

Dig *Italia*

Anno VII, Numero 2 - **2012**

ISSN 1972-6201

Rivista del digitale nei beni culturali

ICCU-ROMA

Il progetto di digitalizzazione degli incunaboli italiani in lingua volgare*

Gianfranco Crupi

Università degli Studi di Roma "La Sapienza"

Una lunga storia quella dell'incunabolistica, che prende le mosse tra la Germania e l'Inghilterra sul finire del XVIII e l'inizio del XIX secolo¹, e che giunge fino ai nostri giorni, con i dispositivi elettronici e digitali che consentono nuove forme di organizzazione, di campionamento e di visualizzazione dei dati e dei documenti. A partire dallo strumento catalografico d'eccellenza della tradizione incunabolistica, il *Gesamtkatalog der Wiegendrucke* (GW), disponibile online da ormai un decennio² e che ad oggi conta più di 35.000 record, metà dei quali (corrispondenti ai volumi relativi agli autori con iniziali J-Z) riproducono in formato digitale le schede della versione manoscritta del catalogo³; e poi il repertorio *The Incunabula Short Title Catalogue* (ISTC)⁴, realizzato dalla British Library con il contributo di istituzioni bibliotecarie di tutto il mondo, che raccoglie e unifica i principali cataloghi nazionali, a partire dal nostro Indice Generale degli Incunaboli (IGI). Ad essi, vanno affiancati altri importanti strumenti di sussidio alla conoscenza dei primi prodotti della tipografia manuale (relativi al frontespizio, alle legature, ai possessori ecc.), su cui non è il caso, in questa sede, di soffermarsi oltre. Un discorso a parte meritano invece le collezioni digitali, che si costituiscono intorno a specifici *corpora* testuali, vere e proprie biblioteche, strutture di servizio organizzate, in cui le collezioni sono al centro di un sistema di relazioni, che supportano l'intero ciclo di vita dei documenti digitali e i servizi creati per l'accesso e per il recupero delle informazioni.

* L'articolo anticipa in sintesi le risultanze di un più articolato contributo sul progetto, che è in corso di stampa.

¹ La storia dell'incunabolistica, soprattutto novecentesca, è stata oggetto, anche di recente, di puntuali ricostruzioni storico-critiche, a partire dall'appassionante monografia di Edoardo Barbieri, *Haebler contro Haebler. Appunti per una storia dell'incunabolistica novecentesca*, Milano: Università Cattolica del Sacro Cuore, 2008.

² <<http://www.gesamtkatalogderwiegendrucke.de>>. Cfr. Ninon Suckow – Werner Klarkowski, *Die Datenbank "Gesamtkatalog der Wiegendrucke"*, «Zeitschrift für Bibliothekswesen und Bibliographie», 51 (2004), n. 4, p. 200-206.

³ A partire dal volume X del GW, per ogni scheda è fornita l'indicazione della fonte impiegata per la descrizione.

⁴ <<http://www.bl.uk/catalogues/istc/index.html>>. L'ultimo aggiornamento di ISTC, che risale all'8 gennaio 2008, registra 29.777 edizioni.

Tralasciando le tante iniziative di digitalizzazione, relative ai patrimoni di singole biblioteche, e concentrandosi su quelle che riguardano interi corpora testuali relativi al libro antico, spiccano la *Verteilte digitale Inkunabelbibliothek* (VdIB) e la collezione di incunaboli italiani in lingua volgare.

La *Verteilte digitale Inkunabelbibliothek*, nata nel 2003 dalla collaborazione tra la Universitäts- und Stadtbibliothek Köln e la Herzog August Bibliothek di Wolfenbüttel (Niedersachsen), rende disponibile in rete il patrimonio incunabolistico posseduto dalle due biblioteche (circa mille esemplari)⁵. Un solido apparato di metadattazione, un efficiente ed efficace dispositivo di visualizzazione e un'interfaccia amichevole che ospita agili strumenti di navigazione (sebbene solo in lingua tedesca), rende la VdIB un sicuro punto di riferimento nel settore e un prezioso strumento a disposizione degli studiosi. Ad essa complementare è, inoltre, la raccolta di collezioni digitali della Bayerische Staatsbibliothek che conta oggi quasi 900.000 documenti digitalizzati, tra cui molti incunaboli⁶.

L'altra collezione di riferimento, oggetto di questo articolo, è la raccolta degli incunaboli italiani in lingua volgare realizzata da Biblioteca Italiana (BibIt) dell'Università La Sapienza di Roma⁷.

Il progetto

Quando nel 2007 il Dipartimento di Italianistica e spettacolo⁸ dell'Università degli studi di Roma "La Sapienza" avviò il progetto di digitalizzazione degli incunaboli italiani in lingua volgare⁹, la base dati ISTC (Incunabula Short Title Catalogue)¹⁰ ne registrava 2347. A un più attento esame bibliografico, il corpus fu ulteriormente circoscritto, poiché 24 incunaboli risultavano attribuiti da ISTC a biblioteche che non li possedevano più o che dichiaravano di non averli mai posseduti e 111 edizioni furono escluse perché falsi incunaboli; perché gli esemplari disponibili non erano riproducibili a causa del loro critico stato di conservazione o perché appartenenti a collezioni private o a fondi all'epoca inaccessibili; infine, perché in latino e non in lingua volgare.

⁵ <<http://inkunabeln.ub.uni-koeln.de/>>.

