Martin Siefkes

Mutual Attraction and Common Interests: The digital humanities and multimodality research have found each other (but will it last?)

Abstract (German)

Dieser Beitrag untersucht den Überschneidungsbereich zwischen den Digital Humanities und der Multimodalitätsforschung. Beides sind interdisziplinäre Ansätze, die mittlerweile in weiten Teilen der Geisteswissenschaften eine Rolle spielen. In den Digital Humanities gewinnen Projekte, Tools und methodologische Ansätze, die den traditionellen Fokus auf Schriftsprache hinter sich lassen, an Relevanz. Digitale Editionen, multimodale Korpora, interaktive Visualisierungen, virtuelle Rekonstruktionen und der Einsatz von Computer Vision sind Beispiele für diesen Trend.

Die Multimodalitätsforschung hat sich in den letzten Jahrzehnten ausgehend von linguistischen und semiotischen Wurzeln entwickelt. Sie besitzt die Theoriegrundlage, methodologische Reflexion und analytischen Werkzeuge, um die vielfältigen Zeichenprozesse zu beschreiben, die multimodale Artefakte und den Umgang mit ihnen kennzeichnen. So verändern sich die semiotischen Eigenschaften multimodaler Texte und Artefakte durch die Digitalisierung (etwa wenn ein Faksimile gescannt oder ein Gebäude virtuell rekonstruiert wird), und dies beeinflusst die weiteren Interpretationsprozesse (etwa wenn mit Deep-Learning-Algorithmen Muster in digitalisierten Sammlungen erkannt und diese als Evidenz für oder gegen Forschungshypothesen interpretiert werden). Der Beitrag geht auch darauf ein, inwieweit die Digital Humanities von Methoden und Technologieentwicklung der Data Science profitieren können.

Abschließend stellt der Artikel das neu entwickelte Multimodal Digital Humanities Framework (MDHF) vor. Mit seiner Hilfe lassen sich zahlreiche Aspekte von DH-Projekten in Einzel- und Vergleichsstudien oder in Korpusanalysen untersuchen. Es bietet eine breite Palette von Anwendungsmöglichkeiten, wie z. B. die Definition von Standards, die Spezifikation von Anforderungen im Planungsprozess, oder die Evaluierung laufender oder abgeschlossener Projekte. Für Forschende bietet es unter anderem die Möglichkeit, diachrone Untersuchungen zur Entwicklung von digitalen Editionen, digitalen Ausstellungen, DH-Visualisierungen oder anderen Angeboten der Digital Humanities durchzuführen.

Keywords. Digital Humanities, digitale Edition, Multimodalität, Semiotik, Multimodal Digital Humanities Framework, Data Science, Machine Learning, Computer Vision, Visualisierung

Abstract (English)

This contribution investigates the area of overlap between the digital humanities and multimodality research, two interdisciplinary approaches that exert their influence across the humanities. An increasing number of projects, tools, and methodologies in the digital humanities are moving

beyond the traditional focus on language and written text. Digital editions, multimodal corpora, interactive visualisations, virtual reconstructions, and the use of computer vision methods are only some of these developments.

Well-founded upon a basis in linguistics and semiotics, multimodality research offers the theoretical depth, methodological reflection, and analytical tools needed to describe the manifold interacting sign processes that are relevant for multimodal cultural artefacts. It can explain and analyse the change of semiotic properties through processes of digitisation, both of the initial multimodal texts or artefacts themselves (e.g. when a facsimile is scanned or a building virtually reconstructed), and of the processes of interpretation (e.g. when patterns are detected in digitised collections with deep learning algorithms, and these patterns are then interpreted as empirical evidence for or against research hypotheses). The potential of a deeper connection between the digital humanities and data science will be touched upon.

Finally, the article introduces the newly developed Multimodal Digital Humanities Framework (MDHF), an detailed analytical tool suitable for both individual and corpus-based analysis of DH projects. It offers a wide range of applications, such as defining standards to ensure comparability across projects, or formulating design specifications during the planning process. It can also be used to evaluate ongoing or finished projects, or to conduct diachronic research on the development of digital editions, digital exhibitions, DH-themed visualisations, or other digital humanities services.

Keywords. digital humanities, digital edition, multimodality, semiotics, Multimodal Digital Humanities Framework, data science, machine learning, computer vision, visualisation, ménage à trois

Introduction: two in love

This contribution will demonstrate that there is a hidden romance going on between the digital humanities and multimodality research. Both research areas – the term "field" may not be justified, at least not for the digital humanities – can be characterised as interdisciplinary endeavours that are driven by methodological innovation, and whose relevance and rapid success has been fuelled by changes in contemporary culture and media landscapes.

