



**Citation:** A. Baldi (2019) Teoria e forme del testo digitale, a cura di Michelangelo Zaccarello, postfazione di H. Wayne Storey, Roma, Carocci editore, 2019, pp. 229. *Lea* 8: pp. 495-500. doi: <https://doi.org/10.13128/LEA-1824-484x-11001>.

**Copyright:** © 2019 A. Baldi. This is an open access, peer-reviewed article published by Firenze University Press (<https://oajournals.fupress.net/index.php/bsfm-lea>) and distributed under the terms of the Creative Commons Attribution – Non Commercial – No derivatives 4.0 International License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited as specified by the author or licensor, that is not used for commercial purposes and no modifications or adaptations are made.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Competing Interests:** The Author(s) declare(s) no conflict of interest.

## Recensione a *Teoria e forme del testo digitale*, a cura di Michelangelo Zaccarello, postfazione di H. Wayne Storey, Roma, Carocci editore, 2019, pp. 229

*Alberto Baldi*

Università degli Studi di Firenze (<[alberto.baldi@unifi.it](mailto:alberto.baldi@unifi.it)>)

### *Abstract*

The collection of essays edited by Michelangelo Zaccarello covers some important aspects of digital textuality and how it has impacted the production, distribution and study of literary texts. Areas covered include: the characteristics of electronic texts, encoding standards, as well as new writing and reading modalities, the quality and conservation of digital libraries, mass digitalization procedures and implications for author's rights.

*Keywords:* mass digitalization, ebook, Information Quality, OCR, digital text

“Di testualità digitale si parla ormai da una quarantina d’anni”, scriveva Roncaglia già nel 2012 (429), a dimostrazione di come i nuovi supporti informatici abbiano da subito acceso il dibattito attorno alla consolidata nozione di testualità. Il testo è stato infatti svincolato dalla materialità cartacea e questo, al netto delle “discussioni abbastanza sterili legate a preferenze e idiosincrasie individuali” (*ibidem*), ha favorito una duplice riflessione sui possibili vantaggi o svantaggi tecnici ed epistemologici derivanti dalla progressiva smaterializzazione. Per quanto il tema dell’informatizzazione del testo abbia natura multidisciplinare – interessando, oltre all’informatica, la sociologia della lettura, la pedagogia, le scienze della comunicazione, le neuroscienze, la biblioteconomia – è nell’ambito degli studi filologico-letterari, laddove i testi, come ha giustamente notato Fabio Ciotti, “sono tanto i principali oggetti quanto i principali strumenti di studio” (2011, 63), che la questione ha assunto carattere di urgenza.

La testualità digitale rappresenta da sempre uno dei principali argomenti delle Digital Humanities<sup>1</sup>. Alle riflessioni teoriche sull'evoluzione del concetto di testo – spesso a partire dalla celebre definizione data da Cesare Segre di testo come “tessuto linguistico del discorso”<sup>2</sup> – nelle discipline letterarie è andato di pari passo un vertiginoso incremento delle soluzioni tecniche a disposizione degli studiosi. Per far luce sulla complessità del tema è opportuno rifarsi a una tripartizione proposta da Celentano, Cortesi, Mastandrea, in cui “l’informatica testuale” viene suddivisa in “documentaria, editoriale ed ermeneutica”: “La prima si occupa della creazione di banche dati e sistemi di interrogazione, la seconda della rappresentazione dei testi a stampa, a video o su altri media, la terza, invece, si dedica alla relazione fra il testo e le sue interpretazioni” (2004, 48).

Il ruolo dell’informatica negli studi letterari differisce quindi a seconda che il lavoro su e per i testi abbia finalità archivistico-documentarie, editoriali (con uno sguardo prioritario all’eccdotica, alla variantistica digitale<sup>3</sup>, alle prospettive di una “filologia collaborativa e cooperativa”<sup>4</sup>), o ermeneutiche<sup>5</sup>. Per quanto efficace, si tratta, com’è ovvio, di una distinzione meramente teorica – giacché i vari approcci informatici agli studi letterari sono tra sé fortemente interrelati –, ma utile, tuttavia, a fare chiarezza all’interno di una disciplina in continua evoluzione.

