Graph Data-Models and Semantic Web Technologies in Scholarly Digital Editing

Schriften des
Instituts für Dokumentologie und Editorik

herausgegeben von:

Band 15

Schriften des Instituts für Dokumentologie und Editorik — Band 15

Graph Data-Models and Semantic Web Technologies in Scholarly Digital Editing

edited by

Elena Spadini, Francesca Tomasi, Georg Vogeler

2021

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über **http://dnb.d-nb.de/** abrufbar.

Digitale Parallelfassung der gedruckten Publikation zur Archivierung im Kölner Universitäts-Publikations-Server (KUPS). Stand 5. Dezember 2021.

Herstellung und Verlag: Books on Demand GmbH, Norderstedt
ISBN: 978-3-7543-4369-2
Einbandgestaltung: Stefan Dumont nach Vorarbeiten von Johanna Puhl und Katharina Weber
Satz: LuaTEX, Bernhard Assmann

Contents

Projects and Editions

Appendices

Introduction

Elena Spadini, Francesca Tomasi

This volume explores possible interactions between digital texts, the graph data-model, scholarly editions, and the semantic web. Combinations of these objects and concepts, first explored in the past decades, remain experimental to date, and represent one possible area of development for the field of digital scholarly editing.

The conceptual and technological stack adopted in most digital editions today[1], as well as in the tools, software, and platforms used to produce them, is based on the centrality of the document, on hyperlinks within closed databases, and on hierarchical text encoding. This *status quo* is challenged by the need to represent multiple perspectives on the same textual work, and by the diffusion of knowledge graphs to organize the relationships between data in a semantically explicit and computationally readable way.

Our experience in creating what we call today a Semantic Scholarly Digital Edition (SSDE) started with the project *Vespasiano da Bisticci, Lettere* (Tomasi 2013), where, in the first release, we proposed a model for managing a documentary approach to the edition through semantic web technologies (Tomasi 2012). Working on the notion of 'knowledge site', by the time the edition reached the third version, it proposed a complete Linked Open Data publication of the collection of letters (Tomasi 2020), by also focusing on the topic of reproducibility of an SSDE[2], as a way to enable dialogue within the scholarly community.

Scholarly edition of historical documents have often been pioneering in exploring the possibilities offered by the graph data-model, and by semantic web technologies (Meroño-Peñuela et al. 2014). In a process of "extension of indices" (Vogeler 2020, 44), they make use of available conceptual models and ontologies (e.g., CIDOC-CRM and FOAF) for the modelling of named entities like events, persons, and places, particularly relevant in this kind of editions.

In scholarly editing, the identification and linking of semantic entities often needs to be coupled with the representation of other aspects of a textual work, such as intertextuality, style, textual variation, as well as the expression of linguistic, palaeographic and codicological features. Scholarly editors and digital humanists are not yet equipped to represent these phenomena in the growing world of knowledge graphs, because of the relative lack of dedicated formal models, standards, and pieces of

[1] Good catalogues of digital editions are available in Franzini et al. (2016-) and Sahle (2008-).

[2] See Tomasi 2020, section *Documentation* available at <https://dharc-org.github.io/vespasiano-da-bisticci-letters-de/documentation/index.html>.

Graph Data-Models and Semantic Web Technologies in Scholarly Digital Editing. Ed. by Elena Spadini, Francesca Tomasi and Georg Vogeler. Schriften des Instituts für Dokumentologie und Editorik 15. Norderstedt: Books on Demand, 2021. 1–6.

infrastructure[3]. The reflections and contributions in these proceedings aim to address these needs. Towards this same end, for instance, we proposed a model for the relationships between the different entities concerned by genetic criticism (documents, publications, dossiers), in the context of the scholarly edition of the complete works of Gustave Roud (Christen and Spadini 2019).

Some challenges are indeed common to all kinds of scholarly editions – literary, historical or others. Among them is the need to represent the editorial assertions, or the interpretation, as explicitly as possible, and the requirement for referencing each element (portion of the text, interpretation act, or external entity) in a unique and permanent manner. Provenance-aware models and standards for machine consumption of textual resources are fundamental to addressing these demands, and both topics are well represented in this volume.

