

## **Alcuni problemi teorici e pratici posti dalle *digital libraries***

Raul Mordenti

(Università degli Studi di Roma "Tor Vergata", Italia)

**Abstract** In this article we will point out the substantial difference between 'digital edition' and 'informatic edition': the first one (DH = Digital Humanities) is the simple digitalisation of a manuscript or a printed text; the second one (HC = Humanities Computing) is an edition created by taking into account the potential of the computer itself and with the intention of exploiting the texts completely. This implies that the informatic edition is designed and built in accordance with the basic laws that allow us to give instructions to a computer (completeness, expediency, univocity, no contradiction, no redundancy, etc.). The key word of an informatic edition is encoding, which is the basis for the scientificity of the edition. Moreover, the possibility of using the computer and its potential to enjoy the text depend on it. Finally, we will deal with the different ways of disseminating the texts, noting some radical differences between the traditional editorial circuit, which is typical of the printed text, and the new one established by computers and new technologies (first of all, the indefensibility of copyright).

**Sommario** 1 Per una definizione più rigorosa di DH e HC. – 2 La nuova biblioteca digitale: caratteristiche, problemi e necessità.

**Keywords** Digital humanities. Digital library. Text. Copyleft. Humanities computing.

### **1 Per una definizione più rigorosa di DH e HC**

Vorrei sollevare due soli problemi, che mi sembra siano già emersi – sia pure in forma implicita e indiretta – in molti interventi: 1) il problema di una definizione teorica più precisa della disciplina che (ne siamo noi coscienti oppure no) stiamo fondando, e 2) il problema del rapporto della nostra collettiva biblioteca digitale con quella cartacea e gutemberghiana di cui essa prenderà il posto.

Non sembri trascurabile lo sforzo di una precisazione terminologica dato che essa comporta necessariamente anche una importante precisazione concettuale. La distinzione fondamentale che si propone è quella fra HC (Humanities Computing) e DH (Digital Humanities).

Dico subito che la necessità di tradurre in italiano queste due espressioni non ci aiuta, giacché mi sembra che fra noi manchi a tutt'oggi una traduzione condivisa di queste espressioni inglesi (una traduzione potreb-

be forse essere, rispettivamente: 'Informatica Umanistica' e 'Umanistica Digitale', ma si aspettano proposte migliori); così come non ci aiuta il fatto che la formula DH (forse semplicemente per la sua comodità) sembra essere entrata ormai nell'uso comune, fino a partecipare di molti acronimi di importanti istituzioni del nostro settore, sia di ambito europeo, come EADH (The European Association for Digital Humanities), che tuttavia ha mutato il nome originale (risalente al 1973) di ALLC (Association for Literary and Linguistic Computing), sia anche di ambito mondiale, come ADHO (Alliance of Digital Humanities Organizations); ma mi piace notare che la nostra italiana AIUCD (Associazione per l'Informatica Umanistica e la Cultura Digitale), costituita a Firenze nel 2011, ha scelto il suo nome, dopo un lungo dibattito, proprio per unire i due ambiti, cioè - se ho ben inteso - quello più specifico e proprio dell'Informatica Umanistica e quello più vasto della Cultura Digitale.

Il punto è che, come ci ha insegnato Tito Orlandi, esiste effettivamente una differenza sostanziale fra il 'Digitale' (da cui DH, Digital Humanities) e il 'Computazionale' (da cui HC, Humanities Computing). E proprio il nostro campo della filologia e delle edizioni, si presta a chiarire, come meglio non si potrebbe, di quale differenza si tratti. Un'edizione digitale, o digitalizzata, è la semplice trasformazione di un oggetto analogico (nel nostro caso un manoscritto o una stampa o altro) in un oggetto digitale, e questo - almeno a partire dall'invenzione degli scanner - può anche avvenire senza alcun intervento critico da parte dello studioso, più o meno come avveniva per fare una fotocopia.

Un'edizione computazionale (che io chiamerei semplicemente informatica) è invece un'edizione che si svolge a partire dalle potenzialità euristiche della computazione (cioè della macchina di Turing) con l'intenzione di sfruttarle pienamente per migliorare la conoscibilità del testo. Ciò comporta due diversità fondamentali fra HC e DH (che non significano affatto - vorrei essere del tutto chiaro su questo punto - una qualche gerarchia fra HC e DH).