Soprattutto per i primi dei due ambiti individuati, l’antologia curata da Michelangelo Zaccarello appare oggi come un esaustivo e prezioso ragguaglio sui principali temi di riflessione sviluppatasi attorno al testo digitale, raccogliendo dieci contributi di studiosi angloamericani. Sin dalla introduzione del curatore (“La svolta digitale nelle pratiche di lettura e di studio del testo digitale”, 13-31), si anticipano i punti della *editorial theory* che saranno toccati nel volume: “esempi di ricerca su forme e aspetti meno noti della testualità digitale” (Zaccarello 2019, 30), volti a metterne in luce le indubbie problematicità ma anche a evidenziare le nuove possibilità che ne derivano.

Tra i nomi antologizzati, troviamo alcune delle voci più note nel panorama delle Digital Humanities, come Jerome McGann, autore di uno dei capisaldi sul testo nell’epoca del digitale (McGann 2001), che apre e chiude il volume con i suoi “Il testo ‘sociale’ fra volontà d’autore e pubblico” (33-42) e “Ritorno alla filologia. La memoria del passato nel contesto digitale” (197-207). Nel primo dei due saggi, risalente al 1983, McGann esplora le differenze metodologiche tra un editore di testi classici e un editore di testi moderni: un contributo solo in apparenza avulso dal tema del volume e che invece appare assai puntuale, anche alla luce dei nuovi strumenti informatici, nel problematizzare il “paradigma editoriale della volontà [d’autore]” (McGann in Zaccarello 2019, 41) e nel riconsiderare il ruolo della “socializzazione” del testo ai fini della sua definizione. L’importanza di questa proposta e la sua attualità nel contesto digitale è del resto già confermata in “Il contesto ‘collaborativo’ degli studi letterari e la dimensione ‘sociale’ delle edizioni scientifiche” di Peter Robinson (115-134), che proprio dagli studi di McGann e da quelli di Donald McKenzie muove per riflettere sulle prospettive dell’eccdotica sociale. L’idea di edizione “sociale”, frutto cioè della cooperazione e dello scambio *online* all’interno di una comunità di collaboratori, che Robinson riconduce alla proposta di Siemens, Timney, Leitch *et al.* 2012, è qui ridimensionata alla luce del fatto che la costituzione di un testo definitivo

<sup>1</sup> A titolo d’esempio e limitatamente all’area italiana, si segnalano Gigliozzi 1997; Fiormonte 2003; Ciotti 2007; Tomasi 2008, in particolare capp. 4 e 5; Fiormonte, Numerico, Tomasi 2010, in particolare capp. 2 e 3; Orlandi 2010; Roncaglia 2010; Perilli, Fiormonte 2011; Anichini 2013 [2010]; Fiormonte 2018; Stella 2018, in particolare cap. 2.

<sup>2</sup> Segre 1981, 269; ma si veda anche Segre 1979 (in particolare 23-37), e, in proposito, Mordenti 2007, 129-166.

<sup>3</sup> Cfr. ancora Fiormonte 2003 e il suo pionieristico progetto *Digital Variants*, <digitalvariants.org> (11/2019).

<sup>4</sup> Cfr. l’“Introduzione” in Boschetti 2018.

<sup>5</sup> A tal proposito si rimanda a Ciotti 2017 e ai capp. 4-5 di Stella 2018.

dipenda pur sempre dall'autorità del singolo, limitando la componente "sociale" alla fase di post-pubblicazione e allo "sviluppo di un dibattito continuo, fatto di commenti e discussioni, agevolate e incoraggiate dai social media" (Robinson in Zaccarello 2019, 133).