The experiment we carried out on Bufalini's notebooks (Daquino et al. 2019a) pursued these same issues, and gave us the opportunity to propose a complete workflow for SSDEs.[4] Starting from the text transcription (a non-neutral action, which required the modelling of both the digital edition features and the XML/TEI markup), we then moved onto the data modelling phase, bearing in mind the need to reuse existing ontologies, and the importance of assuming a provenance-aware approach. We adopted the same approach for the RDF transformation, which required us to use named graphs and the nanopublication model (Groth et al. 2010). Finally, we developed a web application, which was useful not only to serve and exploit RDF data, but also to abide by the principle of using a data-driven visualization to enhance user-experience.

There is still a lot of work to do in order to more-fully exploit the possibilities of this vast domain, and this collection of papers aims to face these demanding challenge. With this volume, we hope to propose some interesting suggestions for doing research in the field of graph data-models and semantic web technologies in scholarly digital editing.

The topics addressed in this proceedings explore this subject from different perspectives: of the infrastructures, of the formal models, and of real-life project implementations. The first of these perspectives, that is, the need for further infrastructure

---

[3] The graph data-model has been successful in the representation of textual variants (Schmidt and Colomb 2009), notably thanks to the Gothenburg model for automatic collation, but the related MVD (multi-version document) format has not been widely adopted. For what concerns formal models, for citations and intertextuality see the Citation Typing Ontology (CiTO) (Peroni and Shotton 2012) and the Intertextual Relationships Ontology for literary studies (INTRO) (Oberreither 2020); for linguistics, cf. for example Chiarcos et al. 2018; Tittel et al. 2018; Franzini 2019; for ancient and medieval documents, see Sharing Ancient Wisdoms Ontology (SAWS) (2012), Linked Ancient World Data Ontology (LAWD) (Cayless 2014) and CRMtex (Murano and Felicetti 2020).

[4] See Daquino et al. 2019b, section *Introduction* available at <http://projects.dharc.unibo.it/bufalini-notebook/introduction>.

to cover existing gaps, and advancements that are being made in this domain, are explored in this volume by the contributions of Boot and Koolen, Cayless and Romanello, Neill and Schmidt, Prosser and Schloen, and Vogeler. They touch upon long-standing issues as diversified as APIs for texts, friendly GUIs, the integration of TEI and RDF, data granularity, and technologies affordance. Boot and Koolen consider the question of how to derive RDF triples from a TEI encoded text and where to store them; their edition server serves the triples to any annotation tool, in order for users to analyse or enrich an edition. The data are modelled according to an ontology for the editorial domain. The paper by Cayless and Romanello addresses the gap in the available digital infrastructure of services for handling the resolution of URIs for texts and their parts. Some real-life uses cases for such services are introduced, and a proposal of their three main components – a registry service, an identifier resolution service and a document metadata scheme – is presented and documented with examples. Neill and Schmidt introduce the SPEEDy editor, its data model for annotations, and its user interface. SPEEDy is a standoff editor, through which any range of text, including overlapping ones, can be enriched with one or more annotations. Prosser and Schloen describe the OCHRE system, which has been successfully used for around a decade to manage text corpus projects. OCHRE, built as a graph database, is flexible and customizable, thanks to a semi-structured data model in which atomized data is made available for different kinds of arrangements and links. The paper by Vogeler focuses on the technological affordances of the XML hierarchical tree model, the RDF graph representation, and stand-off annotations, exploring the potential of virtuously combining the three.

