DE GRUYTER OLDENBOURG

Doing Digital Scholarly Editing

Anna Ananieva*, Sandra Balck and Jacob Möhrke

The Study of Historical Travelogues from a Digital Humanities Perspective: Experiences and New Approaches

Abstract: This article explores how Digital Humanities methodologies can be applied to historical travelogues and highlights the potential of these technologies to provide new insights into historical data. It summarizes the experiences acquired through various research tasks undertaken by the interdisciplinary project team "Digital Editions of Historical Travelogues" (DEHisRe), funded by the German Research Council (DFG) from 2021 to 2024. Based on an iterative case study of handwritten travel journals from the early 19th century, the authors outline the six-step "Life Cycle of Digital Editing". This circular workflow incorporates best practices in the field while aiming to produce sustainable digital editions of historical travelogues.

Keywords: historical travelogues; digital humanities; digital scholarly edition

Introduction to the Case Study

The field of Digital Humanities (DH) has evolved significantly, bringing new technical challenges in producing digital scholarly editions. These challenges include the need for specialized software, adherence to precise standards, and specific technical skills. The new approaches affect the identification of individuals, locations, and objects within the edited material, the enrichment of text with annotations, and the simultaneous display of digital copies of handwritten originals along with the transcripts and comments. Moreover, automated and semi-automated tasks such as handwritten text recognition (HTR) and named entity recognition (NER) have the potential to enhance efficiency and accuracy in scholarly research.

*Corresponding author: Anna Ananieva, Leibniz Institute for East and Southeast European Studies, Regensburg, Germany, E-mail: ananieva@ios-regensburg.de. https://orcid.org/0000-0003-1584-2692
Sandra Balck, Freie Universität Berlin, Berlin, Germany
Jacob Möhrke, Leibniz Institute for East and Southeast European Studies, Regensburg, Germany

The question of how to make travelogues digitally accessible for historical research was addressed in the research project "Digital Editions of Historical Travelogues" (*Digitale Editionen Historischer Reiseberichte*, DEHisRe). Funded by the German Research Foundation (*Deutsche Forschungsgemeinschaft*, DFG), the project was based at the Leibniz Institute for East and Southeast European Studies (IOS) in Regensburg. From October 2021 to September 2024, the project team explored how to transform, process, and code historical travel accounts in a manner that enables scholars to tackle complex research questions regarding the travel practices and mobile culture of the past.[1]

Central to this endeavor was the digital edition of the handwritten travel journals written by Franz Xaver Bronner (1758–1850), a German-speaking writer, journalist, and professor of physics, who traveled from Aarau in Switzerland to Kazan on the Volga in 1810, returning from Russia in 1817 (Figure 1). Bronner's extensive travel narratives offer a rich, multifaceted view of everyday life in Central and Eastern Europe in the early 19th century (Beyer-Thoma 2017). His detailed yet little-known travel journals from the holdings at the Aargau State Archive (StAAG NL.A-0019) also reflect his engagement with a number of contemporaries in politics and culture, including literary figures such as Christoph Martin Wieland.

Figure 1: Travel Itinerary of Franz Xaver Bronner in 1810 and 1817. Source: Hermann Beyer-Thoma (QGIS Cloud Free 2022).

---

1 For further details, see the project's websites: "DEHisRe: Digitale Editionen Historischer Reiseberichte." https://dehisre.ios-regensburg.de and "DeHisRe01." https://github.com/dehisre01/ (accessed 20 June 2024).

With the study of Bronner's travel journals, the DEHisRe team aims to achieve several goals. First, the study provides access to the previously unpublished records of a German migrant from Switzerland and makes this historical source available to a wide audience through a digital scholarly edition. In addition to the editorial and text-critical tasks, the project team deals with the issues of cultural and migration history, the development of transport and technology, and intercultural herme-neutics. This allows the detailed reconstruction and examination of various aspects of historical travel culture, such as travel routes and means of transport, customs duties and travel costs, as well as post station equipment in the vast European area, including its northeastern and southeastern regions. The traveler's specific view of the people and landscapes he observes along the way only comes to light through the edited travelogue, which will thus be made available for intersectional discussions.

Second, the processing of the manuscript serves as a case study for future digital editing of historical travelogues. The DEHisRe project explores the use of digital editions as a basis for analyzing historical travel accounts and for testing which tools are suitable for transcribing, annotating, and visualizing these accounts. The study seeks to answer the following questions: What sustainable technical solutions and low-threshold tools are currently available for the digital edition of historical trav-elogues? How can researchers use these tools effectively to make handwritten source material machine-readable? And more generally, how can digital resources be created from historical sources to effectively provide new knowledge and research opportunities? These questions have driven the work on the case study of Franz Xaver Bronner's handwritten travel journals (Ananieva et al. 2024; Balck et al. 2023). The purpose of the following discussion is to offer a detailed insight into the various areas of work undertaken by the interdisciplinary project team DEHisRe.

