

Manuscripta mediaevalia

Un progetto di digitalizzazione che ha radici profonde e integra risorse di origine diversa

Il sito *Manuscripta Mediaevalia* è realizzato in collaborazione dalla Staatsbibliothek zu Berlin, dal Bildarchiv Foto Marburg e dalla Bayerische Staatsbibliothek di Monaco, ed è sovvenzionato dalla Deutsche Forschungsgemeinschaft (DFG), l'ente che per conto dello stato sovrintende alla ricerca in Germania.¹ Il nuovo progetto attinge a una ricca tradizione catalografica assommando informazioni molto varie e strutturalmente diverse tra loro, in primo luogo quelle provenienti dai cataloghi di manoscritti prodotti in area tedesca. Questi strumenti, che vanno dall'opera di Valentin Rose ai più moderni cataloghi promossi, a partire dal 1960, dalla stessa DFG, sono ben noti per il loro carattere esaustivo, e occupano una larga parte degli scaffali dedicati ai repertori sui manoscritti all'interno delle biblioteche specializzate in quest'area.² Ma il progetto *Manuscripta Mediaevalia* sfrutta ampiamente anche altre fonti, come per esempio la storica attività della Berlin-Brandenburgische Akademie der Wissenschaften, che possiede circa 20.000 descrizioni manoscritte realizzate tra il 1904 e il 1945.³ Tale documentazione costituisce il risultato dell'attività di numerosi studiosi, tutti elencati – istituzioni e collaboratori – in una delle pagine web dedicate all'archivio.⁴ Insomma il nuovo sito, disponibile al pubblico all'indirizzo <[\[diaevalia.de\]\(http://www.manuscripta-mediaevalia.de\)>, è un edificio che ha radici profonde, si erge con forme monumentali e, forse proprio per questi presupposti e per la sua ambiziosa pretesa di integrare risorse di origine diversa, offre all'utente dei percorsi interni complessi e di non immediata comprensione, che cercherò qui di districare almeno in parte.](http://www.manuscripta-me-</p>
</div>
<div data-bbox=)

Osserviamo innanzitutto che 175 cataloghi cartacei prodotti dalla DFG, affascinanti monumenti di un'attività scientifica ben sostenuta finanziariamente, sono oggi interamente riprodotti in forma digitale in *Manuscripta mediaevalia*, insieme ad altri cataloghi di patrocinio diverso e ad alcuni di raccolte estere. Dalla pagina principale del sito, attraverso il link *Handschriftenkataloge online*, si accede all'elenco dei singoli volumi riprodotti, distribuiti per città in un utilissimo *catalogus catalogorum* che può essere letto consequenzialmente. A ognuna di queste descrizioni bibliografiche è collegato il relativo sommario che comprende i rinvii automatici alle singole pagine dell'opera. Con perfetta fedeltà all'originale cartaceo sono considerati an-

che i fogli bianchi; inoltre i link alle pagine riproducono la numerazione (o le numerazioni successive) del volume e sono suddivisi in modo da rispecchiarne la struttura. Per esempio, nel caso del catalogo dei codici latini della raccolta Hamilton di Berlino, di Helmut Boese, il sommario distingue le pagine iniziali del volume, quelle che contengono la prefazione, le concordanze delle segnature e la bibliografia abbreviata, il catalogo vero e proprio (con l'indicazione dei manoscritti considerati in ognuna) e gli indici finali.

Ogni catalogo dunque può essere consultato e sfogliato avanti o indietro come se si trattasse dell'originale a stampa: una vera biblioteca digitale di una parte cospicua dei nostri repertori a stampa, che può far risparmiare allo studioso la visita ai depositi reali, spesso incompleti e talvolta ingiustamente sottratti alla consultazione a scaffale aperto. Robert Giel, responsabile del progetto per la Staatsbibliothek zu Berlin, giustifica così questa scelta di creare un archivio di immagini: "All'inizio del progetto *Manuscripta mediaevalia* abbiamo provato a recuperare i testi dei cataloghi con un sistema di Optical Character Recognition, ma, a causa della complessità dei caratteri e dell'impaginazione variabile, il software produceva l'1% di errori, cioè circa 30 errori per ogni foglio.

Anche la copiatura manuale delle descrizioni, trattandosi di 50.000 pagine a stampa, non si prospettava come un'alternativa possibile. L'unica soluzione è stata acquisire con lo scanner delle immagini in formato TIFF e trasferirle in formato JPEG. Ora queste immagini possono essere sfogliate e ad esse si può accedere attraverso un indice strutturato".

