

Laura Righi

# Digital Dark Ages: The Role of Medieval Corpora in the Context of the Digital Humanities and Religious Studies

(doi: 10.17395/112876)

Cristianesimo nella storia (ISSN 0393-3598)

Fascicolo 3, settembre-dicembre 2023

**Ente di afferenza:**

*Università di Modena e Reggio Emilia (Unimore)*

Copyright © by Società editrice il Mulino, Bologna. Tutti i diritti sono riservati.

Per altre informazioni si veda <https://www.rivisteweb.it>

**Licenza d'uso**

Questo articolo è reso disponibile con licenza CC BY NC ND. Per altre informazioni si veda <https://www.rivisteweb.it/>

# Digital Dark Ages: The Role of Medieval Corpora in the Context of Digital Humanities and Religious Studies

*Laura Righi*

In recent years, the debate on the role and methodologies of the digital humanities has seen considerable development, including in the specific – but disciplinarily vast – domain of Religious Studies. Even if it is a recent debate, its premises are based on epistemological questions and assumptions whose history it's important to outline. In this context, a great contribution could be provided by the research conducted on medieval textual corpora. Through the study of some cases, starting from Roberto Busas' Index Thomisticus up to ongoing research projects, this contribution presents some trends and specificities of the analysis and publication of medieval sources in the digital environment. Aiming at discussing innovations and limits of this research field, and what can be its contribution to the ongoing debate on digital religious studies.

Keywords: Middle Ages, Digital Humanities, Religious Studies

The expression «Digital Dark Age» was coined in the 1990s to refer to the fear that, in the future, there could be a substantial documentary loss due to the rapid obsolescence of digital media and their lack of reliability, which would make our age essentially a dark one<sup>1</sup>. This

\* This work was supported by the PNRR project Italian Strengthening of ESFRI RI RESILIENCE (ITSERR) funded by the European Union – NextGenerationEU (CUP: B53C22001770006). The views and opinions expressed, however, are those of the author, and do not necessarily correspond to those of the European Union or of the European Commission. Neither the European Union nor the European Commission can be considered responsible for the views and opinions expressed in this work.

<sup>1</sup> An example of this assessment is T. Kuny, *A Digital Dark Ages? Challenges in the Preservation of Electronic Information*, in *63rd IFLA (International Federation of Library Associations and Institutions) Council and General Conference, 1997*, available at <https://archive.ifla.org/IV/ifla63/63kuny1.pdf> (23 January 2024).

is a historically inaccurate expression in that it draws parallels with the image of the Middle Ages as lacking in learning, a dark age, which is now widely refuted by historians<sup>2</sup>. In this contribution, I instead intend to use the expression to present the opposite idea, that is, to emphasize the great impetus that the study of the Middle Ages and medieval texts has given to the development of the digital humanities and the opportunities that are still latent in the relationship between medieval technologies and sources.

This paper aims at contributing to an already rich and detailed field of research. The debate concerning the role of the digital humanities – or humanities computing, as some have long preferred to call it – has indeed had great momentum in recent years, including in the broad, albeit particular, field of research of religious studies<sup>3</sup>. In fact, the first volume published in the «Introductions to Digital Humanities – Religion» series, dates back to 2019 and will surely be a reference point for future debates and developments in the disciplines involved<sup>4</sup>.

<sup>2</sup> Even worse is the more literal and misguided Italian translation of the expression: *medioevo digitale*.

<sup>3</sup> For a discussion of this naming choices, see D. Buzzetti, *Alle origini dell'informatica umanistica: Humanities Computing e/o Digital Humanities?*, in «Atti della Accademia Nazionale dei Lincei – Classe di Scienze morali, storiche e filologiche. Rendiconti», nona serie, 30, 1-2, 2019, sezione monografica, *Il museo virtuale dell'informatica archeologica*. Atti della “Segnatura” (13 dicembre 2017), a cura di P. Moscati e T. Orlandi, pp. 71-103; P. Svensson, *Humanities Computing as Digital Humanities*, in «Digital Humanities Quarterly», 3, 3, 2009, available at <https://dhq-static.digitalhumanities.org/pdf/000065.pdf> (25 January 2024).