Ongoing Exchanges with the International Research Community

Research on historical travel, particularly the study of travelogues, is a popular field in literary and cultural studies. Scholarship focuses on cultural aspects of traveling, such as the emergence and rules of the literary genre of the travelogue, reception research, cultural contacts and stereotypes, as well as gender aspects. Practical and technical aspects of traveling, including routes, roads, personal equipment, net-works, the quality of inns, money, currencies, costs, and means of transport, are also important. Moreover, relationships between certain groups of people, such as stu-dents, artists, and scientists, are considered. In most cases, the focus is not on the

journey itself, but on the destinations and the perceptions of the people doing the traveling.[2]

Due to this widespread interest in historical travel research, there has been an increasing focus on the historical sources of spatial mobility, particularly travelogues. The digital transformation of material resources and travel accounts presents both new challenges and opportunities. To make digital editions of travel records analyzable, data needs to be identified and enriched with additional information, and narratives of the travel events need to be explicitly modeled. This leads to our specific research question: How can materials such as travel reports be prepared and encoded in a way that makes it possible not only to present the content digitally, but also to use it to answer complex research questions about cross-European travel practices and mobile culture?

Focusing on Eastern and Southeastern Europe, the project team organized two workshops in Regensburg to test the effectiveness of various methods and tools currently being used for historical travel research. These workshops involved a wide range of DH experiments and discussions with teams from other research projects. The "Opening Historical Travelogues" workshop took place on 20–21 July 2023 and focused on text markup, data modeling, and visualization.[3] On 1–2 February 2024, the workshop "From Place to Place" continued the discussion and explored new perspectives on the digital editing of travel narratives more broadly.[4] Through our interaction with other experts, we identified critical questions and assessed the latest advancements in historical travel research and digital editing of travelogues. As a result, we developed a "Life Cycle of Digital Editing" that we are currently using in our research. Finally, the outcome of three years of collaborative work was shared at an interdisciplinary conference titled "Cultures of Travel: Historical Travel Practices and Digital Humanities", marking the completion of the DEHisRe research project. This conference took place in Tallinn on 29–30 August 2024.[5] It provided a

---

2 For a recent overview, see Gruber, Doris. "Europeans Encounter the World in Travelogues, 1450–1900." *Europäische Geschichte Online (EGO)*, 8 March 2022. https://www.ieg-ego.eu/de/threads/europa-und-die-welt/kuenste/doris-gruber-europeans-encounter-the-world-in-travelogues-1450-1900 (accessed 20 June 2024).

3 Möhrke, Jacob, and Corwin Schnell. "Tagungsbericht: Historische Reiseberichte öffnen: Textaus-zeichnung, Datenmodellierung und Visualisierung." *H-Soz-Kult*. 9 November 2023. www.hsozkult.de/conferencereport/id/fdkn-139736 (accessed 20 June 2024).

4 For the workshop program, see "Workshop 'Von Ort zu Ort: Digitales Edieren von Reisenarrativen – Neue Perspektiven.'" DEHisRe project website. 21 November 2023. https://dehisre.ios-regensburg.de/workshop-von-ort-zu-ort-digitales-edieren-von-reisenarrativen-neue-perspektiven/ (accessed 20 June 2024).

5 For the conference programme, see "Conference 'Cultures of Travel: Historical Travel Practices and Digital Humanities'." DEHisRe project website. 23 May 2024. https://dehisre.ios-regensburg.de/international-conference-cultures-of-travel-historical-travel-practices-and-digital-humanities/ (accessed 20 June 2024).

comprehensive platform to explore the intersections of historical travel practices and Digital Humanities more broadly, spanning from the Middle Ages to the present day, with a focus on the Eastern European region in a global context.

Throughout the final conference, the participants examined the sociocultural and technical conditions of spatial mobility, including who traveled, why, and how. They inspected the impact on mobility of a range of factors – economic, political, and social – and how this, in turn, impacted the travelers' experiences. The conference also explored the ways in which travels were documented in text and images, and how mobile people composed these self-testimonies. In addition, the conference discussed how travel experiences were documented through text and images, including journals, letters, drawings, paintings, and photographs. The accuracy and reliability of these sources, as well as their influence on perceptions of travel and the external world, were critically assessed. In particular, participants at the conference explored how digital methods could be utilized here.

