Meandering Interpretations

Augmented Text Interpretation through Interactive Visualizations

Master Thesis

by

Philipp Proff

Matr. No. 4044954

Supervisors:

Prof. Albert Lang

Prof. Dr. Marc Pfaff

Prof. Dr. Marian Dörk

M.A. Design & Computation

University of the Arts Berlin

Technical University Berlin

Berlin, April 2025

DECLARATION OF AUTHORSHIP

I hereby declare that I have written this master thesis by myself, without the help of third parties and exclusively using the sources and aids listed. I have marked as such all passages that are taken from the sources and aids used, either unchanged or in terms of content. I further declare that I have not yet submitted the work in the same or a similar form to any other examination authority.

_____ _____

Place, Date Signature

ABSTRACT

The interpretation of texts is a central practice in most academic disciplines. In contrast to leisure reading, academic reading is characterized by distinct methods and practices for text analysis, e.g., close reading. In this regard, this work has explored the design space for utilizing interactive visualizations for synergizing algorithmic text analysis and human interpretation in a digital environment. A "research through design" methodology was employed to produce tangible artifacts during the process, which acted as mediators for discussing the design decisions and their ability to convey interpretative agency. A workshop was conducted to explore the usage of graphical elements and variables during an exemplary text interpretation process. Outcomes suggested the employment of lines and arrows and variables such as position and size to support the interpretation process visually. As a result of these outcomes and the insights from the prior literature research, five key aspects were formulated that should be considered when designing a digital environment that visually augments the interpretation process. During the design process, prototypical implementations provided features for interactive deconstruction of the source text, visual alterations of the created elements as well as interactive annotations. Network visualizations and a timeline further augmented these features and emphasized the investigation of non-linear structures composed out of fragments of the old text. Lastly, a reflection upon these prototypes concluded that a digital text interpretation environment must facilitate a form of hacking that allows to actively alter the original text structure, create novel non-linear constellations out of its fragments and juxtapose them for comparative analysis. Furthermore, algorithmically created visualizations must be visually adjustable to reflect human interpretation and facilitate a dialogical mode.

Keywords: Human-Computer Interaction, Information Visualization, Text Analysis, Natural Language Processing, Hermeneutics

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

In the late 90s, science and technologies philosopher Bruno Latour (1986) observed that almost every scientific endeavor, from calculations, observations, and built prototypes are distilled into narrated forms of text. Today almost all disciplines still use text as the primary scientific exchange format (Gascoigne et al., 2020), although, as text is natural language, it is ambiguous. However, text is powerful as it enables to extend the results by adding various layers of context and therefore steer the outcomes to a certain narrative. Considering this ambiguous quality of texts, a procedure is needed to find meaning in it and use it for further work. Therefore, analysis methods are an important aspect of the interpretation process, which try to establish formalized workflows to create more consistent outcomes. However, French philosopher Etienne Bonnot de Condillac emphasizes the creative aspects of the process:

> [Analysis] merely consists in composing and decomposing our ideas to create new combinations and to discover, by this means, their mutual relations and the new ideas they can produce. (Condillac, 2001, p. 48)

This description seems reasonable but neglects the differences in text mediums, which are crucial for this creative process. Although paper is historically the dominating medium, its physical qualities limit the set of operations that can be done with it, especially the dynamic display of complex, non-linear, text structures. In this regard, the digital medium bears the advantage to be much more flexible.

Building on these possibilities, Ted Nelson (1965), a sociologist and early internet pioneer, proposed a novel way of digital text structure which described a distributed non-hierarchical network of text fragments which couldn't be efficiently represented on paper – the hypertext. In this light, the text interpretation process is fundamentally altered, as these novel distributed non-linear texts challenge the hermeneutical premises which build upon the interpretation of a coherent linear structured texts.

Nevertheless, most digital environments for text interpretation provide a poor repertoire of elements and interactions for investigating and making use of this non-linearity. However, interactive visualizations could offer an intuitive interface to efficiently reconfigure texts (Drucker, 2020). The combination of human and machine analysis offers a great opportunity for a process, which dissolves the text into a non-linear network of interpretative elements. In this light, Elena Esposito (2022) proposes that "the machine operates as a partner making proposals that can direct interpretation in unexplored directions" (p. 14).

Based on these thoughts on text interpretation, this work explores the design space of functionalities and interactive visualizations which support the text interpretation process in a digital environment. The following research questions have guided the work:

(RQ1) How does a digital environment look like that visually augments the text interpretation process?

(RQ2) How can algorithmic text analysis and human interpretation be visually connected in a meaningful way?

To answer these questions, the work is structured into three parts. In the first part the historic developments of the hermeneutics are shortly introduced, as well as three different approaches to text analysis and different methods for visualizing text. Thereafter the second part will give a short insight into the motivation and methodology of this work. Consecutively, the findings of the design research process are presented structured into the outcomes of a workshop and a documentation of the ideation on prototyping process. The second part is wrapped up by a reflection upon the design process discussing decisions and insights. Lastly, a conclusion is drawn from the findings and possible answers to the research questions are proposed.

2. Related Theories and Work

2.1. Hermeneutics

The practice of text interpretation has a rich history of influential philosophies that shaped the process of text exegesis. Hermeneutics, from the Greek word "hermeneu's" – mediation, is known as the philosophy of interpretation and understanding and has a long history with an ever-changing philosophical understanding of the process of interpretation (Figal, 2007). Until the times of the enlightment, western hermeneutics was mainly concerned with religious and judicial texts and how to correctly interpret them. During antiquity and until the 15th century it was established that texts can have a plurality of interpretations which fostered diverse discourse around meaning in texts.

This changed in the 16th century as Martin Luther proposed a return to a literal understanding of the biblical texts which according to him "interprets itself" and does not need human intervention. Due to this understanding, hermeneutics became part of logic and the textual rules, such as rhetoric, grammar and philology, provided the framework to uncover the only true meaning as already inscribed into. As the individual became the center of experience, during the times of the enlightment, the theologist Friedrich Schleiermacher (1838) was the first to investigate texts upon their psychological dimension, studying authors intentions in combination with the linguistic analysis.

Schleiermacher also proposed to open the interpretation practice to all sorts of texts which was previously still restricted to religious and judicial texts. As the process of interpretation was bound to the individual, Schleiermacher understood the process of understanding as iterative, where prior knowledge would inform the next iteration of interpretation, also known as the hermeneutical circle. These thoughts have been extended later on by scholars like Wilhelm Dilthey (1900) and Hans-Georg Gadamer (1975), who added societal and historic context of author and recipient as influential factors.

In conclusion, the theory and practice of interpretation has changed substantially over the last centuries. From an understanding of texts as isolated objects to an understanding of texts as constructed by the psychological, societal and historical contexts of author and recipient. In this light, the hermeneutics can be described as antinaturalistic, which means that it does not adapt to the approach of the empiricists and positivists (Thaller, 2022). But here the question emerges, why this is the case? Interpretation needs prior analysis to foster informed critique and argument, and the developed text analysis techniques have changed greatly over time. This has become especially a topic of conflict in the recent decades where computers and algorithms, built upon the foundations of mathematics and physics, penetrate every scholarly practice. But the humanities, as the one of the main disciplines concerned with hermeneutics, have not crumbled under the developments of digitization and thus, novel strategies and theories have emerged, which foster a more holistic combination of analysis methods where interpretation is at the core.

Taking this into account, this work introduces three forms of text analysis which build upon the proposed division of close-, hyper- and machine reading from Katherine Hayles (2010). They provide a robust base for a wide range of "manual" and algorithmic text analysis techniques, which represent a contemporary toolset for text interpretation.

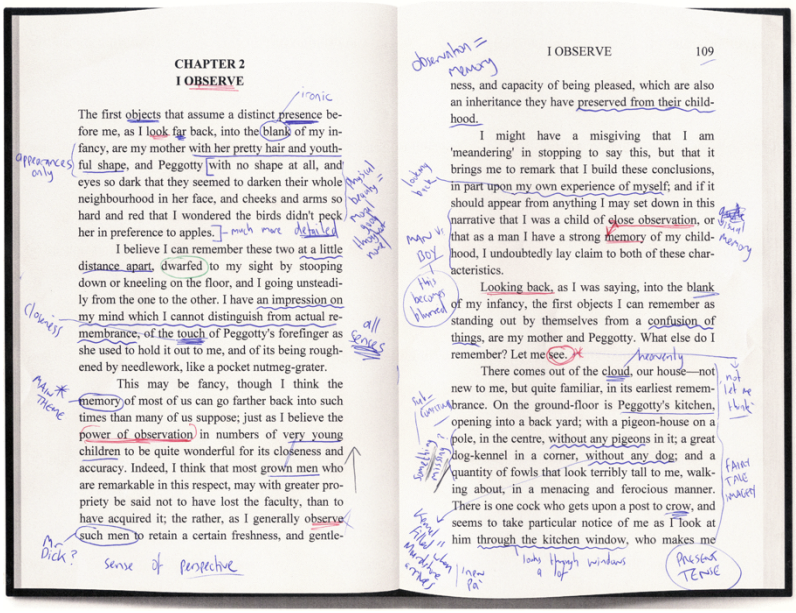
2.2. Different Forms of Text Analysis

Before the different approaches to text analysis are introduced, it needs to be clarified that the act of interpretation is closely connected to analysis. Structured investigations of text provide abstract patterns which must be contextualized and interpreted to make any meaning.

2.2.1. Close Reading

Nancy Boyles (2013) describes close reading as a "reading to uncover layers of meaning that lead to deep comprehension" (p.1). One outstanding work for the introduction of close reading is the book "How to Read a Book" by Mortimer J. Adler and his editor Charles van Doren (1972).

Although the book is quite old, two chapters are still relevant to the contemporary close reading practice: "inspectional reading" and "analytical reading". Inspectional reading has the goal of gaining an informed overview over the structure of the work and its ideas, called "skimming". The text is read at once without stopping for selective analysis of passages, to grasp the overall argument structure. This equips the reader with an overview that provides sufficient knowledge about the text to conduct an in-depth analytical reading of the text at a later stage. Analytical reading, in contrast, is the informed capability to virtually dialogue with the authors and their arguments by critiquing their work. To achieve this, analytical reading employs the "scanning" technique which looks for keywords and sentences with heightened importance. James Jasinski (2001) has distilled the method of analytical reading into four main aspects that are fruitful to be analyzed: individuals and events, words and phrases, structure and style and argument patterns. The knowledge about these elements enables the reader to find different angles from where the text and its statements can be challenged. Although this type of reading is much appreciated for its critical and suspicious approach to texts, Best and Marcus (2009) suggest to return to a more concrete form of close reading which they call "surface reading". They argue that the text itself provides interesting interpretations without employing cryptic chains of inference.



CHAPTER 2
I OBSERVE

The first objects that assume a distinct presence before me, as I look far back, into the blank of my infancy, are my mother with her pretty hair and youthful shape, and Peggotty with no shape at all, and eyes so dark that they seemed to darken their whole neighbourhood in her face, and cheeks and arms so hard and red that I wondered the birds didn't peck her in preference to apples.

I believe I can remember these two at a little distance apart, dwarfed to my sight by stooping down or kneeling on the floor, and I going unsteadily from the one to the other. I have an impression on my mind which I cannot distinguish from actual remembrance, of the touch of Peggotty's forefinger as she used to hold it out to me, and of its being roughened by needlework, like a pocket nutmeg-grater.

