Malte Rehbein: Historical Network Research, Digital History, and Digital Humanities

1. The Scholarly Landscape of Digital Humanities

This essay is about three scholarly paradigms: Historical Network Research (HNR), Digital History, and Digital Humanities (DH). Albeit their grounding ideas, they are not new[1] but all gain new quality, are hence topical and currently discussed widely within the academic world and beyond. Especially the expectations which are tied to them, and the promises they make to reach new territory in research, lead to broad and general attention. On some of these promises, however, opinions differ. On the extreme sides of this dispute, technological ideologists and critical Luddites hardly find a common ground.[2] A situation like this appears to be not untypical in academia when new methodologies, such as proposed under the umbrella of these three paradigms, are introduced.

Beyond a dispute between enthusiasm and strict denial, there are several similarities among HNR and the two DH'es. And they are interconnected in a way that one might easily regard one as the part of the other: HNR is Digital History, Digital History is Digital Humanities,[3] and Digital Humanities embraces (historical) network research.

The overarching Digital Humanities[4] have been defined by Manfred Thaller as the sum of all

[1] Commonly, the year 1947, when Roberto Busa started his seminal work on a word concordance of the works of Tomas Aquinas aided by IBM calculators and punch cards, is regarded as a starting point for what is now known as Digital Humanities (Susan Hockey, "The History of Humanities Computing," in *A Companion to Digital Humanities*, ed. Susan Schreibman, Ray Siemens and John Unsworth, Blackwell companions to literature and culture 26 (Oxford: Blackwell, 2004)). Computational approaches in the Humanities, however, can be traced back further, at least to the 19th century when statistical stylometric calculations where undertaken manually (Manfred Thaller, "Geschichte der Digital Humanities," in Jannidis; Kohle; Rehbein, *Digital Humanities*).

[2] See for instance Matthew K. Gold and Lauren F. Klein, eds., *Debates in the digital humanities* (Minneapolis: Univ. of Minnesota Press, 2012) and Matthew K. Gold and Lauren F. Klein, eds., *Debates in the Digital Humanities 2016* (Minneapolis: University of Minnesota Press, 2016), http://dhdebates.gc.cuny.edu/debates/2. See also Helle Porsdam, "Digital Humanities: On Finding the Proper Balance between Qualitative and Quantitative Ways of Doing Research in the Humanities," *Digital Humanities Quarterly* 7, no. 3 (2013), http://www.digitalhumanities.org/dhq/vol/7/3/000167/000167.html.

[3] Surprisingly, however, the communities of Digital Humanities and of Digital History do not mix very well. The Call for Papers for the DH conference of the Nordic Countries, published in August 2017, describes and addresses this phenomenon which can be observed on an international as well on national levels (e.g. Germany) as follows: "While the number of researchers describing themselves as digital historians is increasing, computational approaches to history have rarely captured the attention of those without innate interest in digital humanities. To address this, we particularly invite presentations of historical research whose use of digital methods advances the overall methodological basis of the field" <http://eadh.org/news/2017/09/01/cfp-digital-humanities-nordic-countries-0>.

[4] A lot of older as well as recent publications try to outline and illustrate the scope of Digital Humanities, such as Susan Schreibman, Ray Siemens and John Unsworth, eds., *A Companion to Digital Humanities,*

attempts to apply information technology to the objects of Humanities scholarship. Further on, he points out that, firstly, the application of such technology serves the purposes of achieving results that would cause otherwise unreasonably high effort and secondly, that this process must be testable intersubjectively in order to have an epistemic value.[5] Digital Humanities as a whole principally embraces many if not all disciplines within the Humanities such as language and literary studies, philosophy, archaeology, art history, and the like as well as potentially Social Sciences. Within this realm, Digital History[6] might be regarded as the subset of Digital Humanities that intersects with the study of history.

The sum of *all* applications of information technology to the Humanities is rather large as not only many disciplines within the Humanities are integrated but also a broad methodological variety is offered.[7] Various disciplines from sciences and engineering, not exclusively Computer Science are included under this umbrella. Even if restricted to the subset of Digital History, the sum of applications ranges from simply accessing a digital surrogate of an archival entity at one's home desk, to digitally editing a mediaeval manuscript in its entire tradition, to the development of a complex procedure in order to automatically classify historical photographs employing artificial neural networks. To better cope with this enormous range of Digital Humanities, I propose a twofold definition:

Firstly, DH encompasses all kind of research in the Humanities that gains its findings (partly) from applying computer-based procedures, practices, and tools. In this understanding, Digital Humanities is pure Humanities scholarship, as its objects and questions are those from the Humanities.

Blackwell companions to literature and culture 26 (Oxford: Blackwell, 2004), Willard McCarty, *Humanities computing* (Basingstoke England, New York: Palgrave Macmillan, 2005), http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10254194, and John Unsworth, Raymond G. Siemens and Susan Schreibman, eds., *A new companion to digital humanities,* Blackwell companions to literature and culture 93 (Chichester, West Sussex, UK, Malden, MA, USA: John Wiley & Sons Ltd, 2016), http://onlinelibrary.wiley.com/book/10.1002/9781118680605. For a brief overview see also Malte Rehbein, "Was sind Digital Humanities?," *Akademie Aktuell* 56, no. 1 (2016). As Digital Humanities has begun to appear on the universities' curricula especially in the German speaking countries, a textbook introduces basic DH-methods: Fotis Jannidis, Hubertus Kohle and Malte Rehbein, eds., *Digital Humanities: Eine Einführung* (Stuttgart: J.B. Metzler Verlag, 2017).

[5] Manfred Thaller, "Grenzen und Gemeinsamkeiten: Die Beziehung zwischen der Computerlinguistik und den Digital Humanities" (DHd 2014, Passau, March 27, 2014).

[6] Fundamental for discussing the connotation of related terms: Onno Boonstra, Leen Breure, and Peter Doorn, "Past, present and future of historical information science," *Historical Social Research / Historische Sozialforschung* 29, no. 2 (2004). Particularities of historical research within the realm of Digital Humanities are discussed in Stephen Robertson, "The Differences between Digital Humanities and Digital History," in *Debates in the Digital Humanities 2016*, ed. Matthew K. Gold and Lauren F. Klein (Minneapolis: University of Minnesota Press, 2016). See also Peter Haber, *Digital Past: Geschichtswissenschaft im digitalen Zeitalter* (München: Oldenbourg, 2011) for a wider contextualization of Digital History within the landscape of historical scholarship.

[7] Lorna Hughes, Panos Constantopoulos, and Costis Dallas, "Digital Methods in the Humanities: Understanding and Describing their Use across the Disciplines," in Unsworth; Siemens; Schreibman, *A new companion to digital humanities*. The authors categories network analysis as one of seven "research activities" of the type "analysis" (ibid.: 157).

Secondly, DH encompasses the design, development, and generalization of these computer-based procedures, practices, and tools as well as the study of their underlying theories and models. In this understanding, Digital Humanities is rather an auxiliary science (Hilfswissenschaft)[8] located at the intersection between Humanities and Computer Science. Since its epistemological interest is particularly grounded in the functional question of this intersection, DH possesses, hence, its own objects and questions to study.

In case of the second viewpoint of DH, one might also speak of an "Information Science for the Humanities" or "Humanistic Informatics" (Aarseth). For the realm of historical research, I propose to use the term "Historical Information Science" (HIS) which has been suggested by others earlier on, also in various translations.[9] Onno Boonstra, Leen Breure, and Peter Doorn define Historical Information Science as "the discipline that deals with specific information problems in historical research and in the sources that are used for historical research"[10] and characterize it as "a science of its own, with its own methodological framework".[11]

With the twofold definition of the overarching Digital Humanities in mind, Historical Information Science serves the second aspect of this definition while Digital History supports its first. Taking into account other disciplines of the Humanities which might or might not developed their own specific information sciences, the scholarly landscape of DH can be summarized as in Fig. 1.[12] This model can be read from two directions. From the outer level, its starting point is a specific historical research problem which can be solved better or only (as in Thaller's definition) by applying computer technology (Digital History). Here, HIS serves as an auxiliary science as it provides suitable tools and practices. From the inner, core level, however, the starting point is a methodological problem for which its applicability and generalization in historical research is to be found.[13] Digital Humanities in its generally accepted definitions encompasses the whole model including both viewpoints.

[8] I argue elsewhere for the notion of DH as a "transferring science" (Transferwissenschaft) which "mission" is to evaluate, adopt, develop, and generalize procedures and tools until they are ready-to-use within the Humanities and can be transferred into their methodological canon. This includes also setting up DH curricula and training programmes Malte Rehbein, "Geschichtsforschung im digitalen Raum. Über die Notwendigkeit der Digital Humanities als historische Grund- und Transferwissenschaftwissenschaft," in *Papstgeschichte des hohen Mittelalters: digitale und hilfswissenschaftliche Zugangsweisen zu einer Kulturgeschichte Europas*, ed. K. Herbers and V. Trenkle (2017).

[9] Especially in German, Russian, Dutch as Historische Fachinformatik, istoricheskaya informatika, historische informatiekunde.

[10] Boonstra, Breure and Doorn, "Boonstra et al. 2004," 20.

[11] Ibid. 10.

[12] This model is inspired by Patrick Sahle's "spheres" (Patrick Sahle, "DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities." *DARIAH-DE Working Papers* 1 (2013): 6, http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2013-1.pdf). It is not yet complete.

[13] Boonstra, Breure and Doorn, "Boonstra et al. 2004," 20.