The second perspective, concerning the question of formal models, is central to the contribution of Giovannetti, as well as of Cools and Padlina. These articles share with others a scalable approach to ontology development: overarching conceptual models are defined or reused, and guide the creation of domain- or project-specific models. Among the former, CIDOC-CRM and FRBR form a reference point for representing the semantics of events, cultural objects, and their relationships. This does not mean, however, that they cover the requirements of any scholarly edition: their use is still experimental in this domain, as is shown by the CRMtex (Murano & Felicetti 2020), an extension of the CIDOC-CRM that supports the study of ancient documents, the first version of which was only released in June 2020. This volume should help to expose some of the advantages and limitations of the use of these higher-level ontologies. Cools and Padlina provide a description of the infrastructure using semantic web technologies created by the Swiss National Infrastructure for Editions (NIE-INE, University of Basel) in order to manage eleven digital edition projects. The focus of the paper is on data modelling, to which a scalable approach is applied in order to integrate generic and project-specific ontologies. The contribution by Giovannetti proposes an ontology for capturing the critical apparatus as a knowledge graph, which

enables, by the use of formal semantics, the decoupling of text and interpretation, and the integration of scholarly editions representing textual variance in the linked open data cloud.

The articles by Burrows et al., by Münnich and Ahrend, and by Sippl, Burghardt and Wolff take the third perspective, each reporting on very different real-life projects, from manuscript description to music and medieval charters. All of the scholars involved in these projects (all of which make use of semantic web technologies) saw the benefit of taking a data-centric approach to scholarly editing and manuscript studies in order to enhance the value of the relationships between different kinds of data, as well as among multiple datasets. The paper by Burrows et al. reports on the results of the transformation of the TEI encoded manuscripts catalogues of the Bodleian Library into Linked Open Data. Special attention is devoted to the provenance data, which are only semi-structured in TEI. The paper by Münnich and Ahrend offers a compelling state of the art of the formalization of music as a graph, before providing an overview of the modelling approach chosen in the *Anton Webern Gesamtausgabe*, covering semantic relationships between different areas of the edition, and interoperability with other datasets. The paper by Sippl, Burghardt and Wolff focuses on a digital edition of medieval charters, which is based on an extremely heterogeneous dataset. NLP techniques are used for entities extraction. The graph database chosen is Neo4j, on top of which an exploratory web application is developed. The data model, which makes substantial use of CIDOC-CRM, is illustrated in detail.

To conclude, we believe this volume will be a valuable resource for anyone embarking on the adventure of joining scholarly editing with graph data models both inside and outside the emerging Linked Open Data ecosystem.

Bibliography

Cayless, Hugh, *Linking Ancient World Data Ontology (LAWD)*, 2014 <http://lawd.info/>

Chiarcos, Christian, Émilie Pagé-Perron, Ilya Khait, Niko Schenk, and Lucas Reckling, 'Towards a Linked Open Data Edition of Sumerian Corpora', in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Miyazaki, Japan: European Language Resources Association (ELRA), 2018) <https://www.aclweb.org/anthology/L18-1387>

Christen, Alessio, and Elena Spadini, 'Modeling Genetic Networks. Gustave Roud's Œuvre, from Diary to Poetry Collections', *Umanistica Digitale*, 7 (2019) <https://doi.org/10.6092/issn.2532-8816/9063>

Daquino, Marilena, Francesca Giovannetti, and Francesca Tomasi, 'Linked Data per le edizioni scientifiche digitali. Il workflow di pubblicazione dell'edizione semantica del quaderno di appunti di Paolo Bufalini', *Umanistica Digitale*, 3.7 (2019) <https://doi.org/10.6092/issn.2532-8816/9091>

Daquino, Marilena, Francesca Giovannetti, and Francesca Tomasi, eds., *Paolo Bufalini's Notebook, 1981-1991. The Digital Edition.* (Università di Bologna: /DH.arc, 2019) <http://projects.dharc.unibo.it/bufalini-notebook>

Franzini, Greta, 'Towards Connecting Scholarly Editions to Corpora in the LiLa (Linking Latin) Knowledge Base of Linguistic Resources', in *Lesen und lesen lassen. Edition und Korpus*, Vortragsreihe des IZED im Wintersemester 2019/20 (Wuppertal (Germany): University of Wuppertal, 2019) <https://doi.org/10.5281/zenodo.3613371>

