

L'edizione elettronica di un manoscritto: il modello XML

Flavia Bruni

Storicamente, 1 (2005).

ISSN: 1825-411X. Art. no. 76. DOI: [10.1473/stor311](https://doi.org/10.1473/stor311)

Una "base di dati" inquisitoriale.

Allo scadere del XVI secolo, trascorso ormai un secolo e mezzo dall'invenzione della stampa e poco meno di un secolo dall'altra rivoluzione ad essa strettamente congiunta, ovvero la Riforma protestante, il libro stampato aveva raggiunto una diffusione tanto ampia da non essere paragonabile a nessun mezzo di diffusione culturale dei secoli precedenti, sia a livello quantitativo, e dunque sociale, sia per così dire qualitativo, ovvero per quanto riguardava la vastità degli argomenti trattati. Questo non poteva non destare l'allarme all'interno della Chiesa di Roma, proprio nel momento in cui essa, sulla scorta delle deliberazioni del Concilio di Trento, si accingeva alla riorganizzazione delle strutture e al disciplinamento culturale per affrontare le insidie provenienti dall'interno della cristianità. Era risultato evidente fin dagli esordi della Riforma come una circolazione libraria priva di controllo potesse avere esiti disastrosi per il tradizionale monopolio culturale della Chiesa sull'Occidente, e di conseguenza anche per la costruzione politica che su di esso poneva le sue fondamenta. Prese così lentamente forma quella che sembrava essere l'unica possibile soluzione: la creazione di un ferreo apparato di controllo, con la funzione di vagliare, proibire, ed eventualmente anche epurare l'intera produzione editoriale europea del secolo trascorso, e, non di meno, di quello ormai prossimo.

Concretamente, dopo alcuni conflitti di competenza, tale compito fu assunto al vertice dalla neonata Congregazione dell'Indice, e lo strumento da essa approntato per l'adempimento della propria funzione fu l'indice dei libri proibiti promulgato nell'anno 1596[1]."

Nonostante gli sforzi compiuti per dare massima diffusione alle proibizioni specialmente dopo la promulgazione dell'indice del '96, l'applicazione concreta risultò, com'è ovvio, operazione tutt'altro che facile, anche in quanto era inevitabile delegare parte del compito di controllo ad una rete inquisitoria a livello locale; tale decentramento spiega perciò il sussistere di forti dubbi sulla scrupolosa applicazione dell'Indice, fosse pure più a causa di una non perfetta conoscenza di esso che per deliberata volontà di trasgressione. Dall'intento di controllare l'osservanza delle prescrizioni almeno all'interno degli ordini regolari si originò dunque un'"inchiesta", promulgata dalla Congregazione dell'Indice, che richiese di inviare a Roma le liste complete dei libri posseduti sia dai singoli religiosi, sia dalle biblioteche comuni degli istituti di tutti gli ordini religiosi[2]. Ne risultò una enorme mole di materiale, costituito da centinaia di liste di libri pervenute a Roma in varia forma: liste singole inviate come fogli sciolti, oppure fascicoli numerati e rilegati che rivelano una operazione coordinata all'interno di alcuni ordini o di alcune province. Questa ricchissima e preziosissima documentazione, dopo il periodo napoleonico, è stata raccolta nella Biblioteca Apostolica Vaticana in sessantuno codici estremamente eterogenei, rilegati per ordine di appartenenza e per lo più suddivisi al loro interno per provincia, convento ed eventualmente singolo possessore[3]. La rilevanza di tale fonte sia per la conoscenza delle biblioteche monastiche, sia per l'analisi della produzione e circolazione libraria in Italia tra XVI e XVII secolo è tanto evidente da non richiedere ulteriori sottolineature[4].

Dell'intero *corpus* di codici si sta occupando un gruppo di ricerca, di cui faccio parte, coordinato dal prof. R. Rusconi e denominato RIC (Ricerca

sull'Inchiesta della Congregazione dell'Indice). Il progetto prevede la trascrizione degli elenchi di libri, l'identificazione delle edizioni, effettuata in una prima fase su banche dati *on line* e successivamente su repertori cartacei, e infine l'inserimento in un *database* a cui verrà consentito l'accesso non appena terminato l'inserimento delle informazioni[5].

