ESTS Conference 2024

Budapest

October 2-4, 2024

Abstract booklet

Content

ELLIE BLEEKER; PETER BOOT; LEO JANSEN; BEATRICE NAVA; MARIKEN TEEUWEN

The eDITem model: towards future-proof digital editions

Forty years ago, Hans Zeller published his Baukasten model, which deconstructs scholarly editions into different modular components (Bausteine; 1984, 321). The same principle – to divide a complex entity into smaller separate components – has been applied to digital editions (e.g., Pierazzo 2015, 117; Andrews 2013).

Within the eDITem project, we build upon this modularity principle and propose several edition components which, together, form an edition model expressed in TEI ODD. The eDITem model allows us, in collaboration with software developers, to build both an editorial workflow and publishing infrastructure that facilitates software reuse, diminishes development and maintenance costs, and decreases the need for IT support.

This paper describes the architecture and implementation of the eDITem model, focusing on the formal criteria we adopt to design our edition components. It also discusses how the eDITem model allows us to address several well-known issues regarding the interoperability and long-term accessibility of digital editions.

October 2. 9:45 – 10:15 Session A

WIM VAN MIERLO

Loughborough University, Loughborough

Intelligent Editing: The New Documentary Edition

For decades now, the trajectory of the Digital Scholarly Edition (DSE) has moved along technological innovation. But have our DSEs become smarter? From the editor's point of view, we can certainly answer this question affirmatively. But is that the case also from the user's point of view? To be clear, my aim in this paper is not to dredge up the arguments about the lack of a standard format that the print edition offered and which the DSE so far has not. However, I do want to put forward an argument about the need for creating space to embed existing technological innovation with a view to improving user experience. The area where this is already happening is in the new documentary edition, a new form of documentary which favours networks over the documents of individuals and which uses 'thick description' to situate these documents within their historical and cultural context. By way of demonstration, I will consider a number of case studies to analysing their affordances and their 'smart' use of technology. These cases will include the digital version of The Letters of Charles Harpus and his Circle, edited by Paul Eggert and Chris Vening (2023) and one or two 'micro editions' from the Association for Documentary Editing's journal Scholarly Editing (scholarlyediting.org).

October 2. 10:15 – 10:45 Session A

Katrine Frøkjær Baunvig

Jon Tafdrup

Kirsten Vad

Aarhaus University, Aarhaus

Automating Explanatory Commentary: AI-Driven Approaches in Digital Scholarly Editions

The meso-scaled project Grundtvig's Works (2009-2029) is the largest humanist project in terms of timeframe and funding on Danish ground so far. We believe this obligates us to make all our work freely available in the public domain for any kind of rendition.

We will present a strategy for a sustainable work loop for digital scholarly editions, in which editions are disseminated on multiple interfaces (i.e. internet, digital textual archives, e-books, and even simple PDFs) to ensure as wide access as possible for readers, users, and scholars of all kinds. We envision a firmer tripartition of labour on a regional level that clearly divides responsibility for a) production of data, b) short-term rendition, and c) long-term storage of data in FAIR manner. This work loop ensures great openness of data, but also closure for editors of digital scholarly editions at the end of projects leaving storage and long-term rendition to others.

October 2. 9:45 – 10:15 Session B.

FLOOR BUSCHENHENKE

LAMYK BEKIUS

Track Changes: studying born-digital writing

Contemporary writers largely work in a digital environment, yet the implications of this practice on textual scholarship is understudied. The digital self-archiving practices of authors vary widely and do not default to creating versions. (Buschenhenke, 2016) In the project Track Changes we have collected experimental data from eleven Dutch and Flemish authors. Using keystroke logging (Leijten & Van Waes, 2013), they recorded their digital work process while writing short stories. This corpus has allowed us to design a new visualisation using XML-TEI, which showcases the textual histories and allows for an analysis of the writers' actions through time. (Bekius 2021; 2023) In this paper we will discuss one case study that makes use of this edition to classify the revisions made by three of the writers. Much of textual dynamics was found to take place during the composition of new sentences and on text that was written earlier in the same writing session. This indicates that compared to archival documents, the keystroke files reveal much more of the writing, and thus both forensic and experimental methods are crucial in securing the evidence of writing in the 21st century. (Varasi, 2019; Ries, 2021)

October 2. 10:15 – 10:45 Session B.