Anzitutto solo nel primo caso (HC), e non nel secondo (DH), l'uso dell'informatica si pone un obiettivo propriamente euristico, cioè intende conseguire un significativo plusvalore conoscitivo, che può consistere nel produrre edizioni critiche migliori o nel consentire operazioni analitiche impossibili senza l'informatica, nel definire ontologie adeguate ai nostri campi di ricerca o nel supportare originali analisi semi-automatiche, ecc. È la computazione che fa del computer, come è stato detto, un 'medium cognitivo'. Sembra invece caratterizzare DH in incremento (voglio ripeterlo: altrettanto utile e prezioso) della comodità di accesso ai dati e della loro diffusione, come ad esempio la produzione delle grandi e ormai indispensabili banche dati testuali.

In secondo luogo solo nel primo caso (HC), e non nel secondo (DH), è implicata, e anzi diventa centrale, la procedura della ricerca, più precisa-

mente la riutilizzazione e la ridefinizione delle procedure caratteristiche delle nostre discipline umanistiche che sono chiamate a operare nel nuovo ambiente informatico. Personalmente credo che proprio questo sia il punto più prezioso (e – se posso dirlo – anche entusiasmante) del nostro attuale lavoro, quello in cui le nostre discipline umanistiche e il loro secolare corredo di procedure, venendo messe alla prova nel nuovo ambiente segnato dall'informatica, ritrovano, ma anche modificano e affinano, se stesse.

Penso anche che consista esattamente in questo punto un apporto possibile degli umanisti, che non si limitano dunque più a 'ricevere' e usare qualcosa che proviene dagli informatici ma possono a loro volta fornire qualcosa di inedito, per ipotesi utile allo sforzo comune di incrementare la conoscenza.

Limitandoci al nostro campo che riguarda la filologia e i testi, si potrebbe dire che la Digitalizzazione lavora sull'oggetto della ricerca (ora reso in forma digitale e non più analogica) e si arresta invece di fronte alle porte del testo in quanto tale, cioè in quanto messaggio storico e semantico (accade questo, ad esempio, quando vengono fornite in formato digitale le immagini dei manoscritti), mentre la Computazione intende proprio attraversare quella soglia fatale per incrementare, attraverso l'uso della macchina informatica, le nostre capacità di lettura dell'oggetto testuale.

Estremizzando forse il nostro ragionamento, credo che potrebbero esistere forme di Digitalizzazione che non prevedono a rigore Computazione, mentre la Computazione presuppone sempre la Digitalizzazione. Più precisamente: la Digitalizzazione si pone all'inizio come premessa necessaria della Computazione (dato che la macchina informatica lavora solo su dati digitali) e poi necessariamente anche alla fine (dato che intendiamo comunque proporre in formato digitale i dati da noi elaborati). Ma proprio questa successione lineare, o intreccio, Digitalizzazione-Computazione-Digitalizzazione, in cui però – si noti bene – nessuno dei due momenti è compreso logicamente all'interno dell'altro, ci aiuta a capire che stiamo effettivamente parlando di due cose diverse. Che dunque non vanno confuse, né dal punto di vista logico né dal punto di vista terminologico.

Vorrei aggiungere che tale differenziazione (per me sacrosanta quanto per lo più rifiutata) fra Digitalizzazione e Computazione (o meglio ancora fra ambiente digitale e ambiente computazionale), (a) non coincide con la differenza fra singole procedure, in quanto nella Digitalizzazione può anche essere compresa una parte computazionale, e nella Computazione – come abbiamo appena detto – è sempre compresa una parte digitale; (b) essa dipende dunque dalla considerazione soggettiva dei creatori e degli utenti piuttosto che dalla oggettiva situazione della realtà.

Cercando di essere un po' più esplicito, e rimanendo nell'ambito filologico e linguistico, un primo esempio è dato proprio dalla Digitalizzazione di un testo. Essa si può considerare come la sostituzione di un simbolo con un altro simbolo; ma si dimentica spesso la lezione di Peirce, che un simbolo è