<sup>4</sup> The series «Introductions to Digital Humanities – Religion» is published by De Gruyter and edited by Claire Clivaz, Charles Ess, Gregory Price Grieve, and Sally Prome. Inaugurated with an initial volume devoted to Buddhism, in recent years the series has included several volumes concerning Christianity or specific methodologies and objects of study. Given the centrality of these discussions, it is worth listing the volumes in the series: *Digital Humanities and Buddhism. An Introduction*, ed. by D. Veidlinger, Berlin-Boston, De Gruyter, 2019; *Digital Humanities and Christianity. An Introduction*, ed. by T. Hutchings and C. Clivaz, Berlin-Boston, De Gruyter, 2021; *Digital Humanities and Research Methods in Religious Studies. An Introduction*, ed. by C.D. Cantwell and K. Petersen, Berlin-Boston, De Gruyter, 2021; *Digital Humanities and Libraries and Archives in Religious Studies. An Introduction*, ed. by C.B. Anderson, Berlin-Boston, De Gruyter, 2022; *Digital Humanities and Material Religion. An Introduction*, ed. by E.S. Clark and R. McBride Lindsey, Berlin-Boston, De Gruyter, 2022; and *Digital Humanities and Religions in Asia. An Introduction*, ed. by L.W.C. van Lit and J.H. Morris, Berlin-Boston, De Gruyter, 2024.

The field of religious studies is particularly interesting not only because of the richness, stratification, and plurality of sources (extremely diverse in terms of languages, alphabets and even supports) and the interdisciplinarity of methods and research, but also because of the early attention that religious communities and institutions have paid to media. This interest was in radio and television as well as in the latest ICT (Information and Communication Technologies), proving to be, as Tim Hutchings wrote in 2015, more resourceful than scholars of religions: «Religious practitioners have taken to digital media with greater enthusiasm than academics»<sup>5</sup>.

From 2015 to the present, the state of digital religious studies, has evolved considerably, as has that of digital humanities in general. Since these are areas of great growth, also thanks to the exponential and ongoing development of certain technologies (such as the recent spread of artificial intelligence), some scholars believe that the digital humanities should be considered one of the many areas in which humanities research is developing and not a (new) field of research in its own right<sup>6</sup>. However, it is still useful to talk about digital humanities and, in particular, the specific methodologies that they have produced, given that they are still being defined and tested.

### 1. *From Digital Dark Age to Digital Middle Ages*

If there is one field in digital humanities that has been active for decades, it is medieval research. As early as 2012, in the journal «Digital Medievalist», John Unsworth called medievalists «early adopters» of the methodologies and tools of information technology<sup>7</sup>. These scholars included not only historians of the Middle Ages, but also linguists and philologists, who found fertile ground for their work in medieval texts.

<sup>5</sup> T. Hutchings, *Digital Humanities and the Study of Religion*, in *Between Humanities and the Digital*, ed. by P. Svensson and D.T. Goldberg, Cambridge, MA, MIT Press, 2015, pp. 283-294, here 283.

<sup>6</sup> See *A Companion to Digital Humanities*, ed. by S. Schreibman, R. Siemens and J. Unsworth, Oxford, Blackwell Publishing, 2004.

<sup>7</sup> J. Unsworth, *Medievalists as Early Adopters of Information Technology*, in «Digital Medievalist», 7, 2012, available at <https://journal.digitalmedievalist.org/article/id/6983/> (24 January 2024).

In recent years, digital humanities themselves have begun to be the subject of historical analysis, in some cases thanks to the epistemological tools provided by media studies and media archaeology<sup>8</sup>. The historical reviews and analyses published to date vary widely according to the multiple contexts in which they are set, but they all share the same starting point: the pioneering work initiated by Roberto Busa in the 1940s for the lemmatization of all the works of Thomas Aquinas. Busa's work has undoubtedly altered the future not only of digital humanities but also, in a broader sense, of information technology<sup>9</sup>. Busa's *Index Thomisticus* is unanimously considered a foundational project – also because of the relationship he established with IBM, which provided funding and tools – but after that several fields and disciplines engaged in creating new technologies and tools for publishing and analyzing medieval sources.

Many of the projects that evolved after this first step concerned medieval corpora, and they developed in several directions, with a marked interest in the sources, their accessibility, and new opportunities for textual analysis. It is impossible to retrace all the various research projects that, especially since the 1990s, began working on medieval manuscript sources and the numerous digitalization initiatives launched, particularly following the invention of the World Wide Web. Suffice it here to mention the collections of major cultural institutions, such as the British Library, the French Bibliothèque

<sup>8</sup> So far, attempts at a critical and systematic analysis of the history of digital humanities are still rare; most contributions in this regard consist more of introductions and reviews of the literature. See S. Hockey, *The History of Humanities Computing*, in *A Companion to Digital Humanities*, cit., pp. 3-19; R. Bleier et al., *Digitale Mediävistik und der deutschsprachige Raum*, in «Das Mittelalter», 24, 1, 2019, pp. 1-12; T. Hutchings and C. Clivaz, *Introduction*, in *Digital Humanities and Christianity*, cit., pp. 1-18; D.J. Birnbaum, S. Bonde and M. Kestemont, *The Digital Middle Ages. An Introduction*, in «Speculum», 92, S1, 2017, pp. S1-S38; L.K. Morreale and S. Gilsdorf, *Introduction. The Medievalist, Digital Edition*, in *Digital Medieval Studies. Practice and Preservation*, ed. by L.K. Morreale and S. Gilsdorf, Amsterdam, ARC Humanities Press, 2022, pp. 1-10.