Krista Stinne Greve Rasmussen

Emilie Karner Hansen

Katrine Frøkjær Baunvig

Jon Tafdrup

Aarhaus University, Aarhaus

Modeling the Literary Archive: From Manuscripts to a Database (Using AI)

The meso-scaled project Grundtvig's Works (2009-2029) is the largest humanist project in terms of timeframe and funding on Danish ground so far. We believe this obligates us to make all our work freely available in the public domain for any kind of rendition.

We will present a strategy for a sustainable work loop for digital scholarly editions, in which editions are disseminated on multiple interfaces (i.e. internet, digital textual archives, e-books, and even simple PDFs) to ensure as wide access as possible for readers, users, and scholars of all kinds. We envision a firmer tripartition of labour on a regional level that clearly divides responsibility for a) production of data, b) short-term rendition, and c) long-term storage of data in FAIR manner. This work loop ensures great openness of data, but also closure for editors of digital scholarly editions at the end of projects leaving storage and long-term rendition to others.

October 2. 14:30 – 15:00 Session A.

ALBRECHT HOLFHEINZ

Unlocking Arabic Manuscripts: Imperfect HTR as a Stepping Stone for AI-powered Analysis

This paper explores the potential of utilizing "imperfect" Handwritten Text Recognition (HTR) as a springboard for AI-powered analysis of Arabic manuscripts.

Achieving perfect HTR remains challenging for complex historical documents like Arabic manuscripts. Taking a philological-historical lens rather than a purely engineering one, the paper argues that imperfect HTR output, despite containing errors, can be leveraged for AI-powered content analysis and summarization, thus facilitating the identification of priority materials for further human study. By reducing the time and effort required to identify material that warrants more detailed examination, this approach can free resources for in-depth textual scholarship. To validate this hypothesis, documents previously studied by the author through conventional approaches will be subjected to AI-assisted analysis based on imperfect HTR, comparing the effectiveness of this method with the 'ground truth' of traditional scholarly examination.

The methodology involves a case study focusing on samples from the 'Hofheinz Collection' (University of Bergen), a unique corpus of over 600 manuscripts from the Sudan (18th-20th c.) photographed by the author. Comparative data from Yemeni manuscripts (19th-20th centuries) held at the Italian Foreign Ministry and previously catalogued by the author will also be considered. The study will (1) assess the relative accuracy of using HTR tools (such as Transkribus, OCR4all, OCR-D, and daiR) in an only partially supervised setting for transcribing these documents; and (2) investigate how well AI-assisted analysis can detect noteworthy passages within the resultant data.

The study serves as a proof of concept aiming to expedite the exploration of large Arabic manuscript corpora. This is particularly relevant given the constraints of limited human resources and the time-consuming nature of traditional HTR pre-training.

The author has extensive experience working with Arabic manuscripts from the Sudan, Yemen, and Mali. His expertise includes collecting, transcribing, and analyzing Sudanese and Yemeni manuscripts; overseeing digitization and cataloguing efforts within the Timbuktu Manuscripts

Project; and designing and delivering university courses focused on the identification of Arabic manuscripts from the National Library of Norway and from a private Norwegian collection.

October 2. 15:00 – 15:30 Session A.

EMANUELLE LEBOFFE

RAMÓN VALDÉS GÁZQUEZ

THE DIGITAL CRITICAL EDITION OF *EL CASTIGO SIN VENGANZA*, A CANONICAL PLAY WITH AN AUTOGRAPH MANUSCRIPT AND A COMPLEX TRADITION

The digital critical edition of certain textual traditions presents serious challenges. The polymetric play *El castigo sin venganza*, by the 17th-century Spanish playwright Lope de Vega, has a complex textual tradition. Its autograph manuscript contains *pentimenti*, authorial variant readings, and drafting phases that are related to different witnesses and moments of its tradition (some manuscripts and prints derive from the earliest phases of the autograph, while others derive from the later ones), including, for added complication, different versions of the ending and interventions by censorship. The very construction of the text and the critical apparatus in its genetic and evolutionary levels, and its combination in a digital edition with different images (hyperspectral photography, manuscript and printed witnesses), as well as its own encoding and visualization, are very complex. Achieving good encoding and an intelligible visualization (ideally for a broad audience, as it is a canonical work) presents a true challenge.

October 2. 16:00 – 16:30 Session A.