This may be fancy, though I think the memory of most of us can go farther back into such times than many of us suppose; just as I believe the power of observation in numbers of very young children to be quite wonderful for its closeness and accuracy. Indeed, I think that most grown men who are remarkable in this respect, may with greater propriety be said not to have lost the faculty, than to have acquired it; the rather, as I generally observe such men to retain a certain freshness, and gentle-

ness, and capacity of being pleased, which are also an inheritance they have preserved from their childhood.

I might have a misgiving that I am 'meandering' in stopping to say this, but that it brings me to remark that I build these conclusions, in part upon my own experience of myself; and if it should appear from anything I may set down in this narrative that I was a child of close observation, or that as a man I have a strong memory of my childhood, I undoubtedly lay claim to both of these characteristics.

Looking back, as I was saying, into the blank of my infancy, the first objects I can remember as standing out by themselves from a confusion of things, are my mother and Peggotty. What else do I remember? Let me see.

There comes out of the cloud, our house—not new to me, but quite familiar, in its earliest remembrance. On the ground-floor is Peggotty's kitchen, opening into a back yard; with a pigeon-house on a pole, in the centre, without any pigeons in it; a great dog-kennel in a corner, without any dog; and a quantity of fowls that look terribly tall to me, walking about, in a menacing and ferocious manner. There is one cock who gets upon a post to crow, and seems to take particular notice of me as I look at him through the kitchen window, who makes me

Figure 1: Annotation of a paper-based text.[1]

[1] Reproduced with permission from Kehoe and Gee (2013).

Most academic readers go multiple times over the same text to investigate different aspects and use the gained knowledge for more informed interpretation (Guillory, 2008). As this process is iterative but also incremental, the hermeneutical circle (or rather spiral) is often used as a metaphor for the process (Alonso Schökel, 1998). A much-appreciated technique, to facilitate an active reading process, encompasses the creation of visual markings or annotations (see Figure 1). These active engagements with the text are often used in a conversational mode but also to support cognitive processes (Tversky, 2011). Kalir and Garcia (2021) observe that annotations "can be a way to create, document, and curate new knowledge" and exhibit "a form of self-expression" (p.7). Thus, annotations help to explicitly describe context, capture thoughts and prompt questions for the purpose of an active engagement with texts which fosters factual knowledge and enables inference.

2.2.2. Hyper Reading

In the 1960s Ted Nelson, a sociologist interested in network technologies, introduced the concept of the "hypertext" which he described as follows:

> Let me introduce the word "hypertext" to mean a body of written or pictorial material interconnected in such a complex way that it could not conveniently be presented or represented on paper. It may contain summaries, or maps of its contents and their interrelations; it may contain annotations, additions and footnotes from scholars who have examined it. (Nelson, 1965, p. 96)

This novel "interconnected" quality of the hypertext and differences between a paper-based and digital text medium had lasting effects on the processes connected to text interpretation. Changes in reading strategies were observed, as well as emergence of novel methods and techniques for text analysis.

On the brink of the new millennial, literary scholar James Sosnoski (1999) observed several behavioral changes in reading amongst scholars and students which were connected to hypertext-based reading.

He loosely coined this novel reading mode "hyper-reading", which is defined by eight characteristic techniques: filtering, skimming, pecking, imposing, filming, trespassing, de-authorizing and fragmenting. Most of these techniques describe reading habits that do not investigate texts as a singular connected piece but rather as source for smaller textual fragments. This is especially provoked by the goal of extracting a certain information, which does not require an understanding of the whole text. Additionally, the increased complexity, facilitated by interactivity and hyperlinks, takes a toll on cognitive load, which makes it harder to focus on one linear text at once (Hayles, 2010).

In this regard, John Guillory (2008) emphasizes two techniques from close reading which have shifted in their purpose: "scanning" and "skimming". While skimming was utilized prior to grasp the core ideas of a text for the sake of informing a new iteration of reading, digital skimming is employed to quickly assert if a piece is worth to read at all. Scanning is used to quickly find relevant passages in a text which can be read and extracted without reading the whole piece. Both techniques provide filtering mechanisms that deal with the exponential growth of eligible texts, which is empirically connected to a possible cognitive overload (DeStefano & LeFevre, 2007). Given these aspects, it is no surprise that the reduced attention to long digital texts comes with behavioral changes. The famous eye-tracking study of Jakob Nielsen (2006) provides such an example, where participants exhibited a “F-shaped” reading pattern on visiting a website, that indicates the reduction of cognitive load by employing the prior mentioned techniques. To support both techniques, filtering algorithms are used to enable keyword-based interactive searches.

Furthermore, Katherine Hayles (2010) emphasizes the opportunity to create virtual links between documents and text segments, to employ find novel approaches to interpretation. The analysis of network structures has been a key element in this regard. Additionally, the interconnected documents foster explorative forms of reading, where the source text points to additional resources and references which can be quickly accessed through the internet.

2.2.3. Machine Reading

The algorithmic processing of text can be described as machine reading or natural language processing (NLP). The field of NLP draws on different theories and methodologies from other fields such as physics and mathematics but also more distinct ones like information theory and linguistics (Eisenstein, 2019). An increasing amount of available computational power has pushed this discipline further in the last decades. Emergent patterns become visible through the computational analysis of large text corpora and shed new light on linguistic and cultural phenomena.
Exemplary is the work of Michel et al. (2011), which analyzed millions of books from the Google Books corpora, and then claimed to have founded a new discipline: "culturomics". This new field was equipped with the possibilities to "[…] provide insights about fields as diverse as lexicography, the evolution of grammar, collective memory, the adoption of technology, the pursuit of fame, censorship, and historical epidemiology" (Michel et al., 2011, p. 1).

NLP provides a wide range of text analysis methods. Here we can differentiate between algorithms that assume a context-insensitive and a context-sensitive grammar (Chomsky, 1956). While the first is developed for efficient pattern recognition in character sequences, the latter can calculate semantic relations between text fragments. To find a text by a pattern, one must be able to describe such pattern by a formal description of elements in the text which are previously unknown (Manning & Schütze, 1999). "Regular expressions" are the go-to method for any computer program that deals with search functions. To extract each word in the exemplary sentence "Mary eats ice" one can formally describe the sentence as the expression "\w*" where "\w" indicates to match any word and "*" to match zero or more. This can be applied to every type of string to extract all words in that text.

Another relevant algorithm is the construction of "ngrams". Manning and Schütze (1999) describe ngrams as an entropy phenomena of language, that produces reoccurring combinations of words which can be analyzed by probabilistic methods.

A famous project in this regard is the 1T 5-Gram Corpus by Google which analyzed 1 trillion English words from texts and split them into frequency tables of bigrams, trigrams, 4-grams and 5-grams (Aston & Burnard, 2022).The database provided an interface which then showed the co-occurrences of words over time and exhibited high throughput rates which makes it still relevant today.

If the text must be analyzed in a context-sensitive way, supervised and unsupervised methodologies can be applied during the pre-processing of a text. Especially tagging and labeling are relevant methodologies that contextualize text elements for further exploration of semantics. Tagging a text corpus can have various motivations.
Manning and Schütze (1999) suggest that a tag system should exhibit "predictive features [and] encoding features that will be useful in predicting the behavior of other words in the context" (p. 144). One frequently used algorithm, developed by Eric Brill (1992), analyzes the text structure and tries to find out the words respective part of speech (POS) tag, such as noun, verb or adjective. POS tags are often used to reduce the number of words of interest and remove stop words, such as "and" or "furthermore" that enable fluent reading but hold minimal meaning. A widely used labeled lexical database for the English language is the "WordNet" database which provides POS tags for a huge number of words and additionally provides information about synonyms and lemmatization.

As the semantics of words are hard to pinpoint due to their various occurrences and context sensitivity, the manual annotation and definition of huge corpuses of text and their relations posed a significant problem to NLP researchers. The work of Mikolov et al. (2013) posed a ground-breaking new unsupervised approach of feeding large text corpora into recurrent neural networks to represent words as vectors in multi-dimensional spaces – the so-called word2vec architecture. Due to the possibility to calculate distances between high-dimensional vectors, semantically closely connected words were found to exhibit shorter distances than words that were rather unrelated. The word2vec model was a milestone in deriving semantics between words from unsupervised learning methods, even though the model often lacked context due to its architecture. Additionally, contextual sensitive words were less well represented in these vector spaces.

This problem was tackled by a research team at Facebook who created the "fastText" architecture that was more aware of contextual nuances of meaning (Bojanowski et al., 2016). The fastText architecture posed significant better processing speed and was able to handle the look up of semantics of words which were not in the corpora. Although these models pose interesting use-cases for text analysis purposes, significant critique has been stated due to the inscribed biases. Bolukbasi et al. (2016) have shown that these models exhibit e.g., sexist and racist relations between occupations and used descriptive words for people. Therefore, these models should always be employed with a certain amount of skepticism towards the drawn relationships.

As stated above, there are many forms, manual or algorithmic, of text analysis which can be utilized during the text interpretation process. These forms privilege certain aspects in a text which can be highlighted, extracted and modeled. Human vision provides excellent pattern recognition abilities which is leveraged by interpreting visual forms of text analysis (Mattson, 2014).

2.3. Text Visualization

As text becomes data in the digital age, visualization algorithms bear the possibility to reconfigure the visual appearance of the text. This novel configuration provides a new contextualization of text and therefore offers new information. Risch et al. (2008) define text visualizations as a subclass of information visualization. One of the most cited works on information visualization, describe the purpose of it as "the use of computer-supported, interactive, visual representations of abstract data in order to amplify cognition" (Card et al., 2007, p. 7).

Usually these visualizations follow the "Information Seeking Mantra" of Shneiderman (1996): "Overview first, zoom and filter, then details-on-demand" (p.337), which itself follows a "prevent distortion of data" rule established by Edward Thufte (1998). Text visualizations are historically closely associated with so-called semantic clustering and semantic mapping which are used to "graphically depict the overall conceptual structure of a document collection" (Risch et al., 2008, p. 164).

Ted Nelson (1965) already envisioned such networks of interconnected text samples in his paper on the Hypertext and provided multiple prototypes and theoretical frameworks with his "Xanadu" project. The "xanadoc" filetype was a frontier idea for an interactive visual text analysis tool aimed at investigating cross-references between text fragments (see Figure 2). Many text visualization techniques have been developed to visualize the results of NLP algorithms (Zanini & Dhawan, 2015). However, as most NLP methods are quantitative, visualizations that depict the outcomes of algorithmic text analysis often use techniques from the natural sciences. In this regard Baumer et al. (2022) point out that text visualizations often obfuscate their ambiguous background and display a certain declarative statement due to their symbolic heritage from the natural sciences.