<sup>9</sup> Fr. Busa's project – see [www.corpusthomicum.org](http://www.corpusthomicum.org) (25 January 2024) – and the success of that enterprise has been investigated from many perspectives and by more authoritative voices, which is why I will not dwell further on it here. In addition to the numerous related reviews, see the volume dedicated to Busa, S.E. Jones, *Roberto Busa, S.J., and the Emergence of Humanities Computing. The Priest and the Punched Cards*, New York-London, Routledge, 2016.

Nazionale with its Gallica catalog, and the DigVatLib platform of the Vatican Library.

In the last decade in particular we have seen an explosion of digitalization projects carried out by a variety of small and large institutions within the broad group of GLAMs (galleries, libraries, archives, and museums). This has also been made possible by the new capabilities of automated digital reproduction, which has made large amounts of material available, although often unaccompanied by accurate metadata. The landscape of digitized online sources that researchers are faced with is thus as rich as it is fragmented and it is not always easy to navigate, which is why there is an increasing need for institutes to collect sources and infrastructures that allow access to multiple collections and resources beyond the individual access provided by those producing or preserving the records<sup>10</sup>.

But the aspects that make the potential of ICT of particular interest for the study of medieval corpora are multiple, and are not limited to the possibility of making the images of manuscripts or archival documents available<sup>11</sup>. The initial phase of digitization has been followed over the years by the first projects engaged in structuring data and metadata, leading to the birth and development of digital codicology, understood not only as a discipline dealing with digitizing and making sources available online, but with reflecting on the different ways of managing the work of digital publishing, its evolution, and the methodologies related to it<sup>12</sup>. The approach of digital codicology

<sup>10</sup> Crucial in this respect are European research infrastructures such as DARIAH (Digital Research Infrastructure for the Arts and Humanities) and, in the field of religious studies, RESILIENCE (Religious Studies Infrastructure: toolS, Innovation, Experts, conNections and Centres in Europe). For a fuller overview, see J. Edmond, *Collaboration and Infrastructure*, in *A New Companion to Digital Humanities*, ed. by S. Schreibman, R. Siemens and J. Unsworth, Oxford, Wiley Blackwell, 2016, pp. 54-65.

<sup>11</sup> An important role has been played by the debate concerning digital archives and supports. In Italy in particular, the debate has been opened and led by the work of Tullio Gregory. See, for example, *L'eclisse delle memorie*, a cura di T. Gregory e M. Morelli, Roma-Bari, Laterza, 1994. On the specific issue of archives of historical-religious texts, see A. Melloni, *Archivi storico-religiosi e strumenti di lavoro. Problemi ed esperienze di coniugazione di tecnologia e ricerca in Italia*, in *Le carte della memoria. Archivi e nuove tecnologie*, a cura di M. Morelli e M. Ricciardi, Roma-Bari, Laterza, 1997, pp. 129-143.

<sup>12</sup> Particularly interesting are the works on older digital editions, which are not considered obsolete but rather seen as a specific object of study. See B. Whearty, *Digital Codicology. Medieval Books and Modern Labor*, Stanford, Stanford University Press,

helps to focus, enhance, and standardize the research work underlying digitalization, but at the same time it also helps to plan for further possible advancements, such as the use of medieval manuscripts in dissemination or teaching<sup>13</sup>.

One of the central issues concerning medieval corpora, which had already been treated by Busa, concerns the creation of instruments for critical editions and textual analysis<sup>14</sup>. As with the *Corpus Thomisticum*, whose work continued after Busa's death<sup>15</sup>, more recent research has focused on issues such as semantic analysis and new tools for searching for references, concordances, and allusions within a textual corpus and across multiple corpora<sup>16</sup>. Such work has been made possible by the creation and progressive expansion of large, fully digitized Latin corpora<sup>17</sup>.

When referring to ancient textual corpora, there is another aspect that cannot be overlooked, which again has seen the strong commit-

2022. For an updated introduction to digital publishing and its future challenges, see instead *Medieval Manuscripts in the Digital Age*, ed. by B. Albritton, G. Henley and E. Treharne, New York-London, Routledge, 2021; and *Meeting the Medieval in a Digital World*, ed. by M.E. Davis, T. Mahoney-Steel and E. Turnator, Leeds, ARC Humanities Press, 2018.