TERESA FILIPE

Center of Linguistics of the University of Lisbon (CLUL)

What is a meaningful textual link? Marginalia and exogenesis.

Fernando Pessoa (1888-1935) is the most celebrated author of the Portuguese 20th century literature. The writer's estate is currently held at two different locations: Fernando Pessoa's House (where circa 1,300 volumes of the author's private library are kept) and the Portuguese National Library, Estate 3 (BNP, E3) with over 30,000 documents including books and parts of books (loose pages, covers, dustjackets), reading diaries and other reading notes. Since its full digitization in 2010, Pessoa's private library shows the author's avid activity as a reader and a marginalist. Books and marginalia hold a strong material and textual nexus with the author's estate at BNP. In this talk, I will address some examples of these nexuses and enquiry about ways in which they can be edited, arguing for the relevance of these diverse material to exogenesis.

October 2. 16:30 – 17:00 Session A.

DIRK VAN HULLE

VINCENT NEYT

Sustaining Digital Editions: The Case of the Beckett Digital Manuscript Project

This paper considers the sustainability of the scholarly edition after active development and maintenance has ended. At such a moment of closure, most digital editions are faced with the problem of how they will remain accessible in, say, 50 years' time. The main research data (TEI/XML transcriptions and facsimiles) will of course be stored in research data management services, but we believe the interface is a crucial aspect that should be preserved as well. The user experience and functionalities of the interface are important features of the edition, guiding the user through the material and establishing valuable links between the documents. We are therefore exploring how to keep the edition's interface available for the long term by minimizing server software requirements and security issues. Our case study is the *Beckett Digital Manuscript Project* (www.beckettarchive.org).

October 2. 17:00 – 17:30 Session A.

KATRIN HENZEL

Kiel University Library

The implied user on the empirical test bench

Publicly funded digital editions should be made accessible to a broad public in order to meet the requirements of open science. However, the reality often looks different, which raises the question of who digital editions are made for. What does the ideal user look like and how do they relate to actual user groups? What barriers prevent people from using specific features of digital editions? In my paper, I will focus on methods of user data collection as well as on methods for active user involvement. I will show how such methods can be integrated into projects and what advantages a mixed methods design offers and where its limitations lie. My hypothesis is that scholarly editing projects should take their users into account as early as possible and actively involve them in order to ensure that the edition meets their *actual requirements*.

October 17:30 – 18:00 Session A.

GIUSEPPE FERRARA

Ianus. A Digital Assistant of the Classical Philologist

In ancient Greek and Latin texts, there are frequently passages whose meaning appears corrupted. In this paper, I introduce Ianus, a tool based on a deterministic algorithm, created to help the classical philologist in identifying conjectures for words containing errors resulting from the misinterpretation of letters in majuscule and minuscule scripts and the evolution of language pronunciation. The efficacy of Ianus in generating conjectures compared to classical philologists was evaluated in a single-centre, controlled, non-randomised study. The software was able to track down the expected conjecture in 100% (95% CI, 95%–100%) of cases, versus 25% (95% CI, 17%–35%) of philologists, and to propose it as the first conjecture in 66% (95% CI, 55%–76%) of cases, versus 15% (95% CI, 9%–24%) of philologists. I am currently working on expanding the pool of errors corrected by the tool and employing AI-based approaches to make the software context-sensitive.

October 2. 18:00 – 18:30 Session A.

PAULINA CHORZEWSKA-RUBIK

Source code philology - born-digital poetry and authorial

textual variants

The purpose of this paper is to reflect on the methodology of utilizing the source code of websites as a resource for philology, particularly in the context of establishing variants in verse foreditions of born-digital poems. Through a case study focused on the poetic works of Polish writer Tomasz Pułka, I will demonstrate instances where a critical analysis of website source code, coupled with platform studies, becomes essential for identifying authorial variants, and uncovering traces of the creative process. I would also inquire whether the source code of a poetic publication page, a dynamic document, should be viewed as a source of the author's variants or as an updated digital edition of the poem. I will focus on analyzing intricate verbo-visual compositions where the poet directly shapes the typographic elements of their texts using web-based word processing software.

October 3. 10:15 – 10:45 Session A.