The Challenges of Innovation and Responsibility on the Path from the Digitized to the Digital Edition

The process of digital editing involves meticulously preparing, annotating, and contextualizing texts and documents to make them available for analysis and use in digital environments. It entails more than simply converting physical content into digital form, as it involves establishing a critical framework that documents and reflects on the editing and interpretation of the material. The focus is on being innovative, open, and sustainable while allowing for future updates and corrections. The balance between innovation and sustainability poses a particular challenge, as innovative approaches are often heavily dependent on individual experts or specialized software which will become outdated if it does not undergo continuous maintenance. The application of standards and careful documentation of the work steps is an effective approach to meeting this need for balance (Crompton 2023). Consequently, the workflow and outcomes of the DEHisRe project align with the current European and national initiatives in the Digital Humanities and Digital Editing, such as the Digital Research Infrastructure for the Arts and Humanities (DARIAH) and the National Research Data Infrastructure (*Nationale Forschungsdateninfrastruktur*, NFDI). While working on the Franz Xaver Bronner case study, we explored and tested various methods for the digital edition of historical travelogues. These ranged from manual to semi-automatic approaches for transcription to the use of large language models (LLMs) such as generative pretrained transformers (GPT) for annotation. Some

details and preliminary results are outlined in the following section, highlighting the issues of both innovation and responsibility in digital scholarly editing.

Innovations: AI and Digital Editing

Soon after the DEHisRe team started their work on the case study of Bronner's travelogues, the advent of large language models (LLMs), often synonymous with advancements in artificial intelligence (AI), notably intersected with the domain of digital editing. While machine learning methodologies have long been integral to certain aspects, such as named entity recognition (NER) – a mode of information extraction through predefined categories such as persons, organizations, or locations – the comprehensive impact of LLMs on the editing workflow remains to be fully determined. This unfolding interest in the capabilities of LLMs was increasingly evident at the events organized by the DEHisRe team.

Our project embarked on preliminary explorations using LLMs, in particular employing GPT-4 for NER tasks. The use of LLMs offers some advantages: ease of use, adaptability, and a multilingual capacity that mitigates the challenges posed by the linguistic variations prevalent in historical texts or languages spoken by smaller groups. Preliminary results from utilizing GPT-4 for NER in the analysis of Bronner's travelogue, where traditional training data is scarce, were promising.[6] Moreover, LLMs demonstrate a superior flexibility over traditional machine learning methods in identifying a broader spectrum of entity categories without predefined constraints.

Despite these benefits, the deployment of LLMs, such as the proprietary GPT-4, presented significant challenges related to accessibility, transparency, and sustainability. These models require substantial resources and have a tendency towards so-called "hallucination" – the generation of unfounded content – which poses risks to the reliability, reproducibility, and factual integrity of the outcomes. Although there are strategies to mitigate these limitations, these often compromise the aforementioned ease of use. Nevertheless, the rapid integration of LLMs into digital editing processes, at least on an exploratory level, and their proven potential suggest a pivotal role for these technologies in the future of the field.

---

6 Möhrke, Jacob, Sandra Balck, and Anna Ananieva. 2024. "Zum Einsatz von GPT-4 für NER: Ein Experiment anhand historischer Reisetexte." Abstract of the paper for the Workshop Generative KI, LLMs und GPT bei digitalen Editionen. DHd 2024. Quo Vadis DH? (DHd2024), 29 March 2024. https://zenodo.org/records/10893761 (accessed 20 June 2024).

Responsibility: "Life Cycle of Digital Editing"

By incorporating the best practices in the field, the DEHisRe introduced a six-step "Life Cycle of Digital Editing". This circular workflow is based on an iterative process and is aligned with the "Life Cycle of Historical Information", which has been identified as the six stages through which a piece of historical information passes when digitized: creation, enrichment, editing, retrieval, analysis, and presentation (Boonstra, Breure, and Doorn 2004; Meroño-Peñuela et al. 2015). In reference to this model, the "Life Cycle of Digital Editing" focuses on responsible resource management in Digital Humanities by outlining objectives for sustainable action in digital scholarly editions, including efficient use of available development resources and infrastructure, consistent pursuit of code efficiency using open standards, and transparent documentation for data reusability.[7]

The life cycle begins with the digitization and transcription (1) of the texts to be edited, followed by the modeling and marking (2 & 3) of the data contained in the text. The processed data can provide answers to complex information queries (4), thereby making it possible to analyze them (5). Visualization (6) creates new access to the texts, which allows various perspectives. This may inspire new research questions that trigger new iterations of the cycle, whether with new sources to expand the database or with the same source to expand or optimize digital methods and data models (Figure 2).

