of coordination of research across sites, and the requirement for individual researchers to demonstrate the comparability and relevance of their work in relation to the work of their colleagues.

In Whitley's framework, literary studies is the antipode of post-1945 physics. Literary studies is characterized by high degrees of task uncertainty, and by low degrees of mutual dependence, while physics is configured in the exactly inverted way. In literary studies, research questions are highly individual, and the communication system is weakly formalized. Knowledge is circulated through monographs. These apply a discursive form that is relatively more accessible to lay people than the esoteric mathematical sign systems and highly abstract research objects of the natural sciences. At the same time, in literary studies, it takes particularly long for neophytes to make meaningful contributions to research, since a lot of individually acquired experience is necessary before one can wield an array of largely non-standardized techniques and theories, and make sense of the ambiguous findings. Theory here predominantly fulfils the function of distinguishing individual researchers in a plurality of coexisting approaches, rather than integrating labor conceptually according to shared theoretical priorities, as in physics. According to Whitley, literary studies is thus the exemplar of a 'fragmented adhocracy', in contrast to the 'conceptually integrated bureaucracy' that is physics.

However, the modern organizational form of literary studies should not be seen as the expression of an inherent essence of the field, but as the result of historical differentiation. To better understand how criteria of valid scholarly knowledge, technological instruments, and the organizational forms that we subsume under the label 'literary studies' have shaped each other over time, it is useful to combine Whitley with insights from the field of infrastructure studies.[21] Specifically, I propose to apply Edwards' (2010)

21 In establishing this theoretical link, I also mean to overcome a recurrently highlighted weakness of Whitley's approach, namely its static character and its rather peripheral interest in exploring change in field characteristics (Fuchs, 1993; Zeldenrust & Hagendijk,

notion of knowledge infrastructure to literary studies.

> Knowledge infrastructures comprise robust networks of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds (Edwards, 2010: 17).

Infrastructure according to this definition is not a specific thing (such as an academic department or a faculty), but a relational concept. It is something that occurs when the various institutional arrangements, scholarly practices, and technical standards that constitute the network fall into a workable configuration for its users, i.e. the scholars, students, administrators and support staff who work in and move through literary studies on a daily basis. Infrastructure emerges for people in practice, connected to activities and structures. The appropriate question, then, is not „what is an infrastructure?“, but “when is an infrastructure?”.

Edwards' definition builds on the foundational work of Star & Ruhleder (1996), who argue that infrastructure can be characterized by a number of interrelated features. For one, it is linked to conventions in communities of practice. Infrastructure both shapes and is shaped by those conventions. The use of a specific tool in a given academic field for example may increasingly become part of disciplinary training. To the same degree it will become bound up with the conceptual frameworks of the field. The concept of infrastructure thus is complementary to Jasanoff's (2004) notion of co-production of science and social order: by organizing scientific labor in a specific way, researchers also reproduce criteria of what counts as proper scientific knowledge. Such criteria in turn will inform the technical skills and tools transmitted through research training, which again is instrumental in reproducing criteria of scientific validity etc. Infrastructure moreover is sunk into other structures and social arrangements, thus reaching beyond a single event or local practice. Infrastructure in fact is something that invisibly support tasks, without needing to be assembled or reinvented for each new task. On the other hand, when infrastructure breaks down, it makes itself visible through its absence – think of the temporary chaos that is caused when an organization migrates its email servers to a new format, or when natural disasters interrupt railway connections in a densely populated country. Star & Ruhleder furthermore propose to see infrastructure as distributions of these properties along the axes of the global/individual, and

---

1985).

the technical/social. For example, if an information system is strongly embedded in a large scholarly knowledge infrastructure, its data categories will be generic enough to represent certain aspects of knowledge throughout the discipline. At the same time the information system must be malleable enough to cater to the specific local requirements of more specialized users (Bowker, Baker, Millerand & Ribes, 2010). This means that the degree of organizational integration of a field will influence the possibility to delegate certain tasks to technology. For example, in tightly integrated forms of knowledge production, such as physics, it will be easy to automate certain elements of the research process, since those elements are standardized throughout the discipline. In less tightly integrated fields on the other hand, possibilities to delegate individual work steps to technology on a global scale will be limited. Infrastructure occurs when the tensions between globally valid standards and local contexts, as well as between automated technological processes and tasks performed by human actors can be successfully resolved. An important consequence of this definition of infrastructure is that it develops incrementally - it is not created, it evolves.

So what does the knowledge infrastructure of literary studies look like? While there are currently no empirical studies on this question, we can make a number of preliminary observations on the basis of Whitley's work. Rather than a comprehensive description, I here present a number of aspects of humanistic infrastructure that will play a role in my empirical analysis.

For one, an important feature of a knowledge infrastructure is what its institutions consider legitimate forms of output. In literary studies, this has traditionally been the monograph. A record of monograph publication(s) often is an important factor in tenure and promotion decisions. The monograph implies a high degree of individual theoretical freedom, and a low degree of organizational differentiation of the underlying scholarly work process. Infrastructure in literary studies foresees that the primary process of producing a monograph be the work of a single individual. A decomposition of the research process that leads up to the publication of the monograph is not foreseen.

Furthermore, an important element of literary studies as a knowledge infrastructure is constituted by the totality of its information systems, such as bibliographies, archives, and library catalogues. These have played an important historical role in charting and making accessible the otherwise chaotic universe of print production (Chartier, 1995). Bibliographies define bodies of relevant scholarly knowledge for given subjects, and they traditionally have fulfilled an important ideological

function in defining national literary histories. These information systems operate with relatively generic, bibliographical categories, which have established themselves together with the emergence of print culture from the 16th century onward (Johns, 1998). In this process, bibliographical categories have become seemingly natural ways of describing print production. At the same time, bibliographical categories have become part and parcel of the conceptual deep structure of literary studies: very often, scholarship is organized around such categories as œuvre or author. Poststructuralist critics have famously drawn attention to the abstracting moves that make possible such forms of knowledge in the first place (Foucault, 1979; Barthes, 1978).

As pointed out, infrastructure is relational - the daily work of one person may be the infrastructure of another (Star & Ruhleder, 1996). This has implications for the visibility and social prestige of certain kinds of work (Star & Straus, 1999). In literary studies, there is an established division of labor between scholars on the one hand, and librarians, bibliographers, and archivists on the other. The division is such that the work of the latter is considered a technical service to the work of the former. Their role is thus similar to that of laboratory technicians as analyzed in the seminal paper by Shapin (1989), whose function is critical to the conduct of experimental science, but at the same time largely invisible. Libraries, archives, and bibliographies are infrastructure for scholars in that they constitute a transparent, ready-to-hand instrument that enables and constrains their research.

Furthermore, it is fair to assume that regional scope is an important infrastructural characteristic of literary studies. In fact it might be useful to think of literary studies as consisting not of a single, but of multiple, regional knowledge infrastructures. For example, the scope of bibliographies is frequently a regional one. Bibliographies in Central Europe are often produced by national Academies of Arts & Sciences, which function as an authority that vouches for the reliability of the information they provide. Also, sub-disciplines of literary studies are delimited by language communities. Some of these communities (English, French, German...) are much more influential than others. This means, among many other things, that smaller disciplinary communities, say, Slavic Studies, are likely to possess less disciplinary knowledge about topics that are well studied in larger communities. Therefore, we can often observe that research trends developed in larger disciplinary communities arrive with a certain delay in smaller ones (Dojcinovic-Nesic, 2006), and more generally, that what is

considered legitimate scholarly work varies by disciplinary community.

However, in recent years, we can observe dedicated efforts to create an integrated, pan-European digital research infrastructure for the humanities. Particularly visible projects are DARIAH (Digital Research Infrastructure for the Arts and Humanities) and CLARIN (Common Language Resources and Technology Infrastructure). Describing itself as a 'connected network of people, information, tools, and methodologies', DARIAH (n.d.) presently partners with archival and research institutions in 14 European countries. Its core mission is to enhance and support digitally-enabled research across the humanities and arts. Next to promoting the coordinated development of analytical applications and improved long-term access to digital datasets, DARIAH's activities include the exchange of digital skills and computational research methods. CLARIN (n.d.) similarly aims to build a federation of European data repositories (archives, libraries), service centers, as well as centers of expertise at universities and other research institutions. The CLARIN web portal offers access to datasets and tools for researchers in computational linguistics and related fields, but also for social scientists interested in analyzing large amounts of text-based material.

Ambitious digital infrastructure initiatives such as DARIAH and CLARIN can be seen as interventions in the organizational landscape of the humanities. Originally inspired by similar efforts in the natural sciences (Jankowski, 2009), the goal of DARIAH and CLARIN is to create a technological basis that would allow humanities scholars to access and analyze uncommonly large amounts of digital data in a collaborative fashion, so as to enable them to answer research questions that could not be tackled with traditional means. The first step for such projects typically is the creation of buzz (by researchers, research managers, and policy makers) to attract the interest of funding bodies (Brown, Rappert & Webster, 2000; Kok & Wouters, 2013; Vann & Bowker, 2006). Once granted, project resources buy a degree of independence of scholars from their local organizational environment that allows for a potential reconfiguration of labor. Insofar as digital infrastructure projects constitute major investments, they are also informed by a specific set of managerial values, such as a pervasive systems perspective, sustainability, and avoidance of investment redundancy (Anderson, Blanke & Dunn, 2010; Zorich, 2008). Against this background, the relatively weak degree of integration of labor in the humanities poses a potential obstacle to the goal of interoperability of data and methods (Fry & Talja, 2007). Advocates of digital infrastructure therefore often promote the

identification/disambiguation of shared research methods, information practices, and/or data standards across the various humanities disciplines, which can then serve as technological design principles (Anderson, Blanke & Dunn, 2010). DARIAH for example is linked to the goal of building digital infrastructure around 'methodological commons', i.e. fundamental building blocks of scholarly processes that are shared across countries and disciplines. Another approach is to focus on integration through data standardization. Lynch (2002) and Borgman (2007; 2009) for example argue that humanists should establish strong, once-and-for-all definitions of what constitutes data and what scholarly interpretation, so as to provide a base for the encoding of interoperable metadata in digital libraries.

However, change in knowledge infrastructures can, per definition, only be incremental – as has been widely acknowledged, the uptake of digital research technology in the humanities depends on the extent to which the forms of user engagement they encourage allow to strike a balance with scholarly conventions, existing technical standards, and other elements of the status quo in local contexts (Bowker, Baker, Millerand & Ribes, 2010; Bulger et al., 2011; Wouters, Beaulieu, Scharnhorst & Wyatt, 2013). Successful implementation of digital research technologies hence will depend on the possibility to reconcile individual professional investment of scholars in existing research paradigms with the affordances of digital scholarship. It is this process of emerging organizational forms of labor that I hope to shed light on through my empirical analysis.

Case study & methods

This paper is based on data collected in a collaborative project in literary studies, entitled Women Writers in History (WWIH). The project is funded in the intergovernmental framework for Cooperation in Science and Technology (COST) for the period 2009 to 2013. COST does not fund research directly, but provides support for networking activities, such as meetings, joint conferences and publications. COST Actions are often meant as a preparation for further projects, for example in the European Union's Seventh Framework Programme.

The aim of WWIH is to lay the groundwork for a new history of European women's participation in the literary field before 1900. In its application for funding, WWIH forcefully argues that current literary historiography is still informed by the chauvinistic, canonizing tendencies of 19th century historians, from whose narratives women have been mostly

excluded. In particular, the document observes the lack of coherent empirical data on the activities of women as writers, translators, and mediators in the literary scene across countries and periods, which is exacerbated by a distorting focus on influential national literatures and language communities. Another limitation is the use of small samples of canonical writers in most literary research, which reflects the amount of empirical information that can be processed in single-author, close reading-based research practices. In contrast, WWIH aims to mobilize the combined efforts of its participants to work towards a more substantial and systematic empirical basis. More specifically, the application document promises the delivery of a "prototype of an online research infrastructure" (COST, 2009: 10) through the collaborative use and further development of an existing digital database, which was developed in a preceding project at Utrecht University. The prototype will build on this existing dataset, as well as interlink with other databases (such as DBNL, ECCO, Gallica2). Additionally, it will be further enriched by individual data input by participants.

> The COST Action will mobilize researchers to collectively create tools allowing to have the full benefit of these sources, and to establish direct connections between women's writings and these very diverse reception contexts. (…) Thus, a new instrument (a research infrastructure combining a virtual collaboratory with an online database) allows large scale approach of sources, and generate new (research) material: data about contemporary reception of early women's writing. These are shared, commented and analyzed (…) by quantitative and qualitative approaches, and eventually suggest new questions impossible to be asked up to now (COST, 2009: 6-7).

WWIH is a grass-roots project that has developed out of previous collaboration of predominantly Dutch literary historians. The number of participants has perpetually increased during the course of project – from initially 50 to about 120 researchers from 26 European countries in 2012. The research interests of the participants can be subsumed under the topic 'reception of women writers', but only very roughly so. Individual research interests include topics as diverse as the literary life of Irish nuns in the 16th and 17th centuries, studied empirically on the basis of monastery archives; the business relations between women writers and their publishers on the Iberian peninsula in the 18th century, studied on the basis of correspondence

and archival materials; or the history of a particular Dutch library for women readers in the 19th century, based on an analysis of the book loan files.

The data on which I draw in my analysis were collected in a total number of 24 semi-structured interviews, ranging from 30-120 minutes in length, extensive participant observation in two database-training events and three project meetings of several days each, an analysis of electronic project communication, a survey focusing on obstacles to database uptake (20 respondents), as well as an analysis of documents produced for internal and external use, such as project plans and presentations delivered at scholarly conferences. Data collection was spread over a period of two years, starting from December 2009. Initially, the data were collected according to a grounded theory approach (Charmaz, 2006), with a gradual concentration of interview emphasis on the topical relation employment-research-reputation.[22]

Discussing theory in a pragmatic way

The COST Action was officially launched during a four day meeting at the Huygens Institute in The Hague in November 2009. The kick-off meeting featured a presentation by Stanford historian Franco Moretti, a member of the COST Action's advisory board, and a pioneer in recent attempts to apply large scale quantitative approaches to literary history. Furthermore, next to introductory presentations by the project leader and digital humanities researchers working in other projects at the Huygens Institute, particularly spirited talks were given by two Belgian PhD students who had joined the COST Action immediately after launch. One of the students had developed a database on the reception of Scandinavian writers in the Netherlands, while the other one demonstrated how she had integrated a quantitative book historical component in a qualitative analysis of fairy tale translations. The participants shared a general enthusiasm about the potential benefits of digitally mediated collaboration, in particular the possibility to expand the empirical scope of their individual research. It seemed to tie in with hopes that they had had for a long time – to rewrite literary history from a gender perspective in a way that could not be achieved by any individual member.

22 I wish to thank the participants of the COST Action Women Writers in History for giving me the chance to conduct fieldwork in the context of their project.

> The reason I got involved was that I thought it was a great idea. To me the great value of the idea is that it's just too much work to come up with a big scale picture of women's production in literature in Europe. As we've seen, in order to do a valuable analysis of a text it takes so much work that it's impossible for any single person or indeed for any collaboration to be able to come up with a hermeneutically valuable analysis of the production of women's writing, but I do think it's possible to look at reception at a large scale.[23]

While the kick-off meeting made tangible the general excitement about digitally mediated collaboration, there was at that stage no particular concern with how the many promises that had been made in the application should be put into practice. Things got much more practical on the occasion of the first of the so-called training school events in The Hague in October 2010.

As indicated in the above, the COST Action builds on a preexisting relational database developed for the rather specific research goals of a small group of Dutch literary scholars. The original purpose of the database had been to investigate the reception of women writers in the periodical press in the Netherlands 1800-1900. In the training school events, participants were confronted with an interface that had been developed for that specific research model. It allowed to search the dataset through three interlinked menus: authors, literary works, reception documents. Each of these displayed data according to a number of subcategories, e.g. authors were described through name, dates of birth and death etc. Literary works as well as reception documents in turn were described through bibliographical categories such as title, year of publication and genre. Genre again contained a number of subcategories. The project leader originally intended that the individual participants of the COST Action could use the database in ways that made sense for them. The database should function as a catalyst for new research: while allowing them to retrieve information already stored in the dataset, it should also offer an incentive for the participants to enter more information as they draw comparisons between existing content and the data they work with in their own research. The training school sessions were set up as an opportunity for the participants to familiarize themselves with the database, but also as a forum for deliberating how the original data

23 Personal interview with Marie-Louise Coolahan, 14 November 2010, Madrid.

format could be further developed to better suit the requirements of the expanded group of users.

A first issue that was recurrently brought up during the training school sessions was the problem of defining literary genres in the database. Some categories were simply perceived as overly specific. Others were rooted in genre definitions specific to a particular language, thus raising the question as to how one should apply them to other languages.

> I think about the genres, we need fundamental discussions about that. There are things which I couldn't really identify. For example, contes, ok, that's the French word for fairy tales, but it actually can be a tale (…) I came across this problem again and again, is something a tale or a story, and is that the same as Erzählung in German, I think we need to discuss this in greater detail.[24]

> It's a problem when you find closet drama, but not drama. The other way around, it's ok if you have poetry, but not sonnet, that's ok. That first level has to complete.[25]

Another particularly contentious issue was the definition of reception:

> I think that one of the big problems at the moment, and people don't seem to want to address that, is, how do we define a reception? Is it the translation of a work, or is it a review. (…) I don't think it's clear enough and I think that a lot of colleagues have very different ideas of what counts as reception I think that's going to be a big problem.[26]

The project leader proposed to resolve the potentially very lengthy discussion about definitions of genre and reception by treating data entry as an unproblematic step, the generation of a raw data package that could be further contextualized by prospective users of the database. To her, it was a predominant concern to amass enough material in the database to allow for larger-scale quantitative comparisons. Specifically, she proposed that

24 Field notes by the author, training school event at the Huygens Institute, 12 October 2010, The Hague.

25 Ibid.

26 Personal interview with Marie-Louise Coolahan, 14 November 2010, Madrid.

translations should always be treated as an instance of reception.[27] In line with what had been laid out in the original application document, she moreover suggested that genres should be identified according to definitions used in the historical reception documents, thus providing first-hand access to contemporary readers' perception of the literary work in question. This, however, implies a specific way of using the database, and a specific theoretical choice that the prospective users would have to agree on.

In contrast to definitions of genres and reception, categories such as author name, author gender, language, year of publication, publisher, i.e. categories that are widely used for information retrieval in many different contexts other than literary history, did not attract noticeable contradiction. They were apparently taken for granted. Through the lens of Star & Ruhleder's (1996) theory, the perception of some categories (author, title, publisher) as occupying a lower level of abstraction, and of others (genre and reception) as occupying a higher, domain-specific level of abstraction, can be seen as the result of a historically grown infrastructure. Some categories have over time become established as an unproblematic technicality to enable information retrieval in many contexts. Others have been developed in more specific disciplinary contexts, thus limiting their global applicability.

Yet even within those specific disciplinary contexts of literary studies, definitions of genre and reception are often subject to theoretical debates in monographs and at scholarly conferences. The organization of literary studies in fact is such that academics identify themselves as scholars by developing and defending an individual theoretical position on these matters (Whitley, 2000). The training schools, however, provided a context where definitions of genre and reception are negotiated in a face-to-face way, in order to advance the project. Rather than an opportunity for individual scholars to distinguish their theoretical perspectives, it was a pragmatic requirement to narrow those definitions down to a workable compromise. The training school thus encouraged the participants to think in a more functionalistic way about what is otherwise a continuous theoretical debate.

One example of this 'will to agree' is the recurrent suggestion to resolve the problem of arbitrary categories by resorting to what participants generally considered to be less abstract categories, e.g. if tale is too arbitrary, resort to the more general category of narrative. On several occasions

27 Field notes by the author, training school event at the Huygens Institute, 12 October 2010, The Hague.

participants argued for adopting less specific categories or standardized categories that are used by library catalogues, in order to be able to at least agree on *something*:

> If you can't decide on a global level, let's go on a higher level. If tale is too specific, use narrative.[28]

> If you choose to do the formal genre, stick to the classifications that libraries do: drama, poetry, prose. That's less specific, so it's drama and not closet drama, but it's an international standard.[29]

Another example demonstrating the 'will to agree' promoted by the training school format are situations in which theoretical debates resulting from certain categories are intentionally avoided, in order to get on with the project work. One informant specifically refers to a lengthy theoretical debate that *could* be had about the genre definition of autobiography as something that the project *perhaps should not get into*, since that would endanger the goal of reaching an agreement on data categorization: "There's a vast literature for example on what an autobiography is and on how you define it which we maybe don't want to get into."

However, in spite of this pragmatism, the training school did not really lead to lasting consensus on all categories. Several participants suggested that in the case of contentious categories, decisions should be delegated to a database editorial committee. The responsibility to take specific theoretical-cum-practical decisions thus was delegated to a future organizational entity within the COST Action. However, by the time of writing this paper, the editorial committee has not materialized, and the project continues to work with the original database format.

These difficulties, I suggest, can be seen as a manifestation of the obduracy of the extant knowledge infrastructure. As pointed out in the above, an important characteristic of infrastructure is how it distributes specific activities along the axes local/global, and social/technical. While bibliographical categories are widely agreed-upon as viable abstractions, definitions of genre and reception are considered matters of theoretical debate in literary studies. Whitley (2000) observes that in fragmented

28 Field notes by the author, training school event at the Huygens Institute, 12 October 2010, The Hague.

29 Ibid.

adhocracies, theory does not integrate labor across research sites, but rather distinguishes individual approaches in a variety of co-existing approaches – "In the humanities, one person's data is another's theory" (Borgman, 2009). The perspective of adopting a shared data scheme therefore implies a redistribution of the definition of analytical goals from an individual activity to a technicality, shared on a group level. The formalization of data categories that are usually not formalized on such a high level of abstraction is a necessity for comparative quantitative research, but it also creates a number of tensions in regard to monograph-oriented scholarship.

Delegating data input

Firstly, a stable classification of data categories required users to enter data in a specific format. The latter did not always correspond to what participants considered useful categories from their individual analytical perspectives.

> The fundamental question for me is: To what degree must the COST Action really stick to the initial outline and to what degree is there space to develop concepts, questioning further, taking into account recent developments in Gender Studies, theory of literary history and literary theory?[30]

> I should have liked the tool to be more flexible, among other things so that it would be easier to correct or edit the structure already built when new knowledge alters the picture (so that the picture doesn't have to be clear when you start entering data in the database).[31]

Furthermore, the database required amounts of data input that exceeded what could be easily entered by participants in the course of doing their individual research. Data input started to appear as an overhead to the 'actual' research.

> Like all of us I am concerned about how to make the database to cover enough material and providing an infrastructure to work with. I might be interested in systematically listing diaries written in

30 Anonymous, unpublished survey among COST Action members, October 2011.
31 Ibid.

> Finland (hoping they discuss literature, reading and writing too...?) But I could not do this alongside my full time job. Letters would be great material to study reader's experiences, but the material is so huge...[32]

To be sure, the curation of data always has been part of the activity of scholars. Especially so in the case of COST Action participants, whose research often involves archival work that leads to the discovery of previously unknown literary documents. Scholarly work may thus in practice overlap with the work done by archivists. However, the amount of data curation involved in research will be directly determined by the individual disciplinary research needs of the scholar, and the data can be categorized (implicitly or explicitly) in ways and in amounts that make sense in the individual research context.

In contrast, for many participants, the activity of entering raw data on the reception of women writers into the database did not have an immediate connection to their individual research. For many, it therefore simply stopped being recognizable as a proper scholarly activity. At the first training school event and at an early meeting in Madrid, participants expressed strong concern about the perspective of entering data at a large expense of time ('slave labor'), without taking any 'real' advantage in the shape of producing peer-reviewed publications. They began to call for a solution that would allow them to delegate data input.

> More generally – and this is a very personal opinion – I think that the pure accumulation of more names and titles to prove the presence of women in literary history, at least for the "big" European literatures as German, English, French, etc., is neither a very interesting task for a researcher nor a theoretically challenging objective (...).[33]

> I think it is important also for the Women Writers database to have [data input personnel], because not everybody wants to be on the technological side. (…) If you have a person whose time is not as valuable as yours – to put it that way – to do that, that time-consuming work, that's great, because you can send data and that

32 Anonymous, unpublished survey among COST Action members, October 2011.
33 Ibid.

person knows how to enter them.[34]

An arrangement that developed in light of the difficulty to reconcile main employment context with project work was to delegate large chunks of data input to assistants. Some of those were graduate students supervised by the participants. Others were hired data workers without further interest in WWIH as a project. Drawing on students as a work force in similarly oriented projects is a quite common strategy in the North American digital humanities (Blackwell & Martin, 2009; Zorich, 2008: 30). Anecdotal evidence I collected in various meetings of the COST Action suggests that 'data work' is often downplayed in project applications, since expenses on this 'subordinate' element literally go at the cost of other project aspects, and thus are feared to make funders suspicious.

As I have pointed out, the knowledge infrastructures of literary studies are characterized by a division of labor between bibliographers and scholars, according to which the work of the former appears as a technical service to the work of the latter. Bibliographical work here can be considered a technicality insofar as it constitutes a basic element of an established research model. Bibliographical categories are conceptually bound up with dominant monograph-oriented research practices, and the ability to use bibliographies, library catalogues, and archives, constitutes a widely shared skill among scholars. Therefore, bibliographical work is delegable. Data input in the COST Action had a family resemblance with bibliographical work, and participants found it convenient to delegate responsibility for that labor. To the participants of the COST Action, data input appeared as 'technical', but not according to the theoretical infrastructural meaning of the term that I have just described. As the above quotes illustrate, the participants tended to perceive data input as 'technical' in the sense of involving the use of digital technology, and in the sense of being a type of non-intellectual work that a professor cannot afford to spend much time on. But in contrast to bibliographical categories, the definition of analytical categories (genres and reception) remained contested. Also, it was still unclear which new research questions and insights the database might enable. The differentiation of data input as a separate work step, and its delegation to students and data workers, therefore was a managerial artefact, rather than an effect of infrastructural evolution. This artificial separation had the unintended effect of limiting the diffusion of database skills.

34 Personal interview with Nieves Baranda, 14 November 2010, Madrid.

Interestingly, the student data workers charged with data input did not necessarily think of their job as a purely technical one. Take for example Astrid, a Dutch Mphil student who was hired for data input as part of her research internship. Next to entering data, Astrid also used the database for her own research on the reception of the British writer Ouida in the Netherlands. Her comments make clear that her research practice literally has emerged in conjunction with entering data.

> You are drawing conclusions as you are entering the data. For instance what's really important in my research is that you can see that there is a number of female translators who are working on Ouida, and they are always the same ones. This is something I might not have found out if I had just listed the translations. (…) As you enter the data you are also placing it in a larger network, and by visualizing that info you get a better sense of what you are working with. You create this overview which does not draw conclusions for you, but it helps to see those links which you might have overlooked otherwise.[35]

Astrid pointed out that the quality of the work done by the otherwise indifferent data input personnel that had been hired before was sometimes rather poor, and demanded a lot of post-hoc correction work. Similarly ambiguous was the quality of occasional contributions by some project participants who manifestly had not taken the time to familiarize themselves with the database interface.

> I am not sure how many people are really using the database for their research. I do think that there is a fundamental difference in entering data into the database and using it as a research tool. There have been assistants before me who just received lists of information that they entered into the database, and I am mostly entering information into the database for my own use. Other people can use it as well, but it's mostly what I deem necessary to be entered I enter. (…) You can see that some people just entered data without thinking. Just authors e.g. without adding works, or lists of works without adding reception. That's less useful because you don't get this view of a network that would get if you enter receptions and work and

35 Personal interview with Astrid Kulsdom, 13 October 2010, The Hague.

link them up to each other.[36]

Despite having attended one or more training school events, some of the more advanced project participants on several accounts exhibited only a rather superficial understanding of the phenomenology of the database until well into the funding period. It was for example relatively common for them to confuse the user interface of the database with the underlying data model, a distinction that became clearer only when a possible transition to a new, more smoothly interoperable data model and various visualization techniques was discussed. Many of the more established scholars seemed to think about the database more in terms of flat excel sheets, a format they were familiar with from individual datasets they had created for their PhD theses.

While the more advanced project members thus have tended to apply a distinction of 'technical' data work vs. 'actual' scholarly activities so as to justify delegating the former, it was exactly by getting their hands dirty in data work that student assistants – for whom the project temporarily became the central reference point of their work lives – have managed to combine database skills with substantive research skills. The participants' usage of the term technical, I suggest, is an expression of the fact that data work currently is not widely recognized as a legitimate part of the scholarly skill set. Yet the commonsensical association of 'technical' with technology artificially severs the link between conventions of practice and criteria of valid scholarly work. If the participants could achieve an understanding of database-related work as having deeper implications for the intellectual substance of scholarship, this link could be re-established. Rather than reproducing a fault line between tech-savvy data workers and professors with a more traditional skill set, the project could then become a seedbed for the dissemination of digital skills.[37]

36 Personal interview with Astrid Kulsdom, 13 October 2010, The Hague.

37 It seems that the participants have indeed developed a more differentiated view of data work in the course of the project. On the occasion of the concluding conference of the COST Action in June 2013, a number of participants reported that once they had taken the time to engage in more substantial data input, they experienced interesting surprises, such as unanticipated empirical discoveries, and an overall more 'system-like' perception of literary reception. While the core tension between time-consuming data input and individual research requirements remained tangible, it appears that the separation of data work as a separate work step has become more blurry over time.

What can you expect from a database, and what can it expect from you?
Another source of tension was the fact that the database could not be readily used as a resource for producing a publications, an aspect that made it even more difficult for the participants to combine project work with their individual career requirements. Although intellectually inclined to distrust canonical accounts of literary history, the expectations of many participants towards the WWIH database in fact was shaped by the advantages they associate with professionally curated archives and bibliographies. Archives and bibliographies are infrastructure in the sense that they provide what is normally agreed to be a relatively complete, authoritative body of knowledge. Bibliographies crystallize regional research traditions and delimit the field of relevant knowledge one must possess when embarking on a new research project, be it a paper or monograph. In this arrangement, trust is delegated to archivists and bibliographers.

> You know what a bibliography is? If you want to research Shakespeare, you take a Shakespeare bibliography and you find all information, and it has been checked and organized in some topics and subjects. So you check that and you see 'oh I need to see this and this and this' and then I have to read all those forty or whatever books and then I get some information and some answers to my question, and I know what is known and what is still to be researched.[38]

Archives and bibliographies are as geographically situated as the research practices they enable. Their reliability is jeopardized if the user ventures to transcend the bibliographically instantiated boundaries, a move that could actually be seen to effectuate a breakdown of infrastructure (Bowker & Star, 2000). The very goal of the COST Action of course was to do just that – to collectively transcend regional contexts in which scholarly labor is usually organized, and to work towards the creation of a systematic transnational perspective on the reception of women writers. Yet while leaving vouched-for bibliographical territory was a core premise of the project, it has nevertheless created significant insecurity. Many participants expressed doubts about the reliability and coherence of the database, given its aim to combine empirical materials from a lot of different contexts. The responsibility to solve these issues, however, could no longer be delegated to

38 Personal interview with Nieves Baranda, 14 November 2010, Madrid.

bibliographers, but had to be dealt with internally.

> The problem is, if I go to the database for information that I am familiar with, I can see that it is poor data, and I cannot trust it. It's poor information, something is always missing. (...) From what I've seen in terms of Spanish authors in the WWIH database, it's not good. So, is it the same with Italian sources, with French sources? If I don't trust information, then I have to double-check every information.[39]

A first cause for concern was the fragmentation of the covered data. For example, Dutch women writers were extraordinarily well represented, given the origin of the database in a preceding research project at a Dutch university. The result was odd contrasts with the sparse coverage of other national literature, especially in the case of countries with a wealth of disciplinary knowledge about women writers. Similarly, relatively famous writers often tended to be underrepresented in comparison to much less well known ones. This was a consequence of the unsystematic manner in which scholars would sometimes enter data that they came across in their individual research. The fragmented empirical picture that emerged reflects the heterogeneous empirical foci adopted by the participants.

> At the same time, when having for instance several smaller projects about Italian writers, it will be completely … how do you say, desequillibré? Because there will be too much information about Italy, and the need to filter out overrepresented countries and periods, no?[40]

Furthermore, the database was meant to be used for studying reception. In turn, this implies a bibliographically rather incomplete picture, if one decides to use the database by looking at literary production first.

> (…) there's a number of very famous English authors in the database, and some of their works are in there because their reception is recorded, but not all of their printed works are

39 Personal interview with Nieves Baranda, 14 November 2010, Madrid.

40 Field notes by the author, project meeting at the Huygens Institute in The Hague, January 2012.

included.[41]

It therefore became clear that the project in the short term would not only produce a larger base of information. Instead, given its very empirical scope and the plurality of research interests of the participants, it would also produce more gaps and fragmentation in that base. Combined with the laborious task of simple quality control of data, this made it difficult to imagine that the database could be used as a resource similar to a traditional archive or bibliography any time soon. The implication rather was that more labor investment was necessary to harmonies and better understand the growing dataset.

In contrast to a view in which the database is seen as a resource from which to 'slice off' bits of empirical information, the project leader promoted an understanding of the database as a serendipitous research tool, a perspective in which the limitations of the dataset represent the very object and catalyst of research. Fragmentation and ambiguity of coverage could then be turned from a professional risk into the distinctive features of a project intent to make visible systematically marginalized aspects of literary culture. As a side effect, this view justifies more dedicated expense of labor on quality control, since the value produced in WWIH would not only be the individual publications the participants could use to advance their careers, but also the database itself as an asset that increases the value of the project on the market for competitive funding. Indeed, in preparing a follow-up application for another funding framework, the project leader sent out an email in which appropriate cleaning of the data in the database was presented as a requirement for delivering a competitive funding proposal.

> It will in particular be important for the HERA proposal that we be able to specify that cleaning up of data is being taken care of, but also for any other subproject for which we might try to find funding (...). So I suggest that each of you also check if she has the possibility of taking part in these sessions (van Dijk, 2011).

The organizational contexts of the project and that of literary studies as a field thus tended to promote diverging expectations towards the database, which reflects also the different types of knowledge those contexts aim to

41 Field notes by the author, project meeting at Complutense University Madrid, 13 November 2010.

produce. Given the need to make their work in the project 'count' for their university-based careers, the participants tended to expect the database to be a readily usable, 'complete' source of data (similar to a bibliography) that they could use to write monographs and papers - typically with a focus on small samples of writers. The project framework tended to promote a notion of the database as a complex research object in its own right, which has a much larger scope than any bibliography, but whose empirical coherence and reliability has to be accomplished by the participants themselves. To the project leader, resolving the tension between those diverging views presented itself as a recursive resource problem. If more funding could be acquired, additional assistants could be hired to support the participants in the provision of quality control and data input. This would further increase the attraction of the project to both participants and funding bodies. Yet to deliver a competitive proposal, the project first would have to mobilize enough labor to reach what could be presented as a critical mass of reliable data...

Conflicting demands

In fact, what was useful for the project increasingly appeared as different from what was useful for the participants in their individual research careers. Many participants continued to be enthusiastic about the longer-term perspective of gaining access to a coherent empirical picture of women writers across countries and historical periods. Yet their ability to free up time to advance the database was limited. They often thought of their working hours as divided between the teaching, admin, and research they had to do at their home institutions, with project work coming in as a distinct, rival entity to these workloads.

> I think a lot of people are having that time problem. I wish I had been able to go to the training school, because that sounds like it was an immersive experience, and I now have to do that myself. That's the same problem with most people, most people aren't employed as full time researchers, they are employed as teachers and they are burdened with administration as well.[42]

42 Field notes by the author, project meeting at Complutense University Madrid, 13 November 2010.

The project leader said:

> (...) I think many of the colleagues have been engaging themselves in something European without really feeling that there is a corresponding need to abandon the individual freedom in a way. (…) I'm not sure that I can go on suggesting to colleagues you should use the database. If they don't, I cannot continue endlessly (...) At the end of the 4 years we will need to have our research programs to submit to the FP7 program or something like that. We cannot ask 25 countries to be all included in the follow-up proposal. We will be between 6 and 10, so there will be a natural selection. Those who want to participate in this research proposal will have to show clear adherence to this way of doing research.[43]

Apart from the general issues concerning data input and trust in data, the ability and willingness of the participants to reconcile individual research careers with participation in the project was influenced by a number of more specific, disciplinary factors. These shed light on some particularities of the various knowledge infrastructures in which the participants were embedded, and on the position of the participants within them.

For one, there is an interesting contrast between the career backgrounds of most participants and the project leader. The latter had for several years combined her employment as a teacher at a secondary school with research as a non-tenured member at a University. After that, she fully switched to a career as a researcher based on various grants. For her, grant-funded project frameworks had for a long time been her interface to engage in the disciplinary discourse of literary studies. The COST Action for her was not rival to a tenure-track based research career at a university, instead it was her primary research context.

Also, there is evidence that individual national research systems in which the participants were embedded influenced their ability and willingness to participate. It is has been variously pointed out that the lack of reward for the creation of digital scholarly resources is an obstacle for the creation of digital infrastructure in the humanities, both in Europe and the US (AHRC, 2006; MLA, 2012). The project leader more specifically reported that participants from the UK were particularly vocal about their difficulties to reconcile project work with their academic careers. An explanation could

43 Personal interview with Suzan van Dijk, 8 November 2011, The Hague.

be the importance of the British Research Assessment Exercise, a nationally orchestrated evaluation protocol that strongly values peer-reviewed publications as a performance criterion for tenured researchers also in the humanities (Barker, 2007).

At the same time, involvement in international projects often is seen as an important addition to scholarly CVs. This results in the phenomenon of pro-forma-participation, i.e. scholars participating only to prove international involvement, rather than actually investing labor in a project:

> (…) these things, participating in a research group [WWIH], these things are important in your CV. But sometimes you have to be careful, because sometimes people want to be in a research group, but they don't mean to work. It's like 'well, I'm in a research project, I get some things out of it", but they don't work very hard in it.[44]

Furthermore, individual disciplinary context mattered. Various scholars expressed their view that in the case of countries with a lot of disciplinary knowledge on women writers, the activity of filling the database seemed particularly unattractive. While more empirical research would be necessary to warrant a firm claim, an infrastructural explanation could be that the threshold at which database-related work stops looking technical, and begins to be more co-extensive with proper scholarly activity, is lower in regions whose bibliographical and archival infrastructure is less well developed. There, charting the activities of hitherto unknown women writers in a database more quickly looks like generating new knowledge also from a disciplinary perspective. In contrast, creating a comprehensive dataset in a country such as Germany first requires amassing data that are very likely to be already found in existing bibliographies, thus making that work appear rather redundant. These dynamics played out in interesting ways in the later stages of the COST Action.

Resolving the tensions in local contexts

After the first year attempt to involve 26 countries in the collaboration, the project leader began to envision the possibility of downsizing the project in applications for follow-up grants – i.e., to reduce the number of participating

44 Field notes by the author, project meeting at Complutense University Madrid, 13 November 2010.

countries, and to focus on those that had shown more active participation. Generally, it was smaller countries situated at the geographical and cultural periphery of Europe, such as Slovenia, Serbia, and Norway, that tended to contribute more data to the database. One reason for the discrepancies in data input was that in certain local contexts, the tensions that plagued the COST Action at large could be more easily managed.

An example is Serbia, where a smaller project has been launched that has adopted the database-structure of the COST Action. Biljana, a professor of Literary Studies at Belgrade University, explained that she had managed to combine her involvement with the COST Action with a project funded by the Serbian Ministry of Science. While conceptually inspired by the COST Action's original goals, she had tailored her funding proposal to the needs of her local disciplinary context. In contrast to the COST Action, she had emphasized the aspect of literary production. Given that research on women writers has only a short history in Serbia, the creation of a database on their literary production made eminent disciplinary sense. Furthermore, she explained that it had been an important element in her application to present Serbia not only as a country that received foreign writers, but that also exported its own writers – this was in line with what she knew to be a general strategy of Serbian cultural policy.

> Everything has started with the COST Action, but you see in the COST Action, the database has much more emphasized reception than production. And we do need to put emphasis on production. (…) The first reason is that within our community - there so many women writers who are not only outside the canon, but the problem is that nobody has ever head of them. (…) The second problem, on a more general level, we want to be perceived as those who have written something, not only as a country who has received something. Slavica non leguntur, people do not read Slavic languages (…) The basic idea was to construct this database, which is based on the COST Action's database. They are similar, but we have added what we are interested in, and in the way we are interested in.[45]

The Serbian sub-project adopted the database structure of the COST Action, and while linking its data up to the COST Action's database, it also has an

45 Personal interview with Biljana Dojcinovic, 26 January 2012, The Hague.

independent web-presence. The much reduced scope of the Serbian sub-project reduced many of the challenges that plagued the COST Action's database. Biljana expected that the database would eventually comprise a number of about 30 women writers. While this seems a small dataset compared to the scope originally intended by the COST Action, the number was still impressive in the context of Slavic Studies, given the almost complete absence of disciplinary knowledge about Serbian women writers. The small scope furthermore allowed data entry to be done almost exclusively by two PhD students of Biljana's. Given the reduced size of the sub-project, decisions about the database interface and data categories were much easier to take, and the challenge of creating a trustworthy dataset was much reduced. Lastly, given the small number of authors that should be covered, the participants could continue working on individual writers, while still contributing meaningfully to the sub-project.

Discussion

Early initiatives to create cyberinfrastructure for the humanities were characterized by the paradigm of e-science (Wouters & Beaulieu, 2006). This led to the creation of sophisticated computational tools that were of great interest to computer scientists, but often met with indifference by the scholars who were supposed to use them (Bulger et al., 2011; see also Wouters, Beaulieu, Scharnhorst & Wyatt, 2013). One conclusion that was drawn from these early initiatives was that the prospective users of digital tools should be thoroughly involved in their design (ACLS, 2006; ESF, 2011). The COST Action I study in this paper can be seen as an excellent example of this approach, in that it actually emerged from a bottom-up effort of scholars to use larger-scale quantitative approaches to literary history. However, this has created problems of its own.