NOÉMI PRÓTÁR

RÓBERT KOPCSÁK

PÉTER HORVÁTH

Approaches to the analysis of the Corpus of Hungarian Lyrical Poetry's Song Lyrics subcorpus

This presentation highlights our analysis of the Corpus of Hungarian Lyrical Poetry's Song Lyrics subcorpus, combining qualitative and quantitative methods. The Corpus of Hungarian Lyrical Poetry is a project of the Research Group on Stylistics of Eötvös Loránd University: its Song Lyrics subcorpus features the most played Hungarian songs from 2014 to 2022, consisting of 144 popular lyrics sampled from radio and streaming charts.

Our analysis differentiates the song lyrics by their platform appearance examining language use; and the stylistic and poetic organization of the song lyrics, to determine whether there is a difference between the songs of the two platforms. In the quantitative phase, we conduct vocabulary analyses and text similarity analyses, to see whether the lyrics are distinguishable into groups and whether these groups correspond to the groups of the two platforms. In the qualitative phase, we perform manual style analyses and finally compare the results of both methodologies.

October 3. 10:45 -11:15 Session A.

RÉKA SÓLYOM

Neologisms of Hungarian slam poetry: a corpus-based analysis

KRE BTK, Department of Hungarian Linguistics

Employing a functional cognitive framework, this talk focuses on the neologisms, which can often be considered as hapaxes, in the Slam Poetry Subcorpus of the Research Group of Stylistics, which contains 108 Hungarian slam poetry texts. The hypothesis of the analysis is that in the corpus various types of neologisms can be witnessed, based on their grammatical and semantic structure, e.g., celofóbia, cigarettenet, rehabzó. The aim of the talk is morefold: 1) to analyse the grammatical structure of the neologisms, based on which different subgroups can be distinguished; 2) to analyse the phonological features of the neologisms; 3) to reflect to their semantic features (e.g., metaphorical, metonymical structures); 4) to analyse with AI whether these words and expressions have spread in today's Hungarian language use, or whether they occur only in the subcorpus, and can be considered as hapaxes.

October 3. 11:15 – 11:45 Session A.

PÉTER HORVÁTH

Eötvös Loránd University

Syntactic repetition in canonical Hungarian poetry

Phonological repetition through meter, rhyme, and alliteration is a key feature of poetry. The paper investigates whether repetition can be detected at another level of Hungarian canonical poetry, the syntactic level of poems. The research corpus consists of all the poems of 24 Hungarian canonical authors from the 19th and the first half of the 20th century. The grammatically and metrically annotated poems were retrieved from the ELTE Poetry Corpus (Horváth et al. 2022). To measure the degree of repetition of syntactic structures, a modified version of the moving-average type-token ratio is used (Covington– McFall 2010). This modified method measures the number of different 2-gram, 3-gram, and 4-gram parts of speech in different window sizes. The calculated repetition scores are compared with the values obtained for the ELTE Novel Corpus (Bajzát et al. 2021) used as the reference corpus. The paper also investigates whether there is a correlation between syntactic repetition and the phonological properties of the poems. On the one hand, it presents that the meter of the poems influences the degree of syntactic repetition. On the other hand, by calculating the Jaccard similarity coefficient between the part-of-speech-ngrams of lines, it also investigates whether consecutive rhyming lines are more similar syntactically than consecutive non-rhyming lines.

October 3. 11:45 – 12:15 Session A.

QIANGI HUANG

VANESSA Klomfaß

JULIA NANTKE

FRANK STEINICKE

University of Hamburg, Hamburg

Multimodal modeling in the age of born-digitals: Walter Kempowski's *Ortslinien*

In the project 'Multimodal Modeling of Cultural Artifacts in Digital Space' we apply methods of data restoration, visual computing, natural language processing, machine learning, deep learning (LLMs) and computer vision to explore the born-digital work fragment Ortslinien left behind by the contemporary German author Walter Kempowski (1929–2007) on his Apple Power Mac G3. In our talk we will present our first approaches for automated structuring and corpus analysis. We will especially focus on making use of generative AI and machine learning techniques for data extraction and categorization as well as content exploration and representation.

As *Ortslinien* poses a variety of challenges (from making the original files legible to dealing with the constitutive multimodality and the incompleteness of the work fragment), we assume that the work is an ideal use case for developing concepts for exploring multimodal born-digital cultural material and representing them in a digital scholarly edition.

October 3. 10:15 – 10:45 Session B.