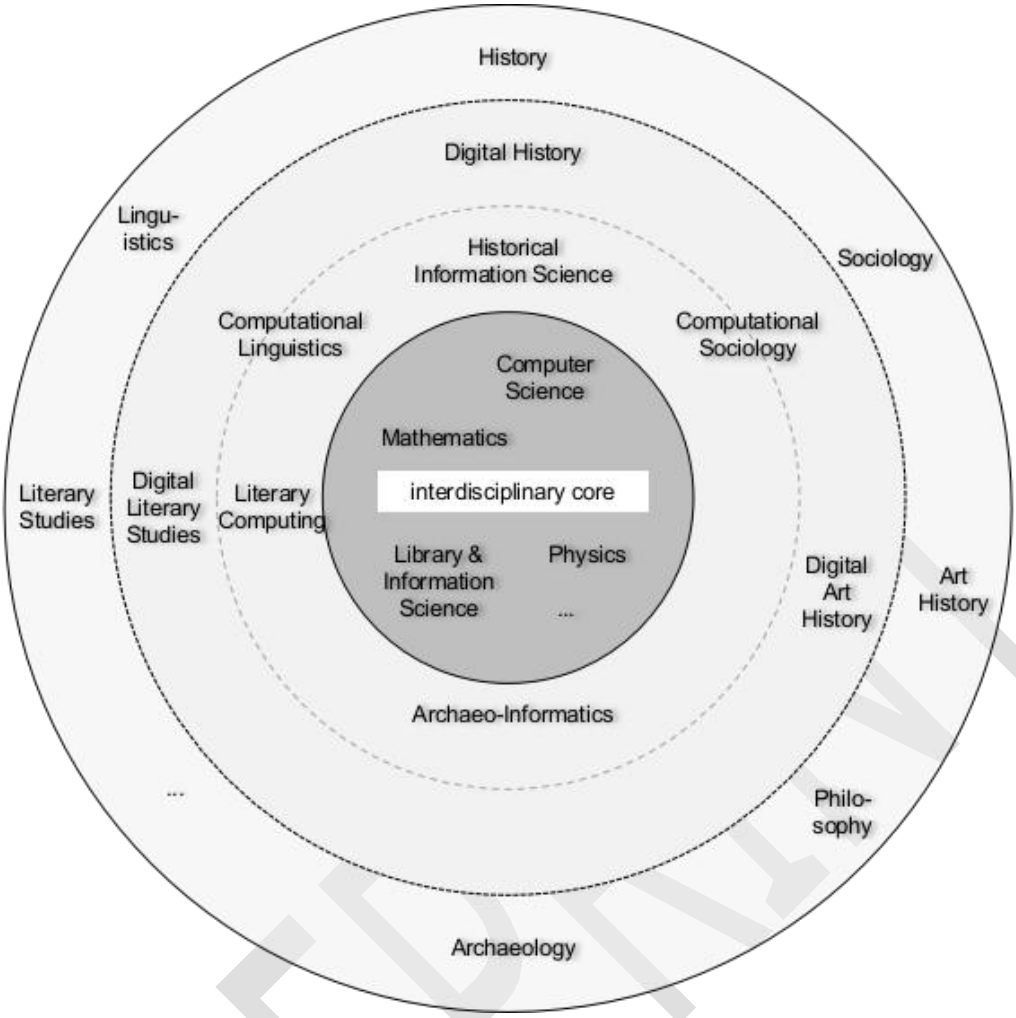

Figure 1. The scholarly landscape of Digital Humanities. An incomplete attempt.

2. Digital Humanities and Network Research

It has been pointed out often[14] that the term "network" (as in HNR) is used in a twofold manner: as a widely used metaphor on the one hand (especially in the contexts of a network society[15], "hyper connectivity"[16], media convergence, and the global internet)[17] and in the

[14] Cf. especially Marten Düring et al., eds., *Handbuch Historische Netzwerkforschung: Grundlagen und Anwendungen,* Schriften des Kulturwissenschaftlichen Instituts Essen (KWI) zur Methodenforschung 1 (Berlin: LIT-Verlag, 2016).

[15] Manuel Castells, *The Rise of the Network Society,* The Information Age: Economy, Society and Culture 1 (Cambridge, MA, Oxford, 1996), Jan van Dijk, *The network society: Social aspects of new media* (London, Thousand Oaks, Calif: Sage Publications, 1999), http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=309855.

[16] Ian Goldin, *Divided Nations: Why global governance is failing, and what we can do about it* (Oxford, 2013), 5.

[17] Currently, the term "network" seems overloaded and used in an inflationary manner, also in history (Marten Düring et al., "Einleitung," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische*

stricter sense, as formalization of semantically annotated nodes and edges grounded on the mathematical graph theory on the other.[18] These semantics may describe communication and interaction among people in which case we are speaking of a social network,[19] but the concept of network research can be applied to other semantics as well.

HNR[20] is understood as a transdisciplinary research paradigm in which a set of methods and techniques, adopted from other disciplines are applied to study social (and other) networks of the past and to embed them into larger contexts of historical research.[21] HNR can principally be applied without using computer technology (and has been done so since its beginnings).[22] However, in order to fully exploit its potential, computer-based procedures and tools, such as for managing and visually analyzing large datasets have become part of the HNR paradigm. Here, at the latest, HNR intersects with HIS, Digital History and DH.

There exists an intersection among these, seen from the perspective of DH, too: networks and network research appear to be of growing interest in the international community of DH.[23] As

Netzwerkforschung, 5). There are cases, however, in which the metaphoric is turned around: the North-American network of people and facilities to help escaped slaves to freedom in Canada since the late 18th Century was called "the underground railway", an expression coined in the 1830s. Here, "railroad" is used as a metaphoric expression to illustrate the activities of the network: "Various routes were lines, stopping places were called stations, those who aided along the way were conductors and their charges were known as packages or freight" (History.com staff, "Underground railroad," http://www.history.com/topics/black-history/underground-railroad).

[18] There seem to be indeed little restrictions in where to apply this research paradigm, see e.g. Milan Janosov, "Network Science Predicts Who Dies Next in Game of Thrones," https://cns.ceu.edu/article/2017-07-08/network-science-predicts-who-dies-next-game-thrones.

[19] In a common definition by James Clyde Mitchell, founding member of the International Network for Social Network Analysis, a social network is a "specific set of linkages among a defined set of persons, with the additional property that the characteristics of these linkages as a whole may be used to interpret the social behavior of the persons involved" (James C. Mitchell, "The Concept and Use of Social Networks," in *Social Networks Urban Situations: Analyses of personal relationships in Central African towns*, ed. James C. Mitchell (Manchester1969), 2).

[20] I am referring to general definitions in Claire Lemercier, "Formale Methoden der Netzwerkanalyse in den Geschichtswissenschaften: Warum und Wie?," *Österreichische Zeitschrift für Geschichtswissenschaft* 23, no. 1 (2012) and Marten Düring and Linda von Keyserlingk, "Netzwerkanalyse in den Geschichtswissenschaften. Historische Netzwerkanalyse als Methode für die Erforschung von historischen Prozessen," in *Prozesse: Formen, Dynamiken, Erklärungen*, ed. Rainer Schützeichel and Stefan Jordan (Wiesbaden: Springer VS, 2015).

[21] This is supported by Thomas S. Kuhn's understanding of a research paradigm: a consensus-based set of procedures that defines what objects to be studied, the kind of questions and hypotheses that are to be asked or tested, how these questions are to be asked, and how findings shall be interpreted. In context of HNR, cf. Wilhelm Heinz Schröder who has characterized Historical Social Research as a transdisciplinary research paradigm (Wilhelm H. Schröder, "Die Buchreihe „Historisch-sozialwissenschaftliche Forschungen" als publizistisches „Flaggschiff" der quantitativen Historischen Sozialforschung in der „Pionierzeit"," *Historical Social Research / Historische Sozialforschung*, Supplement No. 18 (2006): 14).

[22] This is generally true for the wider field of historical computing (Boonstra, Breure and Doorn, "Boonstra et al. 2004," 18).

[23] The programme of the international "Digital Humanities" conference in Montréal 2017 lists 15 presentations and posters that discuss phenomena of literary, social or historical networks. The keyword "network" is listed 47 times, albeit including technical terms such as neuronal networks or network infrastructure (McGill University & Université de Montréal, *Digital Humanities 2017: Conference Abstracts*

an example, the widely received project Republic of Letters and its spin-offs visualize and analyze scholarly exchange of erudite men of the age of Enlightenment by representing correspondence metadata as network graphs.[24] Furthermore, there are several cases in which the mathematical foundation of network graphs are employed for historical studies by applying a different semantic annotation of nodes and edges than that of a social network. In one example, Charlotte Schubert describes the attempt to build a taxonomy of classical textual sources supported by a network in which the longitudinal reception of texts by other authors is represented.[25] In a different setting, the author of this essay studies the diachronic evolution of a continuously revised mediaeval text with the help of an interactive network graph. Here, nodes represent hypothetical stages of the text, edges represent hypothetical developments from one stage of the text to another, and a path from the (chronologically) first stage of the text to its last formulates a hypothesis of how the text evolved over time.[26] More examples in which network graphs are applied to historical research questions beyond HNR in its narrow definition can be found easily.

Apart from historical studies,[27] network research has been employed in Digital Humanities in manifold other manners, for instance in empirical literary studies[28]. Literary scholars also start to describe the configuration of characters in novels or in dramas to form hypotheses about typical patterns for a particular genre[29] by annotating characters as nodes and their co-occurrences or other interactions among them as edges within a network. In linguistics,

(2017)). The expressions “network research” or “network study” do, however, not appear. Networks were not addressed in the fundamental DH-Companions (Schreibman, Siemens and Unsworth, *Schreibman et al. 2004*) and (Unsworth, Siemens and Schreibman, *Unsworth et al. 2016*).

[24] <http://republicofletters.stanford.edu/>.

[25] Charlotte Schubert, “Die Visualisierung von Quellennetzwerken am Beispiel Plutarchs,” *Digital Classics Online* 2, no. 1 (2016), doi:10.11588/dco.2016.1.23825.

[26] Malte Rehbein, “Reconstructing the Textual Evolution of a Medieval Manuscript,” *Literary & Linguistic Computing* 24 (2009).