Media products, such as websites, social media or video games, are becoming more and more complex. Traditional methods of linguistics that investigate language in isolation from other semiotic modes, and media studies approaches that disregard the semiotic complexity of the different signs, codes, and multimodal genres and their interactions, are no longer adequate for analysing contemporary media ecologies (Siefkes in print, Ch. I; Fricke in this volume). Furthermore, the trend towards digitisation of both existing cultural artefacts and newly created cultural products, in combination with recently introduced research methods such as the large-scale analysis of corpora

(which increasingly comprise not just text, but multimodal data of various kinds) with methods of data science, poses new challenges.

It hardly needs to be mentioned that the traditional humanities, both regarding their subject matter(s) and the theories and methods they brought to bear on them, were always closely connected to the means of (cultural) production of the time. These included technologies, social and economic conditions, cultural expectations, and of course the signs and codes developed for producing and reproducing them. Significant changes, often due to the advent of new technologies and their ripples and ramifications throughout the whole network, were often perceived as highly disruptive at the time. It is therefore possible that from a future vantage point, the digital humanities will simply look like the newest update of a cultural system that has always been closely bound to technological means and social ends, and implemented with semiotic systems of various kinds.

In this context, multimodality research and the digital humanities may have been destined to meet, and indeed they have been flirting with each other for some time. Nonetheless, only a minority of multimodality researchers can already claim to have done work in the digital humanities. In the other direction, most digital humanists may never have heard of multimodality research, or if they have, may not possess a clear understanding of what it is or why it could be useful for them. Arguably, however, there is an area where both fields overlap regarding their research domains and questions, tools, and methods of analysis. This has resulted in profitable cooperations, research innovations, and some exchange of methodology and terminology.[1]

It should also be noted that both fields are in some respects complementary to each other. The primary strengths of multimodality research are its solid theoretical foundations, careful terminological distinctions, and empirical focus which have resulted in a number of fairly systematic approaches in recent years (Bateman et al. 2017; Björkvall 2012; Fricke 2012, 2013; Jewitt 2014; Siefkes 2015). These are combined with research methods from corpus linguistics, psychology, media reception studies, and social science in the empirical analysis of all kinds of textuality and communication situations. The main strength of the digital humanities, on the other hand, is their methodological innovation which also leads to new research questions and publication formats, but there have been extensive discussions about the need for a thorough theoretical basis (Bateman 2017; Gold 2012; Svensson 2009).

1 A thorough overview will be given in an upcoming monograph (Siefkes in print), scheduled to appear in the second half of 2021.

While it is clear that the digital humanities have a significant influence on the research practices in the humanities, it is by no means easy to explain what their substantial and lasting contribution to the theories and methods of the humanities may be.

The digital humanities are therefore in need of a theoretical foundation supporting adequate explanations of the changes in interpretation and analysis enabled or facilitated by the new software-based approaches. This is all the more important because these changes are already, or will soon be felt across the entire humanities. Multimodality research can help to answer these questions (Siefkes & Arielli 2018: 151–154). It provides models of transmediation and transcription that describe processes of digitisation as a change of semiotic properties that affects all three aspects of the sign, namely (in Peirce's terminology) the representamen, object, and interpretant. Given that digitisation processes and their influence on culture, its production, reception, and institutional "processing" in the widest sense, are at the heart of the digital humanities, multimodality research may be able to provide the missing theory of the digital humanities. (Siefkes in print, Ch. II.1.)

Mutual attraction: why the digital humanities need multimodality research – and vice versa

Both multimodality research and the digital humanities are new fields or areas of research that have emerged in the last decades. Today's research landscape is both increasingly fragmented with a multitude of semi-distinct research areas, and characterised by an increase of interdisciplinary research collaborations. Under these conditions, it is often advantageous to claim that two research fields or areas should meet. While the perspective and new ideas gained from merging two existing research fields will always be interesting, a certain scepticism is in order, and it should not be taken for granted that such a "merger" will be a lasting success. The present contribution will argue, however, that the digital humanities and multimodality research already have a significant area of overlap in their research practices. This area has aptly been called the "Multimodal Digital Humanities" (O'Halloran 2015). Furthermore, we will claim that both fields stand to gain significantly from a thorough theoretical reflection of the Multimodal Digital Humanities.

There are various aspects in which these two research traditions currently come together:

(1) The digital humanities increasingly focus on phenomena that belong to the subject matter of multimodality research, i.e. on multimodal texts, artefacts and communicative situations.
(2) From the other side, multimodality research increasingly relies on corpus-based methods and software solutions that enable (semi-)automatic annotation and evaluation, adapting tools and methodology from the digital humanities and data science in the process.
(3) Finally, there is an ongoing discussion in the digital humanities about the need for a better theoretical foundation, and multimodality research is a promising candidate.

