Ecdotica

9

(2012)

Alma Mater Studiorum. Università di Bologna
Dipartimento di Filologia Classica e Italianistica

Centro para la Edición
de los Clásicos Españoles

Carocci editore

Comitato direttivo

Gian Mario Anselmi, Emilio Pasquini, Francisco Rico

Comitato scientifico

Edoardo Barbieri, Francesco Bausi, Pedro M. Cátedra,
Roger Chartier, Umberto Eco, Conor Fahy †, Inés Fernández-Ordóñez,
Domenico Fiormonte, Hans-Walter Gabler, Guglielmo Gorni †,
David C. Greetham, Neil Harris, Lotte Hellinga, Paola Italia, Mario Mancini,
Armando Petrucci, Amedeo Quondam, Ezio Raimondi, Roland Reuß,
Peter Robinson, Antonio Sorella, Pasquale Stoppelli,
Alfredo Stussi, Maria Gioia Tavoni, Paolo Trovato

Responsabile di Redazione

Loredana Chines

Redazione

Federico Della Corte, Rosy Cupo, Laura Fernández,
Luigi Giuliani, Camilla Giunti,
Amelia de Paz, Andrea Severi, Marco Veglia

Ecdotica is a Peer reviewed Journal

Ecdotica garantisce e risponde del valore e del rigore dei contributi che
si pubblicano sulla rivista, pur non condividendone sempre e necessariamente
prospettive e punti di vista.

On line:

http://ecdotica.org

Alma Mater Studiorum. Università di Bologna,
Dipartimento di Filologia Classica e Italianistica,
Via Zamboni 32, 40126 Bologna
ecdotica.dipital@unibo.it

Centro para la Edición de los Clásicos Españoles
Don Ramón de la Cruz, 26 (6 B)
Madrid 28001
cece@cece.edu.es
www.cece.edu.es

Con il contributo straordinario dell'Ateneo di Bologna, con il contributo
della Fondazione Cassa di Risparmio in Bologna e della Fundación Aqualogy

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA
CCEE CENTRO PARA LA EDICIÓN DE LOS CLÁSICOS ESPAÑOLES
FONDAZIONE CASSA DI RISPARMIO IN BOLOGNA
AQUALOGY

Carocci editore · Corso Vittorio Emanuele II, 229 00186 Roma · tel. 06.42818417, fax 06.42747931

INDICE

Saggi

PETER ROBINSON, The textual tradition of Dante's *Commedia* and the «Barbi *loci*» 7

ALBERTO CADIOLI, Dare una cronologia alle carte del *Giorno* di Parini. Una riflessione metodologica 39

DOLORES TRONCOSO, Los dos textos de los *Episodios nacionales* 69

PAUL EGGERT, Anglo-American critical editing. Concepts, terms and methodologies 113

PAOLO CHERCHI, Filologia in pericolo. Considerazioni di un *outsider* 125

Foro. *Ecdotica dell'errore. In onore di Michael Reeve*

FRANCISCO RICO, Presentazione 149

MICHAEL REEVE, More on maps 150

PAOLO CHIESA, Una letteratura«sbagliata». I testi mediolatini e gli errori 151

PIETRO G. BELTRAMI, A proposito di errori nella critica del testo romanza 162

GIULIA RABONI, Per una filologia d'autore meno bedieriana 171

Testi

MATTEO VENIER, Francesco Robortello: *Discorso sull'arte ovvero sul metodo di correggere gli autori antichi* 183

Questioni

Paola Italia, Libri che parlano di libri: dentro e fuori 219

Roberta Colbertaldo, La *historisch-kritische Franz Kafka-Ausgabe* 238

Andrea Severi, Se la lezione giusta è quella sbagliata (*Love's Labour's Lost* IV, 2, 92-93) 253

Francesca Tomasi, L'edizione digitale e la rappresentazione della conoscenza. Un esempio: Vespasiano da Bisticci e le sue lettere 264

Rassegne

Érasme de Rotterdam, *Les Adages* (I. Dionigi-F. Citti), p. 287 · *Studi e problemi di critica testuale: 1960-2010. Per i 50 anni della Commissione per i testi di lingua* (S. Albonico), p. 297 · Alberto Varvaro, *Prime lezioni di filologia* (Luca Morlino), p. 312 · *La tradizione della lirica nel medioevo romanzo. Problemi di filologia formale. Atti del convegno internazionale. Firenze-Siena, 12-14 novembre 2009*, a cura di L. Leonardi (S. Martí), p. 319 · Antony Grafton, *The Culture of Correction in Renaissance Europe* (G. Pontón), p. 325 · A. Corveto, *Tipos de imprenta en España* (D. Cruikshank), p. 332 · Fernando Bouza, *Hétérographies. Formes de l'écrit au Siècle d'Or espagnol* (Marta Latorre), p. 336 · Maria Gioia Tavoni, *Circumnavigare il testo. Gli indici in età moderna* (Alberta Pettoello), p. 341 · François Déroche y Valentina Sagaria Rossi, *I manoscritti in caratteri arabi* (Nuria M. de Castilla), p. 345 · *Switching Codes: Thinking Through Digital Technology in the Humanities and the Arts* (Massimo Riva), p. 350 · *Dieci anni di «Per leggere». I generi della lettura. Atti della giornata di studio all'Università Europea di Roma. Indici della prima serie (nn. 1-20)*, a cura di I. Becherucci (M.R. Traina), p. 364

L'EDIZIONE DIGITALE E LA RAPPRESENTAZIONE DELLA CONOSCENZA. UN ESEMPIO: VESPASIANO DA BISTICCI E LE SUE LETTERE

FRANCESCA TOMASI

L'edizione digitale è un processo. E in quanto tale obbliga a riflettere sui diversi «stati», o passi, che traducono le interpretazioni dell'editore critico in un sistema formale. Naturalmente derivano da questa riflessione interrogativi che non solo riguardano l'aspetto tecnologico del processo, ma che soprattutto richiedono la definizione di un modello teorico che risponda alla volontà dell'*interpres* e quindi ai bisogni informativi del lettore. Condividendo l'affermazione di McGann sull'«edizione come attività teoretica»,[1] andranno in primo luogo chiarite le implicazioni che, ai fini della definizione del processo editoriale, la traduzione digitale comporta. Diremo quindi, con Gabler, che «today, as the carrier medium for editions changes from book to the digital medium, all the main *a priori* assumptions about scholarly editions come into question».[2]

Come cambia quindi l'edizione di un testo letterario nel momento in cui la componente digitale diventa veicolo non solo di produzione ma anche di modellazione? Quale ruolo svolge la digitalizzazione, intesa nel senso estensivo di migrazione da un ambiente analogico a quello digitale, nella ridefinizione del procedimento editoriale? Come muta il concetto di edizione quando le tecnologie costringono a riformulare il senso stesso dell'edizione del testo?

Domande che sempre più di frequente popolano l'orizzonte di interesse dell'informatica umanistica (tradizionalmente *humanities compu-*

[1] J. McGann, *Radiant Textuality. Literature after the World Wide Web*, New York-London, Palgrave, 2001; trad. it. *La letteratura dopo il Word Wide Web. Il testo letterario nell'era digitale*, edizione italiana a cura di D. Buzzetti, Bologna, BUP, 2002, pp. 83-107.

[2] H.W. Gabler, «Theorizing the Digital Scholarly Edition», *Literature Compass*, VII/2 (2010), p. 43.

ting[3] o, più di recente, *digital humanities*[4]) che annovera, fra i suoi ambiti di pertinenza, quella che in Italia è stata chiamata filologia computazionale[5] o, a livello internazionale, *digital scholarship* o ancora *electronic textual editing*.[6] Importanti contributi in materia e altrettanto autorevoli studiosi hanno a lungo arricchito la discussione con prospettive critiche stimolanti e propulsive. Crediamo ci sia ancora margine per tornare a ragionare sul senso dell'edizione digitale e sulle potenzialità che le tecnologie del nuovo Web – da quelle connesse al Web semantico[7] a quelle che si possono annoverare nel progetto Linked data[8] – mettono a disposizione, per una riflessione sulle ricadute teoriche, metodologiche e pratiche dell'IT nel settore umanistico e in particolare nel campo delle edizioni di testi. Partendo da un presupposto. L'edizione digitale è un sistema informativo che prevede e determina mutue complesse relazioni fra i suoi elementi costitutivi e che comporta connessioni profonde fra il testo e la rete, vale a dire fra il testo e il suo corredo metatestuale. Una base di conoscenza quindi – concetto che avremo modo di approfondire oltre[9] – che trova nei linguaggi di markup e nei relativi schemi, nei metadati e nelle ontologie una forma di espressione e una modalità di modellazione.