As Whitley (2000) observes, literary studies is configured to enable research according to a qualitative, monograph-oriented model, situated in regional disciplinary contexts. Collaborative project frameworks like the one studied in this paper on the other hand require a certain degree of integration of individual research practices. In the case at hand, this has created a number of tensions between project requirements and disciplinary career requirements of the participants.

Firstly, the project required data input in a specific format, and in amounts that exceeded what was useful in the context of the individual research practices of the participants. An element that increased

organizational flexibility was the modularization of data input as a separate work step (thus mimicking the established division of labor between bibliographers and scholars). While this allowed participants to delegate parts of the seemingly unscholarly labor of data input, it has also limited opportunities for them to find ways of combining data-driven analysis with substantive research skills. Another tension arose from the ambitious scope of the project, which in the short term produced more uncertainty than firm empirical knowledge. The database necessarily remained fragmented, insofar as its coverage was the result of bursts of punctual data input. Rather than a resource participants readily could draw on for the production of publications, the database revealed itself as a potential research object in its own right. The database could not replace bibliographies and archival resources. Instead it implied a wholly different relation between scholar and research technology, one in which the database functions more as a heuristic tool rather than an empirical authority. The theory of infrastructure studies (Bowker & Star, 2000; Edwards, 2010; Star & Ruhleder, 1996) predicts that if such a model would become more widespread, we would also a witness a change in the conceptual deep structure of literary studies. The current canonical literary history is ultimately an expression of a scholarly landscape dominated by the monograph-oriented model. In Whitley's terms, it represents the kind of knowledge one gets if a field is characterized by low degrees of mutual dependence and high degrees of task uncertainty, thus forcing scholars to remain in the scope and conceptual sphere of a bibliographically oriented perspective. A wider diffusion of database-oriented scholarship could effectuate a partial increase in mutual dependence and a reduction in task uncertainty. This would potentially supplement literary history with analytical categories that illustrate relations on a large geographical and chronological scale, such as networks of production and reception. It could also entail a change in the granularity of individual contributions to disciplinary knowledge. Rather than having to provide individually 'complete' narratives (which implies a certain need to appeal to a higher bibliographical authority), scholars could validate their contributions more in relation to the contributions of their colleagues.

As it stands however, the difficulties the participants experienced in the attempt to reconcile their individual research with collaboration in the project mark the distance between two different research models: the monograph-oriented one that currently enjoys almost exclusive dominance, and an emerging research model that encompasses elements of database-related work. Infrastructure studies stresses that skills, conceptual

frameworks, and instruments used in a given field of research are reproduced together with criteria for valid contributions to disciplinary knowledge. By learning how to handle the tools of the trade and produce publishable papers and monographs, neophytes also internalize what it means to do proper scholarship. Inevitably therefore, the activities that are necessary to make a transition from an exclusively monograph-oriented research model towards one that is compatible with database work will look unattractive from the perspective of the status quo. Data input and digital quality control do not look like tasks that a professor, i.e. somebody who is very advanced in the established disciplinary hierarchy, can afford to spend much time on.

An important question for digital humanists, policy bodies and infrastructure designers therefore is, how can those tensions be reduced to a more manageable level, so as to provide interested scholars with opportunities to engage in new forms of research, and thus to broaden the variety of approaches to literary history? This is a challenge that much concerns developers, who tend to react by aiming for more customizable, faceted search engines (personal communication). While these are worthwhile efforts, the problems identified in this paper cannot be completely solved by a 'technological fix'. Even the most customizable search engine must always be based on a data model that is standardized at the back end, thus leaving a certain residual tension between individual hermeneutic freedom of scholars and the need for organizational integration. Another possibility is to think about what funding formats are most appropriate.

One of the findings I present in the above is that the two organizational models proved better reconcilable in instances where organizational authority was conceded to more local contexts. In the case of the Serbian sub-project, the smaller, regional scope of the database reduced the labor needed for data input and quality control. Moreover, the general level of disciplinary knowledge on women writers in a given country or region, as well as their relative degree of 'canonicity', seemed to influence the potential for compromise. For the participants of the sub-project, the generally understudied role of women writers in Eastern European literary history provided particular opportunities to reconcile project interests and individual career interests. This observation can be generalized insofar as the possibility to systematically achieve this sort of compromise will diminish with the scale of a project. Small and medium size projects will be more sensitive to local disciplinary context than large ones, which in turn

will increase their chances to attract additional local funding. If database-oriented project can be seen as interventions into the organizational landscape of literary studies, then small and medium scale projects constitute less radical interventions. They will tend to make the shift of defining analytical goals from a strictly individual activity to a more collective one less steep. Finding a middle ground between individual theoretical freedom and project needs will less overtly feel like a pragmatic concession. At the same time, small to medium size projects will prevent the need for harshly dichotomous divisions of labor between data work and analysis, thus creating better opportunities for digital skill diffusion. In Whitley's terminology, initiatives of this scope provide a means to increase mutual dependence and decrease task uncertainty in more subtle, but also more sustainable ways than very large scale projects.

However, it is here that disciplinary contexts tend to clash with management interests. From the perspective of actors who are interested in creating a pervasive digital infrastructure for the humanities, smaller scale projects may be negatively seen as 'boutique projects' (Friedlander, 2009) or 'data silos' (Zorich, 2008). Such projects may be intellectually beneficial for participating researchers, but as investments in the service layer of future scholarship, they would be considered a waste. Critics of 'boutique projects' often seen them as the result of poor coordination of individual efforts and as distracting funds from larger systemic goals (Borgman, 2009). A problematic implication of this view is that digital humanists are primarily cast as providers of information services, and only in a secondary sense as scholars with a firm grounding in the disciplinary landscape. But it is only in regard to the goal of a pervasive, content-centered digital infrastructure that smaller, research-oriented projects and disciplinary fragmentation of the humanities appear as a problem in the first place. My case study demonstrates that such projects in fact may constitute emergent compromises that allow scholars to strike a balance between investment in extant research models and digital scholarship. What is necessary here is greater appreciation for the evolutionary development of knowledge infrastructures. If there is indeed a critical amount of interest in using digital technology in the humanities, for example to tackle such problems as revising the literary canon, then this endeavor will be a long-term one. Realizing it will entail small steps that may be seen as underwhelming or even as failures when measured against certain policy ambitions.

For one, the preceding case study confirms again my conceptual assumption that the value of particular tools does not depend on inherent features, but on how meaningfully they can be incorporated into a historically grown infrastructure. New features – if not carefully adapted to disciplinary needs – can actually undermine existing scholarly practices. The participants of the COST Action hoped that the shared use of a digital database would allow them to draw together their individual efforts, thus creating knowledge that would be more than the sum of its parts. They were forced to realize, however, that the dedicated use of the database implied an organizational model of scholarly work that was at odds with the epistemic and social structure of their field, effectively making participation in the project a potential risk for disciplinary career development.

To better understand this widely perceived mismatch between project requirements and individual career requirements, I have deepened an analytical theme introduced in chapter 1, namely the conceptualization and valuation of particular forms of scholarly labor (Strauss & Star, 1999; The Virtual Knowledge Studio, 2008). Thereby I have been able to show that intellectual considerations regarding the use of novel scholarly approaches in a field cannot be separated from questions about the relative prestige associated with particular forms of work - the very reproduction of disciplinary paradigms in literary history actually appears to be intertwined with the reproduction of traditional fault lines in the academic labor hierarchy. This can have seemingly paradoxical effects. On the one hand, all participants of the COST Action welcomed the intellectual perspective of engaging in a more data-intensive, comparative mode of scholarship, so as to overcome the limitations of traditional literary historiography. At the same time, those who were employed in advanced research positions at universities also tended to advocate a traditional division of labor between 'scholarly' tasks and 'subordinate' data work, thereby undermining the aim of establishing a new disciplinary paradigm.

These findings could be read as an indication of pronounced inflexibility in the organization of academic work, or perhaps even of a culturally engrained resistance to innovation in the humanities. In fact, ardent proponents of digital scholarship regularly level such charges against their 'traditional' colleagues (ACLS, 2006; Wouters, 2007). It is important to remember, however, that the intelligibility of knowledge depends on the very epistemic and organizational inertia of the scholarly infrastructure (Bowker & Star, 2000; Star & Ruhleder, 1996; Ciborra & Hanseth, 2000). My

analysis has shown that it is by maintaining compatibility to established disciplinary conventions regarding the scope and conceptual formats of research that scholars are able to draw on the work of preceding generations. If, by contrast, different datasets form very different research contexts are simply put together without significant further effort at quality control, harmonization, and the adaptation of existing research practices, information simply does not 'add up'.

Rather than dismissing the difficulties encountered in the COST Action as a sign of individual or structural conservatism, I would therefore draw the following two conclusions. Firstly, a change of focus in current debates about the use of data-intensive research in the humanities - currently often summarized under the label big data – is warranted. The common narrative usually pictures the benefit of big data as a sort of emergent property of bringing together large datasets, and thereby tends to downplay the large amounts of data work that are necessary to harmonize information. This goes particularly for the humanities, since the kind of data that scholars will usually be interested in are unlikely to be produced by any other party (in contrast to, say, economists or social scientists interested in demographic or financial datasets produced for administrative purposes).

Secondly, the work of incorporating new tools in scholarly practices should be performed with an awareness for the epistemic and organizational characteristics of a field, as well as the structure of academic careers. A practical result of such reflexivity in the COST Action has been the adoption of a smaller project format, which has turned out to better suit the disciplinary characteristics of literary history. To be sure, this adaptation of organizational modalities has been a result of coincidental experimentation, rather than intentional design, and we can easily imagine that comparable tensions remain unresolved, or, are resolved differently, in other projects. It thus seems that reflexivity is a choice that scholars can take when incorporating new tools into their practices, but it is not an inevitable one, and neither will it lead to a predictable outcome. In the following chapter, I will further investigate the function of reflexivity through a more in-depth case study of the work required to mutually adapt research practices in a collaborative project involving humanities and computer science.

Chapter 3

Decomposition as practice and process: creating boundary objects in computational humanities[46]

Parenthesis – relation to conceptual framework

The single most commonly heard expectation towards the epistemic advantages offered by digital approaches in the humanities arguably is that of enabling data-intensive research in previously data-sparse intellectual traditions. Some highly publicized initiatives – such as the Harvard-based culturomics project (Aiden & Michel, 2013), or the cultural analytics lab founded by media theorist Lev Manovich - have proposed to apply data-mining and algorithmic analysis to trace developments in language, the arts, and popular culture. Many European and US policy views of digital infrastructure are similarly based on the idea that the provision of datasets and analytical tools will enable the humanities to pose wholly new types of research questions, thus allowing to move beyond the limitations of empirical material and secondary literature 'lone' scholars can cover in the process of writing a monograph. The joint international funding initiative Digging into Data for instance pictures data as a sort of basic empirical layer around which academics from different backgrounds can gather and organize their work (Williford & Henry, 2012). These imaginings are complemented by attempts to retell the very history of the humanities as a chronology of primarily empirical, data-driven research, thus implying that the turn to hermeneutics in the late 19th century is best understood as an accident (Bod, 2014).

Against the background of my conceptual framework of a scholarly infrastructure, in which particular disciplinary cultures of knowledge making are loosely related in a larger work ecology, such accounts warrant critical scrutiny. For one, the commonly heard language of 'enhancing' the humanities by overcoming 'limitations' of 'data scarcity' suggests a problematic hierarchy of empirical exactitude and development, where the humanities have yet to attain the epistemic robustness of the sciences.

46 This chapter has been published as: Kaltenbrunner, W. (2014) Decomposition as Practice and Process: Creating Boundary Objects in Computational Humanities, *Interdisciplinary Science Reviews* 39(2), pp. 143-161.

Secondly, the notion that incorporating data-intensive approaches into the humanities is a simple matter of 'applying' new tools and methods to scholarly problems is based on a reductive epistemological vision, in which data are seen as agnostic with respect to disciplinary cultures. In the preceding chapters I have already analyzed a variety of data-related issues that arise in the attempt to reengineer the scholarly knowledge machine, most prominently that of data work and the role of its valuation for the type of knowledge produced in a field. In the following chapter, I will address data as a problem for interdisciplinary collaboration between computer science and the humanities. Combining infrastructure studies and STS theorizing on scientific method, the characteristic features of modern disciplines – and thus the way they use data to produce knowledge - can be seen as the emergent result of a historical process. Due to repetition of particular research practices over time, elements of research work have become packaged into standardized sequences, for example commonly used data formats, methods, and ways of communicating with peers (Latour, 1987; Fujimura, 1987, 1992). It is by conforming to these specific disciplinary traditions that individual researchers are able to draw on the work done by their predecessors, albeit on the condition that they also adopt the underlying normative assumptions. STS scholar John Law (2004) therefore argues that methods, be they data-intensive or qualitative, are not objective devices for extracting truth out of messy social and cultural contexts, but instead enact a reality according to specific disciplinary conventions. Academic disciplines can in fact be seen to produce implicit and explicit criteria for what individual contributions to the shared body of knowledge must look like (in terms of format, style, methods used) so as to be acceptable within the community of peers. Simultaneously, they provide institutionally embedded facilities (methodological training for undergraduates, available lab equipment, commonly used databases and software) that make it easier to produce exactly such contributions, but not others. Together, established traditions and facilities constitute what Law calls the hinterland of a discipline.

Applied to interdisciplinary work in digital scholarship, this perspective raises both normative and practical questions. If methods enact reality, rather than merely extracting it, normative choices need to be made about what forms of reality to make possible and which to exclude. Practically, it suggests that research at the intersection of humanities and computer science will not be a simple process of 'applying' large datasets to scholarly research questions. Instead, we can expect what Edwards et al.

(2011) call 'science friction', i.e. difficulties in collaboration that arise due to diverging traditions of conceptualizing, analyzing, and practically handling data. Some sort of reflexive work will therefore be necessary to capture differences between disciplinary hinterlands and so be able to create compatibility.

Introduction

Computational humanities, although practised by an international community of scholars since the late 1940s (Busa, 1980, Hockey, 2004), has recently caught the attention of policy makers in both Europe and the USA. Substantial amounts of funding are currently being invested in individual projects as well as in the coordinated creation of cyberinfrastructure, in order to promote computational approaches to the study of literature, art, and history (e.g., Anderson, Blanke & Dunn, 2010; Williford & Henry, 2012). The research published in journals such as *Literary & Linguistic Computing* over the last decades is a testament to the possible interdisciplinary fusion of computer science and humanities scholarship. However, the current policy investments in new computational humanities projects constitute a somewhat different situation. While research published in the above mentioned journal in fact has acquired typical features of disciplinarity over the years — for example, established types of research questions and methodological standards — current policy efforts encourage computational approaches also in other areas of humanistic inquiry with no such history. Moreover, although scholars in the older tradition of computational humanities typically dispose of a hybrid skill set that includes both programming and scholarly domain expertise, the recent wave of collaborative projects tend to operate with a division of labor between computer scientists and humanities scholars (cf. Kaltenbrunner, 2014). Here, the collaborating researchers find themselves in a situation that requires them to work out a viable interdisciplinary arrangement from scratch and in relatively little time.

Policy views on the practice of computational humanities, however, are frequently informed by an insufficiently complex understanding of the dynamics of disciplinarity, research practice, and technology. In keeping with the bold expectations associated with 'big data' and the 'Fourth Paradigm' (Hey et al., 2009), many policy reports on the potential of computational humanities express the expectation that the shared use of large datasets will bring about a unified culture of computationally intensive research (Williford & Henry, 2012; Willikens et al., 2010). Once the required cyberinfrastructure is in place, the assumption is that scholars from all areas of study will naturally find themselves engaging in increasingly intensive teamwork across their respective specialization.

In this study, I subject the process of establishing interdisciplinary collaboration in a computational humanities project to a detailed empirical analysis. My account contrasts with the picture of a smooth, data-driven

synthesis of different fields by making visible the tensions and hard work involved in that process. The analysis is based on my fieldwork in a Dutch computational humanities project, where computer scientists work together with researchers from Indonesian Studies and Network Analysis to study the changing relations between politically influential actors in Indonesia. I make use of two theoretical concepts from Science and Technology Studies (STS) to analyze the collaboration. Firstly, Law's (2004) concept of hinterlands, which describes sedimented socio-material practices that constrain how researchers can structure their research processes in given fields. Combining this perspective with Star & Griesemer's (1989) concept of the boundary object, I theorize the unfolding collaboration as the reflexive search for a viable organizational arrangement that allows the participants to work together without giving up existing disciplinary commitments.

I firstly offer a theoretical discussion in which I introduce in more detail the above mentioned conceptual resources. Turning to my case study, I then describe an initial attempt of the project participants to organize their shared research process around the dataset as an organizational pivot. This, however, created tensions owing to diverging perspectives on the 'nature' of data and their function in different disciplines, which in turn prompted the participants to envision more context-sensitive ways of embedding computational approaches with scholarly practice. In the last section, I critically assess a European cyberinfrastructure initiative that similarly tries to respect the specificities of scholarly practice in the humanities. Here I reflect in particular on the inherent tension between 'mutual shaping' of digital tools and their users on the one hand, and the policy interest in efficient, functionalist design principles on the other.

Case and methods

The Elite Network Shifts (ENS) project was a successful applicant to the Royal Netherlands Academy of Arts and Sciences' newly developed program Computational Humanities. Launching in February 2013, ENS is a collaboration between Indonesianists, network researchers and computer scientists (subdiscipline information retrieval, IR). Its aim is to gain new insights about sociological developments in groups of elite actors in Indonesia, with an empirical focus on two periods of political upheaval. Firstly, the period of decolonization and the rise to power of the military dictator General Suharto (1945–1955). Secondly, the period around the downfall of Suharto and the subsequent democratic reforms (1991–2010).

The empirical basis for this research is constituted by large amounts of digitized newspaper content (in Indonesian, Dutch, and English), which is processed through a combination of natural language processing (NLP) and statistical analysis, so as to filter out names of elite actors and visualize them in networks. ENS thus is based on a specific division of labor: the dataset is a product of the work of the computer scientists (Ridho and Maarten), while it constitutes the empirical basis for the research of the Indonesianists (Jacky, Gerry, Fridus) and network researchers (Vincent and Andrea, who hold doctoral degrees in applied mathematics and physics). The anticipated benefits of collaboration are mutual. Indonesianists and network researchers get access to large amounts of tailor-made empirical material. The IR researchers in turn get the chance to study the search behavior of humanities scholars, which is an important precondition for designing better data extraction algorithms, search engines, metadata etc.

The empirical materials on which I base my analysis of interdisciplinary collaboration in ENS were collected in a variety of ways. Firstly, through participant observation in project meetings, spread over a period of ten months (~25h). Secondly, I conducted a first round of semistructured interviews with all project participants about three months into the project, and then a second round of interviews with one project leader and the three main research participants (Ridho, Vincent, Jacky) eight months after launch. I also collected data from electronic communication within ENS, both by following email conversations and by joining a Zotero group library, which the participants used heavily to exchange project-internal documents, ideas, and papers. My role as an 'embedded' ethnographer was explicitly part of the original research proposal. Although I did not have an employment contract in ENS, I was invited to participate by attending internal meetings. Moreover, I was asked to recurrently present my findings and ideas to the group, thus helping the participants to reflect on the unfolding collaborative process.[47]

Hinterlands and boundary objects

In his provocative book After Method, John Law (2004) introduces the notion of the hinterlands of the social sciences. Such hinterlands are

47 I wish to thank the participants of ENS for allowing me to conduct fieldwork in their project. I am also particularly grateful to Stef Scagliola at the Erasmus Studio Rotterdam, with whom I have had fruitful discussions about the project.

constituted by the often unquestioned methodological apparatus that underlie scholarly knowledge. Examples include both qualitative and quantitative approaches, such as the statistics of the t-test, or the methodological precepts that inform grounded theory. Hinterlands make possible certain forms of knowledge — they both constrain and enable what the researcher can see/say. A research argument that draws attention to one aspect of a phenomenon inevitably pushes into invisibility other aspects. Law argues that there is often a problematic normativity attached to method in the social sciences, in the sense that it is often seen as a secure way to uncovering the actual structures of reality. Law's own assumption about the reality social scientists study is that it produces a surfeit of generative potentials. Method can be used to selectively amplify some of these potentials, thus creating snapshots of a certain reality, without, however, ever exhausting other possibilities. Law's main argument is that social scientists should make an effort to dig into their hinterlands, in order to realize their inevitable blind spots, and to think about what other kinds of knowledge it might be desirable to generate. A number of factors militate against such reflexivity, however. Hinterlands in fact can be seen as sedimented research practices, which have acquired a structure-like quality through repetition over time. These practices become part and parcel of what it means to be a researcher in a given field. Similar to what Bowker & Star (2000) have observed for socio-material infrastructures, hinterlands therefore tend to become invisible to the researchers that inhabit them, precisely because they are so thoroughly embedded with everyday disciplinary routines. But even if a researcher is committed to reflexivity, going against the grain of a given hinterlands can pose a risk for disciplinary career development. Questioning the very foundations of a hegemonic methodological approach often involves the mobilization of significant amount of resources (cf. Latour, 1987), not to mention possible difficulties to get unconventional research published in 'flagship' journals.

Interdisciplinary collaboration constitutes an occasion where different hinterlands are intersected. Each of these comes with established disciplinary ways of structuring the research process, for example in terms of how empirical materials are used, what type of research questions are posed, and the way theoretical frameworks are expressed in specific methods. Interdisciplinary researchers typically expect such collaboration to complement, rather than replace, their disciplinary career development. Their participation thus should contribute to the shared process, but without forcing them to abandon their methodological, theoretical, and praxeological

investment in a given discipline. In ENS for example, participation in the project should not only advance our knowledge of elite shifts in Indonesia, but also result in a PhD thesis for computer scientist Ridho, and in journal publications for the postdocs Jacky and Vincent.

Here it is useful to draw on Star & Griesemer's (1989) concept of the boundary object. A boundary object is an organizational element that connects the activities of actors from different social worlds. Those shared objects are interpretively flexible. They mean different things to different actors, thus satisfying the requirements of their respective social world, but they are robust enough to maintain a certain integrity across those contexts. The concept of the boundary object was originally developed in reaction to Latour and Woolgar's (1979) notion of translation. Translation in this latter sense is a move by which entrepreneurial scientists transform interests of individual actors into goals that are conducive to their own ambitions — translation thus can be seen as the attempt to enroll allies into a socio-material network, which in turn can be used to stabilize a particular scientific fact (Callon, 1986). This is often a competitive process: multiple entrepreneurs try to enroll the same actors for different purposes. However, by virtue of their interpretive flexibility, boundary objects allow different actors to cooperate while maintaining ties to their original social worlds. The translations involved in creating a boundary object thus do not occur along a single axis (actor A tries to enrol actor B by translating her interests, thereby disciplining her), but are multidirectional (various actors try to mutually interest each other in their respective objectives). This multidirectional translation will entail significant amounts of negotiation, given the need to balance diverging actor interests. By definition, the creation of boundary objects is an emergent process that will tend to resist attempts to specify outcomes in too much detail in advance.

Here I am studying the process of establishing a viable collaborative organization in the context of computational humanities. If boundary objects are to emerge, collaborative modalities have to respect strong extant commitments of researchers to their respective hinterlands, rather than simply override them. Romm (1998) has argued that interdisciplinarity always involves a significant amount of reflexivity, insofar as it tends to throw into relief the differences between fields. Developing this point further, I propose to think of the process of creating boundary objects in computational humanities in terms of the metaphor of decomposition: it requires a reflexive effort in which the collaborating actors iteratively discover various aspects of their hinterlands that constrain the possible

modalities of interdisciplinary organization. Ideally, this decomposition leads to a viable arrangement that allows for meaningful collaboration in the project, yet without forcing researchers to simply abandon their original disciplinary affiliation.

My argument thus is in keeping with a key finding of the existing literature on interdisciplinarity, namely that the very establishment of a shared process should be considered a significant product of interdisciplinary research in its own right (Jeffrey, 2003; Lyall et al., 2011). This insight, however, is downplayed by the recent revival of a universalist expectation towards data. Currently, substantial amounts of funding are being invested in the creation of cyberinfrastructure for the sciences and the humanities, both in Europe and the USA (Bulger et al., 2011). The concept of cyberinfrastructure is heavily informed by the assumption that a new paradigm of scientific work is upon us, one in which research is driven by the detection of patterns in large quantities of data. Buzz words that express variants or aspects of this assumption include the 'data deluge' (Hey & Trefethen, 2005), 'big data' or the 'Fourth Paradigm' (Hey, 2009; Williford & Henry, 2012). A key promise of this discourse is that data will speak for themselves if available in sufficient amounts, thus liberating researchers from the constraints of disciplinary methods and theory-building (Anderson, 2008). Such an assumption certainly has managerial advantages, at least on paper. The UK e-science program for example conceptualizes data as a basic, self-identical unit, similar to an atom (Hey & Trefethen, 2002). Importantly, if data are seen as atoms, they can be expected to travel across disciplinary boundaries without losing their integrity. This in turn allows instrumental relations to be specified between producers and (re)users of data on a very large scale. Cyberinfrastructure can then be conveniently conceived as a large data repository, which merely needs to be overlaid with a layer of middleware and interfaces that cater to the more specific disciplinary needs of its users (Wouters & Beaulieu, 2006). The 2010 report on computational humanities by the Royal Netherlands Academy of Arts and Sciences squarely fits this perspective. Research in computational humanities, thus the report, should essentially be organized on top of a shared layer of digital data. The Academy anticipates that the most interesting areas for research is in the formalization of knowledge and perception, for example by the parametric modelling of the interpretive horizon against which pieces of arts are received by their audiences. Such formalization finally allows the humanities to produce knowledge that is 'not hindered by frame problems [resulting] from narrow disciplinary

perspectives' (Willekens et al., 2010, 10). Although conceding that this requires scholars to adapt their current practices, the report suggests that such adaptation constitutes merely an 'enhancement' of the epistemic goals that the humanities have always pursued, namely 'the search for high-level concepts, patterns and motifs in humanities data (Willekens et al., 2010, 11)'. From this perspective, interdisciplinary collaboration between scholars and computer scientists is primarily a matter of undoing the fragmentation of knowledge production into disciplines, which will 'naturally' follow from acknowledging the universality of data.

In this study, I argue that a more complex understanding of interdisciplinary collaboration in computational humanities is required, both to avoid stereotyping the latter as neo-positivist, and for the practical purpose of overcoming fundamental conflicts between disciplines that would otherwise threaten the success of costly initiatives. The analytical framework I have outlined in the above is of course revealing of my own academic socialization in the hinterland of STS, which has a long tradition of providing detailed, ethnographic studies of situated knowledge practices (Latour & Woolgar, 1979; Knorr Cetina, 1981). In this sense, my analytical categories co-produce the very tensions that I argue are characteristic of interdisciplinary collaboration. I would argue, however, that this perspective is selective in a useful way, because it encourages us to take seriously the challenges and concerns researchers are faced with when entering interdisciplinary projects, rather than alienating them by assuming a historically elusive universalism of scientific practice.

Different hinterlands in ENS

As is typical for the early phase of interdisciplinary collaboration (Lyall et al., 2011), the first few months of ENS were characterized by intense exchanges among the participants, during which they familiarized themselves with the specific disciplinary expectations and habits of each other. Through the theoretical lens of Law's approach, these can be seen as revealing of distinctive features of different hinterlands.

A first difference consists in the rhythm and material formats in which traces of the research processes become visible and circulate within the project. For example, the Indonesianists were astonished by the speed and regularity by which PhD candidate Ridho produces NLP algorithm prototypes, thus creating the impression that the work of extracting a dataset from the digitized newspapers was close to being done. This is

evidence of a more experimentally oriented research culture in IR that produces rapid preliminary output, and it contrasts with the one in Indonesian studies, where a scholarly narrative sometimes is crafted over years. In an interview I conducted only shortly after the launch of the project, Indonesianists Gerry and Fridus reflect on Ridho's already significant progress in writing his first academic paper.

> Gerry: In our tradition, a PhD student would not dream of doing a paper in the first five months. They are still thrashing around, reading everything, and deciding what they really want to do. They might write a paper in their third year. Where they begin to synthesize what they have learned, and they would adopt a position on something that we would think is worth adopting a position on.
> Fridus: Part of the papers in computational [science] is... is describing what you're doing. While in our discipline we have to do something and then we can find something and then we have to start analyzing. (...)
> Author: So the research is less prespecified? So what you will do is part of what you have to find out?
> Fridus: Yes.[48]

Furthermore, the hinterlands of IR, network research, and Indonesian studies are characterized by different assumptions about where aspects of knowledge reside. Interestingly, the bibliographical references Jacky posted in the ENS Zotero group library frequently stress the complexity of the political developments under study in contemporary Indonesian history, and that investigating such complexity requires intimate personal familiarity of the analyst with these. The blurb of a standard work in contemporary Indonesian history reads as follows:

> Periods of major political transition are generally so complex as to present the political analyst with one of his most difficult challenges. Indonesia between 1957 and 1959 was no exception. (...) Dr Daniel S. Lev is particularly well qualified to examine the course of Indonesian political developments between 1957 and 1959. Arriving in Indonesia towards the end of this period, he remained there for three years engaged in an intensive study of its political life. His

48 Personal interview with Gerry van Klinken and Fridus Steijlen, March 2013, Leiden.

> monograph constitutes by far the most searching analysis yet to appear of this critically important period (Lev, 1966).

'Data' here are treated as inseparable from the individual, embodied knowledge of the prominent Indonesianist Daniel S. Lev, who in fact used his body as a research instrument (which was physically transported to Indonesia and spent several years there). This is a very different approach from the one envisaged in the project, where data are extracted from newspapers. The assumption underlying the latter approach is that the process of generating data can be partially separated from the process of analysis, which is a practical requirement for the division of labor between IR researchers and Indonesianists.

Furthermore, there are disciplinary differences in the ways researchers reduce the complexity of the studied phenomena. For example, in some social sciences it is common practice to remove outliers from a sample when testing it for statistical significance. In other disciplines this would be frowned upon, since disciplinary interest may exactly lie in what those outliers have to teach us. To give another example, studies on NLP and automated event-coding applied to newspaper content often stress the increased efficiency of automated approaches in comparison to manual coding. One such study (Shellman, 2008), also posted in the ENS group library, suggests that in time, improved coding algorithms will be able 'to capture the event itself'. This claim is based on an underlying theoretical choice about how to define an event, as well as an implicit source-critical assumption ('newspapers record events more or less accurately, therefore large amounts of news data allow most relevant events to be covered'). Some of these choices seem problematic from a social sciences perspective. For example, Jacky recurrently emphasized the importance of thick description, which implies that the significance of an event can only be gauged after the analyst has gained significant familiarity with the specific culture she studies. This is at odds with the a priori definition of an event as in Shellman's paper. Another potential tension could arise from the choice of treating all newspapers as being on equal footing. Traditional source criticism would point to the political bias of individual newspapers, which is also going to influence which actors and events they cover in the first place. This poses a certain challenge to the practice of taking newspaper content out of its original context through algorithmic processing, with the context being represented only by a limited amount of metadata.

Data as a link between hinterlands?

The original research proposal envisaged that the collaboration should be ordered around the dataset as an organizational pivot. Specifically, the document anticipates that about two years into the project, the participating computer scientists will have extracted a dataset from the digitized newspapers (subproject 1), which will then form the basis for the research of the Indonesianists and network researchers (subprojects 2 and 3).

> In subproject 1, we will automatically extract entities and relations between entities from large historical news corpora (...). Subproject 2 will adopt a sociological and historical perspective and use techniques from social network analysis to trace central actors (identified in subproject 1) in their different social relations over time. In subproject 3, we focus on structural properties of social networks and their evolution over time from a statistical physics, complexity point of view.
>
> Based on the extracted relational information, two subprojects are devoted to network analysis. Network analysis faces a trade-off between shallow, quick and effective analysis using minimal language processing tools, or deep but more laborious and risky analysis making extensive use of linguistic analysis. The challenge lies in the sociologically meaningful interpretation of network-analytical results and the emergence of new research questions for mathematical network analysis tools resulting from the empirical study of real-world historic networks. (Oostindie et al., 2011)

The project agenda implies a collaborative path-dependency, and it tends to present a somewhat dichotomous division of labor. First, it foresees that an algorithm will be developed that allows a dataset to be extracted from the digitized newspaper content. This involves custom NLP techniques for the identification of relevant entities in the newspaper articles — in this case, names of actors who are potentially part of elite networks. Moreover, the work of applied mathematician and network researcher Vincent allows co-occurrence based networks to be constructed from these data, i.e. a statistical analysis to single out patterns of recurring links between two actors who are mentioned together in the same sentence of a newspaper article. There is initially no theoretical understanding of what those recurring links mean. Co-occurrence of two actors can be completely random, or it can indeed

point to particular relations that are of interest to students of elite networks, such as kinship or business contacts. To be sure, the plan is that the expert knowledge of the Indonesianists informs the NLP algorithm and the construction of co-occurrence based networks. However, the assumption is that after this step is concluded, the resulting dataset provides a stable empirical basis that allows Jacky to answer her research questions. The proposal assumes that a dataset is made meaningful to the Indonesianists by operationalizing concepts such as 'elite', 'elite circulation', and 'regime transition' as empirically quantifiable data categories, so as to create a bridge between the disciplines. Moreover, the proposal is quite specific as to what kinds of insight the extracted dataset should (must) yield: 'otherwise hidden network relations between key actors', 'correlations between events over time', 'the role of individuals who co-act in different networks at the same time' (Oostindie et al., 2011).

However, this approach proved to be a cause for concern for several of the participants. In fact, project leaders Fridus and Gerry indicated that one of their greatest worries is the possibility that the resulting dataset turns out not to make 'sociological sense' after all.[49] Similarly, at an early project meeting, Indonesianist Jacky voiced her opinion that a quantitative approach might easily brush over many empirical details that can only be appreciated in a hermeneutic case-by-case approach. This could make it difficult to integrate an algorithmically extracted dataset into her personal research.

> In terms of methodologies, I am generally wary of the statistical and quantifiable. To my mind, it can lead to either addressing only those aspects of social phenomena which are easily measured (and so, often, the least interesting) or, if more complex phenomena are studied, there is a danger that the assumptions behind statistical or technical procedures are hidden or insufficiently examined. (...) I want to avoid a situation where I receive a lot of data from Ridho and Vincent and then try to situate it within the detailed contextual knowledge of the Indonesianists (ENS, internal document).

Not least, Jacky reported that upon presenting the original outline of the ENS research proposal to Indonesianist audiences at conferences, she was recurrently faced with the objection that those elite dealings that are of real

49 Personal interview with Gerry van Klinken and Fridus Steijlen, March 2013, Leiden.

interest to scholars are unlikely to be reported in newspapers. Highlighting shifts in, say, the changing frequency and composition of elite names as reported in the news here was generally perceived to be at odds with the conceptual fuzziness of notions such as 'elite' and 'power'.

The source of these tensions, I suggest, ultimately is the diverging function of data in the various hinterlands that are involved in ENS. As argued in the above, data in Indonesian studies is not a clearly differentiated organizational element, but bound up with highly individualized ways of doing research. Generally, and in contrast to the experimental nature of IR with its multiple preliminary research products, the monograph/paper format dominant in the humanities and qualitative social sciences implies a more weakly differentiated research process. When working on a monograph, scholars have considerable freedom in adapting their narrative to new, conceptually relevant insights. It is furthermore assumed that the receiving disciplinary community takes the time necessary to read the entire narrative of a monograph or paper, rather than merely 'extracting' any empirical information it might contain. In addition, while theorizing in IR is more ad hoc and need-driven (e.g. 'how to explain this particular aspect of the search behavior of a user group?'), theory in Indonesian studies normally refers to larger theoretical projects (e.g. 'how are social values reproduced in a society?'), thus indicating that individual research results are primarily meaningful in relation to an encompassing disciplinary discourse. These theoretical projects relate in complex ways to empirical materials and data categories, provided that the latter term is even used in a given hinterland. In many disciplines of the humanities and social sciences it is customary to pose research questions that are of a hermeneutic, open-ended nature, thus defying the possibility of an ultimate, empirically based answer. In other words, the tension arises from treating research problems concerning elite shifts as something that can be uniquely answered by the data that are being produced by computer scientists and network researchers, whereas research problems in Indonesian studies are normally seen as irreducibly open-ended, hermeneutic affairs. The dataset thus is treated as a mechanical 'joint' that connects the work of the computer scientists to the work of the Indonesianists. This, I suggest, creates the rather high expectation that the extracted dataset is (must be) 'sociologically meaningful' in itself.

In contrast, the hinterlands of IR and network research were much easier to intersect, owing to the strong consistency in the understanding of data across those fields. Vincent could integrate the tables of actor names

that resulted from the application of Ridho's NLP algorithms into his own statistical analyses (of network properties such as the amenability of the data to particular clustering methods, the relative density and size of these clusters, as well as their relative persistence over time etc.) without any problem. This allowed Vincent to immediately go about drawing comparisons between 'behavior' of the ENS data and that observed for other types of networks, thus providing him with a clear basis for papers that can be submitted to network research journals. The substantial praxeological continuity between IR and network research is perhaps best reflected in the fact that Ridho provided data in such a format that they could be directly uploaded to R and MatLab, the analysis software Vincent habitually uses — in contrast to Indonesian studies, the data could smoothly travel between the hinterlands of IR and network research.

Discussing a problem closely related the one described in this section, Ramsay (2011) argues that the reason why computational humanists have difficulty breaking into the mainstream of literary criticism is that they often fail to properly embed their computational methods in hermeneutic disciplinary discourse. For example, if one begins to approach the work of Virginia Woolf through algorithmic analysis, a conceptual slippage can easily occur in which statistical criteria of validity replace hermeneutic criteria of validity. Statistical methods that can provide solutions to punctual, often binary questions are then applied to hermeneutic, more open-ended questions. Those criteria are qualitatively different of course, because hermeneutic questions per definition cannot be solved. In ENS, the original assumption was that a dataset, once it is compiled, can help to 'solve' the hermeneutic problems Indonesianists work on. Validity criteria that can be usefully applied to the creation of a dataset (e.g., 'is this a valid formalization of the concept of elite'?) are tacitly extended to apply to research questions in Indonesian studies. As a solution to this kind of problem, Ramsay proposes to strip computational methods of the statistical criteria of validity they are usually associated with. Instead of being associated with a rhetoric of 'solving' problems of interpretation, computational methods could be used to refine or even multiply them. In principle, computational methods could be thoroughly in the service of hermeneutics, rather than their opposite.

Emerging boundary objects

The sometimes tense exchanges in the early phase of ENS thus had made

clear that the original project plan — if followed by the letter — would be rather counterproductive, and that more practical ways of creating interfaces, especially between Indonesian studies and IR, were necessary. Circulating an earlier version of this report in the group played a certain role in this process. Project leader Gerry defined reading an earlier version as 'liberating', in the sense that it made the anxiety that had plagued him during the early months of the project explicit. Again, it is worth pointing out that my choice of framing tensions in ENS through Law's theoretical framework may have had a performative effect. Had I as an analyst had less exposure to STS ethnography with its emphasis on the specificity of disciplinary research practices, or had I had less autonomy in developing my argument, the particular way in which the participants thought about the challenge of collaboration might have been a different one.

A project meeting in June 2013 occasioned an extensive discussion about how to properly integrate Indonesianist research practice with the computational approach as outlined in the original proposal. A first idea suggested by Jacky was to extract a dataset, travel to Indonesia, show it to the actors represented in the dataset, and ask them to comment on it. This would constitute a way of fathoming the limitations of knowledge claims purely based on algorithmic analysis, thus potentially allowing the methodology of elite analysis in ENS to be both developed and questioned through a combination with qualitative interviews. Not least, this could lead to new ways of making academic research in Indonesian studies engage with ongoing societal developments in Indonesia.

A second possibility suggested jointly by Vincent and Jacky is to structure collaboration around 'sentiment analysis'. The principle of this approach is to extract a dataset that not only visualizes networks of elite actors, but that also specifies the modality of their relations through a more fine-grained linguistic analysis of sentences (e.g. actor A attacks actor B). The Indonesianists then could continuously specify their wishes as to which sorts of relations they are interested in, thus potentially allowing for meaningful embedding of the data in their work. The goal here would be to avoid a situation where the Indonesianists are confronted with data extracted by IR and network researchers, together with an overly specific assumption concerning the research questions those data should be used to answer. Of course, this could also entail a restructuring of individual research processes, but this would then be the result of an iterative negotiation process.

Particularly noteworthy moreover is a possibility of integrating co-

occurrence based analysis of the dataset with hermeneutic questions that the participants began to envision about half a year into the collaboration. This approach emerged from a project-wide discussion of a draft report circulated by Jacky, in which she compares a number of theoretically informed ways of conceptualizing the notion of 'elite'. Jacky cites three definitions. Building on Max Weber's institutional sociology, a first definition equates elite essentially with those who are in a structural position to override other people's interest, i.e. actors who occupy top positions in politics, the military, bureaucracy etc. (Mills, 1956). A second definition, adopted from a body of literature that criticizes the nominalistic Weberian approach (e.g., Dahl, 1958), suggests that only those individuals who actually shape political developments — overtly or behind the scenes — should be legitimately considered elites. The resulting list of actors would not necessarily be the same as those who happen to occupy formally high-ranking positions. Thirdly, Jacky outlines a method for identifying elites based on co-occurrence, thereby drawing on Vincent's early experimentation with a subset of the ENS newspaper data. By using a variety of clustering techniques, Vincent was able to identify relatively coherent groups of actors recurrently mentioned in Indonesian newspapers. However, rather than choosing a single 'best' definition of elite, Jacky proposes to build her argument on the comparison of the different approaches. Following the agreement of all members to develop this idea, Jacky then operationalized the first definition of elite by drawing on a number of empirical sources, e.g. a list of the 150 richest businesspeople in Indonesia, as published by the business magazine *Globe Asia*, or by manually compiling a list of all Indonesian ministers and members of parliament. In the next version of the paper, Jacky compared Vincent's networks with the lists she herself had generated using the Weberian approach (Hicks, unpublished). The paper offers observations about the relative degree to which Vincent's networks overlap with her own lists, but also uses the comparison to mutually problematize and question the various definitions of elite. Rather than settling the question 'what is an elite?' through applying a single approach, the comparison thus re-introduces a strong hermeneutic dimension in the discussion.