[27] Further examples include: B. Opheim, “Political Networks and Factions: Online Prosopography of Medieval Scandinavian Sagas,” *History and Computing* 12, no. 1 (2000), David M. Brown, Adriana Soto-Corominas, and Juan L. Suárez, “The preliminaries project: Geography, networks, and publication in the Spanish Golden Age,” *Digital Scholarship in the Humanities*, 2016, doi:10.1093/llc/fqw036, Cornell Jackson, “Using Social Network Analysis to Reveal Unseen Relationships in Medieval Scotland,” *Digital Scholarship in the Humanities*, 2016, doi:10.1093/llc/fqv070, F. Kimura et al., “Visualization of relationships among historical persons from Japanese historical documents,” *Literary and Linguistic Computing* 28, no. 2 (2013), doi:10.1093/llc/fqs045, and M. Bingenheimer, J.-J. Hung, and S. Wiles, “Social network visualization from TEI data,” *Literary and Linguistic Computing* 26, no. 3 (2011), doi:10.1093/llc/fqr020.

[28] See e.g. Peer Trilcke, “Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft,” in *Empirie in der Literaturwissenschaft*, ed. Philip Ajouri, Katja Mellmann and Christoph Rauen (Münster2013), Laura Mandell, “How to Read a Literary Visualisation: Network Effects in the Lake School of Romantic Poetry,” *Digital Studies / Le champ numérique* 3, no. 2 (2012) and Chloe Edmondson, “An Enlightenment Utopia: The Network of Sociability in Corinne,” *Digital Humanities Quarterly* 11, no. 2 (2017).

[29] Fotis Jannidis, “Netzwerke,” in Jannidis; Kohle; Rehbein, *Digital Humanities*.

graphs have long been used to describe relations between words (such as in WordNet).[30] In other works, networks of (co-) authorship or (co-) citation are studied.[31]

3. The Scholarly Landscape Revisited

Besides the methodological and content-based relationship, similarities among HNR and DH can be observed about other aspects of scholarship and academia as well. Current impact to HNR come from case studies, often undertaken by early-stage scholars who might be motivated by the desire to leave trodden paths and to occupy a niche in which they can excel. Bringing together and transferring these start-up-like, often explorative and experimental studies into a larger scholarly context by attempting to consolidate and generalize their (methodological) findings, is, however, not pushed forward by an institution or otherwise motivated extrinsically. It rather stems from the community itself, informally and virtually.[32] This is comparable to the situation of the DH until the early 2000s.[33]

Back to the scholarly landscape mapped earlier on, case studies of HNR can be classified as Digital History because the starting point of their investigation is a historical research question. The generalization of their methodological findings, on the other hand, can be seen as Historical Information Science. Similar to earlier developments of DH, however, in an emerging research paradigm such as HNR, applying methods can hardly be separated from

[30] Other examples in which graph based networks with a different semantic annotation are used for Humanities research include and illustrate the variety of this usage: Tom J. Lynch, "Social Networks and Archival Context Project: A Case Study of Emerging Cyberinfrastructure," *Digital Humanities Quarterly* 8, no. 3 (2014), Jamshid Tehrani, Quan Nguyen, and Teemu Roos, "Oral fairy tale or literary fake? Investigating the origins of Little Red Riding Hood using phylogenetic network analysis," *Digital Scholarship in the Humanities* 31, no. 3 (2016), doi:10.1093/llc/fqv016, Daniel Gamermann et al., "The Small-World of 'Le Petit Prince': Revisiting the Word Frequency Distribution," *Digital Scholarship in the Humanities*, 2016, doi:10.1093/llc/fqw005, Maciej Eder, "Visualization in stylometry: Cluster analysis using networks," *Digital Scholarship in the Humanities* 32, no. 1 (2017), doi:10.1093/llc/fqv061, Valentina Bartalesi et al., "Towards a Semantic Network of Dante's Works and Their Contextual Knowledge," *Digital Scholarship in the Humanities*, 2015, doi:10.1093/llc/fqv044, J. Hu and N. Wang, "Complex network perspective on graphic form system of Hanzi," *Literary and Linguistic Computing* 28, no. 4 (2013), doi:10.1093/llc/fqt057, Florian Kräutli and Matteo Valleriani, "CorpusTracer: A CIDOC database for tracing knowledge networks," *Digital Scholarship in the Humanities*, 2017.

[31] For a methodological overview, cf. Markus Gmür, "Co-citation analysis and the search for invisible colleges: A methodological evaluation," *Scientometrics* 57, no. 1 (2003). As an example for its implementation cf. Markus Eckl, "Von Forschungsteams zur Wissenschaftscommunity: Eine soziale Netzwerkanalyse der wissenschaftlichen Co-Autorenschaften in der Disziplin der Sozialen Arbeit zwischen 1980 und 2014," *Soziale Passagen* 8 (2016).

[32] Such as the website "Historical Network Research" maintained by members of the community: "a platform for scholars to present their work, enable collaboration and provide those new to network analysis with some helpful first information" (<http://historicalnetworkresearch.org/>).

[33] In the much larger area of Digital Humanities, it is still discussed whether DH shall form a scholarly discipline on its own or remain rather a "community of practice" with a low degree of institutionalization (for instance understood so by the sub community "Digital Medievalist", http://digitalmedievalist.org). "The digital humanities is what digital humanists do" Robertson, "Robertson 2016" might be a sufficient definition of such a community.

their critical assessment and contributions to their improvements. Many studies do both. HNF as a transdisciplinary research paradigm can, hence, be integrated into the model along the axis history -- digital history -- historical information science -- interdisciplinary core as shown in Fig. 2.

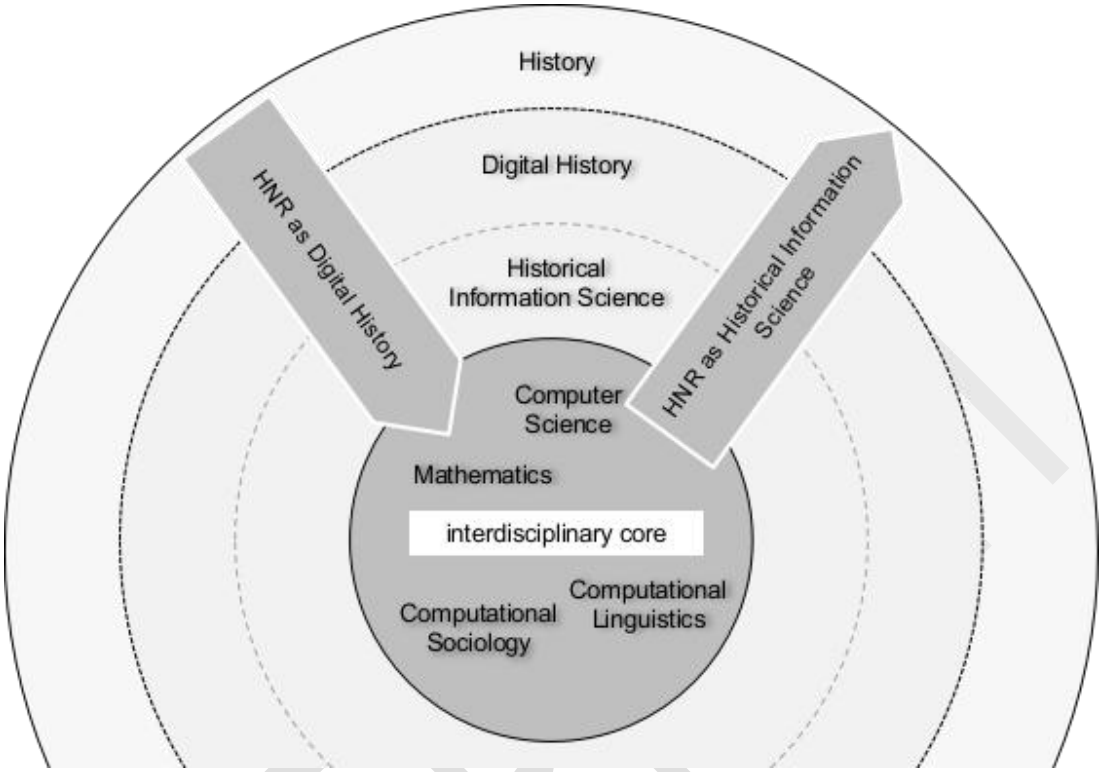

Figure 2. Two-dimensional view on HNR within the scholarly landscape of Digital Humanities. Note how Comutational Linguistics and Computational Sociology have moved into the interdisciplinary core.

At the time of this writing, DH has been earning wider recognition in academia. For this development, it might have been necessary to reach a critical mass of scholars and projects and especially to transfer their methods into the Humanities' canon,[34] but it was probably more the general, societal turn towards digitization in the beginning of this century that has drawn attention to this emerging area of research. As computational methods in the Humanities have been around for several decades, certainly only the availability of ready-to-use software will be a next major step forward towards general acceptance of these methods. Only then, these methods could be applied by the, as it were, masses of scholars who do not possess "arcane" knowledge such as computer programming, database development, or multimedia production.[35] A similar case in the academic past was that of

[34] Something that seems to have been quite successful in the area of digital editing and more so (but still on its way) in the computional method of authorship attribution (Patrick Juola, "Authorship Attribution," *Foundations and Trends in Information Retrieval* 1, no. 3 (2008), doi:10.1561/1500000005). See also the discussion in the final section of this essay.