EWELINA DUBICKA

The John Paul II Catholic University of Lublin

Exploring Research Opportunities in Digital Editions: Genetic Criticism and Tools

With the increasing number of digital editions, new methods of studying them emerge, such as linguistic research conducted using tools from the CLARIN consortium or semantic studies. For researchers of textual genesis, the abundance of easily accessible versions undoubtedly facilitates their work. However, the question remains: how can the documents from these editions be utilized for further research?

The purpose of this presentation is to demonstrate the possibilities of conducting research using genetic criticism methodologies on existing digital editions. Current studies conducted on editions produced by Polish research institutions will be discussed, with particular attention to projects related to 19th and 20th-century literature. The process of designing editions and associated challenges will also be addressed, along with the tools that genetic critics can employ for their ongoing work.

October 3 10:45 – 11:15 Session B.

THORSTEN RIES

University of Texas, Austin

Digital Scholarly Editing and AI as Paradigm Shift of Philology

The potential of "intelligent editions", where AI would play a semi-autonomous or even autonomous role either in the production process or the performance operation of the edition, would mean a major shift of scholarly editing paradigms and the role of the scholarly editor in a context of non-reproducible and non-determinative AI contexts. The reasoning is that while textual scholarship and digital textual scholarship previously focused on minute accuracy, detailed but static encoding (TEI P5 XML, Minimal Computing Editions), comprehensive textual representation and commentary, newer approaches such as WittgensteinOnline try to harness AI for several tasks - at the moment mostly auxiliary and semantic, or as part of HTR tasks (e.g. Dehmel Digital).

But the shift of the role of the scholarly editor actually goes much deeper, as we go forward. For instance, fine-tuned AI may be much better at aligning texts and variants than CollateX, HTR may be very good at creating at least base transcriptions of literary manuscripts at unforeseen speed, and the human written commentary may be completely superseded by a fine-tunable LLM trained on all previous commentary and recent research literature, queriable on demand in a digital edition.

The paradigm shift here may be: the scholarly editor will still do a part of the philological work and resolve the more "interesting" problems, but instead of spending a lifetime on writing the commentary, the job may shift towards quality control of the training input for the local LLM AI and its output for a dynamic system. Of course, on the flipside of the coin, this dynamic AI system is not determinative, its output not reproducible, and its future depends on the stability of the AI models involved. I am looking at this from the angle of Reproducibility and Explainability, which I recently co-edited a special issue of IJDH on.

I would like to make the point that on the one hand, this is an opportunity to challenge our own authoritative bias of the indisputable role of the philological commentary, which could be replaced by a system that is able to take a different theoretical view on the training data upon user request, but on the other, on other levels of the scholarly edition, it messes with the standards of the philological paradigm, as the "Beta" output of the Dehmel Digital edition

shows, where philologically corrected and on the next page hilariously wrong Transkribus transcription appear next to one another in the online edition (see below in PDF).

Within this paradigm shift, further practical and ethical angles play a role: everything that we train into an AI model can then be reused, and we don't know what exactly it learned. This creates, for instance, issues with using AI on archive material that may be protected or confidential - using AI to analyze an author's storage media is virtually impossible by ethical and data protection standards, and would at least involve the complete destruction of the entirely isolated model after single use so the training data cannot leak or be queried by third parties. Mutatis mutandis, these concerns may also apply to metadata and even paper archive material at scale.

October 3. 11:15 – 11:45 Session B.

KIYOKO MYOJO

YASUHIRO SAKAMOTO

How Generative AI will revolutionize dynamic editing: Towards an intelligent edition

Two broad approaches can be taken for editing Kafka's writings: static and dynamic. A static approach attempts to produce a definitive work. A dynamic approach attempts to show the writing process and the incomplete nature of the text.

We are editing "The Trial" to illustrate this dynamic edition. The basic concept, which Myojo presented at the ESTS conference 2022, is to break the traditional framework of chapters and works, create new segments, and then try to arrange these parts in the order of writing. We intend to utilize Deep Learning and Generative AI for text mining and document image analysis to significantly increase the accuracy of the writing chronology. This paper describes how to utilize AI in such dynamic editing and shows how the intelligent edition would be generated.

October 3. 14:30 – 15:00 Session A.