⁶ <<http://mdz1.bib-bvb.de/~mdz/index.html>>. Le politiche di digitalizzazione della Bayerische Stadtbibliothek sono promosse e gestite da uno specifico centro, il Münchener Digitalisierungszentrum (MDZ), ad oggi uno dei migliori e più avanzati laboratori internazionali per la digitalizzazione.

⁷ Il database sarà reso di pubblico dominio entro il 2013.

⁸ Oggi: Dipartimento di studi greco-latini, italiani, scenico-musicali.

⁹ Il progetto, approvato dalla Direzione generale per i beni librari e gli istituti culturali del MiBAC, fu finanziato dalla Fondazione BEIC (Biblioteca europea di informazione e cultura) che ne affidò l'esecuzione al Dipartimento di italianistica e spettacolo (DISp). La convenzione fu sottoscritta il 2 gennaio 2007, da Antonio Padoa Schioppa, per la BEIC, e da Amedeo Quondam, per il DISp.

¹⁰ <<http://www.bl.uk/catalogues/istc/>>.

Dunque, 2212 unità, distribuite in 151 biblioteche, di cui 78 italiane e 73 estere (tra cui la Biblioteca Apostolica Vaticana).

Tuttavia, il numero degli incunaboli realmente acquisiti ammonta a 1711, in quanto per 501 esemplari (poco meno di un quarto del corpus) non si è potuto procedere alla loro digitalizzazione:

- 219 per dichiarata o implicita indisponibilità delle istituzioni, che li posseggono, a partecipare al progetto: 50 relativi a biblioteche italiane, tra cui la Biblioteca Ambrosiana (30); 199 relativi a biblioteche estere, tra cui la British Library (121);
- 130 per l'eccessivo e insostenibile costo richiesto per le riproduzioni: 4 relativi a biblioteche italiane; 126 relativi a biblioteche estere, tra cui la Biblioteca Apostolica Vaticana (69);
- 75 perché la negoziazione non è andata a buon fine: 43 relativi a biblioteche estere, 32 relativi a biblioteche italiane;
- 56 perché l'istituzione conservatrice aveva difficoltà a gestire i processi di digitalizzazione (Biblioteca Capitulare Colombina di Siviglia)
- 18 perché agli ordini inoltrati non ha mai fatto seguito l'invio del materiale richiesto, nonostante i ripetuti solleciti: 18 relativi a biblioteche italiane; 3 relativi a biblioteche estere.

Sono dati che fanno riflettere. Perché fanno da specchio a inerzie culturali, gestionali, professionali, dal momento che i fattori che hanno ritardato, ostacolato e, in parte, pregiudicato la compiuta realizzazione del progetto, sono infatti ascrivibili a diverse cause, tra cui spiccano in modo preoccupante quella *culturale*, relativa soprattutto ad alcune importanti istituzioni internazionali, che hanno manifestato una concezione "proprietaria" del loro patrimonio bibliografico, che contraddice la cultura della rete e del libero accesso alle fonti della conoscenza; e quello relativo alle *competenze professionali* dei bibliotecari, dal momento che l'80% delle biblioteche contattate ha mostrato di avere poca dimestichezza con il vocabolario di base del digitale e le più elementari applicazioni tecnologiche. Ci sono poi altre ragioni, come si è visto: di tipo *procedurale e amministrativo* (l'exasperante ritardo nella risposta - se non il silenzio - da parte delle istituzioni bibliotecarie contattate, tra cui, ancora una volta, spiccano alcune importanti biblioteche internazionali); ed *economico* (la sorprendente divergenza esistente tra le istituzioni che hanno offerto la cessione a titolo gratuito dei documenti digitalizzati e quelle che hanno assurdamamente preteso anche più di 10 euro a immagine).

La gestione del progetto

La gestione dell'intero progetto ha fatto capo alla Fondazione BEIC¹¹, che ha fissato gli standard di riferimento per le procedure di digitalizzazione e di metadattazione e che ne ha verificato l'applicazione, e al Dipartimento di Italianistica e Spettacolo, che ha affidato a Biblioteca Italiana (BibIt)¹², allora presieduta da Amedeo Quondam e da me diretta (fino a luglio 2011), il coordinamento scientifico e l'esecuzione operativa, finalizzata alla creazione di un'organica collezione digitale del libro antico a stampa¹³.

Prioritariamente è stato predisposto un capitolato esecutivo, per la fornitura del servizio da parte delle società che avrebbero dovuto digitalizzare grossi lotti di documenti, e la redazione di un più sintetico documento, a beneficio delle biblioteche che avrebbero potuto provvedere con mezzi propri alla digitalizzazione e metadattazione di pochi esemplari. Nel caso delle biblioteche romane¹⁴, Biblioteca Italiana ha acquisito direttamente in loco le immagini relative a 582 incunaboli, grazie a una convenzione sottoscritta con il CASPUR, che ha messo a disposizione del progetto uno scanner Biblioscan e un operatore tecnico espressamente formato all'uso del dispositivo di scansione¹⁵.

Per l'acquisizione digitale dei documenti sono stati adottati alcuni parametri di riferimento standard, concernenti la generazione dei formati, la risoluzione spaziale, la profondità del colore di ciascuna immagine, nonché il rispetto dei protocolli di acquisizione, relativi all'inquadratura, alla rifilatura dei documenti, al trattamento delle pagine bianche, ai dispositivi utilizzati e alle modalità del loro impiego.

La digitalizzazione

Le immagini prodotte dalla digitalizzazione sono state salvate in formato TIFF con compressione LZW tale da consentire la successiva rielaborazione e il loro riutilizzo, oltre che costituire la copia master finalizzata alla conservazione. Per quanto riguarda la risoluzione spaziale e la profondità del colore si è optato come valori

¹¹ <http://www.beic.it/>.