<sup>13</sup> E. Salvatori, *Edizioni digitali di fonti medievali tra ricerca e pubblico*, in *Il medievista come Public Historian*, a cura di E. Salvatori, Roma, Istituto storico italiano per il medio evo, 2022, pp. 227-250; and, addressing the area of games, S.E. Jones, *New Media and Modeling: Games and the Digital Humanities*, in *A New Companion to Digital Humanities*, cit., pp. 84-97.

<sup>14</sup> D. Buzzetti, *Digital Editions and Text Processing*, in *Text Editing, Print, and the Digital World*, ed. by K. Sutherland and M. Deegan, Burlington, Ashgate, 2009, pp. 45-61; A. Meroño-Peñuela *et al.*, *Semantic Technologies for Historical Research: A Survey*, in «Semantic Web», 6, 6, 2015, pp. 539-564.

<sup>15</sup> *Diachronic Treebanks for Historical Linguistics*, ed. by H.M. Eckhoff, S. Luraghi and M. Passarotti, Amsterdam, John Benjamins, 2020; M. Passarotti, *The Project of the Index Thomisticus Treebank*, in *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution*, ed. by M. Berti, Berlin-Boston, De Gruyter, 2019, pp. 299-319.

<sup>16</sup> An interesting categorization has recently been proposed and argued for by J.-P. Genet, *Le référencement*, in *Digitizing Medieval Sources. L'Édition en Ligne de Documents d'Archives Médiévales*, éd. par C. Balouzat-Loubet, Turnhout, Brepols, 2020, pp. 7-22.

<sup>17</sup> From this point of view, one of the most interesting projects concerning medieval sources is the *Corpus Corporum*, [mlat.uzh.ch/home](http://mlat.uzh.ch/home), in which many of the projects of the main institutions in the publication of medieval Latin sources are being housed: [www.mgh.de/en/digital-mgh/openmgh](http://www.mgh.de/en/digital-mgh/openmgh); [lta.bbaw.de](http://lta.bbaw.de); and [www.mirabileweb.it/home](http://www.mirabileweb.it/home) (25 January 2024).

ment of medieval source scholars. In particular, since the early 2000s, new projects have been developed regarding the transition from the written source to its transcription. Standards, therefore, were first created for the online publication of manuscript documents, followed by the development of software for handwritten text recognition (HTR) and discussion concerning the role and specifics of digital paleography<sup>18</sup>. Medievalists, however, paid much less attention to issues of optical character recognition (OCR) of ancient and modern printed texts – which often need to be digitized and made machine-readable in order to carry out machine learning activities<sup>19</sup> – or to creating and enriching databases and archives with easy-access research tools based on archival-diplomatist principles<sup>20</sup>. Medievalists have mainly focused on the possibilities of HTR because it is one of the most technologically advanced fields<sup>21</sup>. The growth in digital paleography and diplomatics can not only contribute to technologies for automated transcription but also offer an important contribution to the discussion surrounding the organization and publication of online sources, including not just manuscripts but also archival collections<sup>22</sup>.

<sup>18</sup> A project to note is [www.digipal.eu](http://www.digipal.eu) (25 January 2024). See P.A. Stokes, *Computer-Aided Palaeography, Present and Future*, in *Kodikologie und Paläographie im digitalen Zeitalter/Codicology and Palaeography in the Digital Age*, hrsg. von M. Rehbein, P. Sahle und T. Schassan, Norderstedt, Books on Demand, 2009, pp. 309-338; P.A. Stokes, *Digital Approaches to Palaeography and Book History. Some Challenges, Present and Future*, in «Frontiers in Digital Humanities», 2, 5, 2015, available at <https://www.frontiersin.org/articles/10.3389/fdigh.2015.00005/full> (25 January 2024).

<sup>19</sup> The availability of ancient and modern printed texts has increased thanks to the work of companies such as Google Books and Archive.org; however, many of those texts have not been accompanied by an high-quality text digitization work, so they are not machine-readable.

<sup>20</sup> First among these is *Monumenta Germaniae Historica*, which has moved its historical collections to online publication of some sections and, more recently, the creation of the large database of the *Cartae Europae Medii Aevi (5th-15th c.)*, <https://cema.lamop.fr> (25 January 2024).

<sup>21</sup> See M. Kestemont, V. Christlein and D. Stutzmann, *Artificial Paleography: Computational Approaches to Identifying Script Types in Medieval Manuscripts*, in «Speculum», 92, S1, 2017, pp. S86-S109. See also Transkribus, the AI-powered software available for HTR: [readcoop.eu/it/transkribus](http://readcoop.eu/it/transkribus) (25 January 2024).

