

ALBERTO CASADEI

*Nuove frontiere dell'Umanistica digitale*

In

*L'Italianistica oggi: ricerca e didattica*, Atti del XIX Congresso  
dell'ADI - Associazione degli Italianisti (Roma, 9-12 settembre 2015),  
a cura di B. Alfonzetti, T. Cancro, V. Di Iasio, E. Pietrobon,  
Roma, Adi editore, 2017  
Isbn: 978-884675137-9

Come citare:

Url = [http://www.italianisti.it/Atti-di-Congresso?pg=cms&ext=p&cms\\_codsec=14&cms\\_codcms=896](http://www.italianisti.it/Atti-di-Congresso?pg=cms&ext=p&cms_codsec=14&cms_codcms=896)  
[data consultazione: gg/mm/aaaa]

ALBERTO CASADEI

*Nuove frontiere dell'Umanistica digitale\**

*In questo breve contributo vengono indicati, a mò di premessa, alcuni dei possibili sviluppi dell'Umanistica digitale applicata alla letteratura italiana, dopo l'importante stagione che si è aperta negli anni Novanta del secolo scorso.*

Sono trascorsi sessantacinque anni da quando Gianfranco Contini lanciava il suo arguto appello parafrascoliano: “o Italiani, io vi esorto alle Concordanze”<sup>1</sup>. È passata più di un'epoca da quel momento, quando ancora sembrava che un'indagine statistica su un'opera letteraria fosse inutile o eretica; soprattutto, si è moltiplicato il potenziale delle ricerche sistematiche, grazie a un primo passaggio dal manuale-cartaceo al digitale-elettronico, e poi a un secondo che ha reso possibile l'unione di database diversi nella rete web. Ricorderemo tra breve alcune delle tappe che riguardano da vicino il lavoro dell'italianista; ma il problema che ci dovremo porre è soprattutto quello degli eventuali nuovi metodi di ricerca, per arrivare a una rinnovata “repubblica delle lettere”, in grado di accogliere le sfide attuali dell'analisi storico-critica<sup>2</sup>.

In effetti, la disponibilità di concordanze informatizzate, sempre più ampia a partire dagli anni Ottanta e Novanta del secolo scorso, ha reso possibile innanzitutto ricerche sistematiche di tipo lessicale, che si sono rivelate essenziali per gli studi di storia della lingua, ma che hanno avuto una ricaduta anche per la critica: basti pensare, nel campo dell'italianistica, alle possibilità offerte da uno strumento come la *Liz*, la *Letteratura italiana Zanichelli* in CD-Rom, ideata da Pasquale Stoppelli e Eugenio Picchi, che consentiva, già nella sua prima versione del 1993, di indagare i testi di 362 opere di 109 autori da San Francesco a Svevo; il programma di interrogazione DBT, progettato presso l'Istituto di Linguistica computazionale del CNR di Pisa, consentiva non solo di produrre concordanze ma anche di effettuare ricerche di occorrenze, selezionare prefissi, suffissi e qualsiasi stringa di caratteri, generare indici e altro ancora<sup>3</sup>.

Un ulteriore passo, nell'ambito del potenziamento dei dati disponibili, è stato ovviamente segnato dalla creazione della *Biblioteca italiana telematica*. Come recita la pagina di presentazione

\* Questa breve introduzione alla sessione *Metodologie della ricerca. L'umanistica digitale* (9 settembre 2015) è stata da me rivista con aggiunte bibliografiche sino a tutto il 2016. Gli esempi e i riferimenti restano comunque puramente indicativi, specie per la sitografia, in continua evoluzione.

<sup>1</sup> Cfr. G. CONTINI, *Preliminari sulla lingua del Petrarca* (1951), poi in *Varianti e altra linguistica. Una raccolta di saggi (1938-1968)*, Torino, Einaudi, 1970, 184.

<sup>2</sup> Si fa riferimento all'ultimo saggio di uno dei maggiori teorici delle *Digital Humanities*, Jerome McGann: si veda ID., *A New Republic of Letters*, Cambridge (MA), Harvard U.P., 2014. Per un quadro generale e per ampia bibliografia, si vedano i *Companions* pubblicati dalla Blackwell e in parte disponibili *on-line*: <http://www.digitalhumanities.org/companion> e companionDLS. Da vedere anche *The Johns Hopkins Guide to Digital Media*, ed. by M.L. Ryan, L. Emerson, B.J. Robertson, Baltimore, Johns Hopkins U.P., 2014. Per una discussione metodologica aggiornata, cfr. *Debates in the D. H.*, ed. by M.K. Gold e L.F. Klein (Un. of Minnesota P., 2012), ora in *open access* attraverso il sito <http://dhdebates.gc.cuny.edu>. Un portale di grande interesse è quello dell'*Alliance of D.H. Organizations* (<http://adho.org>); in ambito europeo, si veda *European Association for digital humanities* (<http://eadh.org>). Per l'Italia, si veda innanzitutto il sito dell'*Associazione per l'informatica umanistica e la cultura digitale* (<http://www.umanisticadigitale.it>), presieduta da Fabio Ciotti (del quale si veda, con GINO RONCAGLIA, *Il mondo digitale. Introduzione ai nuovi media*, Roma-Bari, Laterza, 2010). Importante poi, in un contesto allargato, il lavoro svolto nell'ambito del progetto europeo *Dariah.eu*: per un'informazione di base, si veda il sito <http://stdl.cnr.it/it/dariah>.

