

The Epistemological Value of the Computational Turn in Scholarly Editing

Cugliana, Elisa

elisa.cugliana@uni-koeln.de
Universität zu Köln, Deutschland
ORCID: 0000-0002-6460-2954

Kuczera, Andreas

andreas.kuczera@mni.thm.de
Technische Hochschule Mittelhessen, Gießen
Deutschland
ORCID: 0000-0003-1020-507X

Grüntgens, Max

gruentgens@fotomarburburg.de
Deutsches Dokumentationszentrum für Kunstgeschichte,
Bildarchiv Foto Marburg, Deutschland

Ward, Aengus

a.m.m.ward@bham.ac.uk
University of Birmingham, Großbritannien
ORCID: 0000-0001-9526-0718

van Zundert, Joris J.

joris.van.zundert@huygens.knaw.nl
Huygens Institute, Amsterdam, Niederlande
ORCID: 0000-0003-3862-7602

Andrews, Tara

tara.andrews@univie.ac.at
Universität Wien, Österreich
ORCID: 0000-0001-6930-3470

Dekker, Ronald Haentjens

ronald.dekker@di.huc.knaw.nl
DHLab / Huygens Institute, Amsterdam, Niederlande
ORCID: 0000-0001-6737-7986

Summary

In the last decade a decidedly computational turn in scholarly editing is noticeable. Texts are not just encoded anymore. To store and execute scholarly decisions increasingly also program code is applied, which thereby becomes a computational expression of scholarly theory. Theorists and practitioners of such computational approaches have

not just put forward scalability and replicability as motivations for computational methods, but also the promise of increased analytical power and on the horizon a changed epistemology. This panel invites some of the most visible proponents of these approaches to reflect on these promises and especially to try to deepen our understanding of the epistemological value that is tied to applying code in the domain of scholarly textual editing.

Introduction

How is the computational turn in scholarly editing fundamentally altering our understanding of the texts that we edit? There is no question that the computational turn has happened over the last decades (cf. Cugliana and Van Zundert 2022; Ward 2022; Van Hulle 2021; Andrews 2019; Buschenhenke 2019; Haentjens, Bleeker and Buitendijk 2018; Ries 2017; Barabucci and Fischer 2017; Hadersbeck et al. 2015; Oldman, Doerr, and Gradmann 2015; Rosenthal 2015; Buzzetti 2002). This panel poses, and tries to address, the question of the long-term implications of the nature of what we are now doing as digital textual scholars. Digital approaches from the recent decades in scholarly editing have produced solid groundwork in theoretical principles and concepts, practical methods and standards, as well as functional tools and processes. Although there is something of a broad consensus emerging in these areas, many of the discussions about the future of digital and computational¹ critical editing focus on the development of specific tools and methods stemming directly from this consensus and not on the long-term implications for the nature of what we do as textual scholars.

THINK, as a research group, looks at the currently dispersed and technically oriented discourse from a wider angle and an open stance. THINK seeks to highlight without prejudice the extent of the possibilities of computational textual analysis that are currently emerging, and by foregrounding to what degree these may, or already do, represent an epistemological shift in the nature of the textual knowledge and understanding we seek by scholarly editing. The discourse initiated by THINK thus questions whether the computationally created editions of the future will merely exist as a quantitatively upscaled means of the scholarly practice that exists today. Or if such computational tools will rather allow us to analyse and represent the knowledge we edit as well as the knowledge we thus produce in some dynamic reciprocal fashion. Or even if their application will constitute a fundamental reorientation of the very object of study itself, making it possible to engage with the object of study in a methodically controlled dialectic that considers both the analog and digital mode of being of the object. The panel we propose is structured in three sub-themes that serve to explore these meta-questions pertinent to the future of computational scholarly editing.

Contents of the Panel

New frontiers: Where does the current methodological frontier of digital and computational textual scholarship lie?

In the past, textual scholars have paid much attention to limits, boundaries, and differences between the book and the digital text. In many respects the frontier of digital philological methodology has moved past these questions of materiality (Kirschenbaum 2008) and feasibility (e.g. Barabucci and Fischer 2017:47-49), not implying that these questions are unimportant, but rather pointing out that the leading edge of methodological thinking seems to have progressed into an area of questioning how scholarly process, observation, interpretation, and even reasoning already is or might be calculated upon algorithmically (Cugliana and Van Zundert 2022). The assumption of these new methodologies seems to be that they also improve in some ways not just our analytics, but also our epistemological thinking, and even our knowledge. This bears new emerging challenges concerning the ways this new kind of thinking should be transmitted to its users and audiences, so to be evaluated. As a matter of fact, if the nature of our work is changing, then there is a need to deepen our study of the language of interfaces (cf. Andrews and Van Zundert 2018) so as to find the best strategies to communicate not only our results but also the very nature of our processes.