<sup>22</sup> C. Williams, *Diplomatic Attitudes: From Mabillon to Metadata*, in «Journal of the Society of Archivists», 26, 1, 2005, pp. 1-24; G. Vogeler, *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*, cit., pp. 37-56; Id., *The Content of Accounts and Registers in their Digital Edition. XML/TEI, Spreadsheets, and Semantic Web Technologies*, in *Konzeptionelle Überlegungen zur Edition von*

A whole other branch concerns data which is rapidly evolving in terms of methodologies and tools: suffice here it to mention the various works of digital and sound visualization applied to the field of architecture<sup>23</sup>, the applications of geolocation<sup>24</sup>, or even the specific field of research concerning the study and preservation of medieval seals<sup>25</sup>.

While I cannot mention all the innovations that have taken place over the past thirty years, it is worth emphasizing how, at this stage, the pioneering research initiated by Busa in the 1940s has grown and led to the systematization of these procedures into stable models and infrastructures for the study of medieval (and other) sources. The rise of the Web was one of the turning points in the development of digital humanities and in the field of medieval studies, allowing not only for easier access to medieval sources but also, more generally, enabling online medieval research, thus changing its nature<sup>26</sup>. At the same time, the creation of shared standards and languages such as the Text Encoding Initiative (TEI) are also considered fundamental steps

*Rechnungen und Amtsbüchern des späten Mittelalters*, hrsg. von J. Sarnowsky, Göttingen, Vandenhoeck & Ruprecht, 2016, pp. 13-42; A. Ambrosio, S. Barret and G. Vogeler, *Digital Diplomatics. Expertise between Computer Science and Diplomats*, in *Digital Diplomatics. The Computer as a Tool for the Diplomatist?*, ed. by A. Ambrosio, S. Barret and G. Vogeler, Cologne, Böhlau Verlag, 2014, pp. 9-14.

<sup>23</sup> For some recent projects in these areas, see S. Antonopoulos *et al.*, *Soundscape of Byzantium*, in «Speculum», 92, S1 2017, pp. S321-S335; B.V. Pentcheva and J.S. Abel, *Icons of Sound: Auralizing the Lost Voice of Hagia Sophia*, *ibidem*, pp. S336-S360; and S.A. Bonde, A. Coir and C. Maines, *Construction-Deconstruction-Reconstruction: The Digital Representation of Architectural Process at the Abbey of Notre-Dame d'Ourcamp*, *ibidem*, pp. S288-S320.

<sup>24</sup> Suffice it to mention the first major digital atlas created, the Digital Atlas of Roman and Medieval Civilization (DARMC), now called Mapping Past Societies (MAPS), [darmc.harvard.edu](http://darmc.harvard.edu) (25 January 2024).

<sup>25</sup> On the study of seals, see M. Bolom-Kotari, K. Rybenská and L. Holakovská, *Sphragistic Sources and Treasures in Archives and Museums of the Czech Republic: Research, Preservation, Digitization*, in *Medioevo digitale. Documenti e archivi/Arte e architettura*, a cura di A. Ambrosio e P. Vitolo, Roma, Viella, 2023, pp. 65-86; J. McEwan, *New Approaches to Old Questions: Digital Technology, Sigillography, and DIGISIG*, in *Digital Medieval Studies*, *cit.*, pp. 33-48.

<sup>26</sup> With respect to the creation of databases and resources, it is also important to mention *The Labyrinth* blog founded in 1993, which was the first to offer medieval scholars a dedicated online portal: [blogs.commonsworld.org/georgetown.edu/labyrinth](https://blogs.commonsworld.org/georgetown.edu/labyrinth) (24 January 2024). Its story is recounted by its founders: D. Everhart and M. Irvine, *Beginnings. The Labyrinth Medieval Studies Website*, in *Digital Medieval Studies*, *cit.*, pp. 11-32.

of the theoretical development of humanities in the digital environment as enduring and shared models of knowledge representation<sup>27</sup>.

Medieval studies are certainly not the only discipline engaged in this area, but some of the projects concerning the Middle Ages that have been launched have opened up new perspectives and fields of research. But what are the specifics of medieval research that have led it to be a leader in projects in this area as well? Or again, what are the characteristics of medieval sources that have led them to be a place for experimentation? Is it possible to trace key elements in the *traditio* and the preservation of medieval textual corpora or in the methodologies developed for their study? To try to answer these questions in the following pages, I will focus on recent and ongoing initiatives which dealt with medieval corpora related to ecclesiastical sources, to highlights different paths and methodology adopted and whether there are particularities in the work regarding medieval textual corpora.

## 2. «A Text Corpus is Not an Archive»: Medieval Archives and Corpora for the History of the Papacy and Ecclesiastical Institutions

In the special issue of the journal «Speculum» dedicated to the digital Middle Ages, Franz Fischer emphasizes some fundamental aspects that should be kept in mind when creating new digital collections and editions, among them the difference between digital archives and digital corpora<sup>28</sup>. In his essay, the author proposes the idea that, when it comes to digital editions, the focus should remain on the text rather than its material aspects, thus prioritizing their annotation and metadation.

