

Limina

Limina

intorno al testo

Rivista diretta da
Ciro Perna ed Elisabetta Tonello

anno II
2025
fascicolo 1

libreriauniversitaria.it
edizioni

Direttori / Editors-in-chief

Ciro Perna (Università della Campania “Luigi Vanvitelli”)
Elisabetta Tonello (Università di Salerno)

Comitato editoriale / Editorial board

Giancarlo Alfano (Università “Federico II” di Napoli)
Zygmunt G. Barański (University of Notre Dame)
Dario Brancato (Concordia University)
Marina Buzzoni (Università di Venezia)
Rosario Coluccia (Accademia della Crusca)
Claudia Corfiati (Università di Bari)
Massimiliano Corrado (Università “Federico II” di Napoli)
Alessio Decaria (Università di Genova)
Federico Della Corte (Università eCampus)
Nathalie Ferrand (ITEM – CNRS Paris)
Maurizio Fiorilla (Università Roma Tre)
Giovanna Frosini (Università per Stranieri di Siena)
Daniela Gionta (Università di Messina)
Rosa Giulio (Università di Salerno)
Cynthia Maria Hamlin (Universidad de Buenos Aires)
Bernhard Huss (Freie Universität Berlin)
Paola Italia (Università di Bologna)
Kosuke Kunishi (Ritsumeikan University)
Rita Librandi (Università di Napoli “L’Orientale”)
Silvia Maddalo (Università della Tuscia)
Alessandra Perriccioli Saggese (Università della Campania “Luigi Vanvitelli”)
Giulia Perucchi (Università di Chieti Pescara)
Chiara Ponchia (Università di Padova)
Fabio Romanini (Università di Ferrara)
Arielle Saiber (Johns Hopkins University)
Federica Sgarbi (Doshisha University)
Franco Tomasi (Università di Padova)
Juan Varela-Portas de Orduña (Universidad Complutense de Madrid)
Paola Vecchi Galli (Università di Bologna)
Zeno Verlato (CNR Firenze)
Mirko Volpi (Università di Pavia)

Comitato di redazione / Publishing copy-editors

Giuseppe Alvino, Federica Cifariello, Federica Maria Giallombardo,
Serena Malatesta, Riccardo Montalto, Beatrice Mosca, Aldo Stabile, Raffaele Vitolo

Direttore responsabile / Legal representative

Mario Lion Stoppato

Limina. Intorno al testo

Rivista digitale open access a cadenza semestrale
pubblicata da libreriauniversitaria.it Edizioni

N. II (2025) *fascicolo 1*

Limina. Intorno al testo è una rivista open access diffusa con licenza Creative Commons CC BY-NC-SA (Attribuzione – Non Commerciale – Condividi allo Stesso Modo).

Limina. Intorno al testo is an open access journal distributed under the Creative Commons CC BY-NC-SA license (Attribution – Non-Commercial – Share Alike).

Rivista registrata presso il Tribunale di Padova N. 8474/2025

ISSN online: 3035-3343



Limina. Intorno al testo è una rivista sottoposta a peer-review

Limina. Intorno al testo is a Peer-Reviewed Journal

Autrici e autori sono gentilmente pregate/i di inviare i loro contributi al comitato scientifico al seguente indirizzo email: prin.limina@gmail.com

Authors who wish to submit their work to *Limina. Intorno al testo*, should send it to prin.limina@gmail.com

Rivista pubblicata con fondi PRIN: PRIN COFIN 2022_Progetto LiMINA (Lost in Manuscripts. Ideas, Notes, Acknowledgments)

Gli autori dei contributi sono responsabili dell'uso delle immagini

Publication Ethics and Malpractice Statement

Limina. Intorno al testo, is a peer reviewed semestral journal committed to upholding the highest standards of publication ethics. In order to provide readers with articles of highest quality we state the following principles of Publication Ethics and Malpractice Statement.

Authors ensure that they have written original articles. In addition they ensure that the manuscript has not been issued elsewhere. Authors are also responsible for language editing of the submitted article. Authors confirm that the submitted works represent their authors' contributions and have not been copied or plagiarized in whole or in part from other works without clearly citing. Any work or words of other authors, contributors, or sources (including online sites) are appropriately credited and referenced. All authors disclose financial or other conflict of interest that might influence the results or interpretation of their manuscript (financial support for the project should be disclosed). Authors agree to the license agreement before submitting the article.

The editors ensure a fair double-blind peer-review of the submitted papers for publication. The editors strive to prevent any potential conflict of interests between the author and editorial and review personnel. The editors also ensure that all the information related to submitted manuscripts is kept as confidential before publishing.

The editors coordinates the editorial board for reviewing the works to be published in *Limina. Intorno al testo*. The reviewers, members of the scientific committee, include experts in the field of higher education, university lecturers and researchers. Each is assigned papers to review that are consistent with their specific expertise.

Reviewers check all papers in a double-blind peer review process. The reviewers also check for plagiarism and research fabrication (making up research data) and falsification (manipulation of existing research data, tables, or images). In accordance with the code of conduct, the reviewers report any cases of suspected plagiarism or duplicate publishing.

Reviewers evaluate manuscripts based on content without regard to ethnic origin, gender, sexual orientation, citizenship, religious belief or political philosophy of the authors. They ensure that all the information related to submitted manuscripts is kept as confidential and must report to the editors if they are aware of copyright infringement and plagiarism on the author's side. They must evaluate the submitted works objectively as well as present clearly their opinions on the works in a clear way in the review form. A reviewer who feels unqualified to review the research reported in a manuscript notify the editors and excuses himself from the review process.

SOMMARIO

Ai <i>limina</i> delle Digital Humanities	7
☞ Beatrice Mosca	

I. SAGGI

Intelligenze artificiali e mondi immersivi: nuovi paradigmi per le Digital Humanities	13
☞ Fabio Ciotti	

Progettazione e modellazione multilineare per edizioni scientifiche digitali.	41
☞ Cristina Marras	

La TEI e la rappresentazione del testo: alcune riflessioni critiche . .	55
☞ Roberto Rosselli Del Turco	

Work in progress: the palimpsest <i>Comedy</i> of the “Dante Poggiali” (BNCF, ms. Pal. 313)	65
☞ Ciro Perna, Elisabetta Tonello	

<i>Divine (virtual) Reality</i> : un mondo virtuale per immergersi nella <i>Commedia</i>	68
☞ Serena Malatesta	

Building and maintaining the crowdsourced archive <i>Dante Today</i> . . 83	
☞ Elizabeth Coggeshall, Arielle Saiber	

Custodire il passato per il futuro: la transizione digitale della Biblioteca Apostolica Vaticana	97
☞ Paolo Manoni	

Transforming manuscripts studies through technology: methods, challenges and future directions 107
☞ Arianna Traviglia

Limina from the virtual unwrapping of the Herculaneum Papyri . . 115
☞ Gianluca Del Mastro

“Voce dei Margini”: le postille di Giorgio Bassani in un’edizione scientifica digitale 125
☞ Angelo Mario Del Grosso, Angela Siciliano

II. SCHEDE BIBLIOGRAFICHE

Lorella Viola, *The Humanities in the Digital: Beyond Critical Digital Humanities*, Cham, Palgrave Macmillan, 2023, XXVI + 173 pp. 155
☞ Denise Tirone

Julianne Nyhan, Geoffrey Rockwell, Stéfan Sinclair, Alexandra Ortolja-Baird (a cura di), *On Making in the digital humanities: The scholarship of digital humanities development in honour of John Bradley*, London, UCL Press, 2023, XX + 279 pp. 161
☞ Denise Tirone



10.69118/2025_L0

Ai *limina* delle Digital Humanities

Premessa

Fin dalle prime sperimentazioni computazionali in ambito umanistico, le Digital Humanities (DH) si sono collocate all'incrocio tra tecnologia e pratiche interpretative, configurandosi ben oltre la semplice applicazione di strumenti digitali, ma come trasformazione radicale dei modi di produrre, rappresentare e comprendere il sapere. Le DH rappresentano oggi un paradigma multidimensionale di convergenze teoriche, metodologiche e politiche, in cui – come osserva Kirschenbaum – la distinzione tra ciò che esse *sono* e ciò che *fanno* diventa essa stessa oggetto di riflessione. Questo insieme di pratiche e approcci mette in discussione i confini disciplinari tradizionali e richiede una rielaborazione profonda dei paradigmi epistemologici dell'umanesimo contemporaneo. Per la prima volta, ricercatrici e ricercatori si trovano a confrontarsi con metodologie applicate alla gestione di dati complessi, la cui organizzazione e rappresentazione digitale implicano scelte interpretative, formali e operative che, in quanto modellizzazioni, sono costitutivamente ermeneutiche e mai neutrali. Non si tratta solo di tecnologie al servizio dei saperi umanistici o un mero campo di intervento strategico: è un mutamento strutturale del modo in cui il sapere viene prodotto, trasmesso e riconosciuto come tale.

Operare oggi nel contesto delle Digital Humanities comporta sfide significative: i nuovi strumenti digitali permettono infatti di aggregare conoscenze precedentemente disperse e di ricomporre in modo reticolare ambiti del sapere che la tradizione accademica ha trattato come compartimenti stagni. Ridurre le Digital Humanities a un mero strumento di “traduzione” del sapere o a una forma aggiornata di divulgazione rappresenta una semplificazione che ne compromette la portata trasformativa. Queste si configurano come uno spazio teorico e operativo in cui la materialità e l'infrastruttura dei dati diventano elementi centrali di riflessione e progettazione per la costruzione di conoscenza. Le esigenze ermeneutiche che ne derivano superano l'interpretazione del dato verso l'esplicitazione dell'atto interpretativo stesso, reso visibile attra-

verso scelte modellanti e computazionali. La transizione verso l'ambiente digitale non può essere intesa come una trasposizione *verbum pro verbo* del reale, ma come un processo di *trans-ductio*, inteso nel senso di una rimediazione produttiva che, pur trasformando l'oggetto, ne preserva la possibilità di significazione. In quest'ottica, la computabilità del dato umanistico è un'operazione critica che deve essere consapevole della propria performatività epistemica. Lo scopo è duplice: da un lato, elaborare strumenti e tecnologie che rispettino la complessità semiotica e culturale del dato umanistico; dall'altro, promuovere pratiche teoriche e operative che rendano conto della profonda co-implicazione tra contenuto, forma e supporto nel contesto dell'ecosistema digitale contemporaneo.

Ogni rappresentazione digitale è una scelta interpretativa: le Digital Humanities non restituiscono il reale, ma ne modellano la possibilità di comprensione. In questo senso l'ambiente digitale è un *medium* critico, che rende visibili le condizioni materiali e concettuali della conoscenza e ne rifonda le basi attraverso operazioni di astrazione, selezione e costruzione intenzionale. La sfida delle DH sta proprio nel trasformare l'infrastruttura in discorso e il dato in oggetto culturale modellato.

Da queste riflessioni è scaturita, all'interno del progetto *DanteLimina.it*, l'esigenza di interrogarci sulle potenzialità offerte dagli strumenti digitali per la rappresentazione di un oggetto di ricerca tanto peculiare quanto sfuggente: il *limen*. Nei diversi progetti che ruotano attorno a tale concetto, ci si è confrontati con la necessità di individuare metodologie adeguate alla sua resa e visualizzazione in ambiente digitale.

La natura stessa dell'oggetto che ricerchiamo – i *limina* – consente di ricostruire i processi della ricezione e della mitopoiesi dantesca, con uno sguardo che va oltre ciò che è presente nel margine del testo all'interno del libro manoscritto o a stampa. I lettori della *Commedia* hanno stabilito un dialogo attivo con l'opera e con l'autore, lasciando tracce materiali e simboliche di tale interazione. Rappresentare questo *limen* attraverso gli strumenti digitali significa, pertanto, restituire la complessità di un'interfaccia dinamica tra testo, supporto e interpretazione. Ogni visualizzazione si configura così come un atto interpretativo e l'interfaccia diventa un dispositivo critico in grado di rendere esplicite le scelte di rappresentazione sottese al modello.

La necessità di un confronto diretto con esperti, nonché con filologi, storici della letteratura e della lingua che hanno impiegato metodologie e strumenti digitali nelle loro ricerche, ci ha portati a immaginare uno

spazio di dialogo e condivisione. Tale esigenza si è concretizzata nella tavola rotonda organizzata in occasione del I Convegno Internazionale LiMINA presso l'Università degli Studi di Cagliari, che ha rappresentato un primo momento prezioso di scambio e discussione interdisciplinare, ed è proseguita nei successivi incontri presso le biblioteche partner Biblioteca Medicea Laurenziana e Riccardiana di Firenze, Biblioteca Nazionale di Napoli e Biblioteca Nazionale Centrale di Firenze. L'obiettivo era duplice: da un lato, riflettere insieme sugli strumenti teorici e pratici da adottare; dall'altro, cercare soluzioni comuni alle difficoltà emerse durante lo sviluppo del progetto *DanteLimina.it*. Un primo esito di questi incontri è confluito in questo fascicolo, che raccoglie i contributi di coloro che hanno partecipato, studiosi ed esperti con i quali, in diversi contesti, abbiamo avuto modo di confrontarci. I loro interventi non si limitano alla presentazione di metodi o strumenti, ma offrono vere e proprie modellizzazioni operative della conoscenza, che esplicitano la tensione tra astrazione computazionale e specificità del dato umanistico. Questi approcci hanno suggerito paradigmi e prospettive che abbiamo potuto – e si potrebbero ancora – riadattare e integrare all'interno del progetto *DanteLimina.it*.

Il fascicolo si apre con una riflessione di ampio respiro sulle implicazioni metodologiche delle Digital Humanities: in questa direzione si muove Fabio Ciotti, che analizza l'impatto epistemologico dell'introduzione dell'intelligenza artificiale e delle tecnologie immersive, affiancato dal contributo di Cristina Marras, sul rapporto tra dimensione teorica e prassi operativa.

Seguono alcune analisi incentrate sulle criticità strutturali e concettuali della codifica dei *marginalia* testuali in XML/TEI, come evidenziato dalle puntuali osservazioni di Roberto Rosselli Del Turco. A queste si affiancano riflessioni sull'efficacia delle tecnologie digitali nel supportare la ricerca su testimoni particolarmente complessi, come dimostra il caso illustrato da Ciro Perna ed Elisabetta Tonello del manoscritto di Firenze, Biblioteca Nazionale Centrale, Pal. 313. In prospettiva diversa, ma complementare, Serena Malatesta presenta *Divine Reality* all'interno del progetto D.A.N.T.E., che esplora il potenziale delle tecnologie immersive (e in particolare la realtà virtuale) per la valorizzazione didattica e critica del patrimonio dantesco.

Il concetto di *liminarità* si estende, inoltre, alle sue declinazioni più contemporanee e pop, concentrandosi sulla ricezione "estrema" della tradizione dantesca e sulle sue possibilità di fruizione, rielaborazione e riuso

all'interno di archivi digitali: è il caso del progetto *Dante Today*, presentato da Elisabeth Coggeshall e Arielle Saiber.

