Text Mining and Natural Language Processing in the Humanities: A Review of Methods and Applications in Historical Texts, Literature, and Social Media

Pavithra Panduwawala[1]
[1]University of Moratuwa, Sri Lanka
pavithrapanduwawala@hotmail.com[1]

Abstract

The integration of Text Mining and Natural Language Processing into the humanities has significantly transformed the way scholars engage with textual data. As digital archives expand and textual information becomes increasingly available in digitized form, the need for computational methods to process, analyze, and interpret large volumes of unstructured data has grown. This paper presents a comprehensive review of how text mining and NLP are being applied in the fields of historical research, literary analysis, and social media studies, with a focus on both their methodological foundations and practical implementations. The study begins with a literature review that traces the origins and evolution of text mining and NLP technologies, examining how they have been gradually adopted within the humanities. It then explores key tools and techniques such as tokenization, part-of-speech tagging, sentiment analysis, topic modelling, and named entity recognition. These methods have enabled researchers to extract valuable insights from massive textual corpora while maintaining academic rigour and interpretive depth. Through a series of case studies, the paper highlights applications in three major domains: historical texts, where patterns across centuries can be revealed; literature, where authorial styles and genre conventions are analyzed, and social media, where real-time discourse, identity construction, and public sentiment are studied. While the benefits of computational analysis are evident, the research also addresses limitations such as linguistic ambiguity, algorithmic bias, lack of historical linguistic resources, and ethical concerns surrounding data use and interpretation. By providing an interdisciplinary overview, this review contributes to the growing field of digital humanities and encourages a balanced, critical, and context-sensitive use of text mining and NLP techniques. The paper concludes with reflections on emerging trends, ethical frameworks, and future directions for research at the intersection of technology and humanistic inquiry.

Keywords: Text Mining, Natural Language Processing (NLP), Digital Humanities, Historical Texts, Literary Analysis, Computational Linguistics, Topic Modelling, Sentiment Analysis.

Article Number: SCHORES-JSSH-01-010
SchoRes Journal of Social Sciences and Humanities
Volume 1, Issue 2, December 2024
Date Published: 15 January 2025

www.schores.org

1. Introduction

In recent decades, the integration of computational tools into the humanities has led to a paradigm shift in how textual data is analyzed, interpreted, and understood. Among the most transformative of these tools are Text Mining and Natural Language Processing two interrelated methodologies that enable the extraction of patterns, sentiments, structures, and meanings from large corpora of unstructured text (Piotrowski, 2012; Pang and Lee, 2008). These techniques, traditionally rooted in computer science and linguistics, are now being adapted for a wide range of humanistic inquiries. From decoding ancient manuscripts to analyzing modern-day tweets, the digital humanities are increasingly reliant on algorithmic methods to engage with texts in more scalable and nuanced ways (Jockers, 2013; Underwood, 2019).

The core appeal of text mining and NLP lies in their ability to process vast amounts of textual information quickly and systematically. While humanists have always been concerned with close reading and contextual analysis, the exponential growth in digital text archives, historical documents, online literature, and social media conversations has made it impractical to manually examine every available source (Schöch, 2013; Boyd and Crawford, 2012). These new tools allow scholars to combine traditional qualitative insights with quantitative rigour, offering hybrid methodologies that open new avenues for research.

In historical studies, text mining enables the discovery of hidden patterns across centuries of archival data. For instance, researchers can map the evolution of political discourse over time or detect subtle changes in the use of moral vocabulary during different eras (Klein, 2013). In literary studies, NLP helps scholars examine genre conventions, authorial styles, intertextuality, and sentiment across large literary corpora challenges that are difficult to address through conventional close reading alone. Likewise, in social media analysis, researchers use sentiment analysis, topic modelling, and named entity recognition to understand public discourse, cultural trends, and identity formation in the digital age.

Even with its vast potential, the application of these methods in the humanities is not without challenges. Issues such as language ambiguity, historical language variations, context preservation, and data accessibility pose significant barriers (Piotrowski, 2012; Smith, 2007). Furthermore, ethical concerns around privacy, algorithmic bias, and interpretability raise important questions about how digital tools should be used in humanistic inquiry (Bender et al., 2021; Van Zundert, 2012).