Although, as we have seen, projects in the digital environment have been numerous, enduring, and have enabled the creation of models and standards for digital editions, the concern for the most appropriate ways to proceed is still at the heart of the debate. To delve into the specifics of some of these lines of research I have cho-

<sup>27</sup> These watershed elements were highlighted in S. Hockey, *The History of Humanities Computing*, in *A Companion to Digital Humanities*, cit., pp. 3-19.

<sup>28</sup> F. Fischer, *Digital Corpora and Scholarly Editions of Latin Texts: Features and Requirements of Textual Criticism*, in «Speculum», 92, S1, 2017, pp. S265-S287, esp. S280-S281. For the quote in the title at of this section see *ibidem*, p. S280.

sen to focus on some very recent projects dealing with digitalization, making available and preserving large medieval text compilations.

The three projects I will discuss here share one characteristic: they concern documentation produced in the religious-ecclesiastical sphere. Often, these projects are drawn from large collections whose producers (and in some cases also conservators) are ecclesiastical individuals, bodies, or institutions. This is so not only because, as is well known, these are institutions of long duration, with a large documentary production, but also because medieval texts and documents have been the subject of study, collection, and republication for a long time, especially between the 18<sup>th</sup> and 20<sup>th</sup> centuries. This is perhaps one of the most distinctive specificities of medieval textual collections – particularly those concerning medieval church history – that has made them particularly suitable for the application of ICT<sup>29</sup>.

First in order of foundation is the Monasterium.net project, one of the richest collections of late medieval archival documents in Europe. It currently holds 600,000 documents from the Middle Ages through the early modern age<sup>30</sup>. The platform has two distinct classifications: one by archival fond, as established by the preserving entity, and one by collection. It aims to gather – in divisions according to archival sources – all the documents related to the given producing entities, in most cases showing an image of the original document and its description or, in a few cases, its transcription, diplomatic analysis, and *regestum*<sup>31</sup>. It is a collaborative platform that, upon registration, allows external scholars to add and edit metadata under the supervision of an internal project moderator. The decision to include the image of the archive document stems from the origins of this plat-

<sup>29</sup> As early as 1990, Alberto Melloni, for example, reflected on the close relationship that exists between church history and the computer: see A. Melloni, *Church History and the Computer*, in «Computers and the Humanities», 24, 5-6, 1990, pp. 393-395.

<sup>30</sup> This is still an ongoing project and is a steadily growing platform, which has recently found new momentum thanks to a European Research Council grant. A. Ambrosio, *Digital Critical Editions of Medieval Documents on Monasterium.Net. The Digital Edition of the Documents of the Abbey S. Maria Della Grotta*, in *Digitizing Medieval Source*, cit., pp. 69-84; G. Vogeler, "Monasterium.net". *Eine Infrastruktur für diplomatische Forschung*, in «Das Mittelalter», 24, 1, 2019, pp. 247-252; K. Heinz, *Monasterium.net. Auf dem Weg zu einem mitteleuropäischen Urkundenportal*, in *Digitale Diplomatie. Neue Technologien in der historischen Arbeit mit Urkunden*, hrsg. von G. Vogeler, Köln, Böhlau, 2009, pp. 70-77.

<sup>31</sup> The project uses the standards for publishing established by the Charters Encoding Initiative (CEI), [www.cei.lmu.de](http://www.cei.lmu.de) (25 January 2024).

form. Born as a place to make the documentary heritage of some institutions available online, it later became a place to also develop a critical edition of the contents<sup>32</sup>. The concern to publish images of original documents does not seem to be shared by many other projects dealing with archival records; this is also due to the requirements of technical and reproduction rights that the publication of an image of the original document entails<sup>33</sup>.

A different approach is that taken by APOSCRIPTA, a database launched in 2017 that aims to become one of the largest textual databases of papal documents from the earliest times to the modern era. Currently, the database contains more than 25,000 texts, mostly late medieval pontifical letters. In contrast to other databases, it pays no attention to the place of preservation of the original document (so no images are given), since it intends to create a new collection of documents to replace or supplement the *Regesta Pontificum Romanorum* collection created by August Potthast between 1874 and 1875. The database and interface still appear to be under construction, but it relies on the TELMA online critical edition platform and flows into the larger *Cartae Europae Medii Aevi (5th-15th c.)* (CEMA) database, two of the most impressive initiatives launched for publishing medieval documents<sup>34</sup>. An essential undertaking would be the dialogue and interoperability of this project with other active ones concerned with the digital publication of papal letters and their collections of *regesta*, the creation and publication of which are still in progress<sup>35</sup>.

