Information Technologies for Medieval Studies: Some Recent Experiences in Italy

by Roberto Delle Donne

Curial Sources and Digital Humanities. Long term projects in international comparison

ed. by Irmgard Fees, Claudia Märtl, Andreas Rehberg, Jörg Voigt

Firenze University Press

Information Technologies for Medieval Studies: Some Recent Experiences in Italy

by Roberto Delle Donne

The article highlights the relations between Italian medieval studies and digital humanities, starting from researches conducted on Vatican documents and other ecclesiastical sources since the 1980s; then it goes on to discuss some more recent projects, based on various types of sources. A multi-faceted debate emerges, in which the experiences of historians and editors of medieval sources are intertwined with those of archivists and librarians, to face the incessant transformation of Information Technology, from the first "Read-Only" Web to the Web 2.0, from the Semantic Web to the challenges posed to historians by research data and Linked Open Data technologies.

L'articolo ricostruisce i rapporti degli studi medievistici italiani con le *digital humanities*, muovendo dalle ricerche condotte sulla documentazione vaticana e su altre fonti ecclesiastiche a partire dagli anni Ottanta del secolo scorso, per passare poi a discutere di alcuni progetti più recenti, basati su diverse tipologie di fonti. Emerge un dibattito a più voci, in cui le esperienze degli storici e degli editori di fonti medievali si intrecciano a quelle maturate nel mondo degli archivi e delle biblioteche, per affrontare le incessanti trasformazioni della *Information Technology*, dai primi siti in rete al Web 2.0, dal Web semantico alle recenti sfide lanciate agli storici dai dati della ricerca e dalle tecnologie dei *Linked Open Data*.

Middle Ages; 20th-21st Century; Italy; Ecclesiastical sources; Digital humanities; Semantic web; Linked Open Data.

Medioevo; Secoli XX-XXI; Italia; Fonti ecclesiastiche; Informatica umanistica; Web semantico; Linked Open Data.

The relationship between historians and computers lies at the crossroad of the evolution of information and communication technologies, the paths of historiography and the transformations of its methods, the modifications in the markets of scholarly communication and in the criteria for assessing the quality of publications. This is not the place to deal with this knot of problems and the changes that the advent of digital era has provoked, even in Italy, in the processes of research, selection, study, interpretation and criticism of the sources – which have been consolidated over the last two centuries, making historiography the study of the past based on verifiable documentary

evidence[1]. Nor is it possible to discuss the transformations in the forms of publication and dissemination of historical knowledge that took place with the emergence of the Internet[2].

I will therefore limit myself to reporting what I believe are the most relevant Italian experiences with digital technologies, in order to let emerge, through their analysis, the metamorphosis of digital source editions in the last twenty years, their ongoing developments and future perspectives. I will present them starting from the closest ones, by theme and character, to the *Repertorium Germanicum* and to the *Repertorium Poenitentiariae Germanicum*, the two German repertoires that offered the starting point for this debate with many voices on the online medieval initiatives in the different European countries. The two German internet editions clearly show that digital humanities, if used in a conscious way, can give effective answers to the needs of historians who have long been involved in large-scale survey and publication projects of historical sources. Since in my contribution the two German databases constitute a constant term of comparison for the Italian experiences, I will now mention their peculiarities.

1. *Curial sources and two German editorial projects*

It is well known that the opening of the Vatican Secret Archives, in 1881, urged the creation of historical institutes of various nations in Rome that initiated the systematic inventorying and publishing of pontifical sources[3]. In this cultural climate, the *Repertorium Germanicum* was conceived by Ludwig Quidde during his stay at the Prussian Historical Institute in the last decade of the nineteenth century, to register, in the form of *regesta*, all the documents that refer to Germans, to German areas and their churches, stated in «the Vatican register series and Cameral holdings, from the Great Schism (1378) to the Reformation (1517)»[4]. The *Repertorium Poenitentiariae Germanicum* was instead launched on the initiative of Ludwig Schmugge in the eighties of the twentieth century, to publish «the full text of all supplications directed to the Penitentiary, relating to people, churches and places in the German Empire»[5]. Both projects are based on curial documentation and use the DENQ (*Digitale Editionen Neuzeitlicher Quellen*) platform, de-

[1] On historical sources in internet still interesting considerations in Vitali, *Passato digitale.*
[2] I have covered this topic in Delle Donne, *Un intreccio di iniziative scientifiche.*
[3] Esch, *Leone XIII.*
[4] Ludwig Quidde was at the Prussian Historical Institute of Rome between 1892 and 1894. He was also a prominent politician in Germany in the early decades of the twentieth century. In 1927 he was awarded the Nobel Peace Prize. On him see the volume: *Friedensnobelpreis und historische Grundlagenforschung.*
[5] The *Repertorium Penitentiariae Germanicum* has just ended with the publication of the eleven volume: *Hadrian VI. 1522–1523.*

veloped by Jörg Hörnschemeyer at the German Historical Institute in Rome[6]. Designed to manage the entire process of document publishing in a network environment, from the reproduction of the original text to its digital re-contextualization, from the online diffusion to the possibility to interrogate the different semantic cores in single documents and in extensive textual *corpora*, DENQ is entirely based on XML (Extensible Markup Language) technologies, from the "declarative" encoding of texts according to the TEI standard (Text Encoding Initiative) to their indexing and management in a native XML database such as eXist, to the use of PHP (Hypertext Preprocessor) and Java modules for the Query-Engine[7]. The unusual foresight of choices made as far back as 2004, which still appear to be the most appropriate ones, should be emphasized.

We can therefore say that the two German repertoires are characterized by three main features: 1) the attention to a geographical area of national and even transnational extent; 2) the absolutely systematic nature of data collection; 3) the advanced level of the technologies involved. In Italian universities and research centers there are no similar projects on Vatican sources, which aim and achieve all the three just-mentioned features, but some excellent initiatives exist that do leverage at least some of these characteristics[8].

2. *Sources for the* Lombardia Sacra *project*

Around 1985 Giorgio Chittolini initiated an extensive project at the University of Milan to identify and publish the papal and notary documentation relating to ecclesiastical people and institutions in fifteenth century Milan and Milanese political society[9]. He aimed to reconstruct the relations between the Roman curia and the Milanese government, the structures of local churches, the activities and operations of diocesan offices and urban chap-

---

[6] DENQ was initially developed together with the German Historical Institute of London: Hörnschemeyer, Schmugge, *Uno spoglio moderno delle fonti vaticane.*

[7] On the project: Matheus, *Grundlagenforschung aus Leidenschaft.* On DENQ characters: Hörnschemeyer, *DENQ.*

[8] Naturally, in this context, I will not speak of the many initiatives concerning Vatican and ecclesiastical sources, which have been carried out in Italy without envisaging any use of information technology; purely as an example, to get an idea of the number of contributions relating to the 13th century only: Paravicini Bagliani, *Il papato nel secolo XIII.* I will not even refer to major projects conceived in other European countries, such as the *Regesta Pontificum Romanorum. Italia Pontificia*, an all-German project even if at first Raffaello Volpini and Luigi Schiaparelli participated (Johrendt, *Papsturkundenforschung in Italien*); in recent years, the Austrian *Monasterium.net* (http://monasterium.net/mom/home) has also involved numerous Italian researchers. For the same reason, I will not talk about the extraordinary digital heritage of the Vatican Secret Archives in Vatican City, which now consists of about 7 million images: < http://asv.vatican.va/content/archiviosegretovaticano/it/attivita/ricerca-e-conservazione/ acquisizioni-digitali.html >.

[9] Chittolini, Belloni, *Fonti notarili.* For the edition of Vatican sources launched under the project: Battioni, *Censimento ed edizione di documenti pontifici.*

ters, the competence and the operating procedures of episcopal courts. Using prosopographic analysis, Chittolini intended to create a *Lombardia Sacra* that systematically mapped churches and monasteries, constituting an *onomasticon* of clerics and religious people during the Visconti-Sforza period in fifteenth century[10]. In other words, he planned to prepare a research tool to collect information and prosopographical data, similar to the repertoires relating to diocesan churches realized in other countries, such as *Germania Sacra*, *Helvetia Sacra*, *Fasti Ecclesiae Anglicanae*, *Fasti Ecclesiae Gallicanae* and others.

The investigation began with the examination of the so-called "chamber documents" concerning the provision of ecclesiastical benefices and, in particular, with the *Liber annatarum*[11], to move on then to Penitentiary documents, to Supplications and to Pastoral visitations[12], thanks also to the collaboration with the Germanic Historical Institute in Rome and the Italo-Germanic Institute in Trento. The systematic research for papal sources in the Vatican Secret Archives, in the State Archives of Rome, in the Archives Nationales de France, was accompanied by an intense exploration activity of the Ducal Archive and the Notarial Archive in the State Archive of Milan, as well as of several local ecclesiastical archives[13]. The documentary examinations have made it possible to identify about five hundred names of ecclesiastics and to gather information on the period of their presence in the main Milanese chapters and often on how they obtained benefices. The results of this research were partly published online in June 2001 in «Reti Medievali Rivista» presenting a tabular list of the Milanese canons in the early Sforza age (1450-1499); also a collection of biographical profiles written by Cristina Belloni, Elisabetta Canobbio, Marzia De Luca, Miriam Ferrari, Marco Lunari, Patrizia Merati, Paolo Ostinelli, Fausto Ruggeri, Elena Salanti, Francesco Somaini, was then published[14].

To make known the results of a wide, systematic and original research, Chittolini and his group have therefore relied on articles and printed books. Therefore, the use of digital humanities tools was limited to structuring some data in a table published in HTML (HyperText Markup Language), the most common "procedural" coding language in the web aimed at the precise defi-

---

[10] Chittolini, I*ntroduzione*. On the project: Belloni, *A proposito di una recente edizione*.
[11] *I "libri annatarum" di Pio II e Paolo II*; *I "libri annatarum" di Sisto IV (1471- 1484)*; *I "libri annatarum" di Innocenzo VIII (1484-1492)*; *I "libri annatarum" di Alessandro VI (1492-1503)*. The *Libri annatarum* have been widely used by individual scholars also outside the Lombard context; I will only mention the fundamental study by Roberto Bizzocchi on the political and ecclesiastical history of Tuscany: Bizzocchi, *Chiesa e potere*. I *Libri annatarum* are also the basis of Li Pira, *La collazione dei benefici ecclesiastici*, on the Dioceses of Southern Italy between the Angevin and the Aragonese age.
[12] I just mention: *Le suppliche alla Sacra Penitenzieria Apostolica provenienti dalla diocesi di Como (1438-1484)*; *I "registra supplicationum" di Pio II (1458-1464)*; *La visita pastorale di Gerardo Landriani*; Belloni, *Visite pastorali milanesi*.
[13] *La rubrica degli atti di Albertolo Griffi*.
[14] Chittolini, Belloni, *Fonti e repertori per la storia milanese*.

nition of the graphic display of documents[15]. Instead, they did not resort to XML, which not only allows to meet the duration and preservation requirements of documents, facilitating the conversion without considerable loss of information, but also permits to define the logical structure of a text and its components with a series of tags, opening up effective research possibilities within historical sources.