Franzini, Greta, Peter Andorfer, and Ksenia Zaytseva, *Catalogue of Digital Editions: The Web Application*, 2016 <https://doi.org/10.5281/zenodo.1250796>

Groth, Paul, Andrew Gibson, and Jan Velterop, 'The Anatomy of a Nanopublication', *Information Services & Use*, 30.1–2 (2010), 51–56 <https://doi.org/10.3233/ISU-2010-0613>

Meroño-Peñuela, Albert, Ashkan Ashkpour, Marieke van Erp, Kees Mandemakers, Leen Breure, Andrea Scharnhorst, and others, 'Semantic Technologies for Historical Research: A Survey', *Semantic Web — Interoperability, Usability, Applicability*, 6.6 (2015), 539–64 <https://doi.org/10.3233/SW-140158>

Murano, Francesca, and Achille Felicetti, *Definition of the CRMtex*, 2020 <http://www.cidoc-crm.org/crmtex/ModelVersion/version-1.0-0>

Oberreither, Bernhard, *Intertextual Relationships Ontology (INTRO)*, 2020 <https://w3id.org/lso/intro/currentbeta#>

Peroni, Silvio, and David Shotton, 'FaBiO and CiTO - Ontologies for Describing Bibliographic Resources and Citations', *Journal of Web Semantics*, 17 (2012), 33–43 <https://doi.org/10.1016/j.websem.2012.08.001>

Sahle, Patrick, *A Catalog of Digital Scholarly Editions*, 2008- <https://www.digitale-edition.de/>

Schmidt, Desmond, and Robert Colomb, 'A Data Structure for Representing Multi-Version Texts Online', *International Journal of Human-Computer Studies*, 67.6 (2009), 497–514 <https://doi.org/10.1016/j.ijhcs.2009.02.001>

Sharing Ancient Wisdoms Ontology (SAWS), 2012 <https://ancientwisdoms.ac.uk/method/ontology/>

Tittel, Sabine, Helena Bermúdez-Sabel, and Christian Chiarcos, 'Using RDFa to Link Text and Dictionary Data for Medieval French', in *Proceedings of the 6th Workshop on Linked Data in Linguistics* (presented at the LDL-2018, Myazaki, Japan, 2018) <http://lrec-conf.org/workshops/lrec2018/W23/pdf/10_W23.pdf>

Tomasi, Francesca, 'Digital Editions Between Embedded Markup and External Representation. A Case Study: Vespasiano da Bisticci's Letters', in *Dall'informatica umanistica lle culture digitali*, ed. by Fabio Ciotti and Gianfranco Crupi, Quaderni Digilab, 2 (Sapienza Università Editrice, 2012), 201–18 <http://www.editricesapienza.it/sites/default/files/Quad_DigiLab_Informatica_Umanistica_Culture_Digitali.pdf>

———, ed., *Vespasiano da Bisticci, Lettere. A Knowledge Site* (Università di Bologna: /DH.arc, 2020) <https://web.archive.org/save/https://projects.dharc.unibo.it/vespasiano/>

———, ed., *Vespasiano da Bisticci, Lettere. A Semantic Digital Edition*, 2013 <https://doi.org/10.6092/unibo/vespasianodabisticciletters>

Vogeler, Georg, 'Digital Edition of Archival Material - Machine Access to the Content: On the Role of Semantic Web Technologies in Digital Scholarly Editions', in *Digitizing Medieval*

Sources ? L'édition en ligne de documents d'archives médiévaux, ed. by Christelle Balouzat-Loubet (Turnhout: Brepols, 2020), 37–56 <https://doi.org/10.1484/M.ARTEM-EB.5.117327>

Infrastructures and Technologies

Formal Models

Projects and Editions

Appendices

Biographical Notes

Thomas Ahrend (University of Basel, Switzerland – thomas.ahrend@unibas.ch) studied Musicology, Philosophy and Literary Studies in Frankfurt a. M. and Berlin. He received his MA 1996, and his PhD 2005 at Technische Universität Berlin with a dissertation on the instrumental music of Hanns Eisler. 1997–2010 member of the editorial staff of the Hanns Eisler Gesamtausgabe in Berlin. Since September 2010, member of the editorial staff of the Anton Webern Gesamtausgabe at Musikwissenschaftliches Seminar at University of Basel.