The third project is the *Corpus Synodaliūm*, a textual database of more than 1,400 provincial and local synods issued between 1200 and 1500, access to which is made possible by the PhiloLogic4 plat-

<sup>32</sup> See Ambrosio, *Digital Critical Editions*, cit., pp. 69-84.

<sup>33</sup> The choice of, for example, the Claustra database is different, [www.ub.edu/claustra/ita/info/sobre\\_el\\_projecte\\_claustra](http://www.ub.edu/claustra/ita/info/sobre_el_projecte_claustra) (25 January 2024).

<sup>34</sup> Established in 2005, TELMA is a platform that aims to create a space for the online publication of documents belonging to corpora of various natures: [telma.hypotheses.org](http://telma.hypotheses.org) (25 January 2024). CEMA, on the other hand, is a collection of papers that have been published and digitized (which contains 250,000 items) to enable their comparison and study through tools of linguistic analysis and the aggregation of the data contained in these texts: [cema.lamop.fr](http://cema.lamop.fr) (25 January 2024).

<sup>35</sup> See A. Kuczera, *Regesta Imperii Online. Von der Internetpräsentation zur Internetedition*, in *Digitale Diplomatie*, cit., pp. 84-90; *Gallia Pontificia*, [qdemo.perspectiviana.net/fr/galliapontificia/start](http://qdemo.perspectiviana.net/fr/galliapontificia/start) (25 January 2024); *Regesta Pontificum Romanorum Online*, [www.papsturkunden.de/EditMOM/database-information.do](http://www.papsturkunden.de/EditMOM/database-information.do) (25 January 2024).

form<sup>36</sup>. The platform allows the use of rather simple text analysis tools associated with some some tools for the geolocalisation and the analysis of words occurrences that can improve the research. The simplicity of the search tool also makes it, in contrast to other text search tools, an easy-to-access database that can be used to search a single text or to search for individual terms, names etc. The database does not offer the possibility of checking the original manuscripts but provides an otherwise impossible overview of the spread of canon law throughout the European territory in the late Middle Ages, probably giving impetus to new critical editions of the texts (there is no critical apparatus here) as well as the discovery of new witnesses and texts<sup>37</sup>. As identified by the project's initiator, the need for building a collection of synodal texts arises from the need to answer research questions that have remained unanswered due to the fragmented state of studies and sources. Building such a new collection is still necessary despite the vast amount of documentation that has been put together by scholars in the modern era: they are the research tools with which all researchers begin (thanks also to the numerous digitalization projects carried out on these volumes) and from which digital publication projects also inevitably start, which is why their collation, selection, and critical points deserve discussion<sup>38</sup>.

The approach, choices, and goals of these projects concerning medieval documentary corpora are very different from one another, but all three emphasize the central role of catalogues, collections, and past editions. The dependence of digitalization projects on tools created in the modern era deserves to be further explored, for it raises quite a few questions with respect to the theoretical foundations of digital edition projects.

<sup>36</sup> R. Dorin, *Corpus Synodaliūm. Local Ecclesiastical Legislation in Late Medieval Europe*, in *BodoArXiv*, 23 March 2021, available at [doi.org/10.34055/osf.io/qx2ve](https://doi.org/10.34055/osf.io/qx2ve) (25 January 2024); see also [corpus-synodalium.com](https://corpus-synodalium.com) (25 January 2024), and R. Dorin, *The Bishop as Lawmaker in Late Medieval Europe*, in «Past & Present», 253, 1, 2021, pp. 45-82.

<sup>37</sup> Indeed, this instrument is part of a thriving publishing environment. In addition to the series that continues to publish updated editions of the councils – *Conciliorum Oecumenicorum Generaliumque Decreta*, ed. G. Alberigo et A. Melloni, 7 voll., Turnhout, Brepols, 2007 – there are also initiatives to digitize modern collections, as in the case of the 18<sup>th</sup>-century work by G.D. Mansi, *Sacrorum Conciliorum Nova et Amplissima Collectio*, available at [mansi.fscire.it/immagine/9428](https://mansi.fscire.it/immagine/9428) (25 January 2024).

<sup>38</sup> R. Dorin, *Corpus Synodaliūm: Medieval Canon Law in a Digital Age*, in *Digital Medieval Studies*, cit., pp. 49-76.

### 3. «The Database is the Theory»

The last pages have highlighted that projects concerning medieval sources have characterized the entire history of digital humanities since before they existed<sup>39</sup>. From being an experimental area, works on medieval sources have become the perfect field to test and establish standards for the development of the ensuing projects, although the theoretical debate on the methodologies and specificities of digital publications, research tools, and databases creation is still ongoing<sup>40</sup>.