3. *Sources in XML: the* Codice diplomatico della Lombardia medievale

Among the participants of the first hour to the Chittolinian project we find, however, Michele Ansani, who, since the end of the twentieth and the beginning of the twenty-first century, was going through one of the most interesting European experimentations with digital documentary editions. To him we owe the *Codice diplomatico della Lombardia medievale (secoli VIII-XII)*, that was started in 2000 to map and publish in a critical edition the Lombard documentary heritage from the first centuries of the Middle Ages up to the mature municipal age[16]. In his imposing work, the documents produced by ecclesiastical bodies (monasteries, chapters, canons, etc.) are clearly predominant[17].

Ansani, who has developed a careful methodological reflection during his intense editorial activity[18], chooses for the markup of texts the "declarative" syntax of the encoding meta-languages known as SGML (Standard Generalized Markup Language) and XML, explicitly providing meta-information on the structure of the text, on its partitions, on the legal documentary functions. He also adopts a system of regularization of person and place names, which allows the automatic construction of indexes, frequency lists and much more. Between adopting the TEI Guidelines and developing an autonomous encoding system, closer to documents, Ansani opts, however, firmly for the second alternative. The reasons for his choice, which is very different from the one made by the editors of *Repertorium Germanicum*, are to be found in the debates on digital humanities that took place in the second half of the Nineties, in which TEI was establishing itself as an encoding model for literary texts, but not yet for other documentary typologies, with respect to which it appeared largely unsatisfactory[19]. Moved by the intention of affirming the needs of the Diplomatic as a discipline that had been consolidated over the

---

[15] The distinction between "procedural" markup languages and "declarative" ones has been widely used in digital humanities for over twenty years: Ciotti, *Il testo*, p. 35.

[16] The project has received funding from both the Ministry of University and Research and the Lombardy Region.

[17] To get an immediate idea, use the *Mappa generale delle edizioni* al *Codice diplomatico della Lombardia medievale* as a path to access to documents: < http://www.lombardiabeniculturali. it/cdlm/edizioni/mappa >.

[18] Ansani, *Diplomatica (e diplomatisti)*; Ansani, *Il Codice diplomatico digitale*; Ansani, *Edizione digitale*; Ansani, Ghignoli, *Testi digitali*.

[19] On the characteristics of the TEI in the second half of the nineties: Siciliano, *Codifiche XML*.

centuries, Ansani preferred to develop a more agile DTD (Document Type Definition), characterized by a high number of markers with few attributes. Therefore, in the Ansani's text-encoding we find the information and the meta-information that are not immediately visible to the human eye but that are used by the search engine TReSy (Text Retrieval System). Developed in PHP and in JQuery by the Scuola Normale Superiore in Pisa, it operates on the textual markup and allows unlimited targeted searches of different type and purpose, with chronological, territorial and typological restrictions; the search can also be limited to specific parts of the document, to its formal structure and its internal subdivision.

In Italy the DTD developed by Ansani has been the subject of extensive discussions and has constituted, in fact, until a few years ago, a standard to confront. It was taken up and adapted into many digital edition projects conceived around the 2000s within the group Reti Medievali, such as those dedicated to the treaties between the Commune of Pisa and the South of France (the so-called *Midi*), to the Statutes of Tuscan, Lombard and Venetian cities, to the Angevin documentation[20]. Nonetheless, shortly thereafter, the TEI would have established itself as a reference model for historians and for diplomatists, after the release of the P5 guidelines in 2007, which were much more flexible than the previous ones[21].

In the text encoding it is in fact always necessary to find the right balance between the specific interests of the historian, actively engaged in research and tendentially oriented to a very personalized and granular structuration of data, and the need to standardize them to enable their interoperability with other textual databases, as well as their sharing with a wider circle of scholars.

4. *The ecclesiastical sources and the archivists: the* Ecclesiae Venetae *project*

There have also been Italian projects decidedly oriented towards compliance with international standards, such as *Ecclesiae Venetae*, an initiative that is very different from the previous ones, not only for the chronological period considered, which goes well beyond the Middle Ages, reaching up to the twentieth century, but also because it was born within cultural heritage conservation institutes, as part of a vast operation to safeguard and enhance the historical archives of Venetian dioceses, promoting the census, invento-

[20] *Il fondo pergamene Pisa-Midi*; Zorzi, *L'edizione digitale degli statuti*; for the Angevin documentation see below, note 78.
[21] The first strong sign of attention for the TEI, within the research group that had referred to the experience of Ansani for editions of city Statutes, is to recognize in 2007; the TEI is in fact adopted for the 1264 Vicenza Statute: Salardi, *L'edizione digitale dello Statuto di Vicenza del 1264*.

rying and cataloguing of manuscripts[22]. After all, in the world of archives, an intense debate on the standardization of the descriptions had already begun in the mid-nineties of the twentieth century, always accompanied by the awareness that it is always necessary to respect the deep-rooted peculiarities of the individual archives. This paved the way for the design of advanced information systems, based on a rigorous archival description conceived as a specific tool for formalized communication of information on archives, record creators and historical contexts of production[23].

The idea of giving life to a project related to Venice and Veneto ecclesiastical documentation found its first realization in 1989 with the creation of ARCA (Historical Archives of the Church of Venice), the database dedicated to the Historical Archive of the Patriarchate of Venice, which assumed definite characteristics of diocesan archive as many parish fonds were deposited into it. Besides ARCA, *Ecclesiae Venetae* was started in 1996 for the dioceses of Padua, Vicenza, Verona, Treviso and Vittorio Veneto[24]. Both projects, which after the year 2000 would have merged, were born by initiative of Francesca Cavazzana Romanelli, with the support of the Ministry of Cultural Heritage and Activities, the Veneto Region, the diocesan archives and the Gladys Krieble Delmas Foundation[25]. *Ecclesiae Venetae* is not the only Italian archival project dedicated to ecclesiastical documentation[26], but it is undoubtedly the largest one regarding the number of medieval surveyed documents and the most influential one, thanks to the uncommon ability of Cavazzana Romanelli to transform it into an opportunity for broad interdisciplinary discussion with historians of ecclesiastical and secular institutions, historians of culture and law, paleographers, diplomatists, archivists and scholars of the history of the book[27]. *Ecclesiae Venetae* now contains over 80,000 record entries and

---

[22] *Ecclesiae Venetae* < http://siusa.archivi.beniculturali.it/cgi-bin/pagina.pl?RicProgetto=ev >.
[23] Ricci, *Gli standard internazionali*.
[24] Cavazzana Romanelli, *L'archivio storico del Patriarcato di Venezia*; Cavazzana Romanelli, *Il progetto "Ecclesiae Venetae"*.
[25] On the reasons and the evolution of the project: Levantino, *Alla scuola di "Arca" e di "Ecclesiae Venetae"*. On the professional and human figure of Cavazzana Romanelli: Varanini, *Archivi ecclesiastici*.
[26] I will only mention the numerous works dedicated to the Apulian ecclesiastical institutions by Domenica Porcaro Massafra, although mainly oriented towards the modern age; they are likewise present in SIUSA, the archival information system of which I will tell soon. I also point out that in 2016 the National Office for Ecclesiastical Cultural Heritage and Building (Ufficio Nazionale per i Beni Culturali Ecclesiastici e l'edilizia di Culto) of the Italian Episcopal Conference (Conferenza Episcopale Italiana) launched the Ecclesiastical Archives Project (Progetto archivi ecclesiastici) to coordinate and support «the activity of reorganization and computerized inventorying of ecclesiastical archives, and in particular of diocesan archives, with the aim of carrying out a map of them and facilitating their use and access [...]»; it further provides for interoperability between the information system of the CEI, BeWeb, with SIUSA and SAN (I will tell about this information system soon): < https://bce.chiesacattolica.it/progetto-archivi-ecclesiastici/ >.
[27] For instance, the very extensive documentation on marriage trials, collected over long years, offered Cavazzana Romanelli the occasion to organize together with the Italo-Germanic Historical Institute in Trento numerous seminars and conferences, published between 2000 and

can be searched through various search keys, both on SIUSA platform (Unified Information System for Archival Superintendencies) and, from 2011, on SAN platform (National Archival System). Inside *Ecclesiae Venetae*, there are the descriptions of conservation institutions and all the fonds contained (archives of the Curies, but also fonds of various types aggregated or annexes, such as bishop's mensa – the incumbents of episcopal office –, parishes, private papers of individual bishops or ecclesiastics, confraternities and associations), the series and any other sub-fonds, up to the individual archival units. An analytical account is given for all the curial archives and also for the other areas of documentation each time identified, on the basis of the international standards for archival description ISAD (G) and ISAAR (CPF)[28]. Originally the database was created according to the paths and the computer program of the *Anagrafe* project; in June 2016 it was entirely migrated to the new open source *Archimista* software[29], to facilitate interoperability with SAN. More in general, a large number of programs have become widespread in Italy over the last decade, both for the archival inventory, and for the management of information systems, prepared by the State Archival Administration[30]. However, even if the software tools used are different and dissimilar are some descriptive paths used, they all share the same conceptual model of document representation, based on the most famous international archival standards[31].