DAVID WRISLEY & HARTLEY ROXANNE MILLER

From HTR to GPT: Repositioning Expert Judgment in Automating Medieval Documentary Editions

Our project is editing a short old French poem known as the Lai de l'ombre (c. 1000 lines). We have used handwritten text recognition (HTR) to automate a first draft of a diplomatic transcription of the poem, but the formal constraints of the automated system have led us to question the norms of transcription, and (silent) normalization prevalent in our field. We have found that automation is not a one time process, but rather a series of critical interactions with documents, which requires us to reconsider the role of the expert in the workflows designed to establish text. In our presentation we will discuss some of the ways we are using GPT4 and varied normalization criteria to create multiple versions of the Lai as open-licensed Open Educational Resources, one of which is a pedagogically-focused, facing-page university edition of the Lai for its Romanian translation (to be published at Polirom).

October 3 15:00 -15:30 Session A.

CLAUDIA ESCH

PIA HOFMANN

YANNIK HERBST

JANA KLINGER

Digital Scholarly Editions in the Making. Reflecting on Workflows

The main challenges of digital scholarly editing seem to contradict themselves: On the one hand, there is a need for standardization of workflows, data models, tools and interfaces, while on the other hand, the complexity and individuality of projects prevent one-fits-all solutions. We will reflect on the challenges of balancing these two goals in our work within different digital scholarly editing projects. Our presentation will not focus on one concrete workflow or pipeline. Instead, we will concentrate on the impact and challenges of generic workflows and their dependency on various **general structural conditions that foster or impede processes of digital scholarly editing**, such as the technical infrastructure, composition of project teams and division of labour between experts from different fields.

October 3. 14:30 – 15:00 Session B.

HANA KOSÁKOVÁ

Actors involved in the work process. The issue offinalizing a literary work and its path to publication.

Department of East European Studies, Faculty of Arts, Charles University

The paper focuses on the issues of various actors (readers, assessors, critics) who - apart from the author - participate in the finalization of a literary work. The issue will be examined on the material of the novel A Czech Dreambook (Český snář), written by Ludvík Vaculík and first published in samizdat in 1981. The research is based on archival materials.

October 3. 15:00 – 15:30 Session B.

PÉTER KASZA

A Window into a Historian's Workshop: The Several TimesRewritten Work of Wolfgang Lazius

Wolfgang Lazius (1514-1565) was one of the outstanding figures of Viennese humanism, and was court historian to Ferdinand I from the mid-1540s until his death. In this capacity, he left behind a vast oeuvre, much of which is still in manuscript. Lazius was a highly conceptual historian, the first to realise that the history of the previously separate areas (Austria, Bohemia, Hungary, etc.) of the emerging Habsburg Monarchy in Central Europe required a unified and common discussion. His wide-ranging survey, the only partially completed Rerum Austriacarum Decades, was born out of this realisation and concept. But how was an early modern work of history produced and written? On the basis of what sources, according to what methods, within what framework?

In the case of Lazius, we are in the rare, almost exceptional situation of having several manuscript versions of his work, or parts of it, so that we can examine not only what Lazius finally wrote, but also the steps he took in doing so, and what he changed in the course of the revisions, and why. Through a few examples, the lecture will seek to give an insight into the workshop of a very conscious historian with a European outlook, examining Lazius' sources and the reasons that led him to repeatedly rework the content, style and structure of Rerum Austriacarum Decades.

This fact, however, also poses special challenges for the editor. If some of Lazius' manuscripts are more than mere versions, if they were written according to a different concept, then a critical edition would be making a mistake if it were to cover up the traces of each phase. More recently, genetic textual criticism offers a theoretical background to how and why we should work with multiple versions of a work, and digital editions provide a reasonable visualisation of this. In the foregoing, I will review the features of Lazius' work that can be presented in their fullness in the context of a digital edition.

October 3. 16:00 – 16:30 Session A.

DR. CHRISTIAN THOMAS

Berlin-Brandenburg Academy of Sciences and Humanities

Since Intelligent Editions are digital, the future for print is derivative:

The digital paradigm in the context of single source, hybrid editions.

I will address the overarching question, "What is an Intelligent Edition?" by combining it with another question from the CfP: "Is there a future for print?". I argue that there is a future for print, but only as one (of several) derivative format(s) in the context of a single source, hybrid publishing strategy—i.e. only if the print edition is derived directly from TEI-XMLencoded documents of an intelligent Digital Scholarly Edition. In the digital paradigm, we overcome concessions to the two-dimensional print paradigm that resulted in oversimplified re-/presentations of textual witnesses, and focus on the encoding of textual, structural and semantic information. To secure a future for print in this context, publishing houses and their vendors need to embrace the TEI-XML-based single-source strategy and handle the transformation more efficiently. I will illustrate current challenges within this ongoing process with examples from two German long-term edition projects on A. v. Humboldt's and Goethe's Biographica.[1]

October 3. 16:30 – 17:00 Session A.

