





# ADVANCES IN DIGITAL SCHOLARLY EDITING





# ADVANCES IN DIGITAL SCHOLARLY EDITING

PAPERS PRESENTED AT THE DIXIT CONFERENCES  
IN THE HAGUE, COLOGNE, AND ANTWERP

edited by

PETER BOOT  
ANNA CAPPELOTTO  
WOUT DILLEN  
FRANZ FISCHER  
AODHÁN KELLY  
ANDREAS MERTGENS  
ANNA-MARIA SICHANI  
ELENA SPADINI  
DIRK VAN HULLE

© 2017 Individual authors

Published by Sidestone Press, Leiden  
[www.sidestone.com](http://www.sidestone.com)

Imprint: Sidestone Press

Lay-out & cover design: Sidestone Press  
Cover illustration: Tessa Gengnagel

ISBN 978-90-8890-483-7 (softcover)  
ISBN 978-90-8890-484-4 (hardcover)  
ISBN 978-90-8890-485-1 (PDF e-book)

# Contents

<b>Welcome</b>	<b>11</b>
<b>Preface</b>	<b>13</b>
<b>Introduction</b>	<b>15</b>
Peter Boot, Franz Fischer & Dirk Van Hulle	
<b>WP1 CONCEPTS, THEORY, PRACTICE</b>	
<b>Towards a TEI model for the encoding of diplomatic charters. The charters of the County of Luna at the end of the Middle Ages</b>	<b>25</b>
Francisco Javier Álvarez Carbajal	
<b>The uncommon literary draft and its editorial representation</b>	<b>31</b>
Mateusz Antoniuk	
<b>Data vs. presentation. What is the core of a scholarly digital edition?</b>	<b>37</b>
Gioele Barabucci, Elena Spadini & Magdalena Turska	
<b>The formalization of textual criticism. Bridging the gap between automated collation and edited critical texts</b>	<b>47</b>
Gioele Barabucci & Franz Fischer	
<b>Modelling process and the process of modelling: the genesis of a modern literary text</b>	<b>55</b>
Elli Bleeker	
<b>Towards open, multi-source, and multi-authors digital scholarly editions. The Ampère platform</b>	<b>63</b>
Christine Blondel & Marco Segala	
<b>Accidental editors and the crowd</b>	<b>69</b>
Ben Brumfield	
<b>Toward a new realism for digital textuality</b>	<b>85</b>
Fabio Ciotti	
<b>Modelling textuality: a material culture framework</b>	<b>91</b>
Arianna Ciula	
<b>Multimodal literacies and continuous data publishing. Une question de rythme</b>	<b>99</b>
Claire Clivaz	

<b>Theorizing a digital scholarly edition of <i>Paradise Lost</i></b>	105
Richard Cunningham	
<b>The digital libraries of James Joyce and Samuel Beckett</b>	109
Tom De Keyser, Vincent Neyt, Mark Nixon & Dirk Van Hulle	
<b>Editing the medical recipes in the Glasgow University Library Ferguson Collection</b>	115
Isabel de la Cruz-Cabanillas	
<b>The archival impulse and the editorial impulse</b>	121
Paul Eggert	
<b>Pessoa's editorial projects and publications. The digital edition as a multiple form of textual criticism</b>	125
Ulrike Henny-Krahmer & Pedro Sepúlveda	
<b>Reproducible editions</b>	135
Alex Speed Kjeldsen	
<b>'... but what should I put in a digital apparatus?' A not-so-obvious choice. New types of digital scholarly editions</b>	141
Raffaella Afferni, Alice Borgna, Maurizio Lana, Paolo Monella & Timothy Tambassi	
<b>Critical editions and the digital medium</b>	145
Caroline Macé	
<b>Scholarly editions of three rabbinic texts – one critical and two digital</b>	149
Chaim Milikowsky	
<b>From manuscript to digital edition. The challenges of editing early English alchemical texts</b>	159
Sara Norja	
<b>Towards a digital edition of the Minor Greek Geographers</b>	165
Chiara Palladino	
<b>Digital editions and materiality. A media-specific analysis of the first and the last edition of Michael Joyce's <i>Afternoon</i></b>	171
Mehdy Sedaghat Payam	
<b>Challenges of a digital approach. Considerations for an edition of Pedro Homem de Mello's poetry</b>	177
Elsa Pereira	
<b>The born digital record of the writing process. A hands-on workshop on digital forensics, concepts of the forensic record and challenges of its representation in the DSE</b>	183
Thorsten Ries	



