



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

## 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.<sup>52</sup> 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.<sup>53</sup>

---

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 as 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).”<sup>54</sup> 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 ESFRI 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.<sup>55</sup>

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.<sup>56</sup> 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.<sup>57</sup> 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 renamed 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 Dombrowski (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.