How they met

Looking at the digital humanities from the perspective of multimodality research, it quickly becomes obvious that in their early stages, they were mostly concerned with written language. This focus on text-based approaches can be partly explained by the fact that computers are extremely good at handling large amounts of written text. For written texts, straightforward and effective character encoding standards (such as ASCII, Unicode, and UTF-8) have long been available, and modern programming languages can deal with the resulting "strings" (variables consisting of a sequence of characters encoded with one of these standards). Especially during the early decades when they were still called "humanities computing" (McCarty 2005) and connected with computational linguistics, the digital humanities were smitten with written text, simply because they knew how to deal with it. Beginning in the 1980s, however, technological developments fundamentally changed the ability of computers to effectively store, analyse, and manipulate images, video, graphics, and multimedia formats of various kinds (Manovich 2001, 2013).

Johanna Drucker has been an influential proponent of approaches in the digital humanities that transcend the textual orientation. As co-founder of the Speculative Computing Laboratory (SpecLab) at the University of Virginia, her work integrates visual modes of communication and design. She has conducted a number of research projects that investigate the role of aesthetics and visual design in the digital humanities (Drucker 2009). Drucker proposes an approach that foregrounds the relationship between the aesthetic dimension of digital media and their use for exploring and experimenting with traditional forms of knowledge and interpretation in the digital humanities.

Katherine Hayles distinguishes between various phases of development of the digital humanities, which are increasingly characterised by "multimodal practices":

> A decade later, the term [digital humanities] is morphing again as some scholars advocate a turn from a primary focus on text encoding, analysis, and searching to multimedia practices that explore the fusion of text-based humanities with film, sound, animation, graphics, and other multimodal practices across real, mixed, and virtual reality platforms. (Hayles 2012: 43)

In an article on the relationship between "Media Studies and the Digital Humanities", Tara McPherson distinguishes between three waves of the digital humanities characterised by different types of DH scholars. The newest incarnation seems to be the "multimodal scholar":

> This third type of digital humanist in effect blends many of the desires and goals of the other "early adopters," the computing humanist and blogging humanist. This emergent breed, the multimodal humanist, brings together databases, scholarly tools, networked writing, and peer-to-peer commentary while also leveraging the potential of visual and aural media that so dominate contemporary life. […] The multimodal scholar explores new forms of literacy that include authoring and analysing visual, aural, dynamic, and interactive media. She also takes her cues from popular culture, imagining what it would be like to immerse yourself in a scholarly argument as you might immerse yourself in a movie or a video game. She investigates what happens when scholarship looks and feels differently, requiring new modes of engagement from the reader/user. (McPherson 2008: 120)

William G. Thomas argues that the digital humanities are developing a new understanding of research that transcends traditional models of scholarship, which may no longer be suitable to the new digital forms of communication. This approach can sometimes lead to conflict with more traditional perspectives in the disciplines, for example regarding what counts as a publication:

> Many digital humanists take the position that digital environments demand multimodal, reciprocal, nonlinear modes of scholarship. Scholars in the disciplines perceive an inherent contradiction between that form of scholarship and criticism, review, and evaluation. (Thomas 2016: 534)

Notably, this conflict runs deeper than the traditional difficulties of interdisciplinary work in an academic landscape dominated by disciplinary boundaries. It concerns the question what counts as a scholarly work in the digital humanities. For example, is an interactive visualization of a museum collection a work of scholarship that deserves to be counted as an original publication?[2] What about the development of a network visualization tool such as Neo4j that allows the exploration of relationships in datasets? The digital humanities tend to answer in the affirmative, at least when the graphics or tool is designed creatively and provides innovative solutions to research questions.

In order to become multimodal, it is not enough for the digital humanities to shift their focus away from texts and books, and to include various other semiotic modes and media in their consideration. For this wider perspective to be successful, changes in theory and methodology are required as well.

> In the 21st century, we communicate in media significantly more varied, extensible, and multiplicative than linear text. From scalable databases to information visualizations, from video lectures to multiuser virtual platforms, serious content and rigorous argumentation take shape across multiple platforms and media. The best Digital Humanities pedagogy and research projects train students both in "reading" and "writing" these emergent rhetorics and in understanding how they reshape and remodel humanistic knowledge. This means developing critically informed literacies expansive enough to include graphic design, visual narrative, time-based media, and the development of interfaces. (Burdick et al. 2012: 10)

Clearly, the digital humanities are already multimodal in a number of ways: they investigate multimodal texts and artefacts, and they present the results of their research and analyses in multimodal ways (for example as interactive visualisations, or as dedicated websites that use different viewing modes). Admittedly, dedicated resources and support for the digital humanities (such as software, tools, research infrastructures and platforms) are often still insufficiently adapted to the needs of the digital humanities. To give an example, while there have been quick advances in computer vision, which promises important advances for the analysis of image databases and should become

2 A good example is the project Coins, an interactive visualisation of the numismatic collection of the Münzkabinett Berlin developed by Flavio Gortana, Franziska von Tenspolde, Daniela Guhlmann and Marian Dörk at the Urban Complexity Lab of the University of Applied Sciences Potsdam, 2017–2018. https://uclab.fh-potsdam.de/projects/coins/ [retrieved 17-04-2021]. The project is explained in Gortana et al. 2018.

increasingly important in art history and media studies in the next years, but there is insufficient software or infrastructure support, or specialised training, for digital humanists interested in deep learning and generally in data science methodology. Given the promises of these methods and tools the DH contexts, this will likely change in the near future.