These features also characterize the softwares used for the reordering, the cataloging and the description of *Ecclesiae Venetae*, while the SIUSA archival information system and the SAN portal intervene at a later stage to harmonize, in view of fruition, the various information describing the archival material at different levels, allowing flexible search paths, capable of restoring both the hierarchical structure of the fonds and the complexity of the archival context. SIUSA, which proposes itself as a primary access point for researching and consulting the Italian public and private non-state archival heritage, preserved outside the State Archives, has a very complex hardware and software structure, capable of managing a huge amount of XML data through an Oracle database, while the language used for the development of the data en-

---

2006 by Silvana Seidel Menchi and Diego Quaglioni: *I processi matrimoniali*. The conference *La memoria delle chiese venete. Archivi diocesani e storiografia* (Padova, Monastero di S. Giustina, January 29th, 1999) should still be remembered; Attilio Bartoli Langeli, Paola Benussi, Liliana Billanovich, Lucio Bonora, Sante Bortolami, Giorgio Chittolini, Giorgio Cracco, Giuseppina De Sandre, Gilda Mantovani, Domenica Porcaro Massafra, Luciano Osbat, Antonio Rigon, and Gian Maria Varanini took part in it: Orlando, *Convegno di studi.*
[28] On the two standards, see the pages of the Central Archives Institute – ICAR (Istituto Centrale per gli Archivi – ICAR), with the related documentation: < http://www.icar.beniculturali.it/index.php?id=54 > e < http://www.icar.beniculturali.it/index.php?id=55 >.
[29] On Archimista characters: Brunetti *et al.*, *Archimista.*
[30] To get an overview of the variety of software used in archives: Aprea, *Uno sguardo sugli strumenti digitali.*
[31] On the standards: Vitali, *La descrizione archivistica.*

try system is JAVA[32]. Since 2011, SIUSA data are periodically loaded into the SAN catalog, which provides instead a unified access to all the Italian archival resources made available on the web by information systems, databases and digital research tools developed at national, regional and local level by the State, the Regions and other public and private entities. By consulting the SAN catalog, it is possible to know which archival records are preserved in Italy, who produced them, where they are stored and how they are accessible. It also offers the possibility of knowing which fonds and series can be consulted directly on the web, in digital libraries and archives created by State archives and by other public and private archival institutions[33]. We could therefore say that SAN is superordinate to SIUSA because it offers researchers a general map of the national archival heritage, pointing them towards more detailed sources of information present in archival systems such as SIUSA.

In SAN the exchange paths, used to generate the XML files that contain the data to be exported, are based on the most common international format standards in archives: EAD (Encoded Archival Description), for the description of archives, and EAC (Encoded Archival Context), for the records creators[34]. The digital documents, accessed via SAN, are provided with descriptive and managerial metadata encoded in XML in a format based on METS (Metadata Encoding and Transmission Standard), the standard developed by the Library of Congress to provide descriptive, structural, technical, and administrative information on digital objects. A few years ago, in 2011, the Italian Norms for the processing of archival authority records for institutions, individuals, families, NIERA (EPF), were released. They represent the transposition of ISAAR (CPF) in Italy and constitute the first descriptive standard of national scope developed in Italy for the realization of homogeneous and comparable descriptions of institutions, people and families, for their univocal identification in every possible role they could have in archival documentation[35]. A SAN development project is also underway, aimed at its publication in Linked Open Data[36].

All these projects are top-level achievements as they pay the utmost attention to the international standards and the high quality of technologies. Nevertheless, they require constant supply of human and financial capital to ensure their proper functioning, while in Italy, in the archival and library sectors, we have been witnessing for too many years a continuous contraction of available resources.

---

[32] The hardware and software structure used for the realization of the SIUSA system is analytically described in Ronca, Romano, *SIUSA*.
[33] On the SAN architecture: Vitali, *The San Portal*.
[34] See the pages of the Central Institute for Archives – ICAR dedicated to ISAAR (CPF), ISAD (G) and METS.
[35] Direzione generale per gli archivi, Commissione nazionale per l'elaborazione del codice normativo per i soggetti produttori d'archivio, *Niera (EPF)*.
[36] < http://san.beniculturali.it/web/san/dati-san-lod >.

It should also be reiterated that the tools of archival description are much more than digital molds or mechanical transpositions of traditional paper tools in a network environment. If their new potentials are undoubted, because they allow to go beyond the borders of a single institution, outlining a virtual space in which descriptions of archives preserved in different institutions can be integrated and put together, not only at local and national level but also internationally; such new and dizzying perspectives require from archivists, who produce the tools, and from historians, who use them, particular precautions. A few years ago, Stefano Vitali observed how the major thrusts for the transformation of archival description practices concern above all the issue of the "production context" of documents. If in the past the records creator (institution, family or person) was conceived as a sort of attribute of the documentation, a component of the set of information that formed its description, now it becomes an entity in itself, whose history, specific characters and structure should be illustrated. In other words, it becomes the fulcrum around which are also aggregated different archives. This new approach, which makes it possible to give a better account of typical phenomena of the archival reality, such as the dispersion among different institutions of the documents produced by a single records creator, determines, more strongly than in the past, an overemphasized centrality of the information on the production context. This perspective reverberates on the architecture of research tools and computerized archival systems, entailing «a necessary reconfiguration of the ways historians conceive and practice research in the archives and interpret its results»[37]. The excessive emphasis on the context of documents production has therefore involved, in some cases, less attention to the actual documental context. Of course, there are examples of migration strategies from analog to digital which are rigorous in the archival description and respectful of the documentary and production context, such as the *Mediceo avanti il Principato* project and the *Diplomatico* one, carried out at the Florence State Archive by Francesca Klein and Stefano Vitali[38].

However, it should never be forgotten that the acquisition of a document in image format never produces a sort of double mirror image in which the original would be immediately reflected, since the decision to select a particular sampling ratio, with a more or less high resolution number of dot per inch (DPI), can lead to a loss or an increase of information with respect to the original. An enrichment or impoverishment of the source can also derive from its possible association with materials of various kinds, from its inclusion in a network of hypertextual relations with other sources, with transcriptions, critical editions and information tools. A few years ago, to underline the distance between the digital representation of sources and the analogical

---

[37] Vitali, *Passato digitale*, p. 77.
[38] Archivio di Stato di Firenze, *Mediceo avanti il Principato*: < http://www.archiviodistato. firenze.it/map/ > and *Diplomatico. Pergamene (sec. VIII-XIV)*, < http://www.archiviodistato. firenze.it/pergasfi/ >.

form from which they originate, Jean-Philippe Genet, Andrea Zorzi and Stefano Vitali proposed to define them as "metasources"[39] – a concept aimed at highlighting the recontextualization process which inevitably conditions the understanding and the interpretation of documentary resources published on the net. Nonetheless, I strongly believe that if the term "metasource" is intended to underline the digital transposition of a source and its recontextualization, which make it «a separate document [...] that is now profoundly different from the original», we could then legitimately use the same term for any reproduction or printed edition of a source – without prejudice to the different potentialities of expansion of the "contexts" made possible by telematic networks.

Going back to *Ecclesiae Venetae*, it should also be pointed out that both software used for cataloging, first *Anagrafe* and then *Archimista*, unlike DENQ, were not originally designed for the management of digitized cultural heritage and for the creation of digital libraries, even if *Archimista* provides some minimal functionalities, as it allows to connect digital objects to each of the main entities (archives, archival units, creators, custodians of the record)[40]. Undoubtedly, *Archimista* is an open source software that can be developed and enriched with new modules and new features, but at present it certainly does not represent the ideal environment for a critical edition of sources in their entirety, in TEI-XML format or not.

The experience gained in the world of archives seems to suggest to historians, with clarity, that it is absolutely essential to combine the representation of specific and "unique" characters of fonds, series, subseries and archival units, with the elaboration of formalized and standardized descriptions of elements that are potentially common to multiple archives, on the basis of norms and conventions shared on an international scale, to achieve their full integration into "archival systems", capable of ensuring the widest possible dissemination.

5. *Vatican Codices for a "data-driven history"?*

In this context an experiment conducted on the Vatican documentation has to be mentioned, which, in perspective, could considerably enrich the digital editions now available on the web as simple images collections with new contents. It is the *In Codice Ratio* project[41], launched in 2016 by paleographers, philologists, historians and computer engineers from the University of Roma Tre, in collaboration with the Vatican Secret Archives. Funded by the Lazio Region, it is aimed at developing technologies for the automated tran-

---

[39] Genet, *Source, Métasource*; Zorzi, *Documenti, archivi digitali, metafonti*, from which the quoted sentence is taken; Vitali, *Passato digitale*, p. 107.
[40] Brunetti *et al.*, *Archimista*, p. 12.
[41] < http://www.inf.uniroma3.it/db/icr/ >.

scription of a vast *corpus* of over eighteen thousand pages that make up the 43 Vatican registers of the thirteenth century, all written «with the same script: the so-called *Cancelleresca*»[42]. The project moves from the observation that the digital libraries mainly provide only images of manuscripts and that the current Handwritten Text Recognition (HRT) procedures are extremely slow and laborious because they require the constant intervention of paleographers. *In Codice Ratio* intends instead to realize «a scalable, end-to-end transcription work-flow [...] based on fine-grain segmentation of text elements into characters and symbols, with limited training effort», starting from «a *corpus* of letters by pope Honorii III»[43]. The project, presented at the 24th international conference on *Knowledge Discovery & Data Mining (KDD)*, aroused strong international interest[44]. Its relevance lies in the application to the Vatican manuscripts the most recent deep learning techniques (convolutional networks, U-Net, sequence2sequence systems), that is algorithms for the automatic learning of the multiple hierarchical levels of data features and their representation. In other words, the *In Codice Ratio* uses a system «that first segments every word into small (possibly overlapping) fragments, then recognizes characters in fragments, and finally the entire word»[45].

Although the role of artificial intelligence experts may seem prominent, it should be stressed that the project has an interdisciplinary character and that the contribution of paleographers and historians remains constant at various levels. The creators of *In Codice Ratio*, however, place their work in the context of a "Data-driven history", which they claim has shown for some years an enormous potential of quantitative lexical analysis of large digitized textual corpora «for the study of cultural change and continuity»[46]. In this regard, reference is made to "culturomics", the computational lexicology developed in 2010 by two researchers from Harvard University, Jean-Baptiste Michel and Erez Lieberman Aiden, using large amounts of textual data drawn from a Google Books *corpus* of over 500 billion words; in the same way, mention is made of the "distant reading" proposed by Franco Moretti, the founder of Stanford Literary Lab in 2010 and one of the most influential literary critic[47].