Ripensare all'edizione di testi letterari in ambiente digitale, dalla modellazione alla produzione, significa riflettere sulla funzione del processo

[3] Il riferimento non può non andare al noto volume di Willard McCarty, *Humanities computing*, London, Palgrave Macmillan, 2005.

[4] Termine ora comunemente utilizzato per riferirsi al concetto (cfr. l'Alliance of Digital Humanities Organizations, http://adho.org). Sullo «shift» del termine si veda: P. Svensson, «Humanities Computing as Digital Humanities», *Digital Humanities Quarterly*, III/3 (2009), http://digitalhumanities.org/dhq/vol/3/3/index.html.

[5] L'allusione è al volume di L. Perilli, *Filologia computazionale*, Roma, Accademia nazionale dei Lincei, 1995.

[6] Questo il titolo scelto dai curatori dell'omonimo volume «supported by the Modern Language Association's Committee on Scholarly Editions and by the Text Encoding Initiative Consortium». J. Unsworth, K. O'Brien O'Keeffe, L. Burnard (eds.), *Electronic Textual Editing*, Modern Language Association of America, 2006. Una versione preliminare del volume si legge sul sito ufficiale TEI: http://www.tei-c.org/About/Archive_new/ETE/.

[7] Sito ufficiale del W3C: http://www.w3.org/standards/semanticweb/. Si veda anche T. Berners-Lee, J. Hendler, O. Lassila, «The Semantic Web», *The Scientific American* (may 2001), pp. 34-43.

[8] Sito ufficiale: http://linkeddata.org/. Si veda anche il primo contributo su Linked Data, il cui ideatore è lo stesso fondatore del WWW: T. Berners-Lee, *Linked Data*, http://www.w3.org/DesignIssues/LinkedData.html (july 2006).

[9] Il termine viene qui usato non in forma tecnica in senso stretto (la *knowledge base* come database, quindi come tecnologia che segue la modellazione logica), ma nel senso di raccolta e collezione di informazioni in grado di esprimere concetti che nascono dalle relazioni fra gli oggetti di un dominio.

editoriale. Potremmo dire che l'edizione di un testo, anche la più tradizionale, si presenta come una collezione di informazioni altamente strutturate o di ipotesi interpretative organiche. Questo significa che l'edizione, a partire dalla delicata fase della trascrizione, è il momento in cui il sistema testo dell'editore incontra il sistema testo(i) dell'autore (ma anche del copista, del tipografo o in generale dell'estensore).[10] E tale incontro produce, come noto, una serie di riflessioni che si manifestano, in prima battuta, in forma di *restitutio textus*. Ma non solo. Il momento editoriale si esplica nella costruzione di un apparato che, oltre a contemplare la componente strettamente filologica – l'apparato delle varianti –, potrebbe prevedere una serie di note storiche, biografiche, bibliografiche, citazionali che arricchiscono il testo stabilito. Diremo che ogni edizione rappresenta «the editor's theory of text».[11]

Ragionando a livello di edizione digitale diremo che l'annotazione di un testo attraverso linguaggi formali di descrizione è un processo che porta alla identificazione degli elementi utili per la creazione di un modello del testo che traduca le ipotesi interpretative dell'editore in modo formale.[12] Tale annotazione enuclea una serie di componenti che hanno la funzione di esprimere formalmente sia le scelte effettuate ai fini della restituzione del testo che della costruzione dell'apparato. Nel processo editoriale dunque la componente digitale costringe a rendere esplicite le scelte dell'editore critico traducendole nel sistema di annotazione. Potremmo parlare di «edizione elettronica come modello del testo».[13] O potremmo anche dire che il modello del testo deriva dalla dimensione digitale della rappresentazione.

Le edizioni digitali in ambito umanistico[14] sono spesso basate su un sistema di annotazione che si fonda su linguaggi di marcatura *embedded* come XML[15] e il suo schema, o modello, TEI.[16] Questo significa che l'an-

[10] Cfr. R. Mordenti, *Informatica e critica dei testi*, Roma, Bulzoni, 2001, pp. 53-82.

[11] Modern Language Association Committee on Scholarly Editions, *Guidelines for Editors of Scholarly Editions*, http://www.mla.org/cse_guidelines (last revised 29 june 2011).

[12] Sui linguaggi di markup e sulle loro implicazioni si veda uno fra i primi contributi prodotti dalla comunità di digital humanities: J.H. Coombs, A.H. Renear, S.J. DeRose: «Markup Systems and the Future of Scholarly Text Processing», *Communications of the ACM*, XXX/11 (1987), pp. 933-947.

[13] T. Orlandi, *Informatica testuale*, Roma-Bari, Laterza, 2010, p. 107.

[14] Una buona rassegna si legge in: P. Sahle (ed.), *A Catalog of Digital Scholarly Editions*, http://www.uni-koeln.de/~ahz26/vlet/index.html.

[15] La specifica del W3C dell'Extensible Markup Language si legge all'indirizzo: http://www.w3.org/XML/.

[16] Sito ufficiale TEI (*Text Encoding Initiative*): http://tei-c.org.

notazione, basata su regole sintattiche esplicite (XML), traduce l'interpretazione in un elemento di vocabolario condiviso (TEI). Attraverso l'abbinata XML/TEI l'editore può quindi esplicitamente dichiarare la funzione logica delle stringhe di caratteri che compongono il testo, arricchendolo di informazioni.[17] L'annotazione, che chiameremo anche markup, è in sé un primo momento di traduzione del linguaggio naturale in linguaggio formale, ma sicuramente è un'attività che non soddisfa tutte le esigenze dell'editore critico e certamente non sfrutta appieno le possibilità del digitale.[18]

Le ragioni di tale insufficienza riguardano tanto il funzionamento stesso dei linguaggi di markup dichiarativo, e i principi che ne stanno a fondamento, quanto le capacità espressive di sistemi embedded e degli schemi relativi.[19] Rispetto al caso che vogliamo qui discutere individueremo tre livelli di intervento critico che vanno oltre il markup e che costringono a ulteriori livelli di formalizzazione.

In prima battuta diremo che l'attribuzione di una componente logica attraverso la marcatura non permette di dichiarare tutte le *features* potenziali che identificano la stringa annotata. Segnalare che un certo segno abbreviativo si può sciogliere secondo una certa modalità non esaurisce il patrimonio informativo su quel segno; allo stesso modo la segnalazione di una lezione variante rispetto alla lezione accettata dall'editore critico non necessariamente traduce il processo che ha deter-

[17] Un elenco di edizioni, e di progetti di *digital libraries* o collezioni digitali in generale, che si basano su XML/TEI si può trovare sullo stesso sito dedicato allo schema: *Projects using TEI*, http://www.tei-c.org/Activities/Projects/.

[18] Sull'insufficienza del markup nella rappresentazione del testo molti sono i contributi critici. Si veda in particolare, nell'ottica della modellazione: D. Buzzetti, *Rappresentazione digitale e modello del testo*, in *Il ruolo del modello nella scienza e nel sapere*, Atti del Convegno (Roma, Accademia Nazionale dei Lincei, 27-28 ottobre 1998), Roma, Accademia Nazionale dei Lincei, 1999, (Contributi del Centro Linceo Interdisciplinare «Beniamino Segre», N. 100), pp. 127-161.

[19] Andiamo dal problema legato al concetto di testo nei linguaggi di markup dichiarativo, derivato dal modello OHCO (*Ordered Hierarchy of Content Objects*) e teorizzato in S.J. De Rose, D.G. Durand, E. Mylonas, A.H. Renear, «What Is Text, Really?», *Journal of Computing in Higher Education*, I/2 (1990), pp. 3-26, alla necessità di aggiungere semantica al markup: A.H. Renear, M. Sperberg-McQueen e C. Huitfeldt, *Towards a semantics for XML markup*, in *DocEng'02: Proceedings of the 2002 ACM Symposium on Document Engineering*, R. Furuta, J.I. Maletic e E. Munson (eds.), (McLean, VA, November 8-9 2002), New York, ACM Press, 2002, pp. 119-126. Ma esistono anche posizioni che prediligono lo stand-off markup per far fronte all'annoso problema dell'overlapping: D. Schmidt, R. Colomb, «A data structure for representing multi-version texts online», *International Journal of Human-Computer Studies*, LXVII/6 (2009), pp. 497-514. In generale si veda: D. Buzzetti, *Digital Editions and Text Processing*, in *Text Editing, Print, and the Digital World*, M. Deegan and K. Sutherland (eds.), Aldershot, Ashgate, 2009, pp. 45-62.

minato quella scelta; o ancora l'individuazione di un nome di persona o di un luogo non dichiara il bagaglio informativo che sta dietro a quella persona o a quel luogo.