Section 2 presents a detailed literature review, tracing the historical development of text mining and NLP techniques and their adoption in the humanities. Section 3 discusses the core methods and tools commonly used in digital humanities research, including linguistic preprocessing, topic modelling, and sentiment analysis. Section 4 explores specific applications in historical texts, literature, and social media, using case studies to illustrate real-world usage. Section 5 outlines the challenges and ethical considerations associated with computational textual analysis. Finally, Section 6 presents the future directions and conclusion, summarizing key findings and proposing avenues for further research.

2. Literature Review

The use of computational methods in the humanities has grown substantially over the last two decades, forming the foundation of what is now known as Digital Humanities. Central to this transformation are Text Mining and NLP**,** which originated in the fields of computer science and computational linguistics. Over time, these techniques have been adapted to humanistic research, allowing scholars to process and analyze vast corpora of text that were previously inaccessible through traditional methods.

2.1. Origins and Evolution of Text Mining and NLP

Text mining and NLP date back to the mid-20th century with early developments in machine translation and computational linguistics. As computer processing power advanced, so did the ability to analyze natural language. In the 1990s, statistical and rule-based methods were developed, and with the rise of the internet, digital text became more widely available, further expanding the possibilities for large-scale text analysis (Pang and Lee, 2008; Piotrowski, 2012). The term "Digital Humanities" began gaining traction in the early 2000s as scholars from literature, history, philosophy, and other fields started integrating digital tools into their research practices (Underwood, 2019). By the 2010s, NLP had become a key component of digital humanities, especially in text-heavy disciplines.

2.2. Integration into Historical and Literary Studies

In historical research, text mining has enabled the analysis of archival materials, parliamentary debates, newspapers, and personal letters. For example, computational approaches are used to detect shifts in language and social trends across large datasets, highlighting patterns that would otherwise go unnoticed through traditional reading (Klein, 2013; Ridge, 2014). Techniques such as word frequency analysis, co-occurrence mapping, and topic modelling have proven particularly effective in these contexts (Jockers, 2013). In literary studies, early work focused on stylometry, or the quantitative study of literary style, while modern approaches use NLP for authorship attribution, genre classification, and sentiment analysis (Craig and Kinney, 2009). Projects such as the HathiTrust Digital Library and Stanford's Literary Lab exemplify the scalability and analytical depth that NLP brings to literary analysis (Jockers, 2013).

2.3. Use in Social Media and Contemporary Communication

The study of social media using NLP is more recent but rapidly evolving. Platforms such as Twitter, Reddit, and Facebook offer rich, unstructured textual data that are ideal for analysis using sentiment detection, named entity recognition (NER), and stance classification (Cinelli et al., 2020). These tools have helped researchers uncover patterns in online discourse, public sentiment during elections or crises, and the spread of misinformation (Boyd and Crawford, 2012; Bender et al., 2021). For instance, during the COVID-19 pandemic, researchers used sentiment analysis and topic modelling to trace public reactions, conspiracy theories, and the spread of misinformation in real time

2.4. Methodological Shifts and Challenges

There has also been a growing awareness of the limitations of these methods in humanistic inquiry. Text mining often struggles with ambiguity, polysemy, and the loss of context, particularly in historical or literary texts where interpretation is highly nuanced (Piotrowski, 2012). Furthermore, many NLP tools are trained on contemporary English corpora and show reduced effectiveness when applied to archaic language or non-standard dialects (Smith, 2007).

Scholars have also raised concerns about the "black box" nature of machine learning algorithms, noting that a lack of transparency can reduce interpretive trust and analytical depth (Bender et al., 2021; Van Zundert, 2012). Ethical issues such as data privacy, algorithmic bias, and the unauthorized use of user-generated social media content are increasingly highlighted in the literature (Bender et al., 2021). These concerns underline the importance of critical awareness and ethical reflection when applying computational tools in humanistic contexts.

From historical archives and literary corpora to real-time social media data, text mining and NLP are reshaping how researchers interpret textual information. Yet, their integration into humanities research is still developing, and debates continue regarding the balance between computational efficiency and interpretive nuance (Underwood, 2019; Klein, 2013)

3. Methods and Tools in Digital Humanities

The integration of text mining and NLP into the humanities depends on a solid methodological framework. This section explores the main components of that framework, including linguistic preprocessing, computational analysis methods, digital tools, and programming libraries. These methods not only enhance the scalability of research but also allow scholars to analyze texts with both depth and precision.