<sup>3</sup> Per una rassegna delle principali tappe di questa fase si veda P.G. BELTRAMI - A. BOCCELLARI, *Banche dati e dizionari on-line. Il "Tesoro della Lingua Italiana delle Origini" e la banca dati dell'italiano antico dell'"Opera del Vocabolario Italiano"*, in *Nuovi media e lessicografia storica*, a cura di W. Schweickard, Tübingen, Niemeyer, 2006, 3-14. Nell'ambito delle concordanze relative alla letteratura italiana, vanno ricordate le numerose coordinate da Giuseppe Savoca, già promotore e presidente del Consiglio scientifico del Gruppo Nazionale di Coordinamento del Consiglio Nazionale delle Ricerche per le *Concordanze della Lingua Italiana Poetica dell'Otto/Novecento* (CLIPON): sulla metodologia, si veda il suo volume *Lessicografia letteraria e metodo concordanziale*, Firenze, Olschki, 2000.

nel sito ufficiale ([www.bibliotecaitaliana.it/progetto](http://www.bibliotecaitaliana.it/progetto)), promossa dal Centro Interuniversitario Biblioteca italiana Telematica (CIBIT), costituito nel 1997 da quattordici università italiane, la Biblioteca Telematica Italiana (BIT) venne presentata a Roma il 2 marzo 2000. Il coordinatore scientifico del programma di ricerca era il professor Amedeo Quondam dell'Università degli Studi di Roma "La Sapienza". La Biblioteca Telematica Italiana, realizzata dal Dipartimento di Italianistica e Spettacolo presso l'Università di Roma "La Sapienza", riuscì a raccogliere e a rendere disponibile in rete per la consultazione una consistente quantità di testi. L'archivio elettronico era inoltre interrogabile tramite il sistema DBT, che consentiva di effettuare ricerche di tipo linguistico e testuale.

In circa vent'anni, la *Bibit* ha raggiunto un'espansione notevolissima, con circa 1600 opere codificate in XML/TEI, criticamente riviste e liberamente accessibili; a questo nucleo se ne affiancano ormai numerosi altri, come quello degli *Scrittori d'Italia*, che propone in versione digitalizzata la storica collana laterziana voluta da Benedetto Croce; quello degli *Incunabili d'Italia*, in collaborazione con l'Istituto Centrale per il Catalogo Unico; quelli relativi alle collezioni speciali, per esempio con digitalizzazioni di manoscritti del *Cortegiano* e del *Furioso*<sup>4</sup>.

Repertori e *database* così ricchi e potenzialmente implementabili senza particolari vincoli sembravano aver coperto il campo delle realizzazioni di umanistica digitale più rilevanti per le ricerche di letteratura italiana, ma non è così. Oltre a nuovi grandi archivi, spesso con la possibilità di consultare direttamente autografi, manoscritti e anche immagini<sup>5</sup>, vanno ora segnalate le nuove sfide riconducibili all'uso sistematico della rete internet, e in particolare quella del cosiddetto *web semantico*: l'immissione dei meri oggetti digitalizzati non è più sufficiente per una buona visibilità e una facilità di ricerca e di accesso in rete, e occorre invece trattare i materiali secondo *ontologie* condivise a livello internazionale, in modo da essere individuabili e connessi anche in futuro, qualunque possa essere l'evoluzione del *World Wide Web*. Al di là quindi delle imponenti acquisizioni di materiali cartacei (per esempio attraverso Google Books)<sup>6</sup>, e alla possibilità di condividere facilmente lavori della ricerca (si pensi a Academia.edu che, partita nel 2008, conta già oltre 25 milioni di utenti in tutto il mondo), è ormai importante rendere riconoscibili i dati immessi e possibili le implementazioni future, grazie all'impiego di marcature *standard* o *scorm* e di sistemi operativi *open-access*. In questo senso, a parte i grandi

<sup>4</sup> A questi nuclei vanno aggiunte numerosi link a preziose *Risorse*, in particolare altri archivi informatizzati incentrati su grandi autori (opere dantesche lemmatizzate in DanteSearch, i *Triumphs* di Petrarca, ecc.), oppure su testi archivistici (di Ferrara, Pisa, ecc.). Dalla pagina <http://www.bibliotecaitaliana.it/collezioni/risorse> sono raggiungibili poi numerosi altri *corpora* e strumenti di ricerca, come l'*Archivio metrico italiano* sviluppato all'Università di Padova: <http://www.maldura.unipd.it/ami/php/ricerca.php>. A fronte di questa evidente rilevanza, va lamentata la ristrettezza di finanziamenti, che sta attualmente rendendo problematico il mantenimento e l'utilizzo di *Bibit*.