We ask the panellists to reflect on some questions in this regard. What is already or may already be calculated algorithmically in the field of philology, and where is the current “computational horizon” for textual scholarship? What do we think computation rather than digitization delivers us? In what ways does this advance our textual scholarship thinking? How do we communicate computed results to various textual scholarship audiences? Do we need new kinds of visualisations and knowledge structures that go well beyond the book metaphor?

The complexities of operationalizing scholarly theory and perspectives in data and process

In related fields such as computational linguistics and computational literary studies a discussion is going on concerning the need for theoretical concepts to be effectively connected with observed data in an unambiguous way (Pichler and Reiter 2022, Bode 2023). The same need exists in scholarly editing. In order to bridge the gap between the theory and its actual instantiations it is necessary to express the former as a discrete set of operations to be performed on the data. What are the claims in this respect of current advanced computational or digital textual scholarship? On a practical level this operationalization raises questions on how to deal with the intricacies and complexities of expressing scholarly observation: what granularity of computatio-

nal description is needed and feasible, and how does a scholar practically engage with such granular complex digital data and information? The levels of indirection required to realise this approach might seem overly complex. How can we augment the granularity of data, its addressability, and tractability in order to break down our theories and create explicit connections between our observations and textual phenomena?

Panellists could reflect on the following questions. How do we make computational and digital data, information, and processes of the highest granularity tractable not just by machines but also by the human scholar? Does computational textual scholarship lead to a hairball problem squared? Are graph based structures of text and annotations a more effective means to express interpretation and annotation of text, and even scholarly reasoning about texts? Are semantic web solutions such as RDF-LOD sufficient? Are current leading-edge graph models and solutions able to embrace an open world assumption in a dialectical fashion and practically relate it to standing textual scholarly practice? Can we augment open world knowledge graphs with a technology supporting logical reasoning to push ahead computational hermeneutics?

The object of research: Are we moving into virtualized textual scholarship?

THINK emerges from a group of scholars who, for the most part, were concerned with the edition of mediaeval manuscript works. The theory and practice of such editing was always necessarily a negotiation between the material and unique instances of manuscript text, and the desire to systematise knowledge of “text” (cf., for instance, Cerquiglini 1999). Editing in print tended to prioritise the latter, in the name of eradicating error or other forms of variance. Digitization and computationalization, on the other hand, exert a strong reifying force on textual scholarship. Involved and intimate working with digital and computational methods tends to substitute the materiality of manuscript sources with the immediacy of digital representation and datafication. We can therefore state that our work is more and more centred on virtualized objects of study, which affects the very nature of our own workflows and methods (cf. Ward 2022). As a matter of fact, digital textual scholarship to an extent virtualizes different classes of knowledge about the work in the same digital space - creative links which might previously have been (only) conceptual and cognitive can now be expressed editorially. Additionally, the statements, annotations and processes of the researcher and editor can now be, must be, explicit. In this dispensation, what then is the epistemological link between the sum of assertions about a (textual) object, and that object itself, however defined? Can it be the case that the very object of study, in this case the work that is presumed to exist in the material object, is itself altered? What does this mean with regard to our objects of research, and how does this impact our epistemological processes?

Thus we ask our panellists: Is the nature of the digital object of research different with respect to the analog object? In what sense? How does our knowledge about the objects of study change when we recognise and start making assertions on digital objects? If we move into an age of simulated textual scholarship, what gains do we see for philological process and knowledge inference? What are the potential gains and pitfalls of the fact that virtualization and digital objects provide the possibility to make observations and processes of philology completely explicit and provenance completely transparent?

Invited Panellists

Prof.in Dr. Tara Andrews Universität Wien, Austria (0000-0001-6930-3470)

Ronald Haentjens Dekker DHLab / Huygens Institute, Amsterdam (0000-0001-6737-7986)

Max Grüntgens Deutsches Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg, Germany

Prof. Dr. Aengus Ward University of Birmingham, United Kingdom (0000-0001-9526-0718)

Dr. Joris van Zundert Huygens Institute, Amsterdam (0000-0003-3862-7602)

Moderator

Jun.-Prof.in Dr. Elisa Cugliana Universität zu Köln, University of Cologne (0000-0002-6460-2954)

Fußnoten

1. In our work (cf. Cugliana and Van Zundert 2022) we understand the difference between digital and computational approach in that the first is primarily interested in digital encoding and representation, while the latter refers to the application of (bespoke) computer code as a tool to operationalize the very activity and processes of scholarly editing.

Bibliographie

Andrews, Tara L. 2019. “Critical Edition as Process: A Digital Model.” In *16th Annual Conference of the European Society for Textual Scholarship - Book of Abstracts*, 5. Málaga: Department of English, French and German Philology of the University of Málaga. https://drive.google.com/file/d/1_Tb6xxA94IocProU09aIgf_eHG8h42zj/view.