3.1. Linguistic Preprocessing Techniques

Text analysis in the humanities begins with preprocessing, which prepares raw textual data for structured analysis. One of the most foundational steps is tokenization, where texts are divided into meaningful units such as words, sentences, or paragraphs. This helps in organizing text for further processing. Another key step is the removal of stop words, which involves eliminating common but non-essential words such as “the”, “and”, or “is”, thus allowing more important terms to stand out (Piotrowski, 2012).

To unify different grammatical forms of a word, researchers apply stemming or lemmatization. While stemming cuts words to their base root (often mechanically), lemmatization considers the word's context to produce more accurate base forms. Another important preprocessing task is part-of-speech (POS) tagging, which assigns grammatical roles (like nouns, verbs, and adjectives) to each word. These steps form the linguistic backbone of any computational text analysis and are implemented through tools such as the Natural Language Toolkit (NLTK), spaCy, and Stanford CoreNLP (Pang and Lee, 2008).

3.2. Analytical Techniques for Textual Interpretation

Once preprocessing is completed, researchers can apply analytical techniques to draw insights from the data. Sentiment analysis is widely used to assess the emotional tone of a text whether positive, negative, or neutral. This is particularly relevant in both literary analysis, where characters' emotional expressions are analyzed, and in social media studies, where public sentiment is tracked. Another common approach is topic modelling, which identifies dominant themes within large text collections. Among topic modelling methods, Latent Dirichlet Allocation (LDA) is especially popular due to its ability to uncover hidden thematic structures in unlabeled data. For identifying people, places, or events within a text, scholars use named entity recognition (NER)**.** This technique enables the automatic classification of named items, providing valuable context in historical and social data analysis.

3.3. Digital Platforms Supporting Humanities Research

Several platforms have been created to help humanities researchers use text mining and NLP methods without requiring advanced programming knowledge. Voyant Tools, for instance, is an online environment that allows users to upload texts and generate visual representations such as word clouds, frequency graphs, and correlation maps. It is commonly used in teaching and exploratory research. **AntConc** is another popular desktop application used for concordance generation and keyword analysis. It allows users to search for specific terms and analyze their usage within context. The Text Analysis Portal for Research (TAPoR) acts as a digital hub, connecting scholars to a wide range of text analysis tools. Meanwhile, large-scale initiatives like the HathiTrust Research Center (HTRC**)** provide access to millions of digitized books and offer cloud-based tools to perform scalable literary analyses across entire corpora.

3.4. Programming Libraries for Advanced Analysis

For more customized or advanced projects, scholars often use programming languages such as Python and R. These languages offer powerful text mining capabilities through various libraries. In Python, tools like NLTK, spaCy, TextBlob, and Gensim are used for tasks ranging from preprocessing

to topic modelling. Machine learning tasks, including classification and clustering, are often performed using scikit-learn. In R, packages like tidytext, quanteda, and tm enable users to clean, model, and visualize textual data in structured workflows (Jockers, 2013). These open-source libraries support reproducibility and transparency, both of which are essential for academic research. Although these tools require some technical skill, online documentation and growing community support make them increasingly accessible to humanities researchers. Text mining and NLP methods rely on a sequence of structured tasks that begin with linguistic preprocessing and extend to deep analytical interpretation. The availability of user-friendly platforms and powerful programming libraries has made it easier for humanities scholars to conduct digital research. As tools continue to evolve, they will further support interdisciplinary collaboration and more sophisticated textual investigations.

4. Applications in Historical Texts, Literature, and Social Media

The interdisciplinary use of text mining and NLP in the humanities becomes most evident in applied research. This section examines how these technologies have been used in three major areas: historical documents, literary corpora, and social media content. Each domain offers distinct challenges and opportunities, reflecting the diversity of textual materials and interpretive goals in humanistic research.