In all of the proposed scenarios, the dataset functions as a boundary object between the subprojects. That is, although allowing all researchers to make use of the dataset, its exact meaning and function differs for each of the participants. The first and the last strategy are structurally similar. In both cases, Ridho first extracts the data through the application of custom

NLP methods, thus allowing him to combine his own doctoral research with his 'instrumental' task of generating empirical material for Vincent and Jacky. The latter two, charged with analyzing the data, can indeed draw on Ridho's work, yet without violating extant commitments to their respective hinterlands. On the one hand, the data extracted by Ridho can be used by Vincent to write papers that interrogate the network characteristics. Simultaneously, this data processing can be used to formulate quantitative statements about the composition and coherence of elites, based on specific operationalizations. On the other hand, Jacky can draw on Vincent's work in her own argument in such a way as to contribute to theorizing in elite studies, namely by using the comparison of different operationalizations to mutually problematize these operationalizations, or by comparing it with the conceptualizations of elite as developed through interviews with the represented actors. The structure of this argument puts to use the computationally extracted data while still accommodating Jacky's hermeneutic agency. The second approach, the one relying on sentiment analysis, has yet to prove its feasibility. Technology here is instrumental in conferring the interpretive flexibility that defines a boundary object. The hope is that fine grained NLP allows for Vincent to conduct his statistical analyses while accommodating the type of hermeneutic questions Jacky is interested in.

Key features of the unfolding collaborative process in the early phase of ENS thus have been reflexivity and emergence. The process is reflexive insofar as it entailed uncovering a problematic assumption in the original research proposal (namely that data can be straightforwardly used as a link between disciplines), as well as highlighting an important constraint (namely that data analysis for Jacky must be combined with a hermeneutically oriented approach that is viable for Indonesianist disciplinary audiences). The process is emergent insofar as a viable collaborative modality has been developed through iteration: from an early attempt to structure organization on the basis of a universalist notion of data towards a more differentiated view and specific solutions for how to make the dataset function as a boundary object. An interesting side effect of this development is that, rather than stabilizing a singular fact about elite shifts in Indonesia, the project is simultaneously producing different knowledge claims, which mutually highlight the limitations of each other: statistically derived claims about elite shifts co-exist with and complement the

hermeneutic reflection on the very meaning of 'elite'.[50] Creating boundary objects in ENS thus entailed a reflection on the epistemological constraints of individual hinterlands, much like the reflexivity Law (2004) calls for. The very diversity of the involved hinterlands in fact functioned as a driver of reflexivity. A collaboration only between IR and network researchers would not have required the same effort at decomposition to arrive at a viable division of labor, given the praxeological continuity between those fields.

Decomposition and cyberinfrastructure design

In the above case study, I have tried to illustrate how the decomposing of disciplinary research processes potentially can yield viable interdisciplinary arrangements in computational humanities. By emphasizing the role of reflexivity and emergence, I have meant to promote an approach that leverages disciplinary specificities in the creation of boundary objects, rather than framing them as a problem that needs to be countered by treating data as atoms. The research policy discourse on cyberinfrastructure initiatives is indeed still very much informed by such a neo-positivist approach, as discussed in the previous theoretical section. At the same time, there are a number of projects that indicate greater awareness of the type of conflicts that this tends to produce. These recent, more context-sensitive initiatives are based on a reflexive approach that bears some similarities to my notion of decomposition. Although this is certainly an interesting development, in this section I critically interrogate to what extent it is actually possible to combine reflexivity and emergence with the creation of large-scale cyberinfrastructure.

A case in point is the agenda-setting paper by Anderson et al. (2010), which illustrates the design principles underlying the European cyberinfrastructure project DARIAH. Anderson et al. (2010: 3782) explicitly acknowledge earlier critique on data-driven approaches to creating digital research facilities for humanities scholarship (Beaulieu & Wouters, 2009). Cyberinfrastructure initiatives, they argue, can only be successful if they acknowledge the specificities of humanities scholarship in comparison to the sciences, as well as the need to involve prospective users in the planning

50 Fujimura (1992) argues that such a co-existence of competing knowledge claims is characteristic of organizing shared research activities around boundary objects. Far from being an obstacle, I suggest, such 'incommensurability' can be seen as a productive irritant for further research.

process. The design of DARIAH itself is based on the interrelated concepts 'methodological commons' and 'scholarly primitives'. The assumption here is that disciplinary research processes can be decomposed in such a way as to filter out processual elements that are shared across all the humanities. Once such elements are identified, they can provide the basis for designing universally applicable digital research tools, dubbed 'methodological commons' (McCarty & Short, 2002). Anderson et al. take their clues from the work of digital humanities pioneer John Unsworth (2000), who first proposed building infrastructural facilities for digital scholarship around 'scholarly primitives'. These include basic functions such as annotating, comparing, referring etc. However, in the understanding of Anderson et al., the primitives identified by Unsworth are still too idiosyncratic for the goal of a really pervasive cyberinfrastructure. They propose to break them down into even more basic processual elements:

> (...) rather than Unsworth's focus on building tools to support discrete practices embodied by the primitives, [this approach] allows us to see scholarly primitives as part of a wider set of activities that could be translated into a set of functions for building a coherent research infrastructure that supports a chain of related activities. For example, we can start to visualize how the scholarly activity of searching, which includes at a lower level of granularity chaining and browsing, and the scholarly activity of collecting, which includes gathering and organizing, could combine to form a linked data infrastructure that allowed researchers to create their own dynamic representations of knowledge from the data deluge that is the Web (Anderson et al., 2010: 3875).

Although not operating with a universalist notion of data, Anderson et al.'s approach thus is still a functionalist one. It presupposes that the totality of scholarly research processes is reducible to a finite set of shared tasks, which can then be partially automated. It is tempting to think of this approach not so much as uncovering pre-existing, universal elements of the scholarly process, but as being instrumental in establishing the universality of these tasks in the first place.

Collins and Kusch (1998) in fact propose thinking about the promulgation of scientific discoveries as an unfolding 'sociology of sameness'. In order to become acknowledged as discoveries, scientific results have to be accompanied by descriptions of repeatable experimental

setups. These indicate how scientific instruments must be configured and used so as to reproduce the 'same' reported findings (Collins, 1985; Shapin & Schaffer, 1985). If deemed convincing, such descriptions help to create collective indifference by the receiving community of scientists to some specific variations in that setup, while at the same time enacting the singularity of the demonstrated phenomenon. This, one could argue, is precisely the effect of designing cyberinfrastructure around scholarly primitives. For example, by assuming that all research processes in the humanities contain the organizational blocks chaining and browsing, Anderson et al. aim to establish a standardized practice of searching. This could create indifference to subtle variations in the practice of searching, but it would also be related to establishing a standard conceptualization of the object that is being searched. The only difference between the promulgation of scientific discovery and the design principle underlying methodological commons is that in the former case, the establishment of sameness is brought about by peer-review and disciplinary acknowledgment, while in the latter case, it is driven by a managerial interest in efficient design of instrumentation. The above case study, however, cautions us to expect substantial friction occurring in the process. Individual disciplinary understandings of 'searching' might be very different from one hinterland to another, provided that they are even explicitly differentiated as an organizational block in the first place. This could create precisely the sort of tension between the need to maintain functioning connections between the newly differentiated process and disciplinary expectations that was initially experienced in ENS, yet with much more limited possibilities to engage in iterative refinement of the decomposition.

To be sure, Anderson et al. (2010) are explicitly concerned with notions of emergence and 'mutual shaping' of technology and its users. Referencing the widely cited literature review by Williams and Edge (1996), they assume that the uptake and actual use of new technology is informed by a complex interaction with the epistemology and social context in a given discipline. Anderson et al. present primitives as a way to organize 'trading zones' around specific digital services and objects (Galison, 1997). Conceptually similar to boundary objects, these trading zones concede potential users some leeway in integrating these shared services and objects into their practices, which the authors assume reduces the friction between extant disciplinary customs and the prespecified purposes. At the same time, Anderson et al. claim that the proposed primitives describe actual basic elements in all humanities research processes. This allows redundant

investments to be avoided in linked-data tools that support the 'same' function.

Trying to reconcile a commitment to mutual shaping with the approach of designing infrastructure around primitives inevitably creates a certain tension. Mutual shaping stresses the emergent character of technology-in-use, whereas primitives are supposedly pre-existing elements of research processes across all fields. Anderson et al. can only deal with this tension by disproportionally emphasizing one specific aspect of the concept mutual shaping. Originally developed to replace techno-deterministic views in which technology is primarily seen to shape the users, the concept stresses that shaping indeed works both ways: technologies shape users in the sense of offering new possibilities; but at the same time, users always adapt technologies to local contexts and needs, rather than following prespecified uses intended by the designers (Bijker et al., 1987). In the use of Anderson et al., mutual shaping primarily seems to mean that scholars are not determined by current disciplinary customs, and that they are in principle free to realize the promise of methodological commons, if only they are willing. Although this is not a straightforward contradiction to mutual shaping, it does confer it a certain normative undertone. Mutual shaping comes to denote something of an encouragement to the users to conform to the holistic managerial values that inform the notion of methodological commons. Ironically, rather than to unpack the tensions that emerge from the attempt to design cyberinfrastructure around discipline-agnostic organizational objects, a theoretical key concept from STS here is applied in a way that tends to black-box those tensions.

Postscript to chapter 3

A core argument of this thesis is that the adoption of digital approaches in the humanities should not be thought of in terms of a 'diffusion' of novel analytical techniques from data-intensive into data-sparse fields. Rather, I have argued that it is best understood as a process of mutual adaptation of technology and infrastructurally embedded, disciplinary cultures. Digital practices will only become properly incorporated if they allow scholars to maintain functioning connections to inert disciplinary conventions. Drawing on a metaphor from infrastructure studies (Edwards, 2010; Edwards et al., 2011), I have suggested to think of the inevitable tension that arises in such reengineering attempts as friction. On the one hand, friction can reach unmanageable levels and so lead practitioners to abandon either their 'home discipline' or the experimentation with digital methods. It can also function as input for a reflexive learning process, however, thus ultimately paving the way for a workable compromise between the infrastructural status quo and novel technological affordances.

My field work in the Elite Network Shifts project has allowed me to study in detail the work that is necessary to accommodate different epistemic traditions in an interdisciplinary collaboration. The process can be summarized in two phases. In a first phase, participants recognized that the division of labor foreseen in the foundational research proposal was based on problematic assumptions, namely on the notion that data can act as a sort of a mechanical link between the task areas of the computer scientists and scholars of Indonesian studies. The proposal foresaw that the hermeneutic question of the humanists can essentially be reframed as an empirical problem, to be solved by sufficient amounts of data. When reporting early findings to their peers, however, the participating indonesiansts had to acknowledge that this approach was not compatible with dominant disciplinary conventions - for other scholars of Asian studies, a paper reporting only the results of a quantitative analysis of patterns in co-occurring names of elite actors did not in itself appear as a viable research contribution, since it failed to engage with ongoing discussions about the proper theoretical conceptualization of political power. In a second phase, the participants therefore began to reflect on differences in the actual role empirical material plays in their specific disciplinary traditions. Increasingly, their aim shifted from generating singular research findings according to the conventions of computer science and network research to juxtaposing and comparing quantitative and hermeneutic ways of framing the underlying notion of power. This has allowed the participants to collaborate in the

project while also respecting disciplinary expectations regarding the format and style of publications, which in turn is a precondition for advancing in the academic career system.

My argument that digital modes of scholarship are most productive when they do not imply a radical departure from the infrastructural status quo may seem counter-intuitive to those who think that the value of novel technology lies exactly in enabling 'breakthrough' type of research. To be sure, new tools may indeed help to generate radically innovate insight. I would argue, however, that such novelty will rarely arise from a paradigm shift that was planned on paper. A more promising policy for innovation in scholarly method is to conceptualize the incorporation of computational techniques in the humanities as a situated activity, with enough room for experimentation and perhaps failed attempts. In such a view, significant changes in the approach or goals of a given project – for example, radically downsizing its empirical scope when the necessary amount of data work cannot be provided, or admitting mutually contradictory knowledge claims within a single collaborative undertaking - would not be considered as a sign of inefficiency or imperfect preparation, but rather as a necessary part of the mutual adaptation of digital tools and different disciplinary practices. Such a policy might also have consequences for the way researchers themselves perceive friction in interdisciplinary work. In fact, whether friction prompts practitioners to quit experimenting with novel methods, or is rather seen as a necessary element of the process, will be strongly influenced by the organizational and administrative culture underlying the collaboration. If significant changes in a project are admissible, chances are higher that practitioners may still work towards interesting results, rather than (perhaps prematurely) abandon costly projects and write them off as failures.

Chapter 4

Infrastructural inversion as a generative resource in digital scholarship[51]

Parenthesis – relation conceptual framework

In the preceding chapters I have argued that the characteristic tensions that arise in the conduct of digital scholarship can only be resolved through reflexive solutions, since these potentially create compatibility between novel technological affordances and the historically developed, disciplinary organization of research. There is no reason to assume, however, that such solutions will be singular. Taking seriously the notion that infrastructure develops in an evolutionary fashion, the mutual adaptation of novel technology and established user practices is more likely to resemble a process of variation and selection. Different actors will propose numerous solutions to infrastructural conflicts, but only some of them will ultimately persist and thereby become part of the infrastructure themselves.

My aim in the following chapter is to theorize and empirically investigate the systemic function of reflexivity in the historical development of the scholarly knowledge machine. To do so, I will adapt a more expansive focus than in the previous case studies. Rather than taking an empirical look at individual projects, I focus on the role of reflexivity in the discourse and practice of digital humanities as such. Bruno Latour (1993) has famously argued that the practice of science has traditionally involved a process of purification, i.e. an effort to retrospectively delete the contingency and messiness of everyday scientific work when circulating research findings. The resulting, purified image has historically served as a source of legitimacy for scientific knowledge, which can thereby claim to be distinct from the less sophisticated and biased opinions of other societal actors, for example in politics or business. Purification has also provided the foundational motivation for ethnographic work in STS – if scientists strive for presenting an idealized account of their research, then anthropologists should try to study scientific work as it unfolds and before its original messiness is deleted (Latour & Woolgar, 1979; Knorr Cetina, 1981). The

51 This chapter has been published as: Kaltenbrunner, W. (2014) Infrastructural Inversion as a Generative Resource in Digital Scholarship, *Science as Culture* 24(1), pp. 1-23.

digital humanities initially seem to subvert this logic. Rather than making invisible the uncertain and contingent practices through which they were generated, digital scholars often use their publications to highlight the very infrastructural conflicts that arise in the everyday conduct of digital project work – for example, difficulties in trying to combine collaborative tool development with a traditional academic career (Flanders, 2011; THATCampCHNM, 2011), or diverging expectations towards data in different scholarly practices (Drucker, 2009). While rarely drawing on STS literature or the sociology of science, the topics and reflexive style of scholarly publications in digital humanities thus often parallel the arguments I present in this thesis.

A critical move I will make in theorizing this phenomenon is to abandon a commonsensical dichotomy of visibility, or, transparency, on the one hand, and invisibility, on the other. Rather than positing the reflexivity in the discourse of digital humanities as the opposite of traditional purification practices in science, I will actually argue that reflexivity itself always entails selectivity in fore- and backgrounding particular elements. Digital scholars draw selective attention to instances of friction that are not routinely discussed in more traditional academic discourse. By putting up for discussion certain infrastructural conventions that pose an obstacle to digital project work, they also destabilize them. This potentially affects the historical development of the scholarly knowledge machine at large. However, insofar as there are many distinct ways of identifying and resolving infrastructural conflicts, the discursive reflexivity in digital humanities can also be analyzed as a site where very different viewpoints on the very purpose of digital scholarship clash. These conflicts are simultaneously intellectual and political, touching for example on questions of desirable research methods, competing views of the function of the humanities in society, as well as the meaning and implications of 'efficiency' in the practical conduct of everyday research. In short, they are conflicts between different ways of reimagining the scholarly infrastructure, and thus ultimately about what forms of knowledge should become possible in the future.

Introduction

Digital humanities (DH) is an emerging field whose practitioners apply digital technology to humanistic research problems. Its manifestations are diverse: from the use of digital annotation tools in the collaborative study of empirical sources (WordHoard, 2004–13), the computational analysis of large corpora of textual data (Michel et al., 2011), to the use of provocative digital performances for exploring the twists and turns of poststructuralist theory (Hansen, 2012). DH has attracted considerable public attention over the last decade. Regular readers of the *New York Times*, for example, will have come across an extensive feature report on DH (Cohen, 2010), or they may have encountered a critical interpretation of the DH in the "Opiniator" column of controversial literary scholar Fish (2012). Characteristically, DH here is referred to as "humanities 2.0", which reflects a general tendency of both observers and practitioners to discuss digital research practices in terms of their implications for how scholarship at large will be conducted in the future.

Indeed, much digital scholarship poses a challenge to defining features of academic life in the humanities. For example, the collaborative practice of most DH work is at odds with the single-author, monograph-oriented research model dominant in many disciplines. Some DH projects, moreover, aim to demonstrate that publications can take the shape not only of articles and books, but also of datasets and digital applications (Hansen, 2012). Strikingly, digital scholars often present such projects in explicitly reflexive accounts. A highly publicized monograph by Kathleen Fitzpatrick, for example, presents an experiment with a new form of online peer review in the format of a deliberately polemical, revisionist account of how the conventional model of peer review came to be seen as the distinctive feature of modern knowledge production. Similarly, the virtual platform alt-academy offers a number of very personal essays in which DH practitioners comment on the relation between academic employment modalities and the historical development of digital scholarship. Such contributions, I suggest, can be considered instances of what Bowker & Star (2000) have called infrastructural inversion, that is, a systematic defamiliarization of routinized academic work that exposes the otherwise invisible inner workings of knowledge production.

In this paper, however, I do not consider the job of the science and technology studies (STS) analyst to be done with the infrastructural inversions that digital scholars perform themselves. Instead, I make those reflexive accounts an object of study in their own right. Rather than

picturing the adoption of digital technology by scholars as a sweeping revolution, I argue that the humanities constitute a socio-material infrastructure that develops in a reflexive process. Specific technological affordances—that is, hypothetical uses of technology, such as computational analysis and digital publishing formats—are not automatically realized, but have to be implemented by actors who are themselves enmeshed in the existing configuration of infrastructure. This often means going against the grain of established routines. Reflexive representations of digital scholarly work, circulated by practitioners, can serve to highlight and problematize such routines, and thus play a role in the evolutionary development of infrastructure. However, since different forms of reflexivity likely open up different paths for development, they also constitute a potential site of controversy. All this makes them a topic relevant to STS research, and to anybody interested in the implications of digital research technology for humanities scholarship as such.

I empirically base my argument on an investigation of four case studies where practitioners of DH circulate reflexive representations of their work settings. This involves the discursive analysis of a variety of online and print sources, such as scholarly monographs, internet forums frequented by digital humanists, and new types of outlets for the publication of digital scholarly applications. My analysis aims to answer the following research questions:

(1) How do actors use infrastructural inversions to promote changes in key aspects of humanities infrastructure, such as peer review, output formats, and the organization of research projects?

(2) How do different ways of defamiliarizing the status quo of infrastructure relate to each other?

(3) What does it mean, theoretically, to think about a defamiliarization of infrastructure as a generative resource for actors?

Infrastructural Inversion as Articulation Work

The adoption of digital technologies in the humanities is related to new ways of funding and organizing research (Borgman, 2007; Gold, 2012; Wouters et al., 2013). DH predominantly takes place in collaborative, grant-funded projects, where scholars team up with professionals from very different backgrounds, such as designers, programmers, and data workers. This implies new ways of bringing research to closure. For example, appropriately funded collaborations can encompass substantial amounts of

data work, thus scaling up the potential empirical scope of research (Kaltenbrunner, 2014). Alternatively, such collaborations allow to present research in provocative digital formats that would be difficult for any single academic to realize (Drucker, 2009). DH thus poses new challenges when compared to traditional modes of scholarly work: it requires managing a collaborative overhead, dealing with different forms and unequal distributions of knowledge in the context of a single project, and worrying about follow-up funding to keep collaborative relations from dissolving.

A concept that is particularly suited to analyze these challenges is articulation work. Originally coined by grounded theory pioneer Anselm Strauss, the term designates the situated activity of meshing distributed elements of labor in cooperative work settings (Strauss, 1985, 1988; Schmidt & Bannon, 1992). As Schmidt & Simone (1996: 4) point out, "to articulate" here means "to put together by joints". Articulation work is distinct from the production tasks in particular work contexts, which are more routinized. The production task of, say, a historian would be to do archival research, and then write a monograph or paper. Articulation work includes everything that is necessary to manage that scholarly process: survey published literature, manage the contingencies of archival work in the face of resource and time constraints, and of course circulating scholarly arguments. Coordinating these tasks can, for example, involve changing one's research question, if it turns out that there are insufficient archival sources to sustain a particular scholarly claim. The need for articulation work thus arises from the necessary underspecification of situated action through formal organizational schemes (Suchman, 1996).

Articulation work is necessary both within and between different levels of work organization. Tasks must be meshed not only on the individual project level, but also between that individual level and the wider community of scholars, as well as the academic (or other) institution hosting the project (Fujimura, 1987). Articulation work between levels includes, for example, the networking among researchers at conferences, which serves to align one's research with the research interests of others. Scholars must also align their work with the interests of employers and/or funding bodies, thus requiring them to promote and justify their ongoing research, and to stay up to date on calls for funding. Given the collaborative organization of much digital scholarship and its almost exclusive reliance on temporary grants, the complexity of articulation work in DH will only tend to increase.

STS scholars have used the concept of articulation work to analyze how researchers in both the humanities and natural sciences manage

inherently uncertain, collaborative processes (Fujimura, 1987; Antonijević et al., 2013). At the same time, insofar as articulation work is constituted by the ad hoc activities that are necessary to "get the job done", it is usually taken to imply only a basic degree of reflexivity of the working individuals with respect to their infrastructural work settings (Strauss, 1988). While actors experienced in a given work setting routinely perform articulation work, thus the argument, it is only in moments in which a disruption to the regular work flow of a project occurs that this work becomes explicitly visible to both observers and actors. Fujimura (1987) suggests that actors are so much concerned with what they consider their production task that they tend to disregard the numerous elements of articulation work as important activities in their own right.

The notion that routinized articulation processes tend to become invisible to actors over time has been further developed by infrastructure studies, inaugurated by Geoff Bowker, Karen Ruhleder, and Leigh Star, herself a PhD student of Strauss (Star & Ruhleder, 1996; Bowker & Star, 2000). The term infrastructure here has a specific theoretical meaning: infrastructure is not a specific thing such as tubes and wires, but a relational state that obtains when actors working in different parts of a historically grown, cooperative work setting achieve a smooth coordination of their individual activities. Particularly important in achieving such coordination are various sorts of classifications, for example, divisions of labor between groups of professionals and standardized ways of exchanging information. These allow for alignment of heterogeneous practices without the need for a singular, hierarchical management structure. Infrastructure, we could say, is the crystallized accumulation of historical articulation work. By implication, infrastructure tends to become transparent over time, precisely because actors interact with and reproduce infrastructure through their daily routines. Bowker & Star argue that to disentangle how technological instruments, conceptual frameworks, and social order in a given infrastructure make possible specific forms of living and knowing, one must interrupt this transparency. There are two ways of achieving such an analytical opportunity: one can systematically defamiliarize particular elements of infrastructure, or one can study it in moments of breakdown. Bowker & Star (2000) call this method infrastructural inversion.

How, then, should we think about the relation between articulation work and inversion? When reading the foundational publications on articulation work and infrastructural inversion cited in the above, one could get the impression that articulation work and inversion are opposites. The

articulation work done by actors in particular work settings, one could reason, is what produces and maintains infrastructure in its relational transparency, while it is the prerogative of an objective ethnographic observer—who is not subject to the professional blindness of the implicated actors—to defamiliarize infrastructure and make articulation work visible in its *real* form and function. This would also mean that infrastructural inversion is something essentially different from the basic reflexivity that is involved in all articulation work in the first place. To be sure, Bowker & Star (2000: 310–311) also discuss cases of infrastructural anomaly, in which actors are forced to develop a particular reflexivity for survival. For example, some actors are marginalized by existing classifications schemes (think of classification by race or gender), or they are simultaneously part of multiple classification schemes that do not properly map onto each other. Maneuvering such anomalous situations requires actors to "juggle" their different memberships, and to find workarounds to infrastructural arrangements geared to exclude them. Bowker & Star propose that attending to such activities is a particularly good starting point for infrastructural inversion by STS researchers, since they require an implicit ability to defamiliarize infrastructure on the part of the studied actors. However, they abstain from defining the reflexivity of actors as a form of inversion.

In this paper, I pick up this is line of thinking, and I supplement it with a theoretical clarification. My argument is that we should resist the temptation to think about infrastructural inversion and articulation work in terms of a dichotomy. Instead, inversion should be conceptualized *as a specific form* of articulation work. Such an approach takes seriously the notion that no instance of inversion— including those performed by STS analysts—uncovers infrastructure as it really is, but always constitutes a situated effort to reconstruct infrastructure. The inverter always selectively highlights certain aspects of infrastructure, and the particular emphasis of a given inversion will often reflect local concerns and individual strategic purposes. Collapsing inversion and articulation work emphasizes that the reflexivity of actors in everyday work settings is not essentially different from the reflexivity of inverting analysts, a notion that resonates well with recent methodological writing on the epistemological status of STS ethnography (Beaulieu et al., 2007; Zuiderent-Jerak, 2007), as well as with an older line of STS research on the role of reflexivity in knowledge production (Woolgar, 1988; Ashmore, 1989). Infrastructural inversion can then be conceptualized as a generative resource actors themselves draw on in developing infrastructure. While some recent scholarship has pointed in a

similar direction (Edwards, 2010: 20–23; Mayernik et al., 2013), the exact mechanisms by which inversions unfold a generative potential have, to my knowledge, not yet been explicitly theorized.

To better understand how actors in digital scholarship wield infrastructural inversion as a generative resource, I analyze four cases of reflexive communication among practitioners of DH, in which they grant a selective look behind the scenes of their everyday work. I will argue that by circulating documents in which they systematically defamiliarize their work contexts—by performing inversions—researchers can promote new ways of meshing efforts of individuals as well as tasks and task areas. As Fujimura (1987) points out, circulating scholarly publications and other kinds of documents, such as letters or email, is a way of aligning the activities on the work level of individual projects with that of the wider community of scholars, employers, and funders. Through such alignment, individual research projects become "doable" economically and in terms of being acknowledged by scholarly peers. The documents analyzed in the following fulfil these functions, but they also carry the additional implication that they are deliberately crafted, reflexive representations of research settings. Not only does the circulation of these documents thus *constitute* a form of articulation work, but the content of the documents itself is explicitly *about* articulation work.

The function of the documents can be illustrated in relation to Fujimura's (1987) concept of packaging. This denotes the practice of compressing unorganized articulation tasks into standardized sequences of standard tasks, which are then assigned as someone's production work. Fujimura provides two examples. A first one is technological instruments that "black-box" certain tasks, which allows a wide variety of (non-specialist) users to accomplish them. Another instance of packaging is the dissemination of manuals describing standardized techniques for particular tasks in a given setting; for example, molecular cloning in oncological lab work. Such manuals enable users with some basic field knowledge to master those techniques by themselves. Packages thus reduce the effort individuals would otherwise need to spend on organizing and coordinating their work autonomously and ad hoc.

I would extend the range of packages to include other ways of reducing the complexity of articulation work in the conduct of scholarship, such as the established form of peer review and the use of conventional ways of framing scholarly arguments and empirical material. Peer review essentially is a widely accepted protocol regulating quality control and the

reproduction of intellectual standards in a field. This absolves individual researchers from having to figure out solutions to these tasks every time they wish to publish a paper. Similarly, the use of conventions of representation in circulating scholarly arguments and empirical material makes it easier for the individual researcher to establish common ground with other actors. Framing a set of spatio-temporal information in Aristotelian–Euclidean categories, for example, facilitates data-sharing with an audience who can be expected to immediately relate to these conventions. In turn, publishing an intellectual argument as a monograph with a prestigious university press makes that argument immediately recognizable as a valid scholarly contribution for peers and employers. Lastly, I would consider it an instance of packaging when a set of articulation tasks is compressed into the responsibility of a single group of actors. Star & Strauss (1999) have described this move as a "disembedding of background work", and they give the job profile of nurses, a classical service profession, as an example.

In most of the cases I analyze, practitioners of DH use inversions to highlight and "untie" existing packages in scholarly infrastructure, which inevitably means releasing the complexity of articulation work that the packages allow to contain. Simultaneously, however, these inversions promote alternative ways of streamlining that articulation work. A particular form of reflexivity may in fact establish new ways of packaging over time, thereby segmenting research practice by re-drawing the boundaries between the visible and the invisible. This again will have implications for the power relations within the academic labor ecology, and ultimately for the forms of knowledge that can be produced in a field (Kaltenbrunner, 2014). An analysis of different forms of reflexivity, therefore, is also a specific method for highlighting the perpetual conflicts through which actors try to shape their shared socio-material environment according to their individual priorities and visions (Bowker & Star, 2000). These conflicts tend to become particularly consequential in times of major infrastructural change, as afforded by the current investments in digital infrastructure and e-science (Edwards et al., 2009).

Here, it is important to take into account the wider context of articulation work. While all inversions are subversive in that they challenge some aspect of field-specific common sense, different inversions will not be equally compatible with the interests of relevant actors beyond the disciplinary domain, such as policy makers or funding bodies. The latter two often see the diffusion of computational methods as an opportunity for

"helping" the humanities to become empirically more robust (Williford & Henry, 2012) and in line with "inevitable" technological developments (Waters, 2013), as well as a solution to the perceived problem of "organizational fragmentation" of the scholarly landscape (ESFRI, 2008). Practitioners can choose to align their vision of DH with these interests, or they may mobilize existing resentment of scholars toward exactly such attempts by research policy to intervene in disciplinary self-governance. These dynamics will affect the likelihood of a given form of reflexivity to establish new, hegemonic ways of packaging articulation work. While not aiming for a comprehensive sociological analysis of DH-internal politics, I will therefore try to situate the inversions I analyze with respect to the political alliances they pursue.

Four Case Studies of Inversion in Digital Scholarship

Methodology and Case Selection

In the following, I investigate four cases in which practitioners of DH perform an infrastructural inversion as a form articulation work. This analysis is based on the discursive study of various types of sources that actors use to communicate and relate with each other: scholarly monographs, blogs, a scholarly journal that only publishes digital applications (*Vectors*), an internet forum for digital humanists employed in non-research positions (alt-academy), as well as a mediaCommons website where scholars can publish and review monograph manuscripts digitally.

I have selected those four cases because they touch on aspects of articulation work that are currently the topic of intense discussions within the community of DH, thus highlighting the nexus between infrastructural organization and disciplinary identity. The significant attention these particular arguments have attracted allow me to make representative statements about the recent development of the field.

The first case study addresses the problematic distinction of "technical" and "intellectual" labor. Digital scholarship involves a manifold of activities that would previously have been considered to lie outside the core business of a scholar, such as coding, data work, and ICT maintenance (The Virtual Knowledge Studio, 2008). The question as to what constitutes scholarly authorship in digitally mediated research, and what proper modalities of employment—two aspects with important consequences for how task areas in the scholarly process are meshed—consequently has occupied an important role in discussions among digital scholars. This is

illustrated in an exemplary fashion by Julia Flanders' contribution to the virtual platform alt-academy (case #1).

Secondly, in contrast to monograph-oriented forms of research, DH usually takes place in collaborative, grant-funded projects. Digital scholars thus face the twofold challenge of a new collaborative overhead in the primary research process, and the need to align their projects with the interests of funding bodies and potential partner institutions such as archives and libraries. However, we can observe the recent emergence of a new type of reflexive organizational literature, namely guidelines and "best practices" in which digital scholars exchange strategies for dealing with the new complexity of articulation work (case #2).

Thirdly, there is a discussion about the role of theory in digital scholarship. Theory plays a role in coordinating and integrating individual contributions to the scholarly knowledge produced by a disciplinary community (Whitley, 2000), thus making it a key element of articulation work. Prominent scholars have argued that DH have been too much concerned with creating digital artifacts such as tools and analytical applications, and not enough with underpinning these activities theoretically (Ramsay & Rockwell, 2012). As a side effect of this neglect of theory, critics such as Johanna Drucker suggest, the DH risk buying into a new form of digital positivism that is tacitly imported together with data-centric methods and tools for quantitative analysis (case #3).

Fourthly, another key theme in the discussions among practitioners of digital scholarship is peer review. The conventional form of peer review has come under critical scrutiny in many fields of research (Campbell, 2006), and DH practitioners have been particularly proactive in exploring digitally mediated alternatives for the provision of scholarly quality control. As an example, I analyse Fitzpatrick's (2011) experiment in peer-to-peer review (case #4).

Julia Flanders: Inverting Divisions of Labor

My first case study draws on an essay by digital humanist Julia Flanders (2011), well known for her work in Brown University's longstanding Women Writers Project. The essay has been published on the internet forum alt-academy. Started as a traditional book project in 2010, alt-academy has developed into a virtual meeting point for humanities-trained individuals in "para-academic" employment (university libraries, ICT service departments, DH centers…). Alt-academics usually work outside the tenure structure and have a professional concern with the development of digital research tools.

Although frequently engaged in cutting edge digital scholarship, many alt-academics feel that due to its "technical" orientation, their work is insufficiently appreciated by university administrators and peers in traditional research positions. Many prominent digital humanists in fact have started their careers in "para-academic" positions, which would imply that these personal histories have shaped the identity of DH as a field (Nyhan, 2012).

An important theme of Flanders' essay is how the deeply cooperative nature of any kind of humanities scholarship makes itself felt once researchers are stripped of their infrastructural embedding. Flanders recounts how her first employer, Brown University's Scholarly Technology Group (STG), was run as a "cost center". This meant that the university covered some expenses, while STG also was expected to attract its own funding through grants and contracts. STG moreover was required to autonomously cover various support activities, such as administration and server maintenance, that enable its actual work. Many of these usually invisible tasks thus became explicitly visible to STG employees for the first time. This experience made Flanders particularly aware of the conventional divisions of labor that are normally used to mesh and coordinate the countless tasks and task areas in the academic labor ecology.

A particularly tricky distinction is that between technical and scholarly work, as Flanders continues to argue in a reflection on her many years of experience as a consultant for various DH projects, such as digitization initiatives in libraries, or digital scholarly editions for university presses. In such consulting relationships, Flanders utilizes expertise in digital technologies as well as substantive scholarly domain knowledge. Both skill sets are required in order to tailor technical solutions to the needs of scholars, managers, and editors. Nonetheless, the consultant must demarcate her services as sufficiently different from the client's own scholarly knowledge, in order to carve out a professional niche for herself. Flanders' strategy as a consultant lies in emphasizing her technical expertise, so as to have a commodity that is interesting for the clients. The latter, in turn, are usually quite happy about not having to bother with "anything digital".

> I suggest that there are at least two forms of valuable knowledge in play. The first is the knowledge that the client values because they are glad they don't have it (or have responsibility for it): they value it in the consultant because it represents what they think they are

> buying. Technical knowledge falls into this category: (…) knowledge of XML, databases, electronic publication systems, digital project management. The second, more problematic category is the knowledge that makes the first type usable to the client—namely, the meta-knowledge through which the consultant grasps the client's subject area. In my case, this includes familiarity with scholarly editing and with methods of literary scholarship, and despite the fact that my technical knowledge would be unusable without it, this knowledge also constitutes a kind of subtle structural irritant in the consulting relationship. Precisely because of its potential value (if I were being considered as a colleague), it must be explicitly devalued here to show that I am not so considered: it creates a necessity for gestures of demarcation by which the boundaries of my role can be drawn, with technical knowledge on the inside and subject knowledge on the outside. (Flanders, 2011).

Ironically, the intervention of the digital consultant into the scholarly process, although framed as merely technical, can bring about a rather consequential reconceptualization of the research object. In the case of the digitization of the New Variorum Shakespeare edition, the consulting work by Flanders resulted in a newly privileged role of XML in the editorial process, which in turn brought to the fore certain features of the Shakespearean text that where not visible before.

> Where in the print production process the editorial manuscript was taken as the most informationally rich artifact in the ecology (…), in the digital process the editorial manuscript is a precursor to that state: the XML encoding brings information structures which are latent or implicit in the manuscript into formal visibility. (Flanders, 2011)

Flanders thus portrays a professional hierarchy between the editor, who is formally in charge of the intellectual content of the edition, and the DH consultant, who is hired for a seemingly subordinate task, namely to provide technical advice. This division of labor also means that editor and consultant relate to the scholarly infrastructure in different ways. Precisely because of her formal prestige, the editor is one degree removed from technical responsibilities. The consultant on the other hand is responsible for "anything digital". A whole task area here is packaged into the

responsibility of a single actor group, thus constituting an example of "disembedded background work" (Star & Strauss, 1999). This reduces the complexity of articulation work the editor would otherwise have to deal with, that is, acquire relevant knowledge of digital editing techniques, and manage the contingencies and organizational overhead their application entails (Fujimura, 1987). However, Flanders suggests that such packaging is problematic, since digital technology effectively alters the power relation between editor and digital consultant. An editor familiar only with traditional methods will not have a good understanding of the affordances of digital technology. Therefore, by ceding responsibility for digital techniques to the consultant, the editor also cedes potentially important design choices.

Flanders' infrastructural inversion serves to highlight and problematize how digital consultants are forced to reproduce an established way of packaging articulation work in digitization initiatives, if they mean to find paying customers. While her experiences are common among her peers, the recent policy and media interest in DH has arguably created a particular strategic opportunity for initiating a broader discussion about academic employment modalities and the valuation of specialist labor. Flanders implies that for digital research technology to unfold its full potential, it must go along with a change in the academic labor hierarchy that currently prevents digital skills from wider diffusion. Her inversion thus rhetorically ties the success of digital scholarship—currently high on the agenda of many policy makers and funding bodies (Williford and Henry, 2012)—to the
need for upgrading the professional status of alt-academics.

Tom Scheinfeldt: Infrastructural Inversion as a Management Technique

As pointed out, digital scholarship is usually organized in collaborative, grant-funded projects that involve professionals with different specializations and accountabilities. This requires practitioners to constantly look out for supplementing and follow-up grants, so as to sustain often fragile, inter-institutional collaborations. Not least, digital research projects frequently depend on other academic actors such as archives, libraries, and heritage institutions for access to digitized source materials and metadata. Collaborative digital scholarship thus presents its practitioners with a bewildering complexity of articulation work, not only on the level of individual projects, but also between the project level and external actors such as funding bodies and potential partner institutions. Degree programs

in the humanities teach students to perform sorts of articulation work required in established formats of scholarship: students learn how to do archival and field work, to survey literature, to produce a clearly defined form of output, and to present at scholarly conferences. This education is not geared toward disseminating skills for dealing with the overhead of articulation work in DH.

However, there is a growing amount of instructional materials on how to go about DH project work available online. An example is a guideline collaboratively assembled by the participants of *one week/one tool 2011*, a workshop on digital project planning and management taught by the director of Georgetown University's Center for History and New Media, Tom Scheinfeldt. The assembled notes compress Scheinfeldt's key lessons, and they have since been circulating among digital scholars as one of the go-to sources for DH project management knowledge (French, 2013).

The function of this guideline is similar to Fujimura's (1987) example of packaging through the dissemination of manuals. Specifically, she refers to a manual describing techniques of molecular cloning, a document so widely used in the field of cancer research that it became known as "the bible". But while this latter document was meant to reduce the time and effort required to teach individuals lower-level skills in well-defined task areas, Scheinfeldt's guideline promotes the establishment of a new class of professionals, the DH project manager. The responsibility of the project manager is to completely absolve the other participants from articulation work that exceeds what is necessary to fulfill their specific production task within a project.

> Project Manager's job is to protect the staff from the PM's job. They shouldn't have to interface with admin, deans, budgets, etc. Not because there's secrecy involved, but because staff should be able to do what they do best: their work. (THATCampCHNM, 2011)

According to Scheinfeldt, the distinctive feature of such managers must be the habit of systematic reflexivity with respect to the work processes they coordinate, and with respect to the wider infrastructural context in which those processes are embedded. This sort of managerial reflexivity cannot be clearly distinguished from infrastructural inversion. In fact, the abstract that precedes the guideline is rather reminiscent of Fujimura's (1987) own analysis, insofar as it highlights the importance of articulating tasks and task areas, and of aligning those articulation efforts between different levels of

work organization.

> This session will consider both the practical, day-to-day work and intangible aspects of managing digital projects in the humanities. Pragmatic lessons will include picking a project, building partnerships and engaging stakeholders, attracting funding, budgeting and staffing, setting milestones and meeting deliverables, managing staff, publicity and marketing, user support, sustainability (…). The session will also consider several intangible, but no less important, aspects of project management, including communication, decision making, and leadership. (THATCampCHNM, 2011)

A crucial aspect of several topics discussed in the manual is the need to "read" the preferences and accountabilities of project-external actors to make a project "doable". This includes, for example, the volatile interests of funding bodies.