Susan Hockey, di cui è riproposto "Le nuove dinamiche di edizione e lettura nell'era digitale" (43-55), pubblicato nel 2003 ma per certi aspetti ancora attuale, oltre a figurare tra i precursori della disciplina (autrice di un manuale già all'inizio degli anni Ottanta, cfr. Hockey 1980), è stata tra i primi a storicizzare le Digital Humanities (si veda Hockey 2004). In questo suo saggio, consapevole delle insanabili differenze tra il *medium* cartaceo e quello digitale, suggerisce di rinunciare al tentativo di tradurre in digitale le modalità di produzione del libro a stampa, sfruttando invece le peculiarità delle edizioni elettroniche per sviluppare nuovi modelli di organizzazione delle informazioni. Le caratteristiche costitutive dei supporti digitali devono essere massimizzate, senza correre il rischio di annullarle cercando un compromesso con l'alternativa cartacea. Essenziale, in tal senso, la codifica dei testi<sup>6</sup>, che "aggiunge intelligenza al testo" (Hockey in Zaccarello 2019, 46) e che costituisce la base per la definizione di quelle "strutture intellettuali" che la Hockey indica come obiettivo primario per il curatore di un'edizione digitale, per garantire all'utente "rotte di navigazione ottimali" e non limitarsi alla "semplice presentazione di un ammasso di materiali" (ivi, 54). Codificare un testo in funzione di un'edizione elettronica ne comporta, quindi, un'interpretazione.

Analogamente, i saggi di Paul Eggert, "L'edizione come 'cantiere' testuale: conservazione, edizione, restauro" (57-70), e di Peter Shillingsburg, "La svolta digitale e lo studio della letteratura nel nuovo contesto (71-84), si soffermano sull'importanza di definire una corretta prassi nella strutturazione dei contenuti di un'edizione digitale. L'idea di "ordered hierarchy of content objects" (DeRose, Durand, Mylonas, *et al.* 1990, 6) è ripresa da Eggert come base per una definizione di testo che sia valida a prescindere dal supporto. Anzi, su tale assunto, è possibile postulare l'esistenza di una materialità anche nel testo elettronico, ossia la sua "completa determinabilità computazionale" (Eggert in Zaccarello 2019, 62). Per ottimizzare le possibilità del libro elettronico, Eggert propone pertanto – di pari passo allo sviluppo di un sistema di codifica *stand-off* che annoti il testo contestualmente alle interrogazioni dell'utente (*Just-in-Time-Markup*) – il modello di "cantiere testuale", un'edizione digitale in cui "il lettore potrà accettare la guida dell'editore [...], il percorso raccomandato attraverso il cantiere testuale" oppure "respingere la guida dell'editore, cercare percorsi inusitati" (ivi, 69). Shillingsburg, invece, consapevole che "gli standard di accuratezza e precisione nella rappresentazione dei testi letterari in quanto arte sono diversi dagli standard che vengono spesso adottati per documenti culturali o corpora testuali per l'analisi linguistica" (Shillingsburg in Zaccarello 2019, 71), evidenzia come i vari livelli da cui è costituita un'edizione digitale – digitalizzazioni, trascrizioni, marcatura e annotazione del testo – siano tra loro in un rapporto di interdipendenza e concorrano in ugual misura alla qualità del prodotto finale.

"La letteratura *born digital*: gli scrittori e il Word Processor" (85-94) di Matthew Kirschenbaum, si sofferma invece sul rapporto di alcuni autori con la videoscrittura – da George R.R. Martin, che utilizza un *software* ormai superato (Wordstar) installato su un computer rigorosamente *offline*, a Don DeLillo e Cormac McCarthy, che invece non hanno mai sostituito le loro macchine da scrivere con un pc –, oltre a segnalare l'urgenza, per i filologi e gli studiosi