Step 1: Creating

The first step aims to create a digital copy of the material object, which forms the basis of the subsequent steps of digital editing. "Creating" describes the transformation of the physical document into digital formats, which begins with the selection of documents to be digitized, followed by the digitization itself and subsequent manual transcription or automated text recognition using suitable tools: "Transcription is the reproduction of a text as it is found in a material document. The aim of transcription is to produce an identical text, an objective reproduction, an exact representation" (Sahle 2013, 251). For our case study, we chose the

---

7 Balck, Sandra, Jacob Möhrke, and Anna Ananieva. "Poster: Digitale Editionen historischer Reiseberichte: Der Kreislauf historischer Informationen." Paper given at the Digital Humanities Day Leipzig 2023. Forum for Digital Humanities, Leipzig, 4 December 2023. https://doi.org/10.5281/zenodo.10301629; Balck, Sandra, Jacob Möhrke, and Anna Ananieva. Abstract of "Responsible and Sustainable Editing: A Life Cycle for Digital Editions of Historical Travelogues (DEHisRe)." Paper given at the conference DH2024 Reinvention & Responsibility, 6–10 August 2024, Washington, D.C.

The first step, creating (1), involves selecting the historical sources to be edited, followed by digitization and then manual or automated transcription (HTR) using appropriate tools. This step aims to produce a digital equivalent of the original document, which serves as the foundation for subsequent actions and helps preserve cultural heritage through digital scholarly editions.
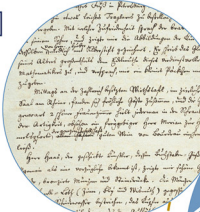
During the data modeling (2) we apply open standards and data models based on the TEI guidelines, the DTA Base Format (DTABf), and the CIDOC-CRM to secure a flawless data exchange, the merging of information from different sources, and the data reuse for extended analyses.

The editing (3) aims to enrich the texts using the previously created data models, while employing methods like NER and authority data from GND, VIAF, and GeoNames for identification and interoperability.
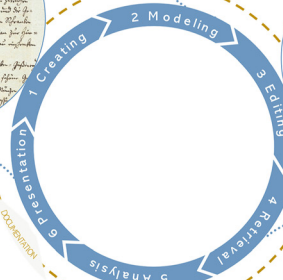
A presentation (6) can be delivered through digital platforms, web applications, and as a print publications. When presenting edited textual material from historical travelogues, spatio-temporal visualizations play a crucial role. These visualizations provide expanded access to the travelogue beyond traditional scholary editions printed on paper. By applying and utilizing established data formats (2 & 3), it becomes possible to visualize the content using a variety of tools. This not only delivers spatio-temporal information, but also allows for linking with other historical travelogues and reusing them in new contexts.

Information retrieval (4) not only serves as the foundation for creating the digital edition but also for subsequent analysis (5) and visualization (6). Tasks at this stage include creating systems and algorithms to extract relevant information from large datasets (using XSLT and XQuery). Based on the modeling (2) and the editing process (3), this step is vital for making the data easily accessible and usable.

Figure 2: A "Life Cycle of Digital Editing". Source: Sandra Balck, Jacob Möhrke, and Anna Ananieva (DEHisRe 2024).

"Transkribus" platform[8] for its ease of use: lower technical requirements reduce barriers and facilitate interdisciplinary collaboration.[9]

In the realm of handwritten text recognition (HTR) – the automated transcription of digitized handwritten documents – current technologies exhibit proficient capabilities when applied to materials with clear layouts and prevalent handwriting styles. Nevertheless, accuracy can be improved through the fine-tuning of existing models. In our case study, we identified a model whose capabilities closely align with the handwriting of the travel reports.

---

8 "Unlock the Past with Transkribus." Transkribus. https://www.transkribus.org/ (accessed 20 June 2024).

9 Balck, Sandra, and Jacob Möhrke. "Digital Edition of Travelogues: The Role That Transkribus Plays." Paper given at the Transkribus User Conference 22. Innsbruck, 29 September 2022. https://youtu.be/fMc7N7_k6hE? (accessed 20 June 2024).