In seconda battuta è innegabile che l'informazione nascosta, che nasce naturalmente dal contesto e dal dominio, ma anche dalle personali conoscenze dell'editore critico, non può emergere attraverso l'annotazione. Soprattutto le relazioni implicite che le stringhe annotate intrattengono (una persona con un luogo, una data con un evento, ma anche una lezione variante con un'altra in qualche modo collegata o un segno grafico con uno affine da un qualche punto di vista) devono essere dichiarate ed espresse in modo formale esternamente al documento, allo scopo di separare il testo da quello che potremmo chiamare, in senso estensivo, l'apparato.

In terzo luogo è evidente che gli elementi dell'annotazione intrattengono, individualmente o a loro volta come relazione, collegamenti con risorse disseminate sul Web ed è quindi opportuno favorire il dialogo fra l'edizione, quindi il testo annotato, e altre risorse digitali relative allo stesso contenuto informativo.

Da questa riflessione derivano tre potenziali livelli di rappresentazione del testo digitale: un livello interno che si esprime sul piano dell'annotazione attraverso linguaggi di markup, un livello esterno che traduce l'annotazione in un sistema di relazioni fra gli elementi marcati e un terzo livello che collega la rappresentazione, interna o esterna, alle risorse disponibili su Web.[20] Passiamo quindi dalla edizione digitale alla base di conoscenza: implicite relazioni diventano asserzioni esplicite attraverso adeguati predicati e gli elementi marcati, assieme alle relazioni identificate formalmente, possono diventare veicolo di dialogo con il WWW.

Non vogliamo quindi ragionare in questa sede su strumenti e applicazioni,[21] ma su un modello teorico che possa essere applicato a situazioni

[20] Si veda anche: M. Barbera, F. Meschini, C. Morbidoni e F. Tomasi, *Annotating digital libraries and electronic editions in a collaborative and semantic perspective*, in *Digital Libraries and Archives*, M. Agosti, F. Esposito, S. Ferilli e N. Ferro (eds.), IRCDL2012, 8th Italian Research Conference on Digital Libraries (Bari, 9-10 febbraio 2012), Revised Selected Papers. Communications in Computer and Information Science (CCIS), Berlin-Heidelberg, Springer Verlag, CCCLIV, 2012, pp. 45-56.

[21] Problema già sollevato da Robinson che individua proprio nella mancanza di strumenti («user-friendly tools») una delle ragioni per la resistenza da parte della comunità degli studiosi alla realizzazione di edizioni digitali. Cfr. P. Robinson, «Where We Are with Electronic Scholarly Editions, and Where We Want to Be», *Jahrbuch für Computerphilologie*, 5 (2003), pp. 125-146; o ancora P. Robinson, «Current Issues in Making Digital Editions of Medieval Texts – or, do Electronic Scholarly Editions have a Future?»,

testuali diverse. Abbiamo milioni di testi marcati[22] e non sappiamo come acquisire conoscenza dal markup se non iniziamo a far dialogare, attraverso sistemi formali condivisi e standardizzati, gli elementi dell'annotazione e a mettere in relazione le edizioni con l'universo del Web. Anche se utilizzato in un contesto diverso da quello che si tenta in questa sede di presentare così afferma Unsworth a proposito delle opportunità legate all'impiego di procedure di *text mining*:

Over the last decade, many millions of dollars have been invested in creating digital library collections: at this point, terabytes of full-text humanities resources are publicly available on the web. Those collections, dispersed across many different institutions, are large enough and rich enough to provide an excellent opportunity ... .[23]

Dobbiamo evitare quindi che progetti di edizioni rimangano chiusi in un circuito di autoreferenzialità, che ne limita le possibilità non solo di dialogo ma anche di reperimento.[24] Metadati e ontologie, quindi vocabolari condivisi e linguaggi formali, forniscono i presupposti teorici, metodologici e tecnici affinché l'edizione diventi un processo, che va preservato nel tempo. Condivisibile quindi la posizione di Cathy Moran Hajo durante il convegno Balisage del 2010:

We don't have a good model in use today for a sustainable XML edition with which we could develop a shared conception of digital editing. There are a lot of silo projects that have been developed for specific sets of documents and that are not broad-minded enough to serve all editions. I think that a lot of these idiosyncratic editions have good ideas, but often do not have sufficient infrastructure to ensure longevity.[25]

Digital Medievalist, I/1 (2005). L'articolo si può leggere all'indirizzo: http://www.digitalmedievalist.org/journal/1.1/robinson/.

[22] Cfr. G. Crane, «What Do You Do with A Million Books?», *D-Lib Magazine*, XII/3 (march 2006), http://www.dlib.org/dlib/march06/crane/03crane.html.

[23] L'affermazione di Unsworth si può leggere in vari luoghi. Il più significativo è quello relativo alla presentazione della storia della ideazione del progetto MONK (*Metadata Offer New Knowledge*). J. Unsworth, *Monk. Background*, http://www.monkproject.org/background.html.

[24] Ne deriva l'esigenza di creare ambienti di aggregazione delle risorse digitali prodotte dalla comunità. Un caso esemplare è quello di *Nines* (*Networked Infrastructure for Nineteenth-century Electronic Scholarship*) che è possibile consultare all'indirizzo: http://www.nines.org/. «NINES is an open source, distributed aggregation of peer-reviewed online scholarly work on nineteenth-century British and American literature and culture».

[25] C.M. Hajo, «The sustainability of the scholarly edition in a digital world», in *Proceedings of the International Symposium on XML for the Long Haul: Issues in the Long-*

Per cercare di esemplificare il processo di modellazione, che trasforma il testo digitale in una base di conoscenza, partiamo da un caso di studio, per analizzare la componente della marcatura, quella delle relazioni fra gli elementi annotati e quella dei collegamenti fra il testo e le risorse disseminate.[26] L'impiego di sistemi standard sia a livello di rappresentazione delle componenti del testo che a livello di costruzione di relazioni è fondamentale per tradurre un'edizione isolata in un corpus informativo che si inserisce in un sistema di dialogo interoperabile.

Diremo quindi che:

> Dopo lo sviluppo delle tecnologie per la *rappresentazione* del testo con l'impiego di marcatori inseriti all'interno della stringa di caratteri che ne costituisce la rappresentazione codificata di base (*embedded markup*), il problema dell'*elaborazione* del contenuto informativo dei dati testuali si presenta in forme nuove.[27]

Possiamo quindi ancora parlare di edizione? O parleremo meglio di collezione di informazioni che trasformano il testo stabilito in una rete extra-testuale?

Per rispondere a questo interrogativo si sono scelti tre fenomeni abitualmente oggetto di intervento interpretativo in un'edizione commentata: le persone, il lessico e i riferimenti bibliografici (autori ed opere).[28]

1. Il caso di studio

Il caso in esame è la collezione delle lettere ricevute e inviate da/a Vespasiano da Bisticci, copista e libraio fiorentino vissuto nel corso del xv secolo. Anzi meglio diremo «princeps omnium librariorum Florentino-

term Preservation of XML, Balisage Series on Markup Technologies, VI (2010), http://www.balisage.net/Proceedings/vol6/print/Hajo01/BalisageVol6-Hajo01.html.

[26] Si veda anche F. Tomasi, «Digital editions between embedded markup and external representation. A case study: Vespasiano da Bisticci's Letters», in *Dall'informatica umanistica alle culture digitali*, Atti del convegno in memoria di Giuseppe Gigliozzi (Roma, 27-28 ottobre 2011), G. Crupi e F. Ciotti (eds.), Quaderni Digilab. II, Roma, Sapienza Università Editrice, 2012, pp. 201-218, http://digilab-pub.uniroma1.it/index.php/Quaderni_DigiLab/issue/view/2.

[27] D. Buzzetti, *Che cos'è, oggi, l'informatica umanistica? L'impatto della tecnologia*, in *Dall'informatica umanistica alle culture digitali*, cit., pp. 103-132.