Andrews, Tara L., and Joris J. Van Zundert. 2018. “What Are You Trying to Say? The Interface as an Integral Element of Argument.” In *Digital Scholarly Editions as Interfaces*, edited by Roman Bleier, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber, and

Gerlinde Schneider, 3–33. Schriften Des Instituts Für Dokumentologie Und Editorik 12. Norderstedt: Books on Demand.

Barabucci, Gioele, and Franz Fischer. 2017. “The Formalization of Textual Criticism. Bridging the Gap between Automated Collation and Edited Critical Texts.” In *Advances in Digital Scholarly Editing: Papers Presented at the Dixit Conferences in the Hague*, Cologne, and Antwerp, edited by Peter Boot, Anna Cappellotto, Wout Dillen, Franz Fischer, Aodhán Kelly, Elena Spadini, and Dirk Van Hulle, 47–53. Leiden: Sidestone Press. <https://www.sidestone.com/books/advances-in-digital-scholarly-editing>.

Bode, Katherine. 2023. “What’s the Matter with Computational Literary Studies?” *Critical Inquiry* 49 (4): 507–29. <https://doi.org/10.1086/724943>.

Buschenhenke, Floor. 2019. “Track Changes: Demarcating Phases in Digital Writing Processes.” In *16th Annual Conference of the European Society for Textual Scholarship - Book of Abstracts*, 7. Málaga: Department of English, French and German Philology of the University of Málaga. https://drive.google.com/file/d/1_Tb6xxA94IocProU09aIgf_eHG8h42zj/view.

Buzzetti, D. 2002. “Digital Representation and the Text Model.” *New Literary History* 33 (1): 61–88.

Cerquiglini, Bernard. 1999. In *Praise of the Variant: A Critical History of Philology*. Baltimore: The Johns Hopkins University Press.

Cugliana, Elisa, and Joris Van Zundert. 2022. “A Computational Turn in Digital Philology.” *Filologia Germanica / Germanic Philology* 14 (1): 43–71. <http://aifg.it/rivista-14/>.

Hadersbeck, Max, Alois Pichler, Florian Fink, Daniel Bruder, and Ina Arends. 2015. “Wittgensteins Nachlass: Erkenntnisse und Weiterentwicklung der FinderApp WiTTFind.” In *DHd2015: Von Daten zu Erkenntnissen. Book of Abstracts*, 187–190. Graz: ZIM/ACDH. <http://gams.uni-graz.at/o:dhd2015.abstracts-vortraege>.

Haentjens Dekker, Ronald, Elli Bleeker, Bram Buitendijk, Astrid Kulsdom, and David J. Birnbaum. 2018. “TAGML: A Markup Language of Many Dimensions.” In *Proceedings of Balisage: The Markup Conference 2018*. Balisage Series on Markup Technologies, Vol. 21. Washington D.C. <https://doi.org/10.4242/BalisageVol21.HaentjensDekker01>.

Kirschenbaum, Matthew. 2008. *Mechanisms: New Media and the Forensic Imagination*. Cambridge (Massachusetts), London (England): The MIT Press.

Oldman, Dominic, Martin Doerr, and Stefan Gradmann. 2015. “Zen and the Art of Linked Data.” In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 251–73. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118680605.ch18>.

Pichler, Axel, and Nils Reiter. 2022. “From Concepts to Texts and Back: Operationalization as a Core Activity of Digital Humanities.” *Journal of Cultural Analytics* 7 (4). <https://doi.org/10.22148/001c.57195>.

Ries, Thorsten. 2017. “The Rationale of the Born-Digital

Dossier Génétique: Digital Forensics and the Writing Process: With Examples from the Thomas Kling Archive.” *Digital Scholarship in the Humanities* 33 (2): 391–424. <https://doi.org/10.1093/lc/fqx049>.

Rosenthal, David S.H. 2015. “Emulation & Virtualization as Preservation Strategies.” Andrew W. Mellon Foundation. https://web.stanford.edu/group/lockss/resources/2015-10_Emulation_&_Virtualization_as_Preservation_Strategies.pdf.

Van Hulle, Dirk. 2021. “Dynamic Facsimiles: Note on the Transcription of Born-Digital Works for Genetic Criticism.” *Variants* 15–16: 231--241. <https://doi.org/10.4000/variants.1450>.

Ward, Aengus. 2022. “Of Digital Surrogates and Immaterial Objects: The (Digital) Future of the Iberian Manuscript in Textual Editing.” *Journal of Medieval Iberian Studies* 14 (1): 41–54. <https://doi.org/10.1080/17546559.2021.2016887>.