Sull'importanza di approcci integrati, soprattutto in contesti complessi che richiedono il dialogo tra competenze disciplinari eterogenee, i due contributi successivi, rispettivamente dedicati al lavoro passato e a prospettive future, offrono uno sguardo complementare sulla relazione tra tecnologie digitali e patrimonio culturale. Quello di Paola Manoni si concentra sul lavoro di digitalizzazione e conservazione svolto presso la Biblioteca Apostolica Vaticana, mentre l'intervento di Arianna Traviglia esplora le prospettive future legate all'impiego delle nuove tecnologie per la tutela e la valorizzazione dei beni culturali.

Sulle recenti innovazioni nelle tecniche di digitalizzazione e nelle modalità di analisi dei materiali, i contributi di Gianluca Del Mastro, Angelo Mario Del Grosso e Angela Siciliano, propongono i risultati delle loro più recenti ricerche, in cui il digitale si configura come una pratica critica e progettuale, capace di modellare nuove forme di accesso e interpretazione del patrimonio.

In chiusura, il fascicolo presenta due schede bio-bibliografiche in dialogo tra loro, che affrontano da prospettive differenti il modo in cui le Digital Humanities si configurino oggi come un ambiente sperimentale per la ridefinizione delle pratiche del sapere. In particolare, emerge come il *making* – l'atto stesso del 'fare' – assuma progressivamente una valenza epistemica autonoma, inscritta in una dimensione sempre più multidisciplinare e corale, in cui l'infrastruttura tecnica e concettuale diventa essa stessa oggetto di riflessione critica.

Ci auguriamo che la lettura di questo fascicolo monografico possa suscitare nei lettori una riflessione vivace e profonda sulle tematiche qui proposte. Il nostro intento è che esso rappresenti uno stimolo per ulteriori sviluppi e confronti, incoraggiando la pratica dell'incontro – e anche dello scontro – con l'alterità, nella consapevolezza che dal dialogo aperto e critico possano scaturire risultati significativi per la ricerca.

In questo orizzonte, la finalità è la costruzione di un terreno comune, dinamico e intuitivo, in cui le diverse competenze cooperino verso una rappresentazione e condivisione del sapere sempre più solida e consapevole del proprio presente, ma, al tempo stesso, orientata a un futuro di convergenze, più che di separazioni, nella ricerca umanistica.

BEATRICE MOSCA

I. SAGGI



10.69118/2025_L1

Intelligenze artificiali e mondi immersivi: nuovi paradigmi per le Digital Humanities

Fabio Ciotti

Università di Roma Tor Vergata
fabio.ciotti@uniroma2.it

Artificial Intelligence and immersive worlds: new paradigms for Digital Humanities

Abstract (ITA)

L'articolo esplora l'impatto epistemologico e metodologico dell'introduzione di tecnologie emergenti – intelligenza artificiale generativa (IA) e tecnologie immersive (realtà aumentata, mista e virtuale) – nel contesto delle Digital Humanities. Nella prima parte si analizzano le tecnologie immersive, evidenziandone il potenziale nella creazione di esperienze multisensoriali e interattive che ridefiniscono la nozione di immersione narrativa tradizionalmente associata alla letteratura. Viene discussa la correlazione teorica tra immersione testuale e digitale, con riferimento a diversi approcci teorico-letterari. La seconda parte esamina l'IA generativa e i Large Language Models (LLM), descrivendone l'architettura tecnica, i meccanismi di funzionamento e le capacità emergenti. L'articolo affronta criticamente il dibattito teorico sulle reali capacità cognitive degli LLM, esaminando posizioni deflazioniste e possibiliste, e propone come gli LLM possano essere considerati modelli rappresentativi dei sistemi culturali attraverso il concetto lotmaniano di "semiosfera". Infine, vengono esplorate le prospettive future derivanti dalla convergenza tra tecnologie immersive e IA generativa, inclusa la produzione automatizzata di contenuti multimediali (ecfrasi inversa, generazione video, world models). L'articolo conclude sottolineando la necessità di un approccio interdisciplinare rigoroso e aperto, capace di sfruttare pienamente il potenziale innovativo di queste tecnologie nella ricerca umanistica contemporanea.

Abstract (ENG)

This article explores the epistemological and methodological implications of introducing new technologies – generative artificial intelligence (AI) and immersive technologies (augmented, mixed and virtual reality) – in the context of digital humanities. The first part analyses immersive technologies, highlighting their potential to create multisensory and interactive experiences that redefine the notion of narrative immersion traditionally associated with literature. The theoretical relationship between textual and digital immersion is discussed with reference to different approaches to literary theory. The second part examines generative

AI and Large Language Models (LLMs), describing their technical architecture, operational mechanisms and emerging capabilities. The article critically addresses the theoretical debate about the actual cognitive capabilities of LLMs, examining both deflationist and possibilist positions, and proposes how LLMs can be considered representative models of cultural systems through Lotman's concept of the 'semiosphere'. Finally, the article explores future perspectives arising from the convergence of immersive technologies and generative AI, including the automated production of multimedia content (inverse ekphrasis, video generation, world models). The article concludes by emphasising the need for a rigorous and open interdisciplinary approach capable of fully exploiting the innovative potential of these technologies in contemporary humanistic research.

Parole chiave / Keywords

Tecnologie immersive, IA generativa, Innovazione digitale / Immersive Technologies, Generative AI, Digital Innovation

Introduzione

Il vasto e multiforme campo delle Digital Humanities¹ è sempre stato soggetto alla pervasiva influenza dell'innovazione tecnologica, adattandosi alle alterne fortune delle singole tecnologie informatiche nel corso degli anni. La capacità trasformativa delle tecnologie digitali, d'altra parte, è innegabile, come lo è il loro impatto sociale e culturale, ed è naturale che chi si occupa di adottare criticamente sistemi e modelli computazionali nello studio della sfera simbolica e culturale subisca il richiamo dell'hype tecnologico di turno. In questa fase storica il ruolo dominante nei "discorsi" sul digitale è rivestito dall'Intelligenza Artificiale (IA) generativa, con particolare riferimento ai Large Language Models (LLM)², e, in misura minore, dalla rinnovata attenzione verso le tecnologie immersive (Realtà Aumentata, Mista e Virtuale), ricomprese sotto la suggestiva formula di *Metaverso* (il termine fa la sua comparsa nel romanzo di Neal Stephenson *Snow Crash*³, ma deve il suo successo al fatto di essere stato adottato come fulcro della strategia di marketing della casa madre di Facebook e altri social media globali, ribattezzata con l'occasione Meta).

L'introduzione e la rapida diffusione di queste tecnologie stanno innescando una trasformazione radicale nelle metodologie delle Digital

1 Si veda Ciotti 2023a.

2 Si veda Ciotti 2023b; Roncaglia 2023.

3 Si veda Stephenson 1992.

Humanities, che trascende in modo significativo la dimensione della semplice innovazione strumentale, per configurarsi come una vera e propria svolta paradigmatica. La convergenza e la sinergia tra i sistemi capaci di generare esperienze multisensoriali immersive e i modelli neurali avanzati dell'IA generativa promette di aprire nuove possibilità sul piano analitico ed ermeneutico, di esplorare dimensioni precedentemente inaccessibili dell'esperienza testuale, offrendo vie alternative per l'interpretazione e la comprensione dei fenomeni culturali. In questo intervento intendo riflettere su alcuni dei fondamenti e delle implicazioni teoriche e metodologiche determinate da tale convergenza tecnologica, rivolgendo una particolare attenzione alle potenziali applicazioni nell'ambito degli studi letterari.

1. Le tecnologie immersive e realtà virtuale nel contesto umanistico

Le tecnologie immersive e la realtà virtuale (VR) hanno fatto la loro comparsa agli inizi degli anni '90, anticipate da numerose prefigurazioni letterarie, in particolare quelle di William Gibson, soprattutto nei romanzi della *Trilogia dello Sprawl*.⁴ – *Neuromancer*, *Count Zero*, *Mona Lisa Overdrive* – e nei racconti connessi, e del movimento Cyberpunk.⁵ Inizialmente, suscitarono un'ondata di entusiasmo, alimentando la promessa di mondi digitali capaci di espandere i nostri sensi e permettere interazioni inedite con la realtà.⁶ Tuttavia, le significative limitazioni tecnologiche dell'epoca determinarono presto una crisi di attenzione e un lungo periodo di relativa oscurità mediatica. Solo negli ultimi anni si è assistito a una parziale inversione di tendenza, trainata dai recenti sviluppi hardware e dalla crescente disponibilità di interfacce immersive a basso costo: visori standalone a prezzi accessibili (es. Meta Quest 3, Pico 4) con prestazioni precedentemente riservate a sistemi high-end; controller aptici avanzati per interazioni naturali con oggetti virtuali; sistemi di *eye-tracking* integrati che abilitano nuove modalità di interazione. La diminuzione dei costi e il miglioramento delle prestazioni dell'hardware VR moderno contribuiscono a un'esperienza più fluida, intuitiva e coin-

4 Si veda Gibson 2019.

5 Si veda Bould 2005; Caronia-Gallo 1996.

6 Si veda Rheingold 1991.

volgente, fondamentale per il suo utilizzo efficace nelle applicazioni DH, dove l'engagement dell'utente e il senso di presenza sono aspetti importanti, ma devono essere commisurati alle risorse limitate tipiche della ricerca umanistica.

In linea generale le tecnologie digitali immersive si articolano in tre tipologie che si articolano lungo il continuum della virtualità definito da Milgram e Kishino (1994), ciascuna delle quali offre diverse possibilità di integrazione percettiva tra elementi virtuali e ambiente fisico. La *Realtà Aumentata* (AR) sovrappone elementi visivi digitali al mondo reale, integrando ed espandendo digitalmente l'ambiente fisico dell'utente in tempo reale. La sua funzionalità di base consiste nel creare connessioni tra il mondo reale e le informazioni generate elettronicamente, arricchendo la percezione umana. La *Realtà Mista* (MR) consente l'interazione in tempo reale tra oggetti virtuali e reali, permettendo la sovrapposizione e l'interazione tra elementi digitali e l'ambiente reale a vari livelli. Essa si basa su metodi di input avanzati e sulla percezione ambientale, adattandosi all'ambiente fisico dell'utente. La *Realtà Virtuale* (VR) crea ambienti digitali tridimensionali completi, offrendo il massimo grado di immersione. Essa fornisce in modo massimale una «esperienza interattiva e immersiva generata da un computer» (Pimentel e Teixeira 1993). La VR è uno spazio percettivo immersivo e interattivo generato dal computer che simula un ambiente tridimensionale dotato di parte delle proprietà fisiche dell'ambiente 'reale', cui si accede con l'ausilio di hardware specializzato come visori e controller, e offre un senso di presenza e la possibilità di interagire ed esplorare mondi virtuali realistici e reattivi.

Dal punto di vista tecnico, i sistemi immersivi si avvalgono di un ecosistema tecnologico complesso e integrato per creare l'illusione di un mondo virtuale coerente e navigabile. Gli *Head-Mounted Displays* (HMD) rappresentano l'interfaccia primaria per la VR, offrendo un completo isolamento percettivo dal mondo esterno e una visione stereoscopica e tridimensionale di quello virtuale attraverso schermi posizionati davanti agli occhi dell'utente, combinati con sistemi di tracciamento del movimento che adattano continuamente la prospettiva visuale ai movimenti della testa e altoparlanti 3D. Nelle applicazioni AR, dove non è necessario l'isolamento, sono molto più diffusi dispositivi come gli *smart glass*, ma anche i comuni smartphone e tablet possono essere funzionali in alcuni contesti. Per la percezione tattile si usano i *data glove*, che consentono di tradurre i movimenti delle mani in azioni all'interno dell'ambiente virtuale, permettendo manipolazioni e interazioni con oggetti si-

mulati, ma nei sistemi low end essi sono sostituiti da controller wireless simili a quelli delle piattaforme di gioco. Questi apparati hardware di base possono essere completati da sensori di movimento per tracciare la posizione dell'utente nello spazio, sistemi audio spazializzati per creare un ambiente sonoro immersivo. Non vanno poi dimenticati i sistemi hardware e software per il *rendering* in tempo reale delle immagini in risposta ai movimenti e alle azioni dell'utente. A questi strumenti si aggiungono i framework di modellazione tridimensionale e simulazione comportamentale che costituiscono l'infrastruttura su cui si costruiscono gli ambienti virtuali. L'esperienza VR è dunque il risultato di una complessa orchestrazione di tecnologie integrate che collaborano per ingannare i sensi e creare un'esperienza percettiva convincente.

Come detto, i concetti chiave che permettono di inquadrare teoricamente l'esperienza VR sono *immersione* e *interattività*. L'immersione si riferisce alla sensazione di presenza all'interno di un mondo fittizio, percepito come reale e autonomo grazie alla simulazione computazionale. Questo mondo virtuale, popolato da oggetti e personaggi specifici, crea uno spazio con cui l'utente può relazionarsi in modo diretto e intuitivo. L'interattività definisce invece la capacità dell'utente di agire all'interno dell'ambiente virtuale, modificando l'esperienza attraverso azioni quali l'esplorazione, la manipolazione e la trasformazione dell'ambiente stesso.

Questi due principi, apparentemente tecnici, hanno profonde implicazioni teoriche e hanno assunto un ruolo significativo anche nei dibattiti della teoria letteraria degli ultimi decenni. La correlazione concettuale tra l'esperienza dell'immersione nel testo narrativo durante l'atto della lettura e quella fornita dai sistemi immersivi digitali appare evidente quando consideriamo modelli e riflessioni teoriche quali la semantica a mondi possibili applicata ai testi finzionali elaborata da teorici come Pavel (1986), Eco (1979), Doležel (1999), Ryan (1991), che concepisce i testi narrativi come dispositivi costruttori di mondi alternativi dotati di proprietà spaziali e temporali specifiche; la teoria della finzione come *make-believe* proposta da Walton (1993), dove i fatti finzionali sono "veri nel gioco appropriato di make-believe", ossia veri nel mondo finzionale rappresentato; la nozione bachtiniana di cronotopo,⁷ intesa come matrice spazio-temporale che struttura la narrazione letteraria. Parallelamente, l'interattività digitale rappresenta un'evoluzione delle teorie sulla costru-

7 Si veda Bachtin 1997.

zione cooperativa del significato tra testo e lettore, manifestando materialmente le intuizioni teoretiche postmoderne e trasformando il testo da elemento statico a costruzione dinamica, come sostenuto dalle teorie sull'ipertesto di George Landow, Stuart Moulthrop, e altri studiosi negli anni iniziali della rivoluzione digitale.⁸

In particolare Marie-Laure Ryan ha sviluppato un'analisi sistematica dei rapporti tra la nozione di realtà virtuale e la narratività nel suo influente volume *Narrative as Virtual Reality* (Ryan 2003; edizione riveduta 2015), dove ha esaminato i parallelismi tra l'esperienza immersiva nei mondi virtuali, l'esperienza del lettore nei mondi narrativi e la stessa evoluzione dei sistemi letterari. La teorica statunitense, infatti, osserva come la storia del romanzo moderno offra una prospettiva illuminante sulle dinamiche culturali della nozione di immersione, evidenziando un'alternanza tra fasi dominate da poetiche dell'immersione e fasi orientate al distanziamento critico e alla meta-riflessione. Il romanzo borghese del XVIII secolo mostrava una caratteristica ambivalenza, utilizzando tecniche illusionistiche per creare mondi finzionali convincenti, ma richiamando contemporaneamente l'attenzione sul processo narrativo stesso attraverso interventi autoriali e strategie metaletterarie. Il romanzo realista del XIX secolo ha privilegiato l'immersione, caratterizzandosi per un narratore progressivamente meno visibile e una focalizzazione intensificata sulle esperienze emotive dei personaggi, creando l'illusione di un accesso diretto e non mediato alla realtà rappresentata. La letteratura modernista del XX secolo ha invece compiuto un deliberato allontanamento dalle strategie immersive, privilegiando approcci metanarrativi, sperimentali e autoriflessivi che problematizzavano la relazione tra testo, realtà e lettore.