The idea that historical reconstructions should be based on data and that large amounts of them, if observed from afar, could make emerge new problems is largely acceptable. Less convincing is the belief that history is

---

[42] Ammirati *et al.*, *In Codice Ratio*.
[43] Firmani, Merialdo. Maiorino, *In Codice Ratio*.
[44] Firmani *et al.*, *Towards Knowledge Discovery*. And its review: *AI tackles the Vatican's secrets*.
[45] Ammirati *et al.*, *In Codice Ratio*, p. 4. Of course, projects with similar aims are underway in Europe: For instance, the Himanis one, which aims at indexing the digital images of ca. 68.000 manuscript charters and documents produced by the Royal French Chancery from 1302 to 1483, uses artificial intelligence to develop «cost-effective solutions for querying large sets of handwritten document images»: < https://www.himanis.org/ >.
[46] Ammirati *et al.*, *In Codice Ratio*, p. 3.
[47] The references are respectively to Michel *et al.*, *Quantitative Analysis*, and Moretti, *Distant Reading*.

"data driven", as if an archive manuscript did not require much more than the recognition of handwritings in order to give rise to the process of historical knowledge construction. Moreover, it should be remembered that for Moretti and his group the effectiveness of "distant reading" is conditioned by the respect of certain conditions[48]: it must be the result of a collective and collaborative activity, based on hypotheses and models made explicit, and throughout inspired by the Popperian principle of falsifiability; scholars should therefore pay attention not only to the results obtained, but also to the procedures for their achievement. Furthermore, according to Moretti, the data does not speak for itself and the "computational criticism", as he prefers to call the "digital humanities", should always be moved by authentically "humanistic" interests, accompanying to the "distant reading" the "close reading" of texts which represent borderline cases, particularly significant because they differ in part from the "patterns" that emerged from the quantitative examination of the sources[49]. Likewise, the *Digital Humanities Manifesto 2.0* of 2009[50] already both included the quantitative and qualitative approach, the "distant reading" based on the automated reading of a large number of texts and the "close reading" that focuses on the careful analysis of a single text, integrating Data Science within the "digital humanities", to create a "human-centered Data Science", focused on the interpretation of machine learning models and on the recognition of an active role for human input into algorithms[51].

Moreover, for some years big data, open data, linked data and linked open data projects have been announced in many countries and domains. Only a few can be remembered here. Many public administrations, driven by the European Community (European Digital Agenda, Action 3), have chosen to make freely accessible on the web as open data the enormous wealth of information they collect and hold by virtue of their institutional roles. In Europe, some public administrations have also identified in the standard technologies of the semantic web, and in particular in the open linked data model, the indispensable tools to give the data (open or not) an identity, to make them connectable and semantically interoperable with each other, as called for by the European Digital Agenda[52].

A strong boost to data sharing and interoperability is also present in the context of scientific research, since the European Commission, the Max-Planck-Gesellschaft and the British Her Majesty Treasury, with their declarations and recommendations, with their policies and their funding programs

---

[48] See the essays collected in the volume Moretti *et al.*, *La letteratura in laboratorio.*
[49] The concept of "pattern" has a key role in Moretti's reflection, because it refers to a data disposition or a data relationship, «behind which there is a form that is repeated» (Moretti et al., *La letteratura*, p. 72).
[50] < https://mbf.blogs.com/files/the-digital-humanities-manifesto-2.0-2.pdf >.
[51] On these topics the bibliography is very broad. I will only mention Bamman, *Interpretability.*
[52] See the European and Italian normative references, the recommendations and the guidelines on the website dedicated to the Open Data by the Agency for digital Italy (Agenzia per l'Italia digitale): < https://www.dati.gov.it/ >.

such as Horizon 2020, encourage scholars to consider the research cycle as a whole and to make it public on the net in its entirety, according to the international standards: from the collection to the classification of data, from their relational structuring to their textual presentation, up to their re-elaboration in the discursive order of an article or a book[53]. In many disciplinary communities, including those in the humanities and social sciences[54], careful consideration has therefore been given to the datasets and to the characters that the data repositories should have, to the way accessibility, interoperability and long-term conservation of data should be ensured, to how they can be made quotable and reusable by others, to the peculiar forms their peer review should assume[55].

Also the research data should be provided with standardized metadata (descriptive, structural, administrative and rights management), which allow their identification and retrieval in a single integrated environment, characterized by high "usability" and full interoperability with the major aggregators of scientific contents operating online, in order to guarantee the widest possible dissemination and the maximum re-use. The most advanced achievements have been made in the field of cultural heritage institutions, and in particular in the world of libraries and archives, in which there have been significant technological innovations aimed at realizing the integration of heterogeneous information sources into a single research platform, able to increase and optimize the possibilities of information retrieval for end-users. From the 1st July 2012, the European digital library *Europeana* exclusively

---

[53] The importance of research data was already highlighted by the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*, promoted by the Max Planck Gesellschaft in 2003: < http://openaccess.mpg.de/Berlin-Declaration >. On the centrality of Open Research Data for the European Commission: Ramjoué, *Open Research Data*; Ramjoué, *Opening up*. An overview of the various initiatives of the European Commission to promote the publication of research data can be found at the URL < http://ec.europa.eu/regional_policy/it/policy/evaluations/data-for-research/ >. For the United Kingdom Her Majesty's Treasury, *Our plan for growth*, p. 46. On the "good practices" of Open Linked Data, Bizer, Heath, Berners-Lee, *Linked Data*; on Open Linked Data and the standard Resource Description and Access (RDA): Bianchini, Guerrini, *Introduzione a RDA*, pp. 52-60.
[54] I just remember the recent article: Borghi *et al., Support Your Data*.
[55] A summary of the various initiatives in an interdisciplinary perspective in Kratz, Strasser, *Data publication*. For archaeological repositories, which provide complex data validation procedures, also through peer review: *The Digital Archaeological Record* (tDAR): < http://www.tdar.org/ >; *Open Context. Web-based research data publication*: < http://opencontext.org/ >. For political and social sciences: *Inter-university Consortium for Political and Social Research* (ICPSR): < http://www.icpsr.umich.edu/icpsrweb/landing.jsp >. So-called Data Journals have also been created for archeology: «Internet Archaeology» < http://www.internetarchaeology.org/ > and «The Journal of Open Archaeology Data» < http://openarchaeologydata.metajnl.com/ >. In Italy, on October 29th, 2014, on the occasion of a meeting dedicated to the theme of Open Data in archeology, during the XVII edition of the "Borsa del Turismo Archeologico" in Paestum, the Open Data Archaeological Manifesto (*Manifesto Open Data Archeologici*, MODA) was presented, promoted by the "Laboratorio di Metodologie Applicate alla Predittività del Potenziale Archeologico" (MAPPA) of the University of Pisa and by the "Gruppo Archeo & Arte 3D" of the University of Rome La Sapienza: < http://www.archeofoss.org/2015/11/moda-manifesto-open-data-archeologici/ >.

exposes metadata in Linked Open Data format under public domain license Creative Commons Zero (CC0), which allows free reuse. In Italy, the main achievement in Linked Open Data is represented, with over two million records, by SHARE Catalog, the platform for the creation and publication of catalogs and open access electronic resources created by the universities of Naples Federico II, Naples L'Orientale, Naples Parthenope, Salerno, Sannio, Basilicata, Salento, together with the company @Cult[56].

Finally, it should be remembered that the possibility of freely accessing the entire cycle of scientific research and its results without incurring subscription costs is one of the main objectives of the *Open Science*, on whose importance the recommendations of the European Commission insist starting at least since 2007, up to the recent identification of three major areas of intervention (*Open access to publications*, *Open research data* and *Open scholarly communication*) and the proposal to implement a road map for the realization of the *European Science Cloud*, adopted by the ministers of European Union on May 29th, 2018[57]. The European Commission Recommendation 2018/790 on access to scientific information and its conservation, which incorporates and reinforces the Recommendation 2012/417, asks member states to provide binding policies to ensure open access and conservation of publications and research data[58]. On May 27th, 2016, the European Competitiveness Council, which brings together the European Union's research, innovation, trade and industry ministers, decided that by 2020 all the results obtained with European funding will have to be made immediately available in open access[59]. On April 4th, 2019, the European Parliament approved, with

---

[56] On *Europeana Linked Open Data*: < https://pro.europeana.eu/page/linked-open-data >. On SHARE Catalogue: Delle Donne, Possemato, *SHARE-CATALOGUE*. Not long ago, the MIBACT (Ministero per i beni e le attività culturali) has launched a similar project at the Central Institute for the Single Catalog (Istituto Centrale per il Catalogo Unico) of Italian libraries, called Meta-FAD: Cerullo, *MetaFAD*. A recent attempt to outline an ontology, with a controlled vocabulary, for the representation of personal relationship in the cultural heritage domains (museums, archives and libraries): Carriero, Daquino, Tomasi, *Convergenze semantiche*.

[57] European Commission, *Communication*, in particular the paragraph 3.1 *A system in transition: new markets, services and players*, which introduces the idea of the *continuum* in the space of scientific information, from raw data to publications, to be made fully accessible on the web. The stages that lead European bodies to promote the Open science are summarized in the 2016 document: Council of the European Union, *The transition towards*. The *Open Science Monitor*, whose implementation was commissioned by the European Commission in 2015, effectively represents the different areas of open science and its developments: < https://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor >. For the *roadmap*: European Commission, *Implementation Roadmap*. For the endorsement of European ministers: < https://ec.europa.eu/info/news/eu-ministers-endorse-commissions-plans-research-cloud-2018-may-29_en >.

[58] European Commission, *Recommendation (EU) 2018/790*.

[59] Council of the European Union, *The transition towards*. A goal which was recently restated by Robert-Jan Smits, former Director-General for Research and Innovation of the European Commission: see the interview given to Roberts, *Open access*. On September 4th, 2018, the goal of making all publications in journals freely available from 1st January 2020 was reiterated in the document called *Plan S*, drawn up by Science Europe together with the European Commission and eleven research funding bodies: Coalition S, *Plan S*.

a large majority of votes, the *Directive on Open Data and Public Sector Information*, which obliges member states to

> support the availability of research data by adopting national policies and relevant actions aiming at making publicly funded research data openly available ("open access policies") following the principle of open by default and compatible with FAIR (Findable, Accessible, Interoperable, Reusable) principles[60].