Tale indicizzazione delle notizie costituisce il secondo fondamentale strumento del sito tedesco. Una prima possibilità di ricerca è già presente alla pagina *Handschriftenkataloge online*. Da qui è possibile interrogare sia le descrizioni bibliografiche dei cataloghi digitalizzati sia le loro singole pagine a partire dall'autore dei singoli volumi (attenzione: non dagli autori delle opere catalogate!) o da una parola compresa nel titolo, dal luogo di conservazione e dalla segnatura di un manoscritto. Tra le risposte, oltre all'archivio digitale costituito dalle descrizioni bibliografiche e dalle riproduzioni dei cataloghi cartacei, compaiono in questo caso anche le schede provenienti dall'archivio della Berlin-Brandenburgische Akademie der Wissenschaften. La logica con cui agisce questa prima maschera di ricerca produce, per esempio digitando il termine "Erfurt" nella prima casella (Autor-/Titelstichwort), risposte di tre tipi:



Cpg 16: Altes Testament Cpg 17: Altes Testament Cpg 18: Altes Testament

Cpg 19: Altes Testament, Bücher Mose, Josua, Richter

Manoscritti digitalizzati della biblioteca dell'Università di Heidelberg, tratti dal sito *Manuscripta mediaevalia*

1) la descrizione bibliografica di un catalogo realizzato da Hans Thurn che descrive codici benedettini conservati alla Universitätsbibliothek Würzburg, tra i quali sono compresi anche codici provenienti da Erfurt;

2) le indicazioni "Erfurt, Domarchiv. Handschriftenarchiv der Berlin-Brandenburgischen Akademie der Wissenschaften - Kasten 92" e "Erfurt, Stadtarchiv. Handschriftenarchiv der Berlin-Brandenburgischen Akademie der Wissenschaften - Kasten 92";

3) una serie di link a singole pagine dello stesso catalogo di Thurn.

Le due indicazioni relative al Domarchiv e allo Stadtarchiv non costituiscono link, e rimandano semplicemente a una collocazione ("Kasten 92") dell'archivio cartaceo della Accademia berlinese, come si può dedurre dal sito della stessa.⁵ Di questo recupero *in fieri* Robert Giel spiega: "Più di 2.000 descrizioni dell'Handschriftenarchiv der Berlin-Brandenburgischen Akademie der Wissenschaften sono state indicizzate presso la Staatsbibliothek zu Berlin tra il 1995 e il 2000, e quindi sono state integrate in *Manuscripta mediaevalia*. Molte di queste descrizioni hanno un grande valore, perché riguardano manoscritti di piccole istituzioni che spesso sono stati distrutti durante l'ultima guerra mondiale. Per il momento non esistono immagini di queste schede, ma intendiamo approntarne le riproduzioni al più presto, in collaborazione con l'Accademia stessa".

Ma tornando alla maschera di ricerca, se il nome "Erfurt" viene digitato all'interno della casella "Bibliotheksort", compaiono solo due risposte, ossia proprio le due notizie che rimandano alla collocazione "Kasten 92" della Berlin-Brandenburgische Akademie der Wissenschaften, mentre vengono

omessi i rimandi al catalogo di Thurn che considera codici provenienti da Erfurt, ma attualmente conservati altrove.

D'altra parte il link *Handschriftendatenbank* presente alla homepage permette di accedere ad altre due maschere di ricerca più complesse, le quali interrogano rispettivamente un archivio di incipit, di struttura e di utilizzo intuitivi, e un

catalogo vero e proprio, che costituisce il nucleo dell'indicizzazione realizzata dal progetto. Qui la ricerca agisce su tutti gli accessi e le notizie esistenti nel database, il quale, per cominciare, ha assorbito le notizie del precedente *Gesamtindex mittelalterlicher Handschriftenkataloge*. Quest'ultimo nacque negli anni Ottanta in forma di microfiche e a partire dal 1992 costituì il contenuto di una base di dati elettronica, quindi venne pubblicato in Internet e oggi è ancora disponibile all'interno del sito del Deutsche Bibliotheksinstitut (DBI), ma non più aggiornato dall'autunno 1999.⁶

"Tramite un procedimento automatizzato di import-export - spiega ancora Robert Giel - circa 300.000 notizie e 200.000 incipit (relativi a 30.000 manoscritti e tratti dagli stessi cataloghi a stampa promossi dalla DFG che sono recuperati in *Manuscripta mediaevalia*) sono stati trasferiti nel nuovo archivio. I dati principali del *Gesamtindex* avevano una triplice struttura gerarchica, del tipo:

Augustinus (LE=lemma)
- *De civitate Dei* (L1)
- - *Excerptum* (L2).