Subsequently, this model was retrained utilizing manually transcribed pages from Bronner's manuscripts. While flawless transcriptions remain elusive with present-day HTR technology, error rates of below 2 % of incorrect characters is within the realm of possibility, provided the handwritten material adheres to certain standards of neatness. Mühlberger et al. (2019) proposed that an error rate threshold of 10 % serves as a pragmatic benchmark to assess the utility of HTR in the transcription process. Beyond this threshold, the labor involved in correcting errors introduced by HTR becomes more time consuming than undertaking manual transcription from the outset.

Step 2: Modeling

During the data modeling step, the digitized document is structured and enriched. This includes defining relationships between data, categorizing information, and creating schemes. To enrich the text with additional information, we use a data model based on the Text Encoding Initiative (TEI) guidelines.[10] The TEI is "[…] a consortium which collectively develops and maintains a standard for the representation of texts in digital form".[11] The TEI guidelines serve as a de facto standard for digital editions. This standard is based on extensible markup language (XML), whose main purpose is to store, transmit, and reconstruct arbitrary data for serialization. The TEI is freely accessible and is used for the digital editing and processing of texts: "The Text Encoding Initiative is the most systematic effort so far to create standards for scholarly memory in an evolving digital culture. The TEI markup language includes […] different elements to satisfy all manner of scholarly needs in the humanities."[12]

The use of open standards creates a consistent and comparable database and enables the lossless exchange of data, the merging of information from different sources, and the reuse of data for extended analyses. We achieve greater standardization by also applying the German text archive basic format (*Deutsches Textarchiv Basisformat*, DTABf), which serves as a guideline for the formatting and structuring of digital texts in German-speaking countries:

---

10 Balck, Sandra, Anna Ananieva, Hermann Beyer-Thoma, Ingo Frank, Jacob Möhrke, and Corwin Schnell. "Poster: Building a Digital Infrastructure for the Edition and Analysis of Historical Travelogues." TEI 2022 Conference Book, 70-71, of a paper given at #TEI2022, Newcastle, 12–16 September 2022. https://doi.org/10.5281/zenodo.7071025.

11 "A Very Gentle Introduction to the TEI Markup Language." Text Encoding Initiative – TEI. 2019. https://tei-c.org/Vault/Tutorials/mueller-index.htm (accessed 20 June 2024).

12 "A Very Gentle Introduction."

> Within the framework of the DTA guidelines, the DTA basic format should […] enable comprehensive text editing and at the same time limit the scope for variation in annotation in such a way that coherence […] is guaranteed […] in order to further minimize ambiguities and consequently misinterpretations of the markup (Haaf et al. 2015).

In addition to the text annotation with TEI, we apply the International Committee for Documentation's (CIDOC) conceptual reference model (CRM) for ontological modeling, which is an event-oriented and widely accepted ontology (Balck et al. 2023). The CIDOC, an entity of the International Council of Museums (ICOM), promotes documentation standards and guidelines for museums.[13] The committee began to develop the CRM in 1996 as an application ontology for the documentation of museum artefacts and was accepted as an International Organization for Standardization's (ISO) standard for data modeling in the cultural heritage sector in 2006.[14] The CIDOC CRM is intended to serve as a transdisciplinary "lingua franca" for the standardization of knowledge on cultural heritage objects and structures and aims to facilitate the exchange of relevant information between museums, libraries, archives, and similar institutions, as well as between different research projects.

Step 3: Editing

The editing step enriches the texts using the previously created data models. The digitized texts are corrected, annotated, and provided with metadata. In addition to manual annotation, methods such as NER and ontological annotation (CIDOC CRM) are used.[15] For example, in the case of Bronner's travel report, a semi-automated marking of dates was implemented which reuses existing software and modifies it to suit the specifics of the case study.[16] The open-source program HeidelTime,

---

13 "Welcome to CIDOC – ICOM International Committee for Documentation." International Council of Museums. https://cidoc.mini.icom.museum/cidoc-icoms-international-committee-for-documentation-provides-the-museum-community-with-advice-on-good-practice-and-developments-in-museum-documentation/ (accessed 20 June 2024).
14 "What is the CIDOC CRM?" Conceptual Reference Model. https://www.cidoc-crm.org/ (accessed 20 June 2024).
15 Balck, Sandra, and Ingo Frank. "Digitale Editionen von historischen Reiseberichten öffnen: Open Text und Open Data mit einheitlicher Textauszeichnung, semantischer Annotation und ontologie-basierter Datenmodellierung." Abstract for the paper given at the conference DHd 2023 Open Humanities Open Culture. 9. Tagung des Verbands Digital Humanities im deutschsprachigen Raum, Trier and Luxemburg, 13–17 March 2023. https://doi.org/10.5281/zenodo.7715261.
16 Möhrke, Jacob. "Semi-automatisierte Annotation von Zeitangaben." DEHisRe Project Website. 10 May 2023. https://dehisre.ios-regensburg.de/semi-automatisierte-annotation-von-zeitangaben (accessed 20 June 2024).

developed at the University of Heidelberg, was used for this purpose.[17] The results were then supplemented with rules adapted to the text source in the form of regular expressions. This procedure combines two rule-based approaches, a general, comprehensive, and powerful model, and then a primitive but highly specific set of rules, which, used together, are equipped to reliably locate dates in Bronner's digitalized handwritten travelogue.

