

Recensione: E. Carbé. 2023. *Digitale d'autore. Macchine, archivi, letterature*. Firenze: Firenze University Press; Siena: USiena Press

Marco Sartor

Università di Parma
marco.sartor@unipr.it

Abstract

Come favorire la conservazione e lo studio di un archivio letterario nativo digitale? Nei tre capitoli che compongono *Digitale d'autore. Macchine, archivi, letterature*, la questione è affrontata da molteplici punti di vista. Nel primo di essi Emmanuela Carbé, dopo aver definito l'oggetto in esame, traccia una presentazione degli aspetti storici, teorici e metodologici, entro cui trova spazio una ricognizione sul rapporto tra scrittori e computer, che cede il passo nel successivo a una rassegna delle principali esperienze nel panorama nazionale e internazionale. L'ultima sezione è invece dedicata all'inquadramento critico di alcune opere a partire dallo studio del materiale archivistico *born digital* conferito dall'autore ad un'istituzione preposta alla sua conservazione.

Parole chiave: digitale d'autore, archivi nativi digitali, archivi letterari, conservazione del digitale, letteratura contemporanea

*How to foster the preservation and study of a born-digital literary archive? In the three chapters that comprise *Digitale d'autore. Macchine, archivi, letterature* the question is tackled from multiple points of view. In the first, after defining the object of investigation, Emmanuela Carbé outlines the historical, theoretical, and methodological aspects, including a definition of the relationship between writers and computers. This gives way in the central part of the volume to a review of the main experiences on both the national and international stage. Finally, the last section is dedicated to the critical framing of a work by investigating the born-digital archive conferred by the author to the institution in charge of its preservation.*

Keywords: digitale d'autore, born-digital archives, literary archives, digital preservation, contemporary literature

Obiettivi e contesto

L'espressione "digitale d'autore" è stata coniata da Emmanuela Carbé nel 2017 per rendere l'inglese *born-digital (literary) archives*. Con essa l'autrice indica un campo d'indagine – quello del

trattamento degli archivi letterari nativi digitali – di recente fondazione, e per questo costituito da territori la cui mappatura è ancora in corso. Lo stadio aurorale di questo fascio di studi si giustifica agevolmente considerando che la diffusione massiva dei personal computer avvenne solo nell'ultimo decennio del XX secolo: un tempo che sembra distare ere geologiche nella vita degli *information studies*, ma che diviene quasi insignificante nella più dilatata scala della trasmissione del sapere, le cui progressioni molto risentono del passaggio generazionale. In questi anni l'uso estensivo del computer investe anche la produzione intellettuale, dopo un'impasse che ha caratterizzato gran parte del decennio precedente. Di quest'ultima rende testimonianza l'indagine di Luciano Curino affidata alle pagine di «Tuttolibri» del 5 giugno 1983, da cui si evince che all'epoca il computer veniva impiegato solo sporadicamente nella scrittura creativa, complice la radicata convinzione che i benefici riguardassero soprattutto le funzioni di calcolo e l'espletamento di formalità burocratico-amministrative ([1]).

Queste coordinate essenziali aiutano a comprendere le ragioni per cui le prime sistematiche riflessioni sul digitale d'autore siano sorte all'incirca due decenni dopo, e segnatamente sul finire della prima decade del nuovo millennio, per lo più in ambienti statunitensi. Si fa riferimento, in particolare, alle attività del gruppo coordinato da Matthew Kirschenbaum e finanziato dal National Endowment for the Humanities (NEH), che nel 2009 pubblicò un primo report sul digitale d'autore compilato di concerto con le principali istituzioni che preservavano materiali *born digital* in aggiunta ai tradizionali “scartafacci” (0). Il tema fu oggetto di un ulteriore approfondimento nella relazione presentata nei mesi successivi alla conferenza annuale iPRES (International Conference on Digital Preservation) e, l'anno seguente, nel convegno *Computer Forensics and Born-Digital Content in Cultural Heritage Collections* (University of Maryland). Quest'ultimo, poi, spianò la via al progetto BitCurator sui problemi della preservazione del digitale, che dal 2014 ha assunto la forma di consorzio.¹ Ma il 2010 vide il proliferare di altre capitali iniziative: fra esse, si ritaglia uno spazio a sé stante *Defining “Born Digital”* di Ricky Erway ([2]), il primo di una lunga serie di *white paper* dell'Online Computer Library Center (OLCL) dedicati alle collezioni nate in formato digitale.² Ben presto l'ambito d'indagine valicò i confini nazionali coinvolgendo le principali istituzioni estere, che talvolta avevano già avviato autonomamente (e precedentemente) alcune riflessioni al riguardo.³