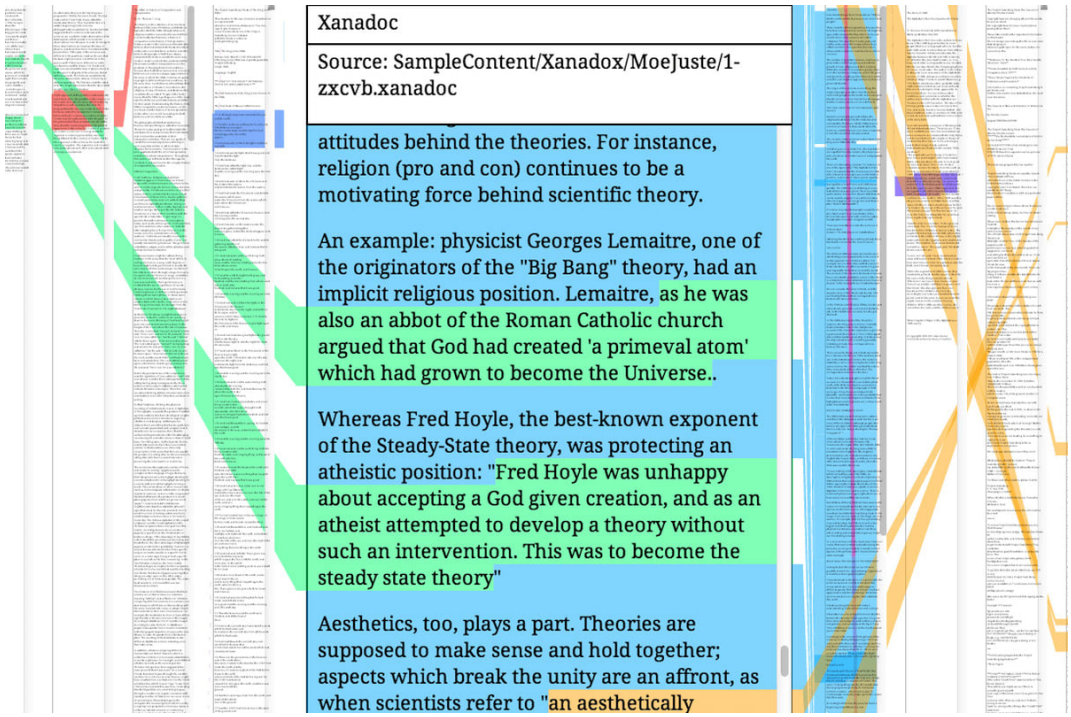


Figure 2: Xanadoc prototype displaying links between synthesized documents.[2]

While in the past, information visualizations were influenced by an empiricist view, inherited from the natural sciences, modern discourses debate this neutrality of data, especially in the case of ambiguous artefacts like text.

[2] Screenshot from https://xanadu.com/xanademos/MoeJusteOrigins.html (12.02.2025).

Johanna Drucker (2014) is a dedicated scholar in this regard: “Visualizations are always interpretations — data does not have an inherent visual form that merely gives rise to a graphic expression” (p. 7). This observation prompts several questions about objectivity and the limitations of visualizations as communication devices.

Despite the ambivalent feelings towards text visualizations in the humanities, Jänicke et al. (2015) have shown that text visualizations are effective in the augmentation of text analysis. A classic text visualization is the word cloud, which maps the word count to the size of the word. The DocuBurst visualization leverages the hyponymy, the relation to a more generic word, information of the WordNet lexical database by displaying hierarchical relations through a radial arrangement towards the most generic word in the center (Collins et al., 2009). Timeline visualizations can reveal patterns in co-appearance of characters (Liu et al., 2013), reconstructing the temporal order of a story where the narrative order is non-linear (Kim et al., 2018) and order the appearance of phenomena along a time axis (Schwan et al., 2019). Data that exhibits non-hierarchical relations between entities is often displayed through graph visualization techniques. Graphs can be used to show semantic networks as seen in Figure 3 or relationships between annotations and markings in texts (Bludau et al., 2020). In order to give an example that incorporates multiple text visualizations for analysis, the multi-modal "Voyant Tools" by Geoffrey Rockwell and Stéfan Sinclair (2017) provides an interface where different visualizations are juxtaposed for the same source text (see Figure 4).

Although all these text visualizations have potential use-cases, an increasing number of them are driven by the visual analysis of only the outcomes and therefore provide little transparency and room for non-directed exploration and personal inflection. Walny et al. (2018) have shown that an active engagement with visualizations foster deeper comprehension and inferential capabilities, which is needed in the context of text interpretation. Therefore, Dörk and Knight (2015) argue for a more playful, open-ended, approach to text visualizations. This is supported by Armaselu & Heuvel (2017) who stress the power of “reconfiguration, reorganization, restructuring” (p.1) for interpretation activities.

These statements emphasize the discursive and ambiguous qualities of text visualizations, which can be fruitful for interpretation.

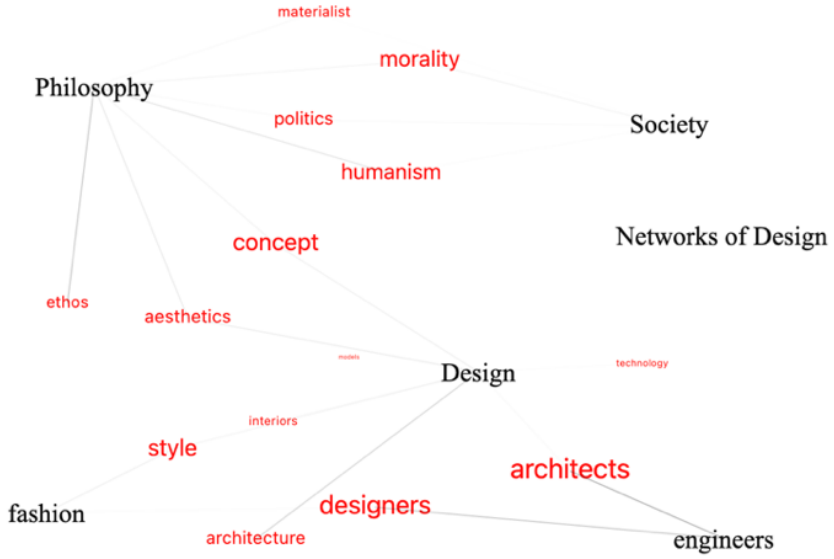


Figure 3: Semantic network between several terms according to a word2vec model.

Figure 4: Voyant Tools interface juxtaposing various text visualizations.[3]

[3] Screenshot from https://voyant-tools.org/ (12.02.2025)

3. DESIGNING AFFORDANCES FOR VISUAL INTERPRETATION

The rise of digital computers and the corresponding emergence of digital text formats has fundamentally altered the way how we work with texts (Hayles, 2012). Cognitive processes have adapted to the exponentially growing amount of information, which is observed in F-shaped reading, and increased skimming and scanning (Nielsen, 2006). Although the behaviors associated with this digital "overload" are often connected to negative effects, scholars like Johanna Drucker (2020) have argued that the digital medium could facilitate new and creative approaches to text analysis and interpretation.

English literature scholar Katherine Hayles (2010) emphasizes the opportunities of digital environments as they enable the efficient juxtaposition of multiple texts and creative analysis methods that utilize visualizations. These methods increase complexity by transforming the former linear into non-linear networked structures which over novel angles on the texts. Lev Manovich extends on this idea, by referencing Howard Rheingold's (2000) "tools for thought", and argues that "computers and software are not just "technology" but rather the new medium in which we can think and imagine differently" (p. 13). This novel perspective sets the focus on the hermeneutical implications of the digital medium, which is characterized by the emancipation of texts from linearity. As digital displays and data-driven interactions fundamentally alter the way we interact with texts, Marshall McLuhan's (1966) famous proclamation "the medium is the message" provokes new questions about the interpretation process of digital texts.

Psychology scholar Barbara Tversky (2011) argues that visual aids have an important role for cognition and thought, as visual language has many analogous characteristics to natural language. Therefore, visualizations can be a powerful tool for schematization and interpretation of meaning, which is illustrated by the annotations and diagrammatic depictions which are often found in the margins of texts (Kalir & Garcia, 2021).

In contrast to speech, text is mostly displayed on a two-dimensional space – the page. Barabara Tversky (2011) emphasizes the advantage of this two-dimensional plane as it enables the utilization of several cognitive processes for interpretation, such as proximity or orientation of elements. Readers often mark and annotate texts to intuitively structure text fragments and thought. Graphic variables can convey meaning in this regard and can help to think visually. Common variables for encoding information into graphic elements are position, color, shape or size (Bertin, 1983). These variables enable effective reading of visualizations as they reduce the cognitive efforts for decoding the information due to well-established connections between visual cues and conceptual ideas (Pinker, 1990). Additionally, the evolution has equipped the brain with an extra portion of pattern recognition (Mattson, 2014). One apparent well-studied phenomena in relation to this is the Gestalt theory, which enables humans to fill in gaps, find continuity or recognize patterns to match with existing knowledge and concepts.

Johanna Drucker (2020) is a key figure in the contemporary discussion around the usage of visualizations for interpretative purposes. She argues that interactive visualizations are dynamic mediators of interpretation and therefore a primary way of knowledge production. Nevertheless, Drucker asserts that most visualizations do not have this quality, although such forms of visualizations would be much needed, especially in the humanities. Therefore, she distinguishes between representational visualizations and non-representational visualizations. While the first is a predetermined one-way mapping between abstract datasets and graphic elements, the second allows interactive manipulations of the dataset and visualization parameters. This enables an open-ended and interactive interpretation process which is interrogative, conditional and subjunctive.

In this regard, it is interesting how to design such interrogative visualizations and how they augment the established interpretation techniques. Therefore, this work has employed an explorative design research methodology to produce tangible prototypes, which act as mediators for a reflection on the design process and the implications of certain decisions during this process.

3.1. Research through Design

A design research process is always situated, the process should be reflective and analyze the design decisions and the parameters that have influenced them. Therefore, this work is subscribed to the approach of Hinrichs et al. (2019) which do not evaluate the outcomes of a design-driven research, but the artefacts, findings and questions that arise during the process. In contrast to the layout of most research endeavors the process is open-ended and produced visualizations acts as aesthetic provocations which act as mediators for discussions upon the design decisions.

Additionally, the visualizations provide certain affordances which greatly influence their usage for interpretative purposes. The term "affordance" was coined by James J. Gibson (1966) to describe the potential usages of an environment, e.g. a doorknob that can be used to open the door but also to hang a jacket on it. As this is not intended by designers, Don Norman (2013) extended this concept upon the specific "signifiers", perceivable clues on how to use an object. As interpretation can be quite individual, affordances of visualizations should exhibit a certain flexibility to foster greater interpretative agency.

As shown in the outcomes of the workshop visual forms of interpretation can produce ambiguous visual artefacts which prompt discussion about their ability to act as interpretative devices. This is underlined by the emergent questions of the workshop about the visual encoding of chain-of-thought, meaning of diagrammatic depictions and non-linear structures. This emphasizes the potential of visualizations for interpretation as they promote visualization of thought and create discursive constellations. Algorithms can support the creation of such visual interpretations, e.g. through counting or querying similar words or text passages. However, this process should be dialogical to foster synergetic effects between the text analysis methodologies of human and algorithm. Therefore, design decisions should be oriented on supporting visual interactions between the different forms of analysis, to foster interpretative transformations.

Although this work is interested in a digital environment, paper-based workshops are an efficient method for producing artefacts which prompt discussion and imagination (Bansemir et al., 2014). The outcomes of such a workshop provide insights into the thought patterns, structuring and visual encodings of participants, which provides important information for the design process.

Therefore, the design research process was structured in the following way: (1) conducting a workshop to explore the utilization of visual aids during the text interpretation process, (2) an experimentation phase with sketches of interfaces and prototypical implementations to iterate on design decisions and lastly, (3) a reflection upon the findings of this process.