<sup>5</sup> Fra gli esempi più recenti, si vedano almeno: *Giovanni Pascoli nello specchio delle sue carte. L'archivio e la casa di Giovanni e Maria Pascoli a Castelvecchio* (<http://pascoli.archivi.beniculturali.it>); l'*Archivio Gozzano Pavese* (<http://www.ad900.it/homesito.asp?IDSito=1>); l'*Archivio Palazzeschi* (<http://www.ad900.it/homesito.asp?idsito=3>). Molti di questi siti sono coordinati sotto il meta-archivio *Carte d'autore*: <http://www.cartedautore.it>. Da citare anche i repertori specificamente dedicati ad aspetti multimediali, come nel progetto *L'Orlando furioso e la sua traduzione in immagini* (<http://orlandofurioso.org>), elaborato dal "Centro di elaborazione informatica di testi e immagini nella tradizione letteraria" (CTL) della Scuola Normale Superiore.

<sup>6</sup> Lanciata nel 2004, questa funzione del browser Google è stata osteggiata ma ha ormai raggiunto (dati 2015) la quota di quasi 26 milioni di libri digitalizzati in tutto il mondo, resi liberamente disponibili se fuori *copyright*; anche le Biblioteche Nazionali Centrali di Firenze e Roma collaborano al progetto dal 2012. Da ricordare anche l'Internet Archive ([www.archive.org](http://www.archive.org)), fondato nel 1996 a San Francisco, e membro dell'Open Book Alliance (<http://openbookalliance.org>): i testi disponibili (anche file multimediali) sono circa 2 milioni. Questi strumenti sono attualmente utili in particolare per il reperimento di edizioni fuori commercio dell'Ottocento e del Novecento (si vedano, a titolo di esempio, quelle dell'*Orlando furioso* reperibili e spesso scaricabili da [www.openlibrary.org](http://www.openlibrary.org)), mentre, per le edizioni rare sino al Settecento, sono fondamentali i servizi digitalizzati *on-line* delle grandi biblioteche nazionali: per l'Italia, oltre ai siti SBN, si veda: <http://www.internetculturale.it/opencms/opencms/it>.

archivi di nuova generazione a livello internazionale, come *Europeana* (<http://www.europeana.eu/portal>), con oltre 53 milioni di oggetti digitali consultabili, sono importanti le nuove ricerche che possono portare a visualizzare in modo inedito materiali diversificati, per esempio per creare reti di relazioni in un determinato periodo storico: è il caso, già piuttosto evoluto, del progetto *Circulation of knowledge and learned practices in the 17th-century Dutch Republic* (<http://ckcc.huygens.knaw.nl>), che permette di superare i confini settoriali (gli epistolari di singoli intellettuali) per creare un'autentica rete dei contatti e degli scambi culturali. Oppure, nell'ambito del progetto PRIN *Per una enciclopedia dantesca digitale*, coordinato dall'Università di Pisa (2013-2016), è stata realizzata una funzione, *Dante Sources*, che permette di interrogare secondo diverse chiavi il repertorio delle fonti primarie delle opere dantesche, con la possibilità di creare grafici e tabelle, nonché di integrare e correggere i dati disponibili<sup>7</sup>.

Si stanno aprendo insomma numerosi campi d'indagine che dovrebbero permettere di passare da una concezione statica dei testi a una dinamica, con la possibilità di legare mondi sinora spesso separati, a livello di umanistica digitale, come quelli della letteratura, delle arti o delle scienze: del resto, la creazione di nuovi progetti collaborativi è stata indicata da tempo come una necessità ineludibile per lo sviluppo dell'Umanistica digitale<sup>8</sup>.

Ma a questo punto, più che mai si impone una riflessione sulle priorità da indicare nella realizzazione di nuovi strumenti importanti per la ricerca letteraria e in specie per l'italianistica. È infatti chiaro che l'aumento esponenziale dei dati disponibili (si potrà parlare di *Big data* umanistici? O meglio ancora di *Rich data*?) impone la realizzazione di tecnologie sempre più sofisticate, sia a livello di *hardware* che di *software*. Inoltre, la fruizione di molte strumentazioni direttamente dal web, con la possibilità di arricchimenti sinora impensabili (per esempio per le nuove edizioni critiche), pone un serio problema di organizzazione della ricerca, specie a livello di gruppi con varie competenze individuali<sup>9</sup>. Ma le problematiche non si limitano a questi aspetti.