Research infrastructures such as DARIAH-DE and many others increasingly try to accommodate the demands of art historians working primarily with images, archaeologists interested in virtual reconstructions, or curators thinking about the long-term preservation of born-digital art. While there is still some way to go until the digital humanities become fully multimodal, there is a rapidly developing international community fascinated by the new possibilities.

Two examples for the Multimodal Digital Humanities: data visualisations and digital editions

In the following, we will take a quick look at two examples for the role of multimodality in the digital humanities, visualisations (cf. Siefkes in print: Ch. III.3) and digital editions (cf. Siefkes in print: Ch. V.1). A first glance at conference programs, journals, or edited collections in the digital humanities confirms that data visualisation is an important topic in the field. In contemporary society, the importance of quickly and grasping complex relations and patterns in data, without necessarily understanding all the mathematical details of these relations and patterns, can hardly be overstated. This is true for professional contexts (e.g. presentations in meetings), for scientific research (papers increasingly include visualisations that are important for understanding the argumentation), and for science communication (visualisations can help to convey complex and abstract information in an easy manner to the general public).

It has therefore become increasingly important to understand what constitutes good information design. Some of the central principles were laid out by Edward Tufte, who stressed the danger of deception produced by misleading graphics (Tufte 1983). A large number of techniques for various purposes and contexts has been developed (Card 2003), and it is important to select carefully those that best fit the data, the research questions for which an answer is to be visualised, and the expertise level of the audience.

From the perspective of multimodality research, visualisations can be understood as transcription from one semiotic mode into another (Mondada

2018); this process is also sometimes called transmediation (Elleström 2014). In most cases, the source for visualisations is numeric data in the form of tables (Jannidis et al. 2017: 334), which are the result of some kind of measurement. However, in the digital humanities there are also other kinds of data that can be directly visualised, for example occurrences of expressions in texts, or collocations between expressions. From a technical perspective, this kind of data can be understood as string variables (corresponding to expressions of written language) and their relations to each other. Graph databases are well suited for storing, and for representing, string variables (e.g. names) as nodes, and their relation to each other (e.g. colloca-tions in a corpus) as edges, where stronger collocations can be visualised as shorter distance or as thicker lines depicting the edges between nodes.

Digital editions (Apollon et al. 2014; Pierazzo 2015; Sahle 2017) are another important area of the digital humanities where multimodality research can be profitably applied. Their presentation of facsimile and other materials alongside transcriptions, reading modes, and technical annotation and coding layers, as well as their often complex bespoke webdesign, makes them highly complex multimodal texts (Rosselli Del Turco 2012). Beyond the interactive design websites or "frontend" interfaces, digital editions often possess a quite complex backend architecture made up of specialised databases (such as Mungo DB) and specific tools for displaying annotated files in different ways, automatically producing collated texts from selected versions, and allowing for synoptic viewing modes.

Digital editions gain their central status for the digital humanities primarily because they are one of the most important and widespread applications. However, current digital edition projects have also been criticised for their tendency to develop tailored solutions from scratch. It seems that currently most digital editions seem to reinvent the wheel; common standards are missing or insufficiently applied, and the sustainability of newly developed technical solutions is not always taken into account (Andorfer et al. 2016). Multimodality research provides a somewhat different perspective, helping to understand the processes of development that are typical for newly introduced media and text genres, which often go through a stage of experimentation and individualistic solutions, before certain solutions are increasingly accepted. In comparison, we may think of the first decades of film or of the World Wide Web; in both cases, a wide range of technical solutions, design features and intended reception processes was tried, until some variants asserted themselves as successful solutions. Similarly, if we conceptualise digital editions as a

multimodal text genre in the early stages of its development (Bateman 2008), we can understand the current diversity and incompatibility of designs as a period of creative experimentation that may have a function in the present period of development of the multimodal genre "digital editions".

A love triangle? – Digital humanities and data science

A number of recent advances in the digital humanities, especially regarding deep learning (e.g. for automatic annotation), computer vision etc., are directly influenced by developments in data science. "Data science" is a fairly new term that has been coined for an area that encompasses advanced statistical learning and algorithmic modelling methods. The term is mostly applied to a new profession and educational profile somewhere between computer science, statistics, and data management consulting. Improved data modelling has increasingly become a hot topic especially for large and medium-sized businesses.

