

Abstracts for the Workshop “Doing Digital Humanities: Bridging the Gap between Quantitative and Qualitative Methods” (22.-23.4., Freiburg)

Organized by Dorothee Birke (Aarhus), Nicole Falkenhayner (Freiburg) and Martin Hinze (Freiburg)

Andreas Langenohl (Giessen)

Theories, Methodologies, Technologies: Does Digitization Create Analytical Idioms?

The concept of ‘analytical idioms’ can be understood as an epistemological perspective focusing on the ways that analyses relate to their objects of investigation, and thus become constituted, prior to any particular analytical claim, be it theoretical, conceptual, methodological, or empirical. In this paper I raise the question what ‘digital humanities’ can mean from this constitutive-theoretical perspective, which will be discussed in three steps.

First, theory. From an idiomatic perspective, ‘digital humanities’ denotes an urge for the current humanities to re-constitute themselves vis-à-vis an increasingly digitized world. This is not just a response or an adaptation, but an attempt to re-constitute the object of investigation in the first place – an object that, however, presents itself already as a second-order reflection on the ways it changes. Many attempts in the ‘digital humanities’ seem to present themselves as ‘humanities of digitization,’ focusing primarily on digitization as a common denominator for a meta-analysis of the processes through which societies and cultures are changing. They thus constitute themselves vis-à-vis an object of investigation that is already a vernacular theory of grand societal and cultural transformation, thus making it inescapable to relate to the world via the detour of vernacular theories, for instance, such of ‘digitization’ or ‘medialization’. ‘Digital humanities,’ thus, would refer to an inquiry into an increasingly self-reflective society and culture that, however, risks to get stuck in a mode of contemporary analysis, focusing solely on present and real transformations. For instance, theories of medial change tend to be dominated by reflections about contemporary media and meta-media, such as the mobile end devices or the internet.

Second, methodology. Unlike most currents in the modernist natural sciences and some strands of the social sciences, the humanities have often relied on a methodology that enters into an indissoluble connection with the concrete object under analysis. Their epistemological gesture is not to analyze ‘data’ with a set of ‘methods’ fixed in advance, but to engage the object of investigation into a dialog with the analytical optics; and further, to challenge the distinction between form and content that is so characteristic of modernist sciences. Digitization, telling from the jargon it uses (like DOI – digital object identifier; content; data set; etc.), might make it necessary to redeem humanist analytical idioms vis-à-vis a digitized environment that is thoroughly modernist, because it requires that everything be turned into data, digital objects, or content.

Third, technology. While digitization obviously refers to technological change, it also raises the question of technique: how do cultural techniques like reading or writing change through digitization, and what does this mean for the humanities? Digitization relates the humanities (and parts of the social sciences) to broader cultural techniques and their normative models, for instance, concerning literacy, education, and understandings of ‘good’ entertainment. As it has traditionally been the humanities that were in charge for the development, dissemination and training of cultural techniques, this is not new in itself, but belongs to the core of the humanist legacy, including humanist critiques of contemporary practices of literacy for which the humanities

must serve as a corrective. However, what might pose a challenge for the humanist gesture to normatively engage cultural techniques is the way that digitization creates demand for cultural meta-techniques. This regards, for instance, the technique to calibrate the relationship between what to put on display and what not, that is, between networking and un-networking. As digital environments tend to be more and more 'interactive,' they involve traces and representations of the subject that 'uses' them. Ideals of education, such as becoming an autonomous subject, might be in need to be supplemented by others, such as becoming an autonomous persona. In other words, under conditions of digitization, cultural techniques enter into an ever more tight relationship with techniques of the (digital) self.

Jan-Christoph Meister (Hamburg)

What's the Point of the Digital for the Humanities?

DH is doubtlessly in. It may have taken some forty years since 1973 when the first conference of the "Association for Literary and Linguistic Computing" (nowadays superseded by the European Association for Digital Humanities, EADH) was held in Oxford – but by 2010, when the leading international event known to practitioners simply as "the DH conference" was held at King's College London, attracting over 400 delegates for the first time ever, what was once considered an undertaking of methodological mavericks had eventually established itself as an integral part of the set of practices and methodologies that inform the Humanities disciplines. Today every second grant proposal submitted to one of the major funding agencies will at the very least hint at some DH component – it's simply a "must have".