For historians, the "research data" may consist of text documents, such as notes, summaries and transcriptions of documents, of spreadsheets, databases, multi-dimensional visualizations and models, images, maps, software and applications, audio and video recordings[61]. In the coming years, medieval studies will have to measure themselves closely with the complex problems of data management and publication on the net. Editors of medieval sources will therefore be called to deal not only with the TEI and the descriptive norms and conventions developed in the world of archives and libraries, but also, more generally, with the semantic web and the Linked Open Data technologies and standards.

6. *Digital humanities and Text Encoding Initiative: two Italian projects*

To grasp the extreme vitality of the Italian experience of digital humanities, however, we must take our eyes off the initiatives born around the ecclesiastical archives and look at the myriad of small and large scientific projects, flourishing within universities and research centers, with the most diverse purposes. Furthermore, Italian medieval studies, in a broad sense, have intertwined with computer technology a dense network of relationships for over fifty years, starting from the pioneering computational linguistic experiments initiated by Roberto Busa with the *Index Thomisticus* in the second post-war period[62]. Then followed years of intense experimentation, also linked to the institution, from the eighties of the twentieth century, of master's and specialization courses in digital humanities, first at the University of Rome La Sapienza, then in those of Florence, Naples Federico II, La Tuscia Viterbo, Pisa, Bologna etc.[63]. Over the past two decades, there have also been many research

---

[60] European Parliament, *Directive on Open Data*, Article 10.
[61] To get an overview of the research data management in social sciences and humanities, within a single university: Schöpfel, Prost, *Research data management*.
[62] Adamo, *Busa*.
[63] Among the historians, I remember Guido Abbattista, Michele Ansani, Pietro Corrao, Alessandro Cristofori, Roberto Delle Donne, Antonella Ghignoli, Rolando Minuti, Serge Noiret, Enrica Salvatori, Pierluigi Totaro, Stefano Vitali, Andrea Zorzi; in a wider disciplinary context, Tito Orlandi, Alberto Cadioli, Paola Castellucci, Fabio Ciotti, Giuseppe Gigliozzi, Raul Mordenti, Franco Niccolucci, Mario Ricciardi, Gino Roncaglia, Francesca Tomasi. The Italian Association for Humanistic Informatics and Digital Culture (Associazione per l'informatica umanistica e la cultura digitale) brings together many scholars of digital humanities.

projects relating to the various aspects of the Middle Ages, which involve the creation of digital archives of texts and / or images. If we take as indicators the data of the Italian Ministry of University and Research (MIUR) related to Research Projects of National Interest (PRIN) funded between 1999 and 2015[64], we observe that about one hundred of the more than three hundred approved projects in medieval studies include the construction of data banks of various kinds, from simple relational databases to the creation of tables connected with spatial georeferenced data via internal links (Geographical Information Systems), to text editions in XML format, to databases of images provided with more or less standardised descriptive and structural metadata. However, no project represents the development of centuries-old initiatives such as the *Repertorium Germanicum.* The longest running Italian project among those still in progress is in fact *Medioevo Latino. Bollettino biblio-grafico della cultura europea da Boezio a Erasmo (6th-15th centuries),* which represents the digital outcome of an innovative enterprise undertaken in the analog world in the 1970s by Claudio Leonardi[65]. The origins of other projects are instead recent and, at best, cannot be traced back earlier than the nineties of the last century.

I will limit myself to mention only some of them that I consider, for reasons that will be evident later, to be particularly relevant, starting with the accomplishments that for technical choices are closer to the *Repertorium Germanicum* and the *Repertorium Poenitentiariae Germanicum,* and then moving on quickly to those more strongly oriented to outline an answer to the complex problems posed by the semantic web.

First of all, I remember the *Vercelli Digital Book,* the online edition of the *Codex Vercellensis,* preserved in the Chapter Library of the Cathedral of S. Eusebio of Vercelli, to which Roberto Rosselli del Turco worked tenaciously for about twenty years. It is a manuscript dating back to the end of the 10th century, containing miscellaneous religious works, in verse and prose. It is a codex of great importance for the studies on Anglo-Saxon language and literature because it is one of the four manuscripts that preserve about 90% of all poetic production in ancient English and is the only one not kept in England[66]. The online edition of the manuscript presents high resolution images, and also offer the digital restoration of damaged sheets; it offers, as well, a selection of texts, both in poetry and prose, presented according to two levels of edition: diplomatic and interpretative[67].

---

[64] < http://prin.miur.it/ >.
[65] It is accessible along with other data banks and journals published by the International Society for the Study of Latin Middle Ages (Società Internazionale per lo Studio del Medioevo Latino) and the Ezio Franceschini Foundation on the platform *Mirabile. Archivio digitale della cultura medievale* < http://www.mirabileweb.it >.
[66] The other three codices are the *Exeter Book,* the *Cotton Vitellius* A XV and the ms. *Junius*: Rosselli Del Turco, *The Digital Vercelli Book.*
[67] *Vercelli Book Digitale*: < http://vbd.humnet.unipi.it/beta2/ >.

Initially the *Vercelli Digital Book* was based on a TEI P4 encoding scheme specifically modified for diplomatic editions of manuscripts; later on, it adopted the P5 version of the TEI, while noting some limitations. The documents currently available in TEI contain the data of the edition, in full or by using inclusion tags in the "main" document to automatically include other XML documents managed with XInclude. Through a transformative style sheet (XSLT), documents are made visible in a Web application called EVT (Edition Visualization Technology), which can be loaded on a server and made available to end-users[68]. *Vercelli Digital Book* also offers tools for the navigation and the study of images and texts, from the magnifying glass to simple filters, from double-page navigation, to the possibility of comparing different versions of the same image, up to the linking of lines of text encoded in TEI to the respective lines in manuscript images. EVT is developed on the basis of open and well documented standards, such as XML, XSLT, HTML, CSS and Javascript; it is also distributed as open source software, to facilitate its reuse and further processing.

It should be emphasized that the Digital Vercelli Book was made entirely by Roberto Rosselli Del Turco, along with a few collaborators and students, whom he introduced to computational philology in the Digital Humanities course he gave at the University of Pisa[69]. It is therefore a realization entirely matured within a department of humanities, both for the encoding of the text in XML and for its publication on the web.

The *Archivio della Latinità Italiana nel Medioevo* (ALIM) was born instead as inter-university project between the universities of Siena-Arezzo, Verona, Palermo, Venice Ca' Foscari, Naples Suor Orsola Benincasa and Basilicata. Conceived in the nineties of the twentieth century, it

> intends to provide online access to all the Latin texts produced in Italy during the Middle Ages. For several centuries, in fact, Latin represented the only language in which, in addition to historical documentation, many of the major creations of thought, science, and literature of the Middle Ages were expressed. Even when national languages imposed themselves in written form, Latin never lost its role and prestige as transnational language[70].

Alongside the literary works we also find diplomatic collections, such as the *Codex diplomaticus Cavensis*.

The IT architecture of ALIM is more complex than that of the *Vercelli Digital Book*, because it was designed and developed by using the Muruca open source framework[71], a modular system for creating, managing and pub-

---

[68] Rosselli Del Turco *et al.*, *Edition Visualization Technology*.
[69] Rosselli Del Turco, *The Digital Vercelli Book*, p. 12. EVT was also used for the Latin texts of the *Codice Pelavicino* and for those in Arabic of the project *Tarsian*: Rosselli Del Turco, *The Digital Vercelli Book*, p. 16. The digital edition of Rotari's Edict with EVT is underway: Buzzoni, Rosselli del Turco, *Verso un'edizione*.
[70] On the project: Ferrarini, *Alim ieri e oggi*.
[71] Muruca is created by the Pisan IT services company Net7: < http://www.muruca.org >. For the IT architecture of the Alim2 platform: < http://it.alim.unisi.it/il-progetto/ >

lishing digital libraries. The application backend, through which it is possible to insert, manage and publish texts, images and data in the digital library, is based on LAMP stack (Linux Apache MySql PHP) and uses the Symfony framework for the authentication to the database and the management of the entities present on the platform. For the full text search engine, an Apache Solr module is used, written in Java, which integrates with the MySql database to generate indexes of the full text. DENQ, the software used for the *Repertorium Germanicum* and the *Repertorium Poenitentiariae Germanicum*, is based on eXist, which performs similar functions using the Lucene search library that creates the indexes starting from the XML files. ALIM manages transcriptions with OxGarage, which allows the conversion of documents encoded in TEI-XML in different formats: txt, pdf, html.

Differently from the *Vercelli Digital Book*, realized in its various phases by a very small number of people, ALIM has involved several research groups, which have worked collaboratively to implement the text encoding in XML on their own, while they have resorted to a private IT company for the development of the platform[72].

The *Vercelli Digital Book* and ALIM, in different ways, have therefore tested the main potentialities of Web 2.0 for digital editions.

7. *Ongoing developments and future perspectives: Semantic Web and Linked Open Data for historical studies*

Otherwise, the so-called Web 2.0 and its developments did not offer only an environment in which every single user can publish and share self-produced content. Much more, Web 2.0 offered an opportunity to shift in the direction of a strictly interconnected research community, oriented not only towards the individual production of research content, but also towards its active and collaborative dissemination and evaluation.

In the last decade, the proliferation of content producers, able to easily publish in internet structured information, was so intense, that we need semantics in order to organize, classify and retrieve such information. In adding semantics (i.e. metadata) to the new web, we have mainly to face a formal approach: XML-based, well-structured ontologies are required in dealing with uniform, authoritative collections of information, such as archives, libraries, structured texts and corpora, in order to create machine-readable semantics. Like the web of hypertext, the semantic web or web of data is constructed with documents on the web. Unlike the web of hypertext, where links are relationships anchors in hypertext documents written in HTML or in XML,

---

[72] This is Net7, which was born from the collaboration of computer scientists and humanists and is very active in participating in competitive bids of the European Community in the field of cultural heritage: < https://www.netseven.it >.

Semantic web refers to a set of best practices for publishing and connecting structured data on the Web: the so-called linked data.

So far, researchers and scholars were mainly left alone to find the most appropriate solutions, hopefully in a collaborative way. Nevertheless, neither *Vercelli Book Digitale* nor ALIM explicitly faces the problems posed by the semantic web.