The editing also includes linking place and personal names with existing databases, which is a resource-efficient solution because it reduces storage use and avoidable data flows. Additional information is outsourced to registers, to reduce redundancy. These registers compiled the preliminary work of historian Hermann Beyer-Thoma and were complemented with authority data records to ensure uniform identification of entities and interoperability between information resources.[18] The technical implementation was via Python, OpenRefine, and ediarum, i.e., using existing tools aimed at sustainable editing. In particular the open-source software ediarum, which is a framework for the XML-editor Oxygen, enables the TEI/XML to be edited in a barrier-free and collaborative manner.[19]

Step 4: Information Retrieval

This step is crucial for the accessibility and usability of the data, as it allows targeted access to the extensive information resources. Information retrieval focuses on finding previously created and structured data. This may include searching for specific documents or text passages, or compiling data based on certain criteria. This step therefore involves developing and applying systems and algorithms to retrieve relevant information from large datasets. In our case study, the XML files, both the travelogue and the registers, are stored in an eXist database, an open-source software that builds on XML technology. This is the basis for ediarum and is directly linked to the editing process. While working with XML documents, we primarily use the programming languages extensible stylesheet language transformations (XSLT)

---

17 "HeidelTime." Institute of Computer Science. University of Heidelberg. https://ds.ifi.uni-heidelberg.de/research/heideltime/ (accessed 20 June 2024).

18 In general, authority data describes a specific entity (such as a person, corporate body, or geographic location) in a rule-based manner to secure its identification. The Integrated Authority File of the German National Library (*Gemeinsame Normdatei*, GND) organizes personal names, subject headings, and corporate bodies applied in library, archive, and museum catalogs. Other libraries maintain their own authority files. The Virtual International Authority File (VIAF) therefore combines multiple name authority files into a single file. GeoNames is a global database containing over 11 million place names.

19 "Digitale Editionen erstellen und publizieren." ediarum. 31 May 2024. https://www.ediarum.org/ (accessed 20 June 2024).

and XQuery for compiling and retrieving information. At the same time, eXist supports XSLT and XQuery and thus facilitates access to and information retrieval from the edited files.

Step 5: Analysis

At the analysis step, the digitized and processed data are examined to provide new insights. This can encompass diverse types of data analysis, including textual, network, spatiotemporal, or statistical evaluation. Advanced analysis tools and algorithms allow large datasets to be processed and complex questions to be answered, something that could not be easily achieved using traditional humanities methods. At the same time, the availability of the resource (as text, object, etc.) allows a low-threshold switch between micro and macro perspectives and can thus be a starting point for both distant reading and close reading approaches (Weitin 2017). Overall, at this step of digital editing, the goal is to identify patterns, trends, or relationships within the data, potentially leading to new research questions that restart the life cycle. In our case, one example emerged regarding overlaps between Bronner's travel route and those of contemporaries with similar destinations. To further explore this, we need at least two editions with geolocated annotations using compatible data formats. Consequently, we initiated a dialogue with the "edition humboldt digital", a pluriannual project edited by the Berlin-Brandenburg Academy of Sciences and Humanities, to address this query (Kraft, Schnee, and Päßler 2020).[20]

Step 6: Presentation

This step involves presenting and disseminating the results by publishing digital editions, creating visualizations, or developing user interfaces for databases. Presentations can be delivered through digital platforms, web applications, and as a print publication. When presenting edited textual material from historical travelogues, spatiotemporal visualizations play a crucial role. These provide expanded access to the travelogue beyond traditional paper scholarly editions. Applying and utilizing established data formats makes it possible to visualize the content using a variety of tools. This not only delivers spatiotemporal information, but also allows content to be linked with other historical travelogues and reused in new contexts.

---

20 Fischer, Gordon, and Christian Thomas. "Humboldt auf Reisen: Chronotopische Zugänge zur edition humboldt digital." Blog vDHd 2021. 28 January 2021. https://vdhd2021.hypotheses.org/292 (accessed 21 June 2024); also see the project website of edition humboldt digital. https://edition-humboldt.de/index.xql?l=en (accessed 20 June 2024).