---

[1] Cf. "Travelling Humboldt – Science on the Move" and its edition Humboldt digital portal, https://editionhumboldt.de/?&l=en and "PROPYLÄEN: Goethe's Biographica", https://goethe-biographica.de/ . Further information on both projects can be found in the projects database of the Union of the German Academies of Sciences and Humanities' Academies Programme at https://www.akademienunion.de/en/research/projectdatabase/travelling-humboldt-science-on-the-move and ../forschungsplattform-zu-goethes-biographica.

ZSÓFIA FELLEGI

Transitioning from Analogue to Born-Digital: MethodologicalShifts in Publishing Medieval and Early Modern CentralEuropean Texts

HUN-REN RCH Institute for Literary Studies

The Bibliotheca Scriptorum Medii Recentisque Aevorum was the critical edition series of the [HAS] HUN-REN RCH Institute for Literary Studies (BSMRAe), publishing the Latin works of unpublished or hard-to-access medieval and early modern Central European authors (mainly Hungarian or authors active in Hungary). The series was founded by László Juhász in 1930, with 39 volumes published under his direction until 1946. The series was revived in 1976 under the title Series Nova. In 2020, the series editors began collaborating with the DigiPhil group at the HUN-REN RCH Institute for Literary Studies with the goal of producing the critical editions in the series in a born-digital format. The first critical edition published in this way is the critical edition of Nicolaus Olahus's correspondence, edited by Emőke Rita Szilágyi: Nicolaus Olahus correspondence (https://digiphil.hu/gallery/olahus/).
In my presentation, I will showcase the challenges and experiences of recent years through concrete examples, and examine the potential opportunities that Artificial Intelligence tools can offer to support philological work.

October 3. 17:00 -17:30 Session A.

MATEUSZ ANTONIUK

Jagellonian University, Krakow

A Man versus the Monster of a Machine. AI, poetry, and genetic criticism

In 1997, a chess match took place in New York, in which world champion Garri Kasparov faced off against a Deep Blue computer designed by IBM. The match was covered by the media, which wove their net around the "global village". In Warsaw, the Polish poet Zbigniew Herbert (1924-1997) read press reports about the chess match and its result (victory for the computer). An insight into the Herbert's archives allows to trace how press reading passed smoothly into the process of composing the poem, eventually published as "Szachy" ['Chess']. In this poem, Herbert interpreted the chess game as a defeat of the human species against artificial intelligence.

In the paper, I intend to combine three topics. Firstly, I will employ genetic criticism's methodology (analysis of the documents of exo and endo-genesis); secondly, I'll argue that the genetic and electronic edition of Herbert's poem unveils a "story" that is interesting in terms of textual scholarship, interpretation and cultural studies; thirdly, I'll show how the modernist imaginary of the 'revolt of the machines' circulates between literature and media discourse.

October 16:00 -16:30 Session B.

EMESE K. MOLNÁR

GÁBOR SIMON

Eötvös Loránd University, Budapest

Is person marking a significant feature of lyrical discourse? A keyness analysis

According to the main objective of the ongoing corpus-based research in the Stylistic Research Group, "person marking constructions constitute characteristic genre-specific patterns in lyrical discourses". The paper investigates the empirical relevance of this statement. Relying on the preliminary version of the ELTE Poetry Corpus, a quantitative analysis is carried out in two different ways to assess the significance of person marking lexicogrammatical patterns in Hungarian poetry. With traditional keyPoS-tag analysis in LancsBox X we want to find out whether these patterns have a keyness status within the Poetry Corpus comparing its subcorpora to each other and using the ELTE Novel Corpus as a reference corpus. We also adopt the two-dimensional keyness analysis to measure the more exact keyness value of person marking constructions in the subcorpora of the Poetry Corpus.

October 3. 16:30 – 17:00 Session B.