Entro questo solco – qui in realtà più abbozzato che ritratto – si situa il volume di Emmanuela Carbé, il cui respiro internazionale non collide con lo sguardo al contesto italiano cui è dedicata la quasi totalità delle pagine. Infatti, se è vero che questa ricerca è un imprescindibile viatico per gli studiosi di letteratura italiana ultra-contemporanea e per coloro che a vario titolo operano con archivi (letterari) nativi digitali, ad essa va altresì riconosciuto il merito di sanare un'aporia. Descrivendo per la prima volta in forma sistematica le esperienze in corso nella penisola italiana con il rigore proprio di chi vi ha preso parte personalmente, la sua diffusione al di fuori del contesto nazionale è auspicata per favorire il riconoscimento estero di alcuni progetti

¹ <https://bitcurator.net>.

² I contributi del gruppo di lavoro sono consultabili all'indirizzo: <https://www.oclc.org/research/areas/research-collections/borndigital.html>.

³ Nel Regno Unito il progetto Paradigm (Personal Archives Accessible in Digital Media) venne avviato dalle biblioteche delle Università di Oxford e Manchester nel 2005, mentre già dal 2000 la British Library incentivò l'utilizzo del termine eMANUSCRIPTS (eMSS) per le acquisizioni *born digital*. Nel contesto italiano, le attività del Centro Interdipartimentale di Ricerca su Tradizione e Traduzione (CIRTI) dell'Università di Cassino (ora non più attivo) presero avvio nel lontano 2003.

particolarmente promettenti – com'è il caso di Pavia Archivi Digitali (PAD) –⁴ e assicurare che la sempre più impellente individuazione di standard e *best practices* tenga conto (anche) delle principali acquisizioni raggiunte in Italia nell'ultimo quindicennio.

Più specificamente, lo studio di Carbé persegue un triplice fascio di obiettivi. In primo luogo, ambisce a fornire una messa a punto di carattere definitorio, stabilendo i fondamenti e la terminologia dell'ambito d'indagine sulla scorta degli studi di pionieri come Matthew Kirschenbaum. Si tratta di un'operazione essenziale, funzionale allo sviluppo di ogni discussione futura e, non da ultimo, alla ricezione di nuovi conferimenti da parte di scrittori. Nel tracciare uno stato dell'arte con una mappatura delle principali esperienze in ambito nazionale e internazionale con focus sull'esperienza pavese, poi, l'autrice pone le basi per una comparazione delle diverse soluzioni adottate, da cui è possibile inferire pratiche comuni e criticità ricorrenti. Da ultimo, la disamina sull'archivio di Francesco Pecoraro (Roma, 1945-vivente) illustra le potenzialità offerte dal digitale d'autore per l'inquadramento storico-critico di un'opera letteraria a partire dallo studio dei suoi materiali genetici.

Articolazione, contenuto e metodologie

Aperto da un'introduzione dove si definisce il concetto di digitale d'autore e chiuso da una postfazione rivolta al futuro, il volume si compone di tre capitoli dedicati rispettivamente al rapporto fra scrittori e computer (pp. 13-23) con annessa panoramica sulle esperienze nazionali e internazionali (pp. 32-49), alla dettagliata illustrazione del progetto Pavia Archivi Digitali (PAD) con attenzione ai principali sviluppi e conferimenti (pp. 51-111) e all'analisi critica di tre opere di Francesco Pecoraro condotta a partire dal vaglio del fondo archivistico conservato presso il centro pavese (pp. 113-159).