¹² <http://www.bibliotecaitaliana.it/>.

¹³ Un'articolata struttura organizzativa ha diretto e controllato, all'interno di Biblioteca Italiana, il corretto adempimento di tutte le fasi lavorative (scientifiche, tecniche e amministrative): Amedeo Quondam, coordinamento scientifico del progetto; Gianfranco Crupi, direzione operativa; Fabio Ciotti, consulenza tecnologica; Francesca Ferrario, coordinamento delle attività di redazione; E. Albano, segreteria di produzione; D. Ranaldi, unità operativa per la generazione dei metadati; I. Macculi, unità operativa per la generazione dei metadati; R. Castrogiovanni, unità di redazione per le attività di backup e conservazione dei dati; M. Scarbaci, unità operativa per la scansione dei testi.

¹⁴ Biblioteca nazionale centrale di Roma (96 incunaboli), Biblioteca Casanatense (247), Biblioteca Corsiniana (199), Biblioteca Angelica (12), Biblioteca universitaria Alessandrina (23), Biblioteca Vallicelliana (5).

¹⁵ L'intera operazione, puntualmente seguita dai responsabili di settore delle biblioteche coinvolte, è avvenuta sotto copertura assicurativa.

minimi di riferimento per una risoluzione di 300 dpi e una profondità del colore a 24 bit. Non sono stati ammessi la riproduzione di immagini interpolate o ricampionate (sia in termini di risoluzione che di profondità del colore), né file TIFF ottenuti tramite trasformazione da file originali compressi (es. da JPEG). Tutti i documenti sono stati digitalizzati rispettando la loro integrità e l'ordine sequenziale delle carte (incluse le pagine bianche) e sono stati acquisiti in modo che ciascuna pagina o facciata fosse interamente compresa nella stessa ripresa digitale, inclusi margini e ogni altro elemento extratestuale. Tuttavia, le immagini sono state rifilate in maniera da evitare l'acquisizione di zone estese al di fuori della pagina, pur garantendo di contenere per intero la pagina stessa e i relativi target, color bar e righe. Ad ogni pagina corrisponde un file immagine denominato con un identificativo univoco, elaborato secondo una sintassi preordinata, e il totale dei file immagini relativi a ciascun volume è stato organizzato in una cartella, definita secondo un formalismo che identifica univocamente il volume digitalizzato sul supporto di archiviazione. Il supporto digitale prodotto, con le opportune garanzie di inalterabilità nel tempo per la conservazione, costituisce la copia master di archivio, riprodotta in triplice copia.

Contestualmente sono stati definiti il set dei metadati tecnici e gestionali e i criteri di archiviazione di ciascun file, all'interno della cartella contenente tutte le immagini, corrispondenti a ciascuna unità bibliografica. Altrettanto è stato fatto dalle società incaricate della digitalizzazione di alcuni lotti di libri, per le procedure di memorizzazione dei documenti digitali e dei relativi file di metadati, che sono stati riversati su supporti magneto-ottici (DVD-R); in particolare, tutti i file XML dei metadati, oltre ad essere contenuti nelle rispettive cartelle delle immagini, sono stati accorpati su un unico supporto magneto-ottico (DVD-R) e la corrispondenza tra file-immagine e file-metadati è stata assicurata dall'adozione della medesima nomenclatura.

Nel caso di grossi lotti, la verifica in corso d'opera sulla correttezza e qualità delle operazioni di acquisizione si è svolta con l'estrazione casuale di un campione di almeno il 10% delle pagine per volume, sul quale è stata effettuata la prova di conformità, comprensiva dell'analisi dei file contenenti la nomenclatura dei file-immagini prodotti (per controllare la presenza di dati omessi, duplicazioni e incongruenze), e l'analisi dei metadati, corrispondenti alle unità digitali dei volumi del campione scelto, per accertare, anche in questo caso, la corrispondenza tra i dati inseriti nei file XML rispetto alle notizie originali degli oggetti digitali del campione determinato.

Nei casi in cui le attività di digitalizzazione e metadattazione non sono state svolte da società specializzate¹⁶, la disseminazione dei testi in un numero così elevato di

¹⁶ Le società che hanno operato per conto di Biblioteca Italiana nella digitalizzazione di consistenti lotti di opere sono: Space S.p.a. - Cultura e innovazione (Biblioteca nazionale centrale di Firenze, Biblioteca Riccardiana, Biblioteca nazionale di Torino, Biblioteca universitaria di Genova); Gap s.r.l. (Biblioteca Braidense); Shylock e-solutions (Biblioteca nazionale Marciana; Biblioteca Fondazione G. Cini).

biblioteche, appartenenti a tipologie istituzionali e, talvolta, a culture bibliotecarie diverse, è stata fonte di non pochi equivoci, dell'introduzione di variabili procedurali e, spesso, della scorretta o incompleta esecuzione dei lavori commissionati: riproduzione di testi diversi rispetto a quelli richiesti; riprese imprecise, sfocate o rovesciate delle immagini; erronea denominazione dei file. Inoltre, la quasi totalità delle biblioteche, fatta eccezione per quelle in cui hanno operato società specializzate, non ha fornito i file di metadati descrittivi e tecnici richiesti e l'attività di metadattazione ha dovuto quindi essere svolta all'interno della redazione di Biblioteca Italiana.