<b>Enduring distinctions in textual studies</b>	<b>187</b>
Peter Shillingsburg	
<b>Blind spots of digital editions. The case of huge text corpora in philosophy, theology and the history of sciences</b>	<b>191</b>
Andreas Speer	
<b>Data driven editing: materials, product and analysis</b>	<b>201</b>
Linda Spinazzè, Richard Hadden & Misha Broughton	
<b>Making copies</b>	<b>213</b>
Kathryn Sutherland	
<b>The Videotext project. Solutions for the new age of digital genetic reading</b>	<b>219</b>
Georgy Vekshin & Ekaterina Khomyakova	
<b>A stemmatological approach in editing the Greek New Testament. The Coherence-Based Genealogical Method</b>	<b>223</b>
Klaus Wachtel	
 <b>WP2 TECHNOLOGY, STANDARDS, SOFTWARE</b>	
<b>What we talk about when we talk about collation</b>	<b>231</b>
Tara L. Andrews	
<b>The growing pains of an Indic epigraphic corpus</b>	<b>235</b>
Dániel Balogh	
<b>The challenges of automated collation of manuscripts</b>	<b>241</b>
Elli Bleeker, Bram Buitendijk, Ronald Haentjens Dekker, Vincent Neyt & Dirk Van Hulle	
<b>The role of digital scholarly editors in the design of components for cooperative philology</b>	<b>249</b>
Federico Boschetti, Riccardo Del Gratta & Angelo Maria Del Grosso	
<b>Inventorying, transcribing, collating. Basic components of a virtual platform for scholarly editing, developed for the Historical-Critical Schnitzler Edition</b>	<b>255</b>
Stefan Büdenbender	
<b>Combining topic modeling and fuzzy matching techniques to build bridges between primary and secondary source materials. A test case from the King James Version Bible</b>	<b>261</b>
Mathias Coeckelbergs, Seth van Hooland & Pierre Van Hecke	

<b>The importance of being... object-oriented. Old means for new perspectives in digital textual scholarship</b>	<b>269</b>
Angelo Mario Del Grosso, Emiliano Giovannetti & Simone Marchi	
<b>Edition Visualization Technology 2.0. Affordable DSE publishing, support for critical editions, and more</b>	<b>275</b>
Chiara Di Pietro & Roberto Rosselli Del Turco	
<b>Compilation, transcription, multi-level annotation and gender-oriented analysis of a historical text corpus. Early Modern Ducal Correspondences in Central Germany</b>	<b>283</b>
Vera Faßhauer	
<b><i>Hybrid scholarly edition</i> and the visualization of textual variants</b>	<b>289</b>
Jiří Flaišman, Michal Kosák & Jakub Říha	
<b>Burckhardtsource.org: where scholarly edition and semantic digital library meet</b>	<b>293</b>
Costanza Giannaccini	
<b>EVI-linhd, a virtual research environment for digital scholarly editing</b>	<b>301</b>
Elena González-Blanco, Gimena del Rio, Juan José Escribano, Clara I. Martínez Cantón & Álvaro del Olmo	
<b>Critical diplomatic editing. Applying text-critical principles as algorithms</b>	<b>305</b>
Charles Li	
<b>St-G and DIN 16518, or: requirements on type classification in the Stefan George edition</b>	<b>311</b>
Frederike Neuber	
<b>Visualizing collation results</b>	<b>317</b>
Elisa Nury	
<b>The Hebrew Bible as data: text and annotations</b>	<b>323</b>
Dirk Roorda & Wido van Peursen	
<b>Full Dublin-Core Jacket. The constraints and rewards of managing a growing collection of sources on omeka.net</b>	<b>333</b>
Felicia Roşu	
<b>Of general and homemade encoding problems</b>	<b>341</b>
Daniela Schulz	
<b>The role of the base manuscript in the collation of medieval texts</b>	<b>345</b>
Elena Spadini	