Sul piano più strettamente teorico Ryan afferma che l'immersione narrativa si articola in multiple dimensioni interconnesse che operano simultaneamente nell'esperienza del lettore. L'immersione spaziale crea un senso di presenza nell'ambiente narrativo, consentendo l'esplorazione mentale degli spazi testuali descritti e la costruzione di una geografia immaginaria in cui collocare l'azione narrativa. L'immersione temporale genera coinvolgimento nella progressione degli eventi, creando tensione narrativa e anticipazione del futuro sviluppo della trama. L'immersione emotiva facilita l'identificazione con i personaggi e la partecipazione em-

8 Si veda Landow 1992; Landow 1994.

patica alle loro esperienze, permettendo al lettore di sperimentare vicariamente stati emotivi complessi. Queste modalità non sono semplicemente additive, ma si integrano in un'esperienza ermeneutica complessa in cui ciascuna dimensione rafforza e modifica le altre. Le narrazioni più efficaci sono quelle che riescono a bilanciare questi diversi tipi di immersione, offrendo un'esperienza cognitivamente ed emotivamente ricca al fruitore.

L'interattività caratteristica dei media digitali, d'altro canto, può essere interpretata come un'evoluzione e materializzazione della concezione postmoderna del significato come processo costruttivo: si è passati da una letteratura orientata a creare un mondo immersivo per il lettore, a una letteratura che sollecita il lettore a diventare un partecipante attivo nella costruzione del significato, fino all'avvento dell'ipertesto e delle narrative interattive come forme di testualità che incorporano strutturalmente questa dimensione partecipativa. D'altra parte, è stato più volte osservato come l'interattività sia connessa a una concezione ludica della testualità, in cui il testo non è più visto soltanto come la rappresentazione di un "mondo" in cui immergersi passivamente, ma come un "gioco" con cui interagire attivamente. In questa prospettiva, l'utente assume il ruolo di un giocatore che esplora, manipola e modifica il testo, partecipando alla costruzione del suo significato. I testi digitali che privilegiano la dimensione ludica rispetto a quella mimetica adottano talvolta una disfunzionalità deliberata, mettendo in discussione la funzione rappresentativa del linguaggio e sottolineando invece la sua natura di sistema di regole manipolabili. L'interattività trasforma così il testo in uno spazio di gioco e sperimentazione, stimolando il fruitore a diventare parte attiva del processo di significazione, superando il ruolo tradizionale di osservatore passivo e distaccato.

La Realtà Virtuale, alla luce di queste considerazioni, trascende la sua natura puramente tecnologica per configurarsi come un dispositivo epistemologico con profonde implicazioni per la teoria letteraria e culturale. Essa permette di materializzare le strutture narrative in spazi navigabili, rendere tangibili le relazioni intertestuali attraverso connessioni spaziali, e visualizzare le stratificazioni semantiche del testo in forma di livelli percettivi sovrapposti. Questa trasformazione determina una riconfigurazione profonda delle modalità attraverso cui il significato testuale viene costruito e interpretato, offrendo nuove possibilità per l'analisi e la comprensione dei fenomeni letterari. Ma anche più rilevante è il valore didattico e di public engagement che essa può assumere, rendendo accessibili testi complessi a pubblici diversificati e potenziando le capacità

didattiche attraverso l'esperienza diretta e l'*embodiment*. Un caso emblematico è rappresentato dai numerosi progetti che utilizzano la VR per ricostruire il mondo della *Divina Commedia*, permettendo agli utenti di esplorare fisicamente gli spazi finzionali descritti da Dante e di materializzare le complesse geografie simboliche del poema rendendo immediata la percezione della sua struttura spaziale; una "tradizione" di cui il progetto D.A.N.T.E. (*Digital Archive and New Technologies for E-content*, <https://dante.dantelimina.it>), coordinato da Ciro Perna e Elisabetta Tonello, rappresenta uno tra i più stimolanti e fondati esempi.

Resta certamente aperta una domanda critica sul valore epistemico di tali applicazioni nel contesto degli studi e non della divulgazione (posto che oggi abbia senso porre tale netta distinzione): la questione se e in che misura queste esperienze immersive apportino conoscenze genuinamente innovative sui testi letterari, o se rappresentino principalmente riformulazioni di comprensioni già acquisite attraverso metodi interpretativi tradizionali. Se da un lato i mondi virtuali possono enfatizzare alcune dimensioni del testo (come la spazialità, il movimento e la materialità degli ambienti descritti), dall'altro rischiano di ridurne la complessità semantica, trasformando in "scene visive" concrete ciò che è primariamente di natura linguistica e concettuale, con tutte le ambiguità e polisemie che questa natura comporta. La visualizzazione, per quanto sofisticata, comporta sempre una selezione e una disambiguazione che può impoverire la ricchezza semantica del testo letterario. Ma a contrastare questo atteggiamento vagamente apocalittico, si può rilevare, da una parte, come la percezione *embodied* e immersiva permetta di evidenziare aspetti spaziali e visuali che lasciano tracce rade nel testo linguistico ma giocano un ruolo importante nei processi cognitivi di creazione e fruizione letteraria; e, dall'altra, come la storia della cultura, e della letteratura al suo interno, sia stata anche una storia di successive rimediazioni e traduzioni semiotiche⁹ in cui questi esperimenti si inseriscono in perfetta continuità.

2. Intelligenza Artificiale generativa e Large Language Models

Passiamo ora a considerare la seconda innovazione tecnologica che sta caratterizzando questi anni recenti: l'Intelligenza Artificiale generativa. Con questa formula ci si riferisce a una classe di sistemi computazio-

9 Si veda Bolter *et al.* 2003.

nali basati sull'impiego di reti neurali artificiali¹⁰ e metodi avanzati di machine learning,¹¹ caratterizzati dalla capacità di produrre contenuti originali a partire dall'analisi di vastissimi corpora con cui sono addestrati. Questa categoria di sistemi include i grandi modelli del linguaggio (Large Language Model, o LLM) come GPT, Claude, Gemini, DeepSeek e LLaMA che hanno rivoluzionato l'elaborazione del linguaggio naturale negli ultimi anni, e sono balzati agli onori delle cronache e al centro del dibattito scientifico. A questi si affiancano i modelli generativi per immagini come DALL-E, Midjourney e Stable Diffusion, che hanno esteso il paradigma generativo al dominio visuale, e ancora più recenti quelli per video come Sora.

Dal punto di vista tecnico, questi sistemi si basano su architetture di reti neurali di grandi dimensioni, addestrate in modo non supervisionato o *self-supervised* su quantità enormi di dati, prevalentemente provenienti dal Web. La loro caratteristica distintiva rispetto ai sistemi di reti neurali artificiali precedenti, essenzialmente dei classificatori, è che essi sono in grado di “generare” contenuti completamente nuovi. Un LLM generativo, infatti, produce testo in risposta a un input testuale (*prompt*), e non si limita (anzi ne è in genere incapace) a replicare o ricombinare gli esempi presenti nei dati di addestramento, bensì sfrutta la competenza linguistica e testuale derivata dall'apprendimento profondo per produrre output originali, spesso con risultati che sorprendono per qualità e coerenza semantica.

A livello astratto, un LLM può essere concettualizzato come un sistema capace di predire, su base probabilistica, la sequenza di token linguistici più appropriata a partire da una sequenza data. Più precisamente, il modello stima la distribuzione di probabilità per tutti gli elementi (detti token) del suo vocabolario, valutando la probabilità che ciascun token possa essere il successore immediato di una sequenza di input $S=t_1, \dots, t_k$. Poi il modello seleziona tra quelli più probabili (in base a metodi stocastici) un singolo token t_{k+1} , generando una nuova sequenza $S_1=S, t_{k+1}$ e reintroduce S_1 come input, in un processo autoregressivo che continua fino al raggiungimento di un punto di terminazione.

Alla base di questo processo computazionale vi è il concetto di *word embedding*, una tecnica che, analizzando un vasto numero di contesti

10 Si veda Mitchell 2022.

11 Si veda Alpaydin 2010.

d'uso, estrapola le proprietà sintattiche e semantiche di una parola, traducendola in un vettore di numeri reali all'interno di uno spazio multidimensionale. Il *word embedding* operationalizza le tesi della linguistica distribuzionale, sintetizzate nell'aforisma del linguista Firth: «You shall know a word by the company it keeps».¹² In questo paradigma, il significato di una parola è determinato dalla totalità dei contesti in cui essa appare, una intuizione che l'adozione di reti neurali ha tradotto in rappresentazioni della semantica lessicale sottoforma di vettori densi di numeri reali. Ma la semantica lessicale non sarebbe sufficiente a sviluppare modelli capaci di comprendere e generare interazioni linguistiche lunghe e complesse. Un aspetto cruciale nel funzionamento degli LLM basati sull'architettura a *transformer* è rappresentato dal meccanismo dell'*attenzione*, una tecnica matematica piuttosto sofisticata utilizzata per determinare il significato di una parola basandosi sul suo contesto linguistico attuale.¹³ In una rete neurale dotata di meccanismo di attenzione, ogni token di input viene elaborato indipendentemente, e il modello utilizza l'attenzione per identificare quali token del contesto sono più rilevanti per la comprensione del significato del token corrente. L'attenzione assegna pesi a ciascun token di input, indicandone l'importanza relativa nel determinare il significato del token in esame. Il risultato è una rappresentazione vettoriale del token che incorpora l'informazione contestuale, utilizzata poi dal modello per generare l'output appropriato.

La generazione dell'output testuale, infine, avviene attraverso un processo di decodifica multifase. Durante questo processo, il modello valuta ogni parola del suo vocabolario in base alla probabilità che essa segua in modo coerente la sequenza di parole esistente nel prompt. Per calcolare questa probabilità, l'LLM moltiplica il vettore di output dei transformer per la matrice delle parole nel suo vocabolario e applica un algoritmo *softmax* al vettore risultante per ottenere la distribuzione di probabilità normalizzata. La scelta finale della parola da aggiungere alla sequenza di input non è deterministica (*greedy decoding*), ma avviene selezionandone una in modo pseudo-casuale a partire da un pool di candidate che superano un certo valore di probabilità. Questo comportamento stocastico è regolato da un iper-parametro denominato "temperatura", che modula il grado di "libertà" e "creatività" linguistica del modello: tem-

12 Firth 1957, p. 179.

13 Si veda Vaswani *et al.* 2017.

perature più alte aumentano la diversità e l'imprevedibilità dell'output, mentre temperature più basse rendono le generazioni più deterministiche e conservative.

Le "capacità" linguistiche degli LLM derivano da un complesso processo di apprendimento articolato in più fasi. Si parte da un massiccio addestramento "auto-supervisionato" di base, in cui il modello apprende a predire la parola finale, opportunamente mascherata, di miliardi di frasi di testo estratte da enormi corpora linguistici. Questo stadio è seguito da fasi successive di "apprendimento supervisionato" su lunghe serie di coppie domanda/risposta prodotte da esperti umani (*Supervised Fine Tuning*), e di "apprendimento per rinforzo" (*Reinforcement Learning from Human Feedback*, RLHF), che costituiscono il *fine-tuning* del modello base. Questa fase è particolarmente importante per allineare il comportamento del modello con le preferenze culturali e i valori preferiti da chi sviluppa il modello, ottimizzando (ma anche censurando) la qualità, utilità e sicurezza delle sue generazioni.

Nonostante siano fondati su principi matematici ben definiti, gli LLM hanno manifestato proprietà emergenti non direttamente riconducibili a tali principi fondativi. Il complesso processo di addestramento su architetture neurali sufficientemente ampie sembra portare all'acquisizione di capacità non esplicitamente previste dall'obiettivo originale di predizione accurata del token successivo in una sequenza linguistica. Tra queste, particolarmente notevole è la capacità di apprendimento contestuale (*in-context learning*) ovvero la capacità di eseguire compiti per cui il modello non è stato specificamente addestrato,¹⁴ solo sulla base di semplici richieste (*zero-shot*), o al massimo di richieste corredate da pochi esempi (*few-shot*); gli LLM, insomma, dimostrano una apparente capacità di generalizzazione del tutto inattesa, anche se ci sono numerosissimi lavori teorici e sperimentali che argomentano a favore o contro queste presunte facoltà emergenti dei modelli (un dibattito che non è possibile approfondire in questa sede).

Gli LLM esibiscono altre notevoli meta-abilità che trascendono le semplici capacità linguistiche. Una capacità significativa è la 'sensibilità ai prompt', ovvero la capacità di essere influenzati dalle richieste specifiche degli utenti, dai suggerimenti impliciti e dalle implicazioni conversazionali. Ad esempio, strategie di *prompting* che chiedono al modello di

14 Si veda Brown *et al.* 2020.

‘agire come’ o ‘impersonare’ un ruolo specifico possono elicitarne risposte radicalmente diverse. Analogamente, richiedere al modello di “pensare passo dopo passo” attraverso tecniche di *Chain-of-Thought prompting* può attivare processi di ragionamento strutturato per compiti complessi, migliorando significativamente le performance in ambiti che richiedono deduzioni multiple. Sfruttando questa ultima capacità sono state sviluppate le più recenti evoluzioni di gran parte dei modelli rilasciati negli ultimi mesi sotto classificati come *Large Reasoning (Language) Models*.¹⁵

Questi fenomeni suggeriscono che gli LLM non sono semplici macchine generatrici di testo basate su correlazioni statistiche; durante l’interazione, essi sembrano impegnarsi, in una certa misura, in un dialogo governato da regole semantiche e pragmatiche complesse, mostrando una sensibilità al contesto e una flessibilità adattiva che richiamano, pur con significative differenze, alcuni aspetti della comunicazione linguistica umana.

3. Il dibattito sulla natura degli LLM

Sin dalle prime evidenze sperimentali delle sorprendenti capacità degli LLM (anche di generazioni precedenti come GPT-3, LaMDA, PaLM), si è acceso un intenso dibattito sulla loro natura e sulle loro reali capacità cognitive. Questo confronto, che vede coinvolti filosofi, informatici, linguisti, psicologi e anche letterati, come chi scrive, verte su questioni fondamentali che trascendono gli aspetti puramente ingegneristici di questi sistemi.