All'interno di un patrimonio già di per sé di enorme rilevanza, occupandomi degli elenchi pervenuti a Roma dalle province dell'Ordine dei Servi di Maria ho potuto analizzare un eccezionale documento, conservato nei ff. 28-47 del manoscritto Vat. Lat. 11286, che riproduce il verbale di una inquisizione compiuta nel convento dei Servi di Maria di Lucca nel novembre dell'anno 1600, nei confronti di un frate di nome Lorenzo[6]. Nella cella del frate, assente al momento della visita inquisitoriale, erano stati rinvenuti circa settecento libri, di cui una buona parte identificabili come proibiti; proprio la descrizione dei singoli libri riportata nella lista testimonia un intervento censorio, probabilmente operato per mano dello stesso Lorenzo, tramite cancellature, raschiature o applicazione di "piastrelli", proprio in corrispondenza di nomi di autori o editori condannati dall'*Index*[7]. Oltre all'elenco dei libri, redatto secondo gli usuali canoni imposti dalla Congregazione, nelle venti carte si trovano altresì parti narrative che descrivono lo svolgimento della *inquisitio* nella cella e nel convento, riportando anche gli interrogatori a cui furono sottoposti i confratelli di Lorenzo.

Di fronte a un documento di tale interesse si è immediatamente delineata l'esigenza di pubblicazione, congiunta con l'idea di effettuare ricerche automatiche mirate all'interno della lista dei libri. Da queste premesse è nato l'esperimento, che sto sviluppando in maniera autonoma e parallela rispetto al lavoro all'interno del sopracitato gruppo di ricerca e che costituisce la mia tesi di dottorato[8], volto ad elaborare una digitalizzazione delle venti carte in questione, avente come fine più concreto e prossimo quello di un'edizione,

sia elettronica sia a stampa, che non trascuri la descrizione accurata dei molteplici aspetti del documento; ma forse di ancor maggiore rilevanza è l'obiettivo concomitante di realizzare un *database* dei circa settecento titoli della biblioteca del frate elencati nel testo, rendendo possibili operazioni e dunque ricerche in maniera automatica.

La modellizzazione. Le entità.

Affinchè la ricerca possa dare dei risultati, la digitalizzazione vera e propria deve essere dunque preceduta da una fase di ricognizione e sistemazione delle caratteristiche del documento: bisogna cioè organizzare un [modello](#) il più possibile corrispondente al reale aspetto della fonte, basandosi sull'attenta osservazione e selezione dei fenomeni ritenuti significativi e delle relazioni tra di essi intercorrenti. Si tratta del passaggio fondante che chiamiamo di *modellizzazione* del documento.

Nel caso dei sopracitati fogli manoscritti, una modellizzazione realizzata tramite marcatura con il metalinguaggio XML può essere la soluzione ottimale, in quanto consente di rendere la complessità dei fenomeni riscontrati in maniera relativamente semplice; all'origine di tale idoneità è la possibilità di stratificare la descrizione su diversi livelli sovrapponendo più DTD[9], lasciando inoltre aperta l'eventualità di aumentare indefinitamente i *tag* da apporre per ottenere una modellizzazione sempre più complessa e ricca di dati.

Adottando lo standard XML è sufficiente rispettare alcune semplici regole di sintassi nell'applicazione dei *tag*, dopo di che questi ultimi, seppure definiti dall'utente per le esigenze di descrizione dello specifico documento, e non desunti dunque da alcun linguaggio prestabilito, potranno essere facilmente convertiti a seconda delle diverse necessità: sarà facile procedere, ad esempio, all'impaginazione su schermo per l'edizione elettronica, trasformando i *tag* personalizzati in quelli standard dell'HTML; o, allo stesso

modo, alla preparazione di un'edizione di tipo "tradizionale", ovvero stampata su carta; o, ancora, alla costruzione di un *database* ordinato delle informazioni contenute nel testo. La potenzialità maggiore di questa procedura deriva infatti senza dubbio dalla possibilità di applicare al documento codificato dei filtri che operino in base ai *tag*. Questi possono essere considerati come delimitatori di campi ricorrenti, di modo che la semplice inserzione di essi trasformi, di fatto, il documento stesso direttamente in un *database*. Mentre in un *database* "tradizionale" avremmo delle schede suddivise in campi, da riempire di volta in volta con i dati, al contrario operando tramite marcatura sono i *tag* ad essere inseriti all'interno del documento e a delimitare i campi, evitando così di "smembrare" la fonte nei diversi campi[10].