tale solo in presenza di un interpretante: esso cioè mette in moto una serie di competenze senza le quali il simbolo non è più tale ma un'inerte forma grafica che sottoposta a procedimenti automatici darà risultati per lo meno equivoci. Per questo la parola-chiave imprescindibile è, nella nostra procedura, codifica. Questo vale per la singola lettera, le cui diverse forme possono avere diversi valori (significati) e dunque devono essere rappresentate da simboli digitali diversi, quanto per le 'parole', che sono simboli utilizzabili in procedure automatiche soltanto se digitalizzate non solo in base alle loro componenti alfabetiche, ma distinguendo le caratteristiche che esse portano con sé nel testo di cui fanno parte. E vale anche per le altre unità di rango superiore, le frasi, le sezioni (paragrafi, capitoli, ecc.), le pagine (valore dell'impaginazione) e così via. Tutto questo non vale se si vuole fornire all'utente soltanto una rappresentazione brutale del testo nel suo senso più banale; e naturalmente anche questo può essere utile per molti usi, ma non per una procedura computazionale, sia pure una semplice estrazione di elementi (peggio se con conseguenti statistiche), perché verranno giudicati equivalenti fenomeni fra sé simili per certi aspetti, ma diversi per altri. In ambito generale umanistico questo momento analitico che precede logicamente la Computazione assume dunque un'importanza fondamentale, perché richiede un'attenzione e una consapevolezza molto speciali, mentre per noi il momento computazionale vero e proprio è relativamente poco complesso. Ma anche qui, sarebbe opportuno valutare le trovate che vanno di moda in confronto alla tradizione di studi, millenaria, che ci appartiene. La stemmatica genealogistica, le cosiddette 'sentimental analysis' e 'network science', siano pure procedimenti computazionali di pieno diritto, ma mi sembrano scorciatoie puerili per pretendere di superare metodologie tradizionali di gran lunga più serie.

## **2 La nuova biblioteca digitale: caratteristiche, problemi e necessità**

Il secondo problema che vorrei sollevare (altrettanto rapidamente e, ahimé, altrettanto sommariamente del primo) riguarda il rapporto, e la differenza, fra la biblioteca digitale che intendiamo costruire insieme e quella gutenberghiana in cui abbiamo finora operato.

Non mi sembra interessante partecipare al gioco mediatico in ordine alla pretesa 'morte del libro' cartaceo: Umberto Eco ha detto sensatamente che il libro è come il cucchiaino, cioè è un oggetto talmente comodo e funzionale che sopravviverà, magari in posizione marginale, anche nel nuovo assetto (i dati recenti sulla precoce crisi di vendita degli e-book sembrano dargli ragione) ed è comunque vero che l'invenzione dell'automobile non ha comportato automaticamente la fine dei cavalli.

Il problema su cui vorrei richiamare l'attenzione è un altro, e riguarda la crisi (e io penso: la fine) del circuito editoriale che si era costruito a partire dal libro di Gutenberg. Dice un adagio caro agli editori che l'editore è colui che non scrive il libro, non lo stampa, non lo distribuisce, non lo vende, non lo recensisce, ma... fa tutto il resto. Se cerchiamo di analizzare questo 'resto', assolutamente decisivo nell'assetto editoriale gutemberghiano, potremmo dire che l'editore (parlo naturalmente dell'editore degno di questo nome, dell'editore vero e proprio): a) progettava/ sceglieva/ programmava il testo, ad esempio inventandosi le collane (come il nostro Manuzio) o addirittura commissionando la scrittura di determinati libri agli autori; b) 'curava' la pubblicazione, intervenendo direttamente sul testo, attraverso l'attività redazionale (e non c'è bisogno di ricordare qui che ruolo abbiano svolto nella cultura italiana i redattori, dal Bembo delle edizioni 'aldine' fino a Calvino o Bassani); c) garantiva la qualità tipografica, l'accuratezza, l'attendibilità, e a volte anche una peculiare bellezza dell'oggetto libro; d) assicurava una mediazione vitale fra il libro e il mercato dei lettori, promuovendo il libro, curando le presentazioni e/o le recensioni, ecc. e, soprattutto, garantendo in un certo senso la qualità del libro attraverso il prestigio del proprio marchio editoriale; e) garantiva infine la reperibilità del libro, e dunque la sua durata, attraverso il proprio magazzino ma soprattutto attraverso il proprio catalogo, che rappresentava la vera gloria di un editore di qualità.

Era l'insieme di queste attività, come si vede assolutamente cruciali, ciò che giustificava e legittimava il profitto dell'editore.

Ora, se esaminiamo il nuovo circuito editoriale che si sta determinato a partire dai prodotti testuali informatizzati, ci rendiamo conto che nessuna di queste funzioni viene più svolta dagli editori (le eccezioni o le tenaci sopravvivenze non interessano questo nostro sommario discorso): a) quasi nessun editore può più progettare, investendo per il medio-lungo periodo, perché le leggi del mercato non glielo consentono; b) le redazioni, il vero centro produttivo delle case editrici di una volta, sono state le prime a essere sciolte (o esternalizzate, che è lo stesso) per risparmiare sui costi; c) della qualità tipografica è meglio non parlare, e non sono poche le case editrici che pretendono dagli autori il libro già impaginato, in PDF, completo di titoli correnti e indici, limitandosi a stamparlo e a fornirlo di ISBN; d) non esiste praticamente più alcuna mediazione di tipo culturale fra libro e lettori (evidentemente l'eventuale pubblicità è tutt'altra cosa), anzi sono sempre più frequenti i casi in cui il libro è pagato, direttamente o indirettamente, dal suo stesso autore, magari usando per questo i sempre più esigui fondi destinati alla ricerca; e) infine - forse questa è per noi la circostanza più negativa - il catalogo non esiste quasi più, la durata del libro è una chimera, i libri durano lo spazio di un mattino e il macero li attende; questa rapida irrimediabilità riguarda addirittura l'opera omnia di alcuni classici (cito per tutti la fondamentale serie delle *Opere complete*