Un primo interrogativo concerne la competenza semantica degli LLM: questi modelli possiedono una vera comprensione del linguaggio, o si limitano a manipolare simboli senza accesso ai loro significati? Questa domanda riecheggia l’esperienza mentale della ‘stanza cinese’ proposto da John Searle negli anni ‘80 (Searle 1980), che metteva in discussione la possibilità che un sistema puramente sintattico potesse sviluppare una genuina comprensione semantica. Un secondo ordine di questioni riguarda la possibilità che gli LLM possiedano stati mentali analoghi a quelli umani: hanno credenze (convinzioni sulla verità o falsità delle proposizioni), desideri (stati motivazionali verso obiettivi) e intenzioni (impegni verso piani d’azione)? Collegata a questo punto è la questione dell’intenzionalità, intesa nella tradizione filosofica come la capacità di avere

¹⁵ Si veda Raschka 2023; Xu *et al.* 2025.

stati mentali “diretti verso” o “a proposito di” qualcosa, una proprietà che tradizionalmente distingue la mente umana dai sistemi puramente meccanici. Infine, si discute se gli LLM possiedano una forma di *agency*, ovvero la capacità di agire in modo autonomo e diretto a scopi, piuttosto che essere semplici strumenti passivi guidati dalle istruzioni umane. Le posizioni in questo dibattito complesso e multidisciplinare possono essere schematicamente divise tra approcci deflazionisti, che tendono a minimizzare le capacità cognitive degli LLM interpretandoli come sistemi di manipolazione simbolica privi di genuina comprensione, e approcci possibilisti o ottimisti, che riconoscono in questi modelli forme emergenti di cognizione che, pur diverse da quelle umane, presentano caratteristiche di genuina comprensione semantica e capacità rappresentazionali.

L'articolo più influente della corrente deflazionista è probabilmente *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* di Bender, Gebru, McMillan-Major e Mitchell (2021). Questo *position paper* articola diverse critiche agli LLM, spaziando da considerazioni teoriche a preoccupazioni etiche, politiche e ambientali. Gli aspetti più propriamente teorici riprendono un precedente lavoro di Bender e Koller (2020), e la tesi centrale sostenuta da questi autori è che il testo generato da un LLM non è fondato su un'autentica intenzione comunicativa, né su un modello del mondo o una rappresentazione dello stato mentale dell'interlocutore. Questa impossibilità deriva dal fatto che i dati di addestramento non includono esperienze di condivisione di pensieri con un ascoltatore in un contesto di comunicazione genuina, né l'architettura del modello possiede i requisiti strutturali per sviluppare tali capacità. Secondo questa visione, sebbene il testo generato appaia sempre più fluente e coerente, la percezione di significato che ne deriva è interamente mediata dalla competenza linguistica umana e dalla nostra predisposizione cognitiva a interpretare atti comunicativi come portatori di significato e intenzione coerenti, indipendentemente dalla loro origine. Se una delle parti della comunicazione (il modello) non possiede significato intrinseco, la comprensione di un significato implicito diventa un'illusione derivante dalla nostra specifica comprensione umana del linguaggio e dalla nostra tendenza a proiettare intenzionalità anche dove questa è assente. In questa prospettiva critica, un LLM è essenzialmente un sistema che “cuce insieme in modo casuale sequenze di forme linguistiche” osservate nei dati di addestramento, basandosi su informazioni probabilistiche sulla loro combinazione, ma senza alcun riferimento al significato: un ‘pappagallo stocastico’ che imita senza comprendere.

A questa posizione deflazionista (oltre che politicamente allarmista) si contrappongono diverse visioni alternative che, pur riconoscendo le limitazioni degli attuali LLM, offrono interpretazioni più sfumate e complesse del loro funzionamento e delle loro potenzialità, almeno in prospettiva.¹⁶ Anche questo dibattito esula dagli scopi e dai limiti di questo articolo, ma riassumendo in modo sintetico la mia posizione, vorrei in primo luogo osservare come la strategia argomentativa alla base dell'argomento del 'pappagallo stocastico' soffra di un eccessivo ricorso a quello che chiamerei l'operatore linguistico 'VERO', di solito inserito in questo schema di proposizione: sia x una facoltà cognitiva umana a piacere; se le performance di un sistema artificiale A su x sono paragonabili (in un qualche senso operationalizzabile) a quelle di un sistema naturale N , allora A è solo in grado di mimare o simulare la "VERA" competenza x di N . Si noti che di solito l'argomentazione non si fornisce alcuna seria spiegazione scientifica su che cosa sia questa "VERA" x .

Si tratta con ogni evidenza di una fallacia del bersaglio mobile (*moving the goalposts fallacy*) – alzare continuamente lo standard di una definizione o di una capacità quando questa viene raggiunta o soddisfatta dal sistema che si sta analizzando, senza mai accettare che l'obiettivo iniziale sia stato davvero raggiunto – fondata su una forma di dualismo cartesiano implicito. Dal mio punto di vista, ritengo che un inquadramento teorico più articolato e intellettualmente produttivo ci venga fornito dall'elaborazione filosofica di Daniel Dennett: considerare gli LLM come 'sistemi intenzionali', ovvero sistemi che possono essere interpretati adottando l'*atteggiamento intenzionale*,¹⁷ la strategia di interpretare il comportamento di un'entità trattandola come se fosse un agente razionale che governa la sua scelta di azione mediante una considerazione delle sue 'credenze' e dei suoi 'desideri'. Non intendo ovviamente sostenere che i modelli linguistici attuali siano effettivamente e intrinsecamente dotati di tutte le proprietà che vorremmo attribuire alla mente umana (facoltà di linguaggio, raziocinio, modello del mondo, teoria della mente, coscienza, empatia...), ma che non c'è nulla di misterioso e irriducibile o nascosto in una insondabile interiorità che permetta di spiegare tali proprietà: tutto ciò che abbiamo sono i comportamenti del sistema, e se l'esame di tali

16 Si veda Chalmers 2023; Lederman-Mahowald 2024; Piantadosi 2024; Millière-Buckner 2024.

17 Si vedano Dennett 1997 e Dennet 1989.

comportamenti ci induce ragionevolmente ad attribuire una di tali facoltà a un agente artificiale, ciò è quanto basta per tale ascrizione. Come dice lo stesso Dennett in un passo della sua recente autobiografia,

la scienza è di per sé una specie di behaviorismo: e una volta che, dato un fenomeno, tutti i comportamenti hanno trovato una spiegazione plausibile, quelli interiori come quelli esterni, quelli macro come quelli micro, non rimane più nulla da spiegare – se non forse perché certa gente si trovasse a disagio per quella spiegazione.¹⁸

4. Gli LLM come modelli dei sistemi culturali

La prospettiva che considera i nuovi sistemi di IA generativa come agenti cognitivi individuali (e che dunque cerca di valutare se essi siano dotati o meno delle proprietà cognitive dei sistemi intelligenti naturali), non è la sola adottabile per cercare di comprendere la loro natura. Anzi, nell'ottica di questo lavoro, che ha lo scopo di indagare il loro impatto nelle Digital Humanities, potrebbe essere più utile e produttivo indagare come gli LLM si relazionino con l'insieme della conoscenza culturale multimodale con cui vengono addestrati. In questa direzione, pur rimanendo fortemente ancorata alla fazione deflazionista, si muove la psicologa dello sviluppo Alison Gopnik. Nel suo articolo *Transmission Versus Truth, Imitation Versus Innovation: What Children Can Do That Large Language and Language-and-Vision Models Cannot (Yet)*,¹⁹ propone di considerare gli LLM essenzialmente come tecnologie culturali, paragonabili, seppur con modalità operative diverse, ai sistemi tradizionali di gestione e disseminazione della conoscenza come le biblioteche, la stampa, e i media etc. Secondo questa visione, gli LLM operano campionando dal patrimonio culturale umano, pur senza possedere una vera capacità di comprensione dei contenuti, ma a differenza degli apparati medialti del passato, non si limitano alla replicazione *verbatim* dei contenuti che memorizzano, ma hanno la capacità di ricombinarli in modi nuovi e pertinenti al contesto. Gopnik sottolinea la distinzione fondamentale tra imitazione e innovazione: i modelli di intelligenza artificiale eccellono nell'imitazione, intesa come capacità di estrarre e replicare pattern esistenti nei dati di addestramento, ma mostrano significative limitazioni nell'innovazione genuina,

18 Dennett 2024, p. 201.

19 Yiu-Kosoy-Gopnik 2023.

che richiede una comprensione profonda e la capacità di interagire creativamente con il mondo fisico e sociale. Secondo questa visione, gli LLM non possiedono la capacità di sviluppare autonomamente nuove idee, non producono nuovi significati e il significato delle loro espressioni è totalmente derivativo, dipendente dai significati presenti nei testi su cui sono stati addestrati; ma proprio per questo si rivelano come dei formidabili strumenti di osservazione e analisi degli universi culturali. Una posizione sostenuta anche da Ted Underwood che osserva come

the immediate value of these models is often not to mimic individual language understanding, but to represent specific cultural practices (like styles or expository templates) so they can be studied and creatively remixed [...] Models of culture are exactly what we need.²⁰

Penso che un utile quadro di riferimento teorico in cui collocare questa modalità di considerare i modelli di IA generativa si possa rintracciare nella teoria semiotica della cultura di Jurij Lotman e della scuola di Tartu.²¹ Il concetto chiave del pensiero lotmaniano a questo riguardo è quello di *semiosfera* (1992): lo spazio semiotico globale all'interno del quale i processi di significazione (semiosi) possono avvenire e la cultura può esistere. Lotman traccia un'analogia potente con la biosfera: come quest'ultima è la condizione necessaria per la vita biologica, così la semiosfera costituisce l'ambiente fondamentale senza il quale nessun segno, testo o sistema di significazione potrebbe emergere o avere senso. Non si tratta quindi di un semplice archivio di segni o testi, ma dell'insieme totale delle loro relazioni e del contesto spaziale e funzionale che li rende possibili e intelligibili. La semiosfera è concepita come uno spazio delimitato, con dei confini che separano il suo 'interno' – lo spazio della cultura specifica, dove i segni sono organizzati e comprensibili – dall' 'esterno', che può essere percepito come non-semiotico, caotico, o appartenente a un'altra semiosfera. Tuttavia, questi confini non sono affatto barriere passive o impermeabili; al contrario, sono zone di intensa attività semiotica. È proprio sulla frontiera che avvengono i processi cruciali di traduzione, filtro, scambio e adattamento tra sistemi diversi, rendendola un'area fondamentale per la dinamica e l'innovazione culturale. Anche internamente, la semiosfera non è uniforme; la sua caratteristica princi-

20 Underwood 2021.

21 Si veda Lotman-Uspenskij 2001.

pale è l'eterogeneità. Al suo interno coesistono e interagiscono una molteplicità di linguaggi, codici e sottosistemi semiotici differenti. Questa struttura interna è anche asimmetrica, organizzata tipicamente attorno a un 'centro' – che rappresenta le strutture dominanti, più stabili e normative della cultura – e una 'periferia', più dinamica, instabile, aperta al cambiamento e al contatto con l'esterno e con elementi marginali interni. Il dialogo e la tensione tra centro e periferia, e tra i diversi sistemi che la compongono, sono motori essenziali dello sviluppo culturale.

La pluralità dei linguaggi, l'intertestualità, la memoria culturale, i processi di traduzione del significato tra sistemi semiotici diversi, l'eterogeneità e perfino le "esplosioni culturali" (momenti di imprevedibile innovazione semiotica) descritti da Lotman trovano sorprendenti corrispondenze nel funzionamento e nei comportamenti dei modelli linguistici: un LLM può essere considerato come un sistema che costruisce ed elabora un modello della semiosfera culturale sottoforma di spazio vettoriale multidimensionale, in cui le relazioni tra i segni e concetti sono rappresentate isomorficamente come proprietà algebriche e geometriche dello spazio vettoriale. Questi modelli, infatti, costruiscono uno spazio semantico dinamico e pluralistico che può essere interpretato come una «superposizione di prospettive culturali»,²² in cui i significati non sono fissati a priori, ma vengono generati, trasformati e connessi durante la produzione di atti linguistici in risposta a specifici prompt e contesti.

È particolarmente significativo notare che Lotman stesso, in un articolo del 1979 intitolato *Culture as Collective Intellect and Problems of Artificial Intelligence*, aveva prefigurato la possibilità di concepire la semiosfera come un intelletto collettivo, suggerendo che questo modello di intelligenza distribuita e collettiva potesse rappresentare un paradigma più esplicito e comprensibile dell'intelletto individuale per lo sviluppo dell'intelligenza artificiale. Infatti, le strutture e i processi dell'intelletto collettivo sono manifestate nel linguaggio della cultura e registrate nei testi, mentre i processi del pensiero individuale rimangono in larga parte inaccessibili all'osservazione diretta:

We must emphasize that collective intellect, as a model for artificial intellect, has several advantages over individual intellect. Since collective intellect is a mechanism created by the history of mankind, it is far more explicit, its procedures are manifest in the language

22 Si veda Kovač *et al.* 2023.

of culture and are recorded in numerous texts, unlike the hidden languages of the human brain.²³

Questa intuizione lotmaniana appare sorprendentemente profetica alla luce degli sviluppi recenti degli LLM, che possono essere interpretati come implementazioni computazionali, per quanto parziali e imperfette, di questa idea di intelletto collettivo materializzato nel linguaggio. E da questo deriva la loro innovativa potenzialità epistemica nel dominio dell'analisi dell'ecosistema culturale. Infatti, gli oggetti e i processi che popolano la semiosfera sono tipicamente "resistenti" all'approccio formalista tradizionale nei metodi computazionali; operazionalizzare e formalizzare (o assiomatizzare) i fenomeni culturali in modo rigoroso si è dimostrato estremamente difficile, se non impossibile, all'interno dei paradigmi computazionali classici. Il paradigma di costruzione della conoscenza sottostante ai sistemi di IA generativa, basato sull'apprendimento distribuito e sull'emergenza di capacità di alto livello da processi computazionali di basso livello, si rivela potenzialmente assai più adatto a trattare questa sfera della realtà rispetto al paradigma formalista basato su regole esplicite e rappresentazioni simboliche. Gli LLM, modellizzando interi sistemi semiotico-culturali attraverso la rappresentazione vettoriale dei token linguistici e delle loro relazioni contestuali, permettono di esternalizzare i processi interpretativi degli oggetti semiotici, distaccandoli dalla dimensione soggettiva della mente ermeneutica umana a cui erano tradizionalmente legati. Questo spostamento dell'attività interpretativa dall'interiorità del soggetto umano all'esteriorità di un sistema computazionale osservabile apre nuove possibilità per lo studio empirico dei processi interpretativi stessi, trasformando l'ermeneutica da pratica puramente riflessiva a oggetto di indagine sperimentale.

