

# 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)

# Introduction

*Peter Boot,<sup>1</sup> Franz Fischer<sup>2</sup>*

*& Dirk Van Hulle<sup>3</sup>*

As the papers in this volume testify, digital scholarly editing is a vibrant practice. Since high quality digital scholarly editions have been around for two decades, it is no longer a new undertaking. In fact, digital scholarly editing represents one of the longest traditions in the field of Digital Humanities – and the theories, concepts, and practices that were designed for editing in a digital environment in turn have influenced deeply the development of Digital Humanities as a discipline.<sup>4</sup> That the field is still experimental in many respects is mainly because the possibilities of digital technologies are in constant flux, and ever expanding. By bringing together the extended abstracts from three conferences organised within the DiXiT project (2013-2017), this volume shows how digital scholarly editing is still developing and constantly redefining itself. So who better to answer the question of what digital scholarly editing entails than a broad selection of the community of scholars who practice it? The common denominator for these three conferences, DiXiT (Digital Scholarly Editions Initial Training Network), is a project funded under the EU's Marie Skłodowska-Curie schemes for researcher training and mobility. The conferences, held in The Hague (2015), Cologne and Antwerp (both 2016) brought together young researchers employed within the DiXiT network with external researchers to discuss the continuing development of digital scholarly editing.

## **Digital scholarly editions initial training network**

Scholarly editing has a long-standing tradition in the humanities. It is of crucial importance within disciplines such as literary studies, philology, history, philosophy, library and information science and bibliography. Scholarly editors were among the first within the humanities to realize the potential of digital media

---

1 peter.boot@huygens.knaw.nl.

2 franz.fischer@uni-koeln.de.

3 dirk.vanhulle@ua.ac.be.

4 For the state of the discourse see: Driscoll and Pierazzo (2016), Pierazzo (2015), Apollon *et al.* (2014), Sahle (2013).

for performing research, for disseminating their results, and for bringing research communities together. Hence, digital scholarly editing is one of the most mature fields within the Digital Humanities. Yet, it is also confronted with many challenges. As a cross-disciplinary field, its practitioners must acquire many competences. But since technologies and standards keep improving at a rapid pace, stable practices are hard to establish. Scholars need both high skills and deep knowledge to face present and future challenges of digital scholarly editing. Before DiXiT there was no dedicated postgraduate programme able to provide the training to form a new generation of those scholars.

The institutions cooperating within the DiXiT network are some of the most respected university groups and academies working within the field of digital scholarly editing, along with partner institutions from the commercial and cultural heritage sectors. Together they represent a wide variety of technologies and approaches to European digital scholarly editing. They have created a robust research and training programme in the core skills of digital scholarly editing, reaching researchers from some 300 European institutions. Some 800 researchers in Europe and beyond participated in about 25 events.<sup>5</sup> In applying for a Marie Skłodowska-Curie Initial Training Network in 2012 the network formulated a number of priorities:

1. Attracting young scholars that are able to develop the mix of competences from the humanities, computer science, and information studies that digital scholarly editions require.
2. Combining resources of the most prominent institutions and enterprises in the field in order to train these scholars in the best possible way.
3. Collaborating seamlessly across international scholarly communities in an increasingly cross-cultural and interdisciplinary field.
4. Intensifying efforts towards standards, interoperability and the accumulation of shared methods.
5. Creating suitable models and core curricula for digital scholarly editing.
6. Improving the publishing workflows, infrastructures and publishing venues for digital scholarly editions.
7. Developing a sustainable infrastructure for improving long-term prospects of digital scholarly editing projects.

Fortunately, the EU commission decided to award the requested grant to the DiXiT network. Since 2013, twelve early stage researchers and five experienced researchers are or have been employed within the network, one or two at each participating institution. Each of the early stage researchers works on his or her thesis (or thesis-size subject), while the experienced researchers have more focussed and shorter appointments. One of the unique characteristics of the Marie Curie

---

5 Participating researchers came from almost all European countries and many non-European countries; DiXiT training events and conferences were held across 11 European countries including two online summer schools in Spanish reaching out to Latin America; see <http://dixit.uni-koeln.de/programme>.

scheme is that researchers cannot apply for jobs in their own countries. Although working in different European countries, they therefore are motivated to reach out to each other. This has created a closely-knit group of researchers that works together very well.