Noi abbiamo rielaborato (e stiamo ancora rielaborando) questi lemmi in due modi: prima definiamo le voci come nomi, titoli, luoghi, informazioni codicologiche; poi correggiamo l'ortografia e creiamo un



collegamento con un authority file, nella maggior parte dei casi con il *Personennamendatei* della Deutsche Bibliothek."

Effettivamente, se si torna al "vecchio" *Gesamtindex*, la semplice strutturazione dei lemmi e la mancanza di un *authority control* saltano subito all'occhio. Ciascun lemma (suddiviso in più parti come si è visto sopra) può essere collegato a una biblioteca e al catalogo relativo (etichetta BI), alla segnatura (etichetta CO) e al foglio (etichetta FO) del codice e infine alla designazione completa del catalogo cui il lemma rimanda (etichetta QU). Oppure altri lemmi possono rappresentare il rimando da una voce a un'altra.

Nel nuovo database di *Manuscripta mediaevalia* a questi stessi dati trasferiti dal *Gesamtindex* sono state affiancate nuove informazioni provenienti da altre fonti: gli indici di nuovi cataloghi sovvenzionati dalla DFG, in parte inseriti manualmente e in parte recuperati dai file utilizzati dagli editori, altri cataloghi scientifici e dati provenienti da nuovi programmi di inventariazione.

"Il software utilizzato per l'inserimento dei dati è HiDA, elaborato inizialmente per il Bildarchiv Foto Marburg e adattato alle esigenze specifiche del nostro progetto.⁷ L'interrogazione in Internet invece si basa su un'esportazione in XML che

viene riversata in un database SQL" dice Robert Giel.

La strutturazione complessiva dei dati così ottenuta permette un'interrogazione web ben più complessa di quella del primitivo *Gesamtindex*. La lunga pagina denominata *Standard-suche* presenta infatti, in una forma graficamente simile a uno schedario, la possibilità di combinare con operatori booleani diverse tipologie di richieste: luogo, istituzione, segnatura antica e attuale in un primo riquadro; nomi di persona e di ente, in forme diverse, nel secondo; titoli, incipit e indicazioni bibliografiche relative al testo nel terzo; un certo numero di informazioni codicologiche e storiche nel quarto; informazioni relative all'iconografia nel quinto; l'indice generale nell'ultimo. A fianco di ogni casella dedicata all'interrogazione è sempre aggiunto un tasto per aprire gli indici relativi delle parole o delle informazioni presenti nel database. Quindi dalla lista dei record ottenuti dalla ricerca è possibile vedere tutti gli accessi legati alla segnatura di un codice e risalire, ancora una volta, alla fonte digitalizzata, quando questa esiste, da cui è stata tratta la notizia.

Per esempio richiedendo il nome "Terentius Afer, Publius" attraverso la casella "Person Normform" della maschera dedicata alla ricerca di persone ➤

ed enti (*Personen und Körperschaften*), si ottengono al presente dodici risposte che costituiscono accessi a varie descrizioni di codici. La prima di queste risposte rimanda a un manoscritto di Hildesheim, di cui viene data la segnatura (Hs 618), il numero di foglio cui è legato il nome del commediografo latino, il ruolo di questa persona nel contesto (che in questo caso è il ruolo di autore), il titolo della sua opera e la fonte da cui è stata derivata la notizia, ossia un catalogo a stampa a più mani dedicato alla Dombibliothek di Hildesheim, realizzato nel 1991. Quindi la procedura informatica permette sia di aprire la pagina digitalizzata di tale catalogo in cui si trova la descrizione di questo manoscritto, sia di vedere tutti gli accessi legati a questa segnatura all'interno del database.

Scegliendo questa seconda possibilità si ritrovano elencati gli accessi relativi alla descrizione fisica del manoscritto e alla sua storia (possessori, data, tipo di scrittura...) e gli accessi rappresentati dai testi tramandati dal codice (autori, titoli, incipit).

In altri casi il tipo di risposte risultanti dalla *Standardsuche* è ancora diverso. Digitando ancora una volta il nome "Erfurt" nella casella "Bibliotheksort" (schermata *Bibliothek*) si ottengono all'inizio delle brevi schede che non rimandano ad al-

cun catalogo a stampa e che comprendono invece città e nome della biblioteca, segnatura del manoscritto, titolo, breve descrizione codicologica (per esempio: "Papier -243 Bl - Paris & Coburg - 1458-1464"), e il rimando all'elenco di tutti gli accessi a tale documento. Quindi, sempre dalla stessa ricerca, si ottengono le schede relative al Domarchiv e allo Stadtarchiv di Erfurt, provenienti dalla Berlin-Brandenburgische Akademie der Wissenschaften, reperibili, come abbiamo visto sopra, anche interrogando a partire dalla pagina *Hanschriftenkataloge online*. Infine si aggiungono gli accessi ad alcuni manoscritti conservati a Erfurt, sempre documentati attraverso le notizie archiviate dalla stessa accademia berlinese che hanno struttura ancora differente: città (Erfurt), segnatura, numero del foglio, autore, titolo, data, collegamento alla posta elettronica dell'accademia per richiesta di informazioni e indicazione completa della fonte (nel nostro caso, per esempio, ancora una volta: "Erfurt, Domarchiv. Handschriftenarchiv der Berlin-Brandenburgischen Akademie der Wissenschaften - Kasten 92").