5. Le frontiere: convergenze tra IA e metaverso

Abbiamo finora trattato separatamente le due linee di innovazione che ci offrono le tecnologie digitali e le loro conseguenze epistemologiche nello studio dei fenomeni culturali e letterari. Ma già oggi si intravedono le potenzialità che possono essere offerte dalla convergenza tra tecnologie dell'immagine, sistemi immersivi e intelligenza artificiale generativa. Potenzialità che si collocano all'interno della dimensione ludica della frui-

23 Lotman 1979, p. 84.

zione testuale, cui già abbiamo fatto riferimento prima, ma che non per questo sono prive di contenuto conoscitivo, come già aveva ampiamente mostrato, da filologo, Jerome McGann sin dagli anni 90, agli albori dell'era del Web, con i suoi numerosi esperimenti di deformazione o *gamification* dei testi letterari presentati poi nel suo libro *Radiant Textuality*.²⁴

Basti pensare alla possibilità di esplorare visualmente i mondi e gli spazi dei testi mediante processi di ecfasi inversa prodotti da modelli generativi multimodali (o *text-to-image*): il gioco esplorativo, in questo caso, diventa strumento operativo di analisi testuale, permettendo di visualizzare interpretazioni multiple dello stesso passaggio descrittivo e rivelando così la pluralità delle letture possibili inscritta nel testo stesso. In modo reciproco, queste metodologie consentono di indagare le diversità delle semiosfere simulate da diversi modelli, di analizzare le divergenze tra le rappresentazioni visive da essi generate a partire dallo stesso input testuale, evidenziando come differenti architetture AI 'leggano' e interpretino il testo in modi distinti, rivelando presupposti impliciti, *bias* interpretativi e differenze nelle priorità semantiche dei vari modelli. Ancora più significativamente, l'ecfasi inversa computazionale permette di studiare sistematicamente la relazione tra descrizione verbale e rappresentazione visiva, aprendo nuove prospettive sulla traduzione intersemiotica, ovvero il passaggio da un sistema di segni (linguistico) a un altro (visivo). La generazione multipla di immagini a partire da un medesimo passaggio testuale crea uno "spazio interpretativo visivo" che può essere analizzato per comprendere le ambiguità semantiche del testo originale, gli elementi descrittivi dominanti e recessivi, e le inferenze culturali e contestuali applicate dai modelli generativi. Ciò consente di evidenziare quali elementi testuali vengono privilegiati o trascurati nella trasposizione visiva, rivelando così le gerarchie implicite nella percezione e interpretazione del testo.

Un'evoluzione significativa di questo approccio è rappresentata dal passaggio dalla generazione di immagini statiche alla produzione di sequenze video a partire da descrizioni testuali. L'introduzione di modelli come Sora di OpenAI (<https://openai.com/sora>) rappresenta un salto paradigmatico in questa direzione, aprendo nuove frontiere per lo studio del testo narrativo. Il video generativo introduce la temporalità come dimensione interpretativa fondamentale, permettendo di visualizzare il ritmo narrativo, le transizioni temporali e le rappresentazioni dinamiche

24 Si veda McGann 2001.

di focalizzazione e punto di vista che nel testo rimangono implicite o sono comunicate attraverso complessi artifici retorici. Una difficoltà particolare per i modelli generativi video consiste nella necessità di mantenere una continuità visiva e narrativa coerente, gestendo implicitamente elementi come causalità e consequenzialità che sono essenziali nella narrazione. Questa necessità li costringe a sviluppare forme di “comprensione” narrativa le cui potenziali applicazioni analitiche sono numerose: lo studio di sequenze descrittive in movimento; l’analisi della costruzione progressiva dello spazio narrativo nel tempo; l’esplorazione delle modalità di transizione tra scene e prospettive narrative diverse.

I *world models* rappresentano infine il paradigma più ambizioso nell’evoluzione dell’IA generativa, transcendendo la semplice generazione di contenuti per mirare alla comprensione e modellazione di interi “mondi” con le loro regole, relazioni e dinamiche interne.²⁵ Un world model genuino dovrebbe possedere una comprensione profonda delle proprietà fisiche e causali del mondo rappresentato: un modello con una rappresentazione di base di un ‘mondo’ (ad esempio, un video di una stanza sporca), dato un obiettivo (una stanza pulita), dovrebbe essere in grado di ideare una sequenza di azioni per raggiungerlo (usare l’aspirapolvere, lavare i piatti, svuotare il cestino) non perché ha osservato specificamente quel pattern nei dati di addestramento, ma perché comprende a un livello più profondo la relazione causale tra sporco e pulizia, e le azioni che permettono di passare da uno stato all’altro. Questi modelli sono addestrati a comprendere e ricostruire le proprietà fondamentali dei mondi rappresentati, inferendo regole, relazioni causali e strutture ontologiche che li governano. Hanno il potenziale per generare ambienti coerenti e interattivi in cui le regole fisiche, sociali e narrative sono consistenti e prevedibili, permettendo forme di simulazione e sperimentazione precedentemente impossibili. L’applicazione di questi world models agli universi narrativi potrebbe consentire la creazione di spazi di simulazione in cui esplorare conseguenze controfattuali, sviluppi narrativi alternativi e interpretazioni divergenti delle ambiguità testuali. Il futuro prossimo delle tecnologie convergenti suggerisce la possibilità di generare interi mondi narrativi navigabili: ambienti virtuali generati dinamicamente a partire da testi letterari, interazioni naturali con personaggi dotati di agency basata su LLM, e ricostruzioni coerenti di ambientazioni, atmosfere e rela-

25 Si veda Ha-Schmidhuber 2018; Garrido *et al.* 2024.

zioni causali testuali. Questa possibilità solleva questioni fondamentali per gli studi culturali e letterari, a partire dal problema della validità epistemica di tali ricostruzioni. Questioni la cui soluzione richiede un serio dialogo interdisciplinare in grado di sviluppare approcci che siano sia tecnicamente solidi sia ermeneuticamente produttivi.

6. Limiti e prospettive

Nonostante le potenzialità innovative che abbiamo ampiamente discusso, è fondamentale riconoscere che i modelli e sistemi generativi attuali presentano significativi limiti epistemici che ne condizionano l'applicabilità nel contesto degli studi letterari e culturali. Un primo limite riguarda l'affidabilità epistemica: questi modelli non producono conoscenza epistemicamente certa e sono soggetti a output erronei o inventati, impropriamente definiti 'allucinazioni' nel gergo tecnico. Queste 'allucinazioni' non sono semplici errori casuali, ma derivano dalle modalità stesse di funzionamento dei modelli probabilistici, che possono generare contenuti plausibili ma fattualmente errati o completamente fittizi, un fenomeno particolarmente problematico nel contesto della ricerca accademica, dove l'accuratezza fattuale è un requisito imprescindibile. Ma ancora più critica è una limitazione strutturale di questi sistemi: la loro "opacità" epistemica. Infatti, non è possibile spiegare in modo certo e completo come funzionano i loro processi inferenziali interni, quali rappresentazioni intermedie costruiscono e quali criteri utilizzano per selezionare determinate interpretazioni rispetto ad altre. Questa caratteristica, nota come il problema della "black box" dell'intelligenza artificiale, pone sfide significative alla validazione scientifica dei risultati ottenuti e alla loro integrabilità all'interno di metodologie critiche tradizionali che valorizzano la trasparenza del processo interpretativo. La validazione delle conoscenze generate dagli LLM rappresenta una sfida epistemologica complessa che richiede approcci diversificati. L'Explainable AI²⁶ e l'interpretazione meccanicistica dei modelli neurali²⁷ offrono potenziali vie per rendere più trasparenti i processi interni di questi sistemi, ma incontrano limiti significativi: problemi di calcolabilità dovuti alla dimensione dei modelli, il fenomeno della

26 Si veda Barredo *et al.* 2020.

27 Si veda Millière 2024; Bereska-Gavves 2024.

sovrapposizione (*superposition*) che rende difficile isolare rappresentazioni specifiche all'interno delle reti neurali, e il problema fondamentale del livello di descrizione appropriato.²⁸

Rimane molto da riflettere e sperimentare sulla natura dei sistemi generativi e sul loro ruolo come potenziali modelli della cognizione umana e dei processi culturali. Ma, indipendentemente da quanto si dimostreranno modelli prossimali della mente umana (una questione che rimane aperta e dibattuta), la loro capacità di modellizzare i sistemi astratti apre enormi spazi di sperimentazione sui fenomeni culturali, letterari e artistici. Le potenzialità analitiche ed ermeneutiche emergenti dalla convergenza tra mondi digitali immersivi e intelligenza artificiale generativa aprono un vasto spazio inesplorato per gli studiosi umanisti capaci di attraversare in entrambe le direzioni i ponti dell'interdisciplinarietà, mantenendo al contempo il rigore metodologico delle proprie discipline d'origine e l'apertura verso linguaggi e paradigmi disciplinari diversi. In questa prospettiva, la convergenza non rappresenta semplicemente l'introduzione di nuovi strumenti nell'arsenale metodologico delle discipline umanistiche, ma l'emergere di un nuovo paradigma epistemologico che richiede una profonda riconsiderazione dei fondamenti teorici e metodologici di queste discipline. Questo processo di ripensamento, lungi dall'impovertire la tradizione umanistica, ha il potenziale per rivitalizzarla e arricchirla, aprendo nuove prospettive su questioni fondamentali relative alla natura del significato, dell'interpretazione, della narratività e dell'esperienza estetica.

Bibliografia

- Alpaydin 2010 = Ethem Alpaydin, *Introduction to Machine Learning*. 2nd ed. Adaptive Computation and Machine Learning, Cambridge (Mass), MIT Press, 2010. <http://mitpress.mit.edu/books/introduction-machine-learning>
- Bachtin 1997 = Michail Michailovič Bachtin, *Estetica e romanzo. Teoria e storia del discorso narrativo*, Milano, Einaudi, 1997.
- Barredo *et al.* 2020 = Alejandro Arrieta Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barba, Salvador García, *Explainable Artificial Intelligence (XAI): Concepts,*

28 Si veda Elhage *et al.* 2022; Scherlis *et al.* 2022.

- taxonomies, opportunities and challenges toward responsible AI*, «Information Fusion», 58 (2020), pp. 82–115.
- Bender *et al.* 2021 = Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, e Shmargaret Shmitchell, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, Association for Computing Machinery, 2021, pp. 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bender-Koller 2020 = Emily M. Bender, Alexander Koller, *Climbing towards NLU: On meaning, form, and understanding in the age of data*, in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 5185-5198. Online: Association for Computational Linguistics.
- Bareska-Gavves 2024 = Leonard Bareska, Efstratios Gavves, *Mechanistic Interpretability for AI Safety – A Review*, arXiv preprint arXiv:2404.14082. <https://doi.org/10.48550/ARXIV.2404.14082>
- Bolter-Grusin 2023 = Jay D. Bolter, Richard Grusin, *Remediation. Competizione e Integrazione Tra Media Vecchi e Nuovi*, a cura di A. Marinelli, trad. in it. da B. Gennaro, Milano, Guerini e Associati, 2023.
- Bould 2005 = Mark Bould, *Cyberpunk*, in *A Companion to Science Fiction*, a cura di David Seed, Oxford, UK: Blackwell Publishing Ltd, pp. 217-238. <https://doi.org/10.1002/9780470997055.ch15>
- Brown *et al.* 2020 = Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei, *Language Models are Few-Shot Learners*, in *Advances in Neural Information Processing Systems 33*, a cura di H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, e H. Lin, Curran Associates, pp. 1877–1901. https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf
- Caronia-Gallo 1996 = Antonio Caronia, Domenico Gallo, *Houdini e Faust: breve storia del Cyberpunk*, Milano, Baldini & Castoldi, 1996.
- Chalmers 2023 = David J. Chalmers, *Could a Large Language Model Be Conscious?*, «Boston Review», 1 (2023). <https://www.bostonreview.net/articles/could-a-large-language-model-be-conscious/>

- Ciotti 2023a = Fabio Ciotti, *Digital humanities. Metodi, strumenti, saperi*, Roma, Carocci editore, 2023.
- Ciotti 2023b = Fabio Ciotti, *Minerva e il pappagallo. IA generativa e modelli linguistici nel laboratorio dell'umanista digitale*, «Testo e Senso» 26 (2023), pp. 289–315. <https://doi.org/10.58015/2036-2293/671>
- Dennet 1989 = Daniel C. Dennet, *The Intentional Stance*, Cambridge (Mass), MIT Press, 1989.
- Dennet 1997 = Daniel C. Dennet, *True Believers: The Intentional Strategy and Why It Works*, in *Mind Design II*, a cura di John Haugeland, Cambridge (Mass), MIT Press, pp. 57-80. <https://doi.org/10.7551/mitpress/4626.003.0003>
- Dennet 2024 = Daniel C. Dennet, *Pensandoci bene: avventure nella filosofia*. Milano, Raffaello, 2024.
- Dolézel 1999 = Lubomír Dolézel, *Heterocosmica: fiction e mondi possibili*, Milano, Bompiani, 2024.
- Eco 1979 = Umberto Eco, *Lector in fabula*, Milano, Bompiani, 1979.
- Ehlage *et al.* 2022 = Nelson Elhage, Tristan Hume, Catherine Olsson, Nicholas Schiefer, Tom Henighan, Shauna Kravec, Zac Hatfield-Dodds, Robert Lasenby, Dawn Drain, Carol Chen, Roger Grosse, Sam McCandlish, Kaplan, Dario Amodei, Martin Wattenberg, Christopher Olah, *Toy Models of Superposition Transformer Circuit Thread*, 2022. https://transformer-circuits.pub/2022/toy_model/index.html
- Firth 1957 = John Rupert Firth, *Papers in Linguistics 1934–1951*, London, Oxford University Press, 1957.
- Garrido *et al.* 2024 = Quentin Garrido, Mahmoud Assran, Nicolas Ballas, Adrien Bardes, Laurent Najman, Yann LeCun, *Learning and Leveraging World Models in Visual Representation Learning*, arXiv (<https://doi.org/10.48550/ARXIV.2403.00504>).
- Gibson 2019 = William Gibson, *Trilogia dello Sprawl*, trad. in it. da G. Cossato, Milano, Mondadori, 2019.
- Ha-Schmidhuber 2018 = David Ha, Jürgen Schmidhuber, *World Models*, arXiv 1803.10122. <https://doi.org/10.48550/ARXIV.1803.10122>
- Kovač *et al.* 2023 = Grgur Kovač, Masataka Sawayama, Rémy Portelas, Cédric Colas, Peter Ford Dominey, Pierre-Yves Oudeyer, *Large Language Models as Superpositions of Cultural Perspectives*, arXiv:2307.07870. <https://doi.org/10.48550/ARXIV.2307.07870>
- Landow 1992 = George P. Landow, *Hypertext: The Convergence of Contemporary Critical Theory and Technology*, Baltimore, Johns Hopkins University Press, 1992.