La studiosa definisce il digitale d'autore come

un'entità o un'aggregazione di entità istanziate su supporto digitale da un'autrice o da un autore, e/o da altri soggetti produttori in interrelazione con l'autrice o l'autore all'interno di un determinato contesto; le entità, così come le eventuali relazioni tra entità, possono costituire gli *eventi* dell'archivio (p. 10).

Già da queste righe si evince la vastità dei materiali accolti entro quest'alveo, come pure il loro carattere eterogeneo: il digitale d'autore, infatti, comprende a pieno titolo i testi e gli avantesti di un'opera letteraria, ma anche le fotografie prodotte dall'autore, le bozze di stampa corrette da quest'ultimo, i suoi interventi in rete (ad esempio i post pubblicati sui social network), le pagine web salvate in locale, e così via. Ne consegue che interrogarsi su cosa preservare e con quali modalità assume un carattere non più eludibile. In realtà, domande siffatte emersero già nel gruppo di lavoro coordinato da Kirschenbaum e vennero discusse nella citata conferenza iPRES del 2009, ma la loro riproposizione in questa sede non è vana se accompagnata dal rilievo di alcune criticità rilevate nelle sperimentazioni del progetto PAD.

⁴ <http://pad.unipv.it>. Nel 2021 il progetto PAD, che ha avuto tra i suoi coordinatori Paul Gabriele Weston, è entrato a far parte del Centro Manoscritti di Giuseppe Antonelli (<http://centromanoscritti.unipv.it>), da cui ha ereditato anche la nuova denominazione (cfr. p. 107).

Sul primo versante (quali materiali conservare) le riflessioni si dispongono lungo due direttrici principali. Il piano più astratto coincide con la definizione dei criteri per la selezione degli autori i cui fondi possono essere accolti dall'istituzione conservatrice. Nell'ambito dell'esperienza pavese, questo argomento è stato affrontato nelle discussioni condotte nel biennio 2013-2014, che portarono a rispondere a interrogativi come: «che cosa andava salvato del nostro tempo? E a quale scopo?» (p. 68). Sono, queste, considerazioni in qualche modo legate alla costituzione del canone, destinate ad acquisire ulteriore rilievo con l'aumento dei conferimenti auspicato nell'immediato futuro. Altrettanto meritevole di attenzione è l'elenco delle tipologie di file ammessi, la cui gestione può farsi delicata se è prevista l'inclusione di documentazione proveniente da profili personali Facebook e Twitter e da scambi e-mail.⁵ Proprio quest'ultimo aspetto consente di spostare l'analisi su un asse più concreto e, in particolare, di rilevare come di norma il materiale conferito dall'autore corrisponda solo a una parte di quello effettivamente posseduto, e ciò non soltanto perché nella maggioranza dei casi lo scrittore è ancora in attività (cfr. p. 69). La stessa selezione autoriale – condotta con tassi variabili di accuratezza e arbitrio – comprende talora documenti trasferiti per errore, file nascosti, *malware* oppure altri documenti volontariamente destinati alla conservazione, ma contenenti informazioni sensibili o riservate tali da impedirne a tempo determinato la consultazione.⁶ Se è chiaro che i file appartenenti al primo nucleo sono da eliminare mentre i secondi da conservare ad accesso chiuso, la situazione prospettata è interessante anche perché scinde l'interrogativo “cosa salvare” dal “cosa pubblicare”, offrendo una testimonianza delle frequenti insidie che si incontrano nel trattamento del digitale d'autore.