> a. Pick something that is interesting to you, but that's not enough of a reason to pick a project. There are other questions you need to ask yourself.
> b. It needs to be something that is fundable. (…)
> c. If it's not fundable, is there a way to modify it so that it is fundable? What adjustments can you make to your grand vision? You need to be flexible (like the willow). (THATCampCHNM, 2011)

Another group of important external actors consists of libraries, archives, and heritage institutions, who often can provide access to digitized collections. At the same time, these institutions are themselves in a process of adapting their function in light of new technological possibilities for storage and dissemination. A key advice by Scheinfeldt is to be aware of the possible new needs of such institutions, which might allow for an individual digital project to engage in a form of trading with much larger partners.

> d. Partnerships are a way to build up your data.
> e. Shoot big in your external partnerships. Do not go thinking, "Library of Congress won't partner with me." You might be part of a more nimble organization than they are. (…) There are some things that they cannot manage to get done on their own. Just because you

> are small does not mean that you do not have something to offer that could be valuable to them. (THATCampCHNM, 2011)

Not least, Scheinfeldt provides some hard-and-fast advice for how to coordinate work on the level of the individual project. His recommendations reflect a strong sense of pragmatism:

> f. Leadership is momentum making. Make sure everyone is always moving forward. If they are not moving forward, you are not leading. (…)
> g. Leaders are first doers. Best collaborations are not about shared decision making, it is about shared doing. (THATCampCHNM, 2011)

Scheinfeldt's guideline promotes an emphatically pragmatic way of dealing with the new complexity of articulation work in digital scholarship. It encourages prospective managers to defamiliarize academic infrastructure in such a way as to see their projects as actors in a larger ecology, which in itself is undergoing change. This entails acknowledging and adapting to the power of funding bodies, and reflecting on individual projects in terms of their instrumental relations with other actors, such as libraries and heritage institutions.

However, the pragmatic managerial style in which Scheinfeldt defamiliarizes the conduct of collaborative digital research, and the readiness by which he accepts key changes in the academic organization, such as the new importance of grants, alienates many traditionally trained scholars. A panel at the Modern Language Association's (MLA) 2013 conference, for example, was entitled *The Dark Side of the Digital Humanities*. It featured a number of renowned scholars who argued that there is a problematic tendency in DH to frame the discussion on the future of research and higher education in an uncritical entrepreneurial discourse. Flush with money from short-term digital project grants, thus the speakers, the DH tends to be complicit with neo-liberal approaches to university governance, thereby sacrificing the critical edge that has characterized much late 20th century scholarship (e.g. Grusin, 2013).

Johanna Drucker/Mark Hansen: Inverting Representational Conventions

There are also instances of digital scholarship, however, that explicitly position themselves in opposition to the entrepreneurial pragmatism of

digital humanists such as Scheinfeldt. An example is the influential work of Johanna Drucker, one of the primary representatives of critical theory in DH. In her book *SpecLab*, Drucker narrates the history of several projects in what she calls speculative computing, undertaken at the University of Virginia in the early 2000s. The label "speculative computing" is meant as a challenge to DH, which Drucker argues is oblivious of the crucial theoretical legacy of poststructuralism and deconstruction. Specifically, she criticizes the notion that the practical constraints of digital scholarship also require epistemological pragmatism.

> Time after time, we saw theoretical understandings subordinated to the practical "requirements of computational protocols". As one of my digital humanities colleagues used to remark, we would go into the technical discussions as deconstructed relativists and come out as empirically oriented pragmatists. (Drucker, 2009: xiv)

A key concern of Drucker is to criticize the conception of data that the humanities in her view tend to import from the natural sciences as they adopt digital technologies for visualization (cf. Drucker, 2011). Drucker argues that such approaches often have the structuralist tendency to treat data as self-identical signifiers. This new form of digital positivism, she suggests, is in fact an ideology that strives to align situated meaning-making in a functionalist way—in Fujimura's (1987) terminology, a specific strategy for packaging articulation work. Her argument moreover recalls Bowker & Star's (2000) well-known infrastructural inversion of representational categories and classification systems: the use of Euclidian geometry and of the Aristotelian definition of time as chronology, Drucker suggests, allows to establish common ground between data-sharing individuals who can be expected to immediately relate to such conventions. This reduces "friction" in the process of exchanging data (Edwards et al., 2011), but at the cost of a creeping reification of those representational categories over time.

Instead, Drucker associates SpecLab philosophically with surrealist "pataphysics", a parody of nineteenth century positivism, and she approaches the challenge of complex articulation work through the lens of poststructuralist literary theory. In the poststructuralist perspective, reading of a sign necessarily entails a creative distortion. Drucker suggests that this distortion should be celebrated, rather than framing it as a problem that hampers distributed collaboration. As a concrete example of this approach, Drucker introduces the project Temporal Modeling, in which she and her

collaborators developed a graphical language to express subjective perceptions of time (Drucker, 2009: 37–64). Standard software for data visualization, she argues, usually comes with certain Aristotelian–Euclidean conventions of representation built into it. The SpecLab team instead started out with conceptualizing ways to represent subjective perceptions of time *before* developing a data structure. Eventually, the team came up with features such as a now-slider, timelines warped by anticipation or anxiety, and special markers to denote emotional inflection of time. While not all of these could be implemented, it is characteristic of the spirit of SpecLab to experiment with new approaches to classifying data, rather than adopting existing ones from information or computer science. The idea here is to emphasize theoretical complexity and open-endedness of research problems. The resulting application intentionally resists easy appropriation by prospective users through shared assumptions about data, instead relegating complexity back to the audience. Drucker recounts the often baffled reactions to SpecLab projects:

> The spirit of play with which we imagined these projects is an essential generative insight. Around conference tables or in public presentations, our projects often provoked the query "Are they serious?" (Drucker, 2009: xix)

This is not to say that Drucker may not in practice apply articulation strategies comparable to those of Scheinfeldt, but she does not foreground any of that in her public presentation of the project. What she does explicitly foreground is poststructuralism and deconstruction as a principle of aligning her work with the work of other scholars. While Temporal Modeling emphatically ignores expectations toward positivist conceptions of data, and thus excludes potential calls for funding that presuppose such an expectation, it clearly seeks to establish common ground with colleagues who share familiarity with this theoretical framework.

Temporal Modeling is an early example of a sort of inversion/articulation in DH that has been further facilitated by the possibility to publish non-discursive digital output. In several ways comparable to Drucker's work is Hansen's (2012) digital application *Shi Jian: time*. The project was published in the journal *Vectors*, an experimental, peer-reviewed scholarly outlet that only accepts digital output. *Shi Jian* is based on the 1,200 photos and 103 videos created by Hansen during a writing sabbatical in Beijing. While the material on display thus is the stuff from

which typical touristic appropriations of visited sights are made, the presentation is geared to undermine such a reception. The application offers an interface through which users can sort the audiovisual material according to a number of different criteria, such as place, time, color shades, and point of view from which a photo was taken. The interface principally allows to explore the collection according to a linear chronological and spatial order, but the multiple alternative ways of displaying the material, which moreover can be overlaid onto each other, encourage users to acknowledge that this is just one out of a spectrum of possibilities. In a discursive introduction, Hansen declares that his goal has been to encourage "experimentation with the 'reference frame' of time", which he hopes "will open up an important conceptual and aesthetic space around questions of how we in the West live time" (Hansen, 2012). Similarly to Drucker, Hansen's project means to "untie" the packaging of articulation work through the use of Aristotelian conventions of representing time, and instead emphasizes how digital technology can be used to multiply ways of framing data conceptually. Again, much like Drucker, this does not mean that there is no attempt to preemptively reduce the complexity of articulation work in the presentation of the project. Hansen manifestly means to establish common ground with his audience through shared understanding of the theoretical framework of poststructuralism and deconstruction, which is referenced in the discursive introduction.

The particular reflexivity advanced by Drucker and Hansen thus creates a contrast to Scheinfeldt's managerial concern with adaptation and efficiency. Combining established frameworks of theoretical critique with digital methods, their inversions outline a vision of digital scholarship more likely to appeal to the apparently numerous humanists who are suspicious of the affinity between DH and the "projectification" of academic life.

Kathleen Fitzpatrick: Inverting Traditional Peer Review

While most digital scholarship takes the shape of collaborative projects, there are also instances that remain closer to traditional formats. An example is the monograph *Planned Obsolescence* by Kathleen Fitzpatrick. Currently the director of Scholarly Communication at the MLA, Fitzpatrick is in an important strategic position for promoting new ways of disseminating scholarly knowledge.

At the heart of Fitzpatrick's argument is an inversion that problematizes traditional peer review and scholarly publishing models as elements of the academic infrastructure. Fitzpatrick (2011: 13-14) argues that

print monographs, as the primary form of scholarly output, are no longer sustainable in light of diminishing budgets even at major academic publishers. To illustrate this point, Fitzpatrick relates how the University of California libraries have switched to purchasing only a single copy of new scholarly monographs, which is then sent around via interlibrary loan. This poses a particular problem for younger scholars, since publishing the works of junior academics is particularly risky for university presses. Moreover, Fitzpatrick argues that the traditional monograph model fails to acknowledge a fact of recent media history, namely a shift in the relation between information and its users from a filter-then-publish approach, characteristic of book production (where editors select manuscripts for publication), to a publish-then-filter approach, characteristic of the Internet (where content is published and only later selected as deserving particular attention). This change, she argues, has caused a shift in the generally accepted definition of epistemic authority, which no longer is with institutionally appointed gatekeepers, but an emergent property of user crowds who sift through large amounts of information. By clinging to the established system of peer review, the humanities allegedly "risk becoming increasingly irrelevant to contemporary culture's dominant ways of knowing" (Fitzpatrick, 2011: 17). Fitzpatrick combines her argument with a revisionist account of the historical foundation of scholarly peer review in the eighteenth century. Citing the historian Biagioli (2002), she emphasizes that modern peer review has its roots in state censorship and in the interest of the Royal Society to protect the privileges of its members. Only later on, peer review was rationalized as the unique quality control mechanism in scholarly knowledge production. The academic elite defending the status quo, Fitzpatrick polemically suggests, may in fact primarily be motivated by the impulse to protect its current position of power. She goes on to discuss evidence of manifest failure of peer review, for example, the influential study by Zuckerman & Merton (1971), in which the institutional affiliation of authors was shown to influence the likelihood that journal editors would accept their submissions.

Peer review and the traditional print monograph, the targets of Fitzpatrick's inversion, can themselves be seen as institutionalized instances of what Fujimura (1987) calls packages. Both serve to reduce the complexity of specific aspects of scholarly articulation work: the former regulates the mechanism of scholarly quality control, and the latter stipulates what scholarly expression must look like in order to be immediately recognizable as a valid contribution to disciplinary knowledge by colleagues and tenure

committees.

Yet Fitzpatrick offers an alternative for containing the complexity of articulation work that is released if conventional forms of quality control and publishing are discarded. *Planned Obsolescence* in fact is an experiment in digital publishing, meant to demonstrate the potential of what she calls peer-to-peer review. Parallel to the formal, anonymous peer review provided by NYU Press, a draft of the book was published in chapter-long postings to a website hosted by mediaCommons, with the possibility for anybody to create an account and comment on the text. In an introductory note to the digital draft, Fitzpatrick (2009) explains that this online conversation between herself and the peers will be "key to [her] revision process". Digital technology was instrumental to this form of quality control in two ways: the open peer review of *Planned Obsolescence* was heavily advertised on Twitter, thus making up for the formal protocol that normally orders the communication between scholars, editors and referees. Moreover, the mediaCommons website has a graphical user interface allowing readers to post publicly visible comments directly next to the text.

However, "untying" an established package of articulation work, and trying to create another one, did not go without some friction. A first shortcoming Fitzpatrick (2011: 191) notes herself in a reflection on the online review experiment is the unequally distributed attention of the referees. Some passages of the online draft received a lot of comments, while others were largely ignored. Moreover, a number of comments primarily consist in enthusiastic applause for a well-put insight, or they digress into longer exchanges among commenters. This communicative register is at odds with Fitzpatrick's explicit request for review-type feedback that can provide the basis for revising the manuscript. Not least, some comments are visibly informed by a certain sense of confusion with respect to the status of the online draft. In the following quote, a commenter preemptively limits the scope of her proposed revisions, since she is unsure about the extent to which Fitzpatrick can even change the draft. The commenter apparently thinks of the online draft as a more or less finished product, rather than a trace of an unfolding writing process.

> I'd like to (. . .) suggest moving what seems to me the key conclusion out of footnote 1.8 and into the body of the text. (I'm assuming, Kathleen, that you can make changes before this goes to hard covers?). (Rowe, 2009)

The reason for such confusion, I suggest, is that peer-to-peer review is still in the process of being defined. Establishing a new protocol for regulating scholarly quality control requires an effort at standardizing chunks of articulation work, a process that will be cumbersome and potentially contested. It is far from obvious that all scholars who declare themselves as advocates of peer-to-peer review share a consensus on its proper modalities. Koh (2013), for example, has recently attacked the editors of the *Journal of Digital Humanities*, who claim to adhere to the goals of peer-to-review. The journal collates and formally publishes existing DH work (applications, tools, discursive arguments) that was previously accessible only online. Koh argues, however, that this specific interpretation of publish-then-filter creates new opportunities for gate-keeping *within* digital scholarship, since it is again the journal editors who fulfil the function of filtering pre-existing "content".

Conclusion

I have introduced my argument by theorizing the relation between the concepts articulation work (Strauss, 1985, 1988) and infrastructural inversion (Bowker & Star, 2000). Articulation work denotes the activities necessary to manage the contingencies that occur in the everyday practice of scholarship. Infrastructural inversion is an analytical shift in perspective that foregrounds the normally taken-for-granted elements that invisibly enable distributed cooperative work. At first sight, the concepts might seem to constitute a dichotomy: articulation work after all is what sustains everyday work routines in their transparent infrastructural-ness, while inversion is a defamiliarizing move performed to interrupt this transparency. I have suggested that it is more useful to think of inversion as a specific modality of articulation work. The particular reflexive perspective that is such a characteristic feature of much writing by digital humanists can then be theorized as a generative resource. Inversions performed by DH practitioners defamiliarize scholarly infrastructure in such a way as to highlight and critique existing traditions of organizing articulation work, while simultaneously promoting alternatives for how to handle that work.

I have provided four cases studies to illustrate this argument empirically. In three of these, actors use inversion to problematize existing ways of streamlining articulation work by sequencing it into standardized packages (Fujimura, 1987). Fitzpatrick (2011) "unties" the package of peer review, which is essentially an established protocol that regulates the communication between scholars/referees. She polemically depicts that

protocol as a suboptimal historical accident, while offering digitally mediated "peer-to-peer review" as an alternative means to streamline the articulation work required for scholarly quality control.

Flanders' (2011) contribution to alt-academy systematically defamiliarizes the division of labor between traditionally trained scholars and digital humanists. Her inversion suggests that this division of labor is not only artificial, but also increasingly problematic: the wholesale packaging and delegation of "anything digital" to DH consultants may be a convenient way of reducing the complexity of articulation work for traditionally trained scholars, but it also means delegating intellectually significant design choices.

The inversions of Drucker (2009) and Hansen (2012) draw attention to how conventional ways of framing empirical material reduce complexity through shared expectations toward data, and they ask whether such reduction is desirable in the first place. Instead they propose theory as a primary interface for aligning individual contributions to scholarly knowledge, and their digital applications intentionally relegate larger bits of that complexity to the audience.

In the case of Scheinfeldt, inversion is a means of tackling aspects of articulation work that have no precedent in more traditional formats of scholarship, namely the organizational challenges presented by grant-funded, collaborative digital projects (THATCampCHNM, 2011). In his view, this requires the creation of a new job profile: that of the DH project manager. A distinctive feature of this new class of professionals according to Scheinfeldt's guidelines is a pragmatic reflexivity with respect to the embedding of DHprojects in a changing academic work ecology.

Through performing and circulating inversions, actors reinterpret the status of quo of infrastructure in light of potentialities, thus paving the way for embedding new tools in particular ways. Yet individual forms of reflexivity express different and sometimes competing visions of digital scholarship. The most obvious faultline in my sample is that between Scheinfeldt and Drucker. While the former's inversion is informed by a strong sense of managerial pragmatism, that of the latter explicitly opposes such pragmatism as an ordering principle. Moreover, while inversions are instrumental in highlighting concrete opportunities for altering specific infrastructural features, the process of containing the complexity of articulation work that is released when existing packages are "untied" prompts substantial, and potentially competing, efforts at creating new standards. Fitzpatrick's experiment, for example, represents not so much a

showcase of a ready-made new form of digital peer review, but rather an ongoing process in which scholars renegotiate the modalities of academic quality control.

Such instances of competition and negotiation draw attention to an aspect that I have only hinted at in this paper, namely the fact that not all inversions are equally opportune politically. Different forms of reflexivity imply distinct possibilities for mobilizing the support of other actors and developments in their environment, thus affecting their chance to establish new hegemonic ways of organizing articulation work. Scheinfeldt's guidelines are geared to reinvent the organization of scholarship in a way that makes it more compatible with broader changes in the organization and governance of academic research, such as the increasing importance of funding bodies and other partners. However, many scholars feel alienated by this prospect, since they perceive it as undermining the critical ambitions of humanistic inquiry. The inversions of Drucker and Hansen accommodate exactly these sensibilities. For them, digital scholarship is not simply a matter of "improving" research by integrating new tools, but also an opportunity for raising questions about the political and epistemological implications of seemingly neutral values such as organizational efficiency.

More research would be desirable to investigate how such alignments affect the restructuring of scholarly knowledge production over time. For example, will the dissemination of guidelines for digital project management indeed promote wider diffusion of digital practices across the humanities? Or will the often polemical call to adapt to new organizational and technological circumstances rather prompt the resistance of traditionally trained humanists, thus keeping digital scholarship a specialist endeavor? Alternatively, will we witness a fragmentation of digital scholarship into distinct theoretical and methodological approaches, a trend that is perhaps foreshadowed by Drucker's critique of "mainstream" DH?

Postscript to chapter 4

In the preceding case studies of the COST Action and Elite Network Shifts project, I have given empirical examples how reflexivity is critical to overcoming instances of infrastructural conflicts. In the case of the COST Action, applied reflexivity has meant experimenting with the size of collaborative project formats. In a field characterized by strong diversity of research questions and intellectual approaches, smaller projects make it easier to identify packageable sequences of work. In the Elite Network Shifts project, it has meant acknowledgment of the different role data play across disciplinary cultures, thus opening up new possibilities of organizing the division of labor between computer scientists and scholars of Indonesian studies. However, to counter a simplistic idea of reflexivity as a panacea that can provide universally accepted solutions, I have in this chapter attempted to study reflexivity itself as an element in the evolutionary development of infrastructure.

Conceptually, I view disciplinary conventions of knowledge production as the result of historical packaging of research work (Fujimura, 1987, 1992; Law, 2004). Disciplinary criteria for what counts as valid questions, methods, and forms of output thus can be seen as the result of particular practices that have been repeated over generations, thereby crystallizing into a scholarly infrastructure (Bowker & Star, 2000; Edwards, 2010; Star & Ruhleder, 1996). Following such conventions makes research more easily feasible economically, in that it allows to draw on the work of one's predecessors without having to reinvent organizational modalities every time anew. At the same time, established conventions make it challenging to engage in radically different research practices. Difficulty in embedding novel tools in the humanities can thus actually be seen as an incongruence between the affordances of digital scholarship on the one hand, and established disciplinary models of organizing scholarly work on the other. In this chapter I have argued that reflexivity in the discourse of digital humanists serves as a means of 'untying' the standardized packages that constitute disciplinary cultures. This allows to reimagine established forms of knowledge production, for example by showing how it is possible to create novel forms of output and conduct collaborative project work in an infrastructure geared towards single-author, monograph-oriented research. Reflexivity thus is not merely the act of creating representations. Instead, this reflexivity also does something, insofar as it creates legitimacy for nascent modes of scholarship as well as disseminating practical knowledge of how to realize them in spite of an inert disciplinary organization. Such

proposed solutions may gradually turn into new conventions in their own right, thereby reshaping the organizational conditions of digital knowledge production for future generations of academics. It is in this sense no exaggeration to say that reflexivity fulfills an evolutionary function in the development of scholarly infrastructure.

At the same time, my focus on the contentious nature of reflexivity has shown that the selection of novel formats of scholarship cannot be thought of as an inevitable survival of the fittest. Different actors in digital humanities propose very different ways of framing and resolving infrastructural friction. Establishing new conventions is thus a matter of active knowledge politics by human actors. These findings also supplement my argument about the instrumental value of reflexivity, as presented in chapters 2 and 3. Each way of framing and tackling infrastructural conflicts - also the ones suggested in this thesis - may become a blueprint for the common practice of digital scholarship in the future. Precisely for this reason, we are well advised to remember the perspectivity of our reflexive thinking, and thus the fact that any solution we now choose may marginalize a host of alternative ways of reengineering the scholarly knowledge machine.

Chapter 5

Digital infrastructure in the humanities: reconfiguring the organization of scholarly tool development

Parenthesis – relation to conceptual framework

Research policy, an important factor in the organization of modern academic knowledge production, has taken something of a backseat in much of the preceding analysis. Except for chapter 1, where I focused on how different actor groups in policy, administration, and academia reimagine Dutch studies as a digitally mediated field, I have concentrated largely on the immediate hands-on interaction of humans with technology. Such an analytical focus is in fact characteristic of most scholarship in infrastructure studies and adjacent fields.[52] Given the strong interest in digital scholarship on the part of policy makers, however, my analysis would be incomplete without a more systematic investigation of how scholarly practices 'on the ground' are informed by strategic considerations and funding arrangements. A suitable empirical entry point for such an investigation are the various ongoing, European and US efforts to build a digital infrastructure for research.

The concept of digital infrastructure adds an important twist to my discussion of specific digital tools in the preceding chapters. The term frames the instruments required to conduct research not as individual artifacts that can be developed and administrated by local actors, but rather as part of a more pervasive organizational layer to support whole fields of research. In many influential policy reports, investment in digital infrastructure is in fact framed as critical for the further course of national economic and scientific development (Atkins et al., 2001; Hey & Trefethen, 2004; ESFRI, 2006, 2008, 2010; ACLS, 2006). However, this also means that a variety of actor groups are implicated as stakeholders – scientific or scholarly users, policy makers, funding bodies, and the wider public that funds research through tax money. When discussed in terms of a digital infrastructure, research tools thus become a boundary object between very

52 Inter alia, this analytical focus is due to a narrowly conceived focus on research as practice, as criticized in the above. For a more extensive critique and reflection on the underlying reasons see Pollock & Williams (2010).

different contexts (Star & Griesemer, 1989), with the priorities of researchers being only one factor among others that impact choices about instrumentation. The question then arises as to how different expectations towards technology will shape the organizational conditions under which individual tools are developed, and ultimately how they will affect the methodological and intellectual organization of the disciplines that the digital infrastructure is meant to support.

The following chapter complements the preceding one in that it attempts to move beyond the scope of individual projects, and instead adopts a systemic perspective on the mutual shaping of scholarly practice and new technologies. In the process of creating digital infrastructure, a newly developed layer of tools is being superimposed on a historically grown landscape of practices and preexisting infrastructural facilities. However, given the many European and North American specificities in academic organization and policy practices, there is good reason to assume that the development of digital infrastructure will be informed by equally specific considerations. Particular approaches to infrastructure development can for example go along with different funding and administration modalities for individual digital projects – development may be coordinated in a relatively centralized way, or rather be based on a more distributed model. This will also have an effect on how scholars in the 'traditional' humanities come to relate to new technology. Development activities may explicitly aim to cover a large bandwidth of fields with perhaps little existing need for digital technology, or rather be geared to support the work of seasoned practitioners of digital scholarship.

A comparative analysis of respective research policies will allow me to draw out the implications of particular concepts of infrastructure for the further organizational and methodological development of digital scholarship. At the same time, by showing how infrastructure is framed differently across countries and regions, the comparison will highlight the cultural situatedness of this seemingly neutral concept, as well as illustrate the long-term strategic choices that are bound up with the incorporation of digital tools into the scholarly knowledge machine.

Introduction

A remarkable development in science and scholarship in the recent decade has been the concerted attempts to create digital infrastructure – or cyberinfrastructure, as it is commonly referred to in the US - for all fields of research. In an influential NSF report published in 2003, cyberinfrastructure is defined as large-scale facilities for the storage, sharing and algorithmic analysis of massive digital datasets. The authors of the report, computer scientist Dan Atkins and his colleagues, vividly argue that creating such facilities will be indispensable if the US science system is to retain its global leadership against the rising BRIC countries in the 21st century. "[I]f infrastructure is required for an industrial economy", Atkins et al. (2003: 5) suggest, "then we could say that cyberinfrastructure is required for a knowledge economy." Since 2007, the European Commission (EC) has spent approximately €1700 million on the development of such technology (EC, 2008), and the NSF (2013) has estimated a respective expenditure of $221 million for the fiscal year 2014 alone.

While the bulk of these investments has been directed to the natural sciences and engineering, there are high-profile undertakings also in the humanities (Anderson, Blanke & Dunn, 2010; ACLS, 2006). Access to an integrated layer of digital instruments, for example for text mining or the algorithmic analysis of large amounts of visual material, is often promised to revolutionize the hermeneutic traditions that characterize many scholarly disciplines (Michel et al., 2011). In this paper, however, I am not primarily interested in how new technologies may change research practices and epistemic frameworks, but rather in their political implications. In traditional organizational formats of science and scholarship, control over the development of research tools was closely tied to the reputation economy within particular fields. Current infrastructure initiatives, by contrast, exemplify a key argument of recent literature on the reconfiguration of national science systems (Nowotny, Scott & Gibbons, 2001; Etzkowitz & Leydesdorff, 2000; Whitley, Gläser & Engwall, 2010), insofar as they entail a profound shift in the relations between researchers, funding bodies, and policy makers. Here, the creation of distributed instruments is often explicitly encouraged by policy makers, with researchers trying to steer the rerouted funding streams in ways that suit their own intellectual and professional ambitions. Analysts will be well advised, however, to avoid picturing these developments as a uniform transition to a singular new way of developing research instrumentation (cf. Hessels & van Lente, 2008; Mirowski & Sent, 2008). Creating digital

infrastructure entails strategic choices that can serve very different purposes, for example the attempt to tackle perceived inefficiencies in the epistemic and social organization of research, or rather to consolidate existing institutional formats. To better understand the implications of such initiatives for the organization of digitally mediated scholarship, it will be important to pay close attention to the specific empirical conditions of infrastructure development in different countries and regions.

As a first step in this direction, I will in this paper present a comparative analysis of current European and US approaches to developing digital infrastructure for the humanities. My analytical goal is twofold. I will firstly show how different groups of actors, such as policy makers, science administrators, and various groups of researchers, compete in establishing a dominant discursive framing of digital infrastructure in their respective national or regional context. Secondly, I will highlight how this framing, once it is instantiated in particular funding and administrative frameworks, mediates the distributed development of digital scholarly tools.

Infrastructure as a discursive interface between policy and scholarship

Over the past fifteen or so years, digital infrastructure initiatives in the US, Europe and Asia have mushroomed (see Jankowski (2009) for an overview). These have provided a rich object of study for social scientists of various disciplines, leading to the publication of numerous edited volumes (Jankowski, 2009; Olson, Zimmerman & Bos, 2008; Dutton & Jeffreys, 2010; Wouters et al., 2013), special journal issues (Jankowski, 2007; Edwards et al., 2009; Ribes & Lee, 2010), and monographs (Borgman, 2007; Hine, 2008). Some of this research addresses digital infrastructure in the humanities. Here, analysts have been particularly interested in the implications of data-intensive analytical instruments for epistemic cultures dominated by hermeneutic approaches (Fry & Talja, 2007), and many studies underline the need for acknowledging such field specificities in the design of technology (Wouters & Beaulieu, 2006; Barjak et al., 2009; de la Flor et al., 2010; Kaltenbrunner, 2014). A growing number of STS researchers moreover is adopting the highly influential framework of infrastructure studies, where infrastructure is conceptualized not as a specific thing, but as a delicate ecology of interrelated socio-technical practices of different user groups (Edwards et al., 2007; Ribes & Lee, 2010). These lines of research have in common that they tend to focus on the micro-level of scholars interacting with digital research tools. Usually, they adopt a constructivist perspective

in the sense of stressing the mutual shaping of infrastructure technology and research practices. Another, smaller strand of research has explicitly set out to produce policy-relevant insight (Barjak et al., 2013; Bos et al., 2007; Meijer, Molas-Gallart & Mattsson, 2012; Voss et al., 2007), for example on questions of appropriate governance of digital infrastructure. The common approach of these studies is to conduct systematic empirical comparisons of ongoing infrastructure projects in order to elicit 'best practices', but usually without problematizing the concept of digital infrastructure itself, e.g. by asking how and why digital infrastructure is conceptualized differently in different countries.

In this paper, I try to combine the constructivist sensibility of the former strand of literature with the policy interest of the latter. Current initiatives to create large-scale digital infrastructure can in fact be seen as concerted interventions into the organization and conduct of humanities scholarship (Beaulieu & Wouters, 2009; Barjak et al., 2013). Traditionally, the dynamics of scholarly and scientific fields used to be determined primarily by disciplinary mechanisms, i.e. peer review, the circulation of influential publications, and the accumulation of intellectual reputation among disciplinary colleagues (Whitley, 2000). The logic here was that researchers who publish papers deemed worthwhile by colleagues got more and more influence and became professors, thus establishing themselves as authorities who control access to academic employment. This also meant that decisions about what type of research instruments (e.g. laboratory and sensor equipment in the natural sciences; bibliographies, lemmatized scholarly editions, textual corpora and archival collections in the humanities) were to be built were usually decided on a national basis, and often reserved for scholars who had accumulated significant renown in their fields. To be sure, such disciplinary self-governance is not synonymous with financial independence – scientists and scholars always had to lobby when in need of larger sums of grant money, and individual national science systems have historically provided somewhat different conditions for the interaction between researchers and funding bodies (Mirowski & Sent, 2008; Whitley, 2010). It was generally uncommon, however, that policy makers and funding bodies would themselves explicitly encourage the development of particular facilities, as in the context of current infrastructure frameworks. This means that the connection between the disciplinary reputation economy and decisions about investment in instruments is reconfigured.[53]

53 For a similar argument, albeit developed into a very different direction, Barjak et al. (2013).

Rather than following the judgment of existing disciplinary elites as a default, choices about tool development are increasingly taken by new, often international coalitions of policy makers, funders, and different groups of researchers. These coalitions are held together by heterogeneous interests. Policy makers may for example hope that the performance and efficiency of research can be improved through providing tools that facilitate collaboration, data sharing, and sophisticated computational analysis. Researchers in turn may put a similar hope in the intellectual benefits of digital instruments, but will also be motivated by new career and funding opportunities. A group of actors that will play a particularly important role in this process are those who have an established expertise in the development and use of digital research technology, for example computational linguists and scholars of humanities computing. The resulting alliances are usually characterized by the pursuit of longer-term, strategic visions of infrastructure (Anderson, Blanke & Dunn, 2010; ACLS, 2006), thus creating distinct economic and political conditions for local tool development, as well as affecting the methodological organization of digitally mediated scholarship in the longer run. It is therefore not sufficient to analyze merely the dynamics of mutual shaping of technology and research on the level of individual scholarly practices. A complementary analysis is necessary to highlight what strategic considerations underlie current infrastructure initiatives, and how they affect the organization of tool development in specific fields.

The formation of infrastructure policy can usefully be thought of in terms of what Hajer (1993) calls a discourse coalition. Such coalitions form around the narrative framing of societal issues, which at the same time conceptualizes possible remedies as well as distributing responsibilities for action. Current debates on digital infrastructure in both Europe and the US typically present the latter as the logical response to the 'advent' of sophisticated information and communication technologies, i.e. as a historical opportunity to lay the basis for future scientific, economic, and cultural success of a country or region. Thinking and talking about research tools as part of such a critical investment indeed transforms their development from a strictly intra-disciplinary matter into an issue that also concerns research policy, funders, and non-expert audiences. Moreover, it suggests that individual research tools should be seen as part of a larger system that needs a comprehensive approach to organization and administration. As Edwards et al. (2013) rightfully argue, creating new infrastructural facilities is actually a process of overlaying an existing,

historically grown ecology of practices with a new set of protocols and accountabilities, thus entailing a partial redistribution of authority, influence, and resources. Individual actors therefore have a strong incentive to promote a discursive construction of the new infrastructure that suits their respective priorities, yet without alienating other stakeholders. Hajer (1993) cautions that such constructions do not exist in a vacuum, but instead draw on familiar tropes and conceptual resources that are imbued with a certain symbolic capital, and that meaningfully relate to how a particular problem has been dealt with in the past (see also Atkins, Held & Jeffares, 2011). In the subsequent empirical analysis, I will argue that the protagonists of current debates on digital scholarship heavily draw on two influential views of infrastructure, thereby trying to steer the shared strategic outlook in a particular direction. The first, more longstanding one is that of infrastructure as a material substratum that enables various kinds of higher-level activity, as for example the railroads that make possible public transportation, or the power grid that provides the foundation for many industrial and corporate production activities. Such a view used to dominate corporate and social scientific thinking about large, distributed information systems, and it has played an important historical role in policy approaches to 'informatization' (Ciborra & Hanseth, 2000; Jensen & Winthereik, 2013). There is also a strong traditional association between this 'modernist' view of infrastructure and the notion of societal, economic, and scientific progress – think of the important symbolic function often fulfilled by prestige projects such national telephone networks, motorways, or water supply systems (see Larkin, 2013). By contrast, drawing on research on large technical systems (Hughes, 1983) as well as the critique of social and technological essentialisms (Haraway, 1989; Latour, 1987), STS scholars (Star & Ruhleder, 1996; Star, 1999; Bowker & Star, 2000) have proposed an influential view in which infrastructure is equated not merely with the material artifacts that constitute it, but simultaneously with material technology and the practices to which it gives rise. The latter, socio-material view of infrastructure parallels more recent computer science approaches to large information systems, which have increasingly moved towards distributed development paradigms (Ciborra & Hanseth, 2000). These two definitions imply different ways of funding and organizing tool development, which can simultaneously be seen as blueprints for how to distribute responsibilities among different groups of researchers, administrators, funders etc. The former, more monolithic definition is compatible with a tightly coordinated, centrally managed implementation process, while the latter, socio-material

conceptualization suggests a more decentralized approach that emphasizes the creative role of local users. The strategic value of these definitions to individual actors will depend on how exactly they are operationalized in a given context (e.g., who counts as a local user?), on the position of the actors within a specific science system, and on their particular intellectual, economic, and political interests. For example, some actors may seek to underline the potential knowledge benefits that could arise from creating economies of scale in the geographical and epistemic organization of scholarship, an approach that resonates with the more traditional view of infrastructure. Alternatively, actors may have an interest in distributing design authority and control over resources across sites, and therefore advocate a definition that is closer to the second conceptualization.

To be sure, we should not assume that the infrastructure discourse – once instantiated in official documents, calls for funding, and administrative frameworks – remains completely stable and will henceforth exert a one-way deterministic force on scholarly practices. Rather, the infrastructure discourse should itself be seen as an interface between actor groups who work under different sets of constraints (see Hajer, 2003). The way a given conceptualization of infrastructure is enacted by science administrators for example may differ from its literal formulation in official policy documents, and scholarly grant recipients will continue to interpret formal policy requirements in ways that suit their respective situation (Atkins, Held & Jeffares, 2011). Such forms of behavior, I would argue, tell us something specific about a dominant discursive construction in their own right – different infrastructure policies after all will require tailor-made actor strategies for adapting them to existing disciplinary and institutional working conditions. Policy makers in turn will try to monitor how formal strategies are enacted in practice, and may choose to adapt the exact goals and modalities of longer-term development projects accordingly. With these analytical caveats in mind, let us review the principal mechanisms by which infrastructure initiatives interfere with disciplinary tool development practices.

Firstly, they reroute money that might otherwise have been spent on traditional disciplinary institutions. Thereby they make it possible to sustain distinct organizational entities that exist partially outside the disciplinary employment system, such as projects or centers for digital scholarship, where humanists often collaborate with computer and information scientists. Academics working in such contexts have a certain freedom from disciplinary obligations such as teaching and publishing. They are also

forced, however, to cultivate a managerial self-awareness that provides the basis for rational use of resources (time, money, personnel) within the project or center. Moreover, participants need to operate with an implicit or explicit idea of the prospective infrastructure users, which may often span a host of different disciplines. Here it should be noted that there is an existing tradition of often project-based computational scholarship in the humanities in many countries, for example conducted in humanities computing centers, linguistics departments, or national heritage institutions (see Zorich, 2008). It is likely that such institutions will be one of the primary beneficiaries of current infrastructure grants, thus potentially affecting their relation to disciplinary university departments. However, the working conditions within particular projects and centers can vary significantly, depending on the amount, dissemination rhythm, and bureaucratic modalities of available grants. A more centralizing discursive construction of infrastructure for example suggests a small number of tightly coordinated, complementary funding opportunities, so at to avoid redundant investment. A more decentralized view is compatible with a variety of parallel funding opportunities, thus accommodating the possibility of very similar projects receiving grant support.

A further way of modulating the social organization of tool development is constituted by coordination mechanisms (cf. Barjak et al., 2013). In more traditional disciplinary contexts, choices about which tools need to be built primarily follow the intellectual considerations of disciplinary elites. Coordination with development activities at other research sites will often be somewhat informal and voluntary, taking place for example in personal meetings at conferences or in peer-reviewed journals. A decentralized approach to infrastructure development will be similar to this arrangement in that it will leave wide leeway to distributed actors. By contrast, a more centralizing vision will tend to operate with explicit managerial instruments, such as roadmaps. The latter will require individual projects to plan their development activities not only in accordance with the intellectual requirements of a discipline, but also with the overarching vision of the grant-giving body. An important feature of infrastructure coordination mechanisms therefore is the extent to which they respect disciplinary logic. For example, a funding framework may give researchers significant authority in choosing which tools and facilities should be developed, or it may intentionally override their judgment in order to counter the perceived disciplinary 'fragmentation'. A situation may therefore arise where disciplinary researchers do not acknowledge the need

for new research instruments that suit an overarching policy vision, or vice versa. Of course, disciplines are themselves not necessarily homogeneous. Another possible scenario is disagreement between different groups of researchers who do not share each other's expectations towards infrastructure.

Coordination mechanisms also have implications for the reproduction of research methods, which in turn is intimately related to the performance of disciplinary identities (Whitley, 2000; Becher & Trowler, 2001). Much current talk about the potential of digital instruments in the humanities in fact speculates on how the latter may widen empirical scope, facilitate the use of algorithmic analysis, and complement hermeneutic approaches with more exact knowledge claims (Cohen, 2010; Williford & Henry, 2012; ESF, 2011). At the same time, there are many cases in which scholars resist or at least ignore the publicized introduction of digital instruments because they perceive it as an uninformed attempt by policy makers and techno-enthusiasts to force on them new ways of going about their work (Piersma & Ribbens, 2013). As I will argue in the subsequent analysis, infrastructure initiatives frequently attempt to coordinate tool development by mapping it onto existing practices and methods in a field. We could say that coordination mechanisms in such cases reify method, in that they treat it as a largely context-independent objective protocol. However, when infrastructural tools become widely available, and if a critical mass of researchers actually takes them up, they may over time affect what counts as an accepted method in a given field. Comparing infrastructure initiatives in terms of their underlying strategic considerations thus is relevant not least because it will allow observations about how they redistribute methodological authority.

Methods

The source materials on which the following analysis is based were collected through a combination of methods that reflect empirical differences in the case studies. As for the European infrastructure projects DARIAH and CLARIN, I studied a large number of policy documents, published by the European Commission and other organizations. Together with the project-related documents (conference presentations, newsletters, scholarly publications) circulated by DARIAH and CLARIN participants, this provides a good insight into the formal goals and internal organization of the two projects. To complement these materials with a less formal view on

day-to-day work, I conducted a series of seven semi-structured interviews with project leaders and 'regular' participants of DARIAH and CLARIN. Given the more community-driven approach to scholarly infrastructure in the US, by contrast, there are fewer official policy documents available (mainly a number of commissioned reports). Other relevant materials (refereed publications and essays by digital humanists, various published interviews with scholars and funding officers) tend to provide an explicitly subjective perspective. In this case, I decided to complement my data collection with four semi-structured interviews with leading scholarly protagonists. The style of my narration also varies with different empirical conditions. European infrastructure initiatives are characterized by a more formalized approach, which results in a lot of acronyms and an important role for organizational actors (ESFRI, EC). My account of corresponding US developments on the other hand is dominated by charismatic individuals, thus reflecting the different way in which infrastructure development has been institutionalized in North America.

European Union: Digital infrastructure as a catalyst for integration

There are currently two large digital infrastructure projects in the humanities in Europe, DARIAH and CLARIN, both jointly funded by the European Commission and a number of individual member states. CLARIN aims to offer centralized access to extensive linguistic corpora, as well as tools for searching and analyzing them. Originally grounded in the community of computational linguistics, CLARIN means to expand its user base to all researchers in the humanities and social sciences with a methodological focus on textual materials (CLARIN, n.d.). The goal of DARIAH is to facilitate access to distributed data repositories and to develop a suite of digital tools that will gradually support all aspects of the scholarly work process in the humanities at large. This includes for example applications for text mining, the collaborative annotation of manuscripts, and the visualization of spatial structures and movement (Anderson, Blanke & Dunn, 2010). The European organizational framework in which the projects operate foresees a two-stage development process, i.e. an approximately three year preparatory phase followed by an equally long implementation or construction phase. This will be accompanied by extensive outreach and training activities. The total construction budget for DARIAH is estimated at €20 million, and that of CLARIN at €104 million (ESFRI, 2010).