4.1. Historical Texts and Archives

In the field of historical research, computational text analysis has opened new possibilities for studying large-scale archival materials. Texts such as court records, newspapers, parliamentary debates, and correspondence have been digitized and made accessible through platforms like *The Old Bailey Online*, *British Newspaper Archive*, and *Europeana*. These resources contain millions of words and span several centuries, making manual reading impractical. Using topic modelling and frequency analysis, historians have been able to trace how concepts like crime, gender, or religion evolved over time. For example, in the analysis of 18th-century British court proceedings, scholars have used word co-occurrence networks to examine how legal language around theft and punishment shifted (Ridge, 2014). Named Entity Recognition (NER) is frequently applied to extract references to people, places, and institutions, enabling researchers to map historical networks and geographies. One of the key benefits in this domain is the ability to identify broad social patterns and institutional changes. However, challenges such as archaic spelling, obsolete vocabulary, and poor-quality digitization often require customized tools or historical dictionaries to ensure accuracy (Piotrowski, 2012).

4.2. Literary Analysis and Interpretation

The use of computational methods in literary studies is one of the most established applications within the digital humanities. Stylometry, authorship attribution, and sentiment analysis are widely used to explore literary style, narrative structure, and thematic development. Projects such as the *Stanford Literary Lab* and the *HathiTrust Research Center* have enabled researchers to examine thousands of novels and poems across different time periods and languages. One of the significant achievements in this area is the identification of authorship in disputed texts. For instance, stylometric analysis has contributed to debates over the authorship of Shakespearean plays and pseudonymous 19th-century novels (Craig & Kinney, 2009). Similarly, sentiment analysis has been used to track emotional progression in novels, helping to understand narrative arcs and character development (Jockers, 2013). Topic modelling has also revealed genre-specific themes in literary texts, such as the recurring motifs in Gothic, Romantic, or Realist literature. By identifying patterns across vast literary corpora, scholars can move beyond anecdotal interpretations and base their

claims on statistical evidence. Nonetheless, interpretive caution is needed, as computational methods cannot fully account for irony, symbolism, or cultural context without human insight.

4.3. Social Media and Digital Discourse

Social media platforms present a unique and increasingly important area for applying text mining and NLP. Unlike historical or literary texts, social media content is produced in real-time, informal, and highly dynamic. Researchers have studied political discourse, social movements, misinformation, public opinion, and online identity using platforms such as Twitter, Facebook, Reddit, and Instagram. Sentiment analysis and stance detection are commonly used to assess public attitudes during elections, protests, or pandemics. For example, studies during the COVID-19 crisis employed sentiment analysis to track public anxiety and misinformation across regions and over time (Cinelli et al., 2020). Named Entity Recognition and clustering techniques have also been used to detect trending topics, influencers, and coordinated campaigns.

5. Challenges and Limitations

While the integration of text mining and NLP in the humanities offers powerful analytical capabilities, it also introduces several challenges. These issues are technical, interpretive, and ethical in nature. Moreover, the effectiveness of these tools depends heavily on the type and quality of data, as well as the researcher's familiarity with computational methods. This section identifies the major limitations faced in historical, literary, and social media contexts and presents them in a comparative format.

5.1. Data Quality and Digitization Issues

One of the most persistent challenges in the digital humanities is the poor quality of source data. Historical texts often suffer from scanning errors, non-standardized spelling, and degraded manuscripts. Optical Character Recognition (OCR), the technology used to convert scanned images into text, is particularly unreliable for older documents with ornate fonts or damage (Smith, 2007). In literature, while many texts are available in digital form, they often lack critical metadata such as publication dates, genres, or author information. Social media data, though abundant, is highly informal and varies in structure, tone, and content. These inconsistencies complicate preprocessing and can introduce significant noise into the analysis, requiring custom cleaning procedures or specialized models.

5.2. Interpretive Complexity and Loss of Context

Another limitation lies in the interpretive scope of computational methods. While machines can detect patterns in language, they struggle with understanding context, irony, metaphor, and cultural nuance. This is particularly significant in literary and historical analysis, where meaning is often deeply embedded in context. For instance, sentiment analysis tools may misinterpret sarcasm as genuine emotion, leading to misleading conclusions (Pang & Lee, 2008). In literary studies, thematic interpretations can vary among human readers, but computational tools tend to fix meaning based on word frequencies or sentiment scores. Therefore, a purely data-driven approach may oversimplify complex humanistic questions unless carefully paired with traditional interpretive methods.