di Boccaccio, curata da Branca per Mondadori, un'edizione che doveva essere per sua natura *für ewig* e che risulta invece assolutamente introvabile già a pochi anni dalla sua pubblicazione).

Tutto ciò, e in particolare il per noi cruciale punto e), cioè la mancata garanzia della durata, risulta enfatizzato in sommo grado per le opere in formato digitale; anzi neanche l'acquisto garantisce il possesso nel tempo del libro digitale, se è vero che spesso se ne compra solo il diritto alla lettura che, in quanto tale, non può essere neanche trasmesso ai nostri eredi.

Per non parlare del problema decisivo rappresentato dalla obsolescenza dei testi digitali, che riguarda non tanto i loro supporti quanto i programmi e le macchine necessari per leggerli. Faccio notare che la biblioteca di pubblica lettura nel vecchio assetto svolgeva un ruolo correttivo del mercato (appunto perché garantiva la lettura gratuita e la conservazione nel tempo del libro), ma essa non è sostituita nel nuovo assetto da alcuna istituzione paragonabile. Già ora, negli scaffali di ciascuno di noi, esistono vecchi floppy disk che risultano inaccessibili ai nuovi computer, determinando una perdita secca di informazioni testuali che nel tempo potrebbe assumere dimensioni catastrofiche. Nel mondo dell'informatica e dei suoi programmi sempre mutanti, è come se si verificasse ogni pochi anni una modificazione della tecnologia testuale paragonabile a quella che segnò il passaggio dal rotolo al volume o dal manoscritto alla stampa. E ogni volta - come sappiamo bene - queste modificazioni tecnologiche della testualità comportarono, assieme alla necessità del 'riversamento' dei testi da una tecnologia all'altra, anche perdite testuali enormi e irreparabili. Ma soprattutto, nel nuovo circuito editoriale segnato dall'informatica, l'editore dove sta? In cosa consiste la sua utilità? Cosa legittima la sua pretesa di trarre profitto dai libri?

Credo, in conclusione, che sia necessario immaginare e implementare al più presto un originale sistema di circolazione e fruizione del libro elettronico che tenga conto delle sue novità e delle sue specifiche caratteristiche e potenzialità, insomma un circuito che non imiti più il vecchio assetto gutenberghiano ma che si svolga *iuxta propria principia*, cioè che sia in grado di garantire di fatto ciò che l'informatica (in questo caso la rete) è per sua natura in grado di consentire, ma dunque in qualche modo pretende che sia realizzato

Mi sembra che ciò significhi anzitutto garantire l'accesso universale e dunque in primo luogo la gratuità, considerando la cultura per quello che effettivamente è, cioè come un bene comune. D'altra parte la cultura è forse l'unico bene che si incrementa con l'uso collettivo e per il quale non vale affatto il principio di esclusività (se una cosa è tua non è più mia, e viceversa), il principio che valeva per i campi e per il bestiame, anzi in questo caso il fatto che molti fruiscano della mia cultura mi arricchisce e non mi impoverisce, esattamente come mi arricchisce fruire della cultura altrui. Questo chiama in causa la mano pubblica, e la necessità di specifiche iniziative e di adeguati finanziamenti pubblici.

Le Università pubbliche, che già oggi ospitano, in modo generoso ma non sistematico e non sostenuto, le nostre biblioteche digitali e i prodotti digitali delle nostre ricerche, devono essere il luogo di questa nuova biblioteca di Alessandria informatica, universale, diffusa e accessibile a tutti.

Mi rendo conto che si tratta di un'affermazione controcorrente in un'epoca che pretende di fare del mercato e delle sue pretese la nuova Legge universale e totalitaria. Ma se è vero - come è stato detto da chi se ne intendeva - che la guerra è una cosa troppo seria per lasciarla fare ai generali, credo che sia ancora più vero che la cultura digitale è una cosa troppo seria per lasciarla nelle mani della spontaneità capitalistica, del mercato e degli interessi delle grandi imprese multinazionali dell'informatica.