There are a variety of options when it comes to the presentation of digital editions, ranging from a customized website to generic open-source platforms such as TEI Publisher or ediarum.WEB. Due to their modular design, these platforms can be modified to a certain extent to meet the specific requirements of the edited material, while still maintaining a standard design framework. This feature simplifies the publication process. However, a major challenge in ensuring the longevity of digital editions is the fast pace of technological evolution along with security concerns. Ensuring that a website remains available at all times requires constant maintenance to address the numerous challenges that may arise. An alternative approach involves making carefully structured and well-documented data available independently, for example, through platforms like GitHub, which allows the control of multiple versions of a file or code. This method guarantees that the data will remain accessible and useful over time without the need for ongoing maintenance. Moreover, if the data complies with widely accepted standards, it can be easily integrated into a preferred visualization platform, thereby enhancing its ability to withstand potential technological changes.

In the context of digital editions, sustainable editing is primarily considered in terms of reusability and long-term availability (Crompton 2023). Initiatives such as Text+ in Germany, which aims at creating a solid data infrastructure for the humanities, have the explicit goal of contributing to greater sustainability of digital data. An obvious way of reducing inconsistency is the use of standards like TEI (Unsworth 2011), which is common in digital editions nowadays and should improve sustainability in the future. The six-step life cycle we presented here adds value to responsible and sustainable editing in two ways: first, by consistently implementing established standards, and second, by integrating them into a workflow that efficiently utilizes and reuses data.

The life cycle approach helps create digital resources from historical sources, providing deeper insights and new research opportunities. Each step contributes to the digital transformation of historical information, maximizing its value for research and teaching while ensuring compliance with the FAIR data principles (findable, accessible, interoperable, and reusable) that the bottom-up, stakeholder-driven, and self-governed initiative GO FAIR seeks to implement.[21]

The iterative nature of the DEHisRE project emphasizes the continuous improvement of digital editions, with transparent documentation ensuring data and pipeline reusability. Continuous refinement through case-by-case replication

---

21 "FAIR Principles." GoFair. https://www.go-fair.org/fair-principles/ (accessed 21 June 2024).

enhances the quality and usability of digital resources, providing sustainable practical guidance for innovative DH approaches along with responsible action.

Conclusion

In the field of Digital Humanities, scholarly digital editing is an ever-evolving area. As stated above, the DEHisRe project team recognized the importance of collaborating with the international research community to conduct exploratory tasks in this sphere. Given our focus, we actively engaged with ongoing projects in digital editing of travelogues to gain insights into the experimental methods used by different editing teams to tackle similar challenges. To conclude, we outline the key issues we discussed with other experts regarding the latest developments in historical travel research and the editing of travelogues.

The process of digitizing historical travel accounts through semi-automated editing processes plays a pivotal role in efficiently transforming texts into digital formats. This digitization not only facilitates the recognition and visualization of historical travel routes, but also opens up new avenues for research by digitally preparing and presenting those routes. However, the task of text annotation and semantic enrichment, particularly through NER, highlights the intricate nature of capturing and categorizing information within historical texts. Despite the complexities involved, these technologies present opportunities to thoroughly scrutinize texts and provide rich contextualization.

Using Semantic Web technologies and the TEI, innovative approaches are being developed to structure and link digital editions. These methods also involve the application of specific ontology design patterns tailored to historical travel datasets, offering solutions to the challenge of integrating digital editions into a broader research context. In addition, an intriguing concept has been developed that explores the representation of digital editions as knowledge graphs, further expanding the potential applications of these technologies.

Incorporating AI techniques into the analysis and research process highlights the potential of deep learning to examine complex datasets and uncover new insights. While it is important to approach automated analysis with a discerning mindset, it offers new possibilities for studying historical texts. Furthermore, the presentation and visualization of digital editions provide an avenue for making historical events and contexts more understandable. Overcoming visualization challenges, such as accurately representing geographical transformations, becomes pivotal in fostering a profound comprehension of historical events.

Overall, advancements in digital technology can open up new opportunities for preserving and analyzing historical texts and making them accessible to a wider

audience. At the same time, critical engagement with the limitations of these tools and methods ensures accurate historical documentation.

References

Ananieva, Anna, Sandra Balck, Hermann Beyer-Thoma, Ingo Frank, Guido Hausmann, and Jacob Möhrke. 2024. "Akademische Mobilität in Zeiten der Napoleonischen Kriege. Edition der Aufzeichnungen Franz Xaver Bronners auf seiner Reise von Aarau an die Universität in Kasan an der Wolga." *Argovia: Jahresschrift der Historischen Gesellschaft des Kantons Aargau* 136 (forthcoming).