Da un punto di vista descrittivo-formale, invece, un aspetto di grande interesse è costituito dalla possibilità di non limitare il set di caratteri a quelli contemplati dall'ASCII 128, senza tuttavia ricorrere a caratteri non standard (ASCII 256), o peggio ad altre fantasiose forme per la descrizione di fenomeni non ritenuti assimilabili a processi di scrittura odierni. È possibile infatti definire una tabella di entità di riferimento, che può comprendere, oltre ovviamente ai caratteri standard dell'ASCII 128, un numero teoricamente infinito di altri fenomeni grafici in esso non rappresentati. Diventa in questo modo possibile riconoscere come entità a sé stanti, da noi "codificate", una gran quantità di fenomeni ascrivibili alla scrittura del XVI secolo, come abbreviazioni, "*tituli*" eccetera, tutti elementi di cui si perderebbe completamente traccia limitandosi a una piana trascrizione diplomatica. Diventa in questo modo plausibile l'obiettivo di realizzare una codifica del documento a partire dalla sua entità fisica, per così dire, e non dalla mera trascrizione del testo in esso contenuto: quest'ultimo infatti, per quanto per determinati fini possa essere riconosciuto come l'elemento di maggiore interesse, non è l'unica fonte di informazione contenuta nel documento stesso. La trascrizione diplomatico-interpretativa, per definizione,

"interpreta", comportando una perdita di informazione, che, se può essere per certi versi irrilevante, potrebbe invece essere completamente fatale per altri, riguardanti, appunto, gli aspetti formali del documento e non il suo contenuto verbale[11].

L'adozione di un modello XML consente dunque una esauriente descrizione del documento, sempre suscettibile di ampliamenti successivi, per di più incorporandola, per così dire, all'interno del documento stesso; al tempo stesso, facilita l'applicazione di procedure automatiche su campi di volta in volta selezionati secondo le esigenze della ricerca. Come detto sopra, il verbale della inquisizione svolta nel convento di Lucca è inserito in uno dei sessantuno codici di un *corpus*, che custodisce nel suo insieme documentazione grossomodo simile, ovvero liste di libri posseduti da regolari negli anni a cavallo tra fine XVI e inizio XVII secolo. È dunque possibile in un primo momento svolgere la ricerca in maniera strettamente focalizzata sul campione prescelto, tentando di compendiarne quanti più aspetti possibile nella maniera ritenuta di volta in volta più adeguata; in una fase successiva, tuttavia, sarebbe plausibile o meglio auspicabile applicare il modello di codifica elaborato, progressivamente ampliandolo e perfezionandolo, ad una documentazione sempre più vasta, arrivando alla costruzione di un *database* unico per l'intero *corpus* di codici: anzi, proprio in questa prospettiva sono da vedere i frutti più interessanti della ricerca.

Note

[1] Vd. J. M. De Bujanda, *Il primo indice romano, 1559*, in: M. Santoro (ed.), *La stampa in Italia nel Cinquecento*, Atti del convegno internazionale, Roma 17-21 ottobre 1989, Roma, 1992; Id. (ed.), *Index des livres interdits*, 10 voll., 1985-1996, in: part. vol. IX: *Index de Rome 1590, 1593, 1596, avec une étude des Index de Parme 1580 et München 1582*, J. M. De Bujanda, U. Rozzo, P. G. Bietenholz, P. F. Grendler (eds.), 1994.