[35] Cf. Boonstra, Breure and Doorn, "Boonstra et al. 2004," 19.

applying statistical methods to larger sets of data and hence bringing also quantitative history to a new level. It came along with the appearance of the Statistical Package for the Social Science (SPSS) software in 1968 which was marked, later on, by sociologist Barry Wellman as the "SPSS revolution"[36]. Wellman illustrates its importance: "Now, we do not have to be giants. We can be ordinary people, using statistical packages to play with data and examine hundreds of analytic possibilities".[37] For DH, comparable developments have been under way for a while, such as web-based text analysis tools, off-the-shelf software for stylometric studies, or customized packages for digital editing. However, a wide break-through has not been reached yet. A thorough understanding of the principles of computer technology is often still required by its applicants, and I am tempted to say, at least partly for the good of scholarship.[38] Another aspect that needs to be discussed further in this context is that of the relation between method and software. Here, one has to be careful to insist on the directional dependency between the two. Matthias Bixler's and Daniel Reupke's argument that only using a software transforms the metaphor of networks into a method[39] should be reversed: the rational of the method does not lie within the software, but the software's rational lies in the method: within the software, the method is operationalized.

Currently, HNR scholarship requires quite a large set of skills, most of which are not taught in regular curricula of history. This, too, is paralleled in Digital Humanities and Digital History and seen as a problem.[40] Apart from the ability to critically interpret "results" from a network visualization within its historical context -- which should stand above all -- scholars working under the paradigm of HNR need to familiarize themselves, each to more or less extent, with theories of social networks, their mathematical, statistical and quantitative foundations, graph theory, visual literacy,[41] layout algorithms including their strengths and restrictions, human

---

[36] Barry Wellman, "Doing It Ourselves: The SPSS Manual as Sociology's Most Influential Recent Book," in *Required reading: Sociology's most influential books*, ed. Dan Clawson (Amherst: University of Massachusetts Press, 1998), 73.

[37] Ibid: 74. In his famous quote: "L'historien de demain sera programmeur ou il ne sera plus" (Emmanuel Le Roy Ladurie, *Le territoire de l'historien,* Bibliothèque des histoires (Paris: Gallimard, 1973), 14), in which he referred to quantitative history only, Emmanuel Le Roy Ladurie did not foresee this development. For him, it was clear that only a deep understanding of the mechanisms behind software will empower the historian to fully exploit its potential and to avoid failures (cf. Manfred Thaller, "Geschichte der Digital Humanities," in Jannidis; Kohle; Rehbein, *Digital Humanities*).

[38] I argue in more detail elsewhere: Rehbein, "Rehbein 2017f".

[39] Matthias Bixler and Daniel Reupke, "Von Quellen zu Netzwerken," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 116–17.

[40] Cf. Martin Stark, "Netzwerkberechnungen. Anmerkungen zur Verwendung formaler Methoden," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 156, Malte Rehbein, "Digitalisierung braucht Historiker/innen, die sie beherrschen, nicht beherrscht," *H-Soz-Kult*, 27.11.2015. In general terms, as digital methods not only those of HNR are arising and getting more widely used, skills required to apply (and to assess and evaluate!) them have to be included in curricula of history, cf. e.g. Rehbein, "Rehbein 2017f", Malte Rehbein and Patrick Sahle, "Digital Humanities lehren und lernen: Modelle, Strategien, Erwartungen," in *Evolution der Informationsinfrastruktur: Kooperation zwischen Bibliothek und Wissenschaft*, ed. Heike Neuroth, Norbert Lossau and Andrea Rapp (Glückstadt: vwh Hülsbusch, 2013).

[41] This aspect cannot be overestimated. Anyone using graphics to communicate research results, a core

computer interaction, and, maybe most importantly, data modelling, formalization, and coding.

A key requirement to a wider acceptance of HNR methods, and this, too, can be observed equally in the broader field of DH, is the fact that generally, all these skills are likewise required by the recipients (i.e. readers) of research results. As Katja Mayer points out: production and dissemination of knowledge are interconnected.[42] This fact in turn requires the applicants of new research paradigms to sensitively write in favour of their audiences, who, as stated above, do not necessarily possess such skills as long as they have not become part of the curricula in history (or other disciplines within the Humanities).[43] However, as scholarly "knowledge" exists only through a mutual agreement among peers,[44] addressing peers and allowing them to test hypotheses and results themselves is an essential, critical step in scholarship. This might also be a necessity if one aims at convincing a broader community of historians about the "added value" that HNR brings into historical research which many might still deny.[45] Hence, to an extent, HNR, as well as Digital Humanities and Digital History, still exist only in particular niches and do not reach out as well as they could.[46]

4. Digital Humanities: An Accelerator for Historical Network Research

As we have seen, various intersections between the research paradigm of Historical Network

element of HNR, must be aware of the effect on the recipient of symbols and their arrangement. Cf. Katja Mayer, "Netzwerkvisualisierungen. Anmerkungen zur visuellen Kultur der Historischen Netzwerkforschung," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 147. For a general introduction see Edward R. Tufte, *The visual display of quantitative information* (Cheshire, Conn. (Box 430, Cheshire 06410): Graphics Press, 1983), for information visualization within DH: Malte Rehbein, "Informationsvisualisierung," in Jannidis; Kohle; Rehbein, *Digital Humanities*.

[42] Katja Mayer, "Netzwerkvisualisierungen. Anmerkungen zur visuellen Kultur der Historischen Netzwerkforschung," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 139.

[43] Cf. Marten Düring and Florian Kerschbaumer, "Quantifizierung und Visualisierung. Anknüpfungspunkte in den Geschichtswissenschaften," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 38 with references to Robert W. Fogel, Konrad Jarausch, and William O. Aydelotte. Arguing from the readers' perspective: N. Thomson, "How to Read Articles which Depend on Statistics," *Literary and Linguistic Computing* 4, no. 1 (1989), doi:10.1093/llc/4.1.6.

[44] Following Thomas S. Kuhn's notion of science (Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago, Ill.: University of Chicago Press; The University of Chicago Press, 1962)), as well as Friedrich Nietzsche's understanding of truth (cf. Wiebrecht Ries, *Nietzsche zur Einführung*, 7th ed. (Hamburg, 2004), 1020–24) and Lothar Kolmer, *Geschichtstheorien* (2008), 68).

[45] For the latter cf. Marten Düring et al., "Einleitung," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 7.

[46] Matthias Bixler, "Die Wurzeln der Historischen Netzwerkforschung," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 59.

Research and the scholarly area commonly named Digital Humanities exist. In this final section, I am arguing that current and past developments within DH can serve as an accelerator for HNR if applied thoughtfully. I am discussing four areas: digitization, automatic annotation, data linkage, and visualization and publication. These areas also partly mirror the four categories of "information problems" in historical research proposed by Boonstra, Breure, and Doorn.[47] A fifth and essential area, that of data modelling, is beyond the scope of this essay and should be addressed separately.

4.1. Digitization

In HNR, oversimplifying, data and information about historical actors, their behaviour and especially interaction among them is collected, annotated and modelled, then visually and/or statistically analyzed. Data and information required is taken directly or indirectly (mediated through editions or research literature) from historical sources which key role for the study of history can hardly be denied. Hence, getting access to sources is often a restricting factor in present day historical research.

Especially the 19th and 20th century has brought us large amounts of edited historical records. But this still is only an extremely small portion of all surviving sources from the past and what is more, as a legacy from historism, it often embraces only those records that represent the "big players" of the past. Yet, especially such types of historical records that are interesting for HNR, serial sources in particular and heterogeneous sources beyond the scope of "big history", are until today hardly transcribed, let alone edited. Tedious visits to archives, finding, reading, and annotating relevant information is hence required should important research questions not remain unanswered. For HNR, too, it has been observed that pragmatic considerations -- and not those of research interests or relevance -- prevail. Various authors point out that only in cases where sources are easily accessible as well as data and information from these sources can be extracted within reasonable time and effort, historical network research is possible or sensible.[48]

Digitization of historical records, that is the process of transforming them into digital representations or surrogates, is one of the corner stones of Digital Humanities.[49] It has been a research effort for long to put historical records efficiently and sustainably into digital

[47] Information problems of historical sources, of relationships between sources, of historical analysis, and of the presentation of sources or analysis Boonstra, Breure and Doorn, "Boonstra et al. 2004," 20–21.

[48] Cf. Christian Marx, "Forschungsüberblick zur Historischen Netzwerkforschung. Zwischen Analysekategorie und Metapher," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 65, (Robert Gramsch, "Zerstörte oder verblasste Muster? Anwendungsfelder mediävistischer Netzwerkforschung und das Quellenproblem," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 69), Matthias Bixler and Daniel Reupke, "Von Quellen zu Netzwerken," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 107.

[49] As well as it has been with the realms of library and information science and archival studies.

format, make them openly accessible, analyzable, and to preserve them for the future. Digitization of archival records follows a new paradigm and offers hence a new quality: their availability does not depend any longer on whether they have been in-depth edited and made available in print (usually with significant investment of time and money) nor on a personal visit to the archives. Instead, continuous digitization of archival records[50] and their availability in the virtual realm certainly brings new impact to HNR and might help overcoming the aforementioned restrictions in research so far seen as necessary due to pragmatic considerations.

Naturally, digitization cannot solve challenges arising from the characteristics of historical research such as incomplete or missing sources: what has not survived from the past cannot be digitized in the present and most likely neither in the future. But digitization is playing a more and more important role in cultural heritage in general and for historical records in particular. After the last "wave" of digitization projects and programmes focused on printed material and rare collections, other types of sources[51] are now gaining relevance in funding schemes, and those might be of high interest for HNR.[52]

While digitization obviously improves access to sources, it does, however, not necessarily help finding relevant information within them. Transforming digital surrogates into machine readable texts is a process comparable to that of palaeographic transcriptions. But in order to transcribe efficiently a large amount of digitized sources, the process has to be automated.[53] For printed texts, this problem has been solved to a good extent with Optical Character Recognition (OCR), albeit the general principle that the older the text, the lower the quality of text recognition, seems to remain valid. Automated recognition of handwritten texts (HTR), however, still is an interdisciplinary research challenge. Recent developments,[54] using,

---

[50] So-called retro-digitization of printed edition as well as of secondary texts have to be mentioned here, too, as both are of importance for HNR.