[28] La necessità di assumere un punto di vista sull'oggetto dell'analisi ha determinato questa scelta selettiva. Ovviamente altri fenomeni potrebbero essere passibili di intervento analogo a quello qui descritto.

rum».[29] Le lettere, non un vero e proprio epistolario quindi, ma una raccolta risultato dell'aggregazione da parte di moderni editori, coprono la gamma degli interessi del libraio e aprono una finestra sulla tipologia delle sue occupazioni, fornendo un quadro nitido, nonostante l'esiguità del *corpus* (costituito da sole quarantaquattro lettere, delle quali ventidue inviate e ventidue ricevute, che coprono il periodo 1446-1498[30]), di alcune questioni fondative dell'umanesimo: le biblioteche degli umanisti e dei mecenati delle principali corti europee, le modalità della circolazione dei codici manoscritti, i rapporti politici, le traduzioni e i volgarizzamenti. Le missive testimoniano relazioni commerciali, rapporti di amicizia, scambi di vedute politiche, narrazioni di fatti storici, discussioni di diatribe teologiche. E la varietà dello status dei corrispondenti rispecchia la molteplicità dei contenuti trattati: amici, maestri, «dipendenti», signori e principi, semplici committenti.

Le lettere di Vespasiano sono ad oggi raccolte in istituti di conservazione diversi e hanno quindi una diversa origine materiale (cfr. nota 32). Il libraio decise di far trascrivere in codice unico solo un piccolo corpus, corrispondente ad un parziale raggruppamento delle lettere ricevute,[31] realizzato forse per compiacere Alessandro di Pierfilippo Pandolfini e i membri dell'Accademia Fiorentina cui il codice era dedicato. Ma si tratta di un caso isolato. Il resto del carteggio rinvenuto fino ad oggi proviene da fonti diverse.[32]

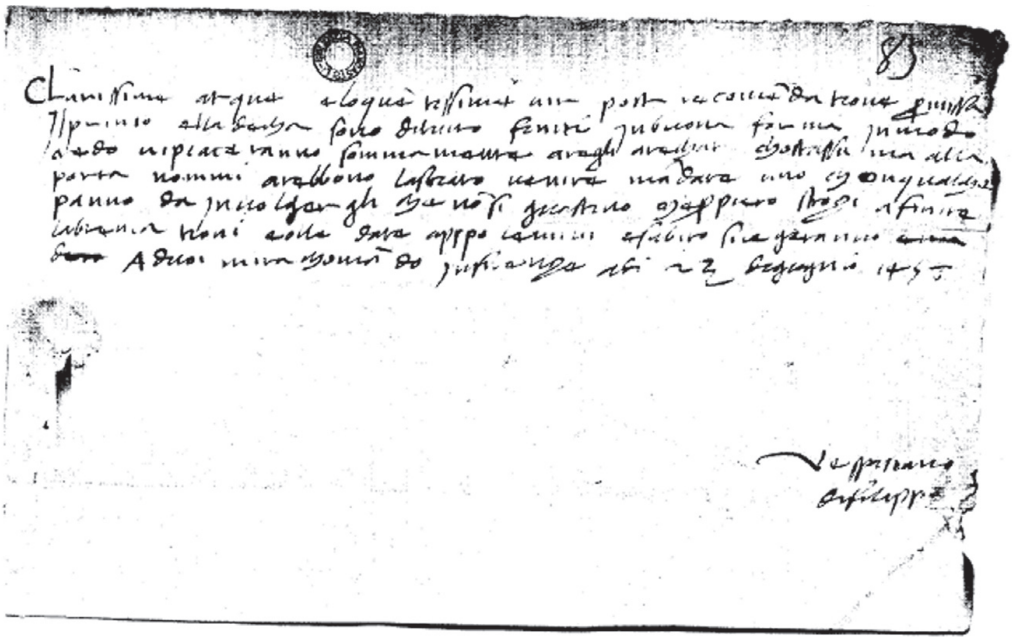
[29] Così lo appella uno dei più prolifici copisti della scuola di Vespasiano, Gherardo del Ciriago (cfr. A.C. de la Mare, *New research on humanistic scribes in Florence*, in *Miniatura fiorentina del Rinascimento 1440-1525*, A. Garzelli (ed.), Firenze, Giunta Regionale Toscana, La Nuova Italia, 1985, App. III, p. 567, nn. 17 e 18).

[30] Vari sono gli editori delle lettere. L'ultimo è il Cagni, che ha realizzato un'edizione pressoché integrale (quarantadue lettere): G.M. Cagni, *Vespasiano da Bisticci e il suo epistolario*, Roma, Edizioni di Storia e Letteratura, 1969, pp. 117-185. Le restanti due lettere sono sempre state pubblicate dal medesimo editore: G.M. Cagni, «Agnolo Manetti e Vespasiano da Bisticci», *Italia Medioevale e Umanistica*, XIV (1971), pp. 293-312. Scopo dell'edizione digitale è di produrre una nuova edizione, partendo dai documenti originali, integralmente disponibili. Il Cagni non rispetta rigorosamente la lingua ma procede ad arbitrarie correzioni ed integrazioni, compreso l'inserimento delle congiunzioni di grado zero.

[31] Si tratta di un codice di dedica, un pluteo della Medicea Laurenziana (Plut. 90 sup. 30) che raccoglie sedici lettere ricevute da Vespasiano.

[32] Oltre al Plut. 90 sup. 30 altre lettere sono confluite in un codice cartaceo miscellaneo: Firenze, Biblioteca Nazionale, Magl. VIII, 1390 (ff. 102, 102v, 103v, 41v, 47v, 48v) che raccoglie, per la maggior parte, note e lettere di Donato Acciaiuoli (sia lettere dell'umanista che lettere scritte per Vespasiano da Bisticci). Un documento è esso stesso un manoscritto di sole 10 carte (Firenze, Biblioteca Nazionale, Ms. Naz. II.XI.34, ff. 1-10v). Altre tre lettere sono carte di codici della Vaticana (Roma, Biblioteca Apostolica Vati-

Analizziamo dunque un autografo di Vespasiano da Bisticci, a Piero de' Medici,[33] che costituisce una parte di una piccola collezione di missive scambiate fra il libraio e l'illustre committente:[34]



Desidero ringraziare sentitamente la Dott.ssa Antonella Imolesi, Responsabile Fondi Antichi, Manoscritti e Raccolte Piancastelli della Biblioteca Comunale «A. Saffi» di Forlì per la gentile riproduzione della carta in oggetto.

cana, Vat. lat. 326, f. 26r-v; Reg. lat. 1612, ff. 157v-158, f. 158-158v). Una lettera si trova in un codice dell'Universitaria (Bologna, Biblioteca Universitaria, Italiano 1452, ff. 347-349). Per quanto riguarda i documenti d'archivio ci interessano quattro fondi: tre dell'Archivio di Stato e uno della Biblioteca Comunale di Forlì: Firenze, Archivio di Stato, *Mediceo avanti il Principato* (1373-1557; 165 pezzi; inventario a stampa) filza XVII n. 165, filza XVII n. 176, filza XII n. 413, filza XXVIII n. 701, filza XXV n. 377. Riproduzione digitale integrale in http://www.archiviodistato.firenze.it/Map/. Firenze, Archivio di Stato, *Carte Strozziane* (1336-1905; 3497 pezzi; inventario) serie I, filza 253 n. 2; serie III, vol. 145 ff. 86-87v; serie I, vol. 137 f. 290r-v; serie I, vol. 137 ff. 288-289v. Firenze, Archivio di Stato, *Notarile Antecosimiano* (1237-1589; 21489 pezzi; inventario) T. 498 allegato primo e secondo all'atto del 6 settembre 1479 e allegato all'atto del 14 febbraio 1479 (il notaio è Ser Leonardo da Colle). Forlì, Biblioteca Comunale, *Autografi Piancastelli*, n. 85 e n. 102. Quattro documenti sono stati trascritti direttamente da due testi a stampa. Si tratta di carte provenienti dal fondo aragonese (cfr. F. Trinchera, *Codice Aragonese*, Napoli, 1866-1874, III voll.), distrutto dai tedeschi nel 1943.

[33] Si tratta della carta conservata nella Busta 85, *ad vocem* Vespasiano da Bisticci, Sez. Aut. Sec. XII-XVIII, Raccolte Piancastelli, Biblioteca Comunale «A. Saffi», Forlì.

[34] Senza dimenticare che Vespasiano fu il principale artefice della realizzazione della libreria medicea. Per la biblioteca di Piero de' Medici cfr. F. Ames-Lewis, *The Library and Manuscript of Piero di Cosimo de' Medici*, New York, Garland Pub., 1984. Sull'inventa-

Clarissime atque eloquentissime Vir, post recomendatione premissa.

Il Prinio[35] e lla Decha[36] sono di tutto finiti in buona forma, in modo credo vi piaceranno sommamente. Aregli arechati chostassù ma alla porta non mi arebbono lasciato venire. Mandate uno chon qualche panno da involgergli, che non si guastino. Meser Piero Strozi[37] ha finite l'abreviationi[38] e holle dato a Pipo[39] <che> le minî, e subito si legaranno. Ad voi mi rachomando.