Peter Boot (Huygens ING, The Netherlands – peter.boot@huygens.knaw.nl) studied mathematics and Dutch language and literature; he wrote his PhD thesis about annotation in scholarly digital editions and its implications for humanities scholarship. He oversaw the creation of the digital edition of the letters of Vincent van Gogh. He is employed as a senior researcher at the Huygens Institute for the History of the Netherlands where he works, among other things, as a consultant in several edition projects.

Manuel Burghardt (University of Leipzig, Germany – burghardt@informatik.uni-leipzig.de) is head of the Computational Humanities Group at Leipzig University. He is interested in the use of digital tools and computational techniques to explore new modes of doing research in the humanities. His most recent areas of research are Sentiment Analysis in the Humanities, Drametrics, Computational Intertextuality, Computational Analysis of Movies and Series and Music Information Retrieval.

Toby Burrows (University of Oxford, United Kingdom – toby.burrows@oerc.ox.ac.uk) is a Senior Researcher in the Oxford e-Research Centre at the University of Oxford, and a Senior Honorary Research Fellow in the School of Humanities at the University of Western Australia.

Hugh Cayless (Duke University, USA - hugh.cayless@duke.edu) is Senior Digital Humanities Developer at the Duke Collaboratory for Classics Computing. Hugh has over a decade of software engineering expertise in both academic and industrial settings. He also holds a Ph.D. in Classics and a Master's in Information Science. He is one of the founders of the EpiDoc collaborative and currently serves on the Technical Council of the Text Encoding Initiative.

Hans Cools (University of Basel, Switzerland – 1961-2021) had a master degree in medicine and a specialization in orthopaedic surgery and traumatology (Universities of Ghent and Antwerp, Belgium, 1997), a bachelor's degree in physical

therapy, and a standalone degree in informatics (1999). Through various research and project management positions, in both companies and academic institutions, he gained expertise in different aspects of the Semantic Web technologies, focusing particularly on formal data modeling and machine reasoning. Those positions were in internationally collaborative research projects in a biomedical setting, mainly of the 5-7th EU Framework Program. Foremost in these projects were semantic interoperability and reusability of data. Since 2016, he worked in the humanities, as knowledge engineer, ontologist, and Semantic Web technology expert, at the University of Basel, as part of the NIE-INE project, which highlights scholarly editing. He (co-)published several articles, and gave workshops on the implementation of Semantic Web technologies in biomedicine and the humanities. He passed away in April 2021.

Francesca Giovannetti (University of Bologna, Italy – francesc.giovannett6@unibo.it) is a second-year PhD student in Digital Humanities at the Department of Classical Philology and Italian Studies, University of Bologna. She received an MA in Digital Humanities from King's College London and a second cycle degree in Digital Humanities and Digital Knowledge from the University of Bologna. She is interested in combining digital scholarly editing with semantic web technologies and in the use of digital technologies in education.

Matthew Holford (University of Oxford, United Kingdom – matthew.holford@bodleian.ox.ac.uk) is Tolkien Curator of Medieval Manuscripts at the Bodleian Library, University of Oxford.

Marijn Koolen (Royal Netherlands Academy of Arts and Sciences - Humanities Cluster, The Netherlands – marijn.koolen@gmail.coml) studied artificial intelligence and wrote his PhD thesis on using hyperlinks in information retrieval algorithms. He has worked on scholarly annotation for digital humanities research and on annotation-related information behaviour and information systems. He works as a researcher and developer at the Humanities Cluster of the Royal Netherlands Academy of Arts and Sciences, where he leads a project on developing annotation support within the *CLARIAH research infrastructure* project.

David Lewis (University of Oxford, United Kingdom – david.lewis@oerc.ox.ac.uk) is a Research Associate in the Oxford e-Research Centre at the University of Oxford.