This vision of infrastructure, reminiscent of 'big science' through its approach to international collaboration, its formal emphasis on centralized project management, and the idea of advancing research by creating large-scale instrumentation, is informed by a specific policy strategy of the EC, in which the development of networked digital facilities for research plays a central role. For more than a decade, European policy makers have pursued the strategic goal of creating an integrated European Research Area (ERA). Their assumption is that the continent's scientific and economic competitiveness would be vastly improved if the organizational fragmentation of European science into a patchwork of individual national research systems could be overcome. According to the EC, the current situation causes duplication of research effort, obstacles to scientific career mobility, and a suboptimal performance in turning basic research into marketable products. The strategic document that first outlined the concept of ERA, published in 2000, presents research infrastructures as an important instrument for bringing about the desired integration (EC, 2000). Two different types of facility are subsumed under this heading: single-sited facilities housing specific instrumentation, as well as distributed electronic infrastructures, offering access to data and tools for analysis and collaboration. The EC ascribes the latter a particularly important role, since it assumes that spatial, institutional, and epistemic fragmentation can be effectively circumvented through creating 'virtual research communities' (EC, 2007). Recent political science literature cautions against treating the EC as an internally homogenous group of technocratic experts who give up ideological and partisan affiliations once they take up office (Favell & Guiraudon, 2009; Georgakakis & Weisbein, 2010) – after all, many top Commission officials have started their careers as professional politicians in national parliaments. Nevertheless, the ERA plan is widely perceived as the ideal-typical expression of a neo-functionalist integration strategy that in various ways challenges existing institutional and political structures in the member states (Edler, Kuhlmann & Behrens, 2003; Guzzetti, 2009). Although several assumptions underlying the ERA have been questioned, for example the notion that research output is straightforwardly maximized through integration of national research systems (Vonortas, 2009), or that competitive European funding will indeed bring about a more homogenous structure of the scientific landscape (Breschi & Cusmano, 2004), a key focus of the recent iteration of the framework programs, Horizon 2020, continues to be scientific collaboration across countries, disciplines, and sectors, facilitated through ICT. Characteristically, and in contrast to the US, current European research

policy on digital infrastructure does not distinguish between science and the humanities – infrastructures for particle physics are conceptualized, planned, and evaluated in the very same committees and reports as those for the humanities. The EC funds infrastructure projects directly during the preparatory phase, which is then followed by an implementation phase in which the member states cover the majority of the costs. Recently, the funding system has been further developed through the creation of the juridical entity European Research Infrastructure Consortium (ERIC). Meant to facilitate legal and administrative negotiations between participating countries, certification in the ERIC framework makes projects legal persons under European law, and it allows them to apply for both European and national funding calls.

It is important to remember, however, that neither DARIAH nor CLARIN are building infrastructures from scratch. Both in fact draw heavily on in-kind contributions by numerous sub-projects, i.e. digital tools, facilities, and expertise generated in previous, nationally based efforts at digital scholarship. In order to properly understand the effect of current European infrastructure projects, it is necessary to take a closer look back at these preceding undertakings.

European digital infrastructures before DARIAH and CLARIN

In most countries, these efforts have originated in humanities computing, a field whose practitioners apply computational methods to research questions in theology, linguistics, history etc. Networking among computational humanists had begun as early as the 1950s (Wisbey 1962; Busa 1980), and while featuring regular transatlantic exchanges, the scene was at that point small enough for most international members to know each other personally. Originally, the use of information technology in the humanities was directly tied to the physical university infrastructure, insofar as computers were unwieldy, centralized mainframes operating with punched cards. Using these facilities required registering for use ahead of time. Mainframe staff could easily keep track of computing operations requested by the users, a type of information that in turn allowed for the development of reusable artifacts, for example word indices (Hockey, 2004). Most users were themselves specialists with a firm grounding in the humanities computing community. However, the advent of PCs made computers a much more widespread tool, also for scholars with no prior knowledge of programming. It allowed for incorporating computers into scholarly practice in many different ways, not only for linguistic or statistical

analysis, but also for word processing or the creation of hypertexts. As a result, many computational humanists became concerned with a perceived risk of fragmentation and duplication of effort (Hockey, 2004). The 1980s therefore witnessed the emergence of a number grass roots standardization endeavors, such as SGML and the Text Encoding Initiative. These provided general guidelines for marking up textual data, thus facilitating interoperability and data reuse among scholars.

Individual national attempts to create digital infrastructure for the humanities in Europe started to take distinct turns in the mid 1990s. In the UK for example, a group of prominent digital scholars set up the Arts and Humanities Data Service (AHDS) in 1995, with funding from the Joint Information Systems Committee (JISC) and the Arts and Humanities Research Council (AHRC). Administrated from King's College London, and building on five university-based hubs, its mission was to collect, catalogue, preserve and promote the re-use of digital resources resulting from research and teaching in the humanities (Greenstein 1998). After having funded the AHDS for twelve years, however, the AHRC decided to discontinue its financial support. The council justified its decision with the argument that British universities by then had developed the capacity to sustain digital data services independently, thus making a national infrastructural investment superfluous (Millet, 2006). Another problem arguably was the difficulty to demonstrate added value of infrastructure to research. When reviewing the AHDS in 2006, the funders AHRC and JISC were particularly interested to know whether the AHDS offered "good value for money", and whether it had made possible any research "which would not have occurred otherwise (AHRC/JISC, 2006)". As Bates (2006) notes, however, the culture of citing digital resources in scholarly disciplines is underdeveloped, thus making it difficult to quantify their intellectual 'impact'. Building digital resources in itself did not count as valid research output in the national research assessment exercise.

Another country with a strong foundation of humanities computing projects, albeit with a historically somewhat different approach to digital infrastructure, is Germany. A number of undertakings, well-known in the international humanities computing community, have existed for almost a decade, for example TextGrid at Tübingen University. Set up in 2006, TextGrid is a so-called Virtual Research Environment that offers access to substantial textual corpora, as well as tools for storage and analysis. While there has never been an attempt to draw individual humanities computing initiatives together in a national infrastructure like AHDS, TextGrid aims to

fulfill an infrastructural function in the sense of convincing other German projects in digital scholarship, mostly based at universities, to adopt its content management software and analysis tools (Textgrid, 2014). Funds for humanities computing facilities like the latter have traditionally been provided by a combination of monies from Länder and Bund. Although public funding has generally been more generous than in the UK, a recent report by the Wissenschaftsrat (2011: 35-6) has critically observed that recent budget cuts in university block funding and the concurrently increasing importance of research grants poses a threat for infrastructure-like facilities such as TextGrid. Too strong a reliance on project-based funding, the Wissenschaftsrat argues, threatens the accessibility and reliability typically associated with infrastructure.

In both Germany and the UK, then, we can observe a relatively strong dependence of community-driven digital infrastructure initiatives for the humanities on a relatively small number of predominantly public funding sources, combined with a trend towards decreasing block funding. The EC has emerged as an important source of funding and political support for digital infrastructure against this background, with the power to instantiate its visions through funding programs such as FP7, Horizon 2020, and the European Structural Funds. In the terminology of actor-network theory, the EC has become an 'obligatory passage point' for digital infrastructure (Callon 1986). The historical perspective also makes clear that there are a few crucial differences between European and national policy makers' expectations towards the function of infrastructure. To the AHRC and JISC for example, expenditure on AHDS was particularly unattractive because it saw infrastructure just as another fixed expenditure on public facilities, such as money spent on maintaining university buildings, but without any particular added value in terms of 'better' or more publications for scholars (AHRC/JISC, 2006). For the EC - which is in a constant competition for authority with national policy actors – digital infrastructure does have an added, political value. By offering specifically configured funding opportunities for digital infrastructure, the EC means to interface directly with disciplinary research communities across Europe, thus requiring them to coordinate the development and use of digital research tools on a supra-national scale, and in a way that circumvents possible 'balkanizing' impulses given by domestic policy actors.

Roadmapping

A key agent in organizing and administrating these infrastructure plans is

the European Strategy Forum for Research Infrastructures (ESFRI), a supranational body constituted following an agreement of the European Council of Ministers and the EC in 2002. ESFRI is staffed with delegates nominated by the member and associate countries, and it has an important influence on the distribution of funding – a recommendation by ESFRI is a precondition for any large infrastructure project to acquire European and increasingly also national grant support. Its main contribution is a periodically updated roadmap to "identify research infrastructure of pan-European significance, as well as emergent new infrastructures (ESFRI, 2006a)."[54] Through the roadmapping process, the EC hopes to ensure a high degree of coordination in the development of infrastructure. All projects, irrespective of their academic field, are described and administrated through the same managerial instrument. This entails a move that Callon (1986) has described as 'translation', i.e. a process of turning disparate elements (the tools, knowledge, and organizational structures created in preceding humanities computing projects) into a new socio-material network.

For one, applying for European funding through participation in ESFRI's roadmap requires framing distinct kinds of infrastructure projects according to shared criteria. Very heterogeneous proposals with complex prehistories, hinted at in the above, are thereby transformed into comparable phenomena that can be conceptually described in terms of their 'relative maturity'. Apart from 29 projects in the natural sciences and engineering, and next to DARIAH and CLARIN, the first iteration of the roadmap includes for example also three social sciences projects (ESFRI, 2006b). These pursue very different and in a certain sense less ambitious goals when compared to the two humanities proposals. SHARE and ESS aim to harmonize and provide centralized access to census and health care data across the member states. CESSDA is a multidisciplinary repository of social sciences data sets, such as survey results and statistical information provided by other public institutions. While the goal of the three social sciences projects thus could essentially be described as general-purpose data harmonization, DARIAH and CLARIN aim to build nothing less than comprehensive research instrumentation for a very large variety of disciplines.

Another seemingly natural category that in fact constitutes an

54 Following the European incentive, many countries have since started to develop their own national roadmaps, which are typically closely aligned with the shared European perspective (ESFRI, 2011).

important outcome of the socio-material translations effectuated through roadmapping is that of implementation. It allows to frame the process of creating infrastructure as a sharply defined phase within a singular project framework, thus making the projects more amenable to administration and evaluation by ESFRI. However, a side effect of such formalization is that the development of infrastructure becomes something that can in principle be thought of as conceptually separate from the characteristic practices and sociology of the disciplinary context in which the prospective users work. The technically connoted term 'implementation' in fact has implications for defining the success criteria of EFSRI projects: it subtly suggests that once the physical facilities are installed and operational – 'implemented' -, users from all disciplines, also the vast majority of humanities scholars with no prior experience in using digital research instruments, will adapt their practices to the rigidities of the newly built infrastructure. Failure is synonymous with lack of 'uptake'.

While the members of the participating humanities projects were naturally happy to get access to a new source of funding (which in some cases, for example in the UK, were direly needed after national funding streams had all but dried up), many of them find the pervasive integration of disciplinary practices through a centrally coordinated, pan-European infrastructure, as envisioned by the EC, to pose a rather steep expectation. Participating in ESFRI required applicants in both projects to make promises about infrastructure comparable to those normally heard in fields with a long tradition in large-scale instrumentation, such as astronomy and physics. At the same time, DARIAH and CLARIN have a rather limited budget for central coordination (an annual amount of €0.4 and 0.6 million respectively (EC, 2013)) and the development of wholly new facilities. Both initiatives in practice adopt a more decentralized approach than originally anticipated in the EC's strategic vision. Much current work consists in gradually integrating in-kind contributions from the constituent national sub-projects, and in encouraging the adoption of the existing digital resources beyond the existing user base.

In trying to coordinate individual tools and development activities on a European scale, DARIAH and CLARIN pursue strategies that reflect their different disciplinary origins. CLARIN has first and foremost been an initiative by computational linguists, a field of research that often involves the algorithmic or statistical analysis of large language corpora. Comparative research has found both computational and 'traditional' linguistics to be atypical when compared to other humanities disciplines,

insofar as there is an unusually strong consensus on methodological standards, theoretical frameworks, and research problems (Whitley, 2000; Fry & Talja, 2007). Integration activities therefore predominantly focus on making accessible large linguistic datasets, yet without entailing the type of fundamental discussions about the nature and purpose of data that frequently occur in digital initiatives in other scholarly disciplines. CLARIN also takes a more formal governance approach when compared to DARIAH. All contributing organizations are classified according to six different types of centers. For example, A centers take on infrastructural responsibility that require particular commitment in terms of funding and maintenance, while B centers merely guarantee access to the resources they themselves offer. The approach here is to specify in great detail what any member organization is expected to contribute. Regardless of the relative methodological consensus within linguistics, CLARIN is faced with the typical problems of infrastructure development (Edwards et al., 2007), namely diverging soft- and hardware standards, reluctance of individual members to accept CLARIN as an overarching organizational reference point, the vagaries of national research policies etc.[55]

DARIAH in contrast targets disciplines such as literary studies, history, and archeology. Many of these are characterized by strong methodological and theoretical plurality, by distinct national research traditions, and by little to no disciplinary tradition of using computational approaches. In this context, the EC's premise of building a pervasive digital infrastructure for the purpose of integrating different fields across Europe acquires missionary overtones – the prospect is to 'bring technology to the humanities'. Several of the DARIAH participants I have interviewed are uncomfortable with this missionary function, since it sometimes results in a certain hostility on the part of the traditional humanists, who feel that they themselves know best what form of infrastructural support they need or do not need.[56] DARIAH director Tobias Blanke expressed his reservation about the idea of 'integration through infrastructure', as well as the centralized approach to coordinating infrastructure development that goes along with it.[57] The Commission, Blanke suggests, has modeled its technological vision on experiences from building monolithic, single-sited facilities such as

55 Skype interviews with Steven Krauwer (15 May 2014), Laurents Sesink (21 May 2014), and Jan Odijk (4 June 2014).

56 Skype interviews with Mirjam Blümm (8 May 2014) and Tobias Blanke (4 June 2014).

57 Skype interview with Tobias Blanke, 7 May 2014.

CERN. While facilitating centralized administration by the Commission, this approach threatens to cut ties between infrastructure projects and the disciplinary landscape of its prospective users. In contrast to the EC's vision of infrastructure as an autonomous organization run by full-time managers (Rizzuto & Wood, 2013), the directors of DARIAH purposefully decided to divide management duties among three individuals, so as to have enough occasion for research and thus retain contact with the scholarly communities.

DARIAH tries to walk the line between disciplinary plurality of theory and methods on the one hand, and the policy expectation towards disciplinary integration on the other, by proposing to organize digital research infrastructure around so-called methodological commons, i.e. tools that can be applied across a large variety of scholarly disciplines. The underlying assumption is that all scholarly work processes can be reduced to a set of basic, universal elements, such as 'discovering', 'annotating', 'comparing', 'referring' (Anderson, Blanke & Dunn, 2010). Using the latter as a principle for coordinating tool development, it is possible to sort existing applications into non-redundant categories, as well as providing a heuristic for identifying gaps in research instrumentation. To be sure, it is not clear whether the basic praxeological elements presupposed by this approach actually exist in the structure of scholarship, or whether they are rather an achievement of the rational development strategy of DARIAH itself. The 'commons' do, however, formally commit the project to an ongoing process of refining its toolset and seeking engagement with users beyond humanities computing, so as to justify its claim of covering the whole bandwidth of research practices. Combining such engagement with enough time and funding, DARIAH might ultimately manage to link up with the institutional reproduction of methods, e.g. through the incorporation of its tool set in undergraduate methodology classes.

Creating organizational flexibility within formal organizational schemes

An interim evaluation of the financial and governance aspects of ESFRI projects, conducted after three years of funding during the preparatory phase, critically remarked that both CLARIN and DARIAH still resemble a network of specialized national projects, rather than a centrally coordinated, European construct widely used across the humanities (EC, 2013). However, both project participants and ESFRI administrators, who are often reputed scientists themselves, make use of informal ways of 'working around' some of the strict assumptions underpinning the roadmap. This creates a degree of

organizational flexibility within the formal scheme.

For one, an important criterion of success for infrastructures will likely be a measure of the distribution and sheer number of users, assessed for example through server log analysis. According to one of my informants, however, indicators such as these can be 'gamed' by formally adding new national sub-projects, which automatically increases the number of users in specific regions. Moreover, rather than mechanically executing idealized Commission policies, ESFRI administrators sometimes take an intentionally benevolent approach to assessing projects already included in the roadmap, since these are seen as existing investments. Milena Žic-Fuchs (2013) for example, a linguist and member of the evaluation working group, publicly argues that although the ESFRI humanities projects may to some extent fall short of an integrated, singular infrastructure, their 'added European value' may still become apparent if evaluation highlights how certain research questions can be tackled even through a relatively loose network of national infrastructures. Evaluation here is difficult to distinguish from demonstrating the value of a funded project. The context-sensitive approach to evaluation advocated by Žic-Fuchs moreover tends to be supported by the social scientific research on digital infrastructure that European policy makers regularly commission to facilitate the implementation process (Barjak et al., 2013; Voss et al., 2007). Most of these studies conclude by encouraging policy makers to respect the "specific demands" of the humanities, and to avoid an overly top-down approach to the development process (Barjak et al., 2009: 596). Over time, such findings and evaluation practices might well contribute to a subtle redefinition of the official policy conceptualization of scholarly infrastructure.

United States: Infrastructure as an emergent property of ongoing digital scholarship

An important event in conceptualizing digital infrastructure for the humanities in the US was the publication of *Our Cultural Commonwealth*, a report commissioned by the American Council of Learned Societies (ACLS, 2006). The authors of the ACLS report constitute a selection of distinguished 'traditional' humanists, information scientists, as well as several influential figures in digital scholarship: John Unsworth (former head of the Institute for Advanced Technology in the Humanities (IATH) at the University of Virginia), who also acted as chairman of the commission, the late Roy Rosenzweig (former head of the Center for History and New Media at

George Mason University), and Jerome McGann (editor of the famous Rosetti Archive). *Our Cultural Commonwealth* is itself conceived as a response to another strategic policy document, namely the hugely influential NSF report by Dan Atkins et al. (2003), in which the popular term 'cyberinfrastructure' was coined. Atkins and his colleagues define cyberinfrastructure as large-scale facilities for the storage, dissemination and collaborative analysis of massive datasets in science and engineering, thus reflecting not least the authors' interest to position their own research in computer science as an enabling, auxiliary discipline for other fields. The Atkins report was widely perceived as a point of reference in the discussion about digital research infrastructure in both Europe and the US (Jankowski, 2009), and it has helped mobilizing significant amounts of funding by the NSF. But while the ACLS report can be seen to take advantage of the attention Atkins et al. had created on the part of policy makers and funders, it also departs from their perspective in a few significant regards. Several commentators have pointed out that the Atkins report presents a somewhat techno-deterministic vision of infrastructure-enabled science, in the sense that it universally equates 'better' science with more computing power, and that it disregards disciplinary specificities and questions of embedding new research tools in established practices (Jankowski, 2009). The ACLS report in contrast adopts a vision of digital infrastructure that is explicitly informed by the work of Star & Ruhleder (1996). As suggested in the theoretical introduction above, this definition of infrastructure is relational. Infrastructure is seen not as a specific thing, but rather as a state that occurs when the various practices of interacting users fall into a workable configuration. This view emphasizes the human expertise connected to material tools, as well as the emergent and evolutionary development of technology in conjunction with practice. Digitally enabled scholarship here is portrayed as a matter of small scale 'tinkering', rather than operating with grids and supercomputers.

In order to contextualize the ACLS report's vision of digital infrastructure, the dominant organizational format of digital scholarship in the US must be taken into account. Similar to the European context, efforts to coordinate digital scholarship on a larger scale predate the current debate on digital infrastructure. These efforts have largely been carried out in campus-based, so-called digital humanities (DH) centers, which usually answer directly to their provost, and which have often originally been set up to serve the special ICT needs of faculty researchers (Clement & Reside, 2011). An important difference to the European context is the pronounced

divide between staff and researchers in the US academic job system. Many digital scholars have started their careers in staff positions, thus preventing them early on from advancing in the professional hierarchy of traditional disciplines, such as Classics or English (Nyhan, 2012). At the same time, American DH centers are often funded through a combination of sources: a certain amount of base funding from the university; commercial revenues, e.g. from subscription fees for the use of digital archives; as well as federal, private, and philanthropic funding. The diversity of important funding sources, both private (IBM, Microsoft, Google) and philanthropic (the Arthur P. Sloan Foundation, the McArthur Foundation, or the Getty Trust, to name but a few), constitutes a difference to the situation in many European countries, where digital scholarship is predominantly funded by a small number of public bodies. According to a widespread organizational practice, various grants from these funding streams are pooled to create a number of stable, but locally defined professional functions within the center, for example a scholar-programmer and a scholar-web designer (Clement & Reside, 2011). Although individual grants are relatively small (typically not exceeding $60000), this organizational practice has historically provided a certain independence for the DH centers, insofar as it has allowed them to draw together money from different sources, yet without tying it to a singular purpose, such as the delivery of a specific product. Instead, the various funding streams could be used to create a center-internal job ecology that allows to combine service functions with intellectual aspirations, i.e. deliver a product but combine that product development with a strong research component (cf. influential digital editions and database projects, such as the Blake Archive or the Brown Women Writers project). Networking among digital humanists in the US has been traditionally very strong, thus creating a quasi-disciplinary structure, yet without formal recognition in the shape of actual university departments. Many now prominent practitioners have spent formative years in a handful of influential institutions (such as IATH, or Brown University's Scholarly Technology Group), from which they have then spread out to other parts of the country, often starting up centers of their own at their new alma mater.

Against this background, the emphasis of the ACLS report on the emergent and evolutionary aspects of infrastructure development makes particular sense. It allows to portray the existing efforts in digital scholarship, conducted at various centers all over the country, as indispensable preparatory work, and the centers themselves as the primary agent in the creation of distributed research technology. The latter in fact is

pictured as something that slowly emerges as a side-effect of these ongoing activities, rather than as something that is created in a singular project, and managed on terms dictated by a centralized policy actor. The ACLS conceptualization of digital infrastructure thus emphasizes the need for more funding for existing DH centers, while simultaneously asserting their organizational and intellectual independence from both funding bodies and local university administrations. After all, according to Star and Ruhleder (1996), "infrastructure is not developed, it evolves".

This strategy can be further illustrated by examining the institutionalization of federal funding for digital infrastructure in the humanities, which is bound up with the history of the very term 'digital humanities'. In 2004, the NSF acted on the recommendations presented in the Atkins report by setting up an Office for Cyberinfrastructure, later on re-named Division of Advanced Cyberinfrastructure. The mission of this new body has been to provide centralized funding and administration for cyberinfrastructure in science and engineering. Following the model of the NSF, the corresponding federal funding body for the humanities, the National Endowment for the Humanities (NEH), set up the Digital Humanities Initiative to provide support for digital infrastructure in 2005. Two years later, the institution was renamed Office of Digital Humanities (ODH) to indicate its permanent character. In contrast to its NSF equivalent, the NEH institution thus carries 'digital humanities' instead of 'cyberinfrastructure' in its title, and it is explicitly positioned as a partner and liaison for the DH communities, rather than a centralized infrastructure reformer.

According to Kirschenbaum (2010), a number of developments that involve both prominent digital scholars as well as NEH officials converged to stabilize the term 'digital humanities'. Firstly, a book project launched by several computational humanists in 2001 was in need of a title. Co-editor Ray Siemens suggested *Companion to Humanities Computing*, which was then the preferred term in the community. The publisher's editorial and marketing team, by contrast, favored *Companion to Digitized Humanities*. Intent to shift emphasis away from mere digitization, and to promote institutional recognition as a discipline, John Unsworth finally convinced the others of the title *Companion to Digital Humanities* (see also Kirschenbaum, 2012). Around the same time (2005), the NEH had decided to set up a small funding initiative to promote digital scholarship, which would eventually become the above mentioned Digital Humanities Initiative. The leader of the initiative, Brett Bobley (2010), recalls picking up the label digital humanities

from his continuous, personal conversations with digital scholars. In an interview, he explains his vision of the ODH as providing support to a conceptually proactive DH community, rather than trying to steer them in a top-down fashion: “Cyberinfrastructure can't be built alone. It is important that the NEH speaks with the community on a regular basis to ensure our funding strategies are best suited to help the field (Smith, 2009).” This approach is also reflected in the funding instruments offered by the ODH. The relatively modest start-up grants (between $5000 and $60000) encourage tool development at more or less established DH centers, since these have both the expertise and facilities to quickly get new digital projects underway.

A short-lived alternative to the ACLS vision of digital infrastructure must be mentioned. In 2008, a coalition of grant officers at the Mellon Foundation, as well as scholars and computer scientists from the University of Chicago and UC Berkeley, launched an infrastructure project that is in many ways reminiscent of the European approach. Perceived as standing “completely outside the DH community” by renowned digital scholar Stephen Ramsay (2013), the initiators managed to combine funding from the two home universities with a substantial contribution by the Mellon Foundation ($2.43 million in total), with the aim of creating a comprehensive set of scholarly resources in a four year project. The underlying approach differed from the ACLS' in that it did not distribute management responsibility across DH centers, but rather concentrated it in the hands of central management team. This, the initiators, hoped, would put an end to the constant “reinventing of the wheel” that they perceived to result from funding many smaller-scale, but dispersed initiatives (Broughton & Jackson, 2008; Ramsay, 2013). However, project Bamboo quickly ran into substantial problems. According to Dombrowksi (2014), numerous scholars attending Project Bamboo workshops felt alienated by its service-oriented approach. The latter entailed 'requirements engineering' sessions, during which software developers asked invited scholars to describe their research practices in an abstract way (verb + direct object), with the aim of designing tools that would uniquely support those practices. Software development here was carried out not by digital scholars, but by computer scientists and software engineers, and in an organizational framework that did not contain any research component. Following a change in management personnel and the financial decision to reduce outreach activities halfway through the project, communication between project staff and prospective users deteriorated even further (Dombrowski, 2014). Observers from within DH

have criticized project Bamboo early on for what they argued was a paternalistic design approach, and for its disregard of experience gained in previous DH projects (Boast, 2009). Ramsay (2013) has also criticized the epistemic implications of Bamboo's vision of infrastructure. In its attempt to avoid 'reinventing the wheel', he suggests, it mistakenly frames the diversity of scholarly approaches as a problem of redundant organization. Ramsay argues instead that in hermeneutic fields of research, knowledge is not primarily gained through reusable instruments that allow for 'solving' research problems more efficiently, but rather through a corresponding diversity of instruments to bring out different nuances of the research object. When Project Bamboo failed to create either substantial facilities by 2011, or a convincing strategy for a follow-up funding period, the Mellon Foundation decided to terminate the project and dissolve its own cyberinfrastructure subdivision. So far, there have been no attempts to emulate the service-oriented approach to infrastructure adopted by Bamboo.

In summary then, the ACLS report can be seen to have de facto established infrastructure development modalities that are very different from the European initiatives. The latter operate with a formalized, policy-mandated coordination mechanism, set up to counter epistemic and geographical fragmentation of national research systems. As a side effect, tool development is partially detached from the disciplinary logic of individual fields, but also not subject to the conceptual authority of a single group of actors. The ACLS approach by contrast serves to consolidate the institutional and intellectual independence of a particular community of researchers – digital humanists -, under the assumption that the tools they create will eventually converge into a layer of reusable facilities that is of benefit to the humanities at large.

Consolidating the DH center

Digital humanists have attempted to fortify their conceptual influence on infrastructure policy and simultaneously strengthen the position of DH centers throughout the 2000s, thereby using the political attention created by both the Atkins and the ACLS report. The 2007 DH Summit at the University of Maryland was widely perceived as a watershed moment in negotiating the relations between DH centers and funding bodies (Cohen, 2007). The two day meeting brought together digital humanists from 17 leading research centers, policy makers and government officials, as well as many philanthropic and private funding bodies, with the goal of developing an infrastructure agenda. In contrast to ESFRI's highly formalized roadmapping

exercise, the summit combined presentations with break-out discussion groups, and thus offered a relatively informal opportunity for personal exchange between scholars and funders.

The strategy of the attending digital humanists clearly was to translate the ongoing work at existing DH centers into terms compatible with the Atkins report, but in such a way as to secure them significant authority over the coordination of technology development. This is perhaps most clearly expressed in John Unsworth's (2007) plenary address, entitled *Digital Humanities Centers as Cyberinfrastructure*. Unsworth's strategy consists in persuading funders that digital infrastructure already exists, and that it manifests itself in the facilities and efforts undertaken at existing DH centers – to more fully develop it, however, the centers need more support. At the same time, the minutes of the summit document that the prospect of *Digital Humanities Centers as Cyberinfrastructure* was to some extent a euphemism. Many of the issues raised in the break-out discussions in fact reflect the perception of scholars that existing centers are not yet sufficiently networked, and often still too dependent on local campus administrations. Unsworth's promise thus is performative in two senses: not only is it meant to convince funding bodies of the potential of centers to bring about infrastructure, it also implies that the centers have to make an effort to realize this vision.

According to the minutes, the DH practitioners reflected on the need to improve coordination of tool development across individual centers by further increasing networking activities (DH Summit, 2007), not least for political reasons. If centers adopted more explicit coordination strategies, funders would get a stronger sense of supporting the humanities as such, rather than individual scholars. Speaking with a more unified voice could also strengthen the position of digital humanists in the attempt to influence criteria for tenure/promotion, as well as the scholarly grant culture (DH Summit, 2007). A recurrently raised issue was the need to increase not only the sheer number and volume of grants, but also to extend grant duration from two or three to five years, so as to make it possible for the DH centers to engage in longer-term planning. At the same time, it is noteworthy that in contrast to the exclusive emphasis on coordination and integration in the European policy discussion, the summit documents also underline the creative potential that may come with uncoordinated variety, for example with respect to the types of projects and tools undertaken/developed in a given center. Scholarly participants in the break-out groups speak of a trade-off between coordination and variety (DH Summit, 2007): while the former

is desirable insofar as it allows to create economies of scale, the latter is desirable for its innovative potential. Variety here is positively connoted probably because it often translates into local autonomy of individual DH centers.

Managing expectations in an informal policy culture

Given the strategy of coupling tool development rather strongly to the specific disciplinary logic of digital humanities, the ACLS approach to digital infrastructure circumvents some of the difficulties encountered in Europe and in Project Bamboo. For one, it avoids the problem of a 'gap of implementation' insofar as it does not set up the goal of creating widely used technological facilities in a clearly circumscribed project, but instead suggests that infrastructure is what gradually emerges from ongoing work at DH centers. Moreover, prominent digital scholars have from early on attempted to shape not only the definition of digital infrastructure, but also the criteria by which success or failure of respective projects can be gauged. The relatively flat hierarchy and informal communication between digital humanists and funding bodies – exemplified by the DH Summit 2007 – here is an asset for the scholars. In contrast to ESFRI's formalized roadmapping process, it allows to avoid specifying strategic deliverables in a way that might later on backfire, and it creates an opportunity to infuse any promises with certain narrative safeguards.

One characteristic strategy has been to domesticate the possibility of failure. On the occasion of the strategically important DH Summit 2007, where he was faced by an audience that included also a considerable number of funding bodies, John Unsworth (2007) argues that failure of individual digital projects should be conceptualized as an opportunity for learning. What is needed is a culture of honesty, rather than hyperbolic future scenarios. A complementary strategy is to emphasize that building infrastructure is at heart a research endeavor. In a variation on the 'endless frontier' theme by Vannevar Bush, Unsworth portrays infrastructure development as an open-ended, profoundly intellectual process, rather than a provision of clearly specifiable service facilities. This perspective implies that DH center staff should be considered researchers in their own right, and that their intellectual perspective should override short-term, functionalist criteria of value.

In spite of these efforts to manage expectations, there are several aspects about the ACLS vision of infrastructure that continue to be seen in a critical light by some influential observers. These objections essentially are

the same that Project Bamboo had already tried to address. A 2008 report by the Council for Library and Information Resources for example notes a problematic tendency of DH centers to be too much oriented towards conducting research, which hampers the development of reusable facilities. The report criticizes that many DH centers currently resemble overspecialized "silos", thus failing to deliver digital resources that address "community-wide needs" (Zorich, 2008: 4-5). One could finally argue that the ACLS report defends a form of elitism, insofar as it aims to concentrate resources and technological expertise in a few well-established institutions, while limiting access to these resources for scholars with no prior experience in computational techniques. Several academics outside the DH scene have moreover described the latter as particularly cliquish, with regular, rather emotional discussions about what type of research should legitimately be allowed to call itself digital humanities (Pannapacker, 2011).

Discussion

In this paper I have provided a comparative perspective on current initiatives to build digital infrastructure for the humanities in Europe and the US. Thereby I have meant to move beyond analyzing the shaping of technology within individual projects and instead trace in a more encompassing way how dominant research policies mediate the reorganization of disciplinary tool development. An inquiry along such lines has been called for by researchers in STS and neighboring fields (e.g., Ribes & Lee, 2010), but is not commonly undertaken, arguably because of a traditional disciplinary focus on ethnographic descriptions of individual laboratory-like sites as well as a relative analytical neglect of the interaction between research practices and policy practices. Analyzing this interaction, however, becomes increasingly topical as traditional relations between science policy, funding bodies, and researchers are being reconfigured in many countries (Mirowski & Sent, 2008; Whitley & Gläser, 2010). Current infrastructure initiatives illustrate a particular aspect of this development: Choices about tool development here are no longer the prerogative of disciplinary elites, but increasingly follow the shared strategic outlook of coalitions of policy makers, researchers, and funders. I have argued that in this context, the discursive construction of infrastructure acts as an interface between research policy and scholars. Different views of what infrastructure actually is and how it functions have implications for funding and coordinating local tool development, thus making them strategic resource

actors draw on to steer infrastructure policy in particular directions. Insofar as coordination mechanisms often entail rationalizing research practices and methods, the specific modalities of current initiatives may also have an effect on the methodological reorganization of scholarly fields.

In the US, a vocal group of digital scholars with a longstanding experience in developing digital tools have rather successfully mobilized a socio-technical view of networked scholarly instrumentation. This view emphasizes the connectedness of user practices and technology, and it pictures infrastructure as something that develops in an evolutionary fashion. Digital scholars can therefore argue that conceptual and managerial responsibility should be situated at established DH centers. Having historically struggled to combine service functions with research, this strategy has resulted in additional funding and institutional consolidation for the centers, which now have almost exclusive control over development activities. A central assumption of the European Commission, by contrast, is that the creation of digital infrastructure can and should be a catalyst for the integration of national research systems into a more homogenous European Research Area. Digital infrastructure here is pictured primarily as a technical phenomenon that can be built in a number of clearly circumscribed projects, tightly coordinated through the formal instrument of a roadmap. The resulting grant opportunities are taken up by preexisting projects in digital scholarship in various countries, which are often dependent on European political authority and funding, but at the cost of translating their preexisting work into terms compatible with the Commission's policy vision.

The specific ways in which different infrastructure initiatives modulate the social organization of tool development can be further illustrated in relation to older STS research on negotiated judgments of similarity and difference of scientific work. Collins (1985) has argued that key intellectual problems in science, for example the question as to whether a given experiment has been successfully replicated, are never fully determined by purely objective criteria, but always involve negotiation and personal judgment among a core set of reputed researchers. Current infrastructure initiatives reconfigure the relations between researchers, administrators, and funders, thus affecting also the way similarity/difference questions are settled. The European approach to infrastructure, with its strong, policy-mandated emphasis on transnational coordination, requires applicant projects to present strategies for mapping, and thereby limiting, the diversity of practices. In this context, it is no longer primarily the views of a core set of researchers that determine what tools need to be developed,

but rather an amalgam of policy and intellectual rationales. While this creates tensions between tool developers and prospective users, manifesting themselves in a 'gap of implementation' of new technology across the humanities, the central coordination through the roadmap also ensures that no single disciplinary community gains exclusive control over technology development. The socio-technical view advocated by US digital scholars paints diversity of practices as characteristic of infrastructure, as well as emphasizing the emergent development of new technology out of local practices. This implies, however, that the authority to determine what tools constitute useful additions to infrastructure, and which are redundant, should primarily remain with the community of digital scholars, since it is the latter who already dispose of the necessary skills and facilities to develop digital resources. The position of existing elites within DH will thereby tend to be reinforced, thus privileging their technological and intellectual judgment over that of other disciplinary communities. At the same time we should take into account that dominant discursive constructions of infrastructure are malleable and may be redefined over time. The implementation of European infrastructure initiatives for example is constantly monitored by social scientists, who regularly make a case for adapting overarching policy goals to the specific properties of the humanities. There are also indications that ESFRI administrators tolerate a lesser degree of integration of the individual predecessor projects that together constitute DARIAH and CLARIN than originally suggested in the roadmap. US digital humanists, on the other hand, face ongoing criticism that their tool development efforts are not sufficiently oriented to the needs of the wider community of traditional scholars. As can be seen from the deliberations at the DH Summit 2007, they do acknowledge the political need for reacting to such claims, for example by striving for a greater degree of formal coordination across centers.

What do these results in turn mean for the further study of infrastructure by scholars in STS and related fields? For one, they should read them as an encouragement to more explicitly think about their work as a potential source of regulatory knowledge. In the European case, a policy vision of a centrally planned, pervasive infrastructure produces a 'gap of implementation' that is then framed as a research problem for social scientists, commissioned to facilitate technology 'uptake' (Barjak et al., 2013). Such research plays an important supervisory role, in that it is in a position to sanction or criticize the underlying definition of infrastructure. US digital scholars in turn mobilize an existing body of social scientific knowledge as

an expert argument in favor of their simultaneously intellectual and political interests, but do operationalize that knowledge in a rather specific way. While capitalizing on the focus on emergent development that is at the heart of Star & Ruhleder's work, they implicitly privilege existing DH centers as sites of emergent creativity, thus downplaying the significance of more traditional scholarly practices. A second implication is that critical infrastructure scholars should extend their analytical focus from the micro-level of scholars interacting with technology to formal and informal policy settings, so as to take into account the political uses to which their analytical insights and theoretical constructs are put. Much recent STS work operates from the theoretical conviction that dualisms such as nature/culture, or technology/social life, are an artifice that serve the function of 'purification' (Latour 1993). The assumption is that social scientists who are equipped with this insight can perform better analyses of science and technology than, say, traditional sociologists. However, an understudied question is what happens when this STS knowledge in turn starts to travel and is being mobilized by other actors outside the scholarly discourse. In such settings, dualisms as well as non-dualisms are not treated as theoretical problems, but as resources for new translations.

Postscript to chapter 5

In this last chapter I have zoomed in on a central recent development in the history of digital scholarship, namely the concerted investment in the creation of digital infrastructure. In both Europe and the US, there are currently high-profile initiatives underway to create a pervasive technological layer of data and tools for a large variety of disciplines. I have argued that digital infrastructure development takes place against the backdrop of an ongoing reconfiguration of the relations between scholars and policy actors. In contrast to earlier periods where the design of research tools was largely at the discretion of disciplinary elites, infrastructure initiatives present us with a case where researchers, policy makers and funders argue about the authority to take critical choices in regard to instrumentation. The debate revolves around fundamental questions about the proper conceptualization of infrastructural technology, and about how it is best developed and administrated. This meta-discourse about technology and knowledge production could actually be seen as a sort of interface through which scholars and policy renegotiate their relations.

Different ways of conceptualizing digital infrastructure have important implications for the methodological organization of digitally mediated scholarship, for example insofar as they affect judgments about the relative similarity or difference of particular tools. A centralized conceptualization of infrastructure implies a tightly coordinated mechanism for assessing the complementarity of individual development projects, thus increasing the likelihood that relatively similar tools are considered to serve the 'same' method. A more decentralized vision leaves more discretion to local tool developers, and so will lead to a larger diversity of digital approaches. If digital infrastructure indeed becomes involved in the reproduction of disciplinary methods, for example by particular tools being incorporated in undergraduate methodology training and textbooks, the scholarly knowledge machine will be restructured.

To be sure, it is still unclear to what extent these technologies will actually be taken up. In chapters 2 and 3 I have argued that initiatives that aim to change practices on a large scale and in very ambitious ways run the risk of simply being ignored by users, since they imply too radical an incongruence with the existing configuration of the scholarly knowledge machine. This is a challenge particularly for the more centrally coordinated European approach to digital infrastructure. By trying to develop a suite of tools that serves a large bandwidth of academics, often inexperienced in digital scholarship, it is particularly likely to create friction with local

disciplinary practices (Edwards, 2010; Edwards et al., 2011). At the same time, the centralizing European approach also entails the need to involve many different users from as many fields as possible, thereby affording them the possibility to shape the nascent technology in a critical phase of its development. US initiatives, by contrast, are characterized by a more decentralized paradigm, in which digital infrastructure is seen as an emergent property of ongoing work at established centers of expertise. This approach avoids the problem of 'implementation' and the attendant friction, but it is arguably less democratic in that it concentrates significant control over resources and design choices in the hand of a relatively small group of renowned digital scholars.

Conclusion

The topic of this thesis, broadly conceived, has been the question as to how the development and use of digital research instruments is related to changes in the organization and conduct of scholarly inquiry. Although computational methods in the humanities have a long history, it is only recently that they have attracted attention beyond their traditional user communities. Digital humanities, the presently common shorthand for such approaches, is serving as a prism through which various actors – digital and traditional scholars, policy makers, as well as the media – imagine the future of the humanities at large. Practitioners are also heavily involved in current initiatives to build digital infrastructure, an undertaking that mobilizes significant amounts of funding, and one that has potentially long-term effects on future research practices. In the first part of this concluding chapter, I will summarize my conceptual framework and empirical findings, according to the order of my five original research questions. In a second part, I will draw conclusions that further develop my conceptual framework, as well as discussing some practical implications. Finally, on the basis of this, I will suggest directions for future research.