Volgendo lo sguardo alle modalità di conservazione del patrimonio digitale, emergono altre sfide che necessitano di essere discusse in seno alla comunità scientifica internazionale. Un ordine di considerazioni di assoluto rilievo riguarda l'accessibilità nel tempo, che si declina sotto vari aspetti. Per i formati non più utilizzati o supportati, la soluzione adottata dalla biblioteca MARBL (poi Stuart A. Rose Library) per i documenti di Salman Rushdie realizzati con l'applicativo ClarisWorks ([5][4]) risulta ampiamente condivisibile: la scelta di percorrere sia «la strada dell'emulazione del computer» che quella della «migrazione dei file in formati aperti e duraturi» (p. 34) contempera i difetti di un approccio con i vantaggi dell'altro. Tuttavia, il problema dell'obsolescenza interessa anche il DAMS (Digital Asset Management System) impiegato nella supervisione delle fasi di acquisizione e conservazione dell'archivio. Gli applicativi sviluppati da Primo Baldini per il progetto PAD, come il software gestionale QUANDO (Quality Control for Archiving and Networking of Digital Objects), sono oggetto di una continua implementazione, ma come garantire il loro funzionamento anche con il venir meno del supporto economico che ne ha permesso la genesi e lo sviluppo? L'interrogativo rende manifesta l'urgenza di un piano di “eredità digitale” che le istituzioni sono chiamate a mettere a punto per evitare che gli sforzi

⁵ Già Michel Foucault nel noto saggio *Che cos'è un autore* [1969] si era interrogato su ciò che era conveniente pubblicare di uno scrittore: «Fra i milioni di tracce lasciate da una persona dopo la sua morte, come definire un'opera? La teoria dell'opera non esiste, e coloro che ingenuamente intraprendono la pubblicazione delle opere non posseggono una simile teoria, il che paralizza ben presto il loro lavoro empirico» ([3]: 5).

⁶ L'argomento è ribadito con puntualità dall'autrice, che non manca di registrare come «attraversare un archivio personale come quello di Pecoraro, che contiene una parte quantitativamente importante del tavolo digitale originario, signific[hi] entrare non solo nel lavoro ma anche nella vita di una persona e accedere a una quantità importante di dati personali» (p. 161). Su questo, vd. anche [4]: 215.

profusi in decenni di attività e ricerche vengano vanificati da congiunture nefaste che mettono a rischio la loro sopravvivenza. È lodevole rimarcare a questo proposito l'adozione nell'ambito del progetto PAD di un sistema distribuito affidato ai server dell'Università di Pavia che coinvolge anche la sede distaccata di Cremona, sì da prevenire – almeno parzialmente – le conseguenze di eventuali calamità naturali (sistema di *disaster recovery*, cfr. p. 58).

La sollecitudine prestata alla conservazione dei file sul lungo periodo deve essere rivolta anche ai processi di acquisizione degli stessi. Carbé ammette che nell'ambito dell'esperienza pavese «ogni archivio acquisito in questi anni ha rappresentato un caso a sé» (p. 110): ciò non costituisce un punto debole dal momento che solo la sperimentazione di pratiche differenti consente di individuare le più efficaci. La maggior parte dei conferimenti ha previsto il trasferimento dei file dal computer dello scrittore ad un disco esterno, ma non è difficile isolare modalità alternative. Esempio paradigmatico è quello del trasferimento di Gianrico Carofiglio: avvenuto il 25 maggio 2011 e concretizzatosi con l'invio di sei file in allegato a un'e-mail (p. 62), ha determinato la perdita di alcuni metadati come la data di creazione del file, sostituita (cioè sovrascritta) da quella della trasmissione. Più in generale, occorre avvertire che le indicazioni cronologiche associate a un file non sono sempre attendibili anche quando le procedure di conferimento sono condotte nel rispetto delle *best practices* perché non di rado gli scrittori incollano in nuovi file le copie di precedenti versioni: è il caso di Franco Buffoni (p. 94) e di Francesco Pecoraro (p. 150).⁷ La possibilità di intervistare gli autori e chiedere loro chiarimenti in merito alla genesi e alla composizione degli archivi digitali permette di ovviare ad alcune di queste criticità. Altre azioni, invece, come il trasferimento di un file in un volume che non supporta la cronologia delle modifiche, sono irreversibili e determinano la perdita di una porzione più o meno consistente di dati. Sicché il monito di Thorsten Ries⁸ per cui ogni azione (compresa la mera apertura di un file) lascia una traccia nel sistema e può avere profonde conseguenze sul fondo stesso dovrebbe fungere al contempo da stimolo per conferire la debita attenzione alle discussioni relative alle pratiche di acquisizione e gestione degli archivi *born digital*.