L'ultima fase del flusso operativo prevedeva la creazione di cartelle il cui nome era in modo univoco rappresentato dal codice ISTC, corrispondente all'unità bibliografica digitalizzata. Tutte le cartelle hanno la medesima struttura e contengono il blocco dei file in formato .tiff, il blocco dei file in formato .jpeg (generati da Biblt) e il file corrispondente di metadati.

L'insieme delle opere e dei loro metadati sono conservati in un server e in diverse unità di backup esterne oltre che nei DVD originali forniti dalle biblioteche.

Il processo di metadattazione

Le fasi operative seguite per la creazione del file di metadati relativo a ciascuna unità bibliografica, hanno riguardato la creazione dei metadati descrittivi, che com'è noto forniscono informazioni relative al contenuto intellettuale della risorsa digitale; la generazione dei metadati tecnici, relativi alle procedure di acquisizione, archiviazione dei documenti e alla loro manutenzione tecnica; e la creazione di quelli strutturali, che descrivono la struttura interna della singola risorsa. Ciascuna di queste fasi, di cui si dà sinteticamente conto nel documento posto in appendice, si è svolta secondo specifici e autonomi flussi di produzione, confluiti poi nella generazione di un unico file in formato xml.

Ogni opera digitalizzata è stata corredata di un documento XML METS¹⁷, che ne veicola i metadati descrittivi, tecnici e amministrativi. Scelta obbligata direi, quella dello standard METS, dal momento che – com'è noto – esso risponde all'esigenza di disporre logicamente tutti gli elementi informativi relativi a ciascuna unità bibliografica e al loro insieme organizzato in una collezione, in virtù del fatto che, dal punto di vista sintattico, METS è un'applicazione XML con cui è possibile descrivere formalmente la struttura logica e fisica di una risorsa informativa digitale, di esprimere i diversi insiemi di metadati a essa associati e di collegarli agli oggetti digitali che la costituiscono.

¹⁷ Metadata Encoding Transmission Standard (METS), <<http://www.loc.gov/standards/mets/mets-home.html>>.

Un ulteriore vantaggio è dato dal fatto che lo schema METS ha una struttura modulare e flessibile, che consente di adottare diversi schemi e vocabolari per esprimere i vari generi di metadati, senza prescriverne nessuno in particolare. Inoltre, la sempre più ampia disponibilità di sistemi e applicativi XML “conformant” consente di utilizzare direttamente i dati in formato METS per lo sviluppo di applicazioni e servizi di accesso alle risorse di un sistema di biblioteca digitale. La sua ormai attestata diffusione in numerosi e importanti progetti di biblioteca digitale in tutto il mondo testimonia sia la qualità intrinseca dello schema sia la consapevolezza che l’adozione di standard per la rappresentazione dei dati costituisce un elemento fondamentale per garantire l’accessibilità universale e la preservazione delle risorse informative in ambiente digitale. In particolare, il profilo METS elaborato per il progetto degli incunaboli italiani in lingua volgare, prevedeva che a ogni volume digitalizzato corrispondesse un record METS, contenente i dati di carattere gestionale; che i metadati descrittivi fossero limitati a una breve indicazione relativa al volume in formato XML MODS¹⁸, così da permetterne l’individuazione in modo non ambiguo; che i metadati tecnici, relativi alle immagini, fossero espressi in formato MIX¹⁹; e, infine, che l’intero set venisse automaticamente convertito nello standard MAG²⁰.

Il corpus Biblt

Il *corpus* Biblt conta 1711 incunaboli volgari, di cui 1613 conservati nelle biblioteche italiane e 98 in quelle straniere.

Il testo più antico registrato nel *corpus* è datato approssimativamente al 1469 (ISTC if00052700): si tratta di un in-quarto, contenente il commento di Federigo da Venezia all’*Apocalisse* di San Giovanni con le glosse di Nicolai de Lyra e pubblicato a Roma o a Napoli da un non meglio identificato “stampatore dell’*Apocalisse*”, per il quale sono stati proposti i nomi di Sixtus Riessinger e di Ulrich Han.

I dati tipografici (luogo di stampa, tipografo, data), quando presenti, sono di diversa tipologia: *certi* (se stampati o ricavabili dall’esemplare), *ricostruiti* (se ricavati da analisi bibliografiche, bibliologiche, filologiche, ecc.), *incerti* (quando, sempre sulla scorta della scheda descrittiva di ISTC, si hanno più ipotesi attributive):

	Dati certi	Dati ricostruiti	Dati incerti	Sine notis
Luogo di stampa	869 pari al 50,9%	778	56	2
Tipografo	801 pari al 46,9%	728	131	45
Data	876 pari al 51,3%	719	95	15

¹⁸ Metadata Object Description Schema (MODS), <<http://www.loc.gov/standards/mods/>>.