5.3. Technical and Accessibility Barriers

Many humanities scholars face barriers related to computational skillsets. Tools like Python, R, or even advanced platforms like Gensim or spaCy require programming knowledge that is not traditionally taught in humanities disciplines. Although user-friendly tools like Voyant exist, they offer limited analytical depth. In contrast, more powerful tools often lack intuitive interfaces and require training in data science (Klein, 2020). Moreover, computational projects often demand considerable computing resources, which may not be accessible to all institutions or researchers, especially in underfunded departments.

5.4. Ethical Considerations and Data Privacy

When applying text mining to social media data, ethical concerns around privacy and consent become critical. Although much online content is public, the individuals who generate that content do not always expect it to be used for academic research. Researchers must navigate issues of anonymisation, consent, and the potential misuse of findings (boyd & Crawford, 2012). Furthermore, NLP models may carry biases if they are trained on skewed or non-representative datasets. This can perpetuate historical inequalities or reinforce stereotypes, particularly in studies involving gender, race, or nationality (Bender et al., 2021). To better illustrate these challenges across the three domains explored earlier historical texts, literature, and social media a summary table is provided below.

Table 1: Comparative Overview of Challenges Across Textual Domains

Challenge	Historical Texts	Literary Texts	Social Media
Data Quality	OCR errors, archaic language	Missing metadata, formatting inconsistencies	Informal language, slang, unstructured data
Interpretive Limits	Context loss in old references	Ambiguity in themes and symbolism	Misinterpretation of irony, memes, emojis
Technical Barriers	Need for custom scripts	Limited programming familiarity	High data volume and dynamic content
Ethical Issues	Intellectual property concerns	Reuse of copyrighted literary content	Privacy, consent, bias in datasets
Tool Accessibility	Limited tools for specific formats	Low compatibility with literary annotation	Need for APIs and real-time data tools

With the growing success of text mining and NLP in the humanities, researchers must remain aware of their limitations. Data quality, interpretive accuracy, technical skill gaps, and ethical sensitivity all influence how findings should be understood and applied. A balanced approach combining computational rigour with humanistic insight is essential for overcoming these limitations and ensuring responsible use of digital methods in scholarly research.

6. Future Directions in Digital Humanities

The intersection of text mining, NLP, and the humanities is rapidly evolving, offering transformative possibilities for future scholarship. As digital tools grow more advanced and interdisciplinary collaboration expands, the digital humanities are poised to become even more influential in how we interpret, preserve, and engage with texts. This section explores key areas of growth and innovation that are likely to shape the future of this dynamic field.

6.1. Multilingual and Cross-Cultural Analysis

One of the emerging priorities in digital humanities is the development of robust tools for analyzing multilingual and culturally diverse texts. Most current NLP models and corpora are heavily biased towards English, which limits research on texts in languages such as Arabic, Urdu, Swahili, or indigenous languages. As the field expands, efforts are being made to create cross-linguistic models that can interpret grammar, syntax, and cultural idioms with comparable accuracy. Projects like CLARIN (Common Language Resources and Technology Infrastructure) are working to create multilingual repositories and language models, making it possible for researchers to conduct comparative studies across national literatures, colonial texts, and global historical narratives. These efforts not only decentralize the field but also encourage inclusivity and the exploration of underrepresented voices in historical and literary archives.

6.2. Integration of AI and Deep Learning

While traditional NLP methods such as topic modelling and sentiment analysis remain foundational, the future of textual analysis lies in more sophisticated machine learning and deep learning approaches. Transformer-based models such as BERT, RoBERTa, and GPT have shown exceptional capabilities in understanding context, generating summaries, and answering questions with human-like nuance. In digital humanities, these models can enhance text interpretation by recognizing subtle literary features, metaphorical language, or shifts in authorial tone. Deep learning can also improve historical document restoration, automatic translation, and semantic search in large archives. However, their successful implementation in the humanities will require model fine-tuning on domain-specific corpora and greater transparency to address interpretability concerns.