Infatti, sembra ormai ineludibile una nuova riflessione sul concetto stesso di testo, ancora fondato su paradigmi linguistico-strutturalisti, e in pratica su un presupposto di unitarietà ancora simboleggiato dalla forma-libro, ma almeno in parte superato: basti pensare alle nuove forme di interazione autori-lettori nel realizzare narrazioni *on-line*, che peraltro riproducono, sia pure con le ovvie diversità tecnologiche, modalità di lunghissima durata nella fase dell'oralità<sup>10</sup>.

Ancora più in concreto, il problema per l'analisi critica è ora quello di riuscire a pensare anche a forme di orientamento dei dati, che siano certamente statistici e digitalizzati, ma possano interagire con i *qualia*, cioè i sensi riconoscibili attraverso un'interpretazione di secondo grado (l'*analogico* umanistico). Sono insomma sempre più necessarie 'macchine ermeneutiche', che

<sup>7</sup> Il link è <http://perunaenciclopediadantescadigitale.eu/dantesources/index.html>. Sulle specifiche tecniche, oltre al sito si veda più oltre il contributo presentato in questo fascicolo. Il progetto ha vinto il riconoscimento nel DH Awards 2015, sezione *Tool or suite of tools*: <http://dhawards.org/dhawards2015/results>.

<sup>8</sup> Per il settore delle *Digital Humanities* si veda la traduzione italiana di A. BURDICK, J. DRUCKER, P. LUNENFELD, T. PRESNER, J. SCHNAPP, *Digital Humanities*, ed. orig. 2012, trad. it. *Umanistica Digitale*, Milano, Mondadori, 2014: sito web <http://hyperstudio.mit.edu>. Si veda anche F. TOMASI, *Metodologie informatiche e discipline umanistiche*, Roma, Carocci, 2008. Sul versante della preparazione di materiali in *epub*, cfr. M.T. LUPIA, M. TAVOSANIS, V. GERVASI, *Editoria digitale*, Torino, Utet Università, 2011: ma la diffusione degli ebook arricchiti, che sarebbero fondamentali per proporre nuovi commenti stratificati e multimediali, incontra numerosi ostacoli, tanto che bisognerebbe pensare a reti di università per la realizzazione di prototipi, da diffondere poi in versioni differenziate (anche *open source*).

<sup>9</sup> In questo senso, è già molto significativa la nuova impostazione seguita da Paolo Trovato per realizzare la nuova edizione critica del poema dantesco, su cui si veda da ultimo E. Tonello-P. Trovato (a cura di), *Nuove prospettive sulla tradizione della «Commedia». Seconda serie (2008-2013)*, Padova, Libreriauniversitaria.it, 2013. Si veda inoltre D. FIORMONTE, *Scrittura e filologia nell'era digitale*, Torino, Bollati Boringhieri, 2003.

<sup>10</sup> Su questi aspetti mi permetto di rinviare alla mia voce *Letteratura e web* nell'*Enciclopedia Italiana*, Appendice IX, Roma, Treccani, 2015, vol. J-Z, 46-49, con altra bibliografia. Si vedano anche l'ultima edizione di G.P. LANDOW, *Hypertext 3.0: Critical Theory and New Media in an Era of Globalization*, Baltimore, Johns Hopkins U.P., 2006<sup>3</sup>, e, per un inquadramento aggiornato, A. Buffardi, D. de Kerckhove, *Il sapere digitale. Pensiero ipertestuale e conoscenza connettiva*, Napoli, Liguori, 2011.

possano tener conto delle scalarità cognitive e non solo dei macrodati quantitativi: il raffinamento degli algoritmi per le ricerche stilistiche potrà essere un altro campo di grande importanza<sup>11</sup>, nella consapevolezza che, al momento, è sempre necessaria la complessità delle competenze e delle intuizioni umane per riuscire, come avrebbe detto un grande filologo, a conoscere sempre meglio il conoscibile in ambito letterario (e non solo).

---

<sup>11</sup> Torna sulla questione, con altra bibliografia, P. ITALIA, *Editing Novecento*, Roma, Salerno Ed., 2013. Un caso recente e interessante è quello del *Diario postumo* attribuito a Eugenio Montale, ma con ogni probabilità confezionato da Annalisa Cima: molto interessanti, anche per l'impiego di specifici algoritmi, i risultati di un Convegno tenutosi a Bologna l'11 novembre 2014 e i cui Atti sono stati editi col titolo *Montale e pseudo-Montale. Autopsia del "Diario postumo"*, a cura di F. Condello, V. Garulli, F. Tomasi, Bologna, Bononia U.P., 2016.