- Landow 1994 = *Hyper/text/theory*, a cura di George P. Landow, Baltimore, Johns Hopkins University Press, 1994.
- Lederman-Mahowald 2024 = Harvey Lederman, Kyle Mahowald, *Are Language Models More Like Libraries or Like Librarians? Bibliotechnism, the Novel Reference Problem, and the Attitudes of LLM*, arXiv. <https://doi.org/10.48550/arXiv.2401.04854>
- Lotman 1979 = Jurij Michajlovič Lotman, *Culture as collective intellect and the problems of artificial intelligence*, «Russian Poetics in translation», 6 (1979), pp. 84–96.
- Lotman 1992 = Jurij Michajlovič Lotman, *La semiosfera: l'asimmetria e il dialogo nelle strutture pensanti*, a cura di S. Salvestroni, Venezia, Marsilio, 1992.
- Lotman-Uspenskij 2001 = Jurij Michajlovič Lotman, Boris Andreevič Uspenskij, *Tipologia della cultura*, Milano, Bompiani, 2001.
- McGann 2001 = Jerome McGann, *Radiant Textuality: Literature after the World Wide Web*, Palgrave, Macmillan, 2001.
- Milgram, Kishino 1994 = Paul Milgram, Fumio Kishino, *A Taxonomy of Mixed Reality Visual Displays*, «IEICE Trans. Information Systems E77-D», 12 (1994), pp. 1321-1329.
- Millière 2024 = Raphaël Millière, *Philosophy of Cognitive Science in the Age of Deep Learning*, «WIREs Cognitive Science», 15/5 2024, e1684. <https://doi.org/10.1002/wcs.1684>
- Millière-Buckner 2024 = Raphaël Millière, Cameron Buckner, *A Philosophical Introduction to Language Models – Part I: Continuity With Classic Debates*, arXiv. <https://doi.org/10.48550/ARXIV.2401.03910>
- Mitchell 2022 = Melanie Mitchell, *L'intelligenza artificiale. Una guida per esseri umani pensanti*, Torino, Einaudi, 2022.
- Pavel 1986 = Thomas G. Pavel, *Fictional Worlds*, Cambridge (Mass), Harvard University Press, 1986.
- Piantadosi 2024 = Steven T. Piantadosi, *Modern language models refute Chomsky's approach to language*, in *From fieldwork to linguistic theory: A tribute to Dan Everett*, a cura di E. Gibson e M. Poliak, Berlin, Language Science Press, 2024.
- Pimentel-Teixeira 1993 = Ken Pimentel, Kevin Teixeira, *Virtual Reality: Through the New Looking Glass*, New York, Intel/Windcrest, 1993. <https://books.google.it/books?id=ErrumAEACAAJ>

- Raschka 2023 = Sebastian Raschka, *Understanding Reasoning LLMs*, in *Ahead of AI* (blog). 16 aprile 2023. <https://magazine.sebastianraschka.com/p/understanding-reasoning-llms>
- Rheingold 1991 = Howard Rheingold, *Virtual reality*, New York, Summit Books, 1991.
- Roncaglia 2023 = Gino Roncaglia, *L'architetto e l'oracolo. Forme digitali del sapere da Wikipedia a ChatGPT*, Roma-Bari, Laterza, 2023.
- Ryan 1991 = Marie-Laure Ryan, *Possible Worlds, Artificial Intelligence, and Narrative Theory*, Bloomington, Indiana University Press, 1991.
- Ryan 2003 = Marie-Laure Ryan, *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*, Baltimore, Johns Hopkins University Press, 2003.
- Ryan 2015 = Marie-Laure Ryan, *Narrative as Virtual Reality 2: Revisiting Immersion and Interactivity in Literature and Electronic Media*, Baltimore, Johns Hopkins University Press, 2015.
- Scherlis *et al.* 2022 = Adam Scherlis, Kshitij Sachan, Adam S. Jermyn, Joe Benton, Buck Shlegeris, *Polysemanticity and Capacity in Neural Networks*, arXiv. <https://doi.org/10.48550/ARXIV.2210.01892>
- Searle 1980 = John R Searle, *Minds, Brains, and Programs*, «Behavioral and Brain Sciences», 3/3 (1980), pp. 417-424. <https://doi.org/10.1017/S0140525X00005756>).
- Stephenson 1992 = Neal Stephenson, *Snow crash*, New York, Bantam Books, 1992.
- Underwood 2021 = Ted Underwood, *Mapping the Latent Spaces of Culture*, in *The Stone and the Shell* (blog). 21 ottobre 2021. <https://tedunderwood.com/2021/10/21/latent-spaces-of-culture/>
- Vaswani *et al.* 2017 = Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, e Illia Polosukhin, *Attention Is All You Need*, arXiv:1706.03762. <https://doi.org/10.48550/ARXIV.1706.03762>
- Walton 1993 = Kendall L. Walton, *Mimesis as Make-Believe: On the Foundations of the Representational Arts*, Cambridge (Mass), Harvard University Press, 1993.
- Xu *et al.* 2025 = Fengli Xu, Qianyue Hao, Zefang Zong, Jingwei Wang, Yunke Zhang, Jingyi Wang, Xiaochong Lan, Jiahui Gong, Tianjian Ouyang, Fanjin Meng, Chenyang Shao, Yuwei Yan, Qinglong Yang, Yiwen Song, Sijian Ren, Xinyuan Hu, Yu Li, Jie Feng, Chen Gao, Yong Li, *Towards Large Reasoning Models: A Survey of Reinforced*

Reasoning with Large Language Models, arXiv. <https://doi.org/10.48550/ARXIV.2501.09686>

Yiu-Kosoy-Gopnik 2023 = Eunice Yiu, Eliza Kosoy, Alison Gopnik, *Transmission versus truth, imitation versus innovation: What children can do that large language and language-and-vision models cannot (yet)*, «Perspectives on Psychological Science», 19/5 (2025). <https://doi.org/10.1177/17456916231201401>



10.69118/2025_L2

Progettazione e modellazione multilineare per edizioni scientifiche digitali

Cristina Marras

*Consiglio Nazionale delle Ricerche,
Istituto per il Lessico Intellettuale Europeo e Storia delle Idee, Roma
cristina.marras@cnr.it*

Multilinear design and modelling for digital scientific editions

Abstract (ITA)

Questo breve contributo si concentra su aspetti di metodo e sull'interrelazione tra dimensioni teoriche e pratiche nel campo dell'Umanistica Digitale, con particolare riguardo alla filosofia. La discussione riguarda nello specifico le Edizioni Scientifiche Digitali, come un interessante e fruttuoso campo di riflessione, ricerca e applicazione. Attraverso una serie di esempi viene descritto come il rapporto tra analogico e digitale si traduce nelle attuali pratiche di edizione e negli obiettivi generali di fornire versioni affidabili, attendibili e utili del patrimonio testuale e documentale. Ma non solo, la mutua trasformazione tra discipline umanistiche e scienze computazionali (l'Umanistica Digitale) trova piena realizzazione nella ricerca e nello sviluppo di piattaforme collaborative e infrastrutture dinamiche e integrate.

Abstract (ENG)

This brief article focuses on aspects of method and the interrelationship between theoretical and practical dimensions in the field of Digital Humanities, with reference to philosophy. The discussion will focus specifically on digital scholarly editions as an interesting and fruitful field of reflection, research and application, and will describe, through a series of examples, how the relationship between analogue and digital translates into current publishing practices and the general goals of providing reliable, trustworthy and useful versions of textual and documentary heritage. In addition, the mutual transformation between the humanities and computational sciences (Digital Humanities) is fully realised in the research and development of collaborative platforms and dynamic, integrated infrastructures.

Parole chiave / Keywords

Infrastrutture, Umanistica Digitale, Edizioni Scientifiche Digitali, piattaforme collaborative, modelli / Infrastructure, Digital Humanities, digital scholarly editions, collaborative platforms, models

Introduzione

Il processo di informatizzazione della conoscenza, anche attraverso la digitalizzazione delle metodologie e delle pratiche di ricerca, ha contribuito a mettere in discussione alcuni statuti disciplinari. La pluralità intrinseca alla complessità digitale richiede che, insieme a conoscenze specialistiche, si impieghino nello stesso tempo approcci trasversali al di là delle consuete e consolidate distinzioni settoriali e istituzionali. La multidisciplinarietà che ne deriva non va letta tanto come un indebolimento delle specificità disciplinari, quanto come un'opportunità per comprendere ed esplorare il potenziale della ricerca, per affrontare il cambiamento e la complessità, e per elaborare l'inevitabile tensione tra tradizione e innovazione.

In questo contributo¹ mi concentro su aspetti di metodo e sull'interrelazione tra dimensioni teoriche e pratiche nel campo dell'Umanistica Digitale (= UD), con particolare riguardo alla filosofia.² Mi concentrerò sulle Edizioni Scientifiche Digitali (= ESD), come un interessante e fruttuoso campo di riflessione, ricerca e applicazione; attraverso una serie di esempi discuto come il rapporto tra analogico e digitale si traduca nelle attuali pratiche di edizione oltre a perseguire gli obiettivi generali di fornire versioni e rappresentazioni affidabili, attendibili e utili del nostro patrimonio testuale e documentale come base per la ricerche avanzate e innovative nelle discipline umanistiche.³ Oltre ciò, la mutua trasformazione tra discipline umanistiche e scienze computazionali, l'umanistica digitale appunto, trova piena realizzazione nella ricerca e nello sviluppo di piattaforme collaborative e partecipate e nelle infrastrutture che offrono servizi e applicazioni dinamici e integrati.

Questo articolo è organizzato in tre sezioni e una conclusione. Questa prima sezione introduttiva è seguita da *Modelli multilineari per le edizioni scientifiche digitali* (2), in cui mi chiedo brevemente se e quali pratiche e categorie siano specifiche dell'attività di ESD in parti-

1 Questo contributo mantiene il tono colloquiale del mio intervento alla tavola rotonda al I Convegno Internazionale *LiMINA Lost in Manuscripts. Ideas, Notes, Acknowledgments* (Cagliari, 3-4 luglio 2024).

2 Come anche in altre mie riflessioni mi riferisco a un uso di "filosofia" nel senso più ampio e generale di indagine speculativa, di una forma di conoscenza che pone domande e riflette sul mondo e sull'essere umano, indaga il senso dell'essere e dell'esistenza, tenta di definire la natura e analizza le possibilità e i limiti della conoscenza (Marras 2022).

3 Sahle 2003; Sahle 2016.

colare quando sono coinvolti testi filosofici. La terza sezione presenta alcuni esempi di progetti, applicazioni e servizi per ESD in filosofia. La conclusione (3) riassume i punti principali affrontati in questo articolo, evidenziando gli aspetti più critici e fruttuosi che aprono la strada a ulteriori discussioni.

1. Modelli multilineari per le edizioni scientifiche digitali

La filosofia è un campo di ricerca storico e teorico che, anche in altri contesti, ho definito plurilingue⁴ almeno per due ragioni: la prima, perché la filosofia ha sempre dialogato con altre discipline non solo umanistiche ma anche, per esempio, con le scienze naturali, la matematica, la fisica, solo per citarne alcune; la seconda ragione, poiché la filosofia risponde alla sfida delle tecnologie, coniuga un approccio storico e problematico con l'apertura ad altre visioni e linguaggi che possono contaminarla. Oltretutto, l'introduzione dell'intelligenza artificiale generativa (GenAI) richiede sofisticate indagini tecniche e filosofiche.

Nella definizione di architetture di piattaforme e infrastrutture, e nella modellizzazione di testi filosofici nel contesto delle edizioni digitali, si possono rilevare alcune specificità. Da un punto di vista teorico, per esempio, la filosofia intercetta la tensione tra la necessità di preservare il multilinguismo dei contenuti e garantirne e facilitarne l'accessibilità, problematizza gli standard di annotazione quando si tratta, per esempio, di annotazione concettuale nella storia della filosofia. A ciò si legano una serie di questioni pratico-applicative, solo per citarne alcune: rispettare l'obiettivo di un mark-up del testo neutro dal punto di vista linguistico, assumere approcci semasiologici (analisi e individuazione delle forme significanti) per sviluppare successivamente indagini semantiche nella costruzione di tesauri o vocabolari concettuali, conciliare l'attività interpretativa di selezione e organizzazione delle entità appartenenti all'universo del discorso (dominio) che si intende analizzare (ontologia).⁵

La filosofia dovrebbe, infatti, essere in grado di proporre modelli convergenti e integrati di organizzazione e rappresentazione della conoscenza.⁶ Da qui, dunque, la necessità di una maggiore integrazione della

4 Marras 2018a; Marras 2022.

5 Tardella-Russo 2017.

6 Ciula *et al.* 2023; Ciula *et al.* 2018.

filosofia con i metodi e gli strumenti computazionali, per poter modellare contenuti e concetti, integrazione che passa dal perfezionare i servizi e i modelli basati su una procedura di co-progettazione per le ECD. Le ontologie computazionali, ad esempio, senza pretendere di essere esaustive o ridursi a formalizzazioni rigide, consentirebbero una prospettiva interdisciplinare e una visione multiprospettica della conoscenza per affrontare le sfide del mondo digitale. La ricerca filosofica si può far carico non solo del potenziale delle scienze computazionali, ma anche di quello della filosofia stessa, affrontando categorie a volte contraddittorie, come contestualità, indeterminazione, simulazione.⁷

Tornando alle ESD, negli ultimi anni si è finalmente fatta strada l'idea che l'ESD sia un'attività di ricerca in sé e non l'ultimo passo o il semplice risultato di un percorso più pratico che teorico.⁸

Si sta superando l'idea che il pdf di un'opera non sia sufficiente a definire una pubblicazione come digitale e che un'edizione digitale sia un'attività di ricerca vera e propria, complessa e articolata, che ha ampi margini di sperimentazione e innovazione. I perimetri di riferimento si sono allargati, implicano anche la creazione di archivi e biblioteche digitali di fonti testuali sempre più complete, affrontano le problematiche editoriali, non trascurano i problemi della codifica di testi con linguaggi standard, e lo sviluppo di motori di ricerca sempre più raffinati per l'accesso, il recupero, l'archiviazione e il riutilizzo di informazioni e documenti. Attività che sempre di più integrano nei loro processi strumenti di automazione cognitiva.⁹ La stabilità, la sostenibilità, l'accessibilità e gli standard di qualità, ad esempio, sono il minimo richiesto per qualsiasi attività scientifica anche digitale.

Un'edizione digitale non è solo una riproduzione, né un tentativo di restaurare e pubblicare un testo fedele all'originale. Non è neanche una versione di un testo, né il risultato dell'applicazione della trasmissione critica e delle operazioni testuali e critico-testuali. L'elaborazione dialettica, il plurilinguismo, la multidisciplinarietà coinvolgono anche le modalità di diffondere i risultati, le conoscenze e le informazioni in un'ottica di ricerca aperta. Sulla base di questi presupposti, propongo di cambiare prospettiva e, accanto alle pubblicazioni 'lineari' in ecosistemi/piatta-

7 Hrachovec 2008; Pichler *at al.* 2021; per la simulazione si veda Marras-Ciraci 2024.

8 Driscoll 2016; Fotis 2005; Pierazzo 2015.

9 Preferisco usare automazione cognitiva invece di intelligenza artificiale.

forme complesse, come ho discusso finora, di includere pubblicazioni dinamiche, aperte, multilivello e ‘relazionali’. Ciò consentirebbe di:

- ◇ offrire la possibilità di una nuova modellazione dei contenuti e delle sue rappresentazioni ridisegnando le mappe/struttura del testo e aprendo nuove e diverse prospettive di ricerca;
- ◇ aggiungere alla lettura e all’accesso lineare le potenzialità della lettura digitale e di un approccio multilivello al testo, potenziandone la fruizione ipertestuale, interattiva e multimodale e attivando, tra l’altro, la riscrivibilità dei contenuti.
- ◇ contribuire a ridefinire gli standard per le ESD, rendendoli più ampiamente utilizzabili e adattabili alle esigenze specifiche delle diverse comunità scientifiche. Questo non si limita ai software, ma anche agli standard di codifica, all’interfaccia dinamica per le trascrizioni (diplomatiche/interpretative), ai fogli di stile per diversi output, solo per citarne alcuni tra i più basilari. Il punto è elaborare flussi di lavoro (workflow) omogenei, dinamici, adattabili e collaborativi, in alternativa a procedure fisse o rigide.¹⁰

In questo modo si può valorizzare sempre più l’apporto e il trasferimento tecnologico che caratterizza queste attività di ricerca.