Conclusioni

La solida impalcatura del volume e la chiarezza espositiva che lo contraddistingue consentono anche ai lettori digiuni di conoscenze sull'argomento di avere contezza del campionario di sfide che il digitale d'autore pone sul tappeto, come pure dei principali problemi di natura informatica,

⁷ Nel caso di quest'ultimo è inoltre contemplato un cambio del mezzo, dal momento che la prima stesura avviene su un tablet e solo in un secondo momento il contenuto è trasferito su di un computer: «Io in genere giro con il tablet e [*sicil.* una certa cosa] la scrivo lì, dopodiché mi spedisco questi file e li inserisco o non inserisco nel testo che sto scrivendo, o comunque li raggruppo per directory». La dichiarazione è resa nel corso del primo seminario letterario PAD (Università di Pavia, 30 marzo 2016), al quale hanno partecipato anche Gilda Policastro, Clelia Martignoni ed Emmanuela Carbé (<https://youtu.be/uMcrAnuJNx0>).

⁸ Il riferimento è al *keynote speech* intitolato *Born-digital Literary Archives – the Present and the Future of an Emerging Field* che lo studioso ha tenuto nell'ambito del XII convegno annuale AIUCD (Università degli Studi di Siena, 5-7 giugno 2023). Dello stesso, nell'attesa che venga licenziata la monografia *Hybride Textgenese und digitale Forensik* per Universitätsverlag Winter, cfr. lo *Special Issue on Born-Digital Archives* dell'*International Journal of Digital Humanities* (1, 1) curato con Gábor Palkó e, in particolare, il loro saggio introduttivo ([6]).

archivistica e giuridica che avvolgono l'atto del conferimento.⁹ Più nel dettaglio, si può affermare che con questo libro l'autrice affida loro una duplice eredità. Da un lato, Carbé ribadisce a più riprese la necessità di individuare *best practices* e adottare standard condivisi sia per quanto riguarda la conservazione che la fruizione dei materiali contenuti in un archivio letterario *born-digital*:

Il percorso verso la definizione di modelli e standard è insomma ancora all'inizio, e non è ormai più differibile una riflessione all'interno della comunità scientifica per individuare delle buone pratiche e superare quelle criticità che rendono, almeno per ora, la gestione di questi archivi critica e troppo spesso lasciata alla buona volontà di poche persone (p. 110).