<sup>6</sup>La Hockey, come standard di riferimento, indicava già nel 2003 le *Linee guida* proposte dal consorzio TEI (Text Encoding Initiative). L'iniziativa, nata nel 1994, ha nel tempo definito un modello di codifica dei testi per le discipline umanistiche, la linguistica e le scienze sociali. L'ultima versione dello *standard* è stata rilasciata nel giugno 2019 ed è consultabile all'indirizzo <<https://tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf>> (11/2019).

di letteratura, di una più scrupolosa riflessione sulle prassi di scrittura connaturate al mezzo informatico, analogamente a quanto è stato fatto per la scrittura manuale e il libro a stampa. Il tema della scrittura digitale negli autori contemporanei ha aperto nuove prospettive di studio anche in campo archivistico, evidenziando la necessità di creare dei sistemi di conservazione per le “carte” d’autore digitali. Da qui sono scaturiti progetti come PARADIGM<sup>7</sup>, un protocollo per la corretta conservazione dei testi *born digital* sviluppato nel 2005 dalle università di Cambridge e Manchester, o PAD – Pavia Archivi Digitali – dell’Università di Pavia, che dal 2009 si occupa di preservare le memorie digitali di autori contemporanei<sup>8</sup>.

Tra gli ultimi saggi l’attenzione si sposta soprattutto sulle problematiche relative alla digitalizzazione di massa. Contestualmente all’inizio delle grandi imprese di digitalizzazione del patrimonio librario – su tutte, ovviamente, Google Books, la più estesa e la più controversa – l’attenzione degli studiosi<sup>9</sup> si è rivolta alle innumerevoli criticità che ne sarebbero derivate. In parallelo alle questioni di natura legale e ai dubbi circa il ruolo monopolistico giocato da Google, è nata l’urgenza di una riflessione sugli standard qualitativi delle digitalizzazioni, oltre che sulla necessità di un’attenta metadateazione dei libri digitali. Lo stesso Zaccarello, nella sua introduzione, sottolinea come i prodotti dell’editoria accademica *born digital*, che, originariamente concepiti per il supporto digitale, sfruttano le potenzialità offerte da una modalità di lettura non lineare e da una capacità di archiviazione praticamente illimitata, costituiscano una minima parte del patrimonio librario reperibile *online*, giacché “la stragrande maggioranza dei testi che circolano in rete non sono *born digital* ma digitalizzati” (Zaccarello 2019, 14). In “Dal cartaceo al ‘digitale di massa’: biblioteche virtuali, diritto d’autore e il caso Google Books” (95-114) di Maurizio Borghi e Stavroula Karapapa si ripercorrono le varie iniziative di digitalizzazione intraprese, commerciali (Google Books) o no-profit (Internet Archive, Open Library, Hathitrust...), e le questioni sulla tutela del copyright che ne sono scaturite, in conseguenza soprattutto del *modus operandi* di Google che a “un sistema di licenze e permessi *ex ante*” ha preferito stipulare accordi con le grandi biblioteche garantendo agli autori “un regime opzionale di autoesclusione” *a posteriori*. Nel contributo di Diana Kichuk, “Quantità e qualità dei testi online: il problema della digitalizzazione di massa” (135-167), il *focus* è rivolto alla valutazione della qualità dei testi *online* prodotti con operazioni di scansione semi-automatica e di acquisizione mediante tecnologie di OCR (*Optical Character Recognition*). “Molti archivi digitali” – nota la Kichuk – “caricano in rete il testo OCR grezzo” (Kichuk in Zaccarello 2019, 153) etichettandolo come “Full Text”: per questa tipologia di prodotti – in cui si ravvisa una “potenziale utilità per la ricerca o come base per creare un e-book”, nel saggio si propone di introdurre l’inequivocabile dizione di “Raw Text”, predisponendo invece un modello di rielaborazione per quei testi che, prodotti con OCR, siano offerti all’utenza per la lettura “in senso tradizionale” (ivi, 163). Ed è proprio dalla volontà di migliorare la qualità delle risorse testuali disponibili in rete che muove Paul Conway nel suo “La conservazione nel contesto digitale: quale protocollo di qualità per le biblioteche virtuali?” (169-196), riadattando un modello predisposto dallo studioso di Information Quality Besiki Stvilia (2007) su un *corpora* di digitalizzazioni (3000 volumi in totale, da Google Books e Internet Archive), al fine di stabilire un modello di errore per i prodotti digitalizzati e processati con OCR.