Why have the DH been successful in the end? Obviously one must acknowledge that IT based analysis, modelling and managing of phenomena and processes of relevance to mankind have quite simply become an epistemological in as much as a pragmatic corner stone of our contemporary knowledge societies. But so are other theoretically demanding technological and scientific approaches, which the Humanities have by and large been quite happy to ignore nevertheless – quantum physics, nano technology, complexity theory, to name a few.

Against this backdrop my thesis is the following: I believe that there is an intrinsic, philosophical affinity between the digital and the humanities' interest in *aesthesis* – that is, in the project of exploring new ways of seeing and conceptualizing phenomena, be they empirical or be they cognitive and/or mental. I will argue that the digital is, in fact, a conceptual *lingua franca* that can enable us to overcome the present compartmentalization of scientific versus humanistic knowledge domains. The point of the digital for the humanities is thus of a methodological and philosophical order, and not restricted to pragmatic affordances.

Stylianos Chronopoulos (Freiburg)

Re-understanding Our Texts: Scholarly Digitized and Digital Editions

Doing philology in a digital world means, among other things, that we transform the materials we are studying, the texts, so that a) they can be stored in and transmitted through computers and b) they are machine perceivable, machine operable and machine comprehensible. The process of this transformation has up to now taken two forms: a) the digitisation of print scholarly editions and b) the production of born digital editions. Both are processes which involve the transfer of content structured in certain ways from one medium to another. This process of transfer means: a) transformation of the architectures a certain text has in a certain medium, and b) transformation of the ways users interact with the text. The text architectures concern its structures on various levels and depend both on its genre and on its storage and transmission medium. The same two factors are decisive for the ways in which a user interacts with a certain text.

Digitisation of a print edition means primarily transfer of the scholarly model, with which a certain historical document has been represented, in the digital medium. The creation of a born digital edition on the other hand requires that the editor is “free” to create a model fully adapted to the digital medium; but in fact in the phase of transition from print to digital editions it is very difficult to “think out of the book” and develop genuinely new models for presenting a text and its history of transmission.

The first part of my presentation is a general discussion of these issues. I will especially focus on the following question: which possibilities offer the digital medium to create products that enable the critical assessment of texts and their history of tradition? In the second part I use my work on a digitized edition of Pollux’ Onomasticon, a thesaurus-similar Greek dictionary of the 2nd century CE as an example in order to discuss two aspects of the complex process of transfer involved in the creation of scholarly editions in a digital space.

1. The issue of referencing to passages of the same or other texts. Referencing is immediately connected with the perception (or creation) of text structures and involves the very question about what a text is (a structured series of signs independent from its concrete materialisation, a specific edition, a specific object). I discuss the issue from a historical perspective, I address its relevance for digital critical editions and I present the example of a framework for scholarly reference, the CITE architecture (<http://www.homermultitext.org/hmt-doc/cite/>).

2. The issue of annotating implicit information. Editing a text means primarily imposing a specific model on it, that prescribes which information will be presented and how. The process of digitising a print edition involves a very detailed understanding of this model and an attempt to annotate and present not only the explicit but also the implicit information it contains. The punctuation for example or certain typographic arrangements without a clearly defined meaning should be interpreted and encoded unambiguously. In that respect the preparation of a digitised edition is an act of interpretation both of the text and of the print edition. I will exemplify this by four specific cases from the digitisation process of Pollux’ Onomasticon.

Christiane Hadamitzky (Freiburg)
Distant Approaches to Close Reading

Large-scale digitization projects offer great chances for research. They provide access to material which would have been unknown to scholars only decades ago or would have taken years of archival research to compile. However, as a literary scholar predominantly concerned with discourse analysis and close reading of case studies, the large amount of material also gives rise to new challenges.

On the basis of my doctoral research on the heroic in Victorian periodicals, my presentation is going to give an example of hands-on work with large corpora in cultural studies and will show how distant reading can indicate ways into close reading. Victorian periodicals - with both continuous and ephemeral qualities - lend themselves well to approaches of distant reading due to the vastness of material that weekly or monthly publications produce. After a description of my methods of quantitative analysis, my presentation is going to show how these quantitative findings of a large number of texts revealed tendencies for qualitative analysis and pointed towards texts suitable for close readings.

Using the example of genre, I am further going to discuss how the close reading can in turn refer back to more general categories. The combination of distant and close forms of reading can contribute to the understanding of functionalizations of specific text types in a way which the two approaches could not provide separately. Thus, a utilization of both distant and close, quantitative and qualitative methods of analysis can provide not only a gateway into large corpora, but also indicate alleyways into future investigation of both specific and broader issues.