The potential of medieval corpora in advancing the discussion regarding religious studies in the digital environment is diverse and deserves deeper study. Scholars have recognised the importance of these large textual corpora, highlighting their role in elaborating and applying new methodologies and research tools. At the same time, it has been noted that it is almost impossible to extract these data manually. The inability to fully explore these documents through individual research or in small units is certainly due to limitations in the funding of research in this area, but also because of limitations in accessing these sources, which remain physically scattered and not readily available or readable. One aspect, however, should not be overlooked: these are large reservoirs of data that, although more complex to access than others, are unrestricted by copy rights or limitations due to privacy issues. Their content and texts, therefore, can be freely used to experiment with new tools of analysis.

Crucial in this respect is the presence of compendiums, collections, indexes and synthesis compiled by important scholars and erudites of the modern era working extensively on medieval documentary collections. The principles that created these great collections and their legacy in studying medieval corpora should be examined, and the methodologies and possibilities of disciplines such as philology and critical editions of texts, paleography and the diplomatic analysis of medieval documents, and the indexing and analysis of historical data should be discussed.

<sup>39</sup> Susan Smylyan's use of the expression used in the title of this section is reported by J. Bauer, *Who You Calling Untheoretical?*, in «Journal of Digital Humanities», 1, 1, 2011, available at [journalofdigitalhumanities.org/1-1/who-you-calling-untheoretical-by-jean-bauer](http://journalofdigitalhumanities.org/1-1/who-you-calling-untheoretical-by-jean-bauer) (24 January 2024).

<sup>40</sup> F. Gibbs, *Critical Discourse in Digital Humanities, ibidem*, available at <https://journalofdigitalhumanities.org/1-1/critical-discourse-in-digital-humanities-by-fred-gibbs/> (24 January 2024).

It must, therefore, be an opportunity to enhance the collections and the work done during the modern era, but also an opportunity to critically analyze them, identifying the methodologies and objectives behind them, in order to avoid posing old research questions to new technologies. As David J. Birnbaum, Sheila Bonde, and Mike Kestemont point out in the introduction to «Speculum»'s special issue on the Digital Middle Ages: «It is troubling that much new digital medieval work responds more closely to the questions and concerns of nineteenth-century medieval scholarship than those of the twentieth or twenty-first centuries»<sup>41</sup>. From this point of view, there is still a lot of work to be done, and it is precisely this path that needs to be recounted and discussed, in even the most recent debates on digital humanities and religious studies.

Only recently have the works and publications on the digitalization and publishing of historical sources been accompanied by reflection on the media and their role in creating and changing knowledge, and how such role is investigated, applying the principles and methodologies of media studies<sup>42</sup>. It will be essential to analyze the history of digital humanities to imagine their future<sup>43</sup>. A key role in this is played by journals and specialized publishing series, which are centers for debate<sup>44</sup>.

Medieval corpora have thus been the grounds for experimentation, but medieval studies must not miss the opportunity to be the forum of the discourse on developing studies in the digital environment, in ICT, and in creating standards, languages, methodology, metadata, and tools. The numerous past and present experiences can

<sup>41</sup> Birnbaum, Bonde and Kestemont, *The Digital Middle Ages*, cit., p. S31. This echoes similar thoughts advanced a few years earlier by L. Swain, *Past, Present, and Future of Digital Medievalism*, in «Literature Compass», 9, 12, 2012, pp. 923-932.

<sup>42</sup> K.D. Scherff and L.J. Sobehrad, *Introduction. Media Technologies and the Digital Humanities*, in *Media Technologies and the Digital Humanities in Medieval and Early Modern Studies*, ed. by K.D. Scherff and L.J. Sobehrad, New York-London, Routledge, 2023, pp. 1-10.

<sup>43</sup> See *A New Companion to Digital Humanities*, cit. Recently, some reflections also seem to have been emerging on the role of these projects in the development of teaching, an area that has already seen a greater use of technology than others, see *Media Technologies and the Digital Humanities*, cit.; and *Digital Humanities and Libraries and Archives*, cit.

<sup>44</sup> See Birnbaum, Bonde and Kestemont, *The Digital Middle Ages*, cit., which provides a list of the leading specialized journals, such as «Digital Medievalist Journal» and «Digital Philology. A Journal of Medieval Cultures».

make a significant contribution to the debate and development of religious studies<sup>45</sup>.

laura.righi@gmail.com  
Università di Modena e Reggio Emilia  
viale Timavo, 93 – 42121 Reggio Emilia, Italia  
Fondazione per le scienze religiose  
via San Vitale 114 – 40125 Bologna, Italia  
ORCID 0000-0002-5770-0978

<sup>45</sup> This work developed within the research project REVER, which aims precisely at offering new tools and methods for the study of medieval papal documents in the broader framework of RESILIENCE, the research infrastructure for Religious Studies.