<sup>7</sup> <<http://www.paradigm.ac.uk/>> (11/2019).

<sup>8</sup> <<http://pad.unipv.it/storia/>> (11/2019); ma cfr. anche Giordano 2018.

<sup>9</sup> Cfr. almeno Roncaglia 2009, Darnton 2011 e, per lo stato dell’arte del progetto Google Books in Italia, De Pasquale 2019.

Il volume si chiude con un secondo saggio di McGann – un invito al recupero, o, quantomeno, alla conservazione del metodo filologico, considerando l'essere umano come elemento imprescindibile per gestire le risorse informatiche, tanto più "intelligenti" quanto più a rischio ingovernabilità – e con una postfazione di H. Wayne Storey (209-216), che nel digitale non vede un rischio per la prassi filologica quanto, semplicemente, un mutamento di paradigma dettato dal *medium* di riferimento: dal *print thinking* al *digital thinking* a cambiare non saranno i risultati del lavoro filologico, ma "i modi di visualizzarli e di integrarli in un corpo di informazione forse sorprendentemente a servizio di non filologi" (Wayne Storey in Zaccarello 2019, 215). Il digitale, dunque, non come un approccio che snaturi, svalutandolo, il lavoro filologico, ma come un nuovo, ulteriore strumento per il quale raffinare e veicolare i prodotti della ricerca umanistico-letteraria.

Nel complesso, il volume di Zaccarello riesce a coniugare con efficacia la scelta di offrire un'eterogenea serie di spunti sulla tematica della testualità digitale con il tentativo di inquadrare il fenomeno in una prospettiva diacronica, ricomprendendolo all'interno della più ampia riflessione sul rapporto tra l'informatizzazione e gli studi letterari.

#### Riferimenti bibliografici

- Anichini Alessandra (2013 [2010]), *Il testo digitale. Leggere e scrivere nell'epoca dei nuovi media*, Milano-Sant'Arcangelo di Romagna, Apogeo-Maggioli.
- Boschetti Federico (2018), *Copisti digitali e filologi computazionali*, Roma, CNR edizioni, <<https://dspace-clarin-it.ilc.cnr.it/repository/xmlui/bitstream/handle/20.500.11752/OPEN-89/dscs.pdf?sequence=1&isAllowed=y>> (11/2019).
- Carbé Emmanuela (2018), "Digitale d'autore. Un archivio possibile?", in Virna Brigatti, Anna Lisa Cavazzuti, Elisa Marazzi, *Archivi editoriali. Tra storia del testo e storia del libro*, Milano, Unicopli.
- Celentano Augusto, Cortesi Agostino, Mastandrea Paolo (2004), "Informatica umanistica: una disciplina di confine", *Mondo digitale* 4, 44-55.
- Ciotti Fabio (2007), *Il testo e l'automa. Saggi di teoria e critica computazionale dei testi letterari*, Roma, Aracne.
- (2011), "La rappresentazione digitale del testo: il paradigma del markup e i suoi sviluppi", in Perilli, Fiormonte 2011, 63-90.
- (2017), "Modelli e metodi computazionali per la critica letteraria: lo stato dell'arte", in B. Alfonzetti, T. Cancro, V. Di Iasio, E. Pietrobon (a cura di), *L'Italianistica oggi. Ricerca e didattica*. Atti del XIX Congresso dell'ADI – Associazione degli Italianisti (Roma, 9-12 settembre 2015), Roma, Adi editore, 1-11, <[http://www.italianisti.it/Atti-di-Congresso?pg=cms&text=p&cms\\_codsec=14&cms\\_cod-cms=896](http://www.italianisti.it/Atti-di-Congresso?pg=cms&text=p&cms_codsec=14&cms_cod-cms=896)> (11/2019).
- Darnton Robert (2011), *Il futuro del libro*, traduzione di Adriana Bottini, Milano Adelphi (ed. orig. *The Case for Books. Past, Present, and Future*, New York, NY Public Affairs, 2009).
- De Pasquale Andrea (2019), "L'attuazione in Italia del Progetto GoogleBooks", *Digitalia* XIV, 1, 103-113.
- DeRose S.J., Durand D.G., Mylonas Elli, et al. (1990), "What is Text, Really?", *Journal of Computing in Higher Education* I, 2, 3-26.
- Fiormonte Domenico (2003), *Scrittura e filologia nell'era digitale*, Torino, Bollati Boringhieri.
- (2018), *Per una critica del testo digitale. Letteratura, filologia e rete: informatica e discipline umanistiche*, Roma, Bulzoni.
- Fiormonte Domenico, Numerico Teresa, Tomasi Francesca (2010), *L'umanista digitale*, Bologna, Il Mulino.
- Gigliozzi Giuseppe (1997), *Il testo e il computer. Manuale di informatica per gli studi letterari*, Milano, Mondadori.
- Giordano Tommaso (2018), "Quando le carte diventano bit", *Biblioteche oggi* 36, 3-8, doi: 10.3302/0392-8586-201804-003-1.