[2] In un primo momento, la Congregazione aveva richiesto di inviare a Roma solamente le liste dei libri proibiti rinvenuti nelle biblioteche comuni o private di tutti i regolari; ma, in seguito alle difficoltà riscontrate nell'adempimento della richiesta, si ritenne più semplice e soprattutto più efficace, sebbene più laborioso, richiedere le liste complete di tutti i libri posseduti. Vd. M. Dykmans, *Les bibliothèques des religieux d'Italie en l'an 1600*, in: «Archivum Historiae Pontificiae», 24 (1986), 207-255.

[3] Vd. M.-M. Lebreton - L. Fiorani, *Inventari di biblioteche religiose italiane alla fine del Cinquecento (Codices Vaticani Latini. Codices 11266-11326)*, Città del Vaticano, Biblioteca Apostolica Vaticana, 1985; R. De Maio, *I modelli culturali della Controriforma. Le biblioteche dei conventi italiani alla fine del Cinquecento*, in: Id., *Riforme e miti nella Chiesa del Cinquecento*, Napoli, Guida, 1973, 365-381; M. Dykmans, *Les bibliothèques des religieux d'Italie en l'an 1600*, in: «Archivum Historiae Pontificiae», 24 (1986), 385-404.

[4] Per una bibliografia completa e aggiornata sull'argomento vd. R. Rusconi, *I libri dei religiosi nell'Italia di fine '500*, in: «Accademie e biblioteche d'Italia», 78 (2004), 17-38.

[5] Vd. sopra, *art. cit.* nota precedente.

[6] Il documento era già stato segnalato da M. Rosa, *"Dottore o seduttore egregio deggio appellarte": note erasmiane*, in: «Rivista di storia e letteratura religiosa», 26 (1990), pp. 30-33, e A. Barzani, *La memoria di Savonarola. Testi savonaroliani nelle biblioteche dei religiosi alla fine del Cinquecento*

, in: G. Fragnito – M. Miegge (eds.), *Girolamo Savonarola da Ferrara all'Europa*, Firenze 2001, 281-282.

[7] Lorenzo era infatti revisore vescovile; per dettagli, rimando a un mio studio di prossima pubblicazione.

[8] Storia e Informatica, XVII ciclo, Università degli Studi di Bologna.

[9] *Document Type Definition*; è la legenda tramite cui si costruisce un modello in XML, definendo appunto, secondo una specifica sintassi, gli elementi con rispettivi attributi e reciproche relazioni. Si tratta in altre parole della dichiarazione esplicita della struttura del modello. Vd. E. van Herwijnen, *Practical SGML*, Dordrecht-Boston-London, 1990, 27 e segg. Incidentalmente, notiamo che l'uso di più DTD combinate per uno stesso documento permette di sovrapporre più livelli di marcatura anche, per così dire, trasversalmente su diversi campi, evitando il cosiddetto annidamento dei *tag*.

[10] Per una sintetica descrizione delle potenzialità dell' SGML (precursore dell'XML) vd. di nuovo T. Orlandi, *Integrazione dei sistemi di lettura intelligente e banche dati nel progetto Medioevo-Europa*, in: C. Leonardi, M. Morelli, F. Santi (eds.), *Macchine per leggere*, Spoleto, 1994, 41-52, in particolare pp. 42-47 e nota 6 per la bibliografia sull'argomento.

[11] Cfr. T. Orlandi, *Presupposti metodologici dei reciproci contributi tra informatica e bibliologia*, in: E. Esposito (ed.), *Bibliologia e informatica*, Ravenna, 1994, 19-33.

Link

Costruire un modello per il testo

Dovendo affrontare qualunque tentativo di digitalizzazione è opportuno riflettere preliminarmente su cosa si vuole ottenere e dunque su cosa esattamente si vuole codificare. Affinchè un testo sia fruibile per una ricerca automatica è necessario rappresentare tutti i diversi livelli del testo, ovvero tutte le nozioni che sono comprese in quello che abitualmente chiamiamo, con termine in realtà non univoco, "testo".

[Una rappresentazione completamente esaustiva è in realtà impossibile in quanto tende ad infinito; si cercherà però, con buona approssimazione, di compendiare quanti più elementi sia possibile, in particolare tenendo conto delle specifiche esigenze della ricerca che ci si propone di effettuare. Vd. D. Buzzetti, *Rappresentazione digitale e modello del testo*, in: *Il ruolo del modello nella scienza e nel sapere*, Atti del Convegno (Roma, 27-28 ottobre 1998), Roma, Accademia Nazionale dei Lincei, 1999, (Contributi del Centro Linceo Interdisciplinare "Beniamino Segre").]