### **This collection**

Apart from the informal visits, secondments, research stays and a multitude of video sessions, the DiXiT training program was organized in the form of three camps (basic training in digital scholarly editing), followed by three conventions. While the training camps were primarily targeted at the DiXiT researchers, the conventions were set up so as to create an exchange with other researchers. The topics of the conventions were based loosely on the three DiXiT work packages: Theory, Practice, Methods (WP1, Antwerp), Technology, Standards, Software (WP2, The Hague), Academia, Cultural Heritage, Society (WP3, Cologne).

This collection brings together extended abstracts from those conventions. Not all presenters chose to submit their paper for inclusion in this volume. We reorganised the papers along thematic lines: a paper presented in Antwerp that fits best within the ‘Technology, Standards, Software’ chapter is included under that heading.

### **About the papers**

#### *WP1: Theories, Practices, Methods*

As digital publications are reaching a stage of maturity and scholarly editors are becoming increasingly aware of the seemingly endless possibilities of hybrid or fully digital scholarly editions, the impact of the digital medium on the field of textual scholarship has become undeniable. As a result of this ‘digital turn’, textual scholars are now faced with new challenges and opportunities that have called for a re-evaluation of the field’s established theoretical and practical framework. To satisfy this need, DiXiT organized a conference at the University of Antwerp in association with the European Society for Textual Scholarship (ESTS), focussing on this reassessment of the theories, practices, and methods of scholarly editing in general, and of the digital scholarly edition in particular. The subject, broadly covering DiXiT work package 1, also was dealt with in papers given at the other two conferences.

The achievements shown by recent digital scholarly editions demonstrate some of the potential for innovation in the digital medium including their openness and exploratory nature. These projects have developed a wide range of editorial products. That is why a first important subject of these papers is assessing and mapping these different types of digital scholarly editions, ranging from ‘digital archives’ to ‘knowledge sites’. This includes projects with large amounts of material as well as stabilised, authoritative readings of important works from all fields of history and human culture. The apparent distinction between these digital archives and digital editions has been questioned in the past decade (Sahle 2007,

Price 2009), leading textual scholars to argue that there is in fact no impermeable border between the two, but rather a ‘continuum’ (Van Hulle 2009). Armed with tools such as automated collation it is up to the reader or researcher to decide in which capacity the archive/edition is used. In his opening keynote in Antwerp, Paul Eggert (this volume) translated these ideas into a ‘slider model’ where the digital scholarly edition gravitates between editorial and archival impulses – thus setting the tone of the conference, which would go on to explore all gradations in between.

Of course, these different types of editions each have different editorial and technological needs, and so a second important subject in the papers concerns the architecture of digital scholarly editions providing more than simply text. The symbiotic interdependency of mass digitisation and scholarly editing does not only raise the question as to how the practice of scholarly editing can be adapted to enrich this data, but also how text/image-linkage can be employed in modelling transcription methodologies that allow for enhanced studies in palaeography and codicology, or how a digital archive of manuscripts may integrate various kinds of relevant material (including, for instance, an author’s complete personal library). By trying to answer these emerging questions on how digital scholarly editions are modelled and designed, the papers seek to facilitate innovative research into the editorial process, and to acknowledge a shift in the role of the editor who is enabled to focus on the interpretive consequences of variants and versions.

These technological changes also imply that through the production of digital editions, editors may be transformed into encoders and even programmers. To accommodate these new developments we now have to rethink the kind of training, skills and knowledge the new generation of editors will need to acquire. To jumpstart this discussion, a third important group of papers critically examines the goals, functions and usability of digital scholarly editions. In the same vein, the Antwerp conference aptly was closed by Kathryn Sutherland (this volume), who drew attention to a recent reflowering of print editions and connected this to our current digital preoccupation with ‘making copies’, which ‘cannot simply be put down to the fact that the technology allows it – that computers are good facsimile machines; something more is going on’. Sutherland related this new preoccupation to the recent reconception of bibliography as book history and to the digital medium’s capacity as a medium for replicating materialities other than its own. The document underlying the edition has been raised in status thanks to developments in the reproduction of facsimiles, alerting readers to the insecurity of certain editorial conventions and challenging the editorial model in fundamental ways.