[51] To be mentioned, among others, are historical newspapers and serial sources such as account books and matriculation registers. As an example: the diocese of Passau, Germany, is digitizing its complete series of parochial records from the 16th to the 20th century and hence opening them up for genealogical research, social history, and potentially, HNR (cf. Robert Gramsch, "Zerstörte oder verblasste Muster? Anwendungsfelder mediävistischer Netzwerkforschung und das Quellenproblem," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*). As pioneering in terms of systematization, completeness and transparency to be mentioned: Monasterium.net <http://monasterium.net/> and Matricula <http://data.matricula-online.eu/de/>.

[52] Within a pilot "Archivalische Quellen" (archival sources), funded by the German Research Council DFG, efficient procedures have been developed, tested, and should be ready for broad application.

[53] Apart from automation, crowdsourcing has also been tried to transcribe (and annotate) historical records (Melissa Terras, "Crowdsourcing in the Digital Humanities," in Unsworth; Siemens; Schreibman, *A new companion to digital humanities*). There are manifold examples, including the "Transcribe Bentham" project (<blogs.ucl.ac.uk/transcribe-bentham/>) or the data entry project of the Danish censuses 1787-1880 (Nanna F. Clausen, "The Danish Demographic Database— Principles and Methods for Cleaning and Standardisation of Data," in *Population Reconstruction*, ed. Gerrit Bloothooft et al., 1st ed. 2015 (Cham, s.l.: Springer International Publishing, 2015)).

[54] The forthcoming results from the EU-funded project READ should be observed (<https://read.transkribus.eu/>).

among others, artificial neuronal networks[55] and deep learning, seem promising. We might have come closer to a goal desired (maybe) by many: to feed a computer with the digital image of a historical manuscript and to receive, without human interaction, machine readable data of the text that the manuscript mediates. Consequences for historical research would hardly be less immense: millions of handwritten historical records could be made accessible, not only for reading them at one's home desk but far more, for automatically searching within, annotating, linking and analyzing them in any thinkable way with the aid of computer-based procedures. Especially for the field of HNR, this would open vast new terrain.

Another aspect of digitization as a societal phenomenon of the present time will have future impact on HNR: "Born digital" data are becoming more and more relevant for historical research. For instance, with the advent of networked computers in the 1960s, electronic communication such as email has been gaining importance (and is now prevailing). It is already or will soon be object of study also for historians. Surprisingly, unwanted characteristics of historical data such as lost records or questions of authenticity seem to have survived in the digital realm in a similar manner. Archiving these records as well as other media of the World Wide Web such as websites is still an open issue. Due to its ephemeral nature, many of past electronic communication will not be available as historical sources in the future.[56]

With the help of automated text recognition, digitization opens historical research to a new paradigm in which the accurate reading of a historical source is not necessarily central for understanding the past any more. Instead, macroscopic analyses of a vast amount (and maybe also: variety) of sources might prevail under this paradigm. In literary studies which builds far more on printed and hence texts easily made machine-readable than historical studies, this paradigm has been (controversially) discussed for a couple of years under the headline "distant reading".[57] For many applications of this "distant reading" that are currently discussed, semantic enrichment of data and its in-depth annotation, essential for HNR, is, however, not a key requirement. For HNR, further steps, especially those of annotating information are necessary.

4.2. Automatic Annotation

While digitization is improving the availability of historical sources and automated text recognition is making them processable by a computer, further information has to be extracted and formalized systematically from these sources about historical actors, their attributes and

[55] A method of machine learning, motivated from biology as well as another example of using the word "network".

[56] Cf. Terry Kuny, "A Digital Dark Ages? Challenges in the Preservation of Electronic Information," *63rd IFLA Council and General Conference*, 1997.

[57] Franco Moretti, *Graphs, maps, trees: Abstract models for a literary history*, 1. publ (London: Verso, 2005).

relations in order to prepare for HNR.[58] Depending on the research design, this process can happen on one or both of two levels: on the level of metadata (e.g. in a correspondence network: who wrote whom)[59], and on the level of content (e.g. who wrote about who). Generally, the second level is far more challenging to address when it comes to automate this process.

For HNR, procedures to support this process lie within the realm of Historical Information Science. They encompass:

1. (Named) entity recognition (NER), that is annotating a textual string as a signifier for a specific type of information e.g. a person name;
2. Entity resolution, that is disambiguating this string in order to identify a unique member of that type, e.g. a specific person;
3. Complex relation extraction, that is extracting from the text whether and how two persons (or other types of entities) relate to each other; and
4. Contextualizing everything historically in space, time, and semantics.

All these steps underlie general characteristics of historical sources such as variety, ambiguity, vagueness, fuzziness, and scarcity or incompleteness. They have been studied extensively since the 1970s at the latest[60] but have not yet been solved satisfactorily, at least not on a generic level. Progress in data cleaning and annotation of this kind, especially employing methods and tools from Computational Linguistics, has been made in non-historical applications of network research of Digital Humanities. However, data from other disciplines often is less complex than historical data. In literary studies, for instance, extracting actors and their co-occurrence on stage can be automated more easily as the number of actors is limited, their names are known, stage instructions that define relations among actors of usually not open for information widely, and data is complete.[61] Again, one might think of a future machine that is fed with all available sources in machine-readable form together with a research question such as that of an ego-network of historical person X or of a trade network within a particular period. Yet, in annotating data in a way outlined above, profound historical understanding already plays a central role. The key to such historical understanding is

---

[58] Cf. Matthias Bixler and Daniel Reupke, “Von Quellen zu Netzwerken,” in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 104.

[59] The *Kalliope* database of personal papers and manuscripts is an example in which such metadata can be used for rudimentary network analysis (<http://kalliope.staatsbibliothek-berlin.de/>).

[60] Boonstra, Breure and Doorn, “Boonstra et al. 2004,” 35.

[61] This way, in an exemplary case study 465 German dramas between 1731 and 1929 have been automatically analysed in order to create and visually compare networks of characters over time (Frank Fischer et al., “Distant Reading Showcase: 200 Jahre deutsche Dramengeschichte auf einem Blick” (DHd 2016, Leipzig, March 09, 2016), http://dhd2016.de/sites/default/files/dhd2016/files/Distant_Reading_Showcase__465_German_Dramas__DHd2016_Poster.pdf).

contextualization and interpretation, and this is based on human experience (Karl Schlögel) and cannot be delegated to algorithms. Hence, concerns should be expressed against attempts to automatically generate historical networks. To follow the so-called "Big Data"-paradigm[62] and its "end-of-theory"-branch[63] is topical and tempting, but is, for the time being at least, rather misleading as long as it is not connected to theory.[64] The strength of automation in HNR lies, so far, mainly in supporting data preparation and any progress made here will be of benefit for HNR.

Various methods are currently tested on historical texts, structured historical data, and research literature. These attempts are mostly situated in the realm of Computer Linguistics and Natural Language Processing and include supervised and unsupervised Machine Learning.[65] Approaches close to HNR encompass automated coding of occupational strings,[66] place name disambiguation,[67] entity recognition and resolution in cultural heritage collections,[68] and relationship extraction from biographical information systems.[69] This kind of annotated data carries information about historical actors that may be useful for network research.

Speaking of relationship extraction in particular, however, one must be aware that several types of such relations exist for conceptualizing a network.[70] They range semantically from

---

[62] For a discussion of the use of the term "Big Data" in the Humanities cf. Christof Schöch, "Big? Smart? Clean? Messy? Data in the Humanities," *Journal of Digital Humanities* 2, no. 3 (2013), http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/.

[63] „Out with every theory of human behavior. [… The traditional] approach to science — hypothesize, model, test — is becoming obsolete" (Chris Anderson, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete," *Wired Magazine* 16.07 (2008)).

[64] One might, for instance, simply ask why different data sets yield different results for the same research questions. I am discussing this in greater detail on biographical data elsewhere (Rehbein, "Rehbein 2017f").

[65] An overview of state-of-the-art methods in automated processing of structured historical data on the example of genealogy provides Gerrit Bloothooft et al., eds., *Population Reconstruction*, 1st ed. 2015 (Cham, s.l.: Springer International Publishing, 2015).

[66] Ibid., Kevin Schürer, Tatiana Penkova, and Yanshan Shi, "Standardising and Coding Birthplace Strings and Occupational Titles in the British Censuses of 1851 to 1911," *Historical Methods* 48, no. 4 (2015).

[67] Ian Gregory et al., "Geoparsing, GIS, and Textual Analysis: Current Developments in Spatial Humanities Research," *International Journal of Humanities and Arts Computing* 9, no. 1 (2015), doi:10.3366/ijhac.2015.0135.

[68] Seth van Hooland et al., "Exploring entity recognition and disambiguation for cultural heritage collections," *Literary and Linguistic Computing*, 2013, http://llc.oxfordjournals.org/content/early/2013/11/29/llc.fqt067.full.pdf.

[69] Matthias Reinert et al., "From Biographies to Data Curation – the Making of www.deutsche-biographie.de," *Proceedings of the First Conference on Biographical Data in a Digital World*, 2015. The project currently focusses on genealogical relations and those of the type teacher-student.

[70] For a thorough and theoretical discussion from the viewpoint of sociology, cf. Mustafa Emirbayer, "Manifesto for a Relational Sociology," *The American Journal of Sociology* 103, no. 2 (1997), doi:10.1086/231209.

permanent genealogical to temporary relations such as that of employer-employee or an affair up to short-term interactions or punctual meetings, and they encompass complex and disputable notions as such of friendship as well as simple and less disputable co-occurrences of names within a certain range of text.