<b>A tailored approach to digitally access and prepare the 1740 Dutch Resolutions of the States General</b>	351
Tuomo Toljamo	
<b>Editorial tools and their development as a mode of mediated interaction</b>	357
Tuomo Toljamo	
<b>TEI Simple Processing Model. Abstraction layer for XML processing</b>	361
Magdalena Turska	
<b>WP3 ACADEMIA, CULTURAL HERITAGE, SOCIETY</b>	
<b>Edvard Munch's Writings. Experiences from digitising the museum</b>	367
Hilde Bøe	
<b>Crowdfunding the digital scholarly edition. Webcomics, tip jars, and a bowl of potato salad</b>	375
Misha Broughton	
<b>Editing medieval charters in the digital age</b>	383
Jan W. J. Burgers	
<b>Editing copyrighted materials. On sharing what you can</b>	391
Wout Dillen	
<b>What you c(apture) is what you get. Authenticity and quality control in digitization practices</b>	397
Wout Dillen	
<b>The journal al-Muqtabas between Shamela.ws, HathiTrust, and GitHub. Producing open, collaborative, and fully-referencable digital editions of early Arabic periodicals – with almost no funds</b>	401
Till Grallert	
<b>Digital editions of artists' writings. First Van Gogh, then Mondrian</b>	407
Leo Jansen	
<b>Digital editing: valorisation and diverse audiences</b>	415
Aodhán Kelly	
<b>Social responsibilities in digital editing – DiXiT panel. Editing and society: cultural considerations for construction, dissemination and preservation of editions</b>	421
Aodhán Kelly	
<b>Documenting the digital edition on film</b>	427
Merisa Martinez	

<b>Towards a definition of ‘the social’ in knowledge work</b> Daniel Powell	433
<b>Beyond Open Access. (Re)use, impact and the ethos of openness in digital editing</b> Anna-Maria Sichani	439
<b>The business logic of digital scholarly editing and the economics of scholarly publishing</b> Anna-Maria Sichani	449
<b>The social edition in the context of open social scholarship. The case of the Devonshire Manuscript (BL Add Ms 17, 492)</b> Ray Siemens	453
<b>Nowa Panorama Literatury Polskiej (New Panorama of Polish Literature). How to present knowledge in the internet (Polish specifics of the issue)</b> Bartłomiej Szleszyński	463
<b>Digital Rockaby</b> Katerina Michalopoulou & Antonis Touloumis	467

# What you c(apture) is what you get. Authenticity and quality control in digitization practices<sup>1</sup>

*Wout Dillen<sup>2</sup>*

*Paper presented at 'Digital Scholarly Editing: Theory, Practice, Methods' DiXiT Convention, Antwerp, October 5-7, 2016.*

Regardless of whether documents are taking a more prominent place in the edition (see Gabler 2007; Robinson 2013; Pierazzo 2014), it seems inevitable that today's scholarly editors are working more and more with these documents' digitized facsimiles. This may happen already in the research stage of the editing process when digital surrogates are integrated into the editor's workflow and effectively become the documents on which the editor bases her interpretation of the text and of its transmission over time. While these facsimiles are not (and probably indeed should not) be used exclusively, *in lieu* of the original analog source materials, it is nevertheless the case that they come more and more to the forefront as the basis of the editor's analysis and transcription – *precisely because* they are more transportable and durable, and because today's imaging hardware and software is capable of revealing aspects of the documents that are difficult or even impossible to detect with the naked eye.