It should be noted that there is a significant overlap of methodology with adjacent fields, such as artificial intelligence and computer vision (which has extensive applications in robotics, self-driving cars, etc.). While "data science" may, to some degree, be regarded as a fashionable label for a range of technologies that have existed for some time, it should be acknowledged that its somewhat opportunistic mixture of mathematical statistics, computer science, deep learning and AI technology has a far-ranging field of applications. Data science is advancing rapidly wherever large amounts of data are available. This is increasingly the case in many areas of government, business, and research, where digitisation of existing use scenarios as well as advancing technologies and new workflows result in often vast amounts of data. In many cases, it is highly promising (and potentially highly problematic) to analyse this data, either with the aim of optimising processes according to pre-defined demands of the government agency, company, or research institution, or with the more general goal to look for patterns and previously hidden relations between variables.

What does this have to do with multimodality research? It is already noticeable that data science is exerting an influence on the digital humanities, especially in technically demanding areas where large datasets are processed with deep-learning algorithms. Some algorithms and areas of data science and artificial intelligence, such as computer vision, have important applications in the digital humanities, for example in the search and recognition of specific objects, scenes, and motifs in large image databases or on social media streams.

Furthermore, data science provides powerful solutions for problems that arise from the processing of large-scale multimodal data, and for the visualisation of patterns found in big data analyses (O'Halloran et al. 2018).

It is therefore likely that data science and artificial intelligence research will become an important aspect of the Multimodal Digital Humanities. They can help with technologically advanced study designs where large datasets of materials combining different semiotic modes (written text, images, graphics, video, or recorded verbal and gestural interactions) will have to be processed, categorised, annotated and analysed (semi-)automatically. Recently, machine-learning specialists have begun to develop systems for the image-guided translation of written or spoken texts (Caglayan et al. 2020, Bugliarello et al. in review). These can be regarded as first steps towards a multimodal analysis that extracts the information given in different semiotic modes and integrates it into a multimodal content model, where interpretations are reciprocally disambiguated and combined across modes.

Given the rapid growth of the amount and semiotic complexity of data posted on social media and internet platforms, solutions for these problems will be essential for the multimodal digital humanities. Data science and artificial intelligence research should have the methods and tools to achieve these solutions. However, they will have to be guided in interdisciplinary cooperations by digital humanities and multimodality researchers who know what is needed, and understand the complex mechanisms of multimodal meaning construction and intermodal relations (Siefkes 2015).

A marriage contract: the Multimodal Digital Humanities Framework

As outlined above, given the recent developments in the digital humanities, there are good reasons to use theoretical terms and methodological approaches of multimodality research for a reflection and theoretical underpinning of these developments. However, "marrying" the digital humanities to multimodality research is not as straightforward as it may look. In order to understand the difficulties, we can compare the digital humanities to other research fields where multimodality research has been successfully applied, such as film studies, comic studies, musicology, and many others. In all these cases, multimodality research was brought to bear on various existing research questions, or helped to gain new perspectives in disciplines that already possess a sound

theoretical and methodological basis. This can be modified and enhanced by a multimodal perspective.

The digital humanities, however, cut across established research fields. This makes them into more of a methodological toolbox than a research field in the traditional sense, which is defined by a specific subject matter such as films or comics. Furthermore, while they possess a rich arsenal of highly innovative methods, they have not yet achieved an adequate methodological underpinning of what these methods actually do, how they change our understanding of the subject matter(s) they are applied to, and how they influence the problem space of the respective research areas.

In terms of theory, the digital humanities are even less in safe waters. Floundering helplessly, they reach for anything that might keep them from drowning (and becoming "humanities computing" again, a theory-less support service for the various humanities disciplines; the term was intentionally discarded for strategic reasons, cf. Kirschenbaum 2012: 6). This makes it more difficult to adequately integrate a multimodal perspective into the digital humanities. While it is pretty straightforward to apply multimodality research to various disciplines (by looking for multimodal aspects of the texts, artefacts, or communicative situations they already investigate, and applying the theories, terminology, and analytical tools of multimodality research to them), this is more difficult with respect to the digital humanities, which are themselves a "bundle of methods" that have been applied to many different subjects and research questions across the humanities.

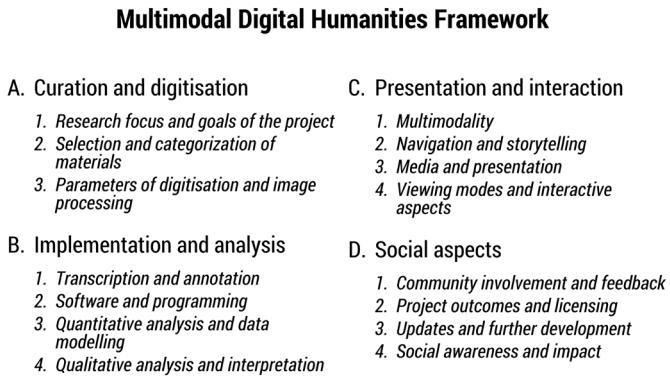

Fig. 1: The Multimodal Digital Humanities Framework (MDHF) proposes a number of dimensions for analysing multimodal artworks.