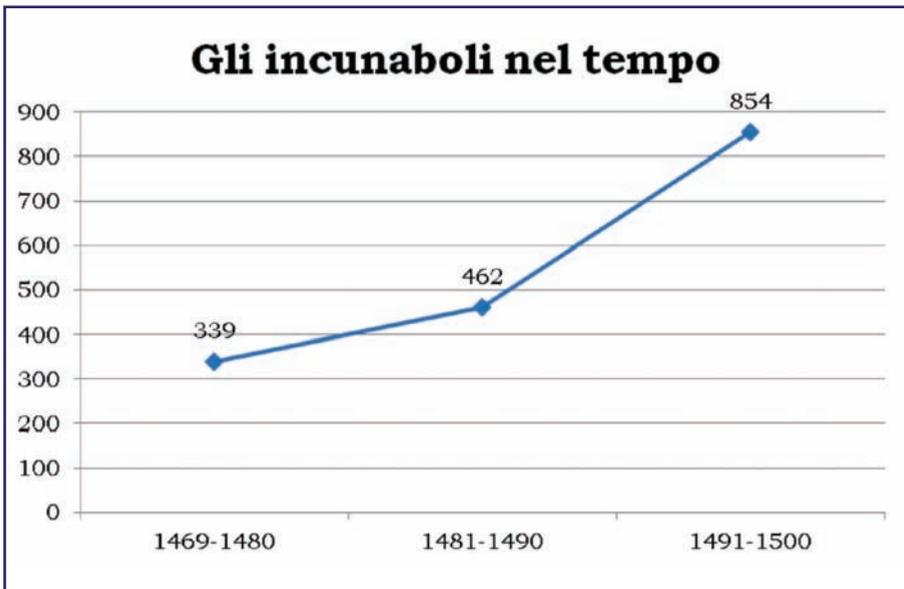
¹⁹ MIX. NISO Metadata for Images in XML Schema, <<http://www.loc.gov/standards/mix/>>.

²⁰ Standard MAG - Versione 2.0.1, <http://www.iccu.sbn.it/opencms/opencms/it/main/standard/metadati/pagina_267.html>.

Faccio presente che nei casi in cui tutti i dati tipografici risultano incerti o non disponibili (edizioni *sine notis*), le relative edizioni sono state escluse dalle proiezioni statistiche di seguito riportate. È noto, inoltre, che tutti i dati qui utilizzati vanno comunque accolti con la opportuna cautela, che va riservata alle indagini di ricerca quantitativa, ancor più, come nel nostro caso, se si tratta di risultanze parziali, relative a un corpus selezionato, sebbene messo spesso a confronto con l'intero corpus della produzione incunabolistica italiana e con quello in lingua volgare, come si evince dalla banca dati ISTC.

Distribuzione cronologica: gli incunaboli nel tempo

Per quanto riguarda la scansione temporale di pubblicazione, la distribuzione delle edizioni per decenni²¹, indica, com'è evidente, quella decisa impennata nell'ultimo decennio, che sancisce l'affermazione dell'arte della stampa che finì per coinvolgere, soprattutto nella sua fase iniziale, numerosi centri urbani anche di piccole dimensioni:

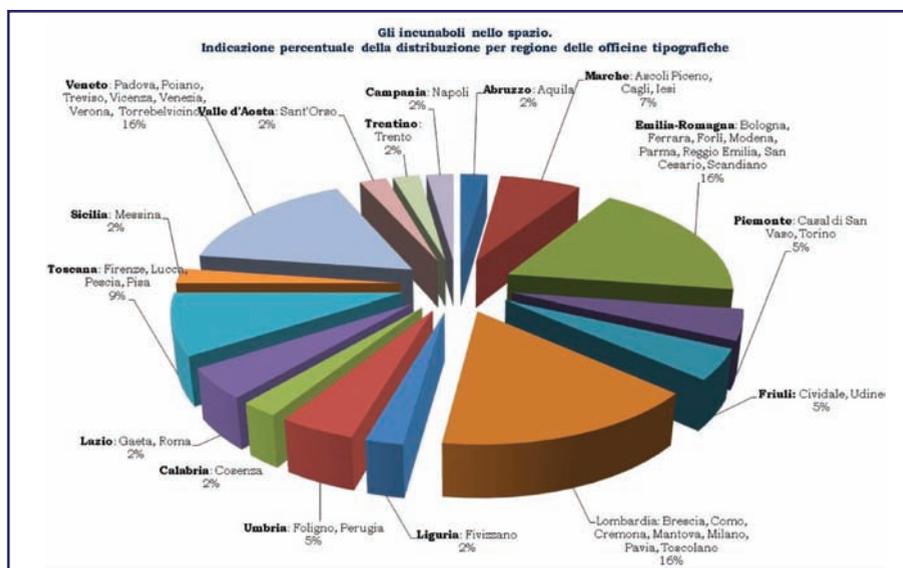


²¹ Prima decade: 1469 (data di pubblicazione del primo incunabolo volgare)-1480; seconda decade: 1481-1490; terza decade: 1491-1500.

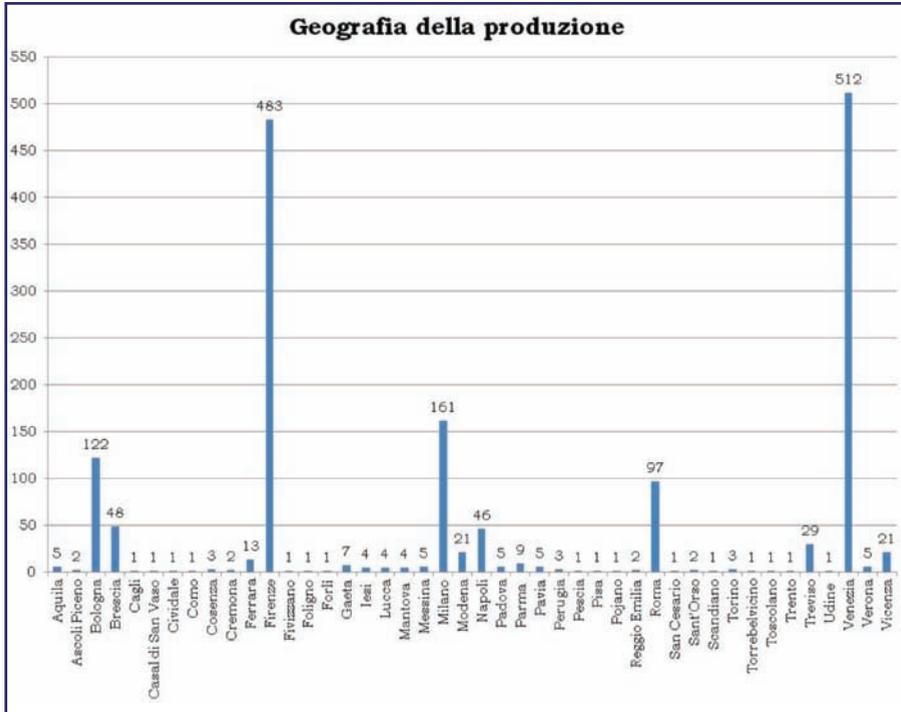
Distribuzione geografica: gli incunaboli nello spazio

