

LE PAROLE DEL NOVECENTO: UN THESAURUS PER GLI ARCHIVI
PROGETTO DEL CONSORZIO BAICR SISTEMA CULTURA

Estratto dalla rivista
«Rassegna degli Archivi di Stato»
nuova serie I (2005), n. 1-2, pp. 241-244

ROMA 2005

Dal gennaio 2004 è stato avviato dal Consorzio BAICR Sistema Cultura uno studio sperimentale dedicato all'ipotesi di un vocabolario controllato come strumento per raffinare la ricerca archivistica sulle basi dati degli istituti che aderiscono alla rete *Archivi del Novecento*¹. La sperimentazione, condivisa nei suoi nodi concettuali e nei suoi obiettivi dagli organismi competenti dall'Amministrazione archivistica, ha potuto contare, grazie ad una convenzione stipulata il 18 dicembre 2003, sul sostegno e sul contributo della Direzione generale per gli archivi e successivamente del Dipartimento per i beni archivistici e librari.

Nel panorama dei progetti di informatizzazione di archivi storici, *Archivi del Novecento* costituisce un esempio di integrazione fra tradizione archivistica e innovazione informatica. La ricchezza della base dati, e il continuo accrescimento del numero degli istituti aderenti al progetto, rendono questo bacino archivistico un insieme di grande interesse, coinvolgendo attualmente gli archivi storici di 46 fra i principali istituti culturali italiani².

Avviato nel 1991, il progetto è fra le prime esperienze di costituzione di una rete archivistica³. Le relazioni fra fondi conservati da istituti diversi arricchiscono il patrimonio di conoscenze e fanno emergere la complementarietà di contenuti e protagonisti della produzione culturale del '900. Gli istituti in rete, utilizzando lo stesso software, condividono le risorse documentarie tramite la partecipazione ad un'unica banca dati archivistica disponibile *on line*.

Ciò che oggi sembra scontato, vale a dire la possibilità di far comunicare archivi diversi e di integrarli in maniera «virtuale», è stato, all'epoca dell'ideazione del progetto, frutto di una precoce intuizione sugli scenari che si andavano delineando con l'affermazione delle tecnologie digitali. Questa prospettiva si offriva non solo come soluzione al problema del ricongiungimento virtuale di carte appartenenti allo stesso soggetto produttore - pur conservate fisicamente in luoghi diversi - ma garantiva anche agli utenti la possibilità di interrogare trasversalmente fonti riguardanti temi affini o complementari.

Dall'intenzione di raffinare le potenzialità di una ricerca trasversale sui fondi archivistici, nasce l'idea di costruire un accesso tematico al complesso di descrizioni archivistiche consultabili *on line*.

Gran parte degli istituti coinvolti conservano documentazione relativa alla

¹ Il gruppo di progetto è composto da Sabrina Auricchio, Patrizia Gabrielli, Simona Luciani e Cristiana Pipitone. I risultati del lavoro sono sistematicamente discussi e revisionati con la supervisione scientifica dei referenti nei diversi ambiti disciplinari: Gianfranco Crupi, Antonio Dentoni-Litta, Guido Melis, Antonella Mulè, Leonardo Musci, Gabriella Nisticò, Mauro Tosti-Croce, Marisa Trigari, Lucia Zannino.

² Per l'elenco degli istituti aderenti e ogni altra informazione sul progetto vedi la pagina web www.archividelnovecento.it.

³ Per una descrizione generale e aggiornata del progetto Archivi del Novecento cfr. G. NISTICÒ, *Un grande progetto della memoria*, in «Millenovecento», 2004, 23, pp. 82-91; cfr. anche: M. CRATA, *L'interconnessione fra i beni culturali come valore per i servizi agli utenti in L'informatizzazione degli archivi storici e l'integrazione con altre banche dati culturali*, Atti della giornata di studio. Trento, 14 dicembre 1998, Trento, Provincia autonoma di Trento. Servizio beni librari e archivistici, 2001, pp. 87-92; L. MUSCI, *Archivi del Novecento*, in *Gli archivi dei partiti e dei movimenti politici. Considerazioni archivistiche e storiografiche*, a cura di S. SUPRANI, San Miniato, Archilab, 2001, pp. 97-104. La prima notizia del progetto è in G. NISTICÒ, *Informatica e archivi virtuali: ipotesi e problematiche nel progetto «Archivi del Novecento»*, in «Archivi per la storia», V (1992), 1, pp. 209-213.