In Firenze, adì 23 di giugno 1458.

Vespasiano di Filippo

[a tergo]

<Clariss>imo atque elo<quenti>ssimo viro Piero <di Chosi>mo de' Medici, Ma<iori su>o honorando.

Concentrandoci sui tre elementi interpretativi sopra menzionati (persone, lessico, opere) possiamo identificare nella lettera: nomi di persona (Vespasiano di Filippo, Piero di Chosimo de' Medici, Piero Strozi, Pipo), luoghi (Firenze), date (23 di giugno 1458), testi e autori classici latini (Prinio, Decha), lessico tecnico della copia (minî, legaranno).

Dalla lettera deduciamo inoltre che: Piero de' Medici ha commissionato alla bottega di Vespasiano un Plinio e le Decadi di Livio e che la copia è sicuramente stata effettuata prima del giugno 1458. Apprendiamo inoltre che Piero Strozzi ha finito le Epitomi e ha consegnato il volume a tal Pipo, identificabile con Filippo Torelli, affinché proceda con le miniature, per poi effettuare la rilegatura del volume.

rio della biblioteca cfr. E. Piccolomini, «Ricerche sulle condizioni e vicende della libreria Medicea», in *Arch. St. Ital.*, serie III, vol. XIX (1874), pp. 115 sgg.

[35] Plinio, *Storia naturale*. Attuale manoscritto Laur. Plut. 82,3.

[36] *Decadi* di Tito Livio (in tre volumi: Laur. Plut. 63, 10-12). Qui si parla della III Deca, attuale Laur. Plut. 63, 12.

[37] Piero di Benedetto Strozzi (1416 ca.-1492), scriba fiorentino. Uno dei copisti più vicini a Vespasiano e descritto dallo stesso come «il più bello scriptore abbi avuto questà età et il più emendato» (si legge nella vita del padre Benedetto Strozzi scritta da Vespasiano, cfr. A. Greco (ed.), *Vite*, Firenze, Sansoni, 1970, II, p. 426). Fu copista molto attivo per la bottega di Vespasiano e forse quello che lavorò più a lungo (fra il 1443 e il 1483); di lui si conoscono 70 manoscritti, di cui solo 5 sottoscritti dallo stesso Piero (per cui cfr. A.C. de la Mare, *Messer Piero Strozzi, a Florentine priest and scribe*, in *Calligraphy and Paleography. Essays presented to Alfred Fairbank on his seventieth birthday*, A.S. Osley (ed.), London, Faber and Faber, 1965, pp. 55-68 e ancora Ead., *New research*, cit., pp. 365-591, in part. pp. 430-431 e App. I, 62, 7-9).

[38] Si intendano le *Epitomi* comprese nella IV Deca.

[39] «Pipo» è stato identificato dalla de la Mare in Filippo Torelli (1409-1468). Lavorò già per Michele Guarducci, nella bottega di via del Proconsolo. Più che miniaturista fu soprattutto specialista in iniziali e bordi (cfr. A.C. de la Mare, *Vespasiano da Bisticci, historian and bookseller*, PhD thesis, London University, 1969, p. 180, n. 45 e p. 188 per il manoscritto liviano, decorato dal Torelli).

La corrispondenza e le lettere, quindi epistolari e carteggi in generale, sono un buon esempio di documento storico, a cavallo fra filologia e archivistica, che fornisce un utile strumento per esemplificare un processo di modellazione.[40]

2. Scopo dell'edizione digitale

L'edizione digitale è dunque il risultato di un processo che individua nell'annotazione il primo «stato» dell'interpretazione. Comunque l'annotazione venga effettuata, embedded o stand-off, la formalizzazione traduce la codifica digitale in un atto «ermeneutico e semiotico».[41] Diremo che l'annotazione determina la possibilità d'estrazione d'informazioni dal testo marcato, ovvero il modello stabilisce le possibilità del processing.

Lo scopo che si intende qui perseguire partendo da un testo marcato è duplice: da un lato si vuole rappresentare l'informazione che è implicitamente connessa all'interno della fonte (persone con testi e autori classici, testi classici come codici descritti attraverso un certo lessico); dall'altro si intende stabilire un sistema di collegamento semantico tra gli elementi annotati in uno specifico contesto relazionale e le risorse disseminate su WWW, utili per descrivere tali elementi (persone, codici, lessico).

Il primo obiettivo è dunque di stabilire una rete di relazioni fra le persone e i codici, ovvero i manoscritti, legati alle persone, che sono identificabili secondo il ruolo o la funzione che su quel manoscritto

[40] Fra i contributi critici relativi all'edizione di collezioni di lettere si veda in particolare: E. Vanhoutte, R. Van den Branden, «Describing, transcribing, encoding, and editing modern correspondence material: a textbase approach», *Literary and Linguistic Computing*, XXIV/1 (2009), pp. 77-98. Fra i casi eccellenti di edizioni digitali di lettere si segnalano: *Vincent van Gogh, The Letters*, Van Gogh Museum and Huygens ING, http://vangoghletters.org/; *The Letters of Matthew Arnold. A digital edition*, C.Y. Lang (ed.), The University of Virginia Press, http://rotunda.upress.virginia.edu/arnold/; *Letters of William Herle Project*, Robyn Adams (ed.), hosted by the AHRC Centre for Editing Lives and Letters, http://www.livesandletters.ac.uk/herle/index.html; *The Linnaean Correspondence*, an electronic edition prepared by the Swedish Linnaeus Society, Uppsala, and published by the Centre international d'étude du XVIIIe siècle, Ferney-Voltaire, http://linnaeus.c18.net/; *Emily Dickinson's Correspondences*, M.N. Smith, L. Vetter (eds.), The University of Virginia Press, http://rotunda.upress.virginia.edu/edc/default.xqy.

[41] D. Fiormonte, V. Martiradonna, *La codifica digitale come atto ermeneutico e semiotico. Il caso di Valerio Magrelli*, in *Mémoire des textes - Textes de mémoire*, Atti del convegno internazionale (Università Paris X-Nanterre, 21-22 ottobre 2005), C. Cazalé (éd.), Paris, Presses universitaires de Paris Ouest, 2008.

esercitano (il copista, il possessore, il richiedente, il miniatore) e collegare quindi i manoscritti con il lessico utilizzato per descriverli. Queste informazioni, che nascono spontaneamente all'atto della lettura, devono essere espresse in modo formale per diventare oggetto di rappresentazione informatica e quindi di interrogazione da parte dell'utente finale.

Secondo obiettivo è creare relazioni fra le persone e le risorse utili per descrivere quelle persone a diversi livelli (da *authority records* a pubbliche prosopografie), manoscritti menzionati con i codici ad oggi esistenti e posseduti dalle biblioteche (cataloghi e descrizioni codicologiche), il lessico con *repositories* di parole tecniche (thesauri specialistici).

Questa modalità di riflessione sul testo riflette una nuova forma di apparato che si traduce in una nuova forma di rappresentazione della testualità e del suo corredo informativo. Ma potremmo anche ipotizzare una filologia del testo digitale che vede nella costruzione di relazioni formalmente dichiarate e nell'apertura al Web attraverso collegamenti semantici una nuova forma di *restitutio textus* come base di conoscenza.

Nei prossimi paragrafi si illustreranno i tre livelli di rappresentazione digitale dell'edizione.

2.1. L'*embedded markup*: l'annotazione XML/TEI

Il modello di markup di seguito illustrato vuole tradurre i tre elementi cardine dell'edizione in un sistema formale. Il modello dichiara quindi in forma esplicita i punti di accesso al contenuto della lettera. Tali punti di accesso rappresentano anche gli indici che costituiscono i possibili percorsi di navigazione dell'utente finale, secondo il principio della navigazione «a faccette»[42]:

a) persone citate nelle lettere;
b) testi citati e manoscritti realizzati dalla scuola di Vespasiano;
c) lessico tecnico della copia e del commercio librario.