Nella tabella e nei grafici seguenti sono visualizzati la distribuzione sul territorio nazionale delle officine tipografiche, che riguarda 44 centri localizzati in 16 regioni, e l'indicazione percentuale della distribuzione per regione delle officine tipografiche, da cui risulta confermato il dato di una significativa concentrazione delle relative attività imprenditoriali soprattutto nelle regioni settentrionali e centrali:

Emilia Romagna	Bologna, Ferrara, Forlì, Modena, Parma, Reggio Emilia, San Cesario, Scandiano
Lombardia	Brescia, Como, Cremona, Mantova, Milano, Pavia, Toscolano
Veneto	Padova, Poiano, Torrelbelvicino, Treviso, Venezia, Verona, Vicenza
Toscana	Firenze, Lucca, Pescia, Pisa
Marche	Ascoli Piceno, Cagli, Iesi
Lazio	Gaeta, Roma
Friuli	Cividale, Udine
Umbria	Foligno, Perugia
Piemonte	Casal di San Vaso, Torino
Abruzzo	Aquila
Calabria	Cosenza
Campania	Napoli
Liguria	Fivizzano
Sicilia	Messina
Valle d'Aosta	Sant'Orso
Trentino	Trento



Il numero di edizioni per centro tipografico mette in evidenza come solo per dieci località questo valore superi le venti unità e che dunque il 90,5% della produzione incunabolistica in lingua volgare (inclusa nel corpus Biblt) è frutto di officine tipografiche attive in dieci città, fra le quali l'unica del Meridione è Napoli. È un rilievo statistico che conferma in proporzione i dati relativi all'intera produzione incunabolistica²²:

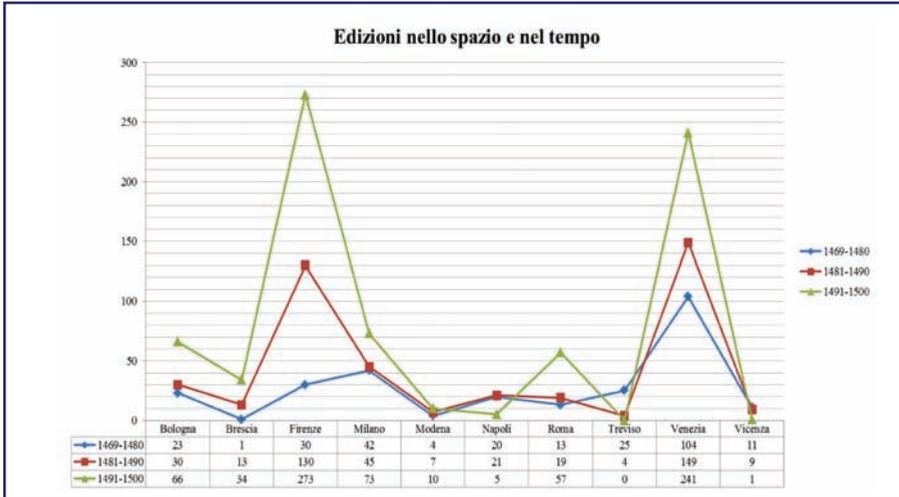


Per le dieci città, che hanno prodotto un numero di incunaboli maggiore di venti (Bologna, Brescia, Firenze, Milano, Modena, Napoli, Roma, Treviso, Venezia, Vicenza), è stata approfondita la cronologia locale. Sulla base dei dati certi e ricostruiti, in nostro possesso, e relativi almeno alla data e al luogo di stampa, se non anche allo stampatore, i primi testi volgari a finire sotto i torchi risultano a Venezia l'edizione del *Canzoniere e Trionfi* di Petrarca per i tipi di Vindelinus de Spira (ip00371000)²³; a Bologna, il *Torneamento fatto in Bologna il 4 ottobre 1470 per ordine di Giovanni Bentivoglio*, un in-quarto stampato da Scipio Malpiglius (if00279300); a Milano, il *Confessionale* di Antoninus Florentinus (ia00842300); a Napoli il *Confessionale* di Pietro Jacopo de Jennaro (ij00218600) edito da Sixtus Riessinger.