In analyzing my empirical material, I have drawn on three conceptual resources: infrastructure studies (Star & Ruhleder, 1996; Bowker & Star, 2000; Edwards, 2010), theoretical literature on disciplinary conventions that emerge from the packaging of material and semiotic relations (Fujimura, 1987; 1992; Law, 2004), and comparative research on the social and intellectual organization of different fields (Whitley, 2000; Becher & Trowler, 2001). At first sight a heterogeneous set, these approaches can actually be used in a mutually complementary way to highlight a range of challenges that accompany the move to digital scholarship. My principal conceptual assumption is that scholarly knowledge is generated within an infrastructure. Drawing on Star & Ruhleder (1996) I define the latter relationally, rather than substantively: infrastructure is not a specific thing, but obtains when interrelated practices fall into a workable configuration. Embedded in this ecology of practices, and co-evolving with it, are material tools, protocols, and standards, for example regarding data and publication formats. These elements are bound up with the conceptual and social structure of scholarly fields. Established data formats for example constrain possible modes of theorizing, and the peer review protocol serves to validate new contributions to the body of scholarly knowledge. Practices moreover are reproduced through institutional mechanisms. Students learn how to

develop scholarly arguments, write books, and use bibliographies to conduct literature searches in their undergraduate classes, and specialized degree programs teach librarians how to organize and curate information in ways that connect meaningfully to scholarly practice. Functioning infrastructure thereby constantly performs what it means to engage in 'proper' scholarship, so that only particular ways of working, particular knowledge claims, and particular forms of scholarly output are recognized as adequate.

Intersecting this perspective with my other two conceptual resources has allowed me to theorize in more detail different aspects of the scholarly infrastructure. The work of Whitley (2000) is useful for highlighting its disciplinary specificities. When compared to most natural sciences, scholarly fields such as history, philosophy, and literary studies are characterized by a relatively weak degree of social and intellectual integration. There is a characteristic plurality of theoretical approaches, which coexist and fuel each other through the conceptual contradictions they create. The dominant format of circulating knowledge products is the monograph, which gives individual writers considerable intellectual and stylistic freedom. Highly integrated fields, by contrast, such as the various sub-disciplines of physics, are characterized by strong agreement on shared theoretical foundations and pressing research problems, and by a tightly coordinated work process that is frequently organized around the use of expensive, large-scale instrumentation. Most natural sciences generate a type of knowledge that is quickly superseded by more recent findings, while scholarly knowledge can retain its relevance to disciplinary audiences for a potentially much longer time. The concept of hinterlands by Law (2004) can be used to theorize in more detail those aspects of an infrastructure that actors draw on and reproduce when they generate new knowledge contributions, i.e. its methodological and epistemic foundations. Law conceptualizes method not as an objective, context-independent protocol that allows to extract hidden meaning out of social realities, but rather as a generative device that selectively amplifies aspects of a research object according to particular epistemic conventions. This emphasizes on the one hand the disciplinary criteria that individual knowledge contributions must resonate with to be considered adequate. It also draws attention to the fact that some contributions are easier to generate than others because they mobilize gradually standardized sequences of articulation work. Articulation work denotes the situated effort necessary to align principally disparate material and semiotic elements in the process of developing an argument (Fujimura,

1987; 1992). It includes for example the collection and curation of snippets of empirical material, but also presenting and defending research at conferences or in negotiations with anonymous journal referees. The practical conduct of articulation work is facilitated if one sticks to established routines – established data formats for example suggest particular ways of ordering and analyzing empirical material, and adherence to methodological conventions reduces the complexity involved in negotiating the closure of particular research problems. In their totality, these packaged routines constitute what Law calls the hinterland of a discipline.

As a guiding metaphor, I have used the image of scholarly infrastructure as a knowledge-producing machine. The metaphor is not meant to suggest a singular monolithic entity, but rather to highlight the historical interconnectedness and complexity of its internal mechanics. Taking inspiration from Edwards et al.'s (2011) complementary image of communication as a lubricant that facilitates data-sharing across geographical and disciplinary distances, I suggest that the machine functions smoothly only when its many moving parts are properly configured and compatible with each other – i.e., evolving, but shared epistemic frameworks, a workable division of labor between different practices, as well as standards and protocols that complement established routines. Digital scholarship, however, presents more or less substantial challenges to this infrastructural balance. It entails new possibilities for collecting and analyzing material, new ways of raising and defending knowledge claims, and uncommon requirements regarding the organization of the research process. If scholarship can be thought of as a complex knowledge-generating machine, the move to digital scholarship constitutes an attempt to swap some of its parts while the machine is running.

Summary of empirical findings

My first research question has been geared to investigate the implications of this view for the mutual shaping of digital research technology and current scholarly practices. *What does it mean to think of scholarship as an inert infrastructure*, I have asked, *and how does this inertia shape the embedding of new research tools in scholarly practice?* In the first exploratory chapter I have investigated the controversial digitization of the bibliography of Dutch Studies (BNTL). Edited for many years by the Huygens Institute of the Royal Netherlands Academy of Arts and Sciences, the use of the BNTL has

traditionally been part of the curriculum of Dutch Studies and thoroughly part of disciplinary practices. Scholars use the bibliography to find sources and delineate a body of relevant scholarly knowledge. In several ways, the bibliographical perspective informs the conceptual deep structure of literary scholarship, for example by continuously performing categories such as 'author' and 'oeuvre'. The BNTL thus can be considered to be part of the infrastructure for scholars in Dutch Studies. The digitization transformed the bibliography from a physical book to be consulted in libraries into an online database. An important change was the managerial decision to save cost by downsizing the editorial team. As a result, the bibliography now covers a smaller number of publications overall – predominantly well-known journals in literary studies and linguistics, but fewer monographs and edited volumes. On the other hand, users now have the possibility to add publications themselves, in case they are not automatically covered. The designers of the digitization plan moreover emphasized the benefits of immediate, automatic updating of the dataset, as well as the possibility to inspect abstracts and full-texts of individual article entries. My empirical investigation revealed that specific ways of asking questions and organizing empirical work had a strong influence on how the digitization was perceived by BNTL users. In the areas of textual criticism and analytical bibliography for example, the more comprehensive, editorially warranted coverage of the old BNTL was an important epistemic precondition. Rather than an overall 'improvement' of the bibliography as suggested by the designers, the digitization constituted a rather significant infrastructural disruption for these research practices.

Such tensions, one might argue, are characteristic of contexts where science administrators and policy makers take important design choices in a top-down fashion. However, my analysis of the COST Action Women Writers in History (Chapter 2), in which technological design is thoroughly grounded in the research interests of scholars, has highlighted problems of a similar nature. The participants had built a career around studying the activities of women writers between the 16th and late 19th centuries, thus complementing more traditional historical accounts that focus on a limited number of heroic male figures. The collaborative use of an online database seemed like a great opportunity to align individual research efforts and so create a more comprehensive empirical picture of 'forgotten' literary history. In spite of similar disciplinary backgrounds and a shared interest in digital technology, the participating scholars encountered significant difficulty when trying to embed the database into their monograph-oriented work

routines. It firstly required integration of very diverse research questions and methods, as well as bringing diverging theoretical perspectives into agreement. Moreover, although the participants had learned to be skeptical about the impression of 'completeness' that is conveyed by traditional bibliographies and literary histories, they nevertheless expected the database to provide a body of reliable knowledge that could be used as a basis for monograph and journal publications. After all, they expected their involvement in the project to help them advance their disciplinary careers. Yet the collaborative model of the COST Action implied a very different relation between scholarly users and research technology. Rather than a readily usable source of knowledge, it required considerable labor input to curate and harmonize the data. The project leader tried to promote a view of the database as not simply the digital equivalent of traditional bibliography, but as a catalyst to stimulate new intellectual approaches which would not necessarily be organized around idiosyncratic monograph-narratives, albeit with limited success.

Equipping tools with better accessibility and participatory features, or designing them in a way that allows for the potential use of substantial datasets and computational power is thus not necessarily perceived as an enhancement by scholars. Instead, insofar as new features are often incongruent with established organizational models of scholarly work, their dedicated use may actually amount to what Bowker & Star (2000) call a breakdown of infrastructure. The digitized bibliography of Dutch Studies for example could be seen as an always up to date, easily accessible database with Web 2.0 functionality. Instead, some established scholars saw it as a detriment for their work, since it failed to provide the reliability and representational function their research depended on. Similarly, the digital database used in the COST Action initially seemed to offer a way of creating exciting empirical insights about forgotten women writers that were impossible to achieve by any single scholar. Quickly, however, participation in the project also began to appear as a risk for individual career development, insofar as the database required a way of organizing scholarly practice very different from the single-author, monograph-oriented model the participants normally operated with. Both case studies thus emphasize that the effective value of individual tools will not so much depend on hypothetical capabilities or abstract epistemological benefits, but rather on the extent to which they can be meaningfully integrated into the scholarly knowledge machine without jamming or producing too much friction. Initiatives that simply ignore this aspect, perhaps carried away by promises

made to create momentum for a given project, or to legitimize predetermined policy decisions, risk a backlash later on. This should not be mistaken for unqualified conservatism or a willful resistance to new technology on the part of the humanities (cf. ACLS, 2006; Wouters, 2007). While the slowly developing configuration of academic practices, as reproduced through undergraduate training and disciplinary reward systems, often appears as a deplorable inflexibility from the vantage point of technological innovators, it also guarantees that new scholarly work meaningfully connects to the extant body of knowledge, for example in terms of recognizable output format, shared tacit knowledge, as well as reliability of empirical sources. The best response to such infrastructural inertia, I suggest, is to address it upfront and without framing it as either inherently positive or negative. This may also constitute a first step in solving potential conflicts between technological affordances and scholarly practice, an aspect that I will elaborate later on.

The academic labor ecology is characterized by numerous interdependencies between distinct task areas. Such distinctions are instrumental in regulating their interaction, because they determine which group of actors needs which skills so as to accomplish its systemic function. An example is the division of labor between research proper and the work of information professionals, such as librarians and archivists. The latter order information according to established categories, thus providing scholars with bibliographical tools (archives, collections, library catalogues...) that can be readily used for producing new knowledge contributions according to disciplinary conventions. The move to digital scholarship, however, is related to shifts in such interdependent relations, with potentially significant effects on the intellectual substance of research. My second research question has therefore been: *How do actors in digital scholarship construct forms of labor as 'scholarly', 'technical', or 'support activities', and how does the distribution of labor make possible certain forms of knowledge, but not others?* In the case of the digitized bibliography of Dutch Studies, a key point of contention was the decision of reducing the editorial team and instead introduce a feature that allows users to individually add publications. This design choice transferred responsibility for bibliographical work from professional documentalists to the scholars who normally rely on the BNTL as a trustworthy, ready-made index of disciplinary knowledge. The ensuing controversy focused inter alia on how the task of guaranteeing the reliability of the bibliography should be distributed – should it be something that the Royal Academy takes responsibility for, or can it legitimately be seen to overlap with the core tasks

of scholars working at universities? To the many critics of the digitization, the incorporation of Web 2.0 features became synonymous with an official acknowledgement that the Royal Academy no longer considered the curation of national cultural heritage as one of its primary responsibilities.

The case of the COST Action further illustrates the intellectual implications of such issues. A core obstacle encountered by the participants was the need to contribute significant amounts of work to enter information into the database, as well as double-check and harmonize the growing dataset. Such work was necessary before participants could start making large-scale, comparative empirical knowledge claims about forgotten women writers. As I have argued, this obstacle presented itself as a recursive resource problem to the project leader: to make the project more attractive to the participants and secure follow-up funding, it was necessary to offer access to as much reliable data as possible, but in order to do so, a significant expense of human effort was required. However, many participants – usually employed as professors or faculty at university departments – tended to consider data work a subordinate technical activity, similar to bibliographic work, that distracted them from their many 'actual' tasks such as research and managerial responsibilities. The definition of data work as a non-intellectual activity was also a common argument to justify its delegation to student assistants, who were for example encouraged to combine their MA or PhD research with data curation in the project. But in contrast to bibliographic work and the monograph-oriented research model it enables, skills and intellectual principles for data-driven historical research are not very far developed among scholars. The only way to establish them is through actually experimenting with databases, algorithmic approaches and visualization tools in a comprehensive way, i.e. delve into the nitty-gritty of their functionalities and limitations. If the step of doing seemingly subordinate data work is delegated, scholars foreclose an opportunity to embed new technology in their research practices and perhaps work towards new conceptual approaches. The COST Action indeed wrestled with a problem of unequally distributed knowledge until its very end. While several student assistants developed meaningful research questions through getting their hands dirty in data work, most of the academically more advanced participants had not achieved such embedding even a few years into the project.

A first conclusion to be drawn from this, casually put, is that big data needs big data work. This aspect is often underestimated in the current hype surrounding the few actually well-funded digitization projects such as

Google Books. Also aside from digitization of sources proper, and in case data are openly available, a significant amount of work is typically necessary to make data compatible, both technically, conceptually, and in terms of their reliability. If this aspect continues to be absent from portrayals of future benefits of digital scholarship, it may result in a certain disappointment on the part of both funders and practitioners. A second conclusion is that the designation of activities necessary to make data usable as either a 'proper' intellectual task or a subordinate technical one is itself often used in a tactical way by actors. Scholars can refer to some kinds of work as technical because it is very different from the kind of activity they are normally rewarded for, and policy makers in turn can refer to data work as scholarly in order to justify reduction of public expenditure. Such dynamics are particularly consequential insofar as the construction of certain forms of labor is directly connected to reproducing or challenging the basic research skills and conceptual deep structure of dominant scholarly paradigms. The bibliographical categories that order the production of documents in literature, public life, and science are considered a subordinate technicality because they are widely accepted. There are many alternative or complementary possibilities for ordering written production that we still have difficulty to imagine, however, simply because the infrastructural conditions to develop them in a concerted fashion are not in place.

These findings, I would argue, warrant reframing the key question in current debates about the scholarly use of data-intensive research methods. At the moment, the question is often formulated like this: how will our understanding of art, history, and social life change if we harness much greater amounts of empirical information than was common in the past? The problem with this formulation is that it locates intellectual agency primarily in the act of analyzing data, and that it downplays the many forms of activity necessary to reproduce disciplinary methods and make data usable. My findings suggest that it would be desirable to rephrase the question: how does our understanding of art, history, and social life change if we begin to *begin to prepare* large amounts of data for interrogation? Such a formulation has a double advantage. It firstly draws attention the conceptual choices made by otherwise invisible data workers, and it emphasizes the effort necessary to mutually adapt scholarly practice and data to each other. Secondly, it encourages a realistic assessment of the investment required to make actual scholarly use of big data, rather than picturing it as an essentially free, added value that somehow arises from 'informatization'.

In summarizing my findings so far, I have emphasized how

infrastructural constraints affect the shaping of new tools, but I have not yet explicitly elaborated the extent to which these are distinctly disciplinary phenomena. Such knowledge is critical because it would give us a sense of the tailoring required to adapt digital tools and infrastructure to the needs of specific domains, and also of how such tailoring could proceed. As a third research question I have asked: *What role does disciplinarity play in the shaping of digital research tools, and how does the use of these tools affect disciplinarity?* Again, the case of the digitized BNTL can serve as a first illustration of what is at stake. The reduction of coverage that accompanied the transformation of the bibliography into an online database concerned, on the one hand, monographs in Dutch literary studies and literary history, i.e. scholarly contributions typically written in the national language and about topics primarily relevant to specialists of Dutch and Flemish literary history, which are not elsewhere indexed in a comprehensive fashion. Moreover, the initial digitization plan meant to completely exclude journals in modern Dutch linguistics, because they are already relatively well covered in other, international, databases. The subsequent discussions show that scholars consistently perceived the digitization to reflect strategic managerial choices about how to position the field of Dutch Studies internationally – critics associated the transformation of the bibliography into a database with reduced coverage and less funds as a testament that the Academy no longer considers the conservation and mediation of knowledge about Dutch and Flemish literature an important scholarly task, and defenders argued that Dutch Studies have to look for a different, more linguistically (and less hermeneutically) oriented international audience anyway.

Chapter 2 shows how the disciplinary configuration of literary history, characterized by low degrees of intellectual and social integration, poses specific obstacles for the originally envisioned usage of the database in the COST Action. Literary historians normally have relatively large leeway in framing their particular research objects, and they can choose a theoretical framework out of a range of coexisting options. This in turn made it very difficult to reconcile differences across individual practices in the framework of the project, even though there was a strong agreement on the basic research goal, as well as shared assumptions such as distrust regarding traditional bibliographies and historical accounts. Integrating individual practices in fact required participants to reach temporary consensus on how to define complex concepts such as genre and reception - theoretical problems that are normally subject to continuous debate among scholars in literary history. In the context of the project, however, these had to be

translated into pragmatic matters for the sake of advancing the collaborative undertaking. A particularly interesting development was the fact that project work could be more easily reconciled with individual career logic in the cases of smaller sub-projects, for example an initiative by a group of Serbian scholars loosely affiliated to the COST Action. The smaller number of participants and a more circumscribed empirical focus made it easier to agree on theoretical assumptions, as well as keeping the sheer amount of data work to a manageable level. This again reduced the need for a strict division of labor between professors and data workers, thus creating more favorable conditions for digital skill dissemination.

Disciplinarity also proved an important constraint in the collaborative project investigated in Chapter 3. Here, a small group of indonesianists, network researchers, and computer scientists tried to harness quantitative analysis of digitized newspaper corpora to better understand shifts in Indonesian elite networks. A key point of contention was the problematic original assumption that the open-ended hermeneutic research problems of the indonesianists (e.g., who and what are social elites, and how do they exert power over other groups?) could be 'solved' by a network analysis of elite actors based on their co-occurrence in newspaper articles. This suited an initially assumed division of labor between the participants, according to which computer scientists were positioned as the producers and conceptual definers of the data, while the humanities scholars were charged with interpreting them. The rationale for this arrangement was to ensure that all participants could contribute to the project but still get something out of it that would advance their individual careers. The proposed collaborative structure, however, turned out to pose a problem for the indonesianists, since disciplinary peers found such a reformulation of hermeneutic questions as an empirical problem for network analysis less than compelling when measured against the conceptual sophistication of dominant disciplinary theorizing. Therefore, the project began to experiment with alternative arrangements in which certain practices did not simply colonize others by imposing their respective assumptions – for example by presupposing that the concept of data current in network research and computer science can simply be transposed onto Indonesian Studies. Rather than using a specific type of quantitative analysis as the ultimate epistemic arbiter, the new goal was to develop an interdisciplinary division of labor in which multiple conceptualizations of the research object of elites could coexist.

Summarizing these results, it is fair to say that early or first time

efforts in digital scholarship are frequently characterized by a conflict between disciplinary expectations towards what a 'proper' knowledge contribution must look like (for example regarding methods, theoretical orientation, or output format) and the initially proposed use of a particular research tool. Digital project frameworks challenge the control over allocation of resources wielded by disciplinary authorities, thus buying the participants a degree of freedom to experiment with approaches deemed unconventional in their field. If they find ways of attracting a steady amount of funding, it is possible that such unconventional practices develop and persist, perhaps cutting ties with the original discipline altogether. Frequently, however, scholars will not be willing to give up their affiliation to traditional disciplinary contexts. In such cases, conflicts between digital affordances and domain-specific expectations can be productive if participants are able and willing to tackle certain ideological, technical, or economic constraints, for example the belief that digital scholarship *should* be about translating 'messy' hermeneutic problems into 'exact' empirical ones, or the fact that some data formats are easier to handle because they are widely accepted. An important strategy is also the adjustment of organizational and funding modalities. Small-scale projects for example can create a relatively circumscribed context for experimenting with novel research tools (i.e., manageable amounts of data work, not too much divergence of theoretical viewpoints among participants), thus creating a niche in which digital practices are easier to reconcile with disciplinary conventions than in larger-scale undertakings. An important general conclusion for future practitioners and policy makers is that the development of digital approaches is best conceptualized as a situated activity – rather than resulting from the implementation of rigid technical or collaborative designs, perhaps conceived by individuals outside the actual projects, viable arrangements are more likely to emerge through practice and over time.

While much of the above research highlights the important role of disciplinary, praxeological, and economic factors in shaping the digital humanities, this does not mean that the configuration of new scholarly practices is merely the mechanical result of intersecting constraints. The way digital scholars tackle conflicts when trying to bring unconventional research to closure is often characterized by deliberate agency, i.e. choices in favor of certain solutions and against others. This dynamics is an important analytical object not only because it shows how actors actively shape emerging technology, but also because it may itself contribute to reshaping

infrastructural constraints over time. I have asked as a fourth research question: *What is the reflexive agency of scholars in the embedding of new tools into their infrastructural work setting?* As suggested in the above, a process of reflexive adaptation was critical to developing a workable division of labor in the early phase of the Elite Network Shifts project. The original collaborative arrangement, according to which a co-occurrence-based dataset created by the computer scientists and network researchers should serve as the basis for 'solving' the hermeneutic problems of the indonesianists led to considerable theoretical criticism by scholars in the community of Asian Studies. The participants subsequently initiated a reflexive discourse, aided by myself as an STS analyst, in which the conceptual and practical ramifications of certain foundational assumptions (e.g., that data is a discipline-agnostic phenomenon) were questioned. While not magically making tensions disappear, this did open up the possibility to imagine alternative collaborative modalities that may well provide a model for how the participants will tackle similar problems in potential follow-up projects.

In Chapter 4, I have focused on the strikingly reflexive style that characterizes many publications circulated by digital scholars, both traditional peer-reviewed ones as well as digital formats. These publications draw attention to the practical conditions under which digital scholarship is conducted, typically by combining research arguments with a discussion of the tensions encountered in an academic environment that favors single-author, monograph-oriented scholarship. I have argued that by making visible aspects of practice that are not routinely addressed in formal scholarly communication, discursive reflexivity selectively 'unties' the standardized packages that together constitute the hinterlands of the humanities. For example, one set of documents grants a look behind the scenes of digital project work to provide practical advice for new entrants (THATCampCHNM, 2011). Topics include strategies for lobbying with funding bodies and partner institutions, but also management techniques for collaborative interdisciplinary work. Frequently, such advice includes a polemical critique of outdated humanistic curricula and inflexible institutional structures (Scheinfeldt, 2011), thus putting up for discussion the basic skills that scholars must possess in the early 21st century. Others again argue that such calls for digital humanists to be more 'realistic' about modern academic life encourages a pragmatic, managerial self-perception that is at odds with the critical function and hermeneutic sensibility of the humanities (Chun, 2013). More specifically, a number of practitioners

(Drucker, 2009; Hansen, 2012) problematize how digital project work is streamlined through reliance on standardized data formats. While useful to facilitate data sharing and reduce the complexity of project work, this tends to reify foundational conceptual assumptions about particular research objects, as well as making scholars dependent on often commercial software products. Ultimately, this form of reflexivity encourages a rather fundamental debate about the function of empirical arguments in digital scholarship, and the extent to which practitioners should sacrifice critical intentions in order to 'get stuff done'.

These empirical results affirm that it is not adequate to think of the development of digital scholarship simply as a diffusion of inherently defined tools into existing research practices, as the more hyperbolic accounts of a digital revolution would have it. Instead, digital scholarship develops through the creative uptake and adaptation of technological possibilities. When practitioners design and use new instruments, they make choices that constitute specific reactions to perceived infrastructural constraints. Individual solutions may in turn crystallize into new conventions for how to conduct digitally mediated research in the future – for example, new standards for how to conceptualize and exchange data, or widely used 'best practices' for dividing labor between collaborators from different disciplines. Reflexivity thus plays a key role in the evolutionary development of the scholarly infrastructure. At the same time, it is important to appreciate the function of human agency in this process. As I have shown, individual speakers propose different ways of framing and resolving tensions, and sometimes these are mutually exclusive. Discursive reflexivity in the digital humanities is consequently also a site of political controversy, where distinct ideas about how to reorganize the scholarly knowledge machine clash.

Most of the findings I have presented so far empirically focus on the level of specific scholarly practices or projects. However, this sidesteps how the economic, organizational, and political conditions for individual projects may themselves be shaped by longer-term, strategic initiatives to create digital infrastructure, which after all constitute an important reason for the current policy interest in digital humanities. My fifth research question is meant to shed light on this interaction from a comparative perspective: *How is infrastructure conceptualized differently across countries, and what role do such conceptualizations play in organizing infrastructure development 'on the ground'?* In my last chapter, I have argued that current infrastructure initiatives reconfigure the organization of disciplinary tool development, thus affecting

also the methodological structure of digitally mediated scholarship in the long run. In traditional modes of research, choices about tool development were largely reserved for disciplinary authorities (Whitley, 2000). Many national science systems, however, have recently entered a process of reconfiguration that challenges the mechanisms of disciplinary self-governance (Nowotny, Scott & Gibbons, 2001; Etzkowitz & Leydesdorff, 2000; Whitley, 2010). Current infrastructure initiatives illustrate this development in that they are based on strategic alliances between policy makers, funders, and researchers. Pursuing often heterogeneous interests, they are nevertheless held together by shared visions of future digital infrastructure. The infrastructure discourse, I have therefore argued, can be seen as an interface between research and policy in a historical period in which their relations have become newly volatile. Insofar as different definitions of networked digital instrumentation also imply specific ways of funding and coordinating tool development, it becomes of strategic importance for individual actors to promote a technological vision that suits their respective priorities.

In Europe, an important actor in infrastructure policy is the European Strategy Forum for Research Infrastructures (ESFRI), a supranational body constituted through an agreement of the European Council of Ministers and the European Commission in 2002. ESFRI stresses the need for centralized coordination of individual development activities, an approach that resonates with the longstanding policy goal to overcome the fragmentation of European research into national science systems. The various projects administrated by ESFRI are grounded in previous, nationally based efforts in humanities computing, but in order to ensure continued funding and political support, they now have to translate their activities into terms compatible with the European Commission's policy agenda. Participating initiatives have to present strategies for limiting the diversity of scholarly approaches supported through digital tools, so as to demonstrate commitment to the overarching goal of de-fragmentation. At the same time, ESFRI pictures infrastructure as something that can be developed in a series of clearly circumscribed projects, i.e. in an organizational context that is principally detached from the primary research process of the prospective users. This approach inevitably produces a certain disconnect between tool developers and the large majority of scholars who are unfamiliar with digital approaches. On the other hand, the emphasis on central coordination also ensures that no single disciplinary community gains exclusive control over technology development. The

situation in the US is different from the European context insofar as there has historically been a larger diversity of different funding sources, public, private, and philanthropic. Although these grant relatively small sums individually, such diversity has made digital humanities centers somewhat independent from the mandate of any single funding body, thus conferring them also a certain intellectual freedom. US digital scholars have rather successfully tried to protect and extend this independence by promoting an explicitly STS-informed definition of infrastructure, which stresses praxeological embedding of tools and the emergent character of technological development. Infrastructure here is not conceptualized as a specific technological artifact to be created in a singular project, but as the skills, tools, and facilities that gradually emerge from ongoing scholarship in digital humanities centers. While avoiding tensions related to the implementation of new tools in more traditional scholarly practices, this arrangement can be criticized for privileging the intellectual and technological choices of existing elites within digital humanities over that of other research communities.

The development of digital infrastructure can thus be seen as a particular example of the reflexive dynamics discussed under the previous heading. Rather than a neutral enhancement of scholarly practice, actors promote definitions of infrastructure informed by normative assumptions about science as well as strategic interests, for example the notion of a highly integrated European Research Area, or that of a network of relatively autonomous digital humanities centers. Once instantiated in funding frameworks and coordination mechanisms, these visions create specific conditions for local efforts in digital scholarship, thus reshaping existing organizational structures of research in their image.

Theoretical and practical implications

Scholarship, then, is in a flux, occasioned by a combination of widely available digital tools and facilities, a public discourse about the possibilities of digital scholarship, and concomitant investments in infrastructure. The contribution of this thesis has been to intersect these developments with a theoretically informed view of how the humanities have developed socially and intellectually into their present configuration. My conceptual influences, derived from various theoretical currents within STS, replace a commonsensical view of infrastructure and tools as characterized by inherent qualities with a socio-technical one, in which technology cannot be

thought of separately from the historically developed work routines in which it is embedded. Local activities always build on an 'installed base', i.e. a more slowly developing foundation of practices and institutional structures that I have characterized in more detail by drawing on Whitley (2000) and Law (2004). Against this background, the recently popular notion of a scientific revolution induced by big data appears to rely on questionable assumptions. More data/more computational power/more collaboration does not straightforwardly result in more or better knowledge, but also in more fragmentation and new conflicts between the status quo and digital affordances. Information, when drawn from different contexts, is unlikely to simply add up to a coherent empirical picture, and trying to enforce an unqualified notion of digital collaboration in a project may actually result in undermining existing collaborative structures. I would venture the following generalized conclusion: given the plurality of theory and methods in the humanities, initiatives that combine ambitious collaborative scale with ambitious praxeological changes will face particular challenges. The changes in practice they demand force a sizeable number of practitioners into opposition to the disciplinary status quo, thus cutting not only certain epistemic ties that depend on compatibility of infrastructural task areas, but also social ties to peer audiences and disciplinary funding and career possibilities. Put differently, the relatively weak degree of intellectual and social integration in the humanities means that practices are overall less standardized than in most natural sciences, thus making it more difficult to identify a substantial group of researchers that can agree on black-boxing certain sequences through a radically new piece of technology or a new organizational format.

To highlight the effort necessary to articulate disparate practices and technological affordances in such a pluralistic field I have used the metaphor of the humanities as a complex, knowledge-generating machine. This image draws attention to the need for reflexive attention to disciplinary history and situatedness when developing its internal mechanics, so as to ensure that engaging in digital scholarship does not create unmanageable friction with the installed base. This can for example mean to make the function of such a fundamental epistemological concept as data a topic for discussion in a project, and to change the role data are given in a collaborative workflow. Reflexivity can also mean rethinking the organizational modalities and goals of a project. For example, downscaling the epistemic scope and sheer number of collaborators in a given project will make it easier to reconcile conflicting demands on scholarly work hours and to reach agreements on

analytical goals. However, even if we acknowledge the need for applied reflexivity in a period of major socio-material reconfiguration, we should be aware that reflexivity can never be neutral. Local solutions to friction may in their own right crystallize into a new status quo over time, thus reshaping the installed base also for other actors.

In the coming years, we are likely to see numerous proposals for solving conflicts and developing the material and social environment of scholarly inquiry, and some of them will become part of the disciplinary hinterlands of future generations. While the exact ways in which such reconfiguration will occur are a problem for further empirical study, my findings do allow to summarize a few critical spots in the scholarly knowledge machine that are bound to play a particularly important role in the process. These can perhaps be imagined as clutches that transmit momentum between moving parts, thus making them privileged points for intervention. The first aspect is the conceptualization and distribution of different forms of labor in the scholarly work process, which I have argued to be instrumental in reproducing the basic conceptual structure and methodological foundation of a field. The reason why it is currently difficult to combine digital scholarly approaches with more traditional ones is a structural conflict between contradicting incentives: the specific acts entailed by digital scholarship are not part of disciplinary training and look too different from what is necessary to advance a traditional academic career. It is possible, however, that the increasing availability of digital tools, combined with ambitious outreach activities and general dissemination of digital skills through other channels, will further reduce that effort that individual scholars have to make to acquire basic skills in the use of databases, data work, and coding. Change in the notion of what constitutes proper intellectual work may be additionally stimulated as younger generations of researchers enter the professoriate. A more negative scenario is equally imaginable, however, in which the increasingly casualized employment of younger academics further reinforce the hierarchical distinction of 'technical' and 'intellectual' activities. Precisely because employment in dedicated research positions is becoming scarcer, the skill set and credentials necessary to achieve it could become even more specialized (cf. Whitley, 2010).

Second, the notion of authorship, which is a critical element in the knowledge machine because it mediates between disciplinary reputation dynamics, employment, and the scholarly publishing industry. Authorship selectively makes visible certain contributions while excluding others - in the

humanities, it has traditionally been equated with publishing monographs. At the same time, conventional definitions of authorship are being challenged through both formal university policies and emerging research practices. According to its 2008 PhD regulations, Leiden University (2008) for example principally accepts PhD dissertations jointly submitted by up to three authors, under the condition that individual contributions are demonstrable and significant. This is arguably the institutional response to the predominance of multi-authored papers in the natural sciences, which thereby become a formal possibility also in the humanities and social sciences. Another emerging trend is a broadening of the notion of authorship to encompass output that is not documents. In many quantitative fields like network research or scientometrics, it has become relatively common to cite technological instruments that bear a particular mark of individual creativity, such as algorithms or specialized software. Complementary developments in digital scholarship are the proliferation of journals that publish digital artifacts.[58] Again, however, we can observe a contradictory trend that has to do with the increasing scarcity of employment in disciplinary institutions. Many prestigious, competitive grant schemes on a national and European level continue to be centered on individual (rather than collaborative) research performance, and application modalities frequently seem to favor single-authored publications in journals or as monographs.

Third, increasing use of digital tools is related to changes in what we could call scholarly transparency practices. STS scholars (Latour & Woolgar, 1979; Knorr Cetina, 1981; Law, 2004) have often pointed out that modern scientific authority partly rests on the convention of deleting the messiness and contingency of the underlying work, i.e. the failed experiments, negotiation and power play among scientists, the effort necessary to translate instrument readings into authoritative statements etc. Traditional laboratory-based science provides specific conditions for accomplishing such purification (Latour, 1993). Most importantly, it takes place in a confined physical space that is accessible only to certified members of an expert community. Digital scholarship, by contrast, leaves more visible traces of the research process, for example through the use of social media, metadata created by tools such as Mendeley, or through applications that take part of the research and publishing process online (e.g., open peer review platforms such as mediaCommons). Such practices have a host of

58 Examples include *Vectors* and the *Journal of Digital Humanities*

advantages. To name but a few, they allow for near real-time discussion between authors and readers of digital publications, and they contribute to disseminating practical knowledge on how to do digital scholarship also among traditionally trained academics. Yet we should not assume that they make research 'visible' in any uncomplicated sense - actors will instead develop new strategies of selectively showing some aspects of their work but not others. This may in turn shape conventions of articulation work for future generations, for example by setting up new informal requirements for presenting stylized images of an unfolding research process in a perpetually uncertain, grant-based funding system.

Fourth, many public science systems have been experimenting with formal research evaluation mechanisms for some time now, since these seem to provide a straightforward way of steering research activities in particular directions (Whitley, Gläser & Engwall, 2010). A particularly pertinent, recent development is the attempt to tailor evaluation modalities to the diverse functions of science and scholarship, such as engagement with non-academic audiences, and the development of reusable software, databases, and other forms of digital output (American Academy of Arts & Sciences, 2014; ESF, 2011). However, expectations towards evaluation as a research policy tool are often based on insufficiently complex understandings of the relevant underlying dynamics. The British Research Assessment Exercise, one of the most radical examples of a regulatory intervention into disciplinary self-governance, was set up to ensure performance-based resource allocation, but also diversity and equality in research. Effectively, it has resulted in concentrating resources in a few elite institutions, the demotion of teaching as a task of universities, and the emergence of a transfer market for highly cited academics (Martin & Whitley, 2010; Mirowski & Sent, 2008). While it is very hard to predict how exactly current attempts at reforming research evaluation will interfere with disciplinary dynamics in the near future, it seems evident that they will have implications also for the further development of digital humanities.

Lastly, the adoption of digital approaches in the humanities will continue to be affected by new funding structures and ongoing development of digital infrastructure. A historically important way of tackling the incongruence between digital approaches and an academic environment configured for more traditional forms of research has been the establishment of centers for digital or interdisciplinary scholarship, i.e. organizational formats that exist partially outside the disciplinary landscape (Unsworth, 2007; Clement & Reside, 2011). Creating and sustaining such institutions

becomes easier as the primarily disciplinary control over academic employment is challenged by funding opportunities specifically meant to encourage digital scholarship (see for example Williford & Henry, 2012; BMBF, 2013). Concerted efforts to create digital infrastructure fulfill a similar function. Respective initiatives in Europe and the US come with specific strategies for coordinating and funding tool development, thus interfering with the disciplinary logic that used to inform choices about research instrumentation. I have suggested that especially ambitious infrastructure projects that aim to cover a large variety of approaches will generate friction due to the divergence of local practices and the intended use of individual tools. Nevertheless, given enough time and funding, they may still manage to link up with institutional reproduction of scholarly methods over time, for example if certain applications attract enough disciplinary attention to be incorporated in undergraduate methodology classes.

All of these aspects warrant broadly inclusive discussion. The latter two, evaluation on the one hand, and funding and development of infrastructure on the other, raise the additional question of participation in formal decision making processes. However, the design of research evaluation protocols on a national or university level is usually not put up for debate on a wider basis, almost as if the notion of excellence in science and (digital) scholarship were self-explanatory. Similarly, current infrastructure visions are usually presented as inherently desirable, with little explicit deliberation of underlying normative choices. Many European decision makers see digital infrastructure as a means to effectuate an integration of national science systems. In the US on the other hand, control over tool development is primarily located at established digital humanities centers, thus privileging the design choices of reputed digital scholars. Largely absent from the discourse is the majority of traditional humanists who are the prospective main users of the new technology. This is not to suggest that a scenario is possible in which all affected parties will be equally happy with the resulting infrastructure plans, and even the most democratically developed strategy may result in unintended effects (cf. Jensen & Winthereik, 2013). Nevertheless, taking inspiration from anticipatory governance models, as for example applied in nanotechnology (Barben et al., 2008), could be useful for working towards a more inclusive arrangement in which prospective beneficiaries of infrastructure are consulted before actual development activities are initiated.

While questions of participation and distributional justice are regularly addressed in STS discussions of infrastructure development

(Edwards et al., 2013), there is typically little attention for how they relate to the specificity and function of humanities knowledge in society. Few observers would disagree that it is desirable for, say, cancer research to be highly integrated intellectually and socially, so as to concentrate investment on a few promising lines of very resource-intensive inquiry. Whether such integration is desirable in the humanities is not so clear. Scholarly knowledge is characterized by and thrives upon the diversity of coexisting intellectual views, and it would not be difficult to argue that such diversity increases the import of the humanities for the rest of society. To what extent should evaluation modalities and longer term investment in digital scholarship be characterized by mechanisms that ensure plurality? This question is worth raising in the context of a possible model of anticipatory governance, since a frequently taken-for-granted design principle of infrastructure is an implicit, unqualified notion of efficiency. While some aspects of infrastructure can and should indeed be designed with efficiency considerations in mind, in other respects, for example methodology and theoretical orientation, such a rationale could result in undermining characteristic virtues of the humanities.

Directions for further research

A first way of further developing the above findings would be to expand significantly the conceptual and empirical scope of some of my original research questions. An important topic that I have only begun to touch is the relation between digitally mediated scholarship in specific projects and the shaping of such project work by policy and funding practices. Current developments in digital humanities occur against the background of a profound reconfiguration of public science systems, popularly summarized in notions such as Mode 2 (Nowotny, Scott & Gibbons, 2001) and the Triple Helix (Etzkowitz & Leydesdorff, 2000). As I have variously shown in the above, grant-based, digital research projects can create niches within disciplinary employment structures. However, they also come with new constraints that scholars in turn will try to adapt to their individual intellectual and professional interests. In what sense is the resulting knowledge informed by these shifting economic conditions, and how do scholars react to reconfiguring relations with policy makers and funding bodies? Such research could provide valuable empirical input for rethinking the commonly applied, and somewhat simplistic, dichotomy of bottom-up and top-down actions in the organizational dynamics of science. The two

notions are usually taken to express a separation between contexts of science policy and the work going on on the scientific shop floor. Infrastructure research, however, suggests a much more complex interplay between the two levels (cf. Edwards et al., 2009; Hepsø, Monteiro & Rolland, 2009), thus raising the question as to why we should keep operating with a hierarchical metaphor. One of the reasons why it has persisted so long arguably is the historical convention of STS and Computer Supported Cooperative Work to engage in ethnographic work in laboratories and office-like settings. Opportunities for ethnographies of Political institutions (with a capital P) on the other hand simply have no disciplinary tradition, and are perhaps also harder to come by. Yet it is exactly the interaction of the two levels that would currently seem to be of particular analytical interest.

Increased empirical scope could also be useful for developing the research questions underlying Chapters 2 and 3, where I study how friction between epistemic perspectives is dealt with in two specific projects. In the humanities, many basic differences between paradigms and theories are never resolved, but instead create an ideally fruitful, intellectual tension. In her influential study of medical practices for the diagnosis and treatment of atherosclerosis, Mol (2002) observes a corresponding multiplicity of ways in which this disease is enacted in different parts of a hospital. Mol in fact argues that various techniques, for example surgical intervention to clear clogged-up arteries, or walking therapy aiming to improve blood flow through physical exercise, bring into being specific ontologies of atherosclerosis, which are interrelated but do not coincide. Friction between them is managed not least by distributing forms of enactment physically across different wards with distinct specializations. This parallels how multiplicity of perspectives is managed in the humanities – scholars can adopt diverging viewpoints on the 'same' object by using the individualistic format of the monograph and by operating in more or less contained national or regional intellectual contexts (often delimited as a language community). It is an open empirical question how the spread of networked research and publication practices, for example the use of databases and augmented journals, will affect the management of multiplicity on a larger scale. More fundamentally, an investigation along such lines would require probing the use of Mol's theoretical work for analyzing knowledge production in the humanities. Can scholarly research practices indeed be seen to enact ontologies? Do historians who write monographs actually create different realities than scholars who apply data-intensive analytical methods, or do they create different representations of the same objects? Is

such a general distinction between ontology and epistemology possible and useful to describe scholarly knowledge production (cf. Lynch, 2014; Aspers, 2014)?

Another fruitful direction for future research would be to intersect the perspective of infrastructure studies with the sociology of expectations. The latter has theorized expectations as future-oriented networks (Borup et al., 2006; van Lente, 2000), i.e. as creating protected niches in which actors can experiment with new practices. It would seem intuitive to integrate this perspective more systematically with the perspective of infrastructure studies, which tries to identify networks that have developed historically. A privileged empirical entry point for such a study could be the widely discussed topic of sustainability of digital applications and infrastructure (ESF, 2011; Berman et al., 2010). Sustainability here denotes strategies for ensuring the continuous accessibility and functioning of digital resources. In a sense, this constitutes a future-oriented complement to my argument about the need for preventing 'jamming' of the scholarly knowledge machine through reflexive adaptation. However, it is striking that sustainability is usually considered only in terms of appropriate funding strategies, data and software formats, as well as legal issues relating to intellectual property and data sharing policies (Berman et al., 2010; David, 2005; Lossau & Peters, 2008). While undoubtedly critical aspects, these discussions would seem to benefit from a historically minded perspective that theorizes sustainability also as a matter of reproducing practices, so as to keep knowledge intelligible and compatible across different historical periods. Put differently, we should expand our notion of obsolescence, which is typically framed as either a purely technical problem or an insidious market strategy, to encompass intellectual and epistemic obsolescence, i.e. the failure to meaningfully relate to historical frames of intellectual practice.