6.3. Visualization and Storytelling Interfaces

As digital humanities become more interactive, the need for advanced visualization tools grows. Future projects are expected to incorporate more dynamic visual storytelling techniques, such as digital timelines, narrative graphs, and geo-spatial mapping. These interfaces help make scholarly findings more engaging and accessible not only to academics but also to the public and policymakers. Examples include projects like “Mapping the Republic of Letters” and “Six Degrees of Francis Bacon,” which use network diagrams to represent historical and intellectual relationships. Future platforms may expand these capabilities using augmented reality and virtual reality to create immersive experiences based on textual data.

6.4. Ethical Frameworks and Responsible AI

As computational methods become more powerful, it is essential to develop ethical frameworks that ensure the responsible use of data and algorithms in the humanities. This includes practices such as bias auditing of NLP models, informed consent in social media research, and transparent reporting of data collection and processing methods. The integration of ethics should not be a separate concern but embedded within every stage of the research process. Journals, universities, and funding bodies are increasingly recognizing the importance of data ethics in digital scholarship and are expected to formalize guidelines that support equitable and inclusive research practices.

7. Conclusion

The integration of text mining and NLP into the humanities represents a significant advancement in how scholars engage with texts, languages, and cultural narratives. These digital methods have introduced new levels of scale, speed, and objectivity to disciplines traditionally grounded in qualitative analysis and critical interpretation. Through this paper, it has been shown that computational methods have had a profound impact across three major domains: historical texts, literature, and social media. In historical research, text mining enables scholars to explore vast corpora that would otherwise be inaccessible due to time and scale constraints. In literary studies, NLP tools such as stylometry, topic modelling, and sentiment analysis have contributed to questions of authorship, genre, and emotional structure. Meanwhile, in the realm of social media, real-time analysis of discourse, sentiment, and community interaction offers valuable insight into contemporary communication and public opinion. Although these achievements, several challenges persist. Issues such as poor data quality, interpretive oversimplification, ethical concerns, and limited technical accessibility continue to limit the full potential of computational methods in the humanities. These limitations emphasize the need for interdisciplinary collaboration, customized tool development, and ethical responsibility in all stages of digital research. Looking forward, the future of digital humanities is shaped by promising developments. The emergence of multilingual NLP, deep learning models, semantic annotation, and interactive visualization tools signals a move toward more inclusive and sophisticated analysis. At the same time, an increased focus on ethical frameworks and critical awareness ensures that this technological progress is used responsibly and equitably.

Ultimately, text mining and NLP are not a replacement for traditional humanistic inquiry but a powerful extension of it. By combining computational precision with interpretive depth, scholars can uncover new patterns, challenge assumptions, and make previously invisible aspects of culture and history visible. As tools and methods continue to evolve, the digital humanities will remain an essential space for innovation at the crossroads of technology and human thought.

References

Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S., 2021, March. On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Boyd, D. and Crawford, K., 2012. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, *15*(5), pp.662-679.

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C.M., Brugnoli, E., Schmidt, A.L., Zola, P., Zollo, F. and Scala, A., 2020. The COVID-19 social media infodemic. *Scientific reports*, *10*(1), p.16598.

Craig, H. and Kinney, A.F., 2009. *Shakespeare, computers, and the mystery of authorship*. Cambridge University Press.

Jockers, M.L., 2013. *Macroanalysis: Digital methods and literary history*. University of Illinois Press.

Klein, L.F., 2013. The image of absence: Archival silence, data visualization, and James Hemings. *American Literature*, *85*(4), pp.661-688.

Piotrowski, M., 2012. *Natural language processing for historical texts*. Morgan & Claypool Publishers.

Pang, B. and Lee, L., 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, *2*(1–2), pp.1-135.

Ridge, M.M. ed., 2014. *Crowdsourcing our cultural heritage*. Ashgate Publishing, Ltd..

Schöch, C., 2013. Big? smart? clean? messy? Data in the humanities. *Journal of digital humanities*, *2*(3), pp.2-13.

Smith, R., 2007, September. An overview of the Tesseract OCR engine. In *Ninth international conference on document analysis and recognition (ICDAR 2007)* (Vol. 2, pp. 629-633). IEEE.

Underwood, T., 2019. *Distant horizons: digital evidence and literary change*. University of Chicago Press.

Van Zundert, J., 2012. If you build it, will we come? Large scale digital infrastructures as a dead end for digital humanities. *Historical Social Research/Historische Sozialforschung*, pp.165-186.