Se nel sistema di ricerca tradizionale è la competenza dello specialista a riassumere tale pluralità di significati, senza che sia strettamente necessario distinguerli in maniera analitica, dovendo operare tramite uno strumento, il computer, che può agire solo in base ad istruzioni precise e ad una competenza esterna, lo studioso deve in qualche modo "trasferire" nel testo le indicazioni che gli derivano dalla sua competenza: bisogna in un certo qual senso che il testo venga caricato di tutte le informazioni necessarie al fine della ricerca. Per restituire al testo lo spessore che gli appartiene in quanto veicolo di informazione su molteplici livelli, la soluzione migliore pare quella suggerita dalla teoria dei sistemi, ovvero di realizzare un "modello" elettronico, per quanto possibile corrispondente all'oggetto (o documento) che si intende digitalizzare.

[Sulla modellizzazione vd. T. Orlandi, *Linguistica, sistemi, e modelli*, in *Il ruolo del modello nella scienza e nel sapere*, Atti del Convegno (Roma, 27-28 ottobre 1998), Roma, Accademia Nazionale dei Lincei, 1999, (Contributi del Centro Linceo Interdisciplinare "Beniamino Segre"), Roma 1999, 73-90; Id., *Ripartiamo dai sistemi*, in *I nuovi orizzonti della filologia. Ecdotica, critica testuale, editoria scientifica e mezzi informatici elettronici*, Conv. Int. 27-29 maggio 1998, Roma 1999 (Atti dei Convegni lincei, 151), 87-101; Id., *Integrazione dei sistemi di lettura intelligente e banche dati nel progetto Medioevo-Europa*, in: C. Leonardi,

M. Morelli, F. Santi (eds.), *Macchine per leggere*, Spoleto, 1994, 41-52, in particolare pp. 42-47 e nota 6 per la bibliografia sull'argomento.]

Possiamo definire "modello" una rappresentazione ordinata che miri a riprodurre la complessità dell'oggetto reale; un esempio tratto dalle arti plastiche è quello che chiamiamo appunto "modellino", che altro non è se non una rappresentazione ordinata di qualcosa che si intende riprodurre in scala. Lo studio del documento deve essere criterio per la scelta del modo per rappresentare l'insieme dei fenomeni e relazioni che lo caratterizzano, in quanto solo da una adeguata posizione dei problemi, in una visione il più possibile unitaria, potranno scaturire adeguate risposte.

[Indietro](#)

Che cos'è l'XML

Per XML si intende qui esclusivamente la conformazione a regole sintattiche secondo lo standard diffuso appunto con il nome di XML; nel caso specifico, per la costruzione del modello si è cercato di elaborare una DTD che si conformasse alle specifiche esigenze della ricerca, discostandosi dunque dalle direttive del molto diffuso modello TEI o TEI light. Da sottolineare è anche che l'adozione della marcatura XML non rappresenta di per sé e tout court la soluzione di qualsiasi problema legato alla codifica o alla realizzazione del modello: si tratta semplicemente di una serie di formalismi descrittivi che, se applicati in maniera corretta e appropriata, possono dare il risultato voluto, purchè venga realizzato preventivamente un adeguato lavoro di modellizzazione basato sull'osservazione dell'oggetto e sugli scopi della ricerca. Potremmo fare un paragone con un contenitore vuoto, che per assumere significato e soprattutto per essere fruibile deve essere riempito con i giusti contenuti. L'XML infatti non è un pacchetto applicativo per l'elaborazione automatica dei dati, e allo stato attuale esistono validatori per controllare che il sistema sia stato ben costruito (well-formed), ma non ancora parser standardizzati per l'elaborazione dei dati in tal modo codificati: i filtri per il trattamento dei dati devono dunque essere realizzati ad hoc, nei migliori casi mediante l'interazione con il linguaggio PHP.

[Indietro](#)