Instead, Reti Medievali Open Archive was created in deep synergy with the world of libraries and in a global effort to deal with semantic problems. It is an Open Access scholarly repository, which covers the whole range of medieval studies: social, economic, political, and institutional history, as well as cultural, religious, and gender representations and practices. RM Open Archive is open to contributions from all scholars who want to maximize the net benefits for scientific distribution and access and welcomes scholarly publications from all historical disciplines (archaeology, philology, palaeography, diplomatics, fine arts, geography, philosophy, literature, law, economics etc.), including historical methodology and didactics[73]. It is under the aegis of many important learned societies, which invite their members to deposit publications: Sismed (Italian Society of Medievalists), CoMUL (Council for Latin Middle Ages and Humanism), Sifr (Italian Society of Romance Philology), AIPD (Italian Association of Paleography and Diplomatics), Mediävistenverband (German Association of Medievalists). There are currently more than 4600 documents in it.

RM Open Archive was set up using a free open source software, EPrints, based on a LAMP (Linux, Apache, MySQL, PHP) architecture and written in PERL (Practical Extraction and Report Language) language by the University of Southampton with the support of a large international community of developers. The exposure of descriptive metadata of all contributions in the Dublin Core schema according to the OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) protocol has proven effective for the dissemination of scientific production and the free access to digital resources for research and teaching. All items deposited in RM Open Archive (articles, books, preprints, research data, multimedia contents etc.) are in fact indexed by harvesters and search engines and provided in bibliographic and full-text databases, in the online library catalogs of universities and research centers worldwide. All the metadata included in RM Open Archive are also transformed into Linked Open Data and periodically uploaded to Zenodo, the open-access repository developed under the European OpenAIRE program, operated by CERN (Eu-

---

[73] < http://www.rmoa.unina.it >. It was realised in the frame of the PRIN 2010-2011 project *Concepts, Practices and Institutions of a Discipline: Italian Medieval Studies in 19th and 20th Centuries*, coordinated by R. Delle Donne at "Federico II" University of Naples: < http://www.medievistica.unina.it >.

ropean Organization for Nuclear Research) in Geneva and supported by the European Commission[74].

However, this disciplinary archive was created in 2014, in agreement with the University Center for Libraries "Roberto Pettorino" at "Federico II" University of Naples, not only to convey several thousand research contributions into the international circuit of open archives, which has over 4100 repositories in the world and which shows an annual increase of over 500,000 units in the number of deposited records[75]. No less decisive was the intention to offer the national and international community of medievalists a platform suitable for responding to the requirements that many funding bodies, both public and private, address to scholars financed by them to deposit research results (articles etc.), within a limited time frame, in open access archives that meet high standards of interoperability and use highly standardized descriptions of resources[76].

In addition, the information classification systems which lie at the heart of the semantic web are formal ontologies, hierarchical-enumerative or analytical-synthetic schemes to order documents developed by experts, expressed in a uniform and rigorous manner and associated with primary information through the use of languages and formalisms, that are rigidly structural and well defined, understandable by machines. In this context, a structured set of linked information is added to the web of documents based on HTML or XML language: the so-called Linked Data, the connected data, based on the Resource Description Framework (RDF) syntax, a series of documented requirements proposed by the W3 Consortium for coding, exchange, reuse, and semantic interoperability. The Linked Open Data therefore constitute a technology and a set of good practices for publishing data on the web in a machine-readable format, interpretable by non-human agents, such as search engines[77].

Even digital editions of historical sources can take advantage of these technologies and good practices, integrating the most widely used model for semantic texts encoding in the humanities, the TEI, with the semantic models of the LOD, based on RDF syntax, adopted also by international content aggregators such as Europeana Digital Library.

Otherwise, by applying standards-compliant TEI markup to a digitized historical source, this acquires a formal as well as an internal structure. Due to the markup structure used, the text already contains many semantic ref-

---

[74] RM Open Archive's metadata are freely available in RDF/XML format: < https://doi.org/10.5281/zenodo.2628263 >.
[75] < http://v2.sherpa.ac.uk/id/repository/3534 >. RM Open Archive is the only disciplinary archive dedicated to medieval studies that exists today.
[76] RM Open Archive also meets the requirements recently introduced by *Plan S* because it complies with OpenAIRE and OpenDOAR: Coalition S, *Technical Guidance*.
[77] There is a very large bibliography on Linked Data. I will just mention: *Current Trends in Semantic Web*; *Linked Data for Cultural Heritage*; Sakr *et al.*, *Linked Data*; Guerrini, Possemato, *Linked data*.

erences, like spatial references, relations between individuals, or conceptual references. From the perspective of the semantic web these references are merely implicit in the text and have to be transformed into explicit RDF semantic annotations to make the historical data usable for semantic web approaches.

This is the goal we have set ourselves for some years at the Department of Humanities of Federico II University in Naples, where we are working on the digital editions of Angevin and Aragonese documents, making statements about resources (in particular web resources) in expressions of the form subject–predicate–object, known as RDF Triples[78]. Furnishing the TEI with a semantics based on a formal ontology could make all the existing historical data available in a machine-readable format, facilitate the management of research using document collections in open and multi-standard contexts, aide interoperability with other relevant standards in the digital cultural heritage context.

These ongoing researches, therefore, move on the one hand along the lines of the historical-philological methodologies of text edition, enriched by the practices of semantic text encoding in XML since the late Nineties of the last century; on the other hand, they are aimed at framing the data present in the texts in an *Open Linked Data Framework*, able to integrate heterogeneous information sources into a single research platform.

In this way, we intend to contribute to the analysis and the reflection on a topic of great relevance for the future of historiography and at the same time to foster the culture of openness in knowledge, science and research.

---

[78] The edition of Angevin documents was started in 2001 by developing a special DTD, which was then transformed into a XML Schema: < http://www.mezzogiornomedievale.unina.it/angioini >. The two editions of Aragonese documents, on which Davide Morra (Federico II University of Naples) and Biagio Nuciforo (University of Basilicata) are working for their PhD theses, are instead based on TEI P5. For both projects, one dedicated to the *Partium* registers of the "Regia Camera della Sommaria", the other to the dispatches from Naples of the Sforza ambassadors, Alfredo Cosco created *Aracne, a XQuery framework*: < https://doi.org/10.5281/zenodo.2647506 >. I point out that in Germany, at the Digital Humanities departement of the "Akademie der Wissenschaften und der Literatur Mainz", was developed a webservice called XTriples (http://xtriples.spatialhumanities.de) «to extract RDF statements out of any HTTP based XML repository using a simple configuration based on statement patterns»; the *Regesta Imperii* used it: Grüntgens, Schrade, *Data repositories*, p. 56. Reflections on these themes also in Ciotti, Tomasi, *Formal Ontologies*.

Works cited

All URLs were checked on April 7th, 2019.

G. Adamo, *Busa, Roberto*, In *Enciclopedia Italiana*, *VI Appendice*, Roma 2000: < http://www.treccani.it/enciclopedia/roberto-busa_(Enciclopedia-Italiana)/ >.

Agenzia per l'Italia digitale, *Open Data*: < https://www.dati.gov.it/ >.

Al tackles the Vatican's secrets, in «MIT Tecnology Review», March 15, 2018: < https://www-technologyreview-com.cdn.ampproject.org/c/s/www.technologyreview.com/s/610530/ai-tackles-the-vaticans-secrets/amp >.

S. Ammirati, D. Firmani, M. Maiorino, P. Merialdo, E. Nieddu, A. Rossi, *In Codice Ratio: Scalable Transcription of Historical Handwritten Documents*, in *25th Italian Symposium on Advanced Database Systems*, Squillace 2017: < http://ceur-ws.org/Vol-2037/paper_11.pdf >.

M. Ansani, *Diplomatica (e diplomatisti) nell'arena digitale*, in «Scrineum», 1 (1999), pp. 1-11: < http://www.rmoa.unina.it/19/ >.

M. Ansani, *Edizione digitale di fonti diplomatiche: esperienze, modelli testuali, priorità*, in «Reti Medievali Rivista», 7 (2006), 2, Art. #1: < http://dx.doi.org/10.6092/1593-2214/140 >.

M. Ansani, *Il Codice diplomatico digitale della Lombardia medievale: note di lavoro*, in «Atti della Società Ligure di Storia Patria», 42 (2002), 1, pp. 23-49: < http://www.rmoa.unina.it/59/ >.

M. Ansani, A. Ghignoli, *Testi digitali: nuovi media e documenti medievali*, in *Les historiens et l'informatique : un métier à réinventer*. Projet ATHIS: Atelier VI. Rome, 4-6 décembre 2008, Roma 2011, pp. 73-86.

G. Aprea, *Uno sguardo sugli strumenti digitali applicati agli archivi: il caso dei software di descrizione archivistica*, in «Bibliothecae.it», 7 (2018), 1, pp. 265-319: < http://dx.doi.org/10.6092/issn.2283-9364/8450 >.

Archivio di Stato di Firenze, *Diplomatico. Pergamene (sec. VIII-XIV)*, a cura di F. Klein, S. Vitali: < http://www.archiviodistato.firenze.it/pergasfi/ >.

Archivio di Stato di Firenze, *Mediceo avanti il Principato*, a cura di F. Klein: < http://www.archiviodistato.firenze.it/map/ >.

D. Bamman, *Interpretability in Human-Centered Data Science*, in *Human Centered Data Science @ Computer-Supported Cooperative Work and Social Computing*, San Francisco 2016: < http://people.ischool.berkeley.edu/~dbamman/pubs/pdf/Bamman_HCDS.pdf >.

G. Battioni, *Censimento ed edizione di documenti pontifici relativi alla provvista beneficiaria delle diocesi padane*, in «Schifanoia. Notizie dell'Istituto di studi rinascimentali di Ferrara», 4 (1987), pp. 151-164.

C. Belloni, *A proposito di una recente edizione di fonti vaticane e di un progetto di ricerca sulle istituzioni ecclesiastiche del ducato di Milano*, in «Nuova rivista storica», 84 (2000), pp. 421-434.

C. Belloni, *Visite pastorali milanesi nella seconda metà del XV secolo*, in *Medioevo dei poteri. Studi di storia per Giorgio Chittolini*, a cura di M.N. Covini, M. Della Misericordia, A. Gamberini, F. Somaini, Roma 2012, pp. 301-336.

C. Bianchini, M. Guerrini, *Introduzione a RDA*, Milano 2014.

C. Bizer, T. Heath, T. Berners-Lee, *Linked Data. The story so far*, in «International Journal of Semantic Web and Information Systems», 5 (2009), 3, pp. 1-22.