GABOR TOTH

Reinventing the Scholarly Monograph in Digital Age

University of Luxembourg

In 2022, De Gruyter Verlag, and the Center for Contemporary and Digital History of the University of Luxembourg formed a partnership to develop a new digitally-born scholarly monograph and edition series. The goal of the new series is to offer a sustainable publishing platform for scholars engaging with digitally born materials, such as for instance webcontent, datasets, and interactive visualizations. The goal of my presentation is to demonstrate the first prototype of the publishing platform, and to discuss the challenges and open questions the development has involved. First, I will discuss the base technology, the Jupyther notebook, underlying the digital platform that C2DH and De Gruyter Verlag jointly develop. Second, I will address the use of research code in a digital monograph. Finally, I will discuss the challenges that the direct inclusion of datasets into a scholarly monograph involves.

October 4. 9:45 – 10:15 Session A.

Krista Stinne Greve Rasmussen

Emilie Karner Hansen

Katrine Frøkjær Baunvig

Jon Tafdrup

Kirsten Vad

Aarhaus University, Aarhaus

Let it go! Setting data and editors free in relation to rendition and long-term storage

The meso-scaled project Grundtvig's Works (2009-2029) is the largest humanist project in terms of timeframe and funding on Danish ground so far. We believe this obligates us to make all our work freely available in the public domain for any kind of rendition.

We will present a strategy for a sustainable work loop for digital scholarly editions, in which editions are disseminated on multiple interfaces (i.e. internet, digital textual archives, e-books, and even simple PDFs) to ensure as wide access as possible for readers, users, and scholars of all kinds. We envision a firmer tripartition of labour on a regional level that clearly divides responsibility for a) production of data, b) short-term rendition, and c) long-term storage of data in FAIR manner. This work loop ensures great openness of data, but also closure for editors of digital scholarly editions at the end of projects leaving storage and long-term rendition to others.

October 4. 10:15 – 10:45 Session A.

ELSA PEREIRA

University of Lisbon – CLUL

Complete Works Editions on Print and Digital Displays

Until recently, the established convention for Complete Works Editions coincided with a printed single text based on the last authorial will. However, the "shift towards textual instability" (Nash, 2003: 12) has led scholarly editors to engage with digital media, enabling readers not only to "navigate the […] complete oeuvre across versions" (Van Hulle, 2022: 225) but to disrupt the rigid "work-by-work approach" (Van Hulle, 2021: 58), conditioned by the materiality of the print medium.

This presentation will survey the main editorial approaches to Complete Works Editions in the analogue and digital paradigms, from arranging canonical reading texts in book form to digital collation and recent *intelligent* undertakings that rely on semantic web technologies to map "creative concurrence" (Van Hulle, 2021) and "genetic networks", revealing "different types of relationships between the texts, intertwining and responding to each other" (Christen & Spadini, 2019: 99).

October 4. 10:45 – 11:15 Session A

SAKARI KATAJAMÄKI

TEI Files as a Starting Point for Dialogue Research

Finnish Literature Society – SKS

About half of Aleksis Kivi's novel Seven Brothers (1870) consists of dialogues written in the manner of dramatic dialogue. The dialogues have puzzled scholars since the novel's publication. According to an early review, Kivi's characters are 'shadows who all speak and act in the same way'. Now that the novel has been transcribed in TEI format for a critical edition, it will be possible to analyse the real differences between the brothers' speech and to examine the dynamics in their dialogues.

In addition to the distribution of the speakers' lines, the paper discusses the order of the speakers in the dialogues and analyses the distribution of personal names and pronouns, which indicate the social dynamics of the brothers. The paper will illuminate these results selectively. The research team behind the paper consists of Niklas Alén, Sakari Katajamäki, and Ossi Kokko †.

October 4. 9:45 – 10:15 Session B.

ADELLE HAY

Digitizing a 'mutilated text'; authorial intent and Anne Brontë

Loughborough University

Since its publication in 1848, Anne Brontë's The Tenant of Wildfell Hall has existed as multiple, very different, editions. Tenant has been working its way back into the literary canon in the UK and US since the 1970s, but many available editions are still based on a heavily edited, 'mutilated' text from 1854. Who has been editing Tenant, why is it so difficult to identify an authoritative edition, and how aware is the reader of its complex textual history? This paper will explore the challenges faced by editors of Tenant looking to produce an edition based on authorial intent, and build a case for the digitization of this long-misunderstood Victorian novel.

October 4. 10:15 – 10:45 Session B.