Andrew Morrison (University of Oxford, United Kingdom – andrew.morrison@bodleian.ox.ac.uk) is a Software Engineer in the Bodleian Digital Library Systems and Services, Bodleian Library, University of Oxford.

Stefan Münnich (University of Basel, Switzerland – stefan.muennich@unibas.ch) studied musicology and communication science at the Technische Universität Berlin, MA 2011 with a thesis on cantional setting in Heinrich Schütz's Becker-Psalter. 2012 research assistant, 2013–2015 research associate of the Felix Mendelssohn Bartholdy. Sämtliche Briefe edition at University of Leipzig (co-editor of vols. 9 & 12). Since October 2015 research associate of the Anton Webern Gesamtausgabe, Basel; received his Doctorate degree in 2020 at the department of musicology at the University of Basel with a dissertation about music notation and its codes.

Iian Neill (Digital Academy of the Academy of Sciences and Literature, University of Mainz - Iian.Neill@adwmainz.de) is a visiting researcher at the Digital Academy of the Academy of Sciences and Literature Department at the University of Mainz, Germany. He is the creator of Codex, a text annotation environment which uses standoff property annotation to generate entities in a graph meta-model. Codex is currently being used to produce a digital edition of the epistles of Hildegard von Bingen at the Digital Academy in Mainz.

Roberta Padlina (University of Basel, Switzerland – roberta.padlina@unibas.ch) studied medieval philosophy at the University of Fribourg, Switzerland, obtaining a doctoral degree in June 2020. She has twelve years of professional experience in the field of Digital Humanities, thanks to which she has been able to work closely with different actors involved in the online publication of open access research. Roberta has worked for several years for e-codices –Virtual Library of Manuscripts in Switzerland and currently coordinates the National Infrastructure for Editions (NIE-INE) project. Roberta's main focus is on the opportunities and challenges that the digital shift poses for traditional education and research institutions, including developing semantic web strategies for scholarly publications and cultural goods.

Kevin Page (University of Oxford, United Kingdom – kevin.page@oerc.ox.ac.uk) is a Senior Researcher in the Oxford e-Research Centre and Associate Member of Faculty in the Department of Engineering in the University of Oxford.

Miller C. Prosser (University of Chicago, USA – m-prosser@uchicago.edu) earned his Ph.D. in Northwest Semitic Philology from the University of Chicago. His academic interests include the social and economic structure of Late Bronze Age Ras Shamra-Ugarit and the use of computational methods for philological and archaeological research. Miller is the Associate Director of the Digital Studies MA program at the University of Chicago where he teaches courses on Data Management and Data Publication for the Humanities. He also works as a

researcher at the OCHRE Data Service of the Oriental Institute of the University of Chicago where he consults with and supports research projects using the Online Cultural and Historical Research Environment (OCHRE). He has also worked as a tablet photographer for the Mission de Ras Shamra (Ugarit) and the Persepolis Fortification Archive Project, employing advanced digital photographic methods such as reflectance transformation imaging, photogrammetry, and high-resolution digital scanning.

Matteo Romanello (Université de Lausanne, Switzerland - matteo.romanello@unil.ch) is Ambizione SNF Lecturer at the University of Lausanne, where he conducts a project on the commentary tradition of Sophocles' Ajax. Matteo is a Classicist and a Digital Humanities specialist with expertise in various areas of the Humanities, including archaeology and history. After obtaining his PhD from King's College London, he worked as a research scientist at EPFL's DHLAB on the Linked Books and Impresso projects, before moving to his current position. He was also teaching fellow at the University of Rostock, researcher at the German Archaeological Institute, and visiting research scholar at Tufts University.