References

Written sources and digital artifacts cited

ACLS (2006) *Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences* (New York: ACLS). Available at http://www.acls.org/uploadedFiles/Publications/Programs/Our_Cultural_Commonwealth.pdf (accessed 2 December 2014).

AHRC (2006) *Peer Review and Evaluation of Digital Resources for the Arts and Humanities*. Available at http://www.history.ac.uk/projects/peer-review. (accessed 5 October 2014).

AHRC/JISC (2006) *AHDS Review and User Survey. Invitation to Tender*. Available at http://www.jisc.ac.uk/media/documents/funding/2006/05/ahds_review.pdf (accessed 7 July 2014).

Aiden, E. and Michel, J.-B. (2013) *Uncharted: Big Data as a Lens on Human Culture* (New York: Riverhead Books).

Akrich, M. (1992) The De-scription of Technical Objects, in: W. Bijker and J. Law (Eds) *Shaping Technology/Building Society. Studies in Sociotechnical Change* (Cambridge, MA: The MIT Press, pp. 205-224).

American Academy of Arts & Sciences (2014) *Humanities Indicators*. Available at http://www.humanitiesindicators.org/ (accessed 2 December 2014).

Anderson, C. (2008, June 23) The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, *Wired*. Available at http://www.wired.com/science/discoveries/magazine/16-07/pb_theory (accessed 27 January 2014).

Anderson, S., Blanke, T., and Dunn, S. (2010) Methodological Commons: Arts and Humanities e-Science Fundamentals, *Philosophical Transactions. Series A, Mathematical, Physical and Engineering Sciences* 368(1925), pp. 3779–3796.

Antonijević, S., Wyatt, S., and Dormans, S. (2013) Working in Virtual Knowledge: Affective Labor in Scholarly Communication, in: P. Wouters, A. Beaulieu, A. Scharnhorst and S. Wyatt (Eds) *Virtual*

Knowledge: Experimenting in the Humanities and the Social Sciences, pp. 57–89 (Cambridge, MA: The MIT Press).

Ashmore, M. (1989) *The Reflexive Thesis: Wrighting Sociology of Scientific Knowledge* (Chicago and London: University of Chicago Press).

Aspers, P. (2014) Performing Ontology, *Social Studies of Science* 43(3), pp. 1-5.

Atkins, D., Droegemeier, K.K., Feldman, S.I., Garcia–Molina, H., Klein, M.L., and Messina, P. (2003) *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue–Ribbon Advisory Panel on Cyberinfrastructure* (Washington, D.C.: National Science Foundation). Available at http://www.nsf.gov/cise/sci/reports/atkins.pdf (accessed 12 September 2008).

Atkinson, R., Held, G., and Jeffares, S. (2011) Theories of Discourse and Narrative: What Do They Mean for Governance and Policy?, in: R. Atkinson, G. Terizakis and K. Zimmermann (Eds) *Sustainability in European Environmental Policy: Challenges of Governance and Knowledge*, pp. 115-130 (London: Routledge).

Baars, M., de Bolster, M., Kamp, E., Kelder, J., Land, R., Peeters, F., and van Zundert, J. (2004, October 27) Med: 0410.22 Forse Bezuinigingen op de BNTL, *Neder–L*. Available at http://www.neder-l.nl/bulletin/2004/10/041022.html (accessed 24 April 2009).

Babeu, A., Bamman, D., Cerrato, L., Crane, G., Friedlander, A. et al. (Eds) (2009) *Working Together or Apart: Promoting the Next Generation of Digital Scholarship* (Washington, DC: Council on Library and Information Resources). Available at http://www.clir.org/pubs/reports/pub145/pub145.pdf.

Barben, D., Fisher, E., Selin, C., and Guston, D.H. (2008) Anticipatory Governance of Nanotechnology: Foresight, Engagement, and Integration, in: E. J. Hackett, O. Amsterdamska, M. Lynch and J. Wajcman (Eds) *The Handbook of Science and Technology Studies*, 3rd ed., pp. 979-1000 (Cambridge, MA: The MIT Press).

Barjak, F., Eccles, K., Meyer, E.T., Robinson, S., and Schroeder, R. (2013) The Emerging Governance of E-Infrastructure, *Journal of Computer-Mediated Communication* 18(2), pp. 1–24.

Barjak, F., Lane, J., Poschen, M., Procter, R., and Robinson, S. (2009) Case

Studies of e-Infrastructure Adoption, *Social Science Computer Review* 27(4), pp. 583–600.

Barker, K. (2007) The UK Research Assessment Exercise: The Evolution of a National Research Evaluation System, *Research Evaluation* 16(1), pp. 3–12.

Bartelse, J. Oost, H., and Sonneveld, H. (2007) Doctoral Education in the Netherlands, in: S. Powel and H. Green (Eds) *The Doctorate Worldwide*, pp. 64-76 (Maidenhead, Berkshire: Open University Press).

Barthes, R. (1978) *Image Music Text*, transl. S. Heath (London: Fontana Press).

Bates, D. (2006) *Peer Review and Evaluation of Digital Resources for the Arts and Humanities* (Swindon: AHRC). Available at http://www.history.ac.uk/sites/history.ac.uk/files/Peer_review_report2006.pdf (accessed 7 July 2014).

Beaulieu, A. (2001) Voxels in the brain: Neuroscience, informatics and changing notions of objectivity, *Social Studies of Science* 31(5), pp. 635–680.

Beaulieu, A. (2004) From brainbank to database: The Informational Turn in the Study of the Brain, *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 35(2), pp. 367–390.

Beaulieu, A. (2010) From Co-location to Co-presence: Shifts in the Use of Ethnography for the Study of Knowledge, *Social Studies of Science* 40(3), pp. 453-470.

Beaulieu, A., Scharnhorst, A., and Wouters, P. (2007) Not another case study: A middle-range interrogation of ethnographic case studies in the exploration of e-science, *Science Technology & Human Values* 32(6), pp. 672–692.

Beaulieu, A., and Simakova, E. (2006) Textured connectivity: An ethnographic approach to understanding the timescape of hyperlinks, *Cybermetrics* 10(1), Available at http://www.cindoc.csic.es/cybermetrics/vollOissl.html (accessed 29 Novemer 2014).

Beaulieu, A., and Wouters, P. (2009) e-Research as Intervention, in: N. W. Jankowski (Ed) *eResearch. Transformation in Scholarly Practice*, pp. 54–69 (London: Routledge).

Becher, T., and Trowler, P.R. (2001) *Academic Tribes and Territories: Intellectual Inquiry and the Culture of Disciplines*, 2nd ed. (Buckingham: Open University Press).

Berman, F., Lavoie, B., Ayris, P., Choudhury, G.S., Cohen, E. et al. (2010) *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information* (La Jolla, CA: Blue Ribbon Task Force on Sustainable Digital Preservation and Access). Available at http://brtf.sdsc.edu/publications.html (accessed 5 May 2010).

Biagioli, M. (2002) From book censorship to academic peer review, *Emergences* 21(1), pp. 11–45.

Bijker, W.E., and Pinch, T.J. (1984) The Social Construction of Facts and Artefacts : or How the Sociology of Science and the Sociology of Technology might Benefit Each Other, *Social Studies of Science* 14(3), pp. 399–441.

Bijker, W.E., Hughes, T.P., and Pinch, T.J. (Eds) (1987) *The Social Construction of Technological Systems* (Cambridge, MA: The MIT Press).

Blackwell, C., and Martin, T.R. (2009) Technology, collaboration, and undergraduate research. *Digital Humanities Quarterly* 3(1). Available at: http://www.digitalhumanities.org/dhq/vol/003/1/000024/000024.html.

Bloor, D. (1976) *Knowledge and Social Imagery* (London: Routledge and Kegan Paul).

BMBF (2013) *Forschungsinfrastrukturen für die Geistes- und Sozialwissenschaften* (Bonn: BMBF). Available at http://www.bmbf.de/pub/forschungsinfrastrukturen_geistes_und_sozialwissenschaften.pdf (accessed 1 December 2014).

Boast, R. (2009) Project Bamboo: Opportunities lost, *Rescite. The blog of noise, interference and diversity.* Available at http://rescite.blogspot.co.at/2009/03/project-bamboo-opportunities-lost.html (accessed 7 July 2014).

Bod, R. (2014) *A New History of the Humanities: The Search for Principles and Patterns From Antiquity to the Present* (Oxford: Oxford University Press).

Boonstra, O., Breure, L., and Doorn, P. (2006) *Past, present and future of historical information science* (Amsterdam: Data Archiving and Networked Services). Available at http://www.oapen.org/download?type=document&docid=353255

(accessed 7 July 2014).

Borgman, C.L. (2007) *Scholarship in the Digital Age: Information, Infrastructure, and the Internet* (Cambridge, MA: The MIT Press).

Borgman, C.L. (2009) The Digital Future is Now: A Call to Action for the Humanities. *Digital Humanities Quarterly* 3(4). Available at: www.digitalhumanities.org/dhq/vol/3/4/000077/000077.html.

Borup, M., Brown, N., Konrad, K., and van Lente, H. (2006) The Sociology of Expectations in Science and Technology, *Technology Analysis & Strategic Management* 18(3/4), pp. 285–298.

Bos, N., Zimmerman, A., Olson, J., Yew, J., Yerkier, J., Dahl, E., and Olson, G. (2007) From Shared Databases to Communities of Practice: A Taxonomy of Collaboratories, *Journal of Computer-Mediated Communication* 12(2), pp. 318–338.

Bourdieu, P. (1988) *Homo Academicus*, transl. P. Collier (Cambridge, MA: Polity Press).

Bowker, G.C. (2000) Biodiversity Datadiversity, *Social Studies of Science* 30(5), pp. 643–684.

Bowker, G.C., Baker, K., Millerand, F., and Ribes, D. (2010) Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment, in: J. Hunsinger, L. Klastrup and M. Allen (Eds) *International Handbook of Internet Research*, pp. 97–117 (Dordrecht: Springer Netherlands).

Bowker, G.C., and Star, S.L. (2000) *Sorting Things Out: Classification and Its Consequences*, (Cambridge, MA: The MIT Press).

Breschi, S., and Cusmano, L. (2004) Unveiling the Texture of a European Research Area: Emergence of Oligarchic Networks Under EU Framework Programmes, *International Journal of Technology Management* 27(8), pp. 747-772.

Broughton, J., and Jackson, G.A. (2008) *Bamboo Planning Project. A Proposal to the Andrew W. Mellon Foundation.* Available at http://www.quinndombrowski.com/sites/default/files/blog/bamboo_planning_project_proposal.pdf (accessed 15 July 2014).

Brown, N., Rappert, B., and Webster, A. (Eds) (2000) *Contested Futures: A Sociology of Prospective Techno-Science* (Aldershot: Ashgate).

Bulger, M., Meyer, E.T., de la Flor, G., Terras, M., Wyatt, S., Jirotka, M., Eccles, K., and Madsen, C. (2011) *Reinventing research? Information Practices in the Humanities* (London: Research Information Network). Available at http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/information-use-case-studies-humanities (accessed 1 December 2014).

Busa, R. (1980) The Annals of Humanities Computing: The Index Thomisticus, *Computers and the Humanities* 14(3), pp. 83–90.

Callon, M. (1986) The Sociology of an Actor-Network: The Case of the Electric Vehicle, in: M. Callon, J. Law and A. Rip (Eds) *Mapping the Dynamics of Science and Technology*, pp. 19–34 (London: MacMillan).

Campbell, P. (2006) Nature Peer Review Trial and Debate, *Nature*. Available at http://www.nature.com/nature/peerreview/ (accessed 30 October 2013).

Causer, T., Tonra, J., and Wallace, V. (2012) Transcription Maximized; Expense Minimized? Crowdsourcing and Editing the Collected Works of Jeremy Bentham, *Literary and Linguistic Computing* 27(2), pp. 119-137.

Charmaz, K. (2006) *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis* (Thousand Oaks, CA: Sage).

Chartier, R. (1995) Lecteurs dans la longue durée: du codex à l'écran, in: R. Chartier (Ed) *Histoire de la lecture: un bilan des recherches, actes du colloque des 29 et 30 janvier 1993, Paris*, pp. 271–283 (Paris: IMEC).

Chun, W.H. (2013) *The Dark Side of the Digital Humanities—Part 3*, paper presented at MLA2013, Boston, MA, January. Available at http://www.c21uwm.com/2013/01/09/the-dark-side-of-the-digital-humanities-part-1/ (accessed 29 October 2013).

Ciborra, C., and Hanseth, O. (2000) *From Control to Drift. The Dynamics of Corporate Information Infrastructures* (Oxford: Oxford Unversity Press).

CLARIN (n.d) *Executive Summary*. Available at http://www.clarin.eu/executive-summary (accessed 27 July 2014).

Clement, T., and Reside, D. (2011) *Off the Tracks: Laying New Lines for Digital Humanities Scholars* (media commons press). Available at http://mcpress.media-commons.org/offthetracks (accessed 14 July 2014).

Cohen, D. (2007, April 19) Digital Humanities Summit Wrap-up, *Blog post*. Available at http://www.dancohen.org/2007/04/19/digital-humanities-

summit-wrap-up (accessed 14 July 2014).

Cohen, P. (2010, November 16) Humanities 2.0. Digital Keys for Unlocking the Humanities' Riches, *New York Times*. Available at http://www.nytimes.com/2010/11/17/arts/17digital.html?pagewanted=all&_r=0 (accessed 29 October 2013).

Collins, E., Bulger, M.E., Meyer, E.T. (2012) Discipline Matters: Technology Use in the Humanities, *Arts & Humanities in Higher Education* 11(1-2), pp. 76-92.

Collins, H.M. (1985) *Changing Order. Replication and Induction in Scientific Practice* (Chicago and London: University of Chicago Press).

Collins H.M., Evans R., and Gorman M.E. (2010) Trading Zones and Interactional Expertise, in: M.E. Gorman (Ed) *Trading Zones and Interactional Expertise: Creating New Kinds of Collaboration*, pp. 7–23 (Cambridge, MA: The MIT Press).

Collins, H.M., and Kusch, M. (1998) *The Shape of Actions. What Humans and Machines Can Do* (Cambridge, MA: The MIT Press).

Contu, A., and Willmott, H. (2003) Re-Embedding Situatedness: The Importance of Power Relations in Learning Theory, *Organization Science* 14(3), pp. 283-296.

COST (2009) *Memorandum of Understanding for the Implementation of a European Concerted Research Action Designated COST Action ISO901: Women Writers in History—Toward a New Understanding of European Literary Culture. Brussels, 11 June*, COST 239/09. Available at http://www.cost.eu/domains_actions/isch/Actions/IS0901 (accessed 1 December 2014).

Dahl, R.A. (1958) A Critique of the Ruling Elite Model, *American Political Science Review* 52(2), pp. 463–469.

DARIAH (n.d) *Welcome*. Available at http://www.dariah.eu/index.php?option=com_content&view=frontpage&Itemid=113 (accessed 27 July 2014).

Davenport, E. (2001) Knowledge Management Issues for Online Organisations: 'Communities of Practice' as an Exploratory Framework, *Journal of Documentation* 57(1), pp. 61–75.

David, P. (2005) Towards a Cyberinfrastructure for Enhanced Scientific Collaboration: Providing its 'Soft' Foundations May Be the Hardest

Part, *Oxford Internet Institute Research Report No. 2*, University of Oxford. Available at http://www.oii.ox.ac.uk/research/publications/RR4.pdf

de la Flor, G., Jirotka, M., Luff, P., Pybus, J., and Kirkham, R. (2010) Transforming Scholarly Practice: Embedding Technological Interventions to Support the Collaborative Analysis of Ancient Texts, *Computer Supported Cooperative Work* 19(3/4), pp. 309-334.

DH Summit 2007 (2007) *Minutes*. Available at https://apps.lis.illinois.edu/wiki/pages/viewpage.action?pageId=22021 (accessed 14 July 2014).

Dojcinovic-Nesic, B. (2006) *GendeRings: Gendered readings in Serbian women's writing* (Beograd: Asocijacija za zhensku inicijativu).

Dombrowski, Q. (2014) What Ever Happened to Project Bamboo?, *Literary and Linguistic Computing* 29(3), pp. 326-339.

Doorenbosch, P. (1993, October 26) Med: 9311.02: Informatie over BNTL Online Retrieval System, *Neder–L*. Available at https://listserv.surfnet.nl/scripts/wa.exe?A2=ind9311&L=neder-l&P=56 (accessed 12 November 2008).

Drucker, J. (2009) *SpecLab. Digital Aesthetics and Projects in Speculative Computing* (Chicago and London: University of Chicago Press).

Drucker, J. (2011) Humanities Approaches to Graphical Display, *Digital Humanities Quarterly* 5(1). Available at http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html (accessed 29 October 2013).

Duguid, P. (2007) Inheritance and Loss? A Brief Survey of Google Books, *First Monday* 12(8-6). Available at http://firstmonday.org/article/view/1972/184

Dutton, W.H., and Jeffreys, P.W. (Eds) (2010) *World Wide Research. Reshaping the Sciences and Humanities* (Cambridge, MA: The MIT Press).

Edler, J., Kuhlmann, S., and Behrens, M. (Eds) (2003) *Changing Governance of Research and Technology Policy: The European Research Area (*Cheltenham: Edward Elgar).

Edwards, P. N. (2010) *A Vast Machine. Computer Models, Climate Data, and the Politics of Global Warming* (Cambridge, MA: The MIT Press).

Edwards, P.N., Bowker, G.C., Jackson, S., and Williams, R. (2009)

Introduction: An Agenda for Infrastructure Studies, *Journal of the Association for Information Systems* 10(5), pp. 364–374.

Edwards, P. N., Jackson, S. J., Bowker, G. C., and Knobel, C. P. (2007) *Understanding Infrastructure: Dynamics, Tensions, and Design*, NSF Report of a Workshop on "History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures". Available at http://hdl.handle.net/2027.42/49353 (accessed 14 July 2014).

Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., Burton, M., and Calvert, S. (2013) *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges* (Ann Arbor: Deep Blue). Available at http://hdl.handle.net/2027.42/97552 (accessed 11 November 2014).

Edwards, P.N., Mayernik, M.S., Batcheller, A., Bowker, G.C., and Borgman, C.L. (2011) Science Friction: Data, Metadata, and Collaboration, *Social Studies of Science* 41(5), pp. 667–690.

Eisenstein, E.L. (1979) *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early Modern Europe*, 2 vols. (Cambridge, UK: Cambridge University Press).

ESF (2011) Research infrastructures in the digital humanities, *Science Policy Briefing*, 42. Available at http://www.esf.org/publications/science-policy-briefings.html (accessed 1 December 2014).

ESFRI (2006a) *European Roadmap for Research Infrastructures* (Luxembourg: Office for Official Publications of the European Communities). Available at http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2006/esfri_roadmap_2006_en.pdf (accessed 14 July 2014).

ESFRI (2006b) *Report of the Social Sciences and Humanities Roadmap Working Group*. Available at http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2006/ssh-rwg-roadmap-report-2006_en.pdf (accessed 14 July 2014).

ESFRI (2008) *Social Sciences and Humanities Roadmap Working Group Report.* Available at http://ec.europa.eu/research/infrastructures/pdf/esfri/esfri_roadmap/roadmap_2008/ssh_report_2008_en.pdf (accessed 26 March 2014).

ESFRI (2011) *Evaluation Report*. Available at http://ec.europa.eu/research/infrastructures/pdf/esfri_evaluation_report_2011.pdf (accessed 14 July 2014).

Espeland, W. N., and Sauder, M. (2007). Rankings and Reactivity: How Public Measures Recreate Social Worlds, *American Journal of Sociology* 113(1), pp. 1–40.

Etzkowitz, H., and Leydesdorff, L. (2000) The Dynamics of Innovation: From National Systems and ''Mode 2'' to a Triple Helix of University–Industry–Government Relations, *Research Policy* 29(2), pp. 109–123.

European Commission (2000) *Towards a European Research Area*, COM/2000/0006 final. Available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52000DC0006 (accessed 14 July 2014).

European Commission (2007) *The European Research Area: New Perspectives*, COM/2007/161 final. Available at http://ec.europa.eu/research/era/pdf/era_gp_final_en.pdf (accessed 14 July 2014).

European Commission (2008) *Developing World-class Research Infrastructures for the European Research Area (ERA). Report of the ERA Expert Group* (Luxembourg: Office for Official Publications of the European Communities). Available at http://ec.europa.eu/research/infrastructures/pdf/ri_era-expert-group-0308_en.pdf (accessed 14 July 2014).

European Commission (2013) *Assessing the Projects on the ESFRI Roadmap. A High Level Expert Group Report.* Available at http://ec.europa.eu/research/infrastructures/pdf/jd-final-aegreport-23sept13.pdf (accessed 14 July 2014).

Favell, A., and Guiraudon, V. (2009) The Sociology of the European Union. An Agenda, *European Union Politics* 10(4), pp. 550-576.

Fish, S. (2012, January 23) Mind your P's and B's: The Digital Humanities and Interpretation, *The New York Times*. Available at http://opinionator.blogs.nytimes.com/2012/01/23/mind-your-ps-and-bsthedigital-humanities-and-interpretation/ (accessed 29 October 2013).

Fitzpatrick, K. (2009) *Planned Obsolescence: Publishing, Technology, and the Future of the Academy* (media commons press). Available at http://mediacommons.futureofthebook.org/mcpress/plannedobsolesce

nce/ (accessed 27 January 2013).

Fitzpatrick, K. (2011) *Planned Obsolescence: Publishing, Technology, and the Future of the Academy* (New York: NYU Press).

Flanders, J. (2011) You Work at Brown. What Do You Teach? *#alt-academy*. Available at http://mediacommons.futureofthebook.org/alt-ac/pieces/you-work-brown-what-do-you-teach (accessed 5 May 2013).

Foucault, M. (1979) What Is an Author?, in: J. V. Harari (Ed) *Textual Strategies: Perspectives in Post-Structuralist Criticism*, pp. 141–160 (Ithaca, NY: Cornell University Press).

French, A. (2013) Re: 27.308 Historical Documents On Project Management, *Humanist Discussion Group* 27(313). Available at http://lists.digitalhumanities.org/pipermail/humanist/2013-September/011241.html (accessed 14 June 2014).

Friedlander, A. (2009) Asking Questions and Building a Research Agenda for Digital Scholarship, in: *Working Together or Apart: Promoting the Next Generation of Digital Scholarship* (Washington, DC: Council on Library and Information Resources). Available at http://www.clir.org/pubs/reports/pub145/pub145.pdf.

Fry, J., and Talja, S. (2007) The Intellectual and Social Organization of Academic Fields and the Shaping of Digital Resources, *Journal of Information Science* 33(2), pp. 115–133.

Fry, J., and Schroeder, R. (2010) The Changing Disciplinary Landscapes of Research, in: W. H. Dutton and P. W. Jeffreys (Eds), *World Wide Research. Reshaping the Sciences and Humanities*, pp. 257–271 (Cambridge, MA: The MIT Press).

Fuchs, S. (1993) A Sociological Theory of Scientific Change, *Social Forces* 71(4), pp. 933–953.

Fujimura, J.H. (1987) Constructing "Do-able" Problems in Cancer Research: Articulating Alignment, *Social Studies of Science* 17(2), pp. 257–293.

Fujimura, J.H. (1992) Crafting Science: Standardized Packages, Boundary Objects, and 'Translations', in: A. Pickering (Ed) *Science as practice and culture*, pp. 168–214 (Chicago and London: University of Chicago Press).

Galison, P. (1997) *Image and Logic: A Material Culture of Microphysics* (Chicago and London: University of Chicago Press).

Georgakakis, D., and Weisbein, J. (2010) From Above and From Below: A Political Sociology of European Actors, *Comparative European Politics* 8(1), pp. 93–109.

Gold, M.K. (Ed.) (2012) *Debates in the Digital Humanities* (Minneapolis: University of Minnesota Press).

Greenstein, D. (1998, December) The Arts and Humanities Data Service Three Years' On, *D-Lib Magazine*. http://www.dlib.org/dlib/december98/greenstein/12greenstein.html. (accessed 14 July 2014).

Grusin, R. (2013) *The Dark Side of the Digital Humanities—Part 2*, paper presented at MLA2013, Boston, MA, January. Available at http://www.c21uwm.com/2013/01/09/dark-side-of-thedigital-humanities-part-2/ (accessed 29 October 2013).

Guzzetti, L. (2009) The 'European Research Area' Idea in the History of the Community Policy-Making, in: H. Delanghe, U. Muldur and L. Soete (Eds) *European Science and Technology Policy. Towards Integration or Fragmentation?*, pp. 64-77 (Cheltenham: Edward Elgar).

Hajer, M. (1993) Discourse Coalitions and the Institutionalisation of Practice: The Case of Acid Rain in Great Britain, in: F. Fischer and J. Forester (Eds) *The Argumentative Turn in Policy Analysis and Planning*, pp. 43-67 (London: Durham).

Hajer, M. (2003) *Deliberative Policy Analysis: Understanding Governance in the Network Society* (Cambridge: Cambridge University Press).

Hansen, M. (2012) Shi jian: Time, *Vectors* 3(2). Available at http://vectors.usc.edu/projects/index.php?project=91&thread=ProjectCredits (accessed 29 October 2013).

Haraway, D. (1989) *Primate Visions: Gender, Race, and Nature in the World of Modern Science* (Routledge: New York and London).

Hepsø, V., Monteiro, E., and Rolland, K. H. (2009) Ecologies of e-Infrastructures, *Journal of the Association for Information Systems* 10(5), pp. 430-446.

Hessels, L. and van Lente, H. (2008) Re-thinking New Knowledge Production: a Literature Review and a Research Agenda, *Research Policy* 37(4), pp 740–760.

Hey, T., and Trefethen, A.E. (2002) The UK e-Science Core Programme and

the Grid, *Future Generation Computer Systems* 18(8), pp. 1017–1031.

Hey, T., and Trefethen, A.E. (2005) Cyberinfrastructure for e-Science, *Science* 308(5723), pp. 817–821.

Hey, T., Tansley, S., and Tolle, K. (Eds) (2009) *The Fourth Paradigm: Data-Intensive Scientific Discovery* (Redmond, WA: Microsoft Research). Available at http://research.microsoft.com/en-us/collaboration/fourthparadigm/ (accessed 27 January 2014).

Hicks, J. 2013) *Old Questions, New Techniques: A Research Note on Identifying Political Elites Computationally*. Unpublished draft.

Hilgartner, S. (1995) Biomolecular Databases: New Communication Regimes for Biology, *Science Communication* 17(2), pp. 240–263.

Hine, C. (2006) Databases as Scientific Instruments and their Role in the Ordering of Scientific Work, *Social Studies of Science* 36(2), pp. 269–298.

Hine, C. (2007) Multi-Sited Ethnography as a Middle Range Methodology for Contemporary STS, *Science, Technology & Human Values* 32(6), pp. 652-671.

Hine, C. (2008) *Systematics as Cyberscience: Computers, Change and Continuity in Science*. (Cambridge, MA: The MIT Press).

Hockey, S. (2004) The History of Humanities Computing, in: S. Schreibman, R. Siemens and J. Unsworth (Eds) *A Companion to Digital Humanities* (Oxford: Blackwell). http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-1&toc.depth=1&toc.id=ss1-2-1&brand=default (accessed 27 January 2014).

Hughes, T.P. (1983) *Networks of Power: Electrification in Western Society, 1880-1930* (Baltimore: Johns Hopkins University Press).

Hutchins, J. (1999) Retrospect and Prospect in Computer-Based Translation, *Proceedings of Machine Translation Summit VII "MT in the Great Translation Era", 13th-17th September 1999, Kent Ridge Labs, Singapore*, pp. 30-44 (Tokyo: AAMT). Available at http://www.hutchinsweb.me.uk/MTS-1999.pdf (accessed 5 December 2014).

Huygens Instituut KNAW (n.d.) *Bibliografie van de Nederlandse Taal- en Literatuurwetenschap.* Available at http://www.bntl.nl/bntl/# (accessed 1

May 2009).

Huygens Instituut KNAW (2006) *Vierde nieuwsbrief BTNL — November 2006*. Available at http://www.huygensinstituut.knaw.nl/index.php?option=com_content&task=view&id=229&Itemid=129&lang=du (accessed 26 April 2009).

Huygens Instituut KNAW (2004a) *Newsletter 5 October*. Available at http://www.huygensinstituut.knaw.nl/index.php?option=com_content&task=view&id=46&Itemid=39&lang=nl (accessed 25 April 2009).

Huygens Instituut KNAW (2004b) *Ontwikkelingsplan digitale bronnen en methoden*. Available at http://www.knaw.nl/nieuws/pers_pdf/CHI-Neerlandistiek.pdf (accessed 25 April 2009).

Jankowski, N.W. (2007) Exploring e-Science: An Introduction. *Journal of Computer-Mediated Communication* 12(2), pp. 542–569.

Jankowski, N.W. (Ed) (2009) *E-Research: Transformation in Scholarly Practice* (London: Routledge).

Jasanoff, S. (Ed) (2004) *States of Knowledge: The Co-Production of Science and the Social Order* (London: Routledge).

Jeffrey, P. (2003) Smoothing the Waters: Observations on the Process of Cross-Disciplinary Research Collaboration, *Social Studies of Science* 33(4), pp. 539–562.

Jensen, C.B. (2007) Infrastructural Fractals: Revisiting the Micro – Macro Distinction in Social Theory, *Environment and Planning D*: *Society and Space* 25(5), pp. 832-850.

Jensen, C.B., and Winthereik, B.R. (2013) *Monitoring Movements in Development Aid: Recursive Partnerships and Infrastructures* (Cambridge, MA: The MIT Press).

Joerges, B., and Shinn, T. (2001) *Instrumentation: Between Science, State, and Industry* (Dordrecht: Kluwer).

Johns, A. (1998) *The Nature of the Book: Print and Knowledge in the Making* (Chicago and London: University of Chicago Press).

Johns, A. (2002) How to Acknowledge a Revolution, *The American Historical Review* 107(1), pp. 106–125

Johns, A. (2009) *Piracy: The Intellectual Property Wars From Gutenberg to Gates* (Chicago and London: University of Chicago Press).

Kaltenbrunner, W. (2014) Decomposition as Practice and Process: Creating Boundary Objects in Computational Humanities, *Interdisciplinary Science Reviews* 39(2), pp. 143-161.

Kaltenbrunner, W. (2014) Infrastructural Inversion as a Generative Resource in Digital Scholarship, *Science as Culture* 24(1), pp. 1-23.

Kaltenbrunner, W. (2014) Scholarly Labour and Digital Collaboration in Literary Studies, *Social Epistemology* 29(2), pp. 207-233.

Kaltenbrunner, W., and Wouters, P. (2010) E-research and methodological innovation in Dutch literary studies, *First Monday* 15(9). Available at http://firstmonday.org/ojs/index.php/fm/article/view/3078/2602.

Kamler, B. (2008) Rethinking Doctoral Publication Practices: Writing From and Beyond the Thesis, *Studies in Higher Education* 33(3), pp. 283-294.

Kirschenbaum, M.G. (2010) What Is Digital Humanities and What's It Doing in English Departments?, *ADE Bulletin* 150, pp. 55–61.

Kirschenbaum, M.G. (2012) Digital Humanities As/Is a Tactical Term, in: M. K. Gold (Ed) *Debates in the Digital Humanities,* pp. 415–428 (Minneapolis: University of Minnesota Press).

KNAW/NWO (2004) *Intentieverklaring inzake de Nederlandse data–infrastructuur in de Maatschappij– en Gedragswetenschappen en de Geesteswetenschappen* (Amsterdam: Data Archiving and Networked Services). Available at http://www.knaw.nl/nieuws/pers_pdf/DANS_intentieverklaring.pdf (accessed 30 March 2010).

Knorr Cetina, K. (1981) *The Manufacture of Knowledge — An Essay on the Constructivist and contextual Nature of Science* (Oxford: Pergamon Press).

Knorr Cetina, K. (1999) *Epistemic Cultures: How the Sciences Make Knowledge* (Cambridge, MA: Harvard University Press).

Koh, A. (2013) The Journal of Digital Humanities: Post-Publication Review or the Worst of Peer Review? *Blogpost.* Available at http://www.adelinekoh.org/blog/2013/08/29/journalofdigitalhumanitie/ (accessed 30 October 2013).

Kok, J., and Wouters, P. (2013) Virtual Knowledge in Family History: Visionary Technologies, Research Dreams, and Research Agendas, in: S. Wyatt, A. Scharnhorst, A. Beaulieu and P. Wouters (Eds) *Virtual Knowledge*, pp. 219–250 (Cambridge, MA: The MIT Press).

Larkin, B. (2013) The Politics and Poetics of Infrastructure, *Annual Review of Anthropology* 42, pp. 327–43.

Latour, B. (1987) *Science in Action: How to Follow Scientists and Engineers through Society* (Cambridge, MA: Harvard University Press).

Latour, B. (1993) *We Have Never Been Modern*, trans. C. Porter (Cambridge, MA: Harvard University Press).

Latour, B., and Woolgar, S. (1979) *Laboratory Life: The Social Construction of Scientific Facts* (Thousand Oaks, CA: Sage).

Law, J. (2004) *After Method: Mess in Social Science Research* (London: Routledge).

Leiden University (2008) *Leiden University PhD Regulations*. Available at http://media.leidenuniv.nl/legacy/PhD%20Regulations%202008.pdf (accessed 1 December 2014).

Lemaine, G., Macleod, R., Mulkay, M., and Weingart, P. (Eds) (1976) *Perspectives on the Emergence of Scientific Disciplines* (Chicago: Aldine).

Lenoir, T. (1998) Shaping Biomedicine as an Information Science, in: M. E. Bowden, T. B. Hahn and R. V. Williams (Eds) *Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems*, pp. 27–45 (Medford, N.J.). Available at http://www.stanford.edu/dept/HPS/TimLenoir/shapingbiomedicine.html (accessed 30 March 2010).

Lev, D.S. (1966) *The Transition to Guided Democracy: Indonesian Politics, 1957–1959* (Ithaca, NY: Cornell).

Liu, A. (2005) The Humanities: A Technical Profession, in: M. Hanrahan and D. Madsen (Eds) *Teaching, Technology, Textuality: Approaches to New Media*, pp. 11-26 (Basingstoke: Palgrave Macmillan).

Lossau, N., and Peters, D. (2008) DRIVER: Building a Sustainable Infrastructure of European Scientific Repositories, *LIBER Quarterly* 18(3/4), pp. 437-448.

Lyall, C., Bruce, A., Tait, J., and Meagher, L. (2011) *Interdisciplinary Research Journeys: Practical Strategies for Capturing Creativity* (London: Bloomsbury).

Lynch, C.A. (2002) Digital collections, digital libraries and the digitization of cultural heritage information, *First Monday* 7(5). Available at

http://firstmonday.org/ojs/index.php/fm/article/view/949/870.

Lynch, M. (2014) Ontography: Investigating the production of things, deflating ontology, *Social Studies of Science* 43(3), pp. 444-462.

Marcus, G.E. (1995) Ethnography in/of the world system: The emergence of multi-sited ethnography, *Annual Review of Anthropology* 24, pp. 95-117.

Martin, B., and Whitley, R. (2010) The UK Research Assessment Exercise, in: R. Whitley, J. Gläser and L. Engwall (Eds) *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and Their Consequences for Intellectual Innovation* (Oxford: Oxford University Press).

Mayernik, M. S., Wallis, J., and Borgman, C. L. (2013) Unearthing the Infrastructure, Humans and Sensors in Field-Based Scientific Research, *Computer Supported Cooperative Work* 22(1), pp. 65–101.

McCarty, W., and Short, H. (2002) *Mapping the Field. Report of ALLC Meeting Held in Pisa, April 2002*. Available at http://www.allc.org/content/pubs/map.html (accessed 30 May 2013).

Meijer, I., Molas-Gallart, J., and Mattsson, P. (2012) Networked Research Infrastructures and their Governance: The Case of Biobanking, *Science and Public Policy* 39(4), pp. 491–499.

Michel, J.-B., Yuan, K. S., Aviva P. A., Veres, A., Gray, M. K. et al. (2011) Quantitative Analysis of Culture Using Millions of Digitized Books, *Science* 331(6014), pp. 176–182.

Mills, W.C. (1956) *The Power Elite* (Oxford: Oxford Press).

Mirowski, P., and Sent, E.-M. (2008) The Commercialization of Science and the Response of STS, in: E. J. Hackett, O. Amsterdamska, M. Lynch and J. Wajcman (Eds) *The Handbook of Science and Technology Studies*, 3rd ed., pp. 635-689 (Cambridge, MA: The MIT Press).

Modern Language Association (2012) *Guidelines for Evaluating Work in Digital Humanities and Digital Media*. Available From http://www.mla.org/guidelines_evaluation_digital (accessed 1 December 2014).

Mol, A. (2002) *The Body Multiple: Ontology in Medical Practice* (Durham, NC: Duke University Press).

Moretti, F. (2005) *Graphs, Maps, Trees: Abstract Models for a Literary History*

(Brooklyn: Verso).

Nelkin, D. (1995) Science Controversies: The Dynamics of Public Disputes in the United States, in: S. Jasanoff, G. Markle, J. Petersen and T. Pinch (Eds) *Handbook of Science and Technology Studies,* rev. ed., pp. 444-457 (Thousand Oaks, CA: Sage).

Nentwich, M. (2003) *Cyberscience: Research in the Age of the Internet* (Vienna: Austrian Academy of Sciences Press).

Nowotny, H., Scott, P., and Gibbons, M. (2001) *Rethinking Science: Knowledge in an Age of Uncertainty* (Cambridge: Polity).

NSF (2013) *FY 2014 Budget Request to Congress.* Available at http://www.nsf.gov/about/budget/fy2014/pdf/EntireDocument_fy2014.pdf(accessed 14 July 2014).

Nyhan, J. (Ed) (2012) Hidden Histories: Computing and the Humanities c. 1965–1985, *Digital Humanities Quarterly* 6(3). Available at http://www.digitalhumanities.org/dhq/vol/6/3/index.html (accessed 14 July 2014).

Olson, G.M., Zimmerman, A., and Bos, N. (Eds) (2008) *Scientific Collaboration On the Internet* (Cambridge, MA: The MIT Press).

Oostindie, G., de Jong, F., Schwegman, M., and de Rijke, M. (2011) *Elite Network Shifts During Regime Change; a Computational Approach to Network Analysis Using Indonesian Language Electronic Newspaper Archives*, unpublished research proposal.

Oudshoorn, N.E., and Pinch, T.J. (Eds) (2003) *How Users Matter. The Co-Construction of Users and Technology* (Cambridge, MA: The MIT Press).

Pannapacker, W. (2011, January 8) Digital Humanities Triumphant? *The Chronicle of Higher Education*. Available at http://chronicle.com/blogs/brainstorm/pannapacker-at-mla-digital-humanities-triumphant/30915 (accessed 15 July, 2014).

Pannapacker, W. (2013, January 5) On "The Dark Side of the Digital Humanities", *The Chronicle of Higher Education*. Available at http://chronicle.com/blogs/conversation/2013/01/05/on-the-dark-side-of-the-digital-humanities/ (acccessed 5 May 2013).

Park, C. (2005) New Variant PhD: The Changing Nature of the Doctorate in the UK, *Journal of Higher Education Policy and Management* 27(2), pp. 189-207.

Pickering, A., and Stephanides, A. (1992) Constructing Quaternions: On the Analysis of Conceptual Practice, in: A. Pickering (Ed) *Science as Practice and Culture*, pp. 139–167 (Chicago and London: University of Chicago Press).

Piersma, H., and Ribbens, K. (2013) Digital Historical Research Context, Concepts and the Need for Reflection, *BMGN. Low Countries Historical Review* 128(4), pp. 78-102

Pipek, V., and Wulf, V. (2009) Infrastructuring: Towards an Integrated Perspective on the Design and Use of Information Technology, *Journal of the Association of Information Systems* 10(5), pp. 306-332.

Pollock, N., and Williams, R. (2010) E-infrastructures: How Do We Know and Understand Them? Strategic Ethnography and the Biography of Artefacts, *Computer Supported Cooperative Work* 19(6), pp. 521-556.

Powell, S., and Green, H. (Eds) (2007) The Doctorate Worldwide (Maidenhead, Berkshire: Open University Press).

Progressive Library Guild Greater Toronto Area (2013, May 13) *Statement on Library and Archives Canada*. Available at http://plggta.org/archives/201 (accessed 1 December 2014).

Ramsay, S. (2011) *Reading Machines: Toward an Algorithmic Criticism* (Chicago: University of Illinois Press).

Ramsay, S., and Rockwell, G. (2012) Developing Things: Notes Toward an Epistemology of Building in the Digital Humanities, in: M. K. Gold (Ed.) *Debates in the Digital Humanities*, pp. 75–84 (Minneapolis: University of Minnesota Press).

Ramsay, S. (2013, July 27) Bambazooka, *Blog post*. Available at http://stephenramsay.us/2013/07/23/bambazooka (accessed 14 July 2014).

Rheinberger, H.-J. (1997) *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube* (Stanford: Stanford University Press).

Ribes, D., and Finholt, T.A. (2009) The Long Now of Technology Infrastructure*:* Articulating Tensions in Development*, Journal for the Association of Information Systems* 10(5), pp. 375-398.

Ribes, D., and Lee, C.P. (2010) Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories, *Computer Supported Cooperative Work* 19(3-4), pp. 231–244.

Rizzuto, C., and Wood, J. (2013) *RAMIRI Handbook*. Available at http://www.ramiri-blog.eu (accessed 1 December 2014).