Currently, there is no general set of tools or process pipeline available to undertake annotating tasks for HNR. As research questions and source material are too heterogeneous,[71] and a broad variety exists of how such relations between people are expressed in the source text (if at all made explicit), it is probably not sensible to develop such tools at all. However, it can be foreseen that computer supported annotation can be used in HNR if tailored to specific types of sources and questions.[72] This kind of generalization on a meso level, that transcends single problems but does not promise "all-inclusive" solutions, can again be understood as a future research task for Historical Information Science and will significantly push forward HNR if transferred to a broader scholarly community.

4.3. Data Linkage

Another benefit that annotated data offer is their capability of instantly creating links to other data sources. With such a network of data, the scope of an HNR research question can potentially be extended. Furthermore, efforts to collect and annotate data can be shared among many. Hence, HNR can benefit from a general trend to make research and other data openly available in a way where relations between data points can be followed across the internet (Linked Open Data, LOD). At the same time, HNR researchers can also contribute to this idea of LOD by providing their data and annotations to the public.[73]

Some considerations have to be made, though. For LOD to work effectively, data has to be interchangeable, that is, it can be (re-)used within a different context. This is usually solved technically (and legally). For LOD to work efficiently (and to an extent, automatically), however, data has to be interoperable.[74] For instance, data points that have the same identifier must refer to the same entity (e.g. historical person) across the whole LOD-universe and vice versa, one entity cannot have different identifiers in different data sources.

---

[71] Christian Marx, "Forschungsüberblick zur Historischen Netzwerkforschung. Zwischen Analysekategorie und Metapher," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 66.

[72] Cf. Robert Gramsch, "Zerstörte oder verblasste Muster? Anwendungsfelder mediävistischer Netzwerkforschung und das Quellenproblem," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung* for a proposal how to categorize HNR research questions in mediaeval studies.

[73] This touches also the question of managing research data which is a general topic in DH and information science. Solutions from there can be reused also by HNR.

[74] For a definition (and distinction) between interchangeability and interoperability, cf. Syd Bauman, "Interchange vs. Interoperability," *Proceedings of Balisage: The Markup Conference 2011*, 2011, doi:10.4242/BalisageVol7.Bauman01, http://www.balisage.net/Proceedings/vol7/print/Bauman01/BalisageVol7-Bauman01.html.

Questions of interoperability and interchange of data have long been studied in Digital Humanities. Especially, the "flagship" data model of DH, that of the Text Encoding Initiative (TEI), a set of recommendations of tag and attribute names and their usages, has addressed them since the mid-1980s[75]. TEI is providing mechanisms to annotate (textual) data in a way that interchange among different data sources is facilitated. Efficient interoperability is still a challenge and promises seem to be unfulfilled, though.[76] Interoperability grounds, as a prerequisite, on a shared understanding and conceptualization of data, information and knowledge[77] and a consistent formalization among potentially heterogeneous sources, and this is difficult if ever to reach.

The same can be said about employing the ideas of LOD for HNR. LOD certainly facilitates data interchange and researchers can manually or semi-automatically draw data from other sources to include them in their own networks. It is still unclear, however, to what extent one can make use of data interoperability in HNR. As mentioned above, a shared conceptualization of data is a prerequisite. Evidently, unique and cross-platform identification of persons is a central part of this shared conceptualization as one has to be sure that the same person X is meant in different data sources. This problem has found its technical solution in authority files.[78]

An example how interoperability based on authority files may work beneficially for HNR is given by the Deutsche Biographie: a network of biographical data of about 740,000 persons from 19 distributed data sources through this mechanism. Thinking this concept to its end, historical relations could be detected and used for HNR that were otherwise not visible. If, for instance, one historical source witnesses a relation between A and B and another source a relation between B and C, then a relation between A and C (with B as a broker) can be inferred even if neither the first source witnesses anything about C nor the second about A. The critical evaluation of this (logical) inference is still the historian's task, though.

Reaching a shared understanding to uniquely identify persons is still a far easier task than that of reaching a shared understanding of what a relation is and what quality it has. Given the various types of possible relations (see above) and different, often subjective notions of relations such as "friendship" or different levels of relevance of a particular relation within different research questions, a general agreement will hardly ever be reached.

---

[75] Nancy Ide et al., "Proposal for Funding for an Initiative to Formulate Guidelines for the Encoding and Interchange of Machine-Readable Texts," http://www.tei-c.org/Vault/SC/scg02.html.

[76] Desmond Schmidt, "Towards an Interoperable Digital Scholarly Edition," *Journal of the Text Encoding Initiative* 7 (2014).

[77] Shared conceptualization is the idea behind "ontologies", proposed by Thomas Gruber (Thomas Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition* 5, no. 2 (1993)). Paired with semantic web technologies, ontologies are one of the current research areas in DH (Malte Rehbein, "Ontologien," in Jannidis; Kohle; Rehbein, *Digital Humanities*).

[78] Mostly commonly the Virtual International Authority File (VIAF), initiated by the Library of Congress and the German National Library and the Gemeinsame Normdatei (GND), maintained by the German National Library.

It is an open issue for the HIS-branch of Historical Network Research to deal with these problems. Questions need to be addressed such as: how to cope with different kinds of types for the same data that would lead to inconsistencies within the own network; how to deal with a categorization used in an external data source that does not fit to one's own research design; how to assess the quality of external data sources and what mechanisms exist in order to improve them or tailor them to specific needs; among different (heterogeneous) data sources, what is the greatest common divisor of a shared conceptualization? Some of these questions are discussed also within the TEI,[79] and a joint discussion would be of mutual benefit, in particular also as TEI-encoded data might serve as data sources for HNR.

4.4. Visualization and Publication

Digitization also brings new forms for communicating scholarly findings. They range from traditional publication formats within new business models (journal articles in Open Access) to rather new formats such as blogging and micro blogging. Scholarly work is beginning to be understood as a process that addresses the public not only at its very end. Findings are being published also in preliminary stages, and experimental formats such as collaborative writing[80] or social editing[81] are developed. The discussion about the future of scholarly publication and communication, another current topic of DH, encompasses also research data for later re-use as a scholarly output.

Scholars of HNR are also confronted with the possibilities (by some seen as opportunities, by others as threats) that new publication paradigms provide. The HNR-platform[82] is an example of a usage for communication and community building. For publishing historical findings, however, it can be observed that the traditional formats of scholarly articles and monographs including static graphics[83] are mostly chosen. As an emerging research paradigm, HNR itself is confronting a traditional scholarly community, that of historians, with a new and disputable methodology. It might be easier to gain acceptance from this community if new methodology is not combined with new forms of publication.

As an analogy to HNR, digital scholarly editing, one of the core areas of Digital Humanities (and Digital History), is about to complete its transformation into the Humanities'

---

[79] John Unsworth, "Computational Work with Very Large Text Collections: Interoperability, Sustainability, and the," *Journal of the Text Encoding Initiative* 1, Issue 1 (2011), doi:10.4000/jtei.215, http://jtei.revues.org/215, Bauman, "Bauman 2011".

[80] E.g. Terry Eagleton, *Marxism and literary criticism*, 2nd ed. (New York: Routledge, 2002), 63.

[81] Kenneth M. Price, "Social Scholarly Editing," in Unsworth; Siemens; Schreibman, *A new companion to digital humanities*.

[82] <http://historicalnetworkresearch.org/>.

[83] Cf. Katja Mayer, "Netzwerkvisualisierungen. Anmerkungen zur visuellen Kultur der Historischen Netzwerkforschung," in Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 149

methodological canon. As a consequence, digital editions might lose their attribute "digital" in the near future as it has become accepted and natural to create and publish scholarly editions digitally. This wide acceptance, however, has come to the cost of a compromise. Most digital editions mimic their printed predecessors more than they could (and should). The medium of publication is different (internet instead of a printed book), but the form remains more or less the traditional one, the one that the long-standing community has been used to use, hence the one that demands less effort to overcome resistance.

This compromise left out vast potential that computational methods offer for publication, e.g. interactive, dynamic visualizations. Peter Shillingsburg's statement from 2006 seem still to be valid: "we have not fully understood or exploited the capabilities of electronic texts."[84] This statement can be transferred to HNR. Especially in the area of Visual Analytics, HNR has the potential to make a large contribution to digital scholarship as well as to historical research when (thoughtfully) exploiting these capabilities, not only as a research tool but also as a way to communicate with the scientific community and lay people. What is required, however, is a progressive conversation about forms of scholarly communication and publication that go beyond academic papers and monographs.

Visualization and Visual Analytics are current trends in digital scholarship; they are not new, but computational methods offer a new quality to them. We must be aware, however, that in HNR, we are never visualizing the historical reality but a hypothetical model of it.[85] As this model is interpretative and subjective, there exists an endless number of potential other models of that reality. For instance, a historian can be uncertain whether a relation between actors A and B existed. Traditionally, it is expected from the researcher to make a decision, based on an accepted historical theory and methodology, or to leave the answer to the question open and document this undecidability. In HNR, however, a formalized statement is required in any case. Within the data, this could be reached by either including the relationship between A and B, or by omitting it, or by attributing the relationship in question with a "certainty weight" of, say, 0.5, meaning that the likelihood that this relationship existed is estimated by the researcher with 50%.[86] Data formalized in this way is processable by a computer as well as it can be used for visualization.

This example of making explicit statements about the data is comparable to that of constituting a particular reading of a text in a scholarly edition. But what happens if the reader of an edition or in our case the viewer of the visualization does not agree with the editor's respectively our researcher's decision? Interpreting the relationship between A and B differently might lead to a different network, consequently to a different model and

[84] Peter L. Shillingsburg, *From Gutenberg to Google: Electronic representations of literary texts* (Cambridge: Cambridge Univ. Press, 2006), http://www.loc.gov/catdir/enhancements/fy0733/2006497553-b.html, 88.

[85] Johanna Drucker, "Humanities Approaches to Graphical Display," *Digital Humanities Quarterly* 5, no. 1 (2011).

[86] A similar mechanism is implemented in the TEI with the @cert-attribute.

hypothesis of the past. But which ways do exist to test the researcher's hypothesis or make up one's own?

