

Lezione-conferenza di Paul Tombeur su
*Le basi dei dati testuali e lessicali per lo studio
dell'Occidente latino:
una vera rivoluzione per i linguisti,
gli storici, i filosofi e i teologi.*

Palermo, Officina di Studi Medievali, 5 aprile 2008

In coordinamento con il Master universitario di I Livello *I mestieri del libro e del documento: conservazione, fruizione, restauro, catalogazione e nuove tecnologie editoriali*, nell'ambito di un programma di attività formative e culturali volte a promuovere la "cultura del libro e del documento" coordinato dall'Officina di Studi Medievali si è svolta, il 5 aprile 2008 presso la sala "Sisto IV" della Basilica di San Francesco d'Assisi, la lezione-conferenza di Paul Tombeur (Université Catholique de Louvain).

Paul Tombeur, nel corso di una lezione d'impostazione prettamente seminariale, ha esposto ai presenti, mediante l'ausilio di un computer collegato a Internet e di un video-proiettore, il risultato delle proprie ricerche applicate alla progettazione di basi di dati per la ricerca testuale nell'ambito degli studi umanistici; per la realizzazione di tali banche dati fruibili *on-line* lo *staff* di ricerca di Paul Tombeur si è avvalso della competenza e della preziosa collaborazione dell'editore Brepols.

Oggi l'abbondanza di nuovi supporti digitali tecnologicamente sempre più avanzati, versatili, economici e dalla diffusione sempre più capillare e rapida, e la conseguente trasformazione in forma digitale (da essi innescata) di tutto quel patrimonio informativo e culturale, che fino a qualche anno fa potevamo soltanto sfogliare nelle grandi biblioteche pubbliche e di prestigiosi enti di ricerca, ha dato vita alla possibilità di un accesso diretto alle informazioni da ogni luogo e da parte di chiunque: infatti, se prima, per effettuare una ricerca testuale, era necessario sfogliare le fonti a stampa sulla materia oggetto dei nostri studi, oggi, invece, si vanno affermando strumenti informatici che rapidamente, grazie all'utilizzo di particolari *softwares*, riescono a compiere complesse ricerche "incrociate" partendo da una parola chiave, offrendo allo studioso molteplici possibilità di ricerca. Questo stimolante scenario tecnologico che lo *Humanities Computing* offre ad un vasto pubblico prende il nome di *Information Retrieval* (reperimento di informazioni): nella pratica tradizionale dell'*Information Retrieval* ogni tipo di ricerca documentale è riportato alle condizio-

ni di una ricerca tramite linguaggio testuale *term-based* (basate sui termini, sulla logica delle parole chiave).

In genere si possono utilizzare due differenti strategie di ricerca:

1. Ricerca mirata in funzione del contenuto (*querying*): classificazione argomentale; strutturazione del contenuto; indicizzazione; dizionari dei sinonimi; lemmatizzazione;
2. Ricerca esplorativa (*browsing*): navigazione; approssimazioni successive; storia e orientamento; ricerca incrementale.

A seguito di una interrogazione, il sistema segnala all'utente il numero di documenti ritrovati rispondenti alle caratteristiche della chiave di ricerca impostata. L'utente può riformulare, specializzare o generalizzare l'interrogazione fino a che il numero di documenti ritrovati non appaia soddisfacente. Tutto ciò rappresenta il sostrato tecnologico che il *CTLO* (*Centre Traditio Litterarum Occidentalium*), sotto la direzione di Paul Tombeur, è riuscito a condensare in un unico strumento digitale di metaricerca, utilissimo per gli studi in campo linguistico, storico, filosofico e teologico: il *CDS* (*Cross Database Searchtool for Latin Databases*). Questa particolare banca dati, fruibile *on-line*, interroga simultaneamente varie banche dati per la ricerca testuale specializzata: la *Library of Latin Texts* (raccolge più di 3000 testi attribuiti a circa 900 autori, dall'antichità alla letteratura neolatina), i *Monumenta Germaniae Historica* (degli oltre 300 volumi di cui è composta questa monumentale opera sono fruibili in digitale, ad oggi, più di 1000 testi), l'*Archive of Celtic-Latin Literature* (una raccolta di testi in latino prodotti in area anglofona dal 400 al 1200) e l'*Aristoteles Latinus Database* (*corpus* completo delle traduzioni medievali delle opere di Aristotele). L'interfaccia offerta all'utente si rivela molto semplice e intuitiva, ma soprattutto usabile e personalizzabile nella propria lingua; essa rappresenta il punto di partenza comune per compiere ricerche nei diversi corpora testuali proposti dal sistema *on-line*.

Altro efficacissimo strumento di ricerca testuale *on-line* è il *DLD* (*Database of Latin Dictionaries*): esso rappresenta un progetto mirato a fornire (ad un pubblico di studenti e specialisti) un unico strumento in grado di interrogare *on-line* tre tipologie di dizionari utili per la traduzione dal latino alle lingue moderne e per l'analisi etimologica dei termini. Vero tratto distintivo è, però, quello di assicurare il collegamento ipertestuale tra questi dizionari e il *CDS* precedentemente descritto (e viceversa): si offre all'utente, cioè, la possibilità di effettuare la ricerca di una parola all'interno di una delle banche dati che formano il *CDS* e di passare istantaneamente dal risultato ottenuto ad una nuova ricerca sul *DLD* e viceversa, dai dizionari del *DLD* al *corpus* del *CDS*.

FABIO CUSIMANO