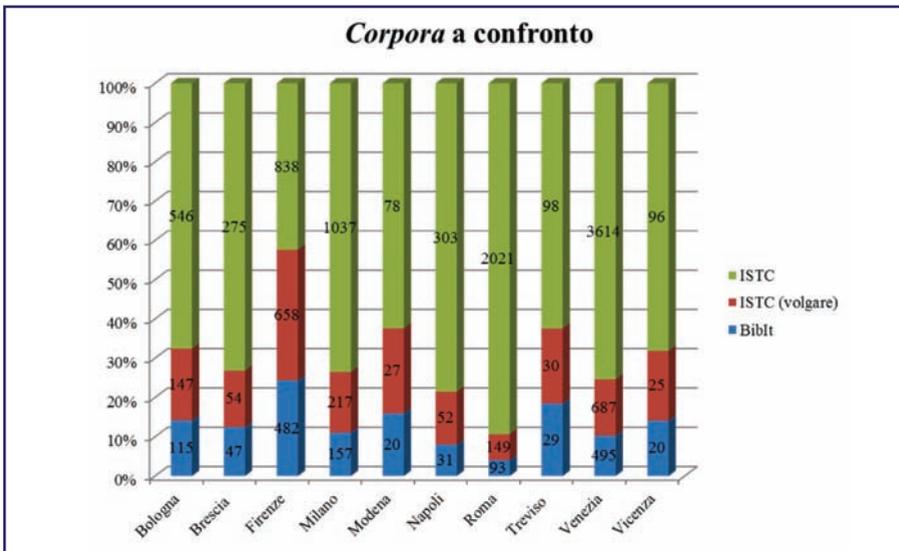
²² Un'analoga proiezione effettuata da Marco Santoro (*Storia del libro italiano*, Milano: Editrice Bibliografica, 2008, p. 68-69), e fondata su un corpus di 7628 edizioni registrate nell'IGI, indica che il 94% della produzione quattrocentesca è concentrata in quindici città.

²³ Si riportano per ciascuna opera i codici identificativi dell'ISTC.

Questa la ripartizione per decenni della produzione di incunaboli volgari nelle dieci città risultate più attive (secondo le proiezioni risultanti dal *corpus* Biblt):



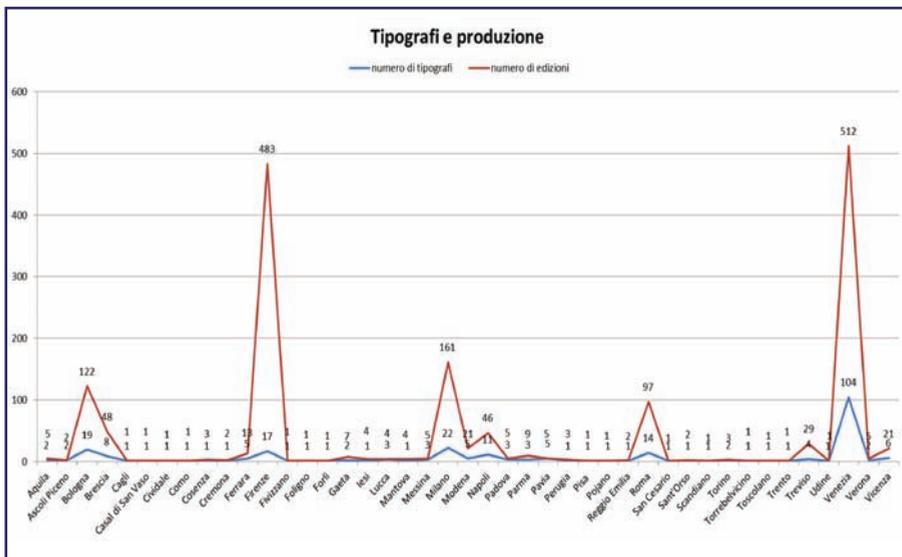
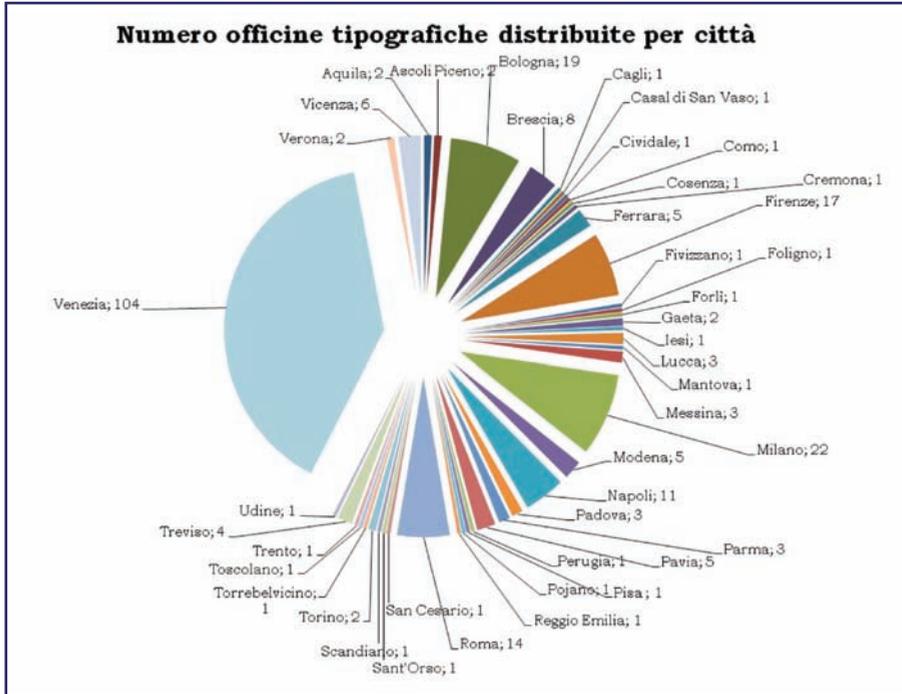
I dati relativi alle dieci città, che hanno prodotto un numero di incunaboli maggiore di venti, sono stati quindi messi a confronto con quelli relativi sia agli incunaboli italiani in lingua volgare presenti in ISTC che all'intero *corpus* ISTC²⁴, e dalla loro comparazione risulta che i rapporti proporzionali restano invariati:



²⁴ Sono state prese in considerazione solo le edizioni che in ISTC presentano per il luogo di stampa un dato o certo o ricostruito.