### *WP2 Technology, Software, Standards*

Technology is an essential ingredient for the digital scholarly edition. The universal availability of computing power makes its production possible, and the global network takes care of its dissemination. Technology, however, continues to develop and that has important consequences for digital editing. The DiXiT application mentioned a few research priorities in that direction, such as the integration of web-based editorial tools into the TEI ecosystem, an investigation into the ways

that TEI and other standards can work together, a publication architecture and the integration of analytical tools (e.g. for text analysis, stylometry and visualisation) into the digital scholarly edition. The first DiXiT convention, held in The Hague in September 2015, was devoted to the subject of technology: how to apply new technologies in editions, how to integrate new technologies in tools, and how to use them for publishing editions. This section of the volume brings together a number of papers with that subject, presented at any one of the DiXiT conventions.

Tools for creating an edition are of course an important subject for digital editors. During the conferences, some presentations focused on fully-fledged environments for digital editing created for specific projects, others presented generic environments for digital editing. Some tools focus on specific preparatory parts of the edition process, such as an analysis of page images. Other papers considered the problem at a higher level of abstraction, discussing for instance the architectural principles for an editorial environment, or the role of the editor in creating such an environment.

The most important technology to be discussed at the meetings was automated collation. This wide interest in collation also was evidenced by the high turnout for the special workshop on the subject of collation held in The Hague in November 2016. In this collection the papers look at the concept of collation, at the problems of collating modern manuscripts, the need for a base manuscript in collation and at the visualisation of collation output. Most of these papers, but not all, use the CollateX collation tool to illustrate their point.

Another subject that remains of vital importance to digital editing is publication technology. Again, the papers take a variety of approaches. Solutions range from the low-tech Omeka platform to a dedicated software platform, the Edition Visualization Technology. A very promising approach is the integration of publication information into the TEI files defining the edition.

A number of other papers discussed technological problems based on specific editions. Examples are the questions of data modelling and linguistic issues in a database of Indian epigraphy, deep (socio-) linguistic annotation in a TEI context and the very complex intertextual relations between medieval capitularies.

Another issue that is often discussed is the question of the edition-as-data versus the edition-as-interface. That too is discussed in this section (as well as under WP1). And there are other papers, not so easily brought under a single heading, such as a paper about font definition in the context of the scholarly edition, a paper about topic modelling in the context of the edition, and a general reflection about tools. Taken together, the papers in this section show the dynamic relation between the fields of scholarly editing and digital technology.

### *WP3: Academia, Cultural Heritage, Society*

Humanities scholarship responds to, explores and brings to light our shared cultural heritage. The vast majority of texts used in scholarly editions are owned by cultural heritage organisations which are increasingly moving towards mass digitisation. Scholarly editing is part of this knowledge ecosystem and contributes to the quality of rich knowledge sites for scholars and the general public.

Papers given at the Cologne Convention provided ample evidence of the diversity and dynamic nature of approaches to editing. These approaches reach from the highest methodological standards, developed over decades of textual criticism dealing with huge text corpora in philosophy, theology and the history of sciences to non-academic communities engaging with historical documents and topics as diverse as burial records, sword fighting and spiders. All these papers and some of those presented at the conventions in Antwerp and The Hague centred on the interrelationship between academia, cultural heritage, and society and how scholarly editors can have a wider impact beyond constituencies typically served by its research.

Accordingly, as a first topic the role of museums, libraries and archives on the one hand, and academic scholars and a wider public on the other has been investigated. Papers explored how the quality of digital images and texts can be measured, ensured and improved and how information from various digitization projects can be implemented into digital critical editions. The emergence of new media and formats of cultural heritage material also require new practices, for instance when using film to document the creation of an edition, creating artworks to represent in a creative deformation of a literary work or when editing electronic literature. Digital preservation of endangered cultural heritage has been addressed as a particular challenge, especially in regions of socio-political conflicts and instability with significant constraints of hardware, software, education, network capacity and power.

Web 2.0 approaches and models of public engagement have been investigated as a second research topic. The recent explosion of social networking sites which encourage wide participation from the public opens up a completely new set of challenges and raises many unanswered questions. New editorial formats such as the 'Social Edition' are aiming at combining traditional scholarly editing practices and standards with recent developments in online social media environments. Enabled and facilitated through media wiki technology and online platforms we see a growing number of community-driven editorial projects establishing scholarly standards and methods for citizen science independently from academic paternalism.

Under a third topic scholars engaged with marketing research and experimentation in order to propose viable publication models reflecting both financial sustainability for exploitation and maintenance on the one hand, and the general interest of the scholarly community in open access on the other. With regard to increasingly image and audio oriented literacies particular attention has been drawn to the rhythm of data publication that should be brought into accordance with human thinking, talking and production of knowledge.