3.2. Workshop

As the workshop was embedded in the greater scope of this work, two questions guided its setup and evaluation:

(1) Which visual encodings and graphic elements are used as a visual means for interpretation?

(2) Which visual encodings can lead to confusion and highly ambiguous interpretations?

3.2.1. Setup

The workshop has been conducted with seven students from various academic backgrounds such as design, philosophy, architecture, cognitive sciences and engineering. The choice for different backgrounds was made on purpose to get a diverse repertoire of approaches and techniques. Dörk et al. (2020) have proposed to produce tangible artefacts during the process which can be the starting point for inspiration and discussion regarding the visual dimensions of the process. Therefore, all participants were asked to document their process on a DIN A3 poster to visually depict the processes of text interpretation.

Different materials and tools were provided such as markers, pens, glue, strings, transparent paper and scissors (see Figure 5). It was emphasized that the text doesn't have to remain in its original DIN A4 format and can be disassembled with scissors. To introduce the participants to the topic of visualization for text interpretation, a short passage from Elena Esposito (2022) was read together. This text emphasized the usage of visualizations as visual provocations, which prompt question rather than providing answers. Afterwards, all participants were asked to read and interpret a short passage of two pages from Bruno Latour (2008) with a time limit of 60 minutes. "Interpretation" was not defined prior in order to provoke a fallback on individually acquired strategies and associations. Additionally, all participants were asked to focus on visually articulating their thought process during the task.

Figure 5: Workshop participants working on their poster with different tools.

3.2.2. Outcomes

During the task it was observed that almost all participants employed marking pens for highlighting words and sections of interest. This could be attributed to the wide-spread teaching of the technique in secondary schools. Participants used various techniques for visually representing their interpretation process.

While the strategy for positioning elements on the poster was diverse the usage of an almost coherent set of graphical elements was apparent. The usage of links and arrows for pointing out relations were one of the most common used graphical elements.
Some participants used transparent paper to create additional layers while others extracted text passages to annotate them in isolated settings. A collage of some of these artefacts can be seen in
Figure 6.

In the following discussion, participants were asked to describe and interpret the posters of other participants. The owner of the discussed poster was asked to be silent until the end, to answer questions and give insights into their process. The layout of the posters converged to mainly two types: either the source text at the center with peripheral annotations and drawings or some snippets of the text arranged to a rhizomatic mind-map (see Figure 7). This division was also observed in the created artefacts where the first group exhibited visually more structured posters while the latter ones were more chaotic.

Another central observation during the discussion was the intended reading direction or sequence of some of the posters, which raised questions about the importance of visual encoding of sequentiality and temporality. This phenomenon was mainly observed in the real-time annotated visualizations as the space was filled spontaneously, leading to arbitrary placed sequences of thought. Questions were raised about the usage of conventions for depicting sequentiality such as left-to-right (in western cultures) or top-down directions. It was also observed that some graphic elements followed a shared symbolic language while others were more arbitrary in use and therefore harder to decode.

Especially arrows stood out as robustly unambiguous and had been interpreted coherently by most participants while the usage of colors caused multiple times misinterpretations. Other formatting practices, e.g. providing bigger text elements for constructing hierarchies or underlining certain elements, were equally good to read, as they build upon shared understandings.

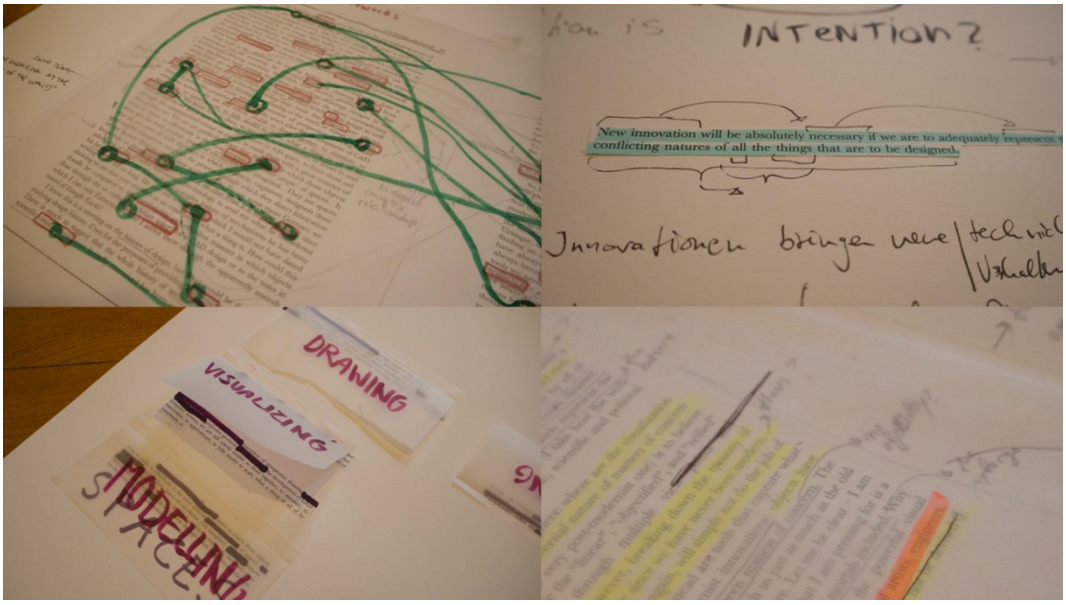


Figure 6: Different text visualizations created by workshop participants.

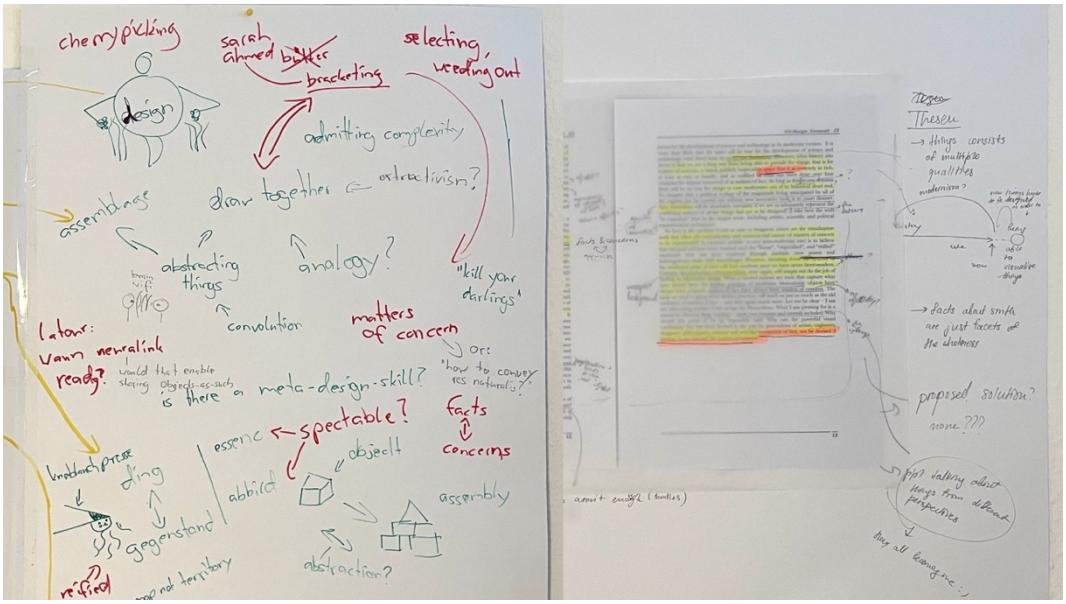


Figure 7: Different poster layouts of workshop participants.

It was emphasized that the usage of visualizations was helpful for the creation of analogies and metaphors to depict abstract concepts more intuitively. Some participants pointed out the non-explicit associations for individuals which transformed visualizations into external cognitive anchors, which were inaccessible to others. It was debated to which extent the visualizations should be accessible to others.

All participants agreed on the importance of a "textual grounding" to ensure a connection between the visualizations and source text fragments. Experimental approaches that physically disassembled the text into snippets or the usage of transparent paper for layering visualizations were received positively (see Figure 6), due to their creative potential to provoke new associations. Participants pointed out the increased appeal of engaging with the text due to the multimodal possibilities. Although the visualizations of each participant exhibited similar sets of graphic elements and variables, the discussion showcased different associations according to the disciplinary backgrounds, provoking arguments about the meaning of symbolics.

Overall, the feedback on the usage of visualizations for the text interpretation process was productive and beneficial to the participants. Regarding the questions, which guided this workshop, five key aspects have been extracted:

(1) Symbols like arrows and circles have a well-established shared foundation of meaning, while colors and abbreviations are ambiguous and hard to decode.

(2) Formatting text elements can provide visual cues for importance of certain elements.

(3) Diagrams were used to visually remix and model interpretation artefacts with existing knowledge.

(4) The chain-of-thought is often uncertain. Visual encodings of sequentiality, like levels of opacity or left-right and top-down, are helpful for guidance.

(5) Visual references to the source text enhance readability, as direct connections can be contextualized more easily.

3.3. Experimentation Process

To further explore these outcomes of the workshop, an experimentation phase was used to create speculative interfaces. These were used as mediators to reflect on the importance of certain design decisions and their influence on the affordances for visual interpretation. As the findings of the workshop have indicated certain aspects of visualizations to be more important, these key points are merged with the theories and methodologies from close-, hyper- and machine reading. As a result, five key aspects have emerged which need to be considered when designing functionalities and visualizations for a digital text interpretation environment:

(1) The format of the source text should be readable by humans and algorithm. As a result, techniques from close reading and natural language processing can be synergized for text analysis.

(2) The structure of the source text should be efficiently transformable to enable comparative analysis of intertextual relationships between the different structures.

(3) All forms of analysis should be done through visual means. Therefore, different interactions must be provided to create visualizations. Additionally, the visualizations must be interactive to turn them into interpretative devices.

(4) To enable a dialogical format between human and algorithm, all elements, which are created by the algorithm should be visually modifiable to reflect the interpretative agency of the human. Furthermore, the possibility to add free form annotations must be provided to augment the visualizations by personal visual inflections and comments.

(5) Affordances should build upon established techniques from engagements with paper (e.g. marking). Furthermore, the design of the affordances should keep the creative process in mind which can benefits from a generative environment which does not need explicit and rigid usage concepts.

3.3.1. Sketches

As the design process produces many associations and conceptual thoughts about an interface, sketching can help to think visually and express ideas. The artefacts created by the workshop participants were also taken into consideration as they provided a repertoire of graphic elements and variables which were also included in the design process. Furthermore, sketching can produce a creative recursive process, where the produced visual artefacts can trigger new associations.

During the beginning of this phase different layouts of the environment were tested on paper to think about limitations of the space and arrangements of the source text and the visualizations. A selection of these sketches can be seen in Figure 8. While the usage of visualizations differed in the sketches a layout stood out as an intuitive arrangement – source text on the left, visualizations on the right. During the process, technical aspects were only partially considered, as the design was foremost interested in a conceptual exploration of interactions and elements. Concepts that provided interesting layouts were further explored in interface mockups with more concrete descriptions of the utilization of elements (see Figure 8). The iterations of prototypical implementations which build upon these sketches are described in the following chapter.