storia politica e culturale del Novecento, tanto che questo può essere considerato il nucleo tematico prevalente. Oltre ai numerosi archivi politici, nell'ambito della storia della cultura e delle idee sono presenti ad esempio anche archivi inerenti la storia della filosofia, della scienza e della produzione artistica. Il contesto generale è quindi di ampiezza e di varietà sufficienti a costituire una casistica lessicale e metodologica di grande rilievo, offrendosi come terreno ideale per lo sviluppo di un prototipo di thesaurus.

Così il progetto «Le parole del Novecento » del Consorzio BAICR si è posto diversi obiettivi, primo fra tutti quello di verificare l'applicabilità in ambito archivistico di strumenti finalizzati al potenziamento della ricerca semantica, senza tuttavia sottovalutare la complessità e i rischi sottesi a tale percorso.

Il rapporto tra archivi e indicizzazione per soggetto è stato preso in considerazione più volte in passato, non senza sollevare dubbi e diffidenze, in particolare per il timore di snaturare le specificità dell'ambito archivistico. Da una parte emergeva la preoccupazione che l'accesso tematico potesse far tornare indietro la disciplina a pratiche di discutibili accorpamenti per materia, dall'altro si pensava di non poter applicare agli archivi strumenti messi a punto in ambito biblioteconomico, senza opportuni adattamenti alla specificità degli archivi.

L'intenzione, nel riprendere il filo delle riflessioni già avviate, è stata quella di verificare sino in fondo le potenzialità e i limiti dell'applicabilità agli archivi degli strumenti di indicizzazione in uso nelle biblioteche.

La principale differenza nell'impiego di strumenti di accesso per materia agli archivi consiste nel fatto che all'interno di una base dati archivistica ciò che deve essere indicizzato è il contenuto delle schede descrittive, piuttosto che il contenuto dei documenti, laddove invece i bibliotecari compiono un'operazione di rappresentazione a partire dal documento stesso (analisi concettuale), per rendere in maniera sintetica e normalizzata il contenuto dell'opera. La differenza è sostanziale, per motivazioni non solo di ordine metodologico, ma anche di natura operativa. L'indicizzazione tematica costituisce una sintesi informativa che viene elaborata su una fonte la quale, a sua volta, è una informazione mediata e rappresentata sinteticamente nella scheda descrittiva compilata dall'archivista («Access points are based upon the elements of description. The value of access points is enhanced through authority control», *ISAD G, II ed., point 1.14*).

Nel corso della sperimentazione sono poi stati presi in esame due aspetti principali emersi dalle riflessioni già avviate in campo archivistico, relativamente al rapporto tra indicizzazione semantica e contesto gerarchico. Da un lato, ci si è confrontati con il presupposto secondo cui il grado di genericità/specificità dei descrittori dovesse riflettere il livello gerarchico della descrizione archivistica: i termini dell'indice a livello del fondo dovevano pertanto risultare più generici dei termini dell'indice rispettivamente a livello di serie, fascicolo, ecc. Dall'altro, è stata superata l'idea che l'indicizzazione dovesse riprodurre, all'interno di una stringa di soggetto, la gerarchia propria del contesto archivistico.