## **Conclusion**

The paper book is not dead. While preparing this publication, some suggested to us that the web might be a better place for extended abstracts such as the ones that we collected here. And of course the web is a wonderful place, digital scholarly editing is thriving because of it. Still, for the focussed reading that scholarship requires,



there is nothing like paper. Conference websites are also likely to disappear, and scattered over the web, unlike paper books. We are proud to offer the world this volume as a testimony to the fecundity of the DiXiT project and the creativity of young scholars facing the challenges and opportunities of the digital realm for the scholarly edition.

## References

- Apollon, D. *et al.* 2014. *Digital Critical Editions*. Urbana: University of Illinois Press.
- Driscoll, M. J. and E. Pierazzo (eds). 2016. *Digital Scholarly Editing – Theories and Practices*. Cambridge: Open Book Publishers.
- Pierazzo, E. 2015. *Digital Scholarly Editing – Theories, Models and Methods*. Farnham: Ashgate
- Price, K. 2009. 'Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?' *Digital Humanities Quarterly* 3.3. <http://www.digitallhumanities.org/dhq/vol/3/3/000053/000053.html>.
- Sahle, P. 2007. 'Digitales Archiv und Digitale Edition. Anmerkungen zur Begriffsklärung.' In *Literatur und Literaturwissenschaft auf dem Weg zu den neuen Medien*, edited by Michael Stolz. Zürich: germanistik. ch, 64-84. Online: [http://www.germanistik.ch/scripts/download.php?id=Digitales\\_Archiv\\_und\\_digitale\\_Edition](http://www.germanistik.ch/scripts/download.php?id=Digitales_Archiv_und_digitale_Edition).
- . 2013. *Digitale Editionsformen*. SIDE 7-9 (3 vols.). Norderstedt: BoD.
- Van Hulle, D. 2009. 'Editie en/of Archief: moderne manuscripten in een digitale architectuur.' In *Verslagen en mededelingen van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde*, 119/2, 163-178.

## Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° 317436.

## List of DiXiT full partners

- Cologne Center for eHumanities (CCeH) – University of Cologne (UoC), Germany
- Swedish School of Library and Information Science – University of Borås (HB), Sweden
- Huygens Institute for the History of the Netherlands (Huygens ING), Department of Textual Scholarship and Literary Studies – Koninklijke Nederlandse Akademie van Wetenschappen (KNAW), Netherlands
- Department of Digital Humanities – King's College London (KCL), United Kingdom
- Centre for Manuscript Genetics (CMG) – University of Antwerp (UA), Belgium

Center for Information Modelling in the Humanities – Graz University (GU),  
Austria

An Foras Feasa – National University of Ireland, Maynooth (NUIM), Ireland

Pôle de Lyon – École des Haute Études en Sciences Sociales (EHESS), France

DigiLab – Università di Roma ‘La Sapienza’ (R1), Italy

IT Services (OXIT) – University of Oxford (UOX), United Kingdom

## **List of DiXiT fellows and affiliated institutions**

### *Early Stage Researchers*

Richard Hadden – An Foras Feasa, National University of Ireland, Maynooth

Tuomo Toljamo – Department of Digital Humanities, King’s College London

Elli Bleeker – Centre for Manuscript Genetics (CMG), University of Antwerp

Frederike Neuber – Center for Information Modelling in the Humanities, Graz  
University

Francisco Javier Álvarez Carbajal – Pôle de Lyon, École des Haute Études en  
Sciences Sociales

Elena Spadini – Huygens Institute for the History of the Netherlands (KNAW)

Misha Broughton – Cologne Center for eHumanities (CCeH), University of  
Cologne

Merisa Martinez – Swedish School of Library and Information Science), University  
of Borås

Daniel Powell – Department of Digital Humanities, King’s College London

Anna-Maria Sichani – Huygens Institute for the History of the Netherlands  
(KNAW)

Aodhán Kelly – Centre for Manuscript Genetics (CMG), University of Antwerp

Federico Caria – DigiLab – Università di Roma ‘La Sapienza’

### *Experienced Researchers*

Wout Dillen – Swedish School of Library and Information Science, University of  
Borås

Linda Spinazzè – An Foras Feasa, National University of Ireland, Maynooth

Magdalena Turska – IT Services (OXIT), University of Oxford

Gioele Barabucci – Cologne Center for eHumanities (CCeH), University of  
Cologne

Roman Bleier – Center for Information Modelling in the Humanities, Graz  
University