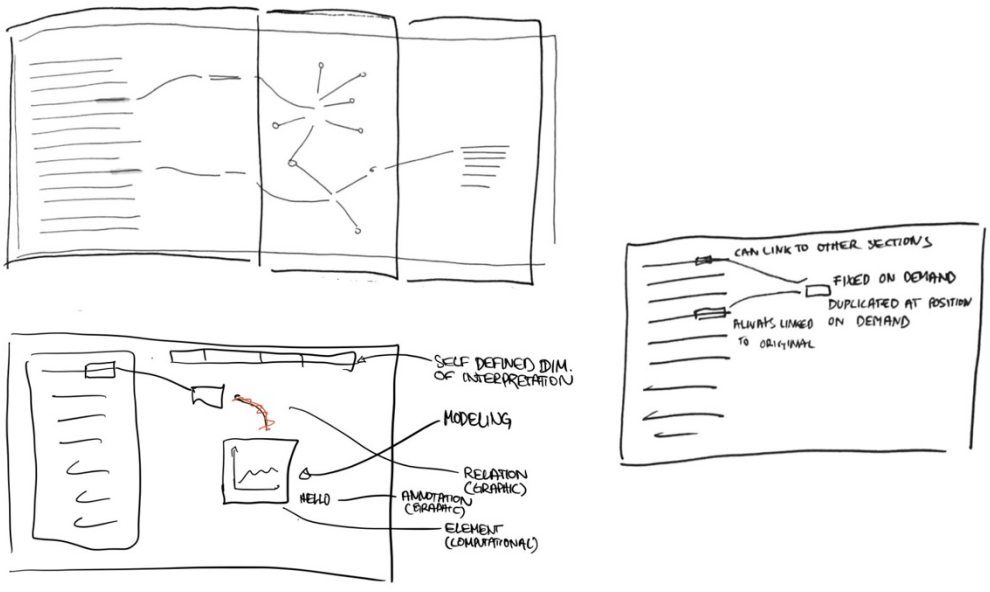


Figure 8: Sketches of possible interfaces and functionalities.

3.3.2. Prototypical Implementations

During the design process prototypical implementations were used to test functionalities and visualizations.[4] A text from Bruno Latour (2008) was used as the textual base for experimenting. Although, several iterations were conducted, only specific iterations will be presented in the following chapters to reflect on more refined functionalities and utilized visualizations. As the development of prototypes is seldom done from scratch, a short introduction to the utilized software packages is given in the next section.

3.3.2.1 Used Software

To develop an application with a certain repertoire of functionalities, in a limited amount of time, one must fall back onto already developed software packages which can be assembled individually. The combination of human and machine-readable text in one environment is most comprehensively implemented by web browsers. The browser-native implemented Hypertext Markup Language (HTML), Cascading Style Sheets (CSS) and JavaScript (JS) provide the base for displaying and manipulating the visual form and interactions between documents in the browser. HTML structures text by creating hierarchies through nested tags such as paragraphs, which can contain headings and other paragraphs.

In terms of visualization algorithms, the D3.js framework proposed by Bostock et al. (2011) provides many application programming interfaces (APIs) to build interactive visualizations. The APIs enable a high level of control and still enable customization during the design process. Additionally, the browser provides some rudimentary functionalities like regular expressions and queries for conducting natural language processing. If the tasks are more complex, the winkNLP JS library provides common algorithms, which can be loaded and used in the browser. Anything that is computationally intensive is commonly offloaded and can be processed in Python. Therefore, Gensim, a well-established Python library, is used to load and inference word2vec and fastText models.

---

[4] Prototypical implementations can be found under https://github.com/elproffesore/masterthesis.

3.3.2.2 Interface Layout

In the context of this work, the central task for the interface was to provide a surface for displaying a source text and at the same time enable to superimpose visual interpretations. In contrast to the portrait orientation of paper-based text, the landscape orientation of digital displays provides significantly more space, however Robert Bringhurst (2004) has shown that a line length of 50-75 characters is ideal to reduce cognitive load and allow an ease in reading. Therefore, the width of the source text was restricted to oblige to this rule and make room for annotations and other visual interpretations.

The placement of the source text is a central decision to make, as it consecutively influences the layout of the application and the possible functionalities. One proposed solution is, to place the source text on the left periphery of the screen to provide a huge space for various visual forms of interpretation. This is especially beneficial for the display of bigger visualizations or spacious modeling of elements. However, the source text takes a literal central role during interpretation and therefore another layout proposal places the source text at the center of the screen. Additionally, the central position simulates the layout of books or other paper-based text formats with blank margins on each side, which can be utilized for placing annotations and visualizations near sections of interest, which is hindered with the left aligned text layout. Both arrangements provided their benefits either in terms of amount of space for expression or resemblance of established layouts (see in Figure 9).

While pagination of long text is common with classic paper-based text formats like books, a simulation of pagination was omitted in benefit for a better display of visual connections among further distanced text fragments. Prototypical implementations used the fixed screen width as the canvas width to ensure the presence of the source text as it provides one of the strongest contextual anchors during the interpretation process. This restriction reduces the complexity of space but also leads to problems of clutter, increasing the cognitive effort for overview and pattern recognition. A digital display allows to extend the canvas into a virtual infinite size which can be leveraged to move elements on the screen out of sight.

One solution to the clutter problem was the implementation of a virtual shift, which moved the visible space to the left and therefore provided the whole screen for modeling and restructuring.

Figure 9: Interface layouts with source text at center or left side of the screen.

These base configurations of the digital interface provide the foundation for the implementation of several interactive elements. During the design and development iterations, two forms of four important elements emerged from the process: extractive markings, free form annotations, network visualizations and the timeline.

3.3.2.3 Extractive Markings

Highlighting text passages of interest is a key technique in close reading. A palette of colors can be used to differentiate categorically or hierarchically between highlighted text segments. This is useful, as key words are mentioned in different contexts across a text and form a conceptual bigger picture. Extractive marking allows to extract a certain text passage (see Figure 10) of interest, which is then placed next to the source text and can be arranged freely on the canvas space. As concepts can be described differently across a text, multiple instances can have different contexts which shape the understanding of a concept.

Because the investigation of relationships among different elements is a common practice in the process of interpretation, the analysis and interpretation of these occurrences is supported by rearranging extractions freely, to model clusters of similar descriptions of e.g., concepts. Lines are culturally established graphic elements which depict relations and paths that connect entities (Tversky, 2011). This was observed in the workshop where almost all participants used lines to connect entities of interest and annotations in the source text. Two important graphic variables for lines are opacity and line thickness. During the design process several levels of line opacities were tested, as they eventually obscure the source text. Line thickness was used to encode interpretative variables such as importance or level of relation. Arrows were used as line caps to visually encode the source of the relation, as either produced by human interaction (pointing towards the extracted word) or an algorithmic created relation (pointing away from the extracted word).

Due to the linear format of the source text, descriptions of concepts or key words can spread out across the full length of a text. This is a problem for comparative tasks. Extracted markings are invisible on a screen that is only able to display a fraction of a text at once without going into miniature font sizes. Therefore, extracted text passages can be "pinned" to a fixed position on the screen, while the text can be scrolled further, enabling a form of parallax. This supports the comparative analysis of multiple, originally distributed, text passages (see Figure 10).

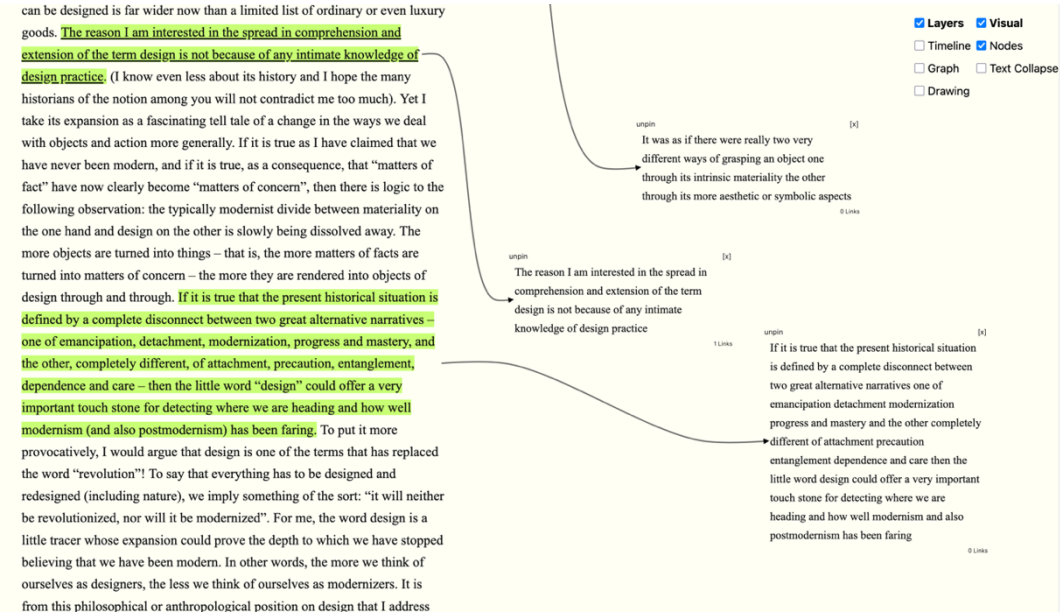


Figure 10: Originally distributed text fragments juxtaposed through pinning.

As the process of highlighting text produces many extractions, a visual hierarchy helps to emphasize certain extractions. To encode such visual hierarchy, altering the size of text elements can provide an intuitive action to highlight elements of interest (see Figure 11). From this perspective, the adjustment is an interpretative act, as the deliberate alteration of graphic elements reflects the agency of the interpreting subject.

Figure 11: Visual hierarchies through graphical variable size.

3.3.2.4 Free Form Annotations

The expressiveness in a digital environment is greatly influenced by the implemented functionalities, graphic elements and variables. The design of addressable graphic elements is hard to achieve as every implementation of predefined elements and relations poses a significantly decreased expressiveness by determining every possible visual form a priori. However, the scalable vector graphics (SVG) format provides a robust framework for the creation of path-based forms. These forms are represented by coordinates which makes them easy to address and alter programmatically. As the browser supports the rendering of SVG, free form input can be directly incorporated into the view. This can be leveraged through the usage of a tablet that enables hand-written input via specialized pencils. Although the forms do not have assigned quantitative or qualitative values, that clearly map to specific interpretations, they can be used to express various forms of annotation such graphic as text or diagrams. A hybrid approach was tested in the prototypes, where relations among extracted words or interesting text passages could be altered or augmented through free form input. As this input format is quite intuitive, individualized annotations and visual schemata can significantly boost expressiveness during the interpretation process (see Figure 12). Additionally, free form annotations also have a strong contrast to typed text which enables better perception of human input.