R. Bizzocchi, *Chiesa e potere nella Toscana del Quattrocento*, Bologna 1987 (Istituto storico italo-germanico. Monografie, 6).

Bleibt im Vatikanischen Geheimarchiv vieles zu geheim? Historische Grundlagenforschung in Mittelalter und Neuzeit. Beiträge zur Sektion des Deutschen Historischen Instituts (DHI) Rom organisiert in Verbindung mit der Westfälischen Wilhelms-Universität Münster Seminar für Mittlere und Neuere Kirchengeschichte 47. Deutscher Historikertag Dresden 30. September - 3. Oktober 2008, hrsg. von M. Matheus, H. Wolf, Rom 2009 (Online-Publikationen des Deutschen Historischen Instituts in Rom), pp. 5-12: < http://dhi-roma.it/fileadmin/user_upload/pdf-dateien/Online-Publikationen/Dresden_Histtag/Hist_Grundlagenforschung_Mittelalter_Neuzeit.pdf >.

J.A. Borghi, S. Abrams, D. Lowenberg, S. Simms, J. Chodacki, *Support Your Data: A Research Data Management Guide for Researchers*, in «Research Ideas and Outcomes», 4 (2018): < https://doi.org/10.3897/rio.4.e26439 >.

D. Brunetti, R. Grassi, M. Savoia, S. Vassallo, *Archimista. Applicativo open-source per la descrizione di archivi storici*, in «Archivi & Computer. Automazione e beni culturali», 1 (2012), pp. 42-56: < http://www.sa-lom.archivi.beniculturali.it/fileadmin/template/allegati/archimista/Archimista_un_software_inventariazione_open_source-A_C.pdf >.

M. Buzzoni, R. Rosselli del Turco, *Verso un'edizione digitale dell'Editto di Rotari*, in *I Longobardi in Italia: lingua e cultura*, Alessandria 2015, pp. 37-85.

V.A. Carriero, M. Daquino, F. Tomasi, *Convergenze semantiche tra musei, archivi e biblioteche. Ontologie per le relazioni interpersonali*, in «JLIS.IT. Italian Journal of Library, Archives and Information Science», 10 (2019), 1, pp. 72-91: < https://www.jlis.it/article/view/12499 >.

F. Cavazzana Romanelli, *L'archivio storico del Patriarcato di Venezia*, in «Notiziario bibliografico. Periodico della Giunta regionale del Veneto», 23 (sett. 1996), pp. 23-25.

F. Cavazzana Romanelli, *Il progetto "Ecclesiae Venetae". L'inventariazione di cinque archivi diocesani del Veneto*, in «Reti Medievali Rivista», 1 (2000), Art. #3: < http://dx.doi.org/10.6092/1593-2214/223 >.

L. Cerullo, *MetaFAD. Sistema di gestione integrata dei beni culturali*, in «Bibelot», 23 (2017), 2, pp. 22-31: < http://riviste.aib.it/index.php/bibelot/article/view/11685 >.

G. Chittolini, I*ntroduzione*, in *Gli Sforza, la Chiesa lombarda, la Corte di Roma. Strutture e pratiche beneficiarie nel Ducato di Milano (1450-1515)*, a cura di G. Chittolini, Napoli 1989 (Europa mediterranea, Quaderni 4), pp. XI-XXI.

G. Chittolini, C. Belloni, *Fonti e repertori per la storia milanese: i canonici delle principali collegiate in età sforzesca*, in «Reti Medievali Rivista», 2 (2001), 1, Art. #6: < http://dx.doi.org/10.6092/1593-2214/229 >.

G. Chittolini, C. Belloni, *Fonti notarili e fonti pontificie per la storia delle diocesi lombarde alla fine del Medioevo*, in *Storia della Chiesa in Europa tra ordinamento politico-amministrativo e strutture ecclesiastiche in Europa*, a cura di G. Vaccaro, Brescia 2005, pp. 181-190.

F. Ciotti, *Il testo e l'automa. Saggi di teoria e critica computazionale dei testi letterari*, Roma 2007.

F. Ciotti, F. Tomasi, *Formal Ontologies, Linked Data and TEI Semantics*, in «Journal of the Text Encoding Initiative», 9 (2016-2017): < https://journals.openedition.org/jtei/1480 >.

Coalition S, *Guidance on the Implementation of Plan S*, Bruxelles 2018: < https://www.coalition-s.org/implementation/ >.

Coalition S, *Plan S*, Bruxelles 2018: < https://www.coalition-s.org/10-principles/ >.

Codice diplomatico della Lombardia medievale (secoli VIII-XII), a cura di M. Ansani, Milano 2000-2019: < http://www.lombardiabeniculturali.it/cdlm/ >.

Codice Pelavicino. Edizione digitale, a cura di E. Salvatori, E. Riccardini, L. Balletto, R. Rosselli del Turco, C. Alzetta, C. Di Pietro, C. Mannari, R. Masotti, A. Miaschi, Pisa 2014: < http://pelavicino.labcd.unipi.it/ >.

Concepts, Practices and Institutions of a Discipline: Italian Medieval Studies in 19th and 20th Centuries, coordinated by Prof. R. Delle Donne, Napoli 2016: < http://www.medievistica.unina.it >.

A. Cosco, *Aracne, a XQuery framework to edit, maintain and publish XML-TEI collections under eXist-db*: < https://doi.org/10.5281/zenodo.2647506 >.

Council of the European Union, *The transition towards an Open Science system*, 9526/16, ANNEX DG G 3 C, Bruxelles 2016: < http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf >.

Current Trends in Semantic Web Technologies: Theory and Practice, edd. by G. Alor-Hernandez, J.L. Sanchez-Cervantes, A., Rodríguez-González, R. Valencia-García, Cham 2019.

R. Delle Donne, *Un intreccio di iniziative scientifiche. Reti Medievali e il futuro della storiografia digitale*, in «Reti Medievali Rivista», 15 (2014), 2, pp. 93-156: < http://dx.doi.org/10.6092/1593-2214/439 >.

R. Delle Donne, T. Possemato, *SHARE-CATALOGUE: Le biblioteche universitarie in Linked Open Data*, in *Bibliotecari al tempo di Google. Profili, competenze, formazione*, Milano 2016, pp. 62-71.

Direzione generale per gli archivi, Commissione nazionale per l'elaborazione del codice normativo per i soggetti produttori d'archivio, *Niera (EPF): Norme italiane per l'elaborazione dei record di autorità archivistici di enti, persone, famiglie*, a cura di E. Fregni, R. Santolamazza, Roma 2011: < http://www.icar.beniculturali.it/index.php?it/169/gestione-documentale-e- archivi-storici#naz >.

Ecclesiae Venetae, a cura di F. Cavazzana Romanelli: < http://siusa.archivi.beniculturali.it/cgi-bin/pagina.pl?RicProgetto=ev >.

A. Esch, *Leone XIII, l'apertura dell'Archivio Segreto Vaticano e la storiografia*, in *Leone XIII e gli studi storici*. Atti del Convegno internazionale commemorativo, Città del Vaticano 30-31 ottobre 2003, a cura di C. Semeraro, Città del Vaticano 2004 (Pontificio Comitato di scienze storiche. Atti e documenti 21), pp. 20-43.

European Commission, *Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee on scientific information in the digital age: access, dissemination and preservation*, Bruxelles 2007: < https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52007DC0056&from=EN >.

European Commission, *Data for Research*, Bruxelles 2018: < http://ec.europa.eu/regional_policy/it/policy/evaluations/data-for-research/ >.

European Commission, *Implementation Roadmap for the European Open Science Cloud*, Bruxelles 2018: < https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf#view=fit&pagemode=none >.

European Commission, *Open Science Monitor*, Bruxelles 2015: < https://ec.europa.eu/research/openscience/index. cfm?pg=home&section=monitor >.

European Commission, *Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information*, Bruxelles 2018: < https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32018H0790&from=EN >.

European Parliament, *Open Data and Public Sector Information*, Bruxelles 2019: < http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P8-TA-2019-0352+0+-DOC+PDF+V0//EN >.

E. Ferrarini, *Alim ieri e oggi*, in «Umanistica digitale», 1 (2017), pp. 7-17: < https://umanistica-digitale.unibo.it/article/view/7193 >

D. Firmani, P. Merialdo. M Maiorino, *In Codice Ratio: Scalable Transcription of Vatican Registers*, in «Ercim News», 111 (2017), pp. 8-9.

D. Firmani, P. Merialdo, M. Maiorino, E. Nieddu, *Towards Knowledge Discovery from the Vatican Secret Archives. In Codice Ratio-Episode 1*, in *Proceedings of the 24th ACM Sigkdd International Conference on Knowledge Discovery & Data Mining*, New York 2018: < https://arxiv.org/abs/1803.03200 >.

Friedensnobelpreis und historische Grundlagenforschung. Ludwig Quidde und die Erschließung der kurialen Registerüberlieferung, hrsg. von M. Matheus, Berlin-Boston 2012 (Bibliothek des Deutschen Historischen Instituts in Rom, 124).

J.-Ph. Genet, *Source, Métasource, Texte, Histoire*, in *Storia & Multimedia. Atti del VII Congresso Internazionale dell'Associazione for History and Computing*, a cura di F. Bocchi, P. Denley, Bologna, Grafis 1994, pp. 3-17.

M. Grüntgens, T. Schrade, *Data repositories in the Humanities and the Semantic Web: modelling, linking, visualizing*, in *1st Workshop on Humanities in the Semantic Web co-located with 13th ESWC Conference 2016*, Anissaras 2016: < http://ceur-ws.org/Vol-1608/paper-07.pdf >.

M. Guerrini, T. Possemato, *Linked data per biblioteche, archivi e musei. Perché l'informazione sia del web e non solo nel web*, Milano 2015.

Hadrian VI. 1522–1523. Text und Indices bearb. von L. Schmugge, Berlin-Boston 2018.

Her Majesty's Treasury, Department for Business, Innovation & Skills, *Our plan for growth: science and innovation*, London 2014: < https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/387780/PU1719_HMT_Science_.pdf >.

J. Hörnschemeyer, *DENQ*, in *Bleibt im Vatikanischen Geheimarchiv*, pp. 13-18.

J. Hörnschemeyer, L. Schmugge, *Uno spoglio moderno delle fonti vaticane. Il Repertorium Germanicum e il Repertorium Poenitentiariae*, in *Religiosa archivorum custodia. IV Centenario della Fondazione dell'Archivio Segreto Vaticano (1612–2012)*. Atti del Convegno di Studi, Città del Vaticano 17–18 aprile 2012, Città del Vaticano 2015, pp. 551–567.