In digital scholarly editing, attempts have been made to use dynamic forms of textual presentations in which the reader is empowered to follow own reading paths, highlight particular aspects of the textual data (at the cost of others) and so on, hence having an interactive tool at hand to verify the editor's decisions as well as to make own hypotheses.[87] Such forms have not yet reached a critical mass and broad acceptance as a consequence of the aforementioned compromise. Naturally, they are disputable, especially because the traditional role model between editor and reader, between researcher and recipient, seem to vanish and responsibilities for research results disavowed. At the time of this writing, I would argue that this discussion whether such new forms of scholarly communication brings scholarship forward or are rather steps backwards is undecided. But as a natural part of research and progress, these questions have to be actively discussed, and HNR can contribute greatly to this debate.

5. Conclusion

In this essay, an attempt has been made to characterise Historical Network Research as a topical research paradigm that is located in the scholarly landscape as an integral part of the realms of Digital Humanities, Digital History, and Historical Information Science. Such an interdisciplinary intersection can be employed for mutual benefit. On the one hand, HNR may profit from computer-based practices, procedures, and tools that are being developed and made ready for general use within the area of Historical Information Sciences. On the other hand, HNR itself offers significant potential to improve digital and otherwise formalized methods which can then be used in other scholarly disciplines under the umbrella of the Digital Humanities. The methodological issues that HNR opens up and attempts to answer may contribute on a general level to the theoretical foundation of computer-based approaches in the Humanities.

HNR as a paradigm contributes also to the methodological canon of scholarship in history. By providing complementary approaches to historical questions, it helps expanding the boundaries of research, offers new perspectives to old questions and potentially opens up new terrain. An anchoring of HNR within the theoris of history as well as those of computer-based Humanities has still to be consolidated, yet.

[87] As an example see Rehbein, "Rehbein 2009b" and the discussion above.

Bibliography

Anderson, Chris. “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.” *Wired Magazine* 16.07 (2008).

Bartalesi, Valentina, Carlo Meghini, Paola Andriani, and Mirko Tavoni. “Towards a Semantic Network of Dante’s Works and Their Contextual Knowledge.” *Digital Scholarship in the Humanities*, 2015, fqv044. doi:10.1093/llc/fqv044.

Bauman, Syd. “Interchange vs. Interoperability.” *Proceedings of Balisage: The Markup Conference 2011*, 2011. doi:10.4242/BalisageVol7.Bauman01. http://www.balisage.net/Proceedings/vol7/print/Bauman01/BalisageVol7-Bauman01.html.

Bingenheimer, M., J.-J. Hung, and S. Wiles. “Social network visualization from TEI data.” *Literary and Linguistic Computing* 26, no. 3 (2011): 271–78. doi:10.1093/llc/fqr020.

Bixler, Matthias. “Die Wurzeln der Historischen Netzwerkforschung.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 45–61.

Bixler, Matthias, and Daniel Reupke. “Von Quellen zu Netzwerken.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 101–22.

Bloothooft, Gerrit, Peter Christen, Kees Mandemakers, and Marijn Schraagen, eds. *Population Reconstruction*. 1st ed. 2015. Cham, s.l.: Springer International Publishing, 2015.

Boonstra, Onno, Leen Breure, and Peter Doorn. “Past, present and future of historical information science.” *Historical Social Research / Historische Sozialforschung* 29, no. 2 (2004).

Brown, David M., Adriana Soto-Corominas, and Juan Luis Suárez. “The preliminaries project: Geography, networks, and publication in the Spanish Golden Age.” *Digital Scholarship in the Humanities*, 2016, fqw036. doi:10.1093/llc/fqw036.

Castells, Manuel. *The Rise of the Network Society.* The Information Age: Economy, Society and Culture 1. Cambridge, MA, Oxford, 1996.

Clausen, Nanna Floor. “The Danish Demographic Database— Principles and Methods for Cleaning and Standardisation of Data.” In *Population Reconstruction*. Edited by Gerrit Bloothooft et al. 1st ed. 2015, 3-22. Cham, s.l.: Springer International Publishing, 2015.

Drucker, Johanna. “Humanities Approaches to Graphical Display.” *Digital Humanities Quarterly* 5, no. 1 (2011).

Düring, Marten, Ulrich Eumann, Martin Stark, and Linda von Keyserlingk. “Einleitung.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 5–10.

———, eds. *Handbuch Historische Netzwerkforschung: Grundlagen und Anwendungen.* Schriften des Kulturwissenschaftlichen Instituts Essen (KWI) zur Methodenforschung 1. Berlin: LIT-Verlag, 2016.

Düring, Marten, and Florian Kerschbaumer. “Quantifizierung und Visualisierung. Anknüpfungspunkte in den Geschichtswissenschaften.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 31–43.

Düring, Marten, and Linda von Keyserlingk. “Netzwerkanalyse in den Geschichtswissenschaften. Historische Netzwerkanalyse als Methode für die Erforschung

von historischen Prozessen." In *Prozesse: Formen, Dynamiken, Erklärungen*. Edited by Rainer Schützeichel and Stefan Jordan, 337–50. Wiesbaden: Springer VS, 2015.

Eagleton, Terry. *Marxism and literary criticism*. 2nd ed. New York: Routledge, 2002.

Eckl, Markus. "Von Forschungsteams zur Wissenschaftscommunity: Eine soziale Netzwerkanalyse der wissenschaftlichen Co-Autorenschaften in der Disziplin der Sozialen Arbeit zwischen 1980 und 2014." *Soziale Passagen* 8 (2016).

Eder, Maciej. "Visualization in stylometry: Cluster analysis using networks." *Digital Scholarship in the Humanities* 32, no. 1 (2017): 50–64. doi:10.1093/llc/fqv061.

Edmondson, Chloe. "An Enlightenment Utopia: The Network of Sociability in Corinne." *Digital Humanities Quarterly* 11, no. 2 (2017).

Emirbayer, Mustafa. "Manifesto for a Relational Sociology." *The American Journal of Sociology* 103, no. 2 (1997): 281–317. doi:10.1086/231209.

Fischer, Frank, Andreas Vogel, Mathias Göbel, Peer Trilcke, Dario Kampkaspar, and Christopher Kittel. "Distant Reading Showcase: 200 Jahre deutsche Dramengeschichte auf einem Blick." DHd 2016, Leipzig, March 09, 2016. http://dhd2016.de/sites/default/files/dhd2016/files/Distant_Reading_Showcase__465_German_Dramas__DHd2016_Poster.pdf.

Gamermann, Daniel, Carmen Moret-Tatay, Esperanza Navarro-Pardo, and Pedro Fernandez de Córdoba Castellá. "The Small-World of 'Le Petit Prince': Revisiting the Word Frequency Distribution." *Digital Scholarship in the Humanities*, 2016, fqw005. doi:10.1093/llc/fqw005.

Gmür, Markus. "Co-citation analysis and the search for invisible colleges: A methodological evaluation." *Scientometrics* 57, no. 1 (2003): 27–57.

Gold, Matthew K. and Lauren F. Klein, eds. *Debates in the digital humanities*. Minneapolis: Univ. of Minnesota Press, 2012.

———, eds. *Debates in the Digital Humanities 2016*. Minneapolis: University of Minnesota Press, 2016. http://dhdebates.gc.cuny.edu/debates/2.

Goldin, Ian. *Divided Nations: Why global governance is failing, and what we can do about it*. Oxford, 2013.

Gramsch, Robert. "Zerstörte oder verblasste Muster? Anwendungsfelder mediävistischer Netzwerkforschung und das Quellenproblem." In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 85–99.

Gregory, Ian, Christopher Donaldson, Patricia Murrieta-Flores, and Paul Rayson. "Geoparsing, GIS, and Textual Analysis: Current Developments in Spatial Humanities Research." *International Journal of Humanities and Arts Computing* 9, no. 1 (2015): 1–14. doi:10.3366/ijhac.2015.0135.

Gruber, Thomas. "A Translation Approach to Portable Ontology Specifications." *Knowledge Acquisition* 5, no. 2 (1993): 199–220.

Haber, Peter. *Digital Past: Geschichtswissenschaft im digitalen Zeitalter*. München: Oldenbourg, 2011.

History.com staff. "Underground railroad." http://www.history.com/topics/black-history/underground-railroad.

Hockey, Susan. “The History of Humanities Computing.” In *A Companion to Digital Humanities*. Edited by Susan Schreibman, Ray Siemens and John Unsworth, 3–19. Blackwell companions to literature and culture 26. Oxford: Blackwell, 2004.

Hu, J., and N. Wang. “Complex network perspective on graphic form system of Hanzi.” *Literary and Linguistic Computing* 28, no. 4 (2013): 660–67. doi:10.1093/llc/fqt057.

Hughes, Lorna, Panos Constantopoulos, and Costis Dallas. “Digital Methods in the Humanities: Understanding and Describing their Use across the Disciplines.” In Unsworth; Siemens; Schreibman, *A new companion to digital humanities*, 150–70.

Ide, Nancy, C. M. Sperberg-McQueen, Robert Amsler, Donald Walker, Susan Hockey, and Antonio Zampolli. “Proposal for Funding for an Initiative to Formulate Guidelines for the Encoding and Interchange of Machine-Readable Texts.” http://www.tei-c.org/Vault/SC/scg02.html.

Jackson, Cornell. “Using Social Network Analysis to Reveal Unseen Relationships in Medieval Scotland.” *Digital Scholarship in the Humanities*, 2016, fqv070. doi:10.1093/llc/fqv070.

Jannidis, Fotis. “Netzwerke.” In Jannidis; Kohle; Rehbein, *Digital Humanities*, 147–61.

Jannidis, Fotis, Hubertus Kohle, and Malte Rehbein, eds. *Digital Humanities: Eine Einführung*. Stuttgart: J.B. Metzler Verlag, 2017.