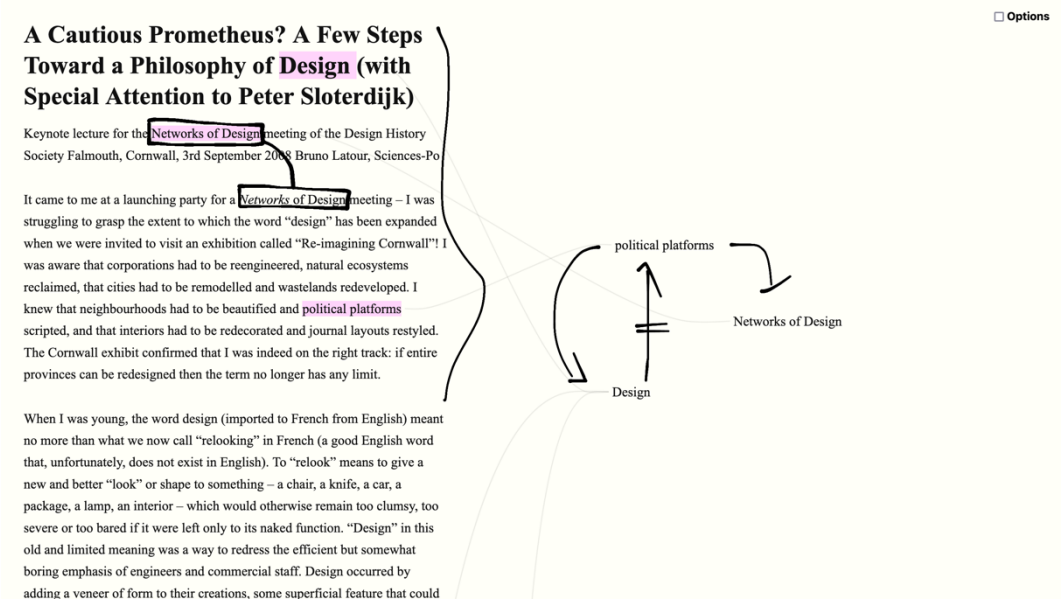


Figure 12: Free form annotations and diagrammatic depiction.

However, the interplay of the static SVG forms and dynamic elements, that have an underlying logic, can cause complication. If extracted text elements are repositioned, annotations and drawings lose their meaning as the context is altered. To address this problem, animated transitions between different constellations and drawings provide an interesting case study for displaying different interpretations of a set of extracted elements.

In conclusion, the synergetic effect between the extracted markings and the possibilities of free form annotations provided an expressive form of input which can foster more individualized processes of analysis through annotations.

3.3.2.5 Semantic Networks

During the usage of close- and hyper reading techniques, the source text can be dissolved into a non-linear network of relations between text fragments of interest. This process can be augmented by algorithms which analyze texts upon structural and semantic similarity. As a result, the extracted relations can be displayed as a graph with visual links between the nodes.

A formal approach to find similar text passages in a source text is the usage of "regular expressions" (described in Chapter 2.2.3). Regular expressions are context-insensitive algorithms, which can find the matches for a formal sequence of characters. As words can contain different inflections of word-stems, other heuristic algorithms such as the Levenshtein distance can be employed to find the same string with different morphologies. Exemplary a network between extracted text passages that contain the word "design" can be seen in Figure 13. As the extraction of elements is done by the interpreting subject, the network is constructed in dialogue with the algorithms as they suggest possible connections among the text passages by highlighting co-appearing words. The network provides a novel non-linear assemblage which exhibits interpretative qualities as it reflects the relations among the curated elements and the algorithmically found similar words.

Although this already provides a great augmentation of the visual interpretation process, words that are morphologically different but share similar semantics are not detectable by such algorithmic procedures.

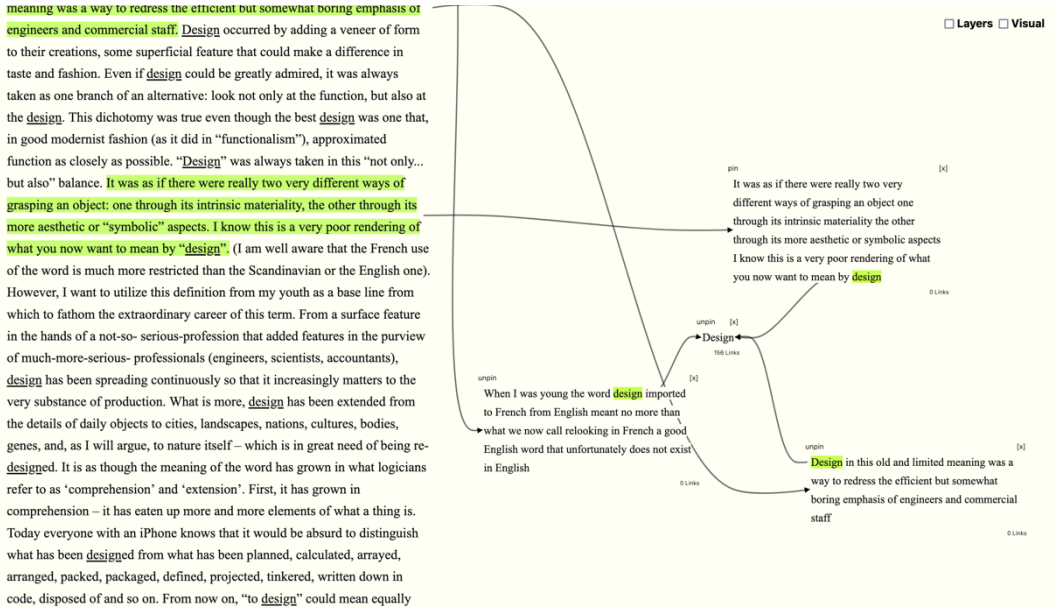


Figure 13: Formal network of relations among similar text elements.

In order to find words and text passages with similar semantics, vector embeddings can be an effective approach. The similarity is calculated based on the cosine distance in vector space between the word embeddings. In the case of this work, the developed prototypes utilized word2vec and fastText models for the semantic analysis. During experimentation two pretrained models were utilized:

- "glove-wiki-gigaword-50" (word2vec model): trained via the unsupervised learning algorithm by Pennington et al. (2014). 50-dimensional vectors trained on Wikipedia and "Gigaword" dataset.
- "cc.en.300" (fastText model): trained via the unsupervised learning algorithm by Mikolov et al. (2017). 300-dimensional vectors trained on "Common Crawl" dataset.

Two approaches were tested for the creation of an algorithmically derived semantic network. The first approach utilized the "most_similar" function provided by the

Gensim API, to get the N most similar words for a given word. Although this produced a large list of candidates, many words were useless, as the model was not aware of the text context. Tweaking the parameters for the output was helpful but provided an extra level of decision-making where most decisions are rather based on intuition.

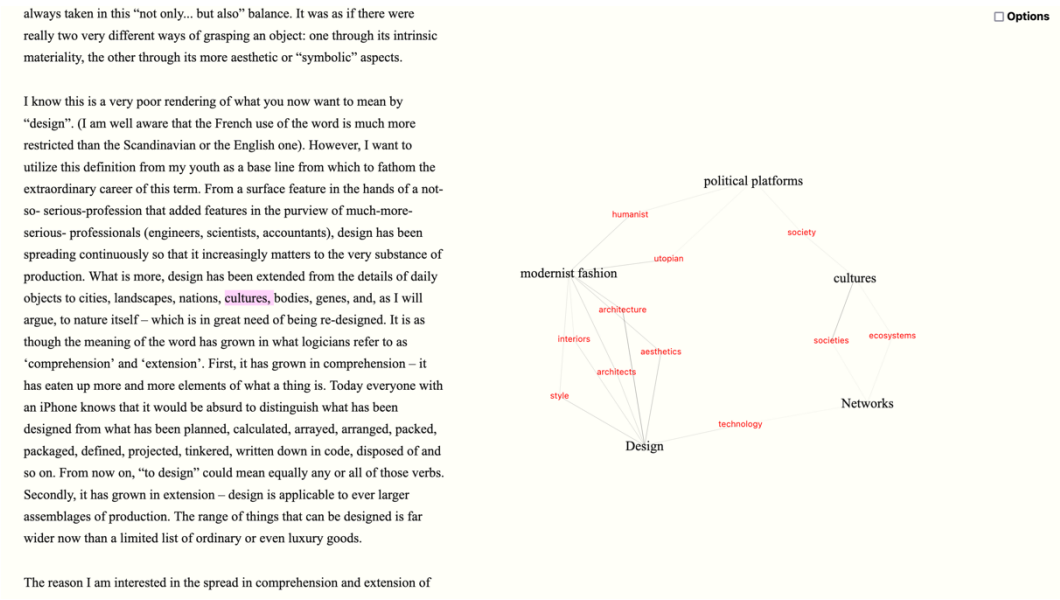


Figure 14: Semantic network between extractions (derived from fastText model).

The second approach didn’t calculate the similarity for every word in the model but only the words that occurred in the source text. To decrease the number of comparisons that need to be made in such a case, stop words were removed from the source text corpora derived by the POS tags. Experimentation with this approach provided significantly more interesting results as the semantically most similar words would also appear in the source text. Algorithmically connected words can act as visual provocations to think about semantic clusters of topics which could be further analyzed. Although this provides an interesting visualization to experiment with, algorithmically created networks increase complexity significantly and could be even regarded as noise, since the words have no further context attached. Furthermore, the display of exponential growing numbers of connections tends to reduce the contrast between relevant and non-relevant relations.

Thus, the construction of a semantic network prompts questions about the role of the interpreting subject. Interactive adjustments of graphic variables such as size and

opacity can help to intuitively highlight elements and declutter the visualization. By deliberately altering the output of the algorithm the interpretative agency of the subject is inscribed and a balance between human and algorithm is restored.

Although this approach provides interesting visual provocations, the two approaches prompt several questions about transparency and traceability of the underlying vector spaces of the models and their inscribed biases. Additionally, without explanation or contextualization it is a hard to understand where the semantic connections come from. It is uncertain if these processes alone add any value, as the process of interpretation is externalized, which is not desired in the case of this work.

During the experimentation different parameters were tweaked, such as the number of similar words which should be incorporated into the network or the threshold a relation must surpass to be relevant. While the tweaks provided better results, the decisions upon a certain threshold are a process of trial and error. This process must be made transparent as it can be seen as an interpretative act to settle on a certain threshold.

3.3.2.6 Timeline

As an in-depth text interpretation process rarely follows the linear flow of the source text, especially when different text analysis techniques are used, temporal relations between visual forms of interpretation are insightful as well. Therefore, an interactive timeline was implemented to investigate the temporal relations among elements via their time of creation and last alteration. As seen in Figure 15, the timeline (on the bottom left) has a cursor for the browsing position and gives additional temporal information about a record. During the interaction all elements get a blur level assigned according to their distance to the cursor. The blur level draws metaphorically from the decay of images where color is slowly getting translucent, indicating the age of an image. This enables to quickly grasp the temporal relations among the elements which can be beneficial for interpreting the importance by order of appearance.

Additionally, every element has a "last changed" parameter, which can also be addressed by the timeline. Although this functionality was interesting for understanding the non-linear process of marking during multiple iterations, the screen didn't allow to

see all elements at once which leads to endless searches while scrolling to see the latest element. A possible solution to this could be the creation of a novel view which arranges elements according to their position in time. This would offer an additional structure of the text interpretation process which doesn't necessarily follow the linear structure of the source text.