In queste righe si coglie la principale ragion d'essere dello studio che, tracciando una sintesi del panorama nazionale, apre al confronto con le esperienze di altri Paesi e invita a una discussione su larga scala quale prerequisito essenziale per raccogliere i frutti degli esperimenti condotti in seno alle istituzioni locali e nazionali. Considerata la recente affermazione dell'ambito d'indagine, Carbé ritrae una situazione in divenire, tesa fra la sperimentazione imperante degli scorsi anni e la sempre più sentita urgenza di sistematizzare quanto fatto finora. Per queste ragioni, *Digitale d'autore* non può essere considerato un traguardo, ma piuttosto un punto di partenza – e segnatamente un punto di partenza imprescindibile – per ogni perlustrazione futura in questo campo. Il suo carattere seminale, che rivela anche l'ottima riuscita dell'operazione, si coglie nelle varie questioni cui allude, destinate ad assolvere un'importanza nodale nel futuro prossimo del digitale d'autore. È il caso dell'impiego di tecniche forensi (p. 162) per un'efficace mappatura e interrogazione dei fondi (soluzione già esperita in alcuni progetti esteri) o del ricorso al Natural Language Processing (NLP) e all'intelligenza artificiale (IA) per la loro esplorazione e l'estrazione di dati (p. 129). Nell'ambito di PAD, tecniche di NLP sono state impiegate in via sperimentale tra il 2019 e il 2020 per l'analisi dei 43,686 file che compongono l'archivio di Francesco Pecoraro. È auspicabile che nei prossimi anni il loro utilizzo sia costante nei fondi caratterizzati da un'elevata consistenza o da un alto tasso di documenti duplicati, rinominati, interpolati, oggetto di diversi backup o dislocati in luoghi differenti (*path*) di uno stesso *hard disk*, nonostante la loro affinità o genesi comune. In questo modo sarà più agevole riconoscere fra migliaia di file le diverse redazioni di un testo o i suoi salvataggi intermedi e ricostruire con ciò la genesi di un'opera. Un'ultima questione su cui andrà posta l'attenzione – non discussa nel volume perché i tempi non sono ancora maturi – riguarda l'interoperabilità fra applicativi di istituzioni differenti. Il conferimento non deve essere necessariamente esclusivo: un autore può condividere uno stesso file con più istituzioni o entrambe possono essere destinatarie di due fondi indipendenti. Dall'analisi di situazioni come queste – che nel futuro potrebbero diventare canoniche – si evince come la sistematizzazione delle esperienze dopo una fase di sperimentazione vada assumendo un carattere non più eludibile. È piacevole constatare che in questo percorso, che si preannuncia ricco di scoperte perché ancora tutto da definire, il volume di Emmanuela Carbé sancisca la prima e fondamentale tappa.

⁹ Alle pp. 53-54 l'autrice riporta le condizioni generali della prima versione del contratto che l'autore era chiamato a firmare all'atto del conferimento dei suoi materiali al progetto PAD. I principali aggiornamenti negli anni successivi, con particolare attenzione all'esclusività del conferimento, sono illustrati nelle pagine successive del capitolo (specialmente alle pp. 68-70).

References

- [1] Curino, Luciano. 1983. «Scusi, lei lo scriverebbe un romanzo con il computer?». *Tuttolibri*, 5 giugno 1983.

Kirschenbaum, Matthew G., Erika Farr, Kari M. Kraus, Naomi L. Nelson, Catherine Stollar Peters, Gabriela Redwine, and Doug Reside. 2009. “Approaches to Managing and Collecting Born-Digital Literary Materials for Scholarly Use.” White Paper to the NEH Office of Digital Humanities Level 1 Digital Humanities Start-Up Grant 98537. <http://dx.doi.org/10.17613/M66937>.
- [2] Erway, Ricky. 2010. “Defining ‘Born Digital.’” Demystifying Born Digital. Online Computer Library Center (OCLC). www.oclc.org/research/activities/hiddencollections/borndigital.pdf.
- [3] Foucault, Michel. 2010. *Scritti letterari*. A cura di Cesare Milanese. 2^a ed. Milano: Feltrinelli.
- [4] Kirschenbaum, Matthew G. 2016. *Track Changes: A Literary History of Word Processing*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.
- [5] Alexander, Benjamin. 2015. “The Salman Rushdie Archive and the Re-Imagining of a Philological E-Volution.” In *Texts, Transmissions, Receptions. Modern Approaches to Narratives*, edited by André Lardinois, Sophie Levie, Hans Hoeken, and Christoph Lüthy, 71–94. Leiden: Brill. https://doi.org/10.1163/9789004270848_006.
- [6] Ries, Thorsten, and Gábor Palkó. 2019. “Born-Digital Archives.” *International Journal of Digital Humanities* 1 (1): 1–11. <https://doi.org/10.1007/s42803-019-00011-x>.