A livello di markup è semplice rappresentare, attraverso adeguati elementi e attributi, questo tipo di dati. Gli esempi che seguono mostrano

[42] Sull'origine della classificazione a faccette nel campo della biblioteconomia si veda S.R. Ranganathan, *Theory of Library Catalogue*, Madras, Madras Library Association, 1938. Sull'uso di questo metodo nel settore dell'architettura dell'informazione si veda p.e. il blog di Luca Rosati, http://lucarosati.it/blog.

l'impiego della TEI P5[43] al caso riportato e descritto poco sopra con riferimenti a file in formato RDF[44] (descritti nella sezione 2.2).

a) Persone – marcatore <persname>. Identificazione dei nomi così come compaiono nella lettera attraverso una sigla che permette di disambiguare tutte le forme varianti dello stesso nome attribuendo un codice identificativo univoco (#iniziale del nome e iniziale del cognome)[45]:

```
<persname ref="people.rdf#PS">Piero Strozi</persname>
<persname ref="people.rdf#FT">Pipo</persname>
<persname ref="people.rdf#VdB">Vespasiano di Filippo</persname>
<persname ref="people.rdf#PdM">Piero di Chosimo de' Medici</persname>
```

b) Codici manoscritti – marcatore <bibl> con elementi author e/o title. Identificazione di autori e testi così come compaiono nella lettera attraverso una sigla, cioè un codice identificativo univoco (#iniziale del nome dell'autore_iniziale del titolo dell'opera_eventuale numero del libro):

```
<bibl ref="manuscripts.rdf#P_SN>
   <author>Prinio</author>
</bibl>

<bibl ref="manuscripts.rdf#L_D_III">
   <title>Decha</title>
</bibl>

<bibl ref="manuscripts.rdf#L_D_IV_E">
   <title>abreviationi</title>
</bibl>
```

c) Termini tecnici <term>. Identificazione di parole tecniche così come compaiono nella lettera attraverso una sigla, cioè un codice identificativo univoco (#abbreviazione):

[43] TEI Consortium (eds.), *Guidelines for Electronic Text Encoding and Interchange*, version 2.2.0 (last updated on 25th october 2012), http://www.tei-c.org/release/doc/tei-p5-doc/en/html/.

[44] RDF è acronimo per 'Resource Description Framework'. La specifica del W3C si legge all'indirizzo: http://www.w3.org/RDF.

[45] Un procedimento analogo si può leggere in: J. Cummings, «The William Godwin's Diaries Project», *Jahrbuch für Computerphilologie*, 10 (2008), http://computerphilologie.de/jg08/cummings.pdf.

<term type="binding" ref="lexicon.rdf#leg">legaranno</term>
<term type="illumination" ref="lexicon.rdf#min">minî</term>

Il markup TEI ci consente quindi di identificare: nomi propri (<persname>), codici manoscritti citati (<bibl>) e relativi autori (<author>) e titoli (<title>), lessico tecnico (<term>) relativo a vari e diversi ambiti (@type="writing|support|binding|illumination").

Il valore dell'attributo @ref (reference) aiuta una prima identificazione delle stringhe presenti all'interno del documento, puntando ad una descrizione dell'elemento conservata in un luogo (in una porzione [identificata da #] di un file) esterno al documento stesso. Nello specifico sono stati creati 3 diversi file .rdf: uno per le persone (people.rdf), uno per i codici manoscritti (manuscripts.rdf) ed un altro per il lessico (lexicon.rdf). Diremo che il valore di @ref è un URI[46] (*Uniform Resource Identifier*) relativo,[47] un identificativo univoco che punta ad una entità conservata in uno dei tre repositories esterni (people.rdf, manuscripts.rdf, lexicon.rdf). Utilizzando il meccanismo degli URIs è possibile creare collegamenti fra il contenuto della lettera ed informazioni descrittive sulle stringhe annotate, conservate in documenti esterni. Il valore dell'attributo @ref, quindi l'URI, è un collegamento alla descrizione formale di quella stringa annotata: per esempio il valore, vale a dire l'URI, people.rdf#PdM punta alla specifica sezione (#) del file people.rdf dove Piero de' Medici, identificato con PdM, viene descritto (cfr. sezione 2.2 punto *a)*).

Se il meccanismo degli URIs ci dà modo di identificare univocamente ogni stringa, e di collegarla ad un documento esterno in cui quella stringa viene descritta, è necessario prevedere un sistema per adempiere al compito primario dell'edizione: costruire relazioni fra gli elementi annotati dotati di un URI. Questi collegamenti devono essere in grado di rispondere a domande del tipo: che relazione esiste fra «Piero de' Medici» (people.rdf#PdM) e uno specifico esemplare della *Storia naturale* di Plinio (manuscript.rdf#P_SN)? E quale relazione esiste fra lo stesso manoscritto (manuscript.rdf#P_SN) e la parola «legaranno» (lexicon.rdf#leg)?

Ma il modello dovrebbe essere in grado di rispondere anche a domande del tipo: quale esemplare della *Storia naturale* è stato realizzato (qual è

[46] Si veda *Uniform Resource Identifier (URI) Activity Statement*, http://www.w3.org/Addressing/Activity.

[47] Diciamo che negli esempi che seguono viene fornito il riferimento al *path* della risorsa dotata di URI. Il valore assoluto dell'URI potrebbe essere: http://edizione-lettere-vespasiano.it/base-di-conoscenza. A cui seguono /people.rdf, /manuscripts.rdf, /lexicon.rdf.

la sua segnatura odierna) e in quale biblioteca è ad oggi conservato? Chi l'ha copiato? Chi l'ha miniato? Chi è Piero de' Medici?

Per rispondere al primo set di domande (cfr. sezione 2.2) è necessario aggiungere semantica alla descrizione del documento. Non è sufficiente quindi classificare le stringhe come persone, codici manoscritti o lessico, ma esprimere formalmente, ed in modo processabile dalla macchina, informazioni aggiuntive sulle stringhe e descrivere le relazioni. Il linguaggio RDF dà la possibilità di esprimere asserti, sotto forma di triple (soggetto, predicato, oggetto) che permettono di esplicitare relazioni fra le stringhe identificate attraverso gli URIs.

Per il secondo set di domande (cfr. sezione 2.3) è necessario creare collegamenti fra l'edizione e risorse su WWW capaci di estendere la descrizione degli elementi annotati. Questo significa agevolare il dialogo fra URIs interni all'edizione e URIs esterni correlati, come vedremo, avvalendosi dei predicati ontologici in uso nei principali modelli concettuali esistenti.

2.2. L'informazione esterna: i file RDF

Una volta che il testo è stato annotato prevedendo forme di dialogo con oggetti esterni è necessario comprendere come esprimere attraverso URIs e RDF le relazioni interne e quindi capire come realizzare i file RDF. Con gli URIs è possibile una prima identificazione univoca di stringhe riconducendole ad una forma accettata.[48] Con RDF è possibile esprimere le relazioni fra gli elementi dotati di un URI. Ma non solo. Ogni elemento dotato di identificazione può intrattenere una relazione di un qualche tipo con una stringa, sia essa un valore letterale o un altro URI.

a) Persone = people.rdf <rdf:description rdf:about="URI">. Le persone possono avere diverse proprietà dedotte dal testo ed essere messe quindi in relazione con i codici menzionati.

ELEMENTI DEDOTTI DAL TESTO	ESEMPIO
Identificazione univoca (frammento di URI)	#PdM

[48] Si veda anche il progetto EATS che usa il concetto di PSI (*Public Subject Indicator*) come strumento per creare *authority files* di nomi di persona. J. Norrish. *EATS: an Entity Authority Tool Set*, NZETC (New Zealand Electronic Text Centre), 2007. URI: http://hdl.handle.net/10063/220. Database: http://authority.nzetc.org/.

Normalizzazione del nome (*entry* per un'*authority record* o definizione di un'*access key*)	Medici, Piero de'
Forme varianti del nome (estratte in automatico per condivisione dell'URI)	Piero, Piero di Cosimo de' Medici, Principe di Firenze
Relazioni con codici manoscritti	possessore-di Plinio, *Storia naturale*; Livio, *Decadi* (III e IV Deca comprese *Epitomi*)

In linguaggio formale possiamo esprimere questi concetti come relazioni attraverso i seguenti asserti:

SOGGETTO	PREDICATO	OGGETTO
people.rdf#PdM	has_normalized_form	Medici, Piero de'
	has_variant_forms	Piero, Piero di Cosimo de' Medici, Principe di Firenze
	is_owner_of	manuscripts.rdf#P_SN manuscripts.rdf#L_D_III manuscripts.rdf#L_D_IV_E

Possiamo quindi associare ad ogni stringa identificata nel documento con <persname> diversi tipi di proprietà: un identificatore univoco (definito attraverso un URI); alcune caratteristiche estratte dal documento stesso (forme varianti del nome); altri elementi dedotti da repertori condivisi (la forma accettata o normalizzata dei nomi); altre proprietà dedotte dalla lettura del documento (la relazione fra la persona e un codice).