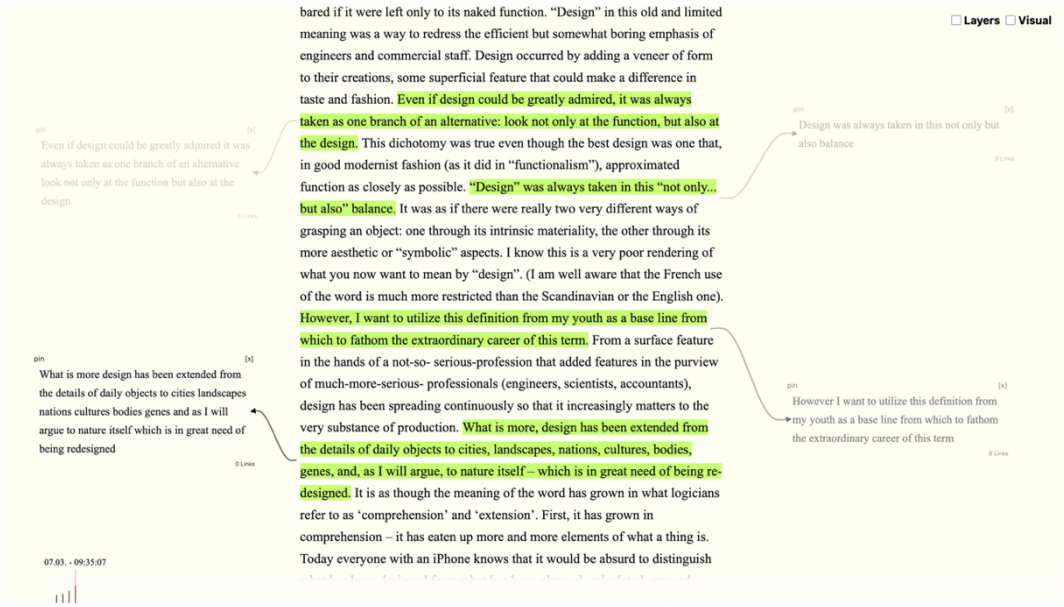


Figure 15: Activated timeline indicates temporal occurrence of elements.

3.4. Reflection

During the research phase of related works, several papers (Esposito, 2022; Hayles, 2010; Kersh & Skalak, 2018) have highlighted the positive aspects of a digital text medium. In this regard, a digital medium could facilitate novel approaches to text analysis, which are nearly impossible with paper, and therefore ask new questions. However, most digital interfaces do not leverage this opportunity but rather only mimic the affordances of paper (e.g., marking and annotation). Although this supports the usage of classical methodologies like close reading, the interfaces do not use their full potential for augmenting. Considering this situation, this design research explored new functionalities to make the interpretation process more interactive and leverage visual analysis.

The prototypical implementations offer several new conceptual functionalities that facilitate an active appropriation of the text through transformation and manipulation of its structure and visual appearance. As the prototypical implementations have a mediative role, this reflection revisits their design process and the aspects that need to be considered to create a digital environment for text interpretation.

3.4.1. Affordances for Visual Interpretation

One apparent advantage of the digital medium is the format of text, which is based on data structures. As the source text of the prototypes was formatted in HTML, every text element was addressable and could therefore be manipulated. If we reconsider Etienne Bonnot de Condillac's (2001) understanding of analysis as a process of composing and decomposing our ideas to create new constellations and investigate their relations, our tools must support this process efficiently. In contrast to paper, the digital medium allows one to efficiently manipulate all textual structures and elements via logic. An interactive interface should facilitate such manipulations by offering affordances that highlight the non-linearity and constructedness of a text and enable changing it.

3.4.1.1 Hacking Structures for Interpretation

As text data does not have an inherent privileged structure or visual form, the manipulation of it becomes an interpretative action. Different structures and visualizations provide insights into relations that emerge from the different structural patterns. The concept of *hacking* comes to mind, a practice where a system is manipulated to do things that were unintended by the authors of the system. In terms of text interpretation, *hacking* questions the authority of the proposed linear text structure and alters it to, e.g., uncover contradictions within argument structures. As this process promotes an active alteration of the text structure, it fosters a deeper understanding of the importance of singular parts and how they constitute the whole. Therefore, hacking texts can be a primary way for interpretation, and as its main objective is to analyze the relation of the parts to the whole, it is also closely connected to the concept of the hermeneutical circle.

As the process is open-ended and iterative, it encourages a playful and explorative form of interpretation as the possibilities to rearrange text structures are almost limitless. Considering these aspects of *hacking* as interpretation, the design process of this work has especially explored the interactive visualizations that support the deconstruction process and produce new visual constellations of the primary text. Since this work is not interested in the "correct" reading of these visualizations but in the usage of them as interpretative devices, the experimental implementations have explored the implications of certain design decisions on affordances and their potential generativity. One apparent important design decision that had a great influence on the affordance of the space was the placement and size of the empty or white space. The design iterations on this subject produced two layouts that produced different affordances. A layout with two peripheral margins evokes associations with affordances of paper and provokes annotations to be placed closely to the source text. In contrast to one singular big margin, which provokes the creation of independent annotations that took themselves for reference. This observation highlights the importance of the design of white space, as it has a great influence on the types of engagement.

3.4.1.2 Margin Characteristics Influence Engagement

In conclusion, peripheral small margins provoke visual interpretations that are context sensitive and in close proximity to the original reference, while one big, joined margin provokes visual interpretations that are independent from the original text. This was also observed in the workshop outcomes, where a whitespace on the margins of a source text produced associative annotations on the side, while blank white spaces provoked mind-map-like structures that didn't orient themselves on the main text. As seen, position and proximity play a crucial role during the visual interpretation process. Especially in the layout with two peripheral margins, the connection to the source text enables establishing context without explicitly restating it. This is often accompanied by highlights of the text, which act as anchors for the annotations that take them for reference. With paper, annotations are bound to a certain position in the linear source text, as the context anchors are immobile.

However, it is most relevant to preserve the context of certain definitions or argument structures throughout the whole linear text for a more critical reception of those. Therefore, these contexts must be mobilized while preserving their original context to support a wider range of comparative analysis tasks.

3.4.1.3 Interactive Deconstruction

Therefore, this work has implemented the functionalities to extract text passages, reposition them, and pin them to a relative position that is not influenced by scrolling. While this process not only enables the active reader to deconstruct the text and rearrange it into new non-linear structures, but it also supports dynamically juxtaposing textual entities that appear early in the text with the ones that appear near the end. As the extractions are automatically visually linked to their source reference, the main context is not lost when repositioning or pinning the element. Many extractions can form a new text structure, which can be visually compared to the original structure, as the source text is juxtaposed with the novel structures on the margin.

In conclusion, the described design decisions support the process of *hacking* the text structure by providing intuitive affordances to deconstruct the old linear text structure, create new non-linear text structures, and enable various forms of comparative analysis through them.

3.4.2. Visual Dialogues with the Algorithm

In terms of a digital text interpretation environment, algorithms play a central role in the processing and display of texts. However, the environment must facilitate a dialogical format during the process of analysis by including algorithm outcomes and human interpretation on a shared plane. Considering these aspects, such a digital environment should employ non-representational visualizations, described by Johanna Drucker (2020), which act as the mediator between human interpretation and digital text data. Non-representational visualizations allow users to interactively alter their own underlying data structures and therefore reflect through the tool.

3.4.2.1 Augmenting Different Reading Methodologies

To leverage this interpretative power of non-representational visualizations in the context of text interpretation, the outcomes of the different methodologies from close, hyper, and machine reading must be rendered visible and modifiable to enable a visual dialogue between their outcomes.

Katherine Hayles (2010) emphasizes the interpretative power of this synergy, as it facilitates a critical engagement with algorithmic outcomes and enables the human to inscribe their own agency during the process. In this regard, Elena Esposito (2022) ascribes algorithms the role of producing "visual provocations," while context and therefore meaning must be established by human actors. As a result, a "hermeneutical circle" is formed where the visual provocations and human interpretation continuously inform the other.

Building on this outline of a visual dialogue between human and machine, the design process must consider the existing methodologies of close, hyper, and machine reading but augment them in a way that their outcomes are rendered visible through interactive visualizations. In terms of this work, two techniques from close reading, highlighting and annotation, were picked to be visually augmented, as they facilitate in-depth analysis, which enables one to formulate their own critique and standpoints. Furthermore, the outcomes of the close reading techniques were made interactive to create non-linear networks, which are part of hyper reading. Lastly, natural language processing algorithms were used to find similar keywords and create semantic networks through the inference of word vector models.

3.4.2.2 Graphic Elements and Variables for Visual Conversing

In terms of the prototypical implementations, the environment provided several affordances to perform and render the described techniques on the same visual level. A small set of graphical elements was used, such as lines and arrows, to make visual nuances of them more meaningful and build upon their established foundation for depicting relationships (Tversky, 2011). Graphic variables, such as opacity, line thickness, size, and position, were used to indicate relationships, temporality, or hierarchies between elements.

These elements and variables provide an expressive repertoire for visualizing interpretative actions, as they already exhibit culturally established semantics but can be visually altered to reflect context-sensitive values. To further enhance the experience of visual interpretation, freeform annotations enable hand-drawn diagrammatic schemata and recontextualization according to the human interpretation. As a result, the expressiveness is greatly enhanced as the meaning is not only encoded in the discrete values of the graphical variables but also in the combination of multiple elements and variables.

However, certain aspects can have negative influences on the visual dialogue process. Especially increased numbers of elements that are visually on the same hierarchy can have a significant impact on cognitive processes, as overview and orientation are impaired. Therefore, it is crucial that the interface design provides functionalities to establish hierarchies and let humans decide on the importance of each element. In this regard, the prototypes of this work find solutions in interactive network and timeline visualizations, which help to group elements and search for semantically similar text segments. Furthermore, all elements can be repositioned, visually altered (according to the described graphic variables) or deleted to let the human subject decide how to deal with the visual provocation of the algorithm.

Drawn together, these functionalities provide the possibility to transform the text interpretation process into a visual augmented dialogue. The communication happens on the visual plane, where semantics are context-sensitive and meaning emerges from the agency of the human subject. Nevertheless, it must be noted that the digital display of elements can still exhibit visual characteristics that do not reflect any semantic meaning but are perceived so, e.g., size or position. Therefore, a critical reception of algorithmically created visualizations is needed.

4. CONCLUSION

Due to the emergence of digital text formats, the process of text interpretation has changed substantially in the last decades. The characteristics of the digital medium offer novel forms of analysis that can shed new light on the constructedness of a text. Visualizations play a crucial role during this analysis phase, as visual expression is closely connected to natural language, and additionally, pattern recognition abilities are an incremental part of cognitive processes. Nevertheless, digital tools most often create static visualizations of algorithmic text analysis outcomes that leave no room for the inscription of human interpretation. Therefore, this work has explored the design of a digital text interpretation environment that is augmented by interactive visualizations. The prototypical implementations have showed some important design aspects which can greatly influence the affordances of such an environment.

In this regard, considering (RQ1), a digital environment should strive to augment established methodologies from text analysis, such as close or hyper reading, and synthesize them through interactive visualizations. Furthermore, the design must provide affordances to *hack* the original text structure and equip the reader with tools to create new structures that offer new angles on the text and therefore new interpretations. This encompasses the design of sufficient white space that affords to be filled by the interpreter. Lastly, a small set of graphical elements and variables, such as lines and position, provide a generative repertoire that is upon culturally established meanings and can therefore be used to intuitively express interpretations visually. Furthermore, considering (RQ2), the analysis and interpretation process must be transformed into a dialogue between algorithm and human. The interactions of humans with a text must be answered by visual provocations of algorithms and vice versa. As a result, humans can leverage the myriad functionalities of digital computing while establishing the context that is necessary to create meaningful outcomes. However, as humans and computers operate differently, this dialogue must be joined on a visual plane, where both actors can read and interpret the interactions of the other. Non-representational visualizations can be the communication device for doing so, as the presentation is done by the algorithm, but the interaction with it creates the meaning.