Balck, Sandra, Hermann Beyer-Thoma, Ingo Frank, and Anna Ananieva. 2023. "Interlinking Text and Data with Semantic Annotation and Ontology Design Patterns to Analyse Historical Travelogues." *Digital Humanities Quarterly* 17 (3). Special Issue "Working on and with Categories for Text Analysis: Challenges and Findings from and for Digital Humanities Practices". http://www.digitalhumanities. org/dhq/vol/17/3/000726/000726.html (accessed 20 June 2024).

Beyer-Thoma, Hermann. 2017. "'Donauwörth - Aarau – Kazan': Die Auswanderungsentscheidung des ehemaligen bayerischen Mönchs Franz-Xaver Bronner im Jahr 1809." *Zeitschrift für bayerische Landesgeschichte* 79 (3): 689–741. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-74075-8.

Boonstra, Onno, Leen Breure, and Peter Doorn. 2004. "Past, Present and Future of Historical Information Science." *Historical Social Research* 29 (2): 4–132. https://doi.org/10.12759/hsr.29.2004.2.4-132.

Crompton, Constance. 2023. "No Boutique or Fashionable Technologies: Project Development, Mentorship, and Sustainability in an Innovation-First World." *Digital Humanities Quarterly* 17 (1). https://digitalhumanities.org/dhq/vol/17/1/000660/000660.html (accessed 20 June 2024).

Haaf, Susanne, Alexander Geyken, and Frank Wiegand. 2015. "The DTA 'Base Format': A TEI Subset for the Compilation of a Large Reference Corpus of Printed Text from Multiple Sources." *Journal of the Text Encoding Initiative* 8. https://doi.org/10.4000/jtei.1114.

Kraft, Tobias, Florian Schnee, and Ulrich Päßler, eds. 2020. "Fragmente des Sibirischen Reise-Journals 1829 von Alexander von Humboldt [= Tagebücher der Russisch–Sibirischen Reise I]." edition humboldt digital, 1–289. Berlin: Berlin–Brandenburgische Akademie der Wissenschaften. https://edition-humboldt.de/H0005449 (accessed 20 June 2024).

Meroño-Peñuela, Albert, Ashkan Ashkpour, Marieke van Erp, Kees Mandemakers, Leene Breure, Andrea Scharnhorst, Stefan Schlobach, and Frank van Harmelen. 2015. "Semantic Technologies for Historical Research: A Survey." *Semantic Web* 6 (6): 539-64. https://doi.org/10.3233/SW-140158.

Mühlberger, Guenter, Louise Seaward, Melissa Terras, Sofia Ares Oliveira, Vicente Bosch, Maximilian Bryan, Sebastian Colutto, et al. 2019. "Transforming Scholarship in the Archives through Handwritten Text Recognition: Transkribus as a Case Study." *Journal of Documentation* 75 (5): 954–76. https://doi.org/10.1108/jd-07-2018-0114.

Sahle, Patrick. 2013. "Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels. Teil 3: Textbegriffe und Recodierung." Norderstedt: BoD.

Unsworth, John. 2011. "Computational Work with Very Large Text Collections: Interoperability, Sustainability, and the TEI." *Journal of the Text Encoding Initiative* 1. https://doi.org/10.4000/jtei.215.

Weitin, Thomas. 2017. "Scalable Reading." *Zeitschrift für Literaturwissenschaft und Linguistik* 47 (1): 1–6. https://link.springer.com/article/10.1007/s41244-017-0048-4.

Bionotes

Anna Ananieva is a Senior Research Fellow at the Leibniz Institute for East and Southeast European Studies in Regensburg working on the project presented in this contribution, Digital Editions of Historical Travelogues (*Digitale Edition Historischer Reiseberichte*, DEHisRe). She has authored two monographs, several edited volumes and journal issues, and has published more than 80 articles and book chapters with a focus on transnational history, European literature of the 18th and 19th century, media history, and Digital Humanities, among others.

Sandra Balck works in Data Engineering at the Freie Universität Berlin. She studied Information Science in Potsdam and was part of the DEHisRe project as a Research Fellow between October 2021 and May 2024.

Jacob Möhrke is a Research Fellow at the Leibniz Institute for East and Southeast European Studies in Regensburg, Germany. He studied Digital Humanities at the University of Regensburg and is currently also working with DEHisRe.