Robins, L., and Kanowski, P. (2008) PhD by Publication: A Student's Perspective, *Journal of Research Practice* 4(2), pp. 1-20. Available at http://www.jrp.icaap.org/index.php/jrp/article/view/136/154.

Robinson, L.S. (1983) Treason our text: Feminist challenges to the literary canon, *Tulsa Studies in Women's Literature* 2(1). pp. 83–98.

Romm, N.R. (1998) Interdisciplinary Practice as Reflexivity, *Systemic Practice and Action Research* 11(1), pp. 48–50.

Rowe, K. (2009, October 2) *Comment on Planned Obsolescence: Publishing, Technology, and the Future of the Academy*. Available at http://mediacommons.futureofthebook.org/mcpress/plannedobsolescence/one/ (accessed 5 March 2013).

Santos, D., and Frankenberg-Garcia, A. (2007) The Corpus, Its Users and Their Needs: a User-Oriented Evaluation of COMPARA, *International Journal of Corpus Linguistics* 12(3), pp. 335–374.

Scheinfeldt, T. (2008, March 13). Sunset for Ideology, Sunrise for Methodology? *Blogpost*. Available at http://www.foundhistory.org/2008/03/13/sunset-for-ideology-sunrise-for-methodology/ (accessed 30 November 2014).

Scheinfeldt, T. (2011) Toward a Third Way: Rethinking Academic Employment, *#alt-academy*. Available at http://mediacommons.futureofthebook.org/alt-ac/pieces/toward-third-way-rethinking-academic-employment (accessed 30 November 2014).

Schmidt, K., and Bannon, L. (1992) Taking CSCW Seriously: Supporting Articulation Work, *Computer Supported Cooperative Work* 1(1), pp. 7–40.

Schmidt, K., and Simone, C. (1996) Coordination Mechanisms: Towards a Conceptual Foundation of CSCW Systems Design, *Computer Supported Cooperative Work* 5(2–3), pp. 155–200.

Schroeder, R. (2008) e-Sciences as Research Technologies: Reconfiguring Disciplines, Globalizing Knowledge, *Social Science Information* 47(2), pp. 131–157.

Shapin, S. (1989) The Invisible Technician, *American Scientist* 77(6), pp. 554–563.

Shapin, S. (1994) *A Social History of Truth: Civility and Science in Seventeenth Century* (Chicago and London: University of Chicago Press).

Shapin, S., and Schaffer, S. (1985) *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life* (Princeton, NJ: Princeton University Press).

Shellman, S.M. (2008) Coding Disaggregated Intrastate Conflict: Machine Processing the Behavior of Substate Actors Over Time and Space, *Political Analysis* 16(4), pp. 464–477.

Shinn, T., and Joerges, B. (2002) The Transverse Science and Technology Culture: Dynamics and Roles of Research–Technology, *Social Science Information* 41(2), pp. 207–251.

Shrum, W., Genuth, J., and Chompalov, I. (2007) *Structures of Scientific Collaboration* (Cambridge, MA: The MIT Press).

Smith, K. (2009) Q&A with Brett Bobley, Director of the NEH's Office of Digital Humanities, *HASTAC*. Available at http://www.hastac.org/blogs/kathleen/qa-brett-bobley-director-nehs-office-digital-humanities-odh (accessed 14 July 2014).

Sonneveld, H. (2010) Doctoral Education in the Netherlands, 2010 – A Brief History and a Silent Revolution, *EUA- Council for Doctoral Education - News*, 10(2-4). Available at http://www.phdcentre.eu/nl/publicaties/documents/EUANewsIssue10_LYpublicatieHS.pdf

Star, S.L. (1999) The Ethnography of Infrastructure, *American Behavioral Scientist* 43(3), pp. 377-391.

Star, S.L., and Griesemer, J.R. (1989) Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39, *Social Studies of Science* 19(3), pp. 387–420.

Star, S.L., and Ruhleder, K. (1996) Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces, *Information Systems Research* 7(1), pp. 111–134.

Star, S.L., and Strauss, A. (1999) Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible work, *Computer Supported Cooperative Work* 8(1), pp. 9–30.

Stichweh, R. (1984) *Zur Entstehung des modernen Systems wissenschaftlicher Disziplinen: Physik in Deutschland 1740-1890* (Frankfurt am Main:

Suhrkamp).

Strauss, A. (1985) Work and the Division of Labor, *The Sociological Quarterly* 26(1), pp. 1–19.

Strauss, A. (1988) The Articulation of Project Work: An Organizational Process, *The Sociological Quarterly* 29(2), pp. 163–178.

Strathern, M. (1991) *Partial Connections* (Lanham, MD: Rowman and Littlefield).

Suchman, L. (1996) Supporting Articulation Work, in: R. Kling (Ed.) *Computerization and Controversy: Value Conflicts and Social Choices*, pp. 407–423 (San Diego, CA: Academic Press).

Swadesh, M. (1952) Lexico-Statistic Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos, *Proceedings of the American Philosophical Society* 96(4), pp. 452–463.

Textgrid (2014) *Das Projekt*. Available at http://www.textgrid.de/ueber-textgrid/projekt (accessed 15 July, 2014).

THATCampCHNM (2011) *Intro to Project Planning & Management*. Available at https://docs.google.com/document/d/1uofGfhCDpu2DK1bKstkEqyK2hwbV4PN4SPmfPIWvk/edit (accessed 29 October 2013).

The Virtual Knowledge Studio. (2008) Messy Shapes of Knowledge, in: E.J. Hackett, O. Amsterdamska, M. Lynch and J. Wajcman (Eds) *The Handbook of Science and Technology Studies*, 3rd ed., pp. 319–351 (Cambridge, MA: The MIT Press).

Tsing, A.L. (2005) *Friction: An Ethnography of Global Connection* (Princeton, N.*J.:* Princeton University Press).

University of Amsterdam (2014) *Doctorate Regulations of the University of Amsterdam*. Available at http://www.uva.nl/binaries/content/assets/uva/en/research/phd/doctorate-regulations-2014.pdf?2832598565441 (accessed 1 December 2014).

Unsworth, J. (2000) *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?* Symposium on Humanities Computing: Formal Methods, Experimental Practice. Available at http://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html (accessed 27 January, 2014).

Unsworth, J. (2007) *Digital Humanities Centers as Cyberinfrastructure*, Plenary Address, DH Summit 2007, Maryland Institute for Technology in the Humanities. Available at http://people.brandeis.edu/~unsworth/dhcs.html (accessed 14 July 2014).

van Lente, H. (2000) Forceful Futures: From Promise to Requirement, in: N. Brown, B. Rappert and A. Webster (Eds) *Contested Futures. A Sociology of Prospective Techno-Science*, pp. 43–64 (London: Ashgate).

Vann, K., and Bowker, G.C. (2001) Instrumentalizing the Truth of Practice. *Social Epistemology* 15(3), pp. 247–262.

Vann, K., and Bowker, G.C. (2006) Interest in Production: On the Configuration of Technology Bearing Labours for Epistemic IT, in: C. Hine (Ed) *New Infrastructures for Knowledge Production*, pp. 71–98 (Hershey, PA: IDEA Group Publisher).

Verkruijsse, P. (2005, January 18) Med: 0501.28 Nogmaals BNTL — 1, *Neder–L.* Available at http://www.neder-l.nl/bulletin/2005/01/050128.html (accessed 12 February 2010).

Verkruijsse, P. (2004, December 1) Med: 0412.05 Redding BNTL, *Neder–L.* Available at http://www.neder-l.nl/bulletin/2004/12/041205.html (accessed 25 April 2009).

Verran, H. (2001) *Science and African Logic* (Chicago and London: Chicago University Press).

Vikkelsø, S. (2007) Description as Intervention: Engagement and Resistance in Actor-Network Analyses, *Science as Culture* 16(3), pp. 297-309.

Vonortas, N.S. (2009) Scale and Scope in Research, in: H. Delanghe, U. Muldur and L. Soete (Eds) *European Science and Technology Policy. Towards Integration or Fragmentation?*, pp. 142-159 (Cheltenham: Edward Elgar).

Voorbij, H. (1999) Onder neerlandici: Ervaringen met de BNTL, *Informatie professional*, 3(12), pp. 28–33.

Voss, A., Mascord, A., Fraser, M., Jirotka, M., Proctor, R. et al. (2007) *e-Research Infrastructure Development and Community Engagement*, paper presented at the UK e-Science 2007 All Hands Meeting. Available at http://www.allhands.org.uk/2007/proceedings/papers/866.pdf (accessed 14 July 2014).

Vrije Universiteit Amsterdam (n.d.) *Promoveren op artikelen*. Available at http://www.rechten.vu.nl/nl/Images/Richtlijn%20Promoveren%20op%20artikelen_tcm22-334043.pdf (accessed 1 December, 2014).

Warhol, R.R., and Herndl, D.P. (Eds) (1997) *Feminisms: An Anthology of Literary Theory and Criticism* (New Brunswick, NJ: Rutgers University Press).

Waters, D.J. (2013) An Overview of Digital Humanities, *Research Library Issues* 284, pp. 3–11. Available at http://publications.arl.org/rli284/ (accessed 6 June 2014).

Whitley, R. (2000) *The Intellectual and Social Organization of the Sciences*, 2nd ed. (Oxford: Oxford University Press).

Whitley, R. (2010) Reconfiguring the Public Sciences: The Impact of Governance Changes on Authority and Innovation in Public Science Systems, in: R. Whitley, J. Gläser and L. Engwall (Eds) *The Intellectual and Social Organization of the Sciences*, pp. 3-47 (Oxford: Oxford University Press).

Whitley, R., Gläser, J., and Engwall, L. (Eds) (2010) *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and Their Consequences for Intellectual Innovation* (Oxford: Oxford University Press).

Whittle, R. (2013) *Gender, Canon and Literary History. The Changing Place of Nineteenth-Century German Women Writers* (Berlin and Boston: De Gruyter).

Whittle, A., and Spicer, A. (2008) Is Actor Network Theory Critique? *Organization Studies* 29(4), pp. 611–629.

Wilkens, M. (2012) Canons, Close Reading, and the Evolution of Methods, in: M. K. Gold (Ed) *Debates in the digital humanities*, pp. 249–258 (Minneapolis, MN: University of Minnesota Press).

Willekens, F., Beaulieu, A., Bod, R., van den Bosch, A., Honing, H., de Jong, F., Mandemakers, L., Schomaker, L., and van Zundert, Z. (2010) *Computational Humanities* (Amsterdam: Royal Netherlands Academy of Arts and Sciences). Available at http://www.mcg.uva.nl/index.html?abstracts/willekens-honing-et-al-2010.html&target (accessed 27 January 2014).

Williams, R., and Edge D. (1996) The Social Shaping of Technology, *Research*

Policy 25(6), pp. 865–899.

Williams, R., and Pollock, N. (2012) Moving Beyond the Single Site Implementations Study, *Information Systems Research* 23(1), pp. 1 – 22.

Williford, C., and Henry, C. (2012) *One Culture. Computationally Intensive Research in the Humanities and Social Sciences. A Report on the Experiences of First Respondents to the Digging into Data Challenge* (Washington, DC: Council on Library and Information Resources). Available at http://www.clir.org/pubs/reports/pub151/pub151.pdf (accessed 27 January 2014).

Winner, L. (1980) Do Artifacts Have Politics?, *Daedalus* 109(1), pp. 121–136.

Wisbey, R.A. (1962) Mechanization in Lexicography, *The Times Literary Supplement* 30 March 1962, repr. in: *Freeing the Mind – Articles and Letters From the Times Literary Supplement, March-June 1962*, pp. 15-18 (London: The Times).

Wissenschaftsrat (2011) *Empfehlungen zu Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften* (Berlin: Wissenschaftsrat). Available at http://www.wissenschaftsrat.de/download/archiv/10465-11.pdf (accessed 15 July, 2014).

Woolgar, S. (2002) Five Rules of Virtuality, in: S. Woolgar (Ed) *Virtual society? Technology, Cyberbole, Reality*, pp. 1–22 (Oxford: Oxford University Press).

Woolgar, S. (Ed.) (1988) *Knowledge and Reflexivity. New Frontiers in the Sociology of Knowledge* (London: Sage).

WordHoard. (2004–13) *WordHoard: An Application for the Close Reading and Scholarly Analysis of Deeply Tagged Text.* Available at http://wordhoard.northwestern.edu/userman/index.html (accessed 29 October 2013).

Wouters, P. (2006) What Is the Matter with e–Science? Thinking Aloud About Informatisation in Knowledge Creation, *Pantaneto Forum*, 23. Available at http://www.pantaneto.co.uk/issue23/wouters.htm (accessed 5 December 2013).

Wouters, P. (2007) *Conservative Culture or Diverging Identities? Studying Resistance to Technology in Academia*, paper presented at the Third International Conference on e-Social Science: Ann Arbor, Michigan, IL, 7-9 October. Available at

http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.100.8644 (accessed 20 April 2015).

Wouters, P., and Beaulieu, A. (2006) Imagining E-Science Beyond Computation, in: C. Hine (Ed) *New Infrastructure for Knowledge Production: Understanding E-Science*, pp. 48–70 (London: Information Science Publishing).

Wouters, P., Beaulieu, A., Scharnhorst, A., and Wyatt, S. (Eds) (2013) *Virtual Knowledge: Experimenting in the Humanities and the Social Sciences* (Cambridge, MA: The MIT Press).

Wyatt, S., and Balmer, B. (2007) Home on the Range: Where is the Middle in Science and Technology Studies?, *Science, Technology & Human Values* 32(6), pp. 619–626.

Wyatt, S., Scharnhorst, A., Beaulieu, A., and Wouters, P. (2013) Introduction to Virtual Knowledge, in: P. Wouters, A. Beaulieu, A. Scharnhorst and S. Wyatt (Eds) *Virtual Knowledge: Experimenting in the Humanities and the Social Sciences*, pp. 1–23 (Cambridge, MA: The MIT Press).

Zeldenrust, S., and Hagendijk, R. (1985) Book Reviews: R. D. Whitley: The Intellectual and Social Organization of the Sciences: 1984. Oxford: Oxford University Press, 319 pages, *Organization Studies* 6(4), pp. 385–386.

Žic-Fuchs, M. (2013) Research Infrastructures in the Humanities: the Challenge of 'Visibility' and 'Impact', in: A. Duşa, D. Nelle, G. Stock and G. G. Wagner (Eds) *Facing the Future: European Research Infrastructure for the Humanities and Social Sciences*, pp. 121-133 (Berlin: SCRIVERO). Available at http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/Tagungsbaende/FACING_THE_FUTURE.pdf (accessed 14 July 2014).

Zorich, D.M. (2008) *A Survey of Digital Humanities Centers in the United States* (Washington, DC: CLIR). Available at http://www.clir.org/pubs/reports/pub143/contents.html (accessed 14 July 2014).

Zuckerman, H., and Merton, R.K. (1971) Patterns of Evaluation in Science: Institutionalization, structure and functions of the referee system, *Minerva* 9(1), pp. 66–100.

Zuiderent-Jerak, T. (2007) Preventing Implementation: Exploring Interventions with Standardization in Healthcare, *Science as Culture* 16(3), pp. 311–329.

Zuiderent-Jerak, T., and Jensen, C.P. (2007) Unpacking 'Intervention' in Science and Technology Studies, *Science as Culture* 16(3), pp. 227-235

Interviews cited

Anonymous personal interview, 10 December 2008, The Hague.
Personal interview with Nieves Baranda, 14 November 2010, Madrid.
Personal interview with Wannie Carstens, 31 October 2008, Antwerp.
Personal interview with Marie-Louise Coolahan, 14 November 2010, Madrid.
Personal interview with Biljana Dojcinovic, 26 January 2012, The Hague.
Personal interview with Elly Kamp, 26 November 2008, The Hague.
Personal interview with Astrid Kulsdom, 13 October 2010, The Hague.
Personal interview with Karina van Dalen–Oskam, 5 October 2008, The Hague.
Personal interview with Suzan van Dijk, 8 November 2011, The Hague.
Personal interview with Gerry van Klinken and Fridus Steijlen, 12 March 2013, Leiden.
Personal interview with Piet Verkruijsse, 10 December 2008, Amsterdam.
Personal interview with Paul Wackers, 15 October 2008, Utrecht.
Personal interview with Henk Wals, 20 October 2008, The Hague.
Skype interview with Tobias Blanke, 4 June 2014.
Skype interview with Mirjam Blümm, 8 May 2014.
Skype interview with Steven Krauwer, 15 May 2014.
Skype interview with Jan Odijk, 4 June 2014.
Skype interview with Laurents Sesink, 21 May 2014.

Summary in Dutch

Deze dissertatie houdt zich bezig met de vraag hoe ontwikkeling en gebruik van digitale onderzoeksinstrumenten gerelateerd zijn aan veranderingen in de organisatie en praktijk van geesteswetenschappelijk onderzoek. Hoewel computationele methoden een lange geschiedenis hebben in dit veld, trekken ze pas sinds kort bredere aandacht. Digital Humanities (vanaf dit punt digitale geesteswetenschappen genoemd), de tegenwoordig gebruikelijke term, dient nu als een lens waardoor verschillende groepen van actoren - digitale en 'traditionele' geesteswetenschappers, beleidsmakers, de media – opnieuw uitvinden wat het betekent om geesteswetenschappelijk onderzoek te doen. In dit proces worden huidige praktijken, institutionele regelingen, en de relaties tussen de verschillende actoren gereconfigureerd. Digitale geestenswetenschappers zijn ook sterk betrokken bij initiatieven voor digitale infrastructuur. Deze initiatieven trekken aanzienlijke financiële middelen aan en hebben mogelijk langetermijneffecten op de methodologische organisatie van de geesteswetenschappen.

De dissertatie bevat vijf hoofdstukken die als aparte tijdschriftartikelen zijn gepubliceerd, maar die ook met elkaar zijn verbonden door een gezamenlijk theoretisch perspectief en complementaire onderzoeksvragen. Mijn conceptuele uitgangspunt is dat wetenschappelijke kennis binnen een infrastructuur wordt gegenereerd. Onder verwijzing naar Star & Ruhleder (1996) definieer ik dit begrip relationeel in plaats van inhoudelijk: infrastructuur is niet een specifiek object, maar ontstaat wanneer samenhangende praktijken in een werkende configuratie vallen. Theoretische kaders, materiële hulpmiddelen, protocollen en standaarden, bijvoorbeeld ten aanzien van data en publicatievormen, zijn ingebed in en geven mede vorm aan deze ecologie van praktijken. Functionerende infrastructuur reproduceert continu wat het betekent om 'serieus' onderzoek te doen, zodat men alleen bepaalde manieren van werken, specifieke kennisclaims, en bepaalde vormen van wetenschappelijke output als adequaat beschouwt.

Het werk van Whitley (2000) is nuttig om de eigenschappen van deze infrastructuur te benadrukken die specifiek gelden voor bepaalde disciplines. Whitley beargumenteert dat de meeste natuurwetenschappen worden gekenmerkt door een grote mate van sociale en intellectuele integratie. In de verschillende deelgebieden van de fysica bijvoorbeeld bestaat een sterke consensus over de theoretische grondslagen en urgente onderzoeksproblemen, en het wetenschappelijke werkproces is strak

georganiseerd rond het gebruik van vaak dure, grootschalige onderzoeksfaciliteiten. Daarentegen worden gebieden zoals geschiedenis, filosofie en literatuurwetenschap gekenmerkt door een relatief zwakke mate van sociale en intellectuele integratie. De dominante manier waarop kennisproducten circuleren is de monografie, die individuele auteurs aanzienlijke intellectuele en stilistische vrijheid geeft. Vaak bestaan er meerdere theoretische benaderingen naast elkaar en deze generen productieve conceptuele tegenstrijdigheden.

Het werk van Law (2004) kan worden gebruikt om in meer detail te theoretiseren over de methodologische en epistemologische fundamenten van een disciplinaire infrastructuur. Law beschouwt een methode niet als een objectief, context-onafhankelijk protocol dat het mogelijk maakt om verborgen betekenis uit de werkelijkheid te extraheren, maar als een generatief apparaat dat selectief en volgens bepaalde conventies aspecten van een onderzoeksobject 'versterkt'. Dit uitgangspunt benadrukt enerzijds de disciplinaire criteria waaraan kennisbijdragen moeten voldoen om als acceptabel te worden beschouwd. Law onderstreept verder dat sommige bijdragen makkelijker te genereren zijn dan andere omdat ze gestandaardiseerde sociotechnische arbeidsprocessen mobiliseren – ingeburgerde routines om gegevens in te delen suggereren bijvoorbeeld bepaalde manieren van het beheer en de analyse van empirisch materiaal, en de naleving van methodologische conventies vermindert complexiteit bij het onderhandelen van kennisclaims. Law noemt het geheel van deze gestandaardiseerde werk-routines het 'achterland' van een discipline.

Als leidende metafoor heb ik de geesteswetenschappelijke infrastructuur met een kennisproducerende machine vergeleken. Het beeld is niet bedoeld om een enkelvoudige monolithische eenheid voor te stellen, maar juist om de historische wederzijdse afhankelijkheid en complexiteit van de interne mechanismen te markeren. De machine functioneert alleen wanneer de vele bewegende delen op de juiste manier geconfigureerd en compatibel met elkaar zijn - dat wil zeggen, evoluerende, maar coherente theoretische kaders, een werkende werkverdeling tussen verschillende infrastructurele taakgebieden, evenals normen en protocollen die gevestigde routines ondersteunen. Digitale geesteswetenschappen brengen echter nieuwe mogelijkheden voor het verzamelen en analyseren van bronnen, nieuwe manieren om kennisclaims te maken en te verdedigen, evenals andere eisen wat betreft de organisatie van het onderzoeksproces. Als onderzoek als een complex kennisgenererende machine kan worden gezien, dan zijn nieuwe digitale benaderingen een poging om enkele onderdelen te

verwisselen terwijl de machine draait.

Mijn eerste onderzoeksvraag richtte zich op de implicaties van deze zienswijze voor de vorming van digitale onderzoekstechnologie: Wat betekent het om de geesteswetenschappen als een sociotechnische infrastructuur te analyseren, en hoe beïnvloedt dit infrastructurele karakter de inbedding van nieuwe onderzoeksinstrumenten in wetenschappelijke praktijken? In het eerste hoofdstuk heb ik de controversiële digitalisering van de Bibliografie van de Nederlandse Taal- en Letterkunde (BNTL) onderzocht. Van oudsher uitgegeven door de Koninklijke Nederlandse Academie van Wetenschappen, is het gebruik van de BNTL een belangrijk deel van onderzoekspraktijken in de Neerlandistiek. Neerlandici gebruiken de bibliografie om bronnen te vinden en een corpus van relevante kennis af te bakenen. Het bibliografische perspectief is op verschillende manieren aan de conceptuele dieptestructuur van de literatuurwetenschap gekoppeld, bijvoorbeeld door het performatieve gebruik van categorieën als 'auteur' en 'werk'. De digitalisering veranderde de bibliografie van een fysiek boek, dat in bibliotheken te raadplegen is, in een online zoekmachine. Een belangrijke bestuurlijke beslissing was de inkrimping van de redactie, waardoor de bibliografie nu een kleiner aantal publicaties beslaat - overwegend bekende tijdschriften en verzamelbundels in de literatuurwetenschap en de historische taalkunde, maar geen monografieën. Aan de andere kant hebben gebruikers nu ook de mogelijkheid zelf publicaties toe te voegen. De ontwerpers van de digitalisering onderstrepen verder de voordelen van realtime bijwerking en de mogelijkheid abstracts van de artikelen te bekijken. Uit mijn onderzoek is gebleken dat specifieke manieren om onderzoeksvragen te stellen en empirisch werk te organiseren een sterke invloed hebben op hoe de digitalisering door gebruikers werd ervaren. Voor onderzoekers in boekgeschiedenis en analytische bibliografie bijvoorbeeld was de door de uitgevers gewaarborgde dekking van de oude BNTL een belangrijke epistemische conditie voor hun werk. In plaats van een objectieve 'verbetering' van de bibliografie door extra functionaliteiten en onmiddellijke bijwerking, vertegenwoordigde de digitalisering in deze onderzoekspraktijken juist een infrastructurele storing.

Men zou kunnen argumenteren dat zulke spanningen kenmerkend zijn voor contexten waar wetenschapsmanagers en beleidsmakers belangrijke ontwerpkeuzes in een top-down manier maken. Mijn analyse van de COST Action Women Writers in History, een project in digitale literatuurgeschiedenis, heeft echter op problemen van soortgelijke aard gewezen. De deelnemers hadden een carrière opgebouwd rond de

bestudering van vrouwelijke schrijfsters tussen de 16e en de late 19e eeuw. Dit ter aanvulling op traditionele historische verslagen die zich richten op een beperkt aantal heldhaftige mannelijke figuren. Het gezamenlijk gebruik van een online database leek een goede gelegenheid om individuele onderzoeksinspanningen te combineren en daardoor een breder empirisch beeld van de 'vergeten schrijfsters' te creëren. Maar ondanks gelijkaardige disciplinaire achtergronden en een gedeeld interesse in digitale technologie, vonden veel deelnemers het moeilijk om de database in hun dagelijkse werkroutines in te bedden. Het gebruik vereiste niet alleen de integratie van afzonderlijke onderzoeksvragen en methoden, maar ook van uiteenlopende theoretische perspectieven. Bovendien werd het snel duidelijk dat een aanzienlijke arbeidsinzet nodig was om van individueel bijgedragen gegevens een betrouwbare, coherente dataset te maken die als basis voor wetenschappelijke publicaties en het werven van verdere financiële middelen gebruikt kon worden.

Deze twee case studies benadrukken dat de effectieve waarde van individuele instrumenten zeker niet alleen afhangt van ingebouwde functies en abstracte epistemologische voordelen, zoals interactieve? elementen, of het mogelijke gebruik van substantiële datasets en rekenkracht. Even belangrijk is de mate waarin de instrumenten in de wetenschappelijke kennismachine geïntegreerd kunnen worden zonder te blokkeren of te veel wrijving te produceren. Dit moet echter niet verward worden met intellectueel conservatisme of een bewuste weerstand tegen nieuwe technologie. Terwijl de zich langzaam ontwikkelende, institutioneel verankerde reproductie van praktijken vaak wordt gezien als een negatief kenmerk van academisch onderzoek, is het juist deze inertie die ervoor zorgt dat nieuwe wetenschappelijke bijdragen zinvol aansluiten bij bestaande kennis, zowel wat betreft conceptuele categorieën en veronderstelde niet-tastbare kennis, maar ook de betrouwbaarheid van empirische bronnen.

De academische arbeidsecologie wordt gekenmerkt door tal van onderlinge afhankelijkheden tussen verschillende taakgebieden. Dergelijke verschillen bepalen welke groep actoren welke vaardigheden nodig heeft om hun functie in het systeem te vervullen. Een voorbeeld is de taakverdeling tussen onderzoekers en informatiebeheerders. Bibliothecarissen en archivarissen ordenen informatie volgens gevestigde bibliografische categorieën, en voorzien onderzoekers daardoor van rechtstreeks bruikbaar empirisch materiaal, bibliotheekcatalogi en andere hulpmiddelen. Digitaal onderzoek betekent echter vaak een verschuiving in deze relaties – de verschillende taken die noodzakelijk zijn om een

wetenschappelijke databank te onderhouden kunnen bijvoorbeeld niet zo makkelijk worden onderverdeeld in informatiebeheer en - analyse. Mijn tweede onderzoeksvraag richtte zich op de epistemische gevolgen daarvan: Hoe onderscheiden digitale geesteswetenschappers vormen van arbeid als 'wetenschappelijk', 'technisch' of 'ondersteunend', en hoe maakt de verdeling van werk bepaalde vormen van kennis mogelijk? Een groot deel van de al besproken controverse rond de BNTL richtte zich op de vraag hoe de verantwoordelijkheid voor het beheer van Neerlandistische kennis verdeeld moet worden - is het een taak die de Koninklijke Nederlandse Academie van Wetenschappen met publieke subsidies moet steunen, of kan het met recht worden beschouwd als een van de kerntaken van wetenschappers ? De COST Action die ik in hoofdstuk 2 analyseerde illustreert verder de intellectuele implicaties van dergelijke kwesties. Een kernobstakel was de noodzaak om substantieel bij te dragen aan de harmonisatie en controle van gegevens in de database. Deze inzet was nodig voordat de deelnemers grootschalige, vergelijkende kennisclaims over vergeten schrijfsters konden ontwikkelen. Echter, veel deelnemers - meestal werkzaam als hoogleraar of universitair docent - beschouwden datawerk als een ondergeschikte technische activiteit, vergelijkbaar met bibliografisch werk, die hen afleidde van hun 'echte' wetenschappelijke verantwoordelijkheden. De definitie van datawerk als een niet-intellectuele activiteit was ook een belangrijk argument voor de grootschalige delegatie ervan aan extra aangetrokken student-assistenten, die bijvoorbeeld hun onderzoekstage of masterscriptie met werk in het project combineerden. Maar in tegenstelling tot bibliografisch werk zijn de algemene vaardigheden en intellectuele principes voor datagedreven historisch onderzoek niet ver ontwikkeld. De enige manier om ze te verwerven is door daadwerkelijk te experimenteren met databases, algoritmische methoden en visualisatiegereedschappen, etcetera. Als het schijnbaar ondergeschikte datawerk wordt gedelegeerd, verdwijnt ook de gelegenheid om huidige intellectuele paradigma's met digitale benaderingen aan te vullen. De COST Action worstelde inderdaad tot aan het einde met het probleem dat vaardigheden duidelijk ongelijk waren verdeeld. Terwijl de student-assistenten snel mogelijkheden vonden de database met hun eigen onderzoek te combineren, bleef dit voor wetenschappers met hogere functies moeilijk.

Populaire uitbeeldingen van digitaal onderzoek, bijvoorbeeld geassocieerd met modewoorden zoals de 'data deluge' en 'big data', suggereren vaak dat een rechtstreekse wetenschappelijke kenniswinst mogelijk is door het gebruik van een overvloed aan makkelijk beschikbare

gegevens. Mijn analyse onderstreept daarentegen de aanzienlijke inzet die in de digitale geesteswetenschappen noodzakelijk zal zijn om informatie compatibel te maken, zowel technisch en conceptueel als qua betrouwbaarheid. Verder blijkt dat de definitie van activiteiten die nodig zijn om gegevens bruikbaar te maken als ofwel een 'echte' intellectuele ofwel een ondergeschikte technische taak vaak op een tactische manier door actoren wordt gebruikt. Wetenschappers kunnen typen werk als technisch definiëren wanneer het erg verschilt van de disciplinaire activiteiten waarvoor ze hoofdzakelijk worden beloond, en beleidsmakers kunnen werk als wetenschappelijk definiëren om vermindering van publieke subsidies te rechtvaardigen. Deze dynamiek is bijzonder relevant omdat de constructie van bepaalde vormen van arbeid direct is verbonden aan de reproductie van basale onderzoeksvaardigheden en de conceptuele structuur van wetenschappelijke kennis.

Tot nu toe heb ik besproken hoe nieuwe instrumenten door verschillende infrastructurele beperkingen worden gevormd, maar zonder expliciet uit te werken in hoeverre dit specifiek disciplinaire verschijnselen zijn. Deze uitwerking is essentieel omdat het ons inzicht zal/kan geven in hoe digitale hulpmiddelen en infrastructuur aan de behoeften van specifieke domeinen aangepast moeten worden. Welke rol speelt disciplinariteit in het vormen van digitale onderzoeksinstrumenten, en hoe beïnvloedt hun gebruik disciplinariteit? Hoofdstuk 2 laat zien hoe de disciplinaire configuratie van literatuurgeschiedenis, gekenmerkt door een lage mate van sociale en intellectuele integratie, specifieke obstakels opwerpt voor het oorspronkelijk beoogde gebruik van de database. Gewoonlijk hebben onderzoekers een relatief grote speelruimte bij het uitstippelen van hun specifieke onderzoeksobjecten en theoretische benaderingen. Dit maakte het moeilijk om individuele praktijken binnen het project met elkaar te verenigen, ondanks een sterke overeenstemming over het algemene onderzoeksdoel. Theoretische problemen die normaal het voorwerp van constante wetenschappelijk discussie zijn, bijvoorbeeld de definitie van complexe concepten zoals genre en receptie, moesten binnen het project in consensus opgelost worden. Een interessante ontwikkeling was dat projectwerk gemakkelijker met de traditionele disciplinaire werkorganisatie te combineren was in het geval van kleinere deelprojecten, zoals een initiatief van een groep van Servische wetenschappers. Door het kleinere aantal deelnemers, evenals een meer afgebakende empirische focus, schiep het deelproject een context waarin de organisatie van databasegedreven onderzoek minder radicaal van de disciplinaire model verschilde, en

daarom bood deze ook betere voorwaarden voor de verspreiding van digitale vaardigheden.

Disciplinariteit speelde ook een belangrijke rol in de in hoofdstuk 3 onderzochte case study. Hier werkten indonesianisten, netwerkonderzoekers en informatici samen in een project met als doel verschuivingen in de netwerken van Indonesische elites te bestuderen door de algoritmische analyse van gedigitaliseerde krantencorpora. Een belangrijk discussiepunt in het project was de oorspronkelijke veronderstelling dat de hermeneutische onderzoeksvragen van de indonesianisten (bijvoorbeeld, wie en wat zijn sociale elites, en hoe oefenen ze macht over andere groepen uit?) in een empirisch probleem vertaald kunnen worden dat op zijn beurt met gegevens uit gedigitaliseerde krantenartikelen te beantwoorden is. Dit paste bij de aangestuurde taakverdeling, waarna de informatici als de producenten van de data fungeerden, terwijl de geesteswetenschappers zich met de interpretatie ervan bezig zouden houden. Dit model leverde echter problemen voor de indonesianisten op. Uit presentaties op academische bijeenkomsten werd duidelijk dat disciplinaire collega's een herformulering van hermeneutische vragen als empirische problemen weinig overtuigend vonden, vooral afgemeten tegen de conceptuele verfijning van dominante sociaalwetenschappelijke theorieën. Het was daarom noodzakelijk om met alternatieve manieren van samenwerking tussen individuele onderzoekers te experimenteren, om te voorkomen dat individuele disciplinaire praktijken simpelweg die van andere discipline koloniseren - bijvoorbeeld door te veronderstellen dat het concept van 'data' zoals gebruikelijk in netwerkonderzoek en informatica gewoon kan worden omgezet naar de indonesianistiek.

Digitale projecten worden dus vaak gekenmerkt door een zekere botsing tussen disciplinaire verwachtingen aan wetenschappelijke kennisbijdragen (bijvoorbeeld wat betreft de vorm van kennisclaims of theoretische oriëntatie) en het aanvankelijk voorgestelde nut van een bepaalde instrument. Als de project-deelnemers in staat zijn doorlopend subsidies zeker te stellen, kunnen zulke onconventionele praktijken zich verder ontwikkelen, en uiteindelijk misschien de banden met de oorspronkelijke disciplines doorsnijden. Maar in veel gevallen zullen wetenschappers hun oorspronkelijke onderzoeksgebied niet willen opgeven, waardoor het noodzakelijk wordt om conflicten tussen disciplinaire verwachtingen en projectwerk op te lossen. Dit kan betekenen dat centrale veronderstellingen heroverwogen moeten worden, bijvoorbeeld de

overtuiging dat digitaal onderzoek 'exacte' empirische antwoorden op 'rommelige' hermeneutische vragen moet geven, of dat digitale projecten altijd van grote schaal profiteren.

Hoewel bovenstaande bevindingen vaak disciplinaire of praktijkgerelateerde factoren in de vorming van de digitale geesteswetenschappen benadrukken, betekent dit niet dat de configuratie van digitaal onderzoek slechts het mechanische gevolg is van elkaar overlappende beperkingen. Wetenschappers maken vaak bewuste keuzes als ze conflicten aanpakken die ontstaan door de incongruentie tussen de infrastructurele status quo en nieuwe technologische mogelijkheden. Deze keuzes kunnen soms leiden tot een hervorming van de infrastructuur en zo bepaalde werkvoorwaarden scheppen voor toekomstige generaties van onderzoekers. Wat is de reflexieve activiteit van wetenschappers in de inbedding van nieuwe instrumenten in hun infrastructurele werkomgeving? Deze onderzoeksvraag beantwoord ik in hoofdstuk 4, waar ik de opvallend reflexieve stijl analyseer die veel publicaties van digitale geesteswetenschappers kenmerkt, zowel traditionele peer-reviewde stukken als nieuwe online vormen. Deze publicaties besteden expliciet aandacht aan de specifieke omstandigheden waaronder ze zelf zijn ontstaan, bijvoorbeeld door het combineren van onderzoeksredeneringen met overwegingen over hoe nieuwe samenwerkingsformaten, specifieke financieringsstructuren, of publicatievormen de inhoud van digitaal onderzoek beïnvloeden. Dit selectief zichtbaar maken van concrete praktijken is een poging om gestandaardiseerde disciplinaire werkroutines expliciet en opnieuw vormbaar te maken. Sommige reflexieve argumenten beogen bijvoorbeeld de status van digitaal werk te veranderen – van iets dat vooral als een dienst voor disciplinaire onderzoekers wordt beschouwd naar een activiteit die belangrijke intellectuele consequenties heeft. Andere redeneringen zijn erop gericht digitale geesteswetenschappers met de organisatorische en bestuurlijke kennis uit te rusten om complexe projecten af te sluiten in de context van een sober financieringsbeleid en een infrastructuur die niet is afgestemd op interdisciplinaire samenwerking. Het hertrekken van de grenzen tussen de zichtbare en onzichtbare aspecten van onderzoekspraktijken wordt regelmatig onderwerp van controverse, omdat? verschillende vormen van reflexiviteit vaak een bepaalde politieke geldigheid hebben. Sommige onderzoekers bekritiseren bijvoorbeeld hardvochtig de pragmatische, bestuurlijke manier waarop anderen over de digitale geesteswetenschappen nadenken. Zij zien dit als expressie van een a-politieke visie op de maatschappelijke functie van de wetenschap,

namelijk een die het kritische commentaar op de sociaaleconomische omstandigheden van onderzoek en hoger onderwijs heeft opgegeven.

Alle bovengenoemde onderzoeksvragen hebben met elkaar gemeen dat ze zich richten op de inbedding van digitale tools in wetenschappelijke praktijken op het niveau van een bepaald project. Echter, dit zegt nog weinig over hoe de economische, organisatorische en politieke voorwaarden voor individuele projecten zelf vorm krijgen door overkoepelende, strategische infrastructuurinitiatieven die op dit moment plaatsvinden in Europa en de VS. Mijn vijfde onderzoeksvraag werpt licht op deze interactie vanuit een vergelijkend perspectief: Hoe wordt infrastructuur in verschillende landen geconceptualiseerd, en welke rol spelen dergelijke conceptualiseringen in het organiseren van digitaal onderzoek op de werkvloer? In traditionele vormen van onderzoek zijn keuzes over instrumentenontwikkeling grotendeels gereserveerd voor disciplinaire elites. Veel nationale wetenschapssystemen zijn echter onlangs een proces ingegaan waarin de mechanismen van disciplinaire zelfbestuur veranderen. Huidige infrastructuur-initiatieven illustreren deze ontwikkeling in zoverre ze gebaseerd zijn op strategische allianties tussen beleidsmakers, financiers en onderzoekers. Vanuit heel heterogene belangen worden deze groepen actoren toch bij elkaar gehouden door gedeelde visies van toekomstige digitale infrastructuur. Omdat verschillende definities ook specifieke manieren van financiering en coördinatie van instrumentontwikkeling impliceren, is het van strategisch belang voor individuele actoren om een visie te bevorderen die bij hun respectievelijke prioriteiten past.

Amerikaanse digitale geesteswetenschappers hebben met succes gepleit voor een sociotechnische visie van infrastructuur als een emergent, evolutionair verschijnsel, waarna conceptuele en bestuurlijke autoriteit bij gevestigde digitale geesteswetenschappelijke centra moet liggen. Terwijl dit problemen met de implementatie van nieuwe instrumenten in traditionele onderzoekspraktijken omzeilt, bevoordeelt deze benadering de intellectuele en technologische keuzes van bestaande elites binnen digitale geesteswetenschappen. Europese initiatieven zijn daarentegen gebaseerd op een meer centraliserende, technologie-gedreven visie van digitale infrastructuur. Dit is in overeenkomst met de beleidsdoelstelling van de Europese Commissie, die is gericht op de integratie van nationale onderzoekssystemen. Dit veroorzaakt een zekere wrijving tussen ontwikkelaars van instrumenten en potentiële wetenschappelijke gebruikers die vaak niet bekend zijn met digitale benaderingen. Aan de andere kant zorgt de nadruk op centrale coördinatie ervoor dat geen enkele disciplinaire

gemeenschap de hoofdzakelijke controle over technologische ontwikkeling krijgt.

Curriculum Vitae

Wolfgang Kaltenbrunner was born on September 9, 1981, in Vöcklabruck, Austria. After completing his secondary education at the HTBLA Vöcklabruck (Federal Higher Technical Institute) in 2001, he moved to Vienna to study Comparative Literature. He obtained his degree with a thesis on science as literary motif in 2007. Subsequently, he successfully applied for the Huygens scholarship from the Netherlands Universities Foundation for International Cooperation, which allowed him to enroll in the Research Master program Cultures of Arts, Science, and Technology at Maastricht University. This program also took him to Amsterdam for an internship in the Virtual Knowledge Studio for the Humanities & Social Sciences. Upon finishing his second Master's degree with a thesis on positivist currents in 19th century literary scholarship, Wolfgang began his doctoral research at Erasmus University Rotterdam in 2009. He briefly interrupted his work in 2011, but returned to finish his thesis at the Center for Science & Technology Studies at Leiden University later that year. Since December 2014, he has worked as postdoctoral researcher in the same institution. Wolfgang has published a number of articles, inter alia on the history of the humanities, and on the implications of digital research tools for scholarly knowledge production.