In conclusion, interactive text visualizations significantly enhance interpretative agency, as they allow to actively reflect interpretation through visual means. Especially the process of dissolving a text into a non-linear structure significantly benefits from visual augmentation, as text data does not have an inherently privileged visual form. Therefore, playful reconfigurations can reveal information about certain relations and under what conditions they exist. Lastly, a balance between algorithmic analysis and human interpretation can be established through a visual dialogue, which synergizes algorithmic efficiency and human specific contexts to generate new knowledge in the process.

Future work could further investigate the interactive visualizations of dynamic text structures with an emphasis on network structures that emerge during the analysis process. Furthermore, the dialogue mode could be further explored regarding a ping pong of visual provocations and human annotations, which allows humans to be in the loop and understand a system by conversing with it.

BIBLIOGRAPHY

Adler, M. J., & Van Doren, C. (1972). *How to Read a Book* (Rev. and updated ed). Simon and Schuster.

Alonso Schökel, L. (1998). *A manual of hermeneutics*. Sheffield Academic Press.

Armaselu, F., & Heuvel, C. van den. (2017). Metaphors in Digital Hermeneutics: Zooming through Literary, Didactic and Historical Representations of Imaginary and Existing Cities. *Digital Humanities Quarterly*, *11*(3). https://www.digitalhumanities.org/dhq/vol/11/3/000337/000337.html

Aston, G., & Burnard, L. (2022). *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh University Press. https://doi.org/10.1515/9780748628889

Bansemir, B., Hannß, F., Lochner, B., Wojdziak, J., & Groh, R. (2014). Experience Report: The Effectiveness of Paper Prototyping for Interactive Visualizations. In A. Marcus (Ed.), *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience* (pp. 3–13). Springer International Publishing. https://doi.org/10.1007/978-3-319-07668-3_1

Baumer, E. P. S., Jasim, M., Sarvghad, A., & Mahyar, N. (2022). Of Course it's Political! A Critical Inquiry into Underemphasized Dimensions in Civic Text Visualization. *Computer Graphics Forum*, *41*(3), 1–14. https://doi.org/10.1111/cgf.14518

Bertin, J. (1983). *Semiology of graphics*. University of Wisconsin Press.

Best, S., & Marcus, S. (2009). Surface Reading: An Introduction. *Representations*, *108*(1), 1–21. https://doi.org/10.1525/rep.2009.108.1.1

Bludau, M.-J., Brüggemann, V., Busch, A., & Dörk, M. (2020). Reading Traces: Scalable Exploration in Elastic Visualizations of Cultural Heritage Data. *Computer Graphics Forum*, *39*(3), 77–87. https://doi.org/10.1111/cgf.13964

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). *Enriching Word Vectors with Subword Information* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1607.04606

Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1607.06520

Bostock, M., Ogievetsky, V., & Heer, J. (2011). D^3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2301–2309.

Boyles, N. (2013). Closing in on Close Reading. *Educational Leadership*, *70*(4), 36–41.

Brill, E. (1992). A simple rule-based part of speech tagger. *Proceedings of the Third Conference on Applied Natural Language Processing* -, 152. https://doi.org/10.3115/974499.974526

Bringhurst, R. (2004). *The elements of typographic style* (3rd ed). Hartley & Marks publ.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (2007). *Readings in information visualization: Using vision to think* (Nachdr.). Morgan Kaufmann.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*(3), 113–124. IRE Transactions on Information Theory. https://doi.org/10.1109/TIT.1956.1056813

Collins, C., Carpendale, S., & Penn, G. (2009). DocuBurst: Visualizing Document Content using Language Structure. *Computer Graphics Forum*, *28*(3), 1039–1046. https://doi.org/10.1111/j.1467-8659.2009.01439.x

Condillac, E. B. de. (2001). *Essay on the origin of human knowledge* (H. Aarsleff, Ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139164160

Dilthey, W. (1900). Die Entstehung der Hermeneutik.

Dörk, M., & Knight, D. (2015). WordWanderer: A navigational approach to text visualisation. *Corpora*, *10*(1), 83–94. https://doi.org/10.3366/cor.2015.0067

Dörk, M., Müller, B., Stange, J.-E., Herseni, J., & Dittrich, K. (2020). Co-Designing Visualizations for Information Seeking and Knowledge Management. *Open Information Science*, *4*(1), 217–235. https://doi.org/10.1515/opis-2020-0102

Drucker, J. (2014). *Graphesis: Visual forms of knowledge production*. Harvard University Press.

Drucker, J. (2020). Visualization and Interpretation: Humanistic Approaches to Display. The MIT Press.

Eisenstein, J. (2019). Introduction to natural language processing. The MIT Press.

Esposito, E. (2022). Reading Images: Visualization and Interpretation in Digital Text Analysis. *Artificial Communication*. https://artificialcommunication.mitpress.mit.edu/pub/jtkikw5z/release/1

Figal, G. (2007). Hermeneutik. In Reallexikon der deutschen Literaturwissenschaft: Neubearbeitung des Reallexikons der deutschen Literaturgeschichte. Bd. I: A – G. Bd. II: H – O. Bd III: P – Z (3rd ed., pp. 29–31). De Gruyter.

Gadamer, H.-G. (1975). *Truth and Method*. Crossroad.

Gascoigne, T., Schiele, B., Leach, J., Riedlinger, M., Lewenstein, B. V., Massarani, L., Broks, P., & Australian National University Press (Eds.). (2020). *Communicating science: A global perspective*. Australian National University Press.

Gibson, J. J. (1966). The senses considered as perceptual systems. Houghton Mifflin.

Guillory, J. (2008). How Scholars Read. *ADE Bulletin*, 8–17. https://doi.org/10.1632/ade.146.8

Hayles, N. K. (2010). How We Read: Close, Hyper, Machine. *ADE Bulletin*, 62–79. https://doi.org/10.1632/ade.150.62

Hayles, N. K. (2012). How we think: Digital media and contemporary technogenesis. The University of Chicago Press.

Hinrichs, U., Forlini, S., & Moynihan, B. (2019). In defense of sandcastles: Research thinking through visualization in digital humanities. *Digital Scholarship in the Humanities*, *34*(Supplement_1), i80–i99. https://doi.org/10.1093/llc/fqy051

Jänicke, S., Franzini, G., Cheema, M. F., & Scheuermann, G. (2015). On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In *Eurographics Conference on Visualization (EuroVis)—STARs* (p. 21 pages). The Eurographics Association. https://doi.org/10.2312/EUROVISSTAR.20151113

Jasinski, J. (2001). *Sourcebook on Rhetoric*. SAGE Publications.

Kalir, R. H., & Garcia, A. (2021). *Annotation*. The MIT Press. https://doi.org/10.7551/mitpress/12444.001.0001

Kehoe, A., & Gee, M. (2013). eMargin: A Collaborative Textual Annotation Tool. *Ariadne*, *71*. http://www.ariadne.ac.uk/issue/71/kehoe-gee/(accessed 15.12.2024)

Kersh, S. E., & Skalak, C. (2018). From Distracted to Recursive Reading: Facilitating Knowledge Transfer through Annotation Software. *Digital Humanities Quarterly*, *12*(2). http://www.digitalhumanities.org/dhq/vol/12/2/000387/000387.html

Kim, N. W., Bach, B., Im, H., Schriber, S., Gross, M., & Pfister, H. (2018). Visualizing Nonlinear Narratives with Story Curves. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 595–604. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2017.2744118

Latour, B. (1986). Visualization and Cognition: Drawing Things Together. *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, *6*, 1–40.

Latour, B. (2008). A Cautious Prometheus ? A Few Steps Toward a Philosophy of Design. *Proceedings of the 2008 Annual International Conference of the Design History Society*.

Liu, S., Wu, Y., Wei, E., Liu, M., & Liu, Y. (2013). StoryFlow: Tracking the Evolution of Stories. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2436–2445. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2013.196

Manning, C. D., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT Press.

Mattson, M. P. (2014). Superior pattern processing is the essence of the evolved human brain. *Frontiers in Neuroscience*, *8*, 265. https://doi.org/10.3389/fnins.2014.00265

McLuhan, M. (1966). *Understanding media; the extensions of man*. Signet Books. https://search.library.wisc.edu/catalog/999567338002121

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, *331*(6014), 176–182. https://doi.org/10.1126/science.1199644

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1301.3781

Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., & Joulin, A. (2017). *Advances in Pre-Training Distributed Word Representations* (arXiv:1712.09405). arXiv. https://doi.org/10.48550/arXiv.1712.09405

Nelson, T. H. (1965). Complex information processing: A file structure for the complex, the changing and the indeterminate. *Proceedings of the 1965 20th National Conference On* -, 84–100. https://doi.org/10.1145/800197.806036

Nielsen, J. (2006, April 16). *F-Shaped Pattern For Reading Web Content (original eyetracking research)*. Nielsen Norman Group. https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content-discovered/(accessed 10.01.2025)

Norman, D. A. (2013). *The design of everyday things* (Revised and expanded edition). Basic Books.

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. In A. Moschitti, B. Pang, & W. Daelemans (Eds.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1162

Pinker, S. (1990). A theory of graph comprehension. In *Artificial intelligence and the future of testing* (pp. 73–126). Lawrence Erlbaum Associates, Inc.

Rheingold, H. (2000). Tools for thought: The history and future of mind-expanding technology (1. MIT Press ed). MIT Press.

Risch, J., Kao, A., Poteet, S. R., & Wu, Y.-J. J. (2008). Text Visualization for Visual Text Analytics. In S. J. Simoff, M. H. Böhlen, & A. Mazeika (Eds.), *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics* (pp. 154–171). Springer. https://doi.org/10.1007/978-3-540-71080-6_11

Rockwell, G., & Sinclair, S. (2017). Hermeneutica: Computer-assisted interpretation in the humanities. The MIT Press.

Schleiermacher, F. (1838). *Hermeneutik und Kritik*. G. Reimer.

Schwan, H., Jacke, J., Kleymann, R., Stange, J.-E., & Dörk, M. (2019). Narrelations—Visualizing Narrative Levels and their Correlations with Temporal Phenomena. *Digital Humanities Quarterly*, *13*(3). https://www.digitalhumanities.org/dhq/vol/13/3/000414/000414.html

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. *Proceedings 1996 IEEE Symposium on Visual Languages*, 336–343. https://doi.org/10.1109/VL.1996.545307

Thaller, M. (2022, May 9). On Hermeneutics and Computing [Billet]. *A Digital Ivory Tower*. https://ivorytower.hypotheses.org/184

Tufte, E. R. (1998). *The visual display of quantitative information* (16th printing, January 1998). Graphics Press.

Tversky, B. (2011). Visualizing Thought. *Topics in Cognitive Science*, *3*(3), 499–535. https://doi.org/10.1111/j.1756-8765.2010.01113.x

Walny, J., Huron, S., Perin, C., Wun, T., Pusch, R., & Carpendale, S. (2018). Active Reading of Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 770–780. https://doi.org/10.1109/TVCG.2017.2745958

Zanini, N., & Dhawan, V. (2015). Text Mining: An introduction to theory and some applications. 19.