Ognuno di questi record va dunque valutato con attenzione per coglierne la valenza e la derivazione, e gli utenti interessati avranno bisogno di qualche tempo per rendersi conto di quali aspettative si

possano avere da questo sito, che sfugge alle classificazioni fin qui rispettate dai cataloghi elettronici dei manoscritti, non proponendo né esclusivamente un catalogo o un inventario, né solo un indice.

Esso, che anzi prevede anche una lista di discussione, una serie di link alle risorse di rete dedicate ai manoscritti medioevali, e una sezione comprendente le immagini di alcuni manoscritti completi, rispetta criteri diversi di lavoro, unificati in primo luogo in base all'esigenza di rendere fruibile il maggior numero possibile di informazioni in un tempo adeguato.

Quanto tempo è occorso per produrre questo risultato? "*Manuscripta mediaevalia* è iniziato nel 1996 – risponde Robert Giel – e abbiamo lavorato per cinque anni. Dapprima abbiamo sviluppato uno strumento informatico, dedicandoci soprattutto alla struttura del database, poi abbiamo recuperato i cataloghi della DFG e quelli più antichi (precedenti al 1960), infine abbiamo aggiunto altri cataloghi a stampa apparsi dopo il 1999 che non erano compresi nel *Gesamtindex*. Ora la fatica maggiore riguarda l'interfaccia Internet, ma complessivamente lo sforzo più consistente dal punto di vista delle ore di lavoro è stato il recupero dei dati, perché questi non potevano essere trattati automaticamente e hanno richiesto

un impegno intellettuale notevole."

E i costi?

"Lasciando da parte le spese dei viaggi e altre voci minori, la parte maggiore dell'investimento è stata usata per pagare il lavoro di ricercatori e studenti. Inoltre tre persone delle istituzioni collaboratrici, me compreso, si sono dedicate al progetto per cinque anni. Credo che il costo totale ammonti a circa di 1 milione di euro".

[Ultima consultazione del sito: 31 agosto 2002]

Note

¹ Il progetto è stato presentato da R. GIEL, "*Manuscripta mediaevalia*". *Handschriften aus deutschen Bibliotheken im Internet*, "Gazette du livre médiéval", 39 (automne 2001), p. 34-40; altre informazioni in IFLA Section on Rare Books and Manuscripts. Newsletter, Summer 2002, disponibile all'indirizzo <<http://www.ifla.org/VII/s18/pubs/summer02.pdf>>.

² Ritorna più volte sulla tradizione catalografica tedesca A. PETRUCCI, *La descrizione del manoscritto*, 2a ed., Roma, Carocci, 2001 (Beni culturali, 24); al panorama storico offerto dal volume di Petrucci si può aggiungere la lettura delle recenti direttive di lavoro elaborate dalla DFG: GRUPPE "WISSENSCHAFTLICHE LITERATURVERSORGUNGS- UND INFORMATIONSSYSTEME", *Neue Konzepte der Handschriftenschiessung*, hrsg. von J. Bunzel, Bonn 06.09.01 (disponibile all'indirizzo <<http://www.dfg.de/foerder/biblio/download/handschriften.pdf>>); hanno recentemente discusso quest'ultimo documento A. CARTELLI – M. PALMA, *Towards the project of an open catalogue of manuscripts*, in *proceedings of the 2002 informing science & IT education joint conference* (Cork, 19-21 giugno 2002), pp. 217-224 (disponibile all'indirizzo <<http://is2002.com>>).

³ <<http://www.bbaw.de/forschung/dtm/HSA/startseite-handschriftenarchiv.htm>>.

⁴ <<http://www.bbaw.de/forschung/dtm/HSA/Bearbeiter.htm>>.

⁵ <<http://www.bbaw.de/forschung/dtm/HSA/HSA-DBI-Ortslist.htm>>.

⁶ <<http://dbix01.dbi-berlin.de:6100/DBI/login.html>>.

⁷ Si possono trovare notizie su questo prodotto all'indirizzo <<http://www.startext.de>>.