Given these challenges, it becomes apparent that it is important to apply multimodality research to the digital humanities in a thorough and general fashion, as both a theoretical foundation and an analytical toolbox. It is necessary to extend and adapt multimodal terminology to the specific range of projects, and the results they produce, that are characteristic for the digital humanities. In order to achieve this, the *Multimodal Digital Humanities Framework* (MDHF) has been developed. It is intended to analyse, describe, make comparable, and evaluate research in the digital humanities on the level of single projects, as well as whole project categories (such as digital editions, data repositories, or interactive visualisations, among others).

The digital humanities are characterised by projects that are interconnected on different levels, for example through the use of platforms, adherence to standards, or the re-use of corpora and digitised materials put under free licenses. Digital humanities projects often develop specific websites, digital platforms, tools and software, or new standards and norms that are closely related to the production and reception of multimodal texts or artefacts. In analysing these outcomes of DH projects, multimodality research really comes into its own. It provides a range of very helpful distinctions, classifications, and methods of analysis that have been thoroughly tested in extensive research on different multimodal text types in a broad range of communicative situations.

This research will be published in the upcoming monograph "Multimodal Digital Humanities" (Siefkes in print), which is scheduled to appear in the fall of 2021. It provides a thorough introduction to the intersection between multimodality research and the digital humanities, presenting an in-depth study of a number of innovative topics and research questions of the digital humanities which are associated with multimodality. On this basis, the categories and annotation schemata of the proposed framework are introduced, explained in detail, and applied to the analysis and evaluation of a corpus of 25 projects belonging to five different categories.

The Multimodal Digital Humanities framework, due to its focus on multimodal texts, is primarily oriented towards analysing the project websites and interactive services that are often the primary outcomes of digital humanities projects (digital editions, online exhibitions, interactive visualisations, online repositories etc.). However, it also includes some aspects of the project organisation and implementation as well, such as the research focus and goals, community involvement and long-term management, or social awareness and impact. This enables a deeper understanding which does not isolate the

resulting websites and other project results from their background and context of development.

The framework includes the four dimensions *Curation and digitization*, *Implementation and analysis*, *Presentation and interaction*, and *Social aspects* (cf. Fig. 1). It is designed in a modular fashion that allows for the inclusion or exclusion of different dimensions and specific annotated features, depending on the research questions of the study or analysis. It covers many features and aspects of the projects, such as research focus and goals, selection and categorization of materials, parameters of digitisation, transcription and annotation, quantitative and qualitative methods of analysis, viewing modes and interactive aspects, navigation and storytelling, community involvement, licensing, as well as social awareness and impact-related questions. For each of these categories, a number of variables with specific options of choice is proposed (cf. Fig. 2).

a. Entry page with the following access points
   1 a main menu
   2 a 'tile' pattern
   3 a list (or lists) of items
   4 a search menu
   5 an interactive map
   6 an animation (e.g. flash)
   7 introductory text(s)
   8 other
b. Number of levels (below entry page)
   1 one
   2 two
   3 three
   4 more than three
c. Main organisation principles of website
   1 topics or themes
   2 different viewing modes
   3 structure of the presented work
   4 different sources
   5 focus on search and user exploration
   6 space
   7 time
   8 people and relations between them
   9 contributing institutions or people
   10 other
d. Main menu
   1 at top of page
   2 at left side of page
   3 at right side of page
   4 elsewhere or movable
   5 with submenus
e. Main menu based on
   1 thematic organisation
   2 types of material
   3 project overview and administration
   4 narrative structure
   5 options for user exploration
   6 space
   7 time
f. Search functions
   1 search window visible on most pages
   2 search function on dedicated page
   3 full-text search
   4 search filters available
   5 detailed expert search
   6 search with logical connectors
g. General texts
   1 introduction to the main topic(s)
   2 detailed project description
h. Symbols and maps
   1 symbols, icons, or colour-coding for orientation on the website
   2 interactive maps
   3 site map

Fig. 2: The coding scheme for framework dimension C2: Navigation and Storytelling. © M. Siefkes, CC-BY SA 4.0.

The framework offers extensive options for the description, analysis, comparison, and evaluation of digital humanities projects and websites. It is suitable for a range of different application scenarios, such as the comparison of two or more projects, the detailed specification of a project during its planning stage, or the evaluation of a finished project against established standards.

Can multimodality research provide the missing theory of the digital humanities?

The Multimodal Digital Humanities Framework can also be understood as a contribution to the discussions about the missing theory of digital humanities. The theoretical foundations of DH are insufficiently developed in comparison to the impressive range of tools, applications, methods and methodological reflection the field has to offer. The framework, however, can only provide a first step towards an adequate theoretical foundation of the digital humanities. With this goal in mind, it combines terminologies and concepts taken from semiotics, linguistics, multimodality research, computer science, and sociology to achieve an adequate description of the diverse research practices and methodological innovations in the digital humanities and the well-developed methodological reflections in multimodality research into a common descriptive framework that allows to connect them with each other.