Geraldine Meaney (Dublin)
Reconciling Close and Distant Reading: Social Network Analysis of the Novel in the Long Nineteenth Century

This paper will examine the methodological issues and critical opportunities that arise from combining digital social network analysis with traditional literary scholarship in relation to the fiction of the long nineteenth century. It will outline the work to date of the 'Nation, Genre and Gender: A Comparative Social Network Analysis of Irish and English Fiction, 1800-1922' project funded by the Irish Research Council. The ultimate objective of the project is to compare gender, genre and the nationality of the author (or setting) in shaping social networks in fiction. Do the social networks mapped out by cumulative interactions between characters in Irish and English fiction differ from one another? Do the social networks represented in fiction differ substantially on the basis of date, genre or gender? The research combines quantitative, computational approaches with critical and interpretative tools. This combination offers new perspectives on well known texts, but also a realistic and judicious form of intense textual engagement with a radically extended canon of fiction, with its diversity of voices, genres and perspectives. This paper will outline the general methodology used and then use *Oliver Twist* and *Phineas Finn* as case studies which demonstrate the benefits of an inclusive approach to nodes within the novels' social networks (including unnamed characters and collectives, for example), the use of search functionality to situate the social imaginary of the novels within key social and political

developments in the Victorian period and to indicate how social network analysis can inform analysis of both gender and ethnicity in fiction.

Ruth Page (Birmingham)

Mixed Methods and Multimodality in the Shared Story: A Case Study in the YouTube Coverage of Oscar Pistorius

In this paper I will introduce the framework I have developed for the analysis of shared stories. Shared Stories have antecedents in collaborative conversational storytelling, in literary adaptations and the retelling of news across different broadcast outlets. Shared stories have four characteristics. They have distributed linearity, are narrated by multiple co-tellers, are richly intertextual and are used to promote shared knowledge between participants. In social media contexts, the analysis of shared stories presents digital humanities scholars with interesting challenges. On the one hand, the scale of the data brought together in a shared story can be very large (consisting of many thousands of comments, or posts), and thus lend itself to semi-automated forms of collection and analysis. On the other hand, the multimodal nature of the data means that qualitative forms of close reading remain invaluable. This is perhaps particularly true in the shared stories found in YouTube, a site in which the spectacular, multimodal potential for online interaction is centre stage (Deumart 2014).

The case study presented here examines the shared stories that emerged in the YouTube-based citizen journalism surrounding the live-streamed trial of South African Paralympian, Oscar Pistorius, for killing his girlfriend, model and television presenter, Reeva Steenkamp. Citizen journalism in YouTube includes many forms of participation, including dissemination of eye-witness accounts, curating online content, responding to the online content in video and text based formats, including remixed parody or written comments. The data in this study comprise a corpus of 601 YouTube videos and their appended comments (totalling 43,092), gathered using Context Miner (Shah 2015) in July 2015, and the further appeal and verdict of November 2015. The videos include the live-streamed trial, news reports, comedians' performances, musical parodies, rants (Lange 2014) and v-loggers' personal responses which critiqued Pistorius and the outcome of the trial. Together, these videos constitute more than a single shared story, and form a shared story spectacle shaped by the medium affordances of YouTube.

The analysis concentrates on the distributed linearity and co-tellership found in the videos and the comments. Both aspects of the shared story are used to question the extent to which the citizen journalism in this case reconfigured or replicated earlier forms of mainstream television news. Using a multimodal adaptation of White's (1998) model of the nucleus^satellite structures found in print news, I show how the video-based responses by online news and v-loggers open up greater opportunities for concession and evaluation. However, analysis of the YouTube template suggests that the mainstream news continues to dominate the production of news in this site, where news companies use the Descriptions to boost their own visibility by coupling the nucleus of the video and its title to satellites which promote audience engagement with the news producers and their channels.

Further, gender-based asymmetries were also identified in the 'satellites' of appraisal, elaboration and contextualisation produced in the comments appended to the videos, where comments by male members of YouTube more than doubled those by female members. The

relative proportion of male commenters varied according to the genre, where commenting on parodies and mainstream news was the most male-dominated. A corpus-based analysis of the comments suggests that far from being conversational, the appraisal in the comments of the mainstream news contained spam-like, sexist 'noise' which disrupted the comment threads of the live-streamed verdict. This strategy both suggested anger with the trial's outcome (and its failure to police masculine violence), but simultaneously reinforced that heteronormative violence in its discursive, pornographic attacks on the female judge and other commenters.

Monika Bednarek (Sydney/Freiburg)

Corpus Linguistic Approaches to the Analysis of Fictional Televisual Narratives

In my contribution to this workshop I will introduce participants to digital methods for the analysis of fictional narratives that come from the discipline of corpus linguistics. Corpus linguistics is an approach to the linguistic analysis of data that uses large computerized collections of text (corpora) and appropriate software to analyse them. Particularly relevant to the study of fictional narratives is a branch of corpus linguistics called corpus stylistics – the corpus linguistic analysis of fictional, usually literary, texts. Stubbs (2005) gives an overview of corpus stylistic investigation and convincingly demonstrates the validity of such research. It is not claimed here that corpus stylistics can show us all aspects of scripted or narrative texts or that it is the only way of studying such texts.

In introducing corpus stylistics, this paper will focus on the corpus linguistic analysis of fictional television narratives, namely contemporary US TV series. It will discuss two corpus linguistic case studies of such TV series, one focusing on characterization and one focusing on plot development. The first case study examines televisual character stability in one particular TV narrative, *Gilmore Girls* (Bednarek 2011), while the second case study focuses on dialogue in seven contemporary crime dramas: *24*, *NCIS*, *Bones*, *The Wire*, *The Shield*, *Breaking Bad*, and *Southland*. The discussion will bring out the affordances and limits of corpus stylistic analysis of such and other narratives.

Jonatan Steller (Leipzig)

"Introduce a little anarchy": Translating Maker Democracy from a Critical Framework into a Publishing Form using Fragmented Web Guides

The academic book is an undemocratic technology as defined by Carroll Pursell: it is hard to obtain due to its steep price and limited run, it is hard to understand because of its ivory tower language and structure, and it is hard to fix in case it does not make sense. Hence when I started work on a critical framework that harnesses concepts of decentralised, collaborative technologies as used by movements such as Occupy, maker, open source and creative commons, the academic monograph was out of the question as the main publishing form. To briefly illustrate the bridge between research content and publishing, this paper illustrates three strands of research: the Maker Democracy framework itself, its exemplary application to Christopher Nolan's Dark Knight trilogy and experiments with an accessible, open source publishing form. The work is rooted in Nathan Jun's proposal for a post-anarchist film theory that reacts to Bourdieu's idea of reification and

resistance. Jun argues that Bourdieu unduly gives political agency to cultural products: post-anarchism states that political action resides only with the producing collective and with the audience. Hence films themselves merely carry potentialities of political action, and the Maker Democracy framework aims to establish a set of such potentialities that are productive in present-day blockbuster fiction (films, video games, novels, music albums and comic books with wide audience appeal and success). In their characterisation of establishment and youthful deviation, these potentialities resemble a subculture as defined by Raymond Williams. There are, however, two important differences: Maker Democracy is a fragmented movement and it only seeks to disrupt, not to oppose the mainstream (neoliberal) ideology.

The framework points out five repeated actions in present-day blockbusters that align to real world practices: creating agency through makership in the context of an established hierarchy, accessing knowledge that is kept secret to spread fear, gaining power through an authentic voice and a network of trust, providing resources to others based on needs instead of on hierarchy, and sharing the control attained in the process. Applied to the Dark Knight trilogy, Bruce Wayne is not characterised as a representation of financial aristocracy with a moral conscience. Instead, he puts the resources available to him into disrupting organised crime (controlling most of Gotham's money), the police and the local political landscape. His avatar Batman provides Wayne with the dynamic, authentic voice that his real identity has been struggling to find; the attempt to make such authenticity permanent through Harvey Dent fails. Since his aim is only to disrupt, he never fully seizes control but returns it to his circle of friends and to the people of Gotham. The merit and dangers of chaos serve as the films' main theme, actualised for example in Gotham as Arkham Asylum in *Begins*, the Joker as a criminal-turned-terrorist in *Dark Knight*, and through a revolution that Catwoman sympathises with in *Rises*.

For the sake of applying the same idea of disruption and to experiment with chaos, I decided to imitate for academic publishing the way open-source software is produced. Text and resources produced for my PhD project are set to be available online on GitHub. From here they are available to fork and enhance if properly attributed. The material is structured as a non-linear collection of approximately 200 web guides, i.e. individual pages organised around topic sentences, analyses and methodology. By making methodology like tagging or still-image interpretation accessible, web guides help make non-STEM research more reproducible. A structure is provided for use as a unified unit of research, but the set of guides