Janosov, Milan. “Network Science Predicts Who Dies Next in Game of Thrones.” https://cns.ceu.edu/article/2017-07-08/network-science-predicts-who-dies-next-game-thrones.

Juola, Patrick. “Authorship Attribution.” *Foundations and Trends in Information Retrieval* 1, no. 3 (2008): 233–334. doi:10.1561/1500000005.

Kimura, F., T. Osaki, T. Tezuka, and A. Maeda. “Visualization of relationships among historical persons from Japanese historical documents.” *Literary and Linguistic Computing* 28, no. 2 (2013): 271–78. doi:10.1093/llc/fqs045.

Kolmer, Lothar. *Geschichtstheorien*. 2008.

Kräutli, Florian, and Matteo Valleriani. “CorpusTracer: A CIDOC database for tracing knowledge networks.” *Digital Scholarship in the Humanities*, 2017.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago, Ill.: University of Chicago Press; The University of Chicago Press, 1962.

Kuny, Terry. “A Digital Dark Ages? Challenges in the Preservation of Electronic Information.” *63rd IFLA Council and General Conference*, 1997.

Le Roy Ladurie, Emmanuel. *Le territoire de l'historien.* Bibliothèque des histoires. Paris: Gallimard, 1973.

Lemercier, Claire. “Formale Methoden der Netzwerkanalyse in den Geschichtswissenschaften: Warum und Wie?” *Österreichische Zeitschrift für Geschichtswissenschaft* 23, no. 1 (2012): 16–41.

Lynch, Tom J. “Social Networks and Archival Context Project: A Case Study of Emerging Cyberinfrastructure.” *Digital Humanities Quarterly* 8, no. 3 (2014).

Mandell, Laura. “How to Read a Literary Visualisation: Network Effects in the Lake School of Romantic Poetry.” *Digital Studies / Le champ numérique* 3, no. 2 (2012).

Marx, Christian. “Forschungsüberblick zur Historischen Netzwerkforschung. Zwischen Analysekategorie und Metapher.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 63–84.

Mayer, Katja. “Netzwerkvisualisierungen. Anmerkungen zur visuellen Kultur der Historischen Netzwerkforschung.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 139–53.

McCarty, Willard. *Humanities computing*. Basingstoke England, New York: Palgrave Macmillan, 2005. http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10254194.

McGill University & Université de Montréal. *Digital Humanities 2017: Conference Abstracts*. 2017.

Mitchell, James Clyde. “The Concept and Use of Social Networks.” In *Social Networks Urban Situations: Analyses of personal relationships in Central African towns*. Edited by James C. Mitchell. Manchester, 1969.

Moretti, Franco. *Graphs, maps, trees: Abstract models for a literary history*. 1. publ. London: Verso, 2005.

Opheim, B. “Political Networks and Factions: Online Prosopography of Medieval Scandinavian Sagas.” *History and Computing* 12, no. 1 (2000): 43–57.

Porsdam, Helle. “Digital Humanities: On Finding the Proper Balance between Qualitative and Quantitative Ways of Doing Research in the Humanities.” *Digital Humanities Quarterly* 7, no. 3 (2013). http://www.digitalhumanities.org/dhq/vol/7/3/000167/000167.html.

Price, Kenneth M. “Social Scholarly Editing.” In Unsworth; Siemens; Schreibman, *A new companion to digital humanities*, 137–49.

Rehbein, Malte. “Reconstructing the Textual Evolution of a Medieval Manuscript.” *Literary & Linguistic Computing* 24 (2009): 319–27.

———. “Digitalisierung braucht Historiker/innen, die sie beherrschen, nicht beherrscht.” *H-Soz-Kult*, 27.11.2015.

———. “Was sind Digital Humanities?” *Akademie Aktuell* 56, no. 1 (2016): 13–17.

———. “Geschichtsforschung im digitalen Raum. Über die Notwendigkeit der Digital Humanities als historische Grund- und Transferwissenschaftwissenschaft.” In *Papstgeschichte des hohen Mittelalters: digitale und hilfswissenschaftliche Zugangsweisen zu einer Kulturgeschichte Europas*. Edited by K. Herbers and V. Trenkle., 2017.

———. “Informationsvisualisierung.” In Jannidis; Kohle; Rehbein, *Digital Humanities*, 328–42.

———. “Ontologien.” In Jannidis; Kohle; Rehbein, *Digital Humanities*, 162–76.

Rehbein, Malte, and Patrick Sahle. “Digital Humanities lehren und lernen: Modelle, Strategien, Erwartungen.” In *Evolution der Informationsinfrastruktur: Kooperation zwischen Bibliothek und Wissenschaft*. Edited by Heike Neuroth, Norbert Lossau and Andrea Rapp, 209–28. Glückstadt: vwh Hülsbusch, 2013.

Reinert, Matthias, Maximilian Schrott, Bernhard Ebneth, and Malte Rehbein. “From Biographies to Data Curation – the Making of www.deutsche-biographie.de.” *Proceedings of the First Conference on Biographical Data in a Digital World*, 2015, 13–19.

Ries, Wiebrecht. *Nietzsche zur Einführung*. 7th ed. Hamburg, 2004.

Robertson, Stephen. “The Differences between Digital Humanities and Digital History.” In *Debates in the Digital Humanities 2016*. Edited by Matthew K. Gold and Lauren F. Klein. Minneapolis: University of Minnesota Press, 2016.

Sahle, Patrick. “DH studieren! Auf dem Weg zu einem Kern- und Referenzcurriculum der Digital Humanities.” *DARIAH-DE Working Papers* 1 (2013). http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2013-1.pdf.

Schmidt, Desmond. “Towards an Interoperable Digital Scholarly Edition.” *Journal of the Text Encoding Initiative* 7 (2014).

Schöch, Christof. “Big? Smart? Clean? Messy? Data in the Humanities.” *Journal of Digital Humanities* 2, no. 3 (2013). http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/.

Schreibman, Susan, Ray Siemens, and John Unsworth, eds. *A Companion to Digital Humanities*. Blackwell companions to literature and culture 26. Oxford: Blackwell, 2004.

Schröder, Wilhelm Heinz. “Die Buchreihe „Historisch-sozialwissenschaftliche Forschungen“ als publizistisches „Flaggschiff“ der quantitativen Historischen Sozialforschung in der „Pionierzeit“.” *Historical Social Research / Historische Sozialforschung*, Supplement No. 18 (2006): 6–23.

Schubert, Charlotte. “Die Visualisierung von Quellennetzwerken am Beispiel Plutarchs.” *Digital Classics Online* 2, no. 1 (2016). doi:10.11588/dco.2016.1.23825.

Schürer, Kevin, Tatiana Penkova, and Yanshan Shi. “Standardising and Coding Birthplace Strings and Occupational Titles in the British Censuses of 1851 to 1911.” *Historical Methods* 48, no. 4 (2015): 195–213.

Shillingsburg, Peter L. *From Gutenberg to Google: Electronic representations of literary texts*. Cambridge: Cambridge Univ. Press, 2006. http://www.loc.gov/catdir/enhancements/fy0733/2006497553-b.html.

Stark, Martin. “Netzwerkberechnungen. Anmerkungen zur Verwendung formaler Methoden.” In Düring; Eumann; Stark; Keyserlingk, *Handbuch Historische Netzwerkforschung*, 155–71.

Tehrani, Jamshid, Quan Nguyen, and Teemu Roos. “Oral fairy tale or literary fake? Investigating the origins of Little Red Riding Hood using phylogenetic network analysis.” *Digital Scholarship in the Humanities* 31, no. 3 (2016): 611–36. doi:10.1093/llc/fqv016.

Terras, Melissa. “Crowdsourcing in the Digital Humanities.” In Unsworth; Siemens; Schreibman, *A new companion to digital humanities*, 420–38.

Thaller, Manfred. “Grenzen und Gemeinsamkeiten: Die Beziehung zwischen der Computerlinguistik und den Digital Humanities.” DHd 2014, Passau, March 27, 2014.

———. “Geschichte der Digital Humanities.” In Jannidis; Kohle; Rehbein, *Digital Humanities*, 3–12.

Thomson, N. “How to Read Articles which Depend on Statistics.” *Literary and Linguistic Computing* 4, no. 1 (1989): 6–11. doi:10.1093/llc/4.1.6.

Trilcke, Peer. “Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft.” In *Empirie in der Literaturwissenschaft*. Edited by Philip Ajouri, Katja Mellmann and Christoph Rauen, 201–47. Münster, 2013.

Tufte, Edward R. *The visual display of quantitative information*. Cheshire, Conn. (Box 430, Cheshire 06410): Graphics Press, 1983.

Unsworth, John. “Computational Work with Very Large Text Collections: Interoperability, Sustainability, and the.” *Journal of the Text Encoding Initiative* 1, Issue 1 (2011). doi:10.4000/jtei.215. http://jtei.revues.org/215.

Unsworth, John, Raymond George Siemens, and Susan Schreibman, eds. *A new companion to digital humanities*. Blackwell companions to literature and culture 93. Chichester, West Sussex, UK, Malden, MA, USA: John Wiley & Sons Ltd, 2016. http://onlinelibrary.wiley.com/book/10.1002/9781118680605.

van Dijk, Jan. *The network society: Social aspects of new media*. London, Thousand Oaks, Calif: Sage Publications, 1999. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=309855.

van Hooland, Seth, Max de Wilde, Ruben Verborgh, Thomas Steiner, and Rik van de Walle. “Exploring entity recognition and disambiguation for cultural heritage collections.” *Literary and Linguistic Computing*, 2013. http://llc.oxfordjournals.org/content/early/2013/11/29/llc.fqt067.full.pdf.

Wellman, Barry. “Doing It Ourselves: The SPSS Manual as Sociology's Most Influential Recent Book.” In *Required reading: Sociology's most influential books*. Edited by Dan Clawson, 71–78. Amherst: University of Massachusetts Press, 1998.