2. Progetti e modelli collaborativi per il trasferimento tecnologico. Alcuni esempi

In campo filosofico i primi interessanti progetti digitali che miravano all’integrazione della ricerca filosofica (teorica e storica) con il digitale, alla costituzione di archivi e di edizioni digitali ad accesso aperto, alla costruzione di ambienti di ricerca multilingue con applicazioni open-source progettate specificamente per la filosofia, risalgono ai primi anni 2000. Il progetto *Discovery* (<http://www.discovery-project.eu/home.html>), coordinato da Paolo D’Iorio, era in questo senso particolarmente innovativo e pionieristico. Mirava a creare *Philosource*, una federazione di biblioteche filosofiche digitali interoperabili, interconnesse con *Philospace*, una rete di applicazioni (desktop) da utilizzare per arricchire i contenuti filosofici e annotarli semanticamente. La confederazione era composta da 4 bi-

10 La rete semantica e la multidisciplinarietà della filosofia eviteranno il rischio di rappresentazioni simboliche e logiche ‘monoculturali’ (Fiormonte 2012).

biblioteche digitali: 1. *NietzscheSource* (<http://www.nietzschesource.org>), nata da un progetto collaborativo originale e all'avanguardia, l'*Hyper-Nietzsche* (<http://www.hypernietzsche.org>), edizione genetica delle opere del filosofo tedesco Friedrich Nietzsche;¹¹ 2. *WittgensteinSource* (<http://www.wittgensteinsource.org>), facsimili ed edizioni del *Nachlass* di Ludwig Wittgenstein; 3. *Daphnet, Digital Archives of PHilosophical Texts on the NET* (<http://www.daphnet.org>), archivio di fonti primarie di filosofia antica, di filosofia moderna e di letteratura critica e strumenti lessicografici interconnessi con le fonti primarie.

I risultati del progetto *Discovery* sono stati successivamente testati e implementati attraverso una serie di esperimenti nell'ambito delle attività del progetto *Agora, Scholarly Open Access Research in European Philosophy*,¹² il cui obiettivo era trasformare la federazione *Philosource* in un archivio multilingue ad accesso aperto per la pubblicazione di fonti primarie e secondarie. La durata e la sostenibilità dovevano essere garantite da un portale LOD che esponeva i metadati nel cloud.

Questi progetti hanno avviato una serie di ricerche, approfondimenti e attività di sperimentazione, sempre esplorando le potenzialità del ESD, avendo un occhio di riguardo verso le esigenze delle comunità di studiosi. Interessante è ad esempio la biblioteca d'autore: importante non solo per la storia della cultura in generale o potremmo dire per la *translatio studiorum*,¹³ ma anche perché la biblioteca «è spesso il luogo in cui avvengono i processi di risemantizzazione che accompagnano la circolazione dei testi».¹⁴

Altri progetti più recenti sono orientati all'analisi, alla mappatura e alla modellazione di corpora di dati della storia della filosofia e della scienza, come quelli condotti, ad esempio, dal *PenceLab* (<https://pencecelab.be/>) e da *The Internet Philosophy Ontology, The InPho Project* (<https://www.inphoproject.org/>). Quest'ultimo utilizza una combinazione di metodi automatizzati di feedback degli esperti per creare un'ontologia computazionale dinamica della filosofia. Si possono menzionare anche i progetti *Ideas at Scale - Towards a Computational History of Ideas (e-Ideas)* (<https://www.nwo.nl/en/projects/277-20-007>), che, come sugge-

11 D'Iorio 2000; D'Iorio 2012.

12 https://www.projectagora.org/wp-content/uploads/2012/09/AGORA_FLYER1.pdf.

13 Per la *translatio studiorum* si veda Gregory 2016.

14 Si veda la sintesi del progetto di Paolo D'Iorio in <https://anr.fr/Project-ANR-15-FRAL-0001>.

risce il titolo, presenta una storia computazionale delle idee, in cui l'esperienza di esperti umanistici è combinata con metodi computazionali per l'elaborazione di modelli concettuali e interpretativi per guidare l'analisi di corpora su larga scala, un'analisi che sarebbe altrimenti impossibile.¹⁵

L'applicazione di metodi quantitativi alla storia del pensiero, e in particolare alla storia della filosofia: «pone sfide specifiche in termini di strategie di raccolta, decisioni di progettazione e annotazione».¹⁶ Vale quindi la pena di notare l'influenza che tali prospettive può avere nella modellazione delle pratiche di ESD.¹⁷

Tutti questi esempi si basano su una forte interoperabilità e questa idea di procedura multilivello e multifattoriale di ESD è, a mio avviso, uno degli approcci più adeguati non solo per trattare i testi filosofici. Inoltre, i risultati di una ESD superano i limiti di un libro stampato a molti livelli, consentendo la collazione di diverse varianti testuali e paratestuali, la pubblicazione di materiali contestuali che arricchiscono l'edizione (mappe, facsimili, immagini, suoni, video, ecc.), la creazione di una rete di interconnessioni, fornendo così complessivamente una stratificazione flessibile dell'apparato critico, come il progetto *DanteLimina.it* ci mostra.

Lo sforzo di confrontare gli schemi dei testi codificati prodotti nell'ambito dell'editing scientifico con le rappresentazioni astratte delle strutture di database create per i sistemi di documentazione del patrimonio culturale nello sviluppo dell'ontologia non è semplice, è il caso, ad esempio, delle connessioni concettuali e storiche e delle divergenze tra i documenti codificati in base allo standard *Text Encoding Initiative* (TEI) e le informazioni fattuali/ipotetiche contenute nei database basati su CIDOC-CRM,¹⁸ che hanno contesti, caratteristiche tecniche, e scopi diversi.

Attualmente sono necessarie piattaforme interattive che consentano un utilizzo più completo e allo stesso tempo più personalizzato e spe-

15 Betti, van den Berg, 2014; Betti, van den Berg 2016.

16 Si veda anche il lavoro del gruppo di ricerca *Distant Reading and Data-Driven Research in the History of Philosophy* (DR2). Questo gruppo si concentra sui dati, sui corpora, sull'analisi comparativa e sulle metriche (<https://dr2blog.hcommons.org>), e riflette sui metodi e sugli strumenti, utilizzando modelli tematici per studiare la storia delle idee. Si veda anche Pasini 2021, p.10.

17 Hall 2008.

18 Ciula-Eide 2014.

cifico per l'utente degli strumenti e dei servizi.¹⁹ Tutto ciò implica una particolare attenzione non solo alla modellazione e alla progettazione, ma alla co-progettazione, orientata e focalizzata, come si diceva, sulle esigenze dei ricercatori. La conseguenza è il raggiungimento di un ESD dinamico, perché qualsiasi edizione accademica statica (anche se regolarmente aggiornata) sarà inadeguata a soddisfare le esigenze degli utenti e la tecnologia in continua evoluzione. La ricerca, l'apprendimento, l'istruzione, il riutilizzo, la lettura non lineare e la flessibilità, nonché l'accesso a studiosi con abilità diverse, superano ciò che un'edizione statica può offrire pur rimanendo la fonte autorevole di riferimento. Ogni settore disciplinare, oltre a seguire i requisiti obbligatori come la standardizzazione (per esempio la fairificazione), o l'integrazione con cloud e repository, progetta i propri elementi e componenti tenendo conto che, sempre più, i testi si rendono disponibili per una varietà di usi,²⁰ ciò che influenza e caratterizza le diverse DSE disciplinari sono dunque non solo la metodologia ma soprattutto gli scopi.

Concludo questa rapida rassegna con uno sguardo verso le infrastrutture di ricerca menzionando brevemente il progetto H2IOSC, *Humanities and Heritage Italian Open Science Cloud*,²¹ per il quale attualmente mi trovo a lavorare. H2IOSC riunisce un cluster collaborativo di infrastrutture di ricerca italiane nel settore umanistico, delle tecnologie del linguaggio e del patrimonio culturale, sostiene la transizione delle infrastrutture di ricerca partecipanti a un ecosistema coerente, altamente integrato e accessibile di risorse²² e moduli di formazione. Intende offrire ai ricercatori e agli utenti, a sostegno delle loro specifiche esigenze di ricerca, un unico punto d'accesso a cataloghi, servizi, strumenti, dataset, software, separatamente o aggregati in workflow. Questo punto d'accesso, il *Marketplace* H2IOSC,²³ ha anche lo scopo di orchestrare e comporre i servizi, elaborare, arricchire, analizzare e confrontare le risorse, oltre i confini dei singoli archivi o repository istituzionali.

19 Pichler 2021.

20 Maciej *et al.* 2021.

21 <https://www.h2iosc.cnr.it>.

22 H2IOSC - Humanities and Heritage Italian Open Science Cloud, è un ecosistema di ricerca, finanziato dall'Unione europea NextGenerationEU – PNRR M4C2 - Codice progetto IR0000029, <https://www.h2iosc.cnr.it/> (cons. 12/01/2025).

23 Cristofaro *et al.* 2024; Sichera *et al.* 2024.

Nel progetto, il nodo italiano di OPERAS, OPen Scholarly Communication in the European Research Area for Social Sciences and Humanities,²⁴ un’infrastruttura di ricerca che supporta la comunicazione scientifica nelle scienze umane e sociali nello Spazio Europeo della Ricerca (SER) coordina il WP5 – *Marketplace*, di cui ho sopra accennato e il WP7 – *Community pilots: innovative cross-domain services and environments*. In relazione a quest’ultimo, le attività progettuali si concentrano sulla pubblicazione accademica aperta, con lo sviluppo di un *Open Resources Publishing Pilot* e di un *Diamond OA Publishing Hub for Scholarly Editions and Learned Journals Pilot*. OPERAS in generale «ridisegna il modo di considerare la ricerca nelle SSH perché, pur seguendo un processo di standardizzazione e normalizzazione (per esempio nella metadatanazione delle risorse) non pensa di appiattare, bensì di esaltare le peculiarità scientifiche di ogni contenuto, garantendone, in linea con i principi FAIR²⁵ e TRUST,²⁶ la rintracciabilità, l’interoperabilità e la sostenibilità». ²⁷ Per questo motivo i marketplace delle piattaforme digitali risultano essere ben più di una vetrina o un catalogo, ma diventano uno strumento che veicola le forme di aggregazione transdisciplinare e lo sviluppo di conoscenza e di tecnologie. Un facilitatore, dunque, della creazione di strumenti, metodi e processi comuni e condivisi: virtuali sì, ma autenticamente plurali, in cui operano nuovi modelli di “interdisciplinarietà digitale”.²⁸

Il progetto rende le risorse disponibili in un ambiente integrato accessibile attraverso una serie di metodi di accesso virtuali e remoti. Gli obiettivi specifici e i risultati attesi del progetto sono la sperimentazione di forme di organizzazione della ricerca basate su un nuovo sistema di produzione, validazione e condivisione della conoscenza gestito direttamente dai ricercatori.

24 <https://www.operas-eu.org/> (cons. 28/01/2025). OPERAS è un’infrastruttura di ricerca distribuita che affronta alcune sfide specifiche, come la diversità linguistica nelle pubblicazioni e nell’accesso ai contenuti, il libero accesso e la gestione di forme specifiche della produzione scientifica (monografie, edizioni critiche, ecc.), l’accesso aperto “Diamond”.

25 FAIR Findable, Accessible, Interoperable, and Reusable, si veda: <https://www.go-fair.org/fair-principles/>.

26 TRUST, Transparency, Responsibility, User focus, Sustainability and Technology.

27 Cristofaro *et al.* 2024, p. 573.

28 Cristofaro *et al.* 2024, pp. 572-576.

Conclusioni

In questa breve riflessione ho discusso quali pratiche e categorie sono specifiche dell'attività delle ESD, attività complessa e articolata che sfida la filosofia su due livelli, quello teorico e quello pratico, evidenziando nuovi ruoli e contributi in scenari multidisciplinari, plurilingue, e multiprospettivi. L'interrelazione e la complementarità tra teoria e pratica interpretano le euristiche contenute nel rapporto tra filosofia e pratiche computazionali nella ricerca e colgono al meglio la complessità dei molteplici processi di edizione scientifica e l'interazione tra interfacce e contenuti.²⁹

Gli esempi presentati nella terza sezione intercettano queste nuove esigenze e prospettive. In particolare, c'è la necessità di investire non solo nell'edizione di corpora o nella raccolta di un autore, ma anche nella ricerca basata su dati, metadati, *corpus-driven* e *corpus-based*,³⁰ di flessibilità dei workflow,³¹ di tenere conto delle esigenze degli utenti. In generale, la co-progettazione e la modellazione dinamica in ESD sono viste come una risposta adeguata alle nuove esigenze di ricerca digitale e, soprattutto, catturano la complessità del dominio testuale e valorizzano il trasferimento tecnologico anche nell'ambito delle discipline umanistiche.

Bibliografia³²

Bastianello *et al.* 2022 = Elisa Bastianello, Alessandro Adamou, Christine Grundig, *Digital Publishing for the Humanities - New Technologies and Ideas*, Slides presentate alla Conferenza Internazionale "Digital Publishing for the Humanities - New Technologies and Ideas", Roma, Bibliotheca Hertziana, Max Planck Institute for Art History, 18-19 ottobre 2022. <https://doi.org/10.17617/3.PEJU8B>

Betti-van den Berg 2016 = Arianna Betti, Hein van den Berg, *Towards a Computational History of Ideas*, in *DHLU 2013, Digital Humanities Luxembourg: Proceedings of the Third Conference on Digital Humanities in Luxembourg with a Special Focus on Reading Historical Sources in the Digital Age*, Lars Wienecke, Catherine

29 Ciraci-Fedriga-Marras 2021; Wachter 2021.

30 Bonino *et al.* 2021.

31 Bastianello *et al.* 2022; Monaco-Puente-Castelo 2013; Plappert 2017.

32 Ultimo accesso alle URL della bibliografia 24 marzo 2025.

- Jones, Marten Düring *et al.* (eds.), Aachen, CEUR Workshop Proceeding, 2016. http://ceur-ws.org/Vol-1681/Betti_van_den_Berg_computational_history_of_ideas.pdf
- Betti-van den Berg 2014 = Arianna Betti, Hein van den Berg, *Modelling the History of Ideas*, «British Journal for the History of Philosophy», 22 (2014), pp. 812-835.
- Bonino *et al.* 2021 = Guido Bonino, Enrico Pasini, Paolo Tripodi, *Metodi quantitativi in storia della filosofia e storia delle idee*. «Intersezioni», 41 (2021), pp. 365-379.
- Ciraci-Fedriga-Marras 2021 = Fabio Ciraci, Riccardo Fedriga, Cristina Marras (a cura di), *Filosofia Digitale*, Milano, Mimesis, 2021.
- Ciula *et al.* 2023 = Arianna Ciula, Øyvind Eide, Cristina Marras, Patrick Sahle, *Modelling between Digital and Humanities. Thinking in Practice*, Open Book Publisher, 2023. <https://www.openbookpublishers.com/books/10.11647/obp.0369>
- Ciula *et al.* 2018 = Arianna Ciula, Øyvind Eide, Cristina Marras, Patrick Sahle (eds.), *Models and Modelling between Digital and Humanities. A Multidisciplinary Perspective*. «Historical Social Research / Historische Sozialforschung Supplement», 31. <https://doi.org/10.12759/hsr.43.2018.4.343-361>
- Ciula-Eide, 2014 = Arianna Ciula, Øyvind Eide, *Reflections on Cultural Heritage and Digital Humanities: Modelling in Practice and Theory*, in *First International Conference on Digital Access to Textual Cultural Heritage (DATECH)*, Madrid, 2014, pp. 35-41. <http://dl.acm.org/citation.cfm?id=2595207>
- Cristofaro *et al.* 2024 = Salvatore Cristofaro, Vittoria Fabiani, Cristina Marras, Enrico Pasini, Pierpaolo Sichera, Mi Yu, *Infrastrutture di ricerca come strumenti di “interculturalità digitale”*, in *Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale AIUCD2024. Quaderni di Umanistica Digitale*, a cura di Antonio Di Silvestro, Daria Spampinato, pp. 572-576, [10.6092/unibo/amsacta/7927](https://doi.org/10.6092/unibo/amsacta/7927)
- D'Iorio 2012 = Paolo D'Iorio, *Qu'est-ce qu'une édition génétique numérique?*, «Genesis» 30|2010, online 30 Mai 2012. <https://doi.org/10.4000/genesis.116>
- D'Iorio 2000 = Paolo D'Iorio, *HyperNietzsche*, Paris, PUF, 2000. <http://www.diorio.info/puf/book/hypernietzsche/le-livre.htm>
- Driscoll 2016 = Matthew James Driscoll, Elena Pierazzo (eds.), *Digital Scholarly Editing: Theories and Practices*, New edition [online],