While it may not be sufficient, multimodality research can be an important part of a theory of the hybrid interpretation processes that form the core of the digital humanities, where computational analysis (often of large corpora or datasets) and human interpretation go hand in hand. From a semiotic perspective, both the computational steps of interpretation and pattern analysis, and the qualitative interpretation of the results which can only be conducted by humans, can be described as various types of sign processes that are combined in multilevel interpretation processes (Siefkes 2011: 22f.).

Importantly, the Multimodal Digital Humanities Framework enables the empirical comparative analysis of DH projects along a range of precisely defined feature dimensions (cf. the corpus analysis conducted in Siefkes in print, Ch. VII). Another application would be the definition of a requirement profile for a digital humanities project, which can be helpful during the development process, taking into account best practices regarding design, standards, and usability, as well as specific demands for the individual project.

Conclusion: these two were meant for each other

Clearly, the digital humanities are no longer "linguocentric" and primarily determined by the traditions of computer linguistics or digital philology. Digital humanities research has become increasingly multimodal, both regarding the investigated materials and research questions, and the methods, tools, and website designs used in the analysis and presentation of these materials. Image corpora are compiled, video and audio materials investigated, and complex media combinations (e.g. on Twitter or Instagram) are no longer outside its reach. While this move towards multimodal texts and artefacts has not yet led to a widespread adoption of the theories and methodologies of multimodality research, there is reason to believe that this is now changing.

Does the future then look bright for a long-term relationship between the digital humanities and multimodality research? Or will this be a fling without lasting consequences? As this contribution has argued, there is some reason to hope for the former. The digital humanities provide valuable computational resources and methodological innovations for multimodality research. They can help, for example, with the investigation of very large and unstructured amounts of raw data found in the wild (i.e. the internet), in the spirit of "big data", as opposed to smaller and to some degree pre-processed datasets, which is what effectively most multimodal corpora still are. The digital humanities are also somewhat permeable and open for further innovation from computer science. They are increasingly influenced by data science, a highly innovative field of research and professional practice situated somewhere between computer science, statistics, and artificial intelligence with widespread real-world applications. By opening itself further to the digital humanities and to data science, we argue, multimodality research, which has always had an empirical focus, can train its extensive theoretical prowess and analytical tools on various practical applications and professional contexts that are currently in high demand. This could be a very fruitful ménage à trois!

By integrating approaches and methods from the digital humanities, multimodality research is forced to reflect on its own theoretical basis and terminological consistency. As everyone knows who has attempted to formalise a theory, or to implement a computational version of an analytical method, such an enterprise mercilessly exposes unclear terminology, inconsistencies, vague hypotheses, or overblown claims. Multimodality research can have a certain tendency towards all of these, especially towards vagueness, which is sometimes optimistically touted as "openness". This ignores the fact that

formulating a theory or approach in precise terms makes it more amenable to criticism, and responsible for its claims. While there is no doubt that not everything can (or should) be formalised, multimodality research can certainly profit from the sometimes rigid demands of the digital humanities. Challenged by new perspectives and expectations, it should be willing to rethink and reformulate its terminology, claims, and analytical methods, and significantly gain both regarding theoretical validity and practical uses.

The digital humanities, in their turn, will profit enormously from the careful theoretical foundations for the investigation of complex mediatised texts and artefacts that have been laid in multimodality research (building on previous decades of theoretical and empirical work in semiotics). The dearth of theory that the digital humanities are afflicted with has been pointed out by many practitioners of the field. Multimodality research can help with that.

In doing so, the DH may come to understand better what they have been doing all along: by digitising and modelling culture on different levels and with different methods, they are changing some, but not all, of the semiotic properties and modes of production and perception of the multimodal texts and artefacts they are concerned with. An arsenal of powerful new methods for categorising, analysing, and interpreting these texts and artefacts has been developed in the digital humanities. These innovations and their consequences can be better understood with the intense theoretical reflection, fine-grained terminology, and wide-ranging analytical frameworks provided by multimodality research. It seems that a successful relationship is in the stars for the digital humanities and multimodality research!

Acknowledgment: This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – SFB 1410.

Bibliography

Andorfer, P., Durco, M., Stäcker, T., Thomas, C., Hildenbrandt, V., Stigler, H., Söring, S., & Rosenthaler, L. (2016). Nachhaltigkeit technischer Lösungen für digitale Editionen. Eine kritische Evaluation bestehender Frameworks und Workflows von und für Praktiker_innen. *Online proceedings of DHd 2016*, http://dhd2016.de/abstracts/panels-006.html [retrieved 18-04-2021].

Apollon, D., Bélisle, C., & Régnier, P. (Eds.). (2014). *Digital Critical Editions*. University of Illinois Press.

Bateman, J. (2008). *Multimodality and Genre. A Foundation for the Systematic Analysis of Multimodal Documents*. Palgrave Macmillan.

Bateman, J. (2017). Multimodale Semiotik und die theoretischen Grundlagen der Digital Humanities. *Zeitschrift für Semiotik*, *39*(1–2), 11–50.

Bateman, J., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, Research and Analysis – A Problem-Oriented Introduction*. De Gruyter.

Björkvall, A. (2012). Multimodality. In J.-O. Östman & J. Verschueren (Eds.), *Handbook of Pragmatics* (pp. 1–19). John Benjamins.

Bugliarello, E., Cotterell, R., Okazaki, N., Elliott, D. (in review). Multimodal Pretraining Unmasked: A Meta-Analysis and a Unified Framework of Vision-and-Language BERTs. To appear in *Transactions of the Association for Computational Linguistics*, 21. Preprint published in *arXiv:2011.15124* [last revised 30.05.2021].

Caglayan, O., Ive, J., Haralampieva, V., Madhyastha, P., Barrault, L., Specia, L. (2020). Simultaneous machine translation with visual context. *Transactions of the Association for Computational Linguistics*, 8, 539–555.

Card, S. (2003). Information Visualization. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook* (pp. 544–582). Erlbaum.

Elleström, Lars (2014). *Media Transformation: The Transfer of Media Characteristics among Media.* Palgrave Macmillan.

Flueckiger, B. (2017). A Digital Humanities Approach to Film Colors. *The Moving Image: The Journal of the Association of Moving Image Archivists*, *17*(2), 71–94.

Fricke, E. (2012). *Grammatik multimodal. Wie Wörter und Gesten zusammenwirken.* De Gruyter.

Fricke, E. (2013). Towards a Unified Grammar of Gesture and Speech: A Multimodal Approach. In C. Müller, A. Cienki, E. Fricke, & D. McNeill (Eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Vol. 1 (pp. 733–754). De Gruyter.

Fricke, E. (in this volume). Multimodality and Semiotic Complexity. In Arne Krause and Ulrich Schmitz (Eds.), Linguistics and Multimodality. *OBST 99*, 57–86.

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

Gortana, F., von Tenspolde, F., Guhlmann, D., & Dörk, M. (2018). Off the Grid: Visualizing a Numismatic Collection as Dynamic Piles and Streams. *Open Library of Humanities*, *4*(2).

Hayles, N. K. (2012). How We Think: Transforming Power and Digital Technologies. In D. M. Berry (Ed.), *Understanding Digital Humanities* (pp. 42–66). Palgrave Macmillan.

Jewitt, Carey (Ed.) (2014). *The Routledge Handbook of Multimodal Analysis*. 2nd ed. Routledge.

Kirschenbaum, M. (2012). What Is Digital Humanities and What's It Doing in English Departments? In M. K. Gold (Ed.), *Debates in the Digital Humanities* (pp. 55–61). University of Minnesota Press.

Manovich, L. (2001). *The Language of New Media*. MIT Press.

Manovich, L. (2013). *Software Takes Command*. Bloomsbury.

McCarty, W. (2005). *Humanities Computing*. Palgrave Macmillan.

McPherson, T. (2008). Introduction: Media Studies and the Digital Humanities. *Cinema Journal*, *48*(2), 119–123.

Mondada, L. (2018). Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction*, 51(1), 85–106.

O'Halloran, K. (2015). Multimodal Digital Humanities. In P. Trifonas (Ed.), *International Handbook of Semiotics* (pp. 389–415). Dordrecht: Springer.

O'Halloran, K. L., Tan, S., Pham, D.-S., Bateman, J., & Moere, A. V. (2018). A Digital Mixed Methods Research Design: Integrating Multimodal Analysis With Data Mining and Information Visualization for Big Data Analytics. *Journal of Mixed Methods Research*, 12(1), 11–30.

Pierazzo, E. (2015). *Digital Scholarly Editing: Theories, Models and Methods*. Ashgate.

Rosselli Del Turco, R. (2012). After the Editing Is Done: Designing a Graphic User Interface for Digital Editions. *Digital Medievalist*, *7*(0).

Sahle, P. (2017). Digitale Editionen. In F. Jannidis, H. Kohle, & M. Rehbein (Eds.), *Digital Humanities. Eine Einführung* (pp. 234–249). Metzler.

Siefkes, M. (2011). Style: A New Semiotic View on an Old Problem. *Kodikas/Code. Ars Semeiotica*, *34*(1–2), 15–25.

Siefkes, M. (2015). How Semiotic Modes Work Together in Multimodal Texts: Defining and Representing Intermodal Relations. *10plus1 – Living Linguistics*, *1*, 113–131.

Siefkes, M., & Arielli, E. (2018). *The Aesthetics and Multimodality of Style*. Peter Lang.

Siefkes, M. (in print), *Multimodal Digital Humanities. Grounding digital research methods in multimodal linguistics and semiotics*. Palgrave Macmillan.

Svensson, P. (2009). Humanities Computing as Digital Humanities. *Digital Humanities Quarterly*, *003*(3).

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Graphics Press.