Un primo passo verso l'interoperabilità è quello di collegare l'URI definito internamente con l'identificatore univoco stabilito in altri progetti e riferito allo stesso contenuto. Riferendo il nostro URI analogo (OWL:same-as) all'identificatore utilizzato in progetti come VIAF,[49] che stabilisce le forme controllate dei nomi, è possibile far dialogare le risorse e creare collegamenti con altri sistemi standard come LCA[50] e MARC.[51] E su questo concetto torneremo nella sezione 3.

[49] *Virtual International Authority File*, http://viaf.org.

[50] *Library of Congress Authorities*, http://authorities.loc.gov.

[51] *Machine-Readable Cataloguing*, http://www.loc.gov/marc.

b) Codici manoscritti = manuscripts.rdf <rdf:description rdf:about="URI">. Allo stesso modo i codici possono essere identificati attraverso un URI e normalizzati utilizzando repertori specifici.[52] Andranno poi ovviamente posti in relazione con le persone con le quali quei codici sono in qualche modo collegati.

ELEMENTI DEDOTTI DAL TESTO	ESEMPIO
Identificazione univoca (frammento di URI)	#P_SN
Normalizzazione	Plinio, *Storia naturale*
Relazioni con persone (possessore, copista, miniatore, richiedente)	richiesto-da Piero de' Medici posseduto-da Piero de' Medici copiato-da Piero Strozzi miniato-da Filippo Torelli

In linguaggio formale possiamo esprimere questi concetti come relazioni attraverso i seguenti asserti:

SOGGETTO	PREDICATO	OGGETTO
manuscripts.rdf#P_SN	has_normalized_form	Plinio, *Storia naturale*
	is_requested_by is_owned_by is_copied_by is_illuminated_by	people.rdf#PdM people.rdf#PdM people.rdf#PS people.rdf#FT

c) Termini tecnici = lexicon.rdf <rdf:description rdf:about="URI">. Anche i termini tecnici, dotati di URI, andranno normalizzati e posti in relazione con i codici a cui si riferiscono.

ELEMENTI DEDOTTI DAL TESTO	ESEMPIO
Identificazione univoca (frammento di URI)	#min

[52] Nel nostro caso l'intento è di utilizzare il thesaurus di *Gnomon*, http://www.gnomon.ku-eichstaett.de/Gnomon/en/Gnomon.html.

Normalizzazione	miniare, miniatura, miniato
Relazioni fra il lessico e il codice a cui il lessico si riferisce	relativo-a *Epitomi* della IV Deca di Tito Livio

In linguaggio formale possiamo esprimere questi concetti come relazioni attraverso i seguenti asserti:

SOGGETTO	PREDICATO	OGGETTO
lexicon.rdf#min	has_normalized_form	miniare, miniatura, miniato
	is_referred_to	manuscripts.rdf#L_D_IV_E

3. I collegamenti con il WWW: la base di conoscenza

Rappresentare contenuti attraverso l'impiego di RDF e di ontologie è un modello che in vari progetti ha dimostrato la sua efficacia.[53] Diremo che il framework RDF, come mezzo per esprimere asserti, associato all'ontologia, come concettualizzazione di un dominio, sono strumenti finalizzati alla rappresentazione della conoscenza implicita dell'editore critico.

Costruire un'ontologia significa, oltre ad usare linguaggi formali come XML e RDF, stabilire classi, definire proprietà, nella forma dei predicati, ed esprimere valori associati a quelle proprietà, rispetto al dominio di analisi. Nel nostro esempio le tre categorie sulle quali abbiamo ragionato

[53] Si veda in particolare il progetto *Henry the III Fine Rolls Project*, CCH (resp.), The National Archives and King's College London, http://www.finerollshenry3.org.uk/home.html. La descrizione del progetto si può leggere in: A. Ciula, P. Spence, J.M. Vieira, «Expressing Complex Associations in Medieval Historical Documents: The Henry III Fine Rolls Project», *Literary and Linguistic Computing*, XXIII/3 (2008), pp. 311-325. Una descrizione degli aspetti implementativi si legge in: J.M. Vieira, A. Ciula, «Implementing an RDF/OWL Ontology on Henry the III Fine Rolls», paper presented at *OWLED* 2007 (Innsbruck, june 6-7 2007), http://www.webont.org/owled/2007/PapersPDF/submission_6.pdf. Altro progetto interessante, che ha condotto alla realizzazione di una specifica ontologia, è SAWS (*Sharing Ancient WisdomS*), http://www.ancientwisdoms.ac.uk/. Il progetto nasce dalla collaborazione fra il King's College di Londra, il Newman Institute di Uppsala e l'Università di Vienna. La descrizione del progetto si può leggere in: S. Dunn *et al.*, «Sharing Ancient Wisdoms: developing structures for tracking cultural dynamics by linking moral and philosophical anthologies with their source and recipient texts», in *Digital Humanities 2012*, Conference Abstracts (University of Hamburg, july 16-22 2012), pp. 176-179.

(persone, termini tecnici e codici manoscritti) sono tre possibili classi di un'ontologia di dominio, dotate di proprietà (i predicati) e relativo valore (gli oggetti). Ogni istanza della nostra classe identificata in modo univoco (il soggetto identificato attraverso URI) intrattiene relazioni di varia natura (le proprietà che diventano predicati), con altre istanze (i valori delle proprietà che diventano oggetti) che possono essere dei valori letterali o a loro volta delle entità dotate di identificativi univoci (vedi asserti nella sezione 2). Passiamo quindi dall'edizione come classificazione e categorizzazione degli elementi annotati (il markup) alla rappresentazione di concetti (l'ontologia) che trasformano il testo in una base di conoscenza.

Semantic relations that are essential in humanities can be expressed by creating an ontology. After storing information in an ontology, consecutive processing, like logic reasoning, creates further possibilities like consistency checks or support by merging data.[54]

Il passaggio dall'elemento marcato alla classe a cui può essere riferito consente una prima generalizzazione dell'appartenenza di una stringa ad un dominio semantico. La costruzione di relazioni fra le classi permette di realizzare un modello che può consentire il recupero di informazione strutturata. Non solo potrò ricercare singole stringhe, ma potrò interrogare la mia edizione alla ricerca di concetti: chi ha miniato un certo codice? chi è il possessore di un certo manoscritto? Ogni istanza di ciascuna classe avrà un contenuto determinato dal contesto in cui compare e sarà portatrice di informazione specifica per il dominio di riferimento.

Ma un'edizione non è conoscenza fino a quando rimane isolata ('siloed') e fino a quando non è in grado di dialogare con altre edizioni. Due sono quindi gli aspetti che ci interessano. Come trasformare l'edizione in un sistema interoperabile e come far dialogare l'edizione con il Web.

Scegliere predicati ontologici *ad hoc* per il tipo di progetto limita l'interoperabilità. Un generico predicato is_related_to, proprietà abitualmente utilizzata per collegare due elementi, è stato trasformato nel nostro esempio in una pluralità di relazioni che spaziano dalla forma normalizzata del nome (has_normalized_form) alle forme varianti (has_variant_forms), dal collegamento fra una persona e un codice (is_illuminator_of) a quello fra un codice e il lessico utilizzato per descriverla (is_described_by). Nulla ci impedirebbe di estendere le relazioni alle date

[54] A. Zöllner-Weber, «Ontologies and Logic Reasoning as Tools in Humanities?», *Digital Humanities Quarterly*, III/4 (2009).

e ai luoghi, tanto per fare un esempio. Per garantire massimo interscambio dell'edizione e quindi dell'ontologia andrà realizzato un *mapping* fra il modello usato e le principali ontologie esistenti nel dominio umanistico, sia a livello di classi che di proprietà.[55] FRBR,[56] DC terms,[57] SKOS,[58] EDM[59] e CIDOC-CRM[60] sono alcune fra le ontologie che si stanno studiando, per stabilire forme di dialogo e di condivisione.[61]

Ma l'informazione interna, rappresentata attraverso il markup, e l'informazione esterna, rappresentata attraverso RDF, devono poter dialogare anche con il mondo del Web. Se infatti persone, codici e lessico intrattengono relazioni interne al documento sono inevitabilmente anche in relazione con quanto è necessario per approfondire le istanze di ciascuna di queste categorie.

Sono quindi in fase di analisi possibili estensioni degli elementi della descrizione e delle loro relazioni. Andranno stabilite relazioni fra l'URI interno al progetto e altri URIs necessari ad approfondire gli elementi della descrizione. Ovviamente anche in questo contesto andranno valutati i predicati più adeguati ad esprimere queste relazioni. Vediamo partitamente come estendere la descrizione di ciascuna delle tre categorie o classi.