And as important as these facsimiles have become for the editor, they are even more important for the user. Once they take their place in the digital scholarly edition, facsimiles are the closest the user will get to an unedited representation of the document's physical and textual features – even though, as Hans Zeller

---

1 The research leading to these results was conducted as part of the author's work on the 'Digital Scholarly Editing and Memory Institutions' project at the University of Borås (Sweden). This is an Experienced Researcher position that is part of the DiXiT network, a Marie Curie ITN which has received funding from the People Programme (Marie Skłodowska-Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° 317436. This funding also allowed the author to participate in the conference.

2 wout.dillen@uantwerpen.be.

already suggested in the 1970s in his influential essay on ‘Befund und Deutung’, a facsimile is not completely unedited or objective (1971). As the only available visual representative of the physical document, the facsimile becomes the document that the user will use to assess the editor’s transcriptions, claims, and arguments about the text. This means that both users and editors are placing a lot of trust in the digital avatars of physical, historical documents they claim to study. Therefore, this paper argued that issues of authenticity and quality control for digital facsimiles are important issues that need to be addressed at the outset of any digital scholarly editing project. Because too often, we take these digital surrogates at face value.

As editors, we pride ourselves on knowing the difference between the two. We know that text is transmitted from document to document, from medium to medium, and that we should consider the effect that the medium has on the text that is transmitted. Documentary aspects like ‘size’, ‘texture’, ‘coloring’, etc. can help us determine how the text was read at the time the document was made. As textual scholars we are very aware that the medium that carries the text has a very specific impact on our interpretation of that text. And we know that by digitizing an analog object, we effectively are transporting it onto a new medium, which inevitably will have an important impact on our understanding of that text. Some aspects, such as the text, typically will be retained – although the readability of that text will depend strongly on the quality of the images. Some aspects, such as the correctness of its colours, can only be hinted at or assessed in relative terms. And some aspects, like the document’s texture, feel, smell, etc. will be lost. In other words, by transporting select aspects of the original object into a new medium, we are creating a new document, that acts as an intermediary between the original on the one hand, and our edited version on the other.

But even as scholarly editors, do we really treat this new, intermediary document with the same due diligence and scrutiny as we do the documents it mediates between? Personal experience tells me otherwise. Working on the Beckett or Brulez editions at the University of Antwerp’s Centre for Manuscript Genetics as part of my PhD, I was transcribing and checking transcriptions of image files that I had never seen the original of. Of course, I was not the real ‘editor’ of those specific modules. And if I were, I would consider it my editorial duty to go to the archives and see the original documents for myself. But even if that were the case, I doubt that I would spend the same time analysing the ins and outs of every page with the same attention to detail as I did when I was transcribing the documents. In many cases such scrutinous inspection is not even possible, because we have to be very careful not to damage the documents while we are studying them. That is one of the reasons why we digitize in the first place. And when I am using someone else’s scholarly edition, I will think about these issues even less, and interact with the facsimiles ‘as if they were the actual documents’ myself.

This is the premise of my current project at the University of Borås, titled ‘Digital Scholarly Editions and Memory Institutions’ that aims to address these issues by calling for a critical assessment of digitization practices. To do so, the project focuses on the moment the analog object is captured – a process that retains as well as discards a substantial amount of information. Therein lies the meaning of this paper’s title. For the facsimiles that we use in our editions, what you see usually

is what you get: you see an image that you try to make sense of in relation to the edited text, an image that promises to be a faithful representation of the source document, but that is rendered without any data or context to back this claim up. And what you see, is really what you capture: those aspects of the document that are possible to represent in the digital medium as we know it – and really only a selection of those aspects, relative to the means and tools the photographer has at her disposal, the standards that she is required to follow, and the decisions that she has to make. It is these issues and the variables they introduce into the digitization process that this project wants to address.