I "libri annatarum" di Pio II e Paolo II (1458-1471), a cura di M. Ansani, Milano 1994.

I "libri annatarum" di Sisto IV (1471- 1484), a cura di G. Battioni, Milano 1997.

I "libri annatarum" di Innocenzo VIII (1484-1492), a cura di P. Merati, Milano 2000.

I "libri annatarum" di Alessandro VI (1492-1503), a cura di M. De Luca, Milano 2006.

Il fondo pergamene Pisa-Midi, a cura di E. Salvatori, Pisa 2003: < http://compass2.di.unipi.it/midi/ >.

I notai della Curia arcivescovile di Milano (secoli XIV-XV). Repertorio, a cura di C. Belloni e M. Lunari, Coordinamento di G. Chittolini, Roma 2004.

I "registra supplicationum" di Pio II (1458-1464), a cura di E. Canobbio e B. Del Bo, Milano 2007.

I processi matrimoniali degli archivi ecclesiastici italiani, a cura di S. Seidel Menchi, D. Quaglioni, IV voll. Bologna, 2000-2006.

Istituto Centrale per gli Archivi – ICAR, *EAC-CPF*: < http://www.icar.beniculturali.it/index.php?id=58 >.

Istituto Centrale per gli Archivi – ICAR, *EAD*: < http://www.icar.beniculturali.it/index.php?id=57 >.

Istituto Centrale per gli Archivi – ICAR, *ISAAR (CPF)*: < http://www.icar.beniculturali.it/index.php?id=55 >.

Istituto Centrale per gli Archivi – ICAR, *ISAD (G)*: < http://www.icar.beniculturali.it/index.php?id=54 >.

Istituto Centrale per gli Archivi – ICAR, *METS*: < http://www.icar.beniculturali.it/index.php?id=59 >.

J. Johrendt, *Papsturkundenforschung in Italien; Chancen und Grenzen der Digitalisierung*, in *Papsturkundenforschung zwischen internationaler Vernetzung und Digitalisierung. Neue Zugangsweisen zur europäischen Schriftgeschichte*, hrsg. von I. Fees, B. Hotz, B. Schönfeld, Göttingen 2015: < https://rep.adw-goe.de/handle/11858/00-001S-0000-0023-9A13-A >.

J. Kratz, C. Strasser, *Data publication consensus and controversies*, in «F1000Research», 3 (2014), 94: < http://dx.doi.org/10.12688/f1000research.3979.3 >.

La memoria delle chiese venete. Archivi diocesani e storiografia (Padova, Monastero di S. Giustina, 29 gennaio 1999), con interventi di E. Orlando, *Convegno di studi. La memoria delle chiese venete. Archivi diocesani e storiografia (Padova, Monastero di S. Giustina, 29 gennaio 1999)*, in «Rassegna degli Archivi di Stato», 60 (2000), 1, pp. 50-60.

La rubrica degli atti di Albertolo Griffi notaio e cancelliere episcopale di Pavia (1372-1420), a cura di R. Crotti e P. Majocchi, Milano 2007.

La visita pastorale di Gerardo Landriani alla diocesi di Como (1444-1445), a cura e con Introduzione di E. Canobbio, Milano 2001.

Le suppliche alla Sacra Penitenzieria Apostolica provenienti dalla diocesi di Como (1438-1484), a cura e con Introduzione di P. Ostinelli, Milano 2003.

L. Levantino, *Alla scuola di "Arca" e di "Ecclesiae Venetae". L'inventariazione di sei archivi diocesani del Veneto*, in *Francesca Cavazzana Romanelli: archivista, storica e organizzatrice di cultura*, Roma 2017: < http://www.icar.beniculturali.it/fileadmin/risorse/Interventi_materiali_contributi/Levantino_10.02.2017.pdf >.

Linked Data for Cultural Heritage, edd. by E. Jones, M. Seikel. London 2016.

F. Li Pira, *La collazione dei benefici ecclesiastici nel Mezzogiorno angioino-aragonese, "Libri Annatarum", I. 1412-1458*, Salerno 2014.

M. Matheus, *Grundlagenforschung aus Leidenschaft oder Vom bleibenden Wert kritischer Editionen. Einleitung*, in *Bleibt im Vatikanischen Geheimarchiv*, pp. 5-12.

Max Planck Gesellschaft, *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*, Berlin 2003: < http://openaccess.mpg.de/Berlin-Declaration >.

J.-B. Michel et al., *Quantitative Analysis of Culture Using Millions of Digitized Books, in Science*, 331 (2010), 6014, pp. 176-182: < http://dx.doi.org/10.1126/science.1199644 >.

F. Moretti, *Distant Reading*, London, New York 2013.

F. Moretti et al., *La letteratura in laboratorio*, a cura di G. Episcopo, Napoli 2019 (Digitalis purpurea, 1): < http://dx.doi.org/10.6093/978-88-6887-051-5 >.

A. Paravicini Bagliani, *Il papato nel secolo XIII: cent'anni di bibliografia*, Firenze 2010.

C. Ramjoué, *Open Research Data in Horizon 2020*, in *Science 2.0 Conference 2014*, Hamburg 2014: < http://www.science20-conference.eu/science-2-0-conference-2014/programme/ >.

C. Ramjoué, *Opening up scientific information in Horizon 2020 and beyond*, in *Dichiarazione di Messina 2.0: la via italiana all'accesso aperto*, Messina 2014: < http://decennale.unime.it/?page_id=588 >.

F. Ricci, *Gli standard internazionali di descrizione archivistica: dalle origini alla seconda versione di ISAD(G)*, in «Rassegna degli Archivi di Stato», 63 (2003), 1, pp. 11-45.

J. Roberts, *Open access to scientific publications must become a reality by 2020. Interview to Robert-Jan Smits*, «Horizon. The EU Research & Innovation Magazine», 23 March 2018: <https://horizon-magazine.eu/article/open-access-scientific-publications-must-become-reality-2020-robert-jan- smits_en.html>.

A. Ronca, G. Romano, *SIUSA. Il software*, in «Centro di Ricerche Informatiche per i Beni Culturali», 11 (2001), 2: *SIUSA - Sistema Informativo Unificato per le Soprintendenze Archivistiche. Genesi e sviluppi di un progetto*, a cura di D. Bondielli, pp. 73-76: < http://siusa. archivi.beniculturali.it/documenti/SIUSA_Genesi_e_sviluppi_di_un_progetto.pdf >.

R. Rosselli Del Turco, *The Digital Vercelli Book*, Torino 2017: < https://www.collane.unito.it/ oa/files/original/142e56744c617c5e459c4e142238290b.pdf >.

R. Rosselli Del Turco, G. Buomprisco, C. Di Pietro, J. Kenny, R. Masotti, J. Pugliese, *Edition Visualization Technology: A Simple Tool to Visualize TEI-Based Digital Editions*, in Journal of the Text Encoding Initiative, 8 (2015), pp. 1-29: < http://dx.doi.org/10.4000/jtei.1077 >.

S. Sakr, M. Wylot, R. Mutharaju, S. Le Phuoc, I. Fundulaki, *Linked Data. Storing, Querying, and Reasoning*, Cham 2018.

V. Salardi, *L'edizione digitale dello Statuto di Vicenza del 1264*, in «Mélanges de l'École française de Rome - Moyen Âge», 119 (2007), 2, pp. 337-345.

J. Schnapp, T. Presner, P. Lunenfeld, J. Drucker, *The Digital Humanities Manifesto*, 2009: < https://mbf.blogs.com/files/the-digital-humanities-manifesto-2.0-2.pdf >.

J. Schöpfel, H. Prost, *Research data management in social sciences and humanities: A survey at the University of Lille (France)*, in «LIBREAS. Library Ideas», 29 (2016): < https://libreas.eu/ausgabe29/09schoepfel/ >.

L. Siciliano, *Codifiche XML tra adesione agli standard e sistemi autonomi. Il caso degli Statuti*, in «Mélanges de l'École française de Rome - Moyen Âge», 119 (2007), 2, pp. 327-335.

Società Internazionale per lo Studio del Medioevo Latino, Fondazione Ezio Franceschini, *Mirabile. Archivio digitale della cultura medievale*, Firenze 2019: < http://www.mirabileweb. it >.

Tarsian: Digital Edition of an Ancient New Testament Arabic Manuscript, ed. by C. Clivaz, Lausanne 2016: < http://tarsian.vital-it.ch/about/ >.

G.M. Varanini, *Archivi ecclesiastici veneziani e archivi di Terraferma nelle ricerche di Francesca Cavazzana Romanelli*, in *Francesca Cavazzana Romanelli: archivista, storica e organizzatrice di cultura*, Roma 2017: < http://www.icar.beniculturali.it/fileadmin/risorse/ Interventi_materiali_contributi/Varanini_10.02.17.pdf >.

Vercelli Book Digitale, a cura di R. Rosselli Del Turco, trascrizione e codifica a cura di, R. Cioffi, F. Goria; software EVT creato da C. Di Pietro, J. Kenny, R. Masotti, R. Rosselli Del Turco, Pisa 2017: < http://vbd.humnet.unipi.it/beta2/ >.

S. Vitali, *La descrizione archivistica al tempo degli standard e dei sistemi informatici*, in *Archivistica. Teorie, metodi, pratiche*, a cura di M. Guercio, L. Giuva, Roma 2014, pp. 179-210.

S. Vitali, *Passato digitale. Le fonti dello storico nell'era del computer*, Milano 2004.

S. Vitali, *The San Portal: A Common Gateway to Italian Archival Resources on the Web*, in *Cultural Heritage on Line. Empowering Users: An Active Role for User Communities*, a cura di C. Cirinnà, M. Lunghi, Firenze 2010, pp. 51-55.

A. Zorzi, *Documenti, archivi digitali, metafonti*, in «Archivi & computer. Automazione e beni culturali», 10 (2000), pp. 274-291.

A. Zorzi, *L'edizione digitale degli statuti dei comuni italiani: qualche riflessione a modo di premessa*, in «Mélanges de l'École française de Rome - Moyen Âge», 119 (2007), 2, pp. 323-326.

Roberto Delle Donne
Università degli Studi di Napoli Federico II
delledon@unina.it