Tipografi e produzione

Un ulteriore dato, particolarmente significativo per la fortuna degli incunabili volgari, è quello relativo alla distribuzione geografica delle officine tipografiche e al rapporto tra produzione editoriale e numero di stampatori:



Un'ulteriore conferma, dunque, circa la distribuzione geografica delle tipografie più attive anche nella produzione di libri volgari, che finirà per interessare prima le città sedi universitarie e poi l'intero territorio nazionale, sebbene non in modo uniforme e con la prevalenza dell'area padana e delle regioni dell'Italia centrale sul resto della penisola.

Conclusioni

Il progetto, non appena sarà reso di pubblico dominio, potrà avere un impatto significativo non solo nell'ambito della ricerca incunabolistica, ma anche in quella storico-letteraria e filologica, consentendo lo studio sistemico e ravvicinato dei libri tipografici in lingua volgare; permettendo così di indagare meglio, avendo a disposizione una così cospicua e omogenea collezione di documenti: la storia delle tipologie paratestuali che porteranno alle più mature prove del Cinquecento; il rapporto tra tradizione manoscritta e tradizione a stampa, di particolare importanza per saggiare la fortuna di autori, opere e generi e di verificare la tenuta di canoni letterari; lo studio più preciso delle tipologie librerie e la loro relazione con quelle dei generi letterari; una più circostanziata indagine sui processi che porteranno, nel corso del Cinquecento, a un'uniformazione linguistica del volgare, attraverso una fissazione di regole tipografiche, che troverà in Venezia il suo centro di irradiazione.

ALLEGATO

Documentazione sulla creazione dei metadati degli incunaboli

Requisiti tecnici necessari alla realizzazione dei metadati.

- 1) Software BMES (elaborato in Biblt)
- 2) Database eXist
- 3) Oxygen XML Author (software per la validazione o modifica dei file xml)
- 4) Internet Explorer (consigliato) o altri browser per l'accesso a internet.
- 5) Sun xVM VirtualBox (crea un ambiente virtuale su cui installare Ubuntu)
- 6) Sistema operativo: Windows XP e Linux Ubuntu (su virtual Machine)
- 7) Programmi per la generazione dei metadati tecnici:
 1. ImageMagick per la generazione dei jpeg
 2. Jhove per estrazione dei metadati tecnici dalle immagini
 3. Saxon 9 per l'applicazione di procedure XSLT
- 8) Fogli di stile XSLT creati per l'uniformazione allo standard Niso Mix 2.0 e per la rielaborazione del documento finale.

Fasi di realizzazione. Si è considerato opportuno procedere all'elaborazione di un gruppo alla volta di metadati. Il gruppo corrisponde sempre al nome della Biblioteca in cui sono presenti gli incunaboli di interesse.

- 1) Copia dei file master su supporto esterno (da collegare alla macchina su cui è installato BMES) e rinomina secondo lo schema *Scan#####.tif*.
- 2) Creazione dei metadati descrittivi tramite software BMES: partendo dal record relativo del catalogo ISTC si descrivono i campi Autore, Altro autore, Stampatore, Anno di pubblicazione tramite il confronto con il catalogo IGI di cui si assumono i dati descrittivi dell'esemplare. Viene creato un file di nome *mets.bibit.#####.xml*.
- 3) Creazione dei metadati strutturali tramite software BMES: vengono presi in considerazione i seguenti elementi strutturali: Coperta, Fogli di guardia anteriori e posteriori, Testo. Il metadato così ottenuto viene memorizzato nel database eXist.
- 4) Riversamento del gruppo dei file di metadati ottenuti sulle macchine con i sistemi operativi Windows e Linux in condivisione fra loro. Sulla macchina Windows saranno posizionati i file da processare e verranno salvati i file processati. Sulla macchina Linux saranno installati i software necessari alla generazione dei dati.
- 5) Creazione dei file derivati Jpeg tramite software ImageMagick: dai file master TIF verranno generate immagini con compressione all'85%.
- 6) Generazione dei metadati tecnici tramite software Jhove: le immagini master TIF e le immagini derivate JPEG vengono processate per l'estrazione automatica dei metadati tecnici. Per ogni immagine viene creato un file nome.estensione.jhove.xml.
- 7) Uniformazione allo standard Niso Mix 2.0: tramite uno specifico foglio di trasformazione XSLT (per ciascuna tipologia di immagine è stato creato un apposito foglio di stile) e il software Saxon 9 per ogni file nome.estensione.jhove.xml si genera un file nome.estensione.mix.xml.
- 8) Assemblamento di tutti i dati ottenuti tramite procedure XSLT e creazione del metadato finale *mets.bibit.#####.xml*: tutti i file nome.estensione.mix.xml vengono assemblati nel documento xml già compilato con i dati descrittivi e strutturali. Anche in questo caso per ciascuna tipologia di immagine processata è stato creato un apposito file di stile XSLT. Il risultato finale sarà un file nome *mets.bibit.#####.xml* completo di metadati descrittivi, tecnici e strutturali.
- 9) Validazione del file in ambiente Oxygen XML Author.

Per tutti i siti web, l'ultima consultazione è avvenuta nel mese di dicembre 2012.