Sandra Schloen (University of Chicago, USA – sschloen@uchicago.edu) is the Manager of the OCHRE Data Service at the Oriental Institute of the University of Chicago, and is the co-designer and developer of the Online Cultural and Historical Research Environment (OCHRE). Trained in computer science and mathematics (B.Sc. University of Toronto; M.Ed. Harvard University), Sandra has spent over 30 years working with technology as a systems analyst, technical trainer, and software developer. A long association with colleagues in the academic community has enabled her to develop a specialty in solving problems in the Digital Humanities where challenges of data capture, data representation and data management abound. Specifically, she has served extensively as a database manager for several archaeological projects in Israel and Turkey, and supports a wide range of research projects at the Oriental Institute and at other universities.

Desmond Schmidt (University of Bologna - desmond.allan.schmidt@gmail.com) has a background in classical Greek philology, information security and eResearch. He has worked on several scholarly edition projects, including the Vienna Wittgenstein Edition (1990–2001), Digital Variants (2004–2008), the Australian Electronic Scholarly Editions project (2012–2013), the Charles Harpur Critical Archive (2014-) and a pilot edition of Gianfrano Leopardi's Idilli (2018-). He currently works on developing practical web-based tools for making, visualising and publishing digital scholarly editions.

Colin Sippl (University of Regensburg, Germany – colin.sippl@ur.de) is currently a project employee at the University Library of Regensburg. Since 2017, he has been working on extending the open access services of the Electronic Journals Library (EZB). More recently, he has started developing and setting up a digital repository for literature, artefacts and experiments relating to the early life sciences based on the Invenio framework. He specialised in textual data mining and the development of media services in the institutional domain.

Elena Spadini (University of Lausanne - elena.spadini@unil.ch) is a postdoctoral researcher at the University of Lausanne. She holds a Ph.D. in Romance Philology from the University of Rome Sapienza (2016) and a M.A. in Digital Humanities from the École nationale des chartes (2014). She was a Marie Curie fellow in the IT Network DiXiT and co-directed the related volume Advances in Digital Scholarly Editing (Sidestone Press, 2017). She published in international journals and taught specialized courses in various European countries in the field of Digital Philology.

Francesca Tomasi (University of Bologna - francesca.tomasi@unibo.it) is associate professor in Archival Science, Bibliography and Librarianship at the University of Bologna (Italy). Her research is mostly devoted to digital cultural heritage, with a special attention to documentary digital edition, and a focus on knowledge organization methods in archives and libraries. She is member of different scientific committees of both associations and journals. In particular, she is President of the Library of the School of Humanities in the University of Bologna (BDU - Biblioteca di Discipline Umanistiche), Director of the international second cycle degree in Digital Humanities and Digital Knowledge (DHDK), President of the Italian Association of Digital Humanities (AIUCD – Associazione per l'Informatica Umanistica e la Cultura Digitale), and co-head of the Digital Humanities Advanced Research Center (/DH.ARC). She wrote about 100 papers and 4 monographs related to DH topics. She is editor and scientific director of several digital scholarly environments.

Athanasios Velios (University of the Arts London, United Kingdom – a.velios@arts.ac.uk) is Reader in Documentation at the University of the Arts London.

Georg Vogeler (University of Graz - georg.vogeler@uni-graz.at) is professor for Digital Humanities at the University of Graz and scientific director of the Austrian Center for Digital Humanities and Cultural Heritage at the Austrian Academy of Sciences. He is a trained historian (Historical Auxiliary Sciences). He spent several years in Italy (Lecce, Venice). In 2011, he became member of faculty at the Centre for Information Modelling at Graz University, where he was nominated

full professor for Digital Humanities in 2016 and head of department in 2019. His research interests lie in late medieval and early modern administrative records, diplomatics (digital and non digital), digital scholarly editing and the history of Frederic II of Hohenstaufen (1194–1250). He was and is part in several national and international research projects related to his research interests.

Christian Wolff (University of Regensburg, Germany – christian.wolff@ur.de) has been Professor of Media Informatics at the Institute for Information and Media, Language and Culture at the University of Regensburg since 2003. He holds a PhD in information science and is a habilitated computer scientist. His research interests include: human-computer interaction, multimedia and web-based information systems, (multimedia) software engineering and information retrieval (in particular information literacy and social media).