Bisogna partire dal presupposto che l'indicizzazione tematica non sostituisce e non deve cercare di riprodurre la struttura archivistica, ma deve basarsi

strettamente sulla descrizione: il livello di analiticità dei descrittori sarà quindi strettamente dipendente da quello delle schede. Ciò che deve essere omogeneo è il rapporto tra descrizione e descrittore e non quello tra descrittori e livelli archivistici. Questo problema assume un notevole rilievo se si pensa di indicizzare insiemi archivistici allargati (come la rete *Archivi del Novecento*) e non omogenei riguardo al livello di descrizione. La prospettiva, infatti, cambia se si vuole indicizzare un singolo fondo, più fondi all'interno di un singolo istituto, o più fondi conservati in istituti diversi. Ed è in quest'ultimo caso o nel caso di vastissimi complessi archivistici che la ricerca per thesaurus raggiunge la sua massima efficacia.

È evidente inoltre che l'approccio tematico non può sostituire la ricerca di tipo storico-istituzionale; né i due metodi possono essere confusi o sovrapposti. La ricerca tramite indici semantici deve essere un complemento rispetto alla ricerca che muove dal soggetto produttore della documentazione.

D'altra parte, come in più occasioni è stato sottolineato, nel lavoro d'archivio si sono sempre prodotti anche indici per materia o delle «cose notevoli», indici che hanno puntualmente rivelato la loro utilità, scontando tuttavia un limite di efficacia proprio per la mancanza di criteri di normalizzazione ben definiti.

Il progetto « Le parole del Novecento » nasce proprio dall'esigenza di disporre di una metodologia a carattere scientifico per elaborare parole chiave e organizzarle in una rete di relazioni semantiche in grado di ampliare e rendere più mirata e fruttuosa la ricerca archivistica *on line*.

L'approccio sperimentale, che ha caratterizzato le fasi di realizzazione del progetto, ha consentito di procedere senza preconcetti né forzature. Il lavoro fin qui svolto ha coperto l'arco temporale di circa 18 mesi, scandito in tre fasi, la terza ancora in corso per le ultime verifiche sul campo. La prima ha preso le mosse dallo studio delle rare esperienze di indicizzazione per soggetto realizzate fino ad oggi in ambito archivistico e dalle pubblicazioni disponibili sul tema, per tostare la loro applicabilità ad una base dati archivistica di grande ricchezza e complessità come quella di *Archivi del Novecento*. Sulla scorta delle esperienze esaminate e di prime simulazioni si è cercato di individuare le soluzioni che potevano essere adottate, come per esempio la scelta tra il metodo di indicizzazione pre- o post-coordinato, o la scelta sul linguaggio dei descrittori, nonché sui legami per la strutturazione del prototipo.

A ciò è seguita la seconda fase di verifica dell'applicabilità ai mondo archivistico delle norme internazionali per l'indicizzazione semantica e degli standard per la costruzione dei *thesauri* (ISO 2788) sino a uno sviluppo di un prototipo di thesaurus iniziato con l'operazione fondamentale di indicizzazione della base dati. Man mano che il vocabolario cresceva, si procedeva alla revisione e agli aggiustamenti della strutturazione gerarchica, organizzando infine i descrittori in categorie o sottosectori specialistici, che hanno lo scopo di facilitare l'orientamento degli utenti nella navigazione all'interno del thesaurus.

La terza fase del progetto ancora in corso è dedicata alla stesura dei criteri metodologici, frutto sia del lavoro di indicizzazione sia delle riflessioni teoriche emerse nel corso di essa. Un'attenzione particolare è stata dedicata alla scelta del registro linguistico, sollecitata dall'esigenza di salvaguardare la ricchezza

espressiva dei termini connotati storicamente e tematicamente e al tempo stesso di predisporre uno strumento facilmente utilizzabile da una utenza non specializzata.

Le possibilità di utilizzare al meglio un simile strumento saranno ampliate in maniera decisiva dalla presenza di un modulo thesaurus per la gestione informatica della ricerca. Il thesaurus è uno strumento per sua natura complesso, che spesso risulta ostico perfino agli specialisti, e che, pur offrendo grandi opportunità per ampliare e raffinare la ricerca, rischia di non essere utilizzato appieno. Un'interfaccia amichevole del thesaurus può aiutare a superare la diffidenza degli utenti, migliorandone il grado di utilizzazione.

CONSORZIO BAICR SISTEMA CULTURA