- Hockey Susan (1980), *A Guide to Computer Applications in the Humanities*, London-Baltimore, Duckworth-Johns Hopkins, 1980.
- (2004), “The History of Humanities Computing”, in Susan Schreibman, R.G. Siemens, John Unsworth (ed.), *A Companion to Digital Humanities*, Malden, Blackwell, <<http://www.digitalhumanities.org/companion/>> (11/2019).
- McGann J.J. (2001), *Radiant Textuality. Literature after the World Wide Web*, New York, Palgrave.
- Mordenti Raul (2007), *L'altra critica. La nuova critica della letteratura fra studi culturali, didattica e informatica*, Roma, Meltemi editore.
- Orlandi Tito (2010), *Informatica testuale. Teoria e prassi*, Bari, Laterza.
- Perilli Lorenzo, Fiorimonte Domenico (2011), *La macchina del tempo. Studi di informatica umanistica in onore di Tito Orlandi*, Firenze, Le Lettere.
- Roncaglia Gino (2009), “Google Book Search e le politiche di digitalizzazione libraria”, *Digitalia* IV, 2, 17-35.
- (2010), *La quarta rivoluzione. Sei lezioni sul futuro del libro*, Roma-Bari, Laterza.
- (2012), “Testualità digitale e forme dell'argomentazione”, *Quaestio* 11, 429-440.
- TEI Consortium (2019), *TEI P5. Guidelines for Electronic Text Encoding and Interchange*, <<https://tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf>> (11/2019).
- Tomasi Francesca (2008), *Metodologie informatiche e discipline umanistiche*, prefazione di Dino Buzzetti, Roma, Carocci.
- Segre Cesare (1979), *Semiotica filologica. Testi e modelli culturali*, Torino, Einaudi.
- (1981), “Testo”, in *Enciclopedia*, vol. XIV, *Temalmotivo-Zero*, Torino, Einaudi, 269-291.
- Siemens Ray, Timney Meagan, Leitch Cara, et al. (2012), “Toward Modeling the Social Edition: An Approach to Understanding the Electronic Scholarly Edition in the Context of New and Emerging Social Media”, *Literary and Linguistic Computing* XXVII, 4, 445-461.
- Stella Francesco (2018), *Testi letterari e analisi digitale*, Roma, Carocci editore.
- Zaccarello Michelangelo, a cura di (2019), *Teoria e forme del testo digitale*, postfazione di H. Wayne Storey, Roma, Carocci editore.