As the project's first case study, I visited the National Library of Sweden in Stockholm to investigate how this institution handles digitization practices: which standards are used, and why?; how are these standards established?; how are they negotiated between different parties?; and how minutely are they followed in practice? The National Library of Sweden was selected for this case study because 1) it has for some time now been actively involved in the development of a detailed digitization strategy, and 2) it prides itself on putting trained photographers in charge of capturing (i.e. digitizing) the source materials. These factors, as well as some preliminary meetings with Lars Björk – the head of conservation at the National Library – suggested that this institution has a relatively high awareness of the issues this project means to investigate, and that its expertise in these areas may be transferred to memory institutions that have a less advanced digitization strategy. The case study would exist of three parts: 1) a document study – including documents with technical information and digitization manuals that the library had drafted as part of its digitization strategy; 2) an observation study – where I followed a photographer for a couple of days while she was capturing the materials, to study how her workflow was executed in practice, as well as her interaction within this workflow with her hardware and software, and with her colleagues; and 3) a series of interviews with people from the institution who were involved in the digitization process in some form or other.

This paper reported on the preliminary results of this first case study at the KB in Stockholm. By mapping the interactions between different agents who are involved in the digitization process, it aimed to refute the notion that digitization is a simple and straightforward process. Instead, the research suggests that different agents put different demands on the digital object – demands that need to be taken into account by the photographer during the moment of capture. And that even when these issues and agreements are taken into account, digitization often is still a process of 'problem solving'. To this end, the paper presented a set of examples to support the claim that in many cases the quality of the digitization will depend highly on the photographer's professional skills and interpretation (arguably much in the same way as the construction of the 'text' of a scholarly edition depends on the editor's professional skills and interpretation). This would suggest that (as in scholarly editing) the 'accountability' of the photographer is much more important and relevant than her 'objectivity' – and that in a scholarly edition, we should try to find a way to incorporate that accountability into the edition, and thereby raise the user's awareness of the degree of interpretation that goes into the process of creating the digital facsimile.

To conclude, the paper introduced the next proposed case studies that will serve to put the findings at the National Library in Sweden in perspective. By bringing these case studies together, I want to examine how aware these memory institutions are of each other's efforts in this area, how their standards are communicated, and whether they are negotiated further on the international level as well. I think that this mapping of the way in which quality measures are negotiated between different agents is an important first step towards getting a better understanding of the relation between the source document and the digital facsimiles that we are displaying in our editions. And I think that this awareness is essential if we really want to be accountable for all the aspects of our edition, and our interpretation of its materials.

## References

- Gabler, Hans Walter. 2007. 'The Primacy of the Document in Editing.' *Ecdotica* 4: 197-207. Accessed 24 October 2016. [https://www.academia.edu/166634/The\\_Primacy\\_of\\_the\\_Document\\_in\\_Editing](https://www.academia.edu/166634/The_Primacy_of_the_Document_in_Editing).
- Pierazzo, Elena. 2014. 'Digital Documentary Editions and the Others.' *Scholarly Editing: The Annual of the Association for Documentary Editing* 35: 1-23.
- Robinson, Peter. 2013. 'Towards a Theory of Digital Editions.' *Variants* 10: 105-31. Accessed on 24 October 2016. [https://www.academia.edu/3233227/Towards\\_a\\_Theory\\_of\\_Digital\\_Editions](https://www.academia.edu/3233227/Towards_a_Theory_of_Digital_Editions).
- Van Hulle, Dirk, and Peter Shillingsbiurg. 2015. 'Orientations to Text, Revisited.' *Studies in References* 59. 1: 27-44.
- Zeller, Hans. 1971. 'Befund Und Deutung. Interpretation Und Dokumentation Als Ziel Und Methode Der Edition.' In *Texte Und Varianten. Probleme Ihrer Edition Und Interpretation*, edited by Hans Zeller and Gunter Martens, 45-89. München: CH Beck.