Persone. Un'istanza della classe persona, oltre ad essere dotata di un URI e di forme varianti del nome, ha diverse proprietà generali: un luogo e una data di nascita e di morte, una biografia, elementi iconografici che la qualificano.[62]

Esempio:

[55] Un esempio di *mapping*: Ø. Eide, C.E. Ore, «TEI, CIDOC-CRM and a possible interface between the two», in *Digital Humanities 2006*, The First ADHO International Conference, Conference Abstracts (Université Paris-Sorbonne, july 5-9 2006), pp. 62-65, http://www.allc-ach2006.colloques.paris-sorbonne.fr/DHs.pdf. Si veda anche: Ø. Eide, C.E. Ore, «TEI and cultural heritage ontologies: Exchange of information?», *Literary and Linguistic Computing*, XXIV/2 (2009), pp. 161-172.

[56] *Functional Requirements for Bibliographic Records*, http://www.ifla.org/publications/functional-requirements-for-bibliographic-records.

[57] *Dublin Core Terms*, http://dublincore.org/documents/dcmi-terms/.

[58] *SKOS Simple Knowledge Organization System Reference*, http://www.w3.org/TR/skos-reference.

[59] *Europeana Data Model*, http://pro.europeana.eu/edm-documentation.

[60] *Cidoc-Conceptual Reference Model*, http://www.cidoc-crm.org/.

[61] Per una descrizione della funzione di ciascuna di queste ontologie si veda il W3C Incubator Group Report, *Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets*, 25 october 2011, http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/.

[62] La TEI prevede una sezione specifica per la gestione di informazioni relative alle persone: TEI Consortium (eds.), «13.3 Biographical and Prosopographical Data», in

<rdf:description rdf:about="#LM">
Access_key: Medici, Lorenzo de'
Same-as VIAF permalink: http://viaf.org/viaf/54169908
Biography: http://it.wikipedia.org/wiki/Lorenzo_de'_Medici
Father-of: Piero de' Medici (people.rdf#PM)
Born-in: Firenze (place.rdf#Fi)
Died-in: Firenze (place.rdf#Fi)
Born-when: 1449 (date.rdf#1449)
Died-when: 1492 (date.rdf#1492)
Iconography: http://www.google.com/search?hl=en&q=lorenzo+de%27+medici&um=1&ie=UTF-8&tbm=isch&source=og&sa=N&tab=wi&ei=_1vcUOeAJqeI4ASikoCYBQ&biw=1146&bih=709&sei=AVzcUNnjIOXV4gTVloHQCQ

Un'ontologia come il CIDOC-CRM può essere uno strumento utile per creare relazioni fra persone, eventi, oggetti, luoghi e date.

Codici manoscritti. Molti manoscritti menzionati sono codici ad oggi posseduti da numerose biblioteche europee. Alla forma normalizzata del nome, come desunta dai repertori, andranno allora associate la segnatura, l'indicazione dell'istituto di conservazione ma anche un'eventuale descrizione codicologica o la disponibilità dell'immagine digitale del codice o del suo full-text.

Esempio:

<rdf:description rdf:about="#PV">
Access_key: Plutarco, *Vite Parallele*
Shelfmark: Laur. Plut. 65, 26
Available-in: Biblioteca Medicea Laurenziana, Firenze http://www.bml.firenze.sbn.it/
Described-in: http://opac.bml.firenze.sbn.it/Record.htm?idlist=1&record=710312453859
Item-in(image): http://teca.bmlonline.it/TecaViewer/index.jsp?RisIdr=TECA0000777138
Item-in (full-text): --

Un'ontologia come FRBR potrebbe essere usata per descrivere la relazione fra il codice inteso come oggetto fisico, quindi l'*item*, dotato di una segnatura – che può includere anche diverse sue rappresentazioni o espressioni (file di testo, file immagine, audio o video) – e l'opera di cui il codice è una specifica manifestazione.

Guidelines for Electronic Text Encoding and Interchange (last updated on 21 december 2011), http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html#NDPERS.

Termini tecnici. Anche il lessico deve essere qualificato. I termini vanno quindi descritti e vanno stabilite relazioni con altri repertori.

Esempio:

<rdf:description rdf:about="#aF">
Access_key: Ad fragmenta
Definition: «Sistema di resa scrittoria opposto alla scrittura "ad volumina". ... (S. Rizzo, *Lessico filologico degli umanisti*, Roma, Edizioni di Storia e Letteratura, 1973, p. 196)».
Type-of: writing
Iperonym: scrittura
Synonim: pecia
Antonym: ad volumina

L'ontologia SKOS fornisce uno strumento per creare thesauri terminologici e stabilire collegamenti fra i termini utilizzati nel documento e gli stessi definiti in altri repertori (come ad esempio la rete lessicale WordNet[63]).

Nell'ambito del Web semantico, quindi di modelli e linguaggi di rappresentazione della conoscenza, un ruolo importante, ormai da alcuni anni, ha assunto Linked Data[64] come progetto e strumento per la gestione di collegamenti fra risorse e quindi per il trattamento dei link. Molti *linked datasets* sono già disponibili sul WWW e possono di conseguenza essere utilizzati in ogni progetto, cioè richiamati al fine di associare istanze dello stesso concetto.[65] Molti URI di risorse e molte asserzioni RDF sono quindi a disposizione della comunità, con lo scopo di favorire il dialogo e lo scambio fra progetti.[66] Rappresentando le informazioni relative alle edizioni utilizzando il meccanismo degli URIs e le modalità espressive di RDF è possibile contribuire a popolare l'universo dei linked data sets, in una prospettiva di interscambio e interoperabilità.

[63] http://wordnet.princeton.edu/.

[64] Un ottimo volume di riferimento è: T. Heath, C. Bizer, *Linked Data: Evolving the Web into a Global Data Space*, San Rafael, CA., Morgan & Claypool, 2011. Si veda anche: C. Bizer, T. Heath, M. Hepp (eds.), «Special Issue on Linked Data», *International Journal on Semantic Web and Information System*, V/3 (2009).

[65] Un interessante progetto in questa direzione si legge in: M.O. Jewell, «Semantic Screenplays: Preparing TEI for Linked Data», in *Digital Humanities 2010*, Conference Abstracts (London, july 7-10 2010), pp. 53-55, http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/pdf/ab-878.pdf/.

[66] SWEO Community Project, *Linking Open Data on the Semantic Web* (last update 23 august 2011), http://www.w3.org/wiki/TaskForces/CommunityProjects/LinkingOpenData/DataSets.

L'edizione digitale diventa quindi una base di conoscenza che permette all'utente finale di reperire informazioni strutturate. Le relazioni latenti sono definite in modo formale, diventando oggetto di query. I commenti, che l'editore mette tradizionalmente in nota, sono espressi tramite RDF e predicati ontologici. Ogni elemento annotato è identificato in modo univoco. L'edizione digitale così realizzata entra in un processo di dialogo con altre risorse digitali correlate.

L'obiettivo della creazione di una base di conoscenza è duplice: da un lato garantisce l'interoperabilità dell'edizione, aperta allo scambio con altri repositories; dall'altro dà agli studiosi uno strumento (liste di autorità espresse in modo formale e accessibili tramite URIs) e un metodo (dal contenuto delle lettere all'informazione esterna e dalle informazioni esterne al testo) che possono essere utilizzati in situazioni e contesti diversi. Un modello che sfrutta le tecnologie del Web semantico per la rappresentazione della conoscenza, e linked data per l'interscambio.[67]

E torniamo quindi all'interrogativo di partenza. Possiamo ancora parlare, a questo punto, di edizione?[68]

Università di Bologna
Dipartimento di Filologia Classica e Italianistica
francesca.tomasi@unibo.it

[67] Si veda anche: G. Tummarello, C. Morbidoni, E. Pierazzo, «Toward Textual Encoding Based on RDF», in *9th ICCC International Conference on Electronic Publishing* (Belgium, Leuven-Heverlee, june 2005).

[68] K.M. Price, «Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?», *Digital Humanities Quarterly*, III/3 (2009).

Progetto grafico e impaginazione: Carolina Valcárcel
(Centro para la Edición de los Clásicos Españoles)

1ª edizione, maggio 2013
© copyright 2013 by
Carocci editore S.p.A., Roma

Finito di stampare nel maggio 2013
dalla Litografia Varo (Pisa)

ISBN 978-88-430-6451-9

Riproduzione vietata ai sensi di legge
(art. 171 della legge 22 aprile 1941, n. 633)

Senza regolare autorizzazione,
è vietato riprodurre questo volume
anche parzialmente e con qualsiasi mezzo,
compresa la fotocopia, anche per uso
interno o didattico.