- Cambridge, Open Book Publishers, 2016. <http://books.openedition.org/obp/3381>
- Fiormonte 2012 = Domenico Fiormonte, *Per una critica culturale delle Digital Humanities*, in *Dall'Informatica Umanistica alle culture digitali. In memoria di Giuseppe Gigliozzi*, Roma, Casa Editrice Università La Sapienza, 2012, pp. 220-241.
- Fotis 2005 = Jannidis Fotis, *Geschichte der elektronischen Editionen*, in *Handbuch germanistischer Editionen*, R. Nutt-Kofuth, B. Plachta (Hrsg.), Tübingen, Niemeyer, 2005.
- Gregory 2106 = Tullio Gregory, *Translatio linguarum, traduzioni e storia della cultura*, Firenze, Leo S. Olschki Editore, 2016.
- Hall *et al.* 2008 = David Hall, Daniel Jurafsky, and Christopher D. Manning, *Studying the History of Ideas Using Topic Models*, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing – EMNLP '08*, 363, Honolulu (HI), Association for Computational Linguistics. <https://doi.org/10/bzz22p>
- Hrachovec 2008 = Herbert Hrachovec, Alois Pichler (eds.), *Philosophy of the Information Society*, in *Proceedings of the 30. International Ludwig Wittgenstein Symposium*, vol. 2, Kirchberg am Wechsel 2007, Frankfurt, Ontos Verlag, 2008.
- Maciej *et al.* = Maryl, Maciej, et al. *D6.5 Report on the future of scholarly writing in SSH*. Project deliverable. <https://doi.org/10.5281/zenodo.4922512>
- Marras-Ciraci 2024 = Cristina Marras, Fabio Ciraci, *Simulatio*, «Lexicon Philosphicum, International Journal for the History of Texts and Ideas», n. 11, 2023-2024. <https://lexicon.cnr.it/ojs/index.php/LP/issue/view/47>
- Marras 2022 = Cristina Marras, *Filosofia digitale? Orizzonti disciplinari*, in *Digital Humanities: metodi, strumenti, saperi*, cura di Fabio Ciotti, Roma, Carocci, pp. 325-336.
- Marras 2021 = Marras Cristina, *Biodiversità ed ecosistema digitale. Per una filosofia plurilingue e multiprospettica*, in *Filosofia Digitale*, a cura di Fabio Ciraci, Riccardo Fedriga, Cristina Marras, Milano, Mimesis, 2021.
- Marras 2014 = Cristina Marras, *Exploring digital environments for research in philosophy. Metaphorical models and sustainability*, in *AIUCD '14, Proceedings of the Third AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem*, a cura di Francesca Tomasi, Roberto Rosselli del Turco, Anna Maria Tammaro,

- Bologna, settembre 18-19 settembre 2014, New York, ACM, 2014.
<https://doi.org/10.1145/2802612.2802639>
- Monaco - Puente-Castelo 2013 = Leida Monaco, Luis Puente-Castelo, *Conditionals in Eighteenth-Century Philosophy Texts: A Corpus-Based Study*, Poster presentato a Corpus Linguistics, Leicester, 2013.
<https://www.researchgate.net/publication/291346120>
- Pasini 2021 = Enrico Pasini, *On some challenges posed by corpus-based research in the history of ideas*, in *DR2 Working Papers 1*, a cura di Guido Bonino, Enrico Pasini, Paolo Tripodi, Torino, Academia University Press, 2021.
- Pichler 2021 = Alois Pichler, *Complementing Static Scholarly Editions with Dynamic Research Platforms: Interactive Dynamic Presentation (IDP) and Semantic Faceted Search and Browsing (SFB) for the Wittgenstein Nachlass*, in *Selected papers from the CLARIN Annual Conference 2020i*, Linköping Electronic Conference Proceedings 180, 2021, pp. 194-207. http://wab.uib.no/alois/Pichler_Clarin2021.pdf
- Pichler *et al.* 2021 = Alois Pichler, James Matthew Fielding, Nivedita Gangopadhyay e Andreas Lothe Opdahl, *Crisscross ontology: Mapping Concept Dynamics, Competing Argument and Multiperspectival Knowledge in Philosophy*, in *Filosofia Digitale*, a cura di Fabio Ciraci, Riccardo Fedriga, Cristina Marras, Milano, Mimesis, 2021, pp. 59-73.
- Pierazzo 2015 = Eleina Pierazzo, *Digital Scholarly Editing Theories, Models and Methods*, London, Routledge, 2015.
- Plappert 2017 = Plappert Garry, *Candidate Knowledge? Exploring Epistemic Claims in Scientific Writing: A Corpus-Driven Approach*, «Corpora», 12 (2017), pp. 425-457. <https://doi.org/10.3366/cor.2017.0127>
- Robinson 2002 = Peter Robinson, *What is a Critical Digital Edition?. Variants*, «Journal of the European Society for Textual Scholarship», 1 (2002), pp. 43-62.
- Robinson 2002 = Peter Robinson, *The One Text and the Many Texts*, «Literary and Linguistic Computing», 15 (2002), pp. 5-14.
- Russo - Tardella 2017 = Ada Russo, Michela Tardella, *TheofPhilo. Thesaurus of Philosophy*, «ILIESI digitale. Memorie», n. 3, ILIESI-CNR, marzo 2017. https://www.iliesi.cnr.it/pubblicazioni/Memorie-03-Russo_Tardella.pdf
- Sahle 2016 = Patrick Sahle, 2. *What is a Scholarly Digital Edition? in Digital Scholarly Editing: Theories and Practices*, Cambridge, Open Book Publishers, 2016. <http://books.openedition.org/obp/3397>.

- Sahle 2003 = Patrick Sahle, *Vom editorischen Fachwissen zur digitalen Edition. Der Editionsprozeß zwischen Quellenbeschreibung und Benutzeroberfläche*, in *Fundus – Forum für Geschichte und ihre Quellen*, Beiheft 2: Quellen und Quelleneditionen im neuen Medienzeitalter, Manfred Thaller (Hrsg.), Göttingen, 2003, pp. 75-102. Online version: <http://webdoc.sub.gwdg.de/edoc/p/fundus/2/sahle.pdf>.
- Sichera *et al.* 2024= Pietro Sichera, Cristina Marras, Enrico Pasini, *Orchestrazione API per workflow applicativi nell'ambito delle Digital Humanities*, Scientific-technical Report. vers.2024, 19 novembre). <https://doi.org/10.5281/zenodo.14187534>
- Smith 2004 = Martha Nell Smith, *Electronic Scholarly Editing*, in *A Companion to Digital Humanities*, Susan Schreibman, Ray Siemens, John Unsworth (eds.), Oxford, Blackwell, 2004. https://companions.digitalhumanities.org/DH/?chapter=content/9781405103213_chapter_22.html
- Vanhoutte 1999 = Edward Vanhoutte, *Where is the Editor? Resistance in the creation of an electronic critical edition*, «Human IT», 1999, pp. 197-214. <https://humanit.hb.se/article/viewFile/226/299>
- van Hulle 1999 = Dirk van Hulle, *Authenticity or Hyperreality in Hypertext Editions. Notes Towards a Searchable Recherche*, «Human IT», 1999, pp. 227–244. <http://www.hb.se/bhs/ith/1-99/dvh.htm>
- Wachter 2021 = Christian Wachter, *Publishing Complexity in the Digital Humanities*, «magazén», 2 (2021), pp. 103-118. <doi:10.30687/mag/2724-3923/2021/03/004>



10.69118/2025_L3

La TEI e la rappresentazione del testo: alcune riflessioni critiche

Roberto Rosselli Del Turco

Università di Torino

roberto.rosselidelturco@unito.it

TEI and text representation: some critical considerations

Abstract (ITA)

Questo contributo affronta alcune delle principali criticità strutturali e concettuali della codifica XML/TEI nell'ambito delle edizioni scientifiche digitali. Sebbene la TEI garantisca agli editori una notevole libertà nella rappresentazione di fenomeni testuali complessi, tale flessibilità si traduce spesso in pratiche di codifica molto eterogenee e in una scarsa interoperabilità tra progetti. Tra i problemi centrali vi sono il modello gerarchico rigido di XML, che non consente la rappresentazione nativa di strutture sovrapposte, e la necessità di strumenti a valle in grado di interpretare markup fortemente diversificati. Il caso dei *marginalia*, esemplare per la sua ricchezza semantica e caratteristiche di marginalità strutturale, viene discusso come esempio concreto delle difficoltà generate da queste dinamiche. Si riflette infine su come le scelte di codifica influenzino le strategie di visualizzazione e sulla necessità di sviluppare strumenti adattabili ai futuri sviluppi tecnologici e standard emergenti.

Abstract (ENG)

This paper addresses some of the structural and conceptual limitations of XML/TEI encoding in the context of Digital Scholarly Editions. While the TEI schemas offers editors considerable flexibility in representing complex textual phenomena, this same flexibility often results in highly divergent encoding practices and limited interoperability across projects. One of the most critical issues is the rigid hierarchical model of XML, which cannot natively represent overlapping structures, but also the need for downstream tools to consistently interpret heterogeneous markup consistently. The case of *marginalia*, paradigmatic of paratexts that are semantically rich but structurally peripheral, is examined to illustrate the practical consequences of these challenges. The paper also reflects on how encoding decisions impact visualization strategies and discusses the importance of designing adaptable tools capable of responding to evolving standards and technologies.

Parole chiave / Keywords

Filologia digitale, annotazione semantica, codifica XML/TEI, edizioni scientifiche digitali / Digital Philology, Semantic Annotation, XML/TEI Encoding, Digital Scholarly Editions

Nel corso di *Digital Philology* che tengo nell'ambito del corso di laurea in *Language Technologies and Digital Humanities* presso l'Università di Torino,¹ ricorro spesso a una metafora per introdurre un concetto fondamentale. Immaginiamo di aver acquistato un certo numero di scatole di cartone di varie dimensioni: inserire una scatola piccola in una più grande, o più scatole, ciascuna contenente al proprio interno altre scatole più piccole, sempre in una più grande, non costituisce alcuna difficoltà. Ciò che invece non è possibile fare è collocare una scatola piccola per metà in una scatola grande e per l'altra metà in un'altra. Si tratta di un principio del tutto intuitivo nella vita quotidiana, ma assai utile per introdurre l'idea di gerarchia strutturale che sottende l'organizzazione degli elementi, veri e propri contenitori testuali per quanto riguarda l'ambito della filologia digitale, nel linguaggio XML. Questa metafora permette inoltre di illustrare una delle principali problematiche nella codifica testuale: la rappresentazione di gerarchie multiple (ingl. *overlapping hierarchies*), resa complessa dall'impossibilità, in XML, di sovrapporre elementi. XML, infatti, si fonda sul principio OHCO (*Ordered Hierarchy of Content Objects*)² e non consente la rappresentazione diretta di strutture gerarchiche alternative rispetto a quella principale (e unica) del documento. La parola chiave, in questo contesto, è proprio "rappresentazione diretta", poiché gli schemi XML della TEI (*Text Encoding Initiative*)³ mettono comunque a disposizione una serie di strategie alternative per gestire gerarchie multiple, e quindi la sovrapposizione di elementi, tra cui: l'uso di elementi vuoti per descrivere la struttura fisica del documento, la segmentazione del testo in più elementi correlati, il ricorso al markup *stand-off*, fino alla tecnica nota come *Trojan horsing*,⁴ che consiste nell'impiego di elementi vuoti

1 <https://www.digitalhumanities.unito.it/do/home.pl>.

2 Renear *et al.* 1996.

3 <https://tei-c.org/>.

4 Sperberg-McQueen 2018.

per indicare i punti di inizio e fine di un brano di testo, simulando l'apertura e la chiusura di un unico elemento XML.

In termini generali, l'annotazione semantica basata sugli schemi XML/TEI si è rivelata uno strumento efficace e particolarmente adatto alla realizzazione di edizioni scientifiche digitali (ingl. *Digital Scholarly Editions*). Dopo anni di crescita costante,⁵ possiamo ormai considerare l'adozione del formato TEI uno standard *de facto*. L'impiego di tale formato rappresenta infatti la scelta più razionale per la creazione di nuove edizioni digitali, ed è ritenuto essenziale anche da quei progetti che adottano approcci alternativi, come ad esempio l'utilizzo di database relazionali per la gestione dei testi, quantomeno come formato di esportazione per raggiungere l'interoperabilità con altri progetti. Alla luce dell'esperienza accumulata nel corso di decenni, dell'ampio numero di DSE che ne hanno verificato l'efficacia, della vitalità della comunità che lo sostiene e della disponibilità al confronto da parte dei suoi editor, si può affermare senza esitazione che la Text Encoding Initiative gode di ottima salute e rappresenta una risorsa solida, affidabile e consolidata per l'intera comunità accademica impegnata nel campo della filologia digitale.

La preminenza assunta dalla TEI rappresenta un importante elemento di stabilità per chi opera nell'ambito della filologia digitale, costituendo di fatto le fondamenta sulle quali si basa la grande maggioranza dei progetti di edizione digitale attualmente in fase di sviluppo. Il rovescio della medaglia, tuttavia, è che, proprio in virtù della sua efficacia nelle codifiche 'leggere' e della sua capacità di risultare quasi sempre 'sufficientemente buona' (ingl. *good enough*) anche per progetti più articolati, l'impianto metodologico e concettuale delle edizioni basate su codifica TEI è rimasto sostanzialmente invariato da oltre vent'anni, pur a fronte di strumenti tecnici notevolmente più evoluti rispetto alla prima generazione di edizioni digitali. Il problema della rappresentazione di gerarchie multiple, ad esempio, continua a sussistere ed è affrontato oggi con soluzioni tecnicamente più sofisticate – si pensi, in particolare, all'impiego del metodo *double-end-point attachment* in markup *stand-off* per l'apparato critico – ma al tempo stesso più complesse e onerose per chi si occupa della codifica. La verbosità intrinseca del linguaggio XML, unitamente alla caratteristica peculiare della TEI di offrire all'utente un elevato grado

5 Si veda ad esempio il catalogo curato da Greta Franzini (Franzini 2017). La web app del catalogