

Quantitative Methods in the Humanities
and Social Sciences

Adelheid Heftberger

Digital Humanities and Film Studies

Visualising Dziga Vertov's Work

 Springer

Quantitative Methods in the Humanities and Social Sciences

Editorial Board

Thomas DeFanti, Anthony Grafton, Thomas E. Levy, Lev Manovich,
Alyn Rockwood

Quantitative Methods in the Humanities and Social Sciences is a book series designed to foster research-based conversation with all parts of the university campus – from buildings of ivy-covered stone to technologically savvy walls of glass. Scholarship from international researchers and the esteemed editorial board represents the far-reaching applications of computational analysis, statistical models, computer-based programs, and other quantitative methods. Methods are integrated in a dialogue that is sensitive to the broader context of humanistic study and social science research. Scholars, including among others historians, archaeologists, new media specialists, classicists and linguists, promote this interdisciplinary approach. These texts teach new methodological approaches for contemporary research. Each volume exposes readers to a particular research method. Researchers and students then benefit from exposure to subtleties of the larger project or corpus of work in which the quantitative methods come to fruition.

More information about this series at <http://www.springer.com/series/11748>

Adelheid Heftberger

Digital Humanities and Film Studies

Visualising Dziga Vertov's Work

 Springer

Adelheid Heftberger
German Federal Archive (Bundesarchiv)
Berlin, Germany

Translation from the German language edition: *Kollision der Kader. Dziga Vertovs Filme, die Visualisierung ihrer Strukturen und die Digital Humanities* by Adelheid Heftberger, © Adelheid Heftberger, 2016. Published by München: edition text + kritik. All Rights Reserved.

ISSN 2199-0956 ISSN 2199-0964 (electronic)
Quantitative Methods in the Humanities and Social Sciences
ISBN 978-3-030-02863-3 ISBN 978-3-030-02864-0 (eBook)
<https://doi.org/10.1007/978-3-030-02864-0>

Library of Congress Control Number: 2018960940

© Springer Nature Switzerland AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For Maksim

Acknowledgements

In international research, one may find both historical and recent activities in formal film analysis, concurrently indicating paths to film scholarly explanations. Only in collaboration and exchange with colleagues is academic research possible and meaningful, and for that reason my first expression of gratitude goes to the so-called Vertovians who form a special circle within that of scholarship and became my friends over the course of time. To these belong, in the first rank, not only Aleksandr Derjabin, John MacKay and Yuri Tsivian but also my project partners Michael Loebenstein, Thomas Tode, Georg Wasner and Barbara Vockenhuber, as well as Matthias Zeppelzauer, Dalibor Mitrović and Maia Zaharieva. My sincere thanks go to Lev Manovich for the inspiring collaboration and support in creating the visualisations. My heartfelt thanks to the Slavic scholars Eva Binder, Natascha Drubek and Barbara Wurm, who were always ready and willing to give me information in response to questions both small and large. No less hearty are my thanks to my colleagues in Russia, who excel in both their comprehensive knowledge of film history and their readiness to be of assistance, Peter Bagrov, Nikolai Izvolov, Sergei Kapterev and Naum Kleiman, as well as to the archive staff in Russia for their willingness to provide information at short notice and the access to research materials and for their encouragement: Elena Kolikova (RGAKFD), Valerij Bosenko (GFF, died 2013) and Oleg Bočkov (GFF). For stimulating conversations on various topics, my special thanks go to Oliver Hanley, Renaud Tschirner, Karen Pearlman, Anna Bohn and Veronika Heftberger. Not the least, I am especially grateful to the Austrian Film Museum for the permission to use illustrations and the Brandenburg Centre for Media Studies (ZeM) for partly funding the English translation. My very special thanks go finally to the translator from the original German, Stewart Tryster, who brought commitment, the absolutely vital film-technical professional knowledge and enjoyment to his work with me.

Contents

1	Introduction	1
	References.	5
2	The Measurement of Aesthetic Phenomena	7
	The Sciences and the Humanities: Culture Clash or Interdisciplinary Potential?.	9
	From “Cyberinfrastructure” to the Digital Humanities	11
	The Digital Humanities in Film Archives and Libraries.	17
	Big Data: Distant Reading	20
	References.	25
3	Annotation and Statistics	29
	Annotation in the Project Digital Formalism	30
	Results of the Annotation	37
	Shot Type	38
	Camera Movement	47
	Statistical Evaluation.	51
	Distribution of Shot Lengths.	54
	Further Studies	59
	References.	63
4	Dziga Vertov’s Films	65
	The Theory of <i>Kinoglaz</i>	65
	Vertov’s Phrase Model	70
	Formal Procedures in Vertov’s Camera and Editing Work	73
	Material and Condition of Prints.	74
	Film Prints in Russian and International Archives	76
	The Individual Works	78
	<i>Kino-Eye</i> : “An Army of Film Observers”.	79
	Available Film Prints	81
	Critical Reception in the 1920s/1930s	82
	Formal Characteristics of the Film	83

<i>Stride, Soviet!</i> The Shot Library	85
Available Film Prints	87
Critical Reception in the 1920s/1930s	88
Formal Characteristics of the Film	89
<i>A Sixth Part of the World: Exotic Images</i>	91
Available Film Prints	92
Critical Reception in the 1920s/1930s	95
Formal Characteristics of the Film	97
<i>The Eleventh Year: Currents of Energy with Voids</i>	99
Available Film Prints	101
Critical Reception in the 1920s/1930s	102
Formal Characteristics of the Film	104
<i>Man with a Movie Camera: The Avant-Garde Experiment</i>	
Without Intertitles	106
Available Film Prints	109
Critical Reception in the 1920s/1930s	111
Formal Characteristics of the Film	114
<i>Enthusiasm: The Laboratory of Hearing</i>	114
Available Film Prints	118
Critical Reception in the 1930s	118
Formal Characteristics of the Film	120
<i>Three Songs of Lenin: Silence in the Film</i>	121
Available Film Prints	122
Critical Reception in the 1930s	125
Formal Characteristics of the Film	126
<i>Lullaby: The Subjective Camera</i>	128
Available Film Prints	130
Critical Reception in the 1930s	130
References	131
5 The Filmic Structure as Visualisation	139
Visualisation of Formal Characteristics of Films	141
Film Is the Art of Taking Time: The Cinemetrics Project	146
Visualisations of Vertov's Film in Cinemetrics	148
Detailed Study of <i>Man with a Movie Camera</i>	153
<i>The Eleventh Year and In the Shadow of the Machine:</i>	
A Work Report from the Archive	160
Potential of the Cinemetrics Platform	161
Visualisations Without Reduction	164
The Visualisation of Vertov's Filmic Structures	167
Shot Lengths	167
Types of Shots	170
Motion Types	172
Visual Characteristics of the Image	175
Representation of Montage	181

Combined Image Information and Shot Length	182
Combined Image Information and Shot Compositions	189
References.	189
6 Charts and Diagrams of Dziga Vertov	193
An Artist in Word, Image and Sound	193
The “Flag Episode” of <i>Kino-Eye</i>	200
Two Episodes from <i>Man with a Movie Camera</i> as a Table.	203
The Visualisation of the “Creative Laboratory”	208
References.	213
7 From Filmic Form to Meaning	215
Movement: Vertov’s Theory of the Interval	220
Movement Within the Image in <i>The Eleventh Year</i>	221
Busy Hands and Poetic Bodies: The Use of Slow Motion and Time-Lapse.	231
The Picture that Runs Backwards	232
Time-Lapse	235
Slow Motion	238
“ <i>The Little Life</i> ”: Faces of Communism	245
<i>Kino-Eye</i> : Observers and Observed	249
<i>The Eleventh Year</i> : Faces as Bearers of Messages	252
Visual Depiction of Rows of Faces in <i>The Eleventh Year</i>	260
Concluding Observations	264
Political Power: Lenin and Stalin as Film Subjects	270
Lenin as Beloved Leader.	272
Stalin in Vertov’s Films.	275
Vertov Between Avant-Garde and Social Realism	291
References.	297
8 Summary and Outlook	301
References.	307
Index.	309

Chapter 1

Introduction



In Vertov's view the mission of cinema was not to present facts, but to explain them. (Tsivian 2004b: 10)

Possible humanities future: Instead of critique, construction. Instead of close reading, patterns. Instead of interpretation, conversation. (Manovich 2012)

Anything we study thoroughly loses value for us, Reger said. We should therefore avoid studying anything thoroughly. But we cannot help studying everything thoroughly, that is our misfortune, by doing so we dissolve everything and ruin everything for us, indeed we have very nearly ruined everything for us already. (Bernhard 1989: 179)

The character from Thomas Bernhard's novel may be speaking about the works of Goethe and Shakespeare but expresses an uneasiness in principle about the analysis of art. I wish, however, to contradict the great Austrian author. If one analytically comes to grips with the oeuvre of the Russian director Dziga Vertov (David Abelevič Kaufman), in the process dismantling his films into the smallest possible units and graphically displaying individual portions in a new manner, one inevitably arrives, in addition to and alongside academic realisations, at a fundamentally deeper understanding of Vertov's films, his ideas and his times. Not only during the director's lifetime was criticism levelled at a certain mysteriousness about his films; it is not rare for bafflement also to be expressed by their audiences of today. Furthermore, analysis is consonant with Vertov's own artistic self-perception, for the director repeatedly represented his creative processes in numbers and tables and was enthusiastic about both formal and technical experiments and developments in film work. A precise knowledge of the filmic resources at his disposal, as well as the techniques and their effects, were all part and parcel of this.

The life of the Jewish documentary film pioneer (1896 to 1954) followed a restless trajectory. His family dispersed in all directions early on, with the multitalented Vertov migrating from his (now) Polish home town Białystok to Russia, where he began a career in newsreels. Moscow and St. Petersburg were the laboratories for an

atmosphere of new departures, encompassing every aspect of life, after the October Revolution. Leading figures such as Aleksandr Rodčenko or Vladimir Majakovskij energetically proclaimed the victory of new forms of expression and the young David Abel'evič Kaufman had plans for film that were just as exuberant and ambitious as were those of his role models for literature, photography or the theatre, and he adopted the artistic pseudonym of Dziga Vertov. His manifestos, articles, speeches and diary entries testify to a creative and polemic intellect that insisted on the independence of art even while fully committed to the service of the Communist cause. Already in his first manifestos, the director and the Kinoks (Vertov's neologism for his fellow campaigners, translatable as "film eyes") proclaimed to the world that the old cinema had to die and they permitted only authentic documentary film to qualify as real film art. Vertov was never to shift away from this deep conviction, even though it was ultimately to mean the slow and painful end of his career. Above all, after Lenin's death and Stalin's assumption of power, his artistic difficulties increased as the new auguries of cultural policy were restrictively enforced. The director was finally physically and mentally broken by the dictates that no longer permitted him to make films according to his own ideas.

Vertov's medium of expression was the film, which, as a technical medium and as a collectively produced art form, rose in the ranks to become the avant-garde paradigm of artistic production in the young Soviet Union. Those engaged in film-making claimed a leading role in the arts at a time in which the old social system was to be overthrown and replaced by a new order. For what could more effectively depict the achievements of the Soviet state, present new visions and surprise and enthuse the people with special effects than film? In the 1920s, throughout the country, the illiteracy rate was high, and the population was a heterogeneous mixture of peoples, differing greatly from one another in terms of language and culture; the political leadership wished to reach them in a joint effort with the film-makers. Due to the conditions of the making and distribution of films, their production and exhibition were easier for the Party to control than other art forms, although most of the film-makers had no intention of working subversively against the state, as such famous directors as Sergei Eisenstein, Vsevolod Pudovkin and Abram Room, as well as the representatives of FĖKS (The Factory of the Eccentric Actor), also saw themselves completely in the service of the Soviet Union. At the 1928 Party conference for cinema workers, the avant-garde representatives asked for guidelines for their film work and received a clear answer, taking up Lenin's statement of 1922: film must be comprehensible to millions of people, for if the population, above all the rural population, does not understand film, then all the agitation and propaganda remain ineffectual. But Vertov was accused of being unable to meet this requirement, although he himself was firmly convinced that his documentary films were grasped by the population.

My decision to adopt a formalistic and quantitative (computer-aided) method was conditioned not only by the object of research and Vertov's own working method but must also be understood against the backdrop of current changes in university research. Knowledge has for a while now been produced, transmitted and stored in completely new ways. It is, above all, the humanities that are currently in

the grip of a transformation that could perhaps fundamentally alter the demarcation lines separating them from other fields and, thus, their very self-image. Beyond this, investigations of this kind principally facilitate accessibility in the field of visual depictions of time-based processes – an area which has its origins in the environment of the Russian Formalists (in the 1920s) and which has gained a new immediacy with technological advances and the computer-aided ability to evaluate complex and extensive data sets. Methods which are supported by quantitative data acquisition and attempts at quantification are being given a new impetus by the integration of computer-based and software-based technology, which could be fruitful both for theoretical investigations and for practical implementation.

In the meantime, even conventional computers have made it possible to manage and visualise far more extensive data sets of individual images in higher quality than was the case only a few years ago. Film material is increasingly available in higher resolution, and automatic video analysis is constantly being further developed. In this area, academic film research is interwoven with digital databases, which no longer display only metadata, but can now be enabled simultaneously to deliver research tools in full image. My contribution is a more modest one, but it is to be hoped that the concrete example of Dziga Vertov – whose work has, not without reason, been compared to a database (Manovich 2001) – will show how computer science and information visualisation can be meaningfully applied to film-scholarly and film-historical analysis. Visualisation could thus be understood as a joint aspiration of artists and film scholars, as the desire to see more than a first glance at the performative, and thus temporally bound, art that film is can ever permit. It is no coincidence that what Vertov's *kinoglaz* concept celebrates as the basic principle of his film theory is precisely the camera as technical marvel, superior in every respect to the human eye and simultaneously capable of both overview and microscopic insight.

This book is based on a body of empirical data created between 2007 and 2010 in the course of the interdisciplinary research project Digital Formalism. This project was financed by the Vienna Science and Technology Fund (WWTF). The project partners were the Interactive Media Systems Group of the Vienna Technical University, the Institute of Theatre, Film and Media Studies of Vienna University and the Austrian Film Museum. Since then the academic landscape has undergone great change; the digital humanities have become institutionally established at Austrian and German universities, research centres have been founded, and associations for digital humanities have been brought into being. Exciting times have thus begun for those researchers who wish to work at the intersections of the disciplines. In 2007 none of these developments could yet be seen; today Digital Formalism is considered a pioneer project for the collaboration of the humanities and computer science, for the application of computer-aided analysis and the interpretation of the data that follows it. Some challenges that the scholars undertook in 2010 (the end of the project) are, however, still present and pressing. How does one arrive at knowledge from the quantitative data and in what form should it be formulated? What contribution can quantitative analysis make to film history? How can the data be depicted in order to enable the work of different disciplines?

In the Digital Formalism project, I created extensive annotations for eight of Vertov's films. I would like to describe both the software used and the process of manual annotation, as well as presenting the results. This data, which depicted Vertov's films to a degree of empirical description never previously achieved, could be put at Lev Manovich's disposal for continuing collaboration. From it, in a stimulating exchange, we jointly developed various forms of visualisation, which are chronologically presented in this volume. These experiments are intended, among other things, as a contribution to a fundamental discussion about the graphic depiction of filmic structures. Alongside early attempts which work with abstract depictions, we were ultimately able to arrive at the so-called direct (reduction-free) visualisations. For Manovich and myself, this method offers an exciting new approach to visual representations of films.

One aim of the project, and my continuing work which followed it, was also to develop the computer-aided methods on the basis of Vertov's writings and thus to develop his own terminology. I carried out the preliminary work for this in the Dziga Vertov Collection at the Austrian Film Museum, the most extensive part of the estate outside Russia, in which many handwritten documents, including letters, poems and diagrams, are preserved. Vertov's own graphs and diagrams constitute a valuable approach for the investigation of montage and rhythm in his films. Some of these documents have already been published, and in some cases annotated; a detailed explanation of the graphic design and concrete purpose as it relates to Vertov's filmic work is, however, for the most part, a gap that has yet to be filled. Using selected examples, an attempt will be made to close that gap at least for a few documents and to gain insight into the systematics of Vertov's recording. Generally, Vertov's films are held in differing versions, the lengths of which, in some cases, deviate from one another considerably.

Vertov's work is particularly suitable for formal investigation, as the director conceived his messages in formal procedures such as shot length, shot size, image composition or intensity of motion. In order to gain meaningful insight into Vertov and his films, this manual or computer-aided data analysis must, however, be coupled with film-historical knowledge and a study of sources. Critical work with sources enables essential information to be gained, for the state in which the film prints have been preserved, the precise analysis of the film material and familiarity with historical film techniques provide testimony regarding the archival policy and political culture of the Soviet Union in the 1920s and 1930s. Part of my work is thus devoted to researching the state of preservation of the prints of Vertov's first eight full-length films, the same body of work which was selected for Digital Formalism. My results will be embedded in a brief description of content and illuminated by the mirror of the contemporary press. In addition, formal results from the annotation will be presented in summary form and discussed.

The central section of my book will include three detailed studies of specific questions dealing with form and content. Through the analysis of the formal procedures and the functions with which Vertov invests them in the respective films, information may be gained regarding their varying significance. Vertov's use of faces in close-up is analysed and visualised in two of his films. I also devote a

detailed study to the interval, a term central to Vertov's film theory. An illustrative study of motion direction and motion intensity enables a closer look at this category, for which the research to date has adhered more to Vertov's writings. Finally, I investigate the different portrayals of the political leaders Lenin and Stalin in Vertov's films.

References

- Bernhard, Thomas. 1989. *Old Masters*. London: Penguin.
- Manovich, Lev. 2001. *The Language of New Media*. Cambridge, MA: MIT Press.
- . 8.12.2012. Twitter.
- Tsivian, Yuri. 2004b. Dziga Vertov and his Time. In *Lines of Resistance. Dziga Vertov and the Twenties*, ed. Yuri Tsivian. Sacile, Pordenone: Le Giornate del Cinema Muto.

Chapter 2

The Measurement of Aesthetic Phenomena



Film art is the art of the abstract word. We are abstract people. Every day splits us into ten activities. That is why we go to the cinema. (Tynjanov 2005: 242)

To say that a work of art may be broken down into smaller units may be a banal statement. It is also nothing new to portray or arrange these building blocks in one or another way, whether as part of the artistic process or due to a technical necessity during production. What for creative artists is an integral component of their work raises methodological questions in scholarship: is it academically even constructive to divide works of art into measurable units, to extract formalised data and transfer them to another system of recording in order to test theses or attain new statements? And if so, for which corpora and approaches would it be conceivable and meaningful? Different theoretical approaches have, after all, led to various methods in the study of film and literature.

Alongside the still strongly represented poststructuralist theories in university humanities research, a trend towards formalistic practice may again be observed. According to Vinzenz Hediger and Markus Stauff, who in 2011 co-edited the special issue “Empirie” for the *Zeitschrift für Medienwissenschaft*, the humanities and media studies are currently located between the pressure to justify themselves towards a traditional disciplinary self-image and the interest in new technologies. These tendencies are influenced on the one hand by other university disciplines society and on the other by society. Hediger and Stauff see only two alternative strategies as a way out: either one elevates the distance to empiricism to the status of a constitutive characteristic, representing the claim to legitimacy of the humanities and cultural studies, or accepting the supposed priority and dominance of empiricism, one qualifies one’s own research with further verifiability in the framework of a quantitative or scientifically precise procedure (Hediger, Stauff 2011: 10).

Whether such reservations towards formalistic or empirical methods can also be established geo-culturally, that is to say, in the sense of different academic traditions in individual countries, is not a question easily answered (also Lepenies 2006). At

least in the German-speaking regions, it seems one is more cautious than open regarding the potential. The film scholar Barbara Flückiger sums up these concerns in her essay “Die Vermessung ästhetischer Erscheinungen” and is not very optimistic:

Whether and how one may or should measure aesthetic objects is the subject of fundamental debate. Anyone wishing to deconstruct the inherent imprecision characteristic of all artistic work into measurable units is easily exposed to suspicion of reductionist positivism. Notwithstanding a long line of attempts, going back to the turn of the 20th century, to place philosophical aesthetics on an empirical-scientific basis, empiricism and aesthetics seem to find themselves still in a field of tension that can barely be overcome. (Flückiger 2011: 44)

The author, herself active in projects served by computer-aided film analysis, expressed, however, the commitment of empiricism to the tradition of the humanities “marked by a higher philosophical aesthetic”. She outlines the history of the empirical approach to aesthetic research in its disciplinary context in order to argue that a new tradition of computer-aided collaborative cooperation with web-based interfaces has developed out of the original psychological approach: the “digital humanities 2.0”. This then advances the humanities-oriented line of work-immanent analysis with formalistic or structuralist methodologies (Lepenies 2006: 45).

In the USA there are apparently fewer reservations; there one is more engaged with closing the gaps between the methodological and theoretical approaches. Thus, for example, Trevor Owens claims that it should not be outlandish for humanities scholars to see the objects of their research as potential data. In their analysis, one could definitely make use of methods that already exist in the humanities, for example, hermeneutics and interpretations. Subsequently, according to Owens, the data could undergo the same procedures one has thus far used for original texts:

We can choose to treat data as different kinds of things. First, as constructed things, data are a species of artifact. Second, as authored objects created for particular audiences, data can be interpreted as texts. Third, as computer-processable information, data can be computed in a whole host of ways to generate novel artifacts and texts which are then open to subsequent interpretation and analysis. Which brings us to evidence. Each of these approaches – data as text, artifact, and processable information – allows one to produce or uncover evidence that can support particular claims and arguments. Data is not in and of itself of evidence but a multifaceted object which can be mobilized as evidence in support of an argument. (Owens 2011)

This approach nonetheless makes it possible to redetermine the relationship between text and data and to discuss it productively. If there is a willingness on the part of the sciences not to understand data by definition as objective entities, the process of data collection and classification is potentially opened to the interests of the humanities. In this way, the definition of a procedure such as “data mining” can be broadened from the general and literal sense of “extracting something useful from a mountain of data”. The term is mainly used in the context of obtaining knowledge from databases, with the aim of deploying marketing strategies efficiently. Though this subject was hitherto mainly taught in information

management, an effort is now being made to apply it meaningfully within the humanities, e.g. the application of data mining in the case of film and television scripts,¹ or the analysis of colour in film.²

The Sciences and the Humanities: Culture Clash or Interdisciplinary Potential?

The acquisition of knowledge is traditionally evaluated differently in the humanities than are the results of research in the natural sciences, while at the same time, the results of research in the humanities are – just as traditionally – viewed with the suspicion of being in principle unscientific. Demarcated from “exploitable” measurable knowledge, the humanities consciously work in categories that are not easily expressed; discourse and the ability to grant validity to divergent convictions profoundly mark these disciplines.

The humanities’ immanent reflection on their own scholarship and methodology is considered quite constraining in interdisciplinary collaboration. The media scholar Lev Manovich expressed this polemically in an interview: “The problem with the humanities [...] is, that people tend to worry too much about what can’t be done, about mistakes, problems, as opposed to just going and doing something” (Williford 2011). Manovich is here already looking at joint research projects of the humanities and the sciences, as naturally software and computers can simplify and accelerate experiments by “trial and error” in data-based disciplines. Ideally, interdisciplinary projects combine the different approaches of the individual disciplines and work together on the problems. Increasingly frequently, therefore, interdisciplinary project applications are being developed and submitted: “The contacts between the humanities and technical studies are shifting the questions of what culture is, what characterises it and what determines its development ever more into the interest of research” (Schneider, Wedell 2004: 7). The humanities not infrequently also profit financially from such projects.

The step to interdisciplinary research can also be perceived as progress. In the context of the by now decades-long trials of strength and establishment of hierarchies at the universities, the focus of which is the humanities (above all, the philologies) and the sciences. One may put this debate in a wide historical context, beginning with the case of the historical pair of opposites in the case of Great Britain. Thomas Kühn thus understands the two-culture controversy as a conflict pattern of long duration,

¹Adam Ganz and Fionn Murtagh have delivered a lecture in Swansea in 2010 entitled “From Data Mining in Digital Humanities to New Methods of Analysis of Narrative and Semantics.” See also the work by Manuel Burghardt (<https://ch.uni-leipzig.de/burghardt/>).

²See Barbara Flückiger’s ERC Advanced Grant “FilmColors” (<http://www.research-projects.uzh.ch/p21207.htm>) and the work done by Niels-Oliver Walkowski and Johannes Pause (<urn:nbn:de:kobv:b4-opus4-25910>).

the origin of which is to be found in Victorian England with Thomas Henry Huxley and Matthew Arnold and which was taken up again in the 1950s by C. P. Snow and F. R. Leavis. Huxley's inaugural address at Mason Science College in Birmingham in 1880, entitled "Science and Culture", initiated the debate and was followed 2 years later by Arnold's reply in the form of the article "Literature and Science". The discourse was then resumed by Snow in 1959, this time with Leavis as sparring partner. Even though Kühn refers to the English politics of higher education and even though he, in accordance with society, perceived universities as a cultural system within the respective national culture, there is still a battle of power in evidence, valid for both the Western European and North American cases (Kühn 2002: 87).

It is the tension between the primacy of educating towards a profession (training) and the non-utilitarian "education" towards being a gentleman. This tension is on the one hand connected to a Utilitarianism aimed at financial success. On the other hand it is linked to a more idealistic contrary position, critical of civilisation and uninterested in the goal-oriented education at the universities, which it even considers damaging. (Ibid.)

The biologist Huxley, in 1880, still had to plead eloquently and, to a certain extent, quite polemically for the recognition of the sciences as a component of culture. His address voiced a close connection between the sciences and the economic exploitability of education, which seems more topical than ever today. However, he warned against a purely practical science and against a one-sided specialisation in the sciences and asserted that pure research could be wholly justified (ibid: 44). Naturally, the arguments have changed since then; pure research is today firmly anchored in faculties of science. Matters have not, then, remained at simple "training", which could be assumed to be the basic prerequisite for integration into university structures (unlike technical colleges).

Through the increasing integration of interdisciplinary projects, with a view, perhaps, to the concept of a "cyberinfrastructure" (also known as "e-science", "e-research" or "e-infrastructure") or in the new discipline of digital humanities, the humanities have themselves been taking an active step in the direction of a redefinition. This includes the trend towards new study programmes that make it possible to acquire basic knowledge in data modelling, the Java programming language or legal frameworks alongside the humanities specialisation.³ This creates possibilities for university education that initiates a convergence of education and training. The question still remains as to whether these two strands can meaningfully be linked, methodically and practically, and whether it ought, in general, to be a university aim increasingly to offer such courses. For Alexander von Humboldt and the German idealists, culture was more than the sum of the knowledge that could be taught; alongside it – as the result of study – one's own character should be cultivated and developed (Berry 2014: 52). Against the Humboldtian model of higher education, a utilitarian transmission of knowledge, which could lose its way in the description of modules and strategies to increase efficiency, is once more in the foreground of the current development (Nida-Rümelin 2006: 70).

³For example, at the Centre for Information Modelling at Graz University, cf. <https://information-smodellierung.uni-graz.at/de/zentrum/>. Last accessed 25 Aug 2015.

The traditional approach in research and publication practice, as well as the self-conception and problem-solving strategy, nonetheless remains, and varying views regarding “exploitable knowledge” arouse conflict in interdisciplinary collaboration. Kühn assumes that the basis of the debate is the conviction of the knowledge elite that one is challenged to find the best way to help society out of its crisis. This can take place only through specialisation and analytical thinking as practised in the sciences or by training and promoting a capacity for systematic thought that is free from practical application (Kühn 2002: 52). In Kühn’s summary, the controversy between Leavis and Snow is later mainly marked by the contrast between liberal-humanistic criticism of civilisation on the one hand and by utilitarian optimistic scientific positivism on the other (ibid.: 43).

From “Cyberinfrastructure” to the Digital Humanities

It should be said at the outset that what precisely one understands by digital humanities is still the subject of lively debate on the Internet, in articles and anthologies, as well as at conferences and discussions. The statements there are as provocative as they are interesting and, for the present, at any rate, largely limited to the Anglo-American sphere. Despite the abundance of publications, some contributions have already achieved “cult” status, some voices have established themselves as leaders, and there are manifestos with great entertainment value.

In 2009 two American scholars published a foundational text, the so-called Digital Humanities Manifesto, which humorously outlines the new discipline. In it, Jeffrey Schnapp, director of the Harvard metaLAB, and Todd Presner, professor of German Languages and Comparative Literature at UCLA, describe, among other things, the attitude of digital humanities to the traditional and the new locations of the acquisition and dissemination of knowledge:

Digital Humanities is not a unified field but an array of convergent practices that explore a universe in which: a) print is no longer the exclusive or the normative medium in which knowledge is produced and/or disseminated; instead, print finds itself absorbed into new, multimedia configurations; and b) digital tools, techniques, and media have altered the production and dissemination of knowledge in the arts, human and social sciences. (Schnapp, Presner 2009)

The debates within the field still, perhaps more than ever, revolve around a possible definition of its own discipline, which moved Dave Parry to call such publications “something of a genre essay” (Parry 2012: 429). The titles of anthologies on this subject, too, for example, *Understanding Digital Humanities* (2012), *Defining Digital Humanities* (2014) or *Debates in the Digital Humanities* (2012), are testimony to the lively discussion among scholars about their own object of study. With few exceptions, however, voices from European academia are (still) lacking. In addition, the contributions seem to come overwhelmingly from the humanities, a few from libraries and archives, and there are only isolated examples from other disciplines, such as information visualisation. The discussion thus seems to remain, for now, in its own field.

The digital humanities emerged, over the course of many years, from the previous *Humanities Computing*, which a reading of *Defining Digital Humanities* makes clear. It is therefore unremarkable that the choice of the term digital humanities was analogous to the introduction of a brand and just as carefully planned. In the provocative article “Digital Humanities As/Is a Tactical Term”, Matthew Kirschenbaum, director of the Maryland Institute for Technology in the Humanities (MITH), argues that one decided on this name because it fulfilled two basic functions: “[It] possessed enough currency and escape velocity to penetrate layers of administrative strata to get funds allocated, initiatives under way, and plans set in motion. On the other hand, it is a populist term, self-identified and self-perpetuating through the algorithmic structures of contemporary social media” (Kirschenbaum 2012: 417). In the meantime, the digital humanities can already look back at an interesting history in several phases of development. Originally the computer was understood as a purely technical support for the traditional humanities scholars and less as “participant” with its own critical potential (Berry 2012: 3). The beginnings of the advancing fundamental restructuring of knowledge and the academic worlds can be helpfully dated to 2003 with the “Atkins Report”. At this time a panel commissioned by the National Science Foundation (NSF) published this groundbreaking report, colloquially named after then chairman Dan Atkins (2003). The NSF, founded in 1950, is an independent American government organisation tasked with promoting basic research across disciplinary borders. In two subsequent reports from the years 2006 and 2007, published by the NSF and the American Council of Learned Societies (ACLS), the term was conclusively defined and disseminated. The ACLS is a private non-profit association of over 71 national scientific organisations, which has an ambitious vision for the NSF, as follows:

At the heart of the cyberinfrastructure vision is the development of a cultural community that supports peer-to-peer collaboration and new modes of education based upon broad and open access to leadership computing; data and information resources; online instruments and observatories; and visualization and collaboration services. Cyberinfrastructure enables distributed knowledge communities that collaborate and communicate across disciplines, distances and cultures. (National Science Foundation 2007)⁴

This NSF report basically formulates four key areas, which were to be urgently dealt with in the period from 2006 to 2010: (1) “high performance computing”; (2) “data, data analysis and visualisation”; (3) “virtual organisations for distributed communities”; (4) “learning and workforce development.” In academic projects “cyberinfrastructure” was primarily to be used in the development of technological processes, in order to approach a solution for the problem of efficient and meaningful networking of data, computers and people with the aim of generating new academic

⁴A further central report assigns the “cyberinfrastructure” a role within the humanities and social sciences, cf. American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences: *Our Cultural Commonwealth. The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences.* 2006. URL: http://www.acls.org/uploadedFiles/Publications/Programs/Our_Cultural_Commonwealth.pdf. Last accessed 27 Aug 2014.

theories and knowledge. With libraries, archives and museums, systems had already been established over centuries to make finding information possible; only bibliographies, finding aids, citation systems and concordances are mentioned here. In publications, too, this information was made available, thus linking the publishers, librarians, archivists and curators with the researchers.

While the existing academic infrastructure was established over a period of centuries as an active participant in “science”, “cyberinfrastructure” developed much faster and more abruptly. For this reason, the NSF has considered it important to involve academics in this process right from the start. Even though this refers only to the American initiative, there are tendencies that may be seen within the framework of the European Union, for example, the construction of the “Europeana” Internet platform with all its subprojects, in particular, at the present time, with the “Horizon 2020” programme.⁵

In this context, the humanities are called upon actively to articulate their requirements and to reflect on the form in which use of the new digital tools is meaningful. Up for discussion, among other things, is where and how, within the humanities, quantitative data may be determinative and meaningfully depicted, alongside the qualitative. The Internet as the bearer of hope for the swift location of publicly accessible information, linked to correct and relevant metadata, is also a great opportunity for humanities scholars, for it is appropriate to develop ontologies and to collaborate on new classifications and standards for the management and utilisation of data records.

In order to be able to arrive at a more precise assessment of what such a contribution could consist, the ACLS was assigned to deal with the specific requirements and tasks of scholars in the humanities and social sciences. The humanities could and should make an active contribution. Their role as a reflective discipline was recognised, for:

after all, science – whose goal is predictive certainty – only has half the picture. Uncertainty (or ambiguity, if you prefer) is the other half, and the humanities and social sciences celebrate that, explore it, tolerate it, and understand it better than the sciences do. Or, at another level, if science and engineering are about what we can do, the humanities and social sciences are about what we should do. (Unsworth: 2004)

The humanities were thus perceived and confirmed as an important contribution to the solution of urgent societal problems, for “the study of history, literature, languages, philosophy and other humanities subjects help us not only to better understand our own nation, but other cultures as well” (ibid.). This may be read as a deft line of argument in a globalised world, which is dependent on understanding between peoples and cultures. The dean of Library and Information Sciences at Brandeis University, John M. Unsworth, additionally states that computational methods have meanwhile come to occupy a fixed and meaningful place in the social sciences, which could in future be expanded in the area of research into the history of literature and art.

⁵ Cf. <http://ec.europa.eu/programmes/horizon2020/>. Last accessed 8 Aug 2018.

Unsworth goes further still and outlines a visionary image of a universally educated humanities scholar, adept not only within the sphere of that discipline but equally in technical knowledge. Only thus can one prevent, for example, that projects are carried on as pure product development:

We will need English majors who have a background in logic, who can handle statistics, who do maths, if we are going to turn out a generation of disciplinary specialists who can bring the accumulated wisdom of the humanities to bear the computational contexts – perhaps in helping build ontologies for scholarly projects in disciplinary contexts, or building tools for data-mining in the context of humanities research. (Ibid.)

Ten years later, it seems that Unsworth's vision has arrived at international universities and extramural research facilities. Above all, it is in the USA that study programmes have been established that, among other things, explore the possibilities of using software and other technological products.⁶ One could also mention Stanford University and King's College London, which are designing study programmes in the framework of digital humanities. With a slight delay, the digital humanities have also arrived in Europe. The foundation of the association Digital Humanities in the German-Speaking World (Dhd) in 2012 marks a milestone in the disciplinary entrenchment of the new subject within the universities and newly established research centres. At the same time, the Association for Literary and Linguistic Computing (ALLC), which had existed since 1973, changed its name to the European Association for Digital Humanities (EADH). The Dhd's first conference in 2014 still tellingly bore the title "Digital Humanities – Bridge Building or Hostile Takeover?", while the second annual conference took place under the considerably more positive motto "From Data to Realisation: Digital Humanities as Mediator Between Information and Interpretation".

In the digital humanities, there is not yet a unity of opinion on whether it is sufficient for metalevel research if scholars understand computer software only to a certain degree. Lev Manovich, founder and director of the Software Studies Initiative in San Diego, is one of the few who speaks out explicitly regarding the necessity of practical training. In his opinion, humanities scholars should be capable, even without integration in interdisciplinary support programmes, of conducting certain analyses independently:

if every data-intensive humanities project required a research scholarship that would make such a collaboration [between the humanities and information technology] possible, we would only be able to advance very slowly. We want humanities scholars to be capable of using data-analysis and visualisation software in their daily work, in order to be able to combine quantitative and qualitative methods in their work. How we get to that point is one of the key questions for "digital humanities." (Manovich 2014: 81)

⁶Lisa Spiro, director of the National Institute for Technology in Liberal Education (NITLE), has put together a very useful document in this context: it lists useful addresses and institutions on topics from workshops, tutorials and best practice documents to the planning of one's own projects (cf. Lisa Spiro: Getting Started in Digital Humanities. In: Digital Scholarship in the Humanities, 14.10.2011. URL: <https://digitalscholarship.wordpress.com/2011/10/14/getting-started-in-the-digital-humanities/>. Last accessed 8 Aug 2018.

In accordance with this idea, he teaches these areas at the City University of New York (CUNY). Manovich’s view is also shared by important American funding bodies. The director of the Office for Digital Humanities at the National Endowment for the Humanities, Brett Bobley, thus went on the record that he would like to see more practical exercises in “digital tools and methodologies for humanities scholarship” (Gavin, Smith 2012: 64) and posed the question: “How many graduate humanities programs include classes on using GIS, 3-D modeling, data analysis, or other methods of scholarship?” (Ibid.) A proposal by film scholar Nick Redfern, in support of basic statistical training in film studies, points in the same direction (2013: 60). He attaches importance to the statement that it is not sufficient to train the students to be users of statistics programmes; statistics, rather, means critical thinking in a comprehensive sense. One could classify Manovich’s courses as part of this direction, including, as they do, not only practical exercises within the study plan but also principle questions regarding the presentability of data and, not least, aesthetics. When one reads the theses of the working programme for the further development of the DHd into the year 2020, it is not quite clear, however, whether the digital humanities are to be seen as an applied branch of information science or still as part of the humanities.

1.1 The digital humanities enrich the traditional humanities conceptually and methodologically – their tools and methods complement the “how” of our practice with an empirically oriented epistemology.

1.2 For all their methodological and theoretical claims, the digital humanities are nonetheless characterised by a pragmatic orientation. The development and preparation of tools of information technology are therefore among their central features. (DHd 2014)

The debate within the digital humanities as to *which* branch of information science could best be deployed here, or whether one is once more dealing with a new discipline, continues controversially, as was impressively shown by a disputation in the framework of the first DHd conference in 2014. One of the protagonists was Manfred Thaller, professor of Computer Science for the Humanities at Cologne University, who represented innovative propositions about the positioning of the humanities towards information science. Text-based research is for Thaller only one of four areas in which he sees potential for cooperation between the humanities and information science. Further sub-groups are factual analysis, the analysis of non-textual information in its broadest sense (this subsumes audiovisual data) and “humanities computer science” (Thaller 2014). Thaller is also doubtful that computer linguistics/corpus linguistics/text mining can be included in the digital humanities: “There are, rather, good reasons to assume that they constitute their own discipline, the methods of which, while certainly useful for the digital humanities, can on no account be considered part of its core” (ibid.).

The new assignment of the name digital humanities can be seen not only as an application of computer technology in the humanities but also as a step towards integration. This goes with a revision of the traditional humanities, the methodological strengths of which are “attention to complexity, medium specificity, historical context, analytical depth, critique and interpretation” (Schnapp, Presner 2009). The conventional institutions and paths of transmitting knowledge are adjudged by the

“Digital Humanities Manifesto” as no longer up-to-date; they must face up to the changes in the academic environment: “Today the old theory/praxis debates no longer resonate. Knowledge assumes multiple forms; it inhabits the interstices and criss-crossings between words, sounds, smells, maps, diagrams, installations, environment, data repositories, tables, and objects” (ibid.). The digital humanities have thus entered a phase of critical self-reflection regarding their field of research and their academic practices. David M. Berry prognosticates about the direction the development will take: “Indeed, we could say that a third-wave digital humanities points to the way in which digital technology highlights the anomalies generated in a humanities research project and which leads to the questioning of the assumptions implicit in such research, for example close reading, canon formation, periodisation, liberal humanism and so forth” (Berry 2012: 5).

What can here already be heard, even if it is not explicitly said, is the circumstance of having for the moment reached the end of the pioneering era, in which the most varied projects and initiatives were welcome (a “big-tent” approach). The trend towards a large number of very varied contributions is presently still continuing, a fact which is correctly seen as problematic. Not only does one lose an overview of conferences and publications, there is also a risk of academic discussion becoming bogged down, as the individual areas are so widely dispersed. An additional problem that should not be underestimated is the choice of suitable peer reviewers who possess the requisite knowledge of the subject, in order to be able to judge the quality of specific themes (Terras 2014: 268).

On the other hand, according to Berry (2012: 5), one meanwhile devotes oneself to the establishment of standards and best practice models, both in research and in the evaluation of knowledge. Not least, this brings with it a discursive convergence with the methods and approaches of the traditional humanities. It is about fundamentally understanding how processes of knowledge transformation in the twenty-first century can take place and be understood by means of computer technology and digitalisation. In doing so, interpretation must remain an important component of the discipline, for:

What the community can do with the results of a digital humanities project is, like art, often outside what a creator or project team might have envisioned for it – and this is where the interpretation becomes important for multivalent digital humanities projects. What does it mean that a database has been structured in a certain way? What are the larger consequences for one design over another? How does a certain project push the boundaries of what we consider acceptable digital humanities work? How can new analytical processes or methodologies be applied in different contexts? These are subjective and interpretative questions that we must openly discuss. (Gibbs 2012)

Whether one smiles at the new discipline, greets it enthusiastically or simply ignores it, it may be stated that in the last several years, a revolution in the acquisition, transmission and storage of knowledge has been on the march, the effects of which we are already feeling: in the funding programmes of the EU or national sponsors and in the open-access movement or the development of the Semantic Web. Against this, the academic structures have remained relatively immobile. The call to the universities made by Schnapp and Presner to tread a productive path to

collaboration with nonacademic institutions and break up the traditional hierarchies can be seen as an example for the innovative demands that are valid for the new discipline.

The Digital Humanities seek to play an inaugural role with respect to a world in which universities – no longer the sole producers, stewards, and disseminators of knowledge or culture – are called upon to shape natively digital models of scholarly discourse for the newly emergent public spheres of the present era (the www, the blogosphere, digital libraries, etc.), to model excellence and innovation in these domains, and to facilitate the formation of networks of knowledge production, exchange, and dissemination that are, at once, global and local. (Schnapp, Presner 2009)

Both authors correctly point out that research certainly takes place in libraries and archives and that these have their own well-functioning networks for the exchange of information. Unfortunately, the traditional hierarchies are still very strongly anchored in people’s consciousness, for example, according to Schnapp and Presner, because the modern university separates research from curation. Curation is thus assigned only a secondary, supporting role, and in this way, curators within the museums, archives and libraries are “sent into exile” (ibid.). In contrast, the digital humanities expressly call for a new definition of the “scholar as curator and the curator as scholar”, (ibid.) which should include the academic activities of institutions such as museums, libraries and archives. Despite the noble claims of both authors, the roles continue to be assigned in a relatively fixed manner, and it is apparently still difficult to consider extramural research. This partially also unconscious exclusion is one that I have already dealt with elsewhere in the context of the situation in film archives and libraries (Heftberger 2014; Heftberger 2018b).

The Digital Humanities in Film Archives and Libraries

The long-term networking of academics with librarians and specialists in the field of information technology is an explicit formulated goal of cyberinfrastructure. The emphasis has hitherto been on text-based artifacts, but the establishment of networks for electronic resources or for management, for exchange and for optimisation of metadata in different fields has also been promoted. Archives and libraries already increasingly use digital tools in order to present parts of their collections meaningfully on the Internet and thus also provide the public with the possibility of searching and researching in their holdings.⁷

In the meantime it has become possible to annotate video files, to provide them with keywords to the precision of a frame, to furnish them with geographical data and to comment on them in free text. The most diverse web applications and database systems are used and developed, which makes it difficult, in these already relatively

⁷For example, the project BFI Filmography at <https://filmography.bfi.org.uk/>, where filmographic data can be visualised. Last accessed 8 Aug 2018.

unstandardised areas, to establish internationally binding standards (metadata, codecs, etc.). Research can take place at the level of content but also at the levels of metadata generation and meaningful web presentations.

Film archives are, depending on their type and national conditions, organised in very different ways. For the most part, however, a division of roles has emerged which differentiates between technical expertise and work with content. Even if this is a gross oversimplification, this differentiation is even more strongly pronounced in the Anglo-American film archives than in those of Eastern Europe, which normally employ a relatively large number of academic staff. Although the term digital humanities is – compared to many libraries – scarcely an issue in film archives, it may be assumed that there, too, a greater permeability between the areas of responsibility is to be expected in the future. It is already the case that different fields are no longer as sharply delineated from one another, as digital data management cannot be restricted to one area of material (e.g. films, photos or documents) and metadata must be managed across the boundaries between collections, among other reasons in order optimally to use technical infrastructures and personnel resources. A great step forwards in the standardisation of metadata necessary for this has been made in recent years, of which more later.

The subdivision into “digital” on the one hand and “analogue” on the other makes, to put it provocatively, no sense anymore, and the traditional image of archivists, too, bringing down reels of film from shelves, is included in the rapid transformation. This applies above all to the staff of audiovisual archives, in which the entire field is still in need of definition, according to Martin Koerber, head of the film collection at the Stiftung Deutsche Kinemathek:

Compared to other heritage archivists, audiovisual archivists, and audiovisual restoration experts in particular are still in a minority position. Often their demands to be accepted as heritage specialists go unappreciated by their institutions and by the heritage fields as a whole. Audiovisual archivists have yet to define their field, and due to the continuing technological change, to constantly redefine who they are and what they do will be a key challenge for the foreseeable future. (Koerber 2013: 46)

The borders between university research positions and curatorial professions in the archives or extramural facilities, which can no longer be so clearly demarcated, have also already been dealt with in the aforementioned “Digital Humanities Manifesto 2.0”. Curatorship is there defined as follows: “Curatorship means making arguments through objects as well as words, images, and sounds” (Schnapp, Presner 2009). Although a similar process of “spatialisation” is taking place here as in the humanities, above all in history, there is still a fundamental difference. Instead of working with language, one works with space, in which physical or virtual objects are arranged. In spite of that, the authors emphasise the similarities and the potential that just such a self-transforming environment can have for a mutual exchange of impulses:

It means becoming engaged in collecting, assembling, sifting, structuring, and interpreting corpora. All of which is to say that we consider curation on a par with traditional narrative scholarship. It is a medium with its own distinctive language, skill sets, and complexities; a medium currently in a phase of transformation and expansion as virtual galleries, learning environments, and worlds become important features of the scholarly landscape. (Ibid.)

But we can and should not restrict the process of curation only to the assembly of a film programme or an exhibition. It could be expanded to the identification, cataloguing and online presentation of archive material such as films and photos from the collections. It is precisely the identification that is often a slowly and sometimes also frustrating process, as personnel resources are limited and specific expertise and the necessary time are lacking. For this reason, film archives with interesting online projects have turned to the public, in order to be able to obtain expertise from specialists outside the institution and beyond the national borders as well. As examples of such crowdsourcing initiatives, one could mention the Deutsche Kinemathek's ambitious LOST-FILMS project⁸ or the Austrian Film Museum's Schlemmer frame collection.⁹ From experience, one can say that a relatively high investment (including financial) is linked to projects such as those mentioned. There is also the risk that identical documents in other institutions have been digitalised and put online with the same intentions without being linked or displayed in a search. The digital humanities could serve to provide an important aid to archives, ensuring and optimising the collection and exchange of information. The generation of metadata or the location of similar subjects across databases and collections could benefit in like manner. Researchers are often interested in specific people, objects or activities that appear in films. In the current environment, the answering of such requests is a task that can scarcely be resolved for archive staff. Apart from personal knowledge, which may naturally be very comprehensive, there are hardly any aids. To formulate it exaggeratedly, a curator who can simultaneously develop digital tools is a dream of the future but would be a desirable enrichment.

This naturally raises the question of where students should acquire the kind of knowledge that embraces both the science of materials and film-specific computer-technical abilities. Much of it can only be acquired through long years of experience. Other parts of it, for example, skills in digitisation, databases, programming or text mining can perhaps not even be covered by university studies in the archival or library fields (even should the will be there) (also Heftberger 2018a).¹⁰ Julia Flanders of the Brown University Library describes, in her informative article "Time, Labor and 'Alternate Careers' in Digital Humanities Knowledge Work", an additional phenomenon: "Most digital humanities work, however – as performed by library staff, IT staff, and other para-academic staff who are not faculty – is conceptualized according to one of the other models: hourly, by FTE, or as an agenda of projects that granularizes and regulates the work in quantifiable ways" (Flanders 2012: 303).

⁸ See <https://www.lost-films.eu/>. Last accessed 8 Aug 2018.

⁹ See https://www.filmmuseum.at/en/collections/special_collections/schlemmer_frame_collection. Last accessed 8 Aug 2018.

¹⁰ One could, at any rate, mention the information and library science courses at the FH Potsdam and the HU Berlin, which, however, do not specialise in film, as well as those at the Film University Babelsberg *KONRAD WOLF*, the Goethe University in Frankfurt am Main, the HTW Berlin and Amsterdam University, which all have different areas of specialisation.

She points out the circumstance that there is a tendency in extramural institutions to define assignments in the framework of digital humanities (e.g. databases, online presentations, also perhaps long-term archiving) as projects and to farm them out to third parties. Although such a procedure is understandable, a chance is being missed: the opportunity to build in-house resources which would also enable engagement with collection content alongside the digital infrastructure and information management. Conversely, there are many academics who have in the meantime acquired a technical competence especially for humanities problems. In both cases, important knowledge is always being removed from the institution when the project is completed and must then be “purchased” again.

To sum up, and in reference once more to Schnapp and Presner, one could say that the ideas in the manifesto are illuminating and perhaps even revolutionary, especially as far as the redefinition of the relationship between universities and other knowledge-producing institutions goes. However, a realistic look at the situation shows there is room for improvement. The same goes for the democratic call to break up long-established hierarchies, for example, through the integration of independent researchers. The current debates in the digital humanities do not necessarily seem to be particularly interested in these topics but are devoted more to the question of academic publication possibilities in the digital age, the staking out of disciplinary boundaries and overlaps with the traditional humanities or methodological discussions, such as quantitative analysis versus hermeneutics. Although the importance of such questions is not disputed, the inclusion here of new partners in the discussion could be fruitful.

Big Data: Distant Reading

Digital collections, as they are today, alter not only the individuals who deal with them; they “also deliver destabilising quantities of knowledge and information which lack the regulatory power of philosophy – which, as Kant explained, guarantees that institutions remain rational” (Berry 2014: 53). If we understand Berry to mean that the digital humanities are called upon actively to accept the so-called big data,¹¹ this is not without practical obstacles, depending on which data interest us (Manovich 2014: 79). Whether the humanities can really design a philosophy which – and this is the crux – can also show a regulatory effect outside the academic world is doubtful. There are already too many economic interests involved which manage the collection and storage of “social data”. Even though on the one hand there is a sheer oppressive overflow of data, the demand nonetheless virulently continues for more digitised documents, including those of the film heritage. In this

¹¹ For a definition cf. https://en.wikipedia.org/wiki/Big_data. Last accessed 8 Aug 2018, Manovich also refers to Wikipedia.

book I shall be discussing only a small part of the available volume of data and concentrate more on why there are reservations in the humanities about going from a “close reading” to a “distant reading” or to use both in parallel.

The technical possibilities for computer-aided analysis and perhaps even interpretation represent a radical innovation for the humanities. Still more radical is the approach of analysing not one work after another but hundreds or even thousands of human creations simultaneously and at a speed no human could achieve. Parallel to this is a variety of methodologies, in which interdisciplinary approaches mutually stimulate and complement each other. The nature and extent of human participation in digital humanities projects is therefore the subject of controversial discussion within the humanities. For her overview article “How We Think”, the literature scholar N. Katherine Hayles interviewed her colleagues about the possibilities and limits of the digital humanities. The author wanted to find out how the use of digital technology influenced and changed the thinking of her peers in the humanities. Hayles was interested above all in the changes in the individual conditions of work in the framework of the respective university environments: “How engagements with digital technologies are affecting the assumptions and presuppositions of humanities scholars, including their visions of themselves as professional practitioners, their relations to the field, and their hopes and fears for the future” (Hayles 2012: 42). Hayles considers the results, which took a variety of forms, astounding and arranged them in large units: scale, critical/productive theory, cooperation, databases, multimodal scholarship, code and future trajectories (ibid.: 43).

For Hayles, the statements dealing with “scale” were in this context the most important. An autocatalytic process was already set in motion by the mere fact *that* we use digital technologies: “The more we use computers, the more we need the large-scale analysis they enable to cope with enormous data sets, and the more we need them, the more inclined we are to use them to make yet more data accessible and machine-readable” (ibid.: 48). Without evaluating, Hayles describes databases as the cultural expression of the new era, which transforms itself in a similar fashion to them: “Databases are not necessarily more objective than arguments, but they are different kinds of cultural forms, embodying different cognitive, technical, psychological and artistic modalities and offering different ways to instantiate concepts, structure experience, and embody values” (ibid., also Vesna 2007).

The basic tendency within the digital humanities, to grasp the object of research, among other things, as quantifiable, analysable and visualisable elements of data, would appear to contribute to the impetus of the new discipline. At least in American universities there are already signs of a desire for an improvement in the testability of statements in the humanities. A polemic commentary by the philosopher John Holbo could be seen as a representative of the current atmosphere: “If the answer is that literary scholars take the undesirability of quantification for granted, whereas everyone else takes its desirability for granted, the literary folks are flat out of luck” (Holbo 2011: 9).

With the volume of data that is suddenly available, the potential questions that can be directed at the material are multiplied. Theoretically, many sources are meanwhile available with which to support or falsify theses. The choice of scale for the penetration of an area, as already mentioned, also plays a large role; for Hayles this is a good opportunity to rethink established “bad habits” in everyday university life. She criticises the college practice, widespread at American universities, for the same books always to be placed on the reading lists for literature study. This could lead to a loss of the knowledge of how the “canon” is to be distinguished from “normal” literature, as it is precisely because of its unusual construction that it was singled out. Hayles’ criticism is not directed at the fact of the reading being restricted to a small number of selected works but at the fact that in consequence the questions raised become similar. Computer-aided analysis could potentially open other perspectives, as the body of work is more extensive.

Someone who has only read those texts will likely have a distorted sense of how »ordinary« texts differ from canonized works. By contrast, as Gregory Crane observes, machine queries enable one to get a sense of the background conventions against which memorable works of literature emerge. Remarkable works endure in part because they complicate, modify, extend and subvert conventions, rising about the mundane works that surrounded them in their original contexts. Scale changes not only the quantities of texts that can be interrogated but also the contexts and contents of the questions. (Hayles 2012: 46)

One of the most vehement and polemical advocates for dealing “rationally” with works of art is the literature scholar Franco Moretti. He expressed himself clearly in favour of a formal quantitative approach in literary research, encompassing both computer-aided statistical evaluations and thematic data collections. In this context, one of his key terms is “international cooperation”, for without it, according to Moretti, quantitative work is simply inconceivable. Not only because the task of creating all the records with one’s own labour would be interminable but also because the data is in this way independent of the person who created it (Moretti 2007: 5). In his opinion, data acquisition must initially be independent of any interpretative propositions. At the same time, what is also being called into question is one of the most privileged terms in the humanities – the process of reading. For Hayles the question arises of the extent to which one can also speak of an algorithm as a form of reading. One of the most radical pioneers of the digital humanities is the philosopher of science Timothy Lenoir, for whom a computer programme is better than analysis by a human, as no preconceived opinions filter the “material”, culminating only in the expected results. Any human intervention, according to Lenoir, is less expedient, if what one seeks are objective results. As he himself polemically puts it: “I am totally against ontologies” (Hayles 2012: 46). Algorithms, however, are written according to previously determined categories.

Moretti pits the traditional perusal of “close reading”, that is, the thorough and detailed reading and understanding of a text, against his concept of “distant reading”. He is primarily interested in literary processes over large periods of time, for which he gathers individual data in a graphic representation. Thus, for example, a change of method is required in order to be able to research nineteenth-century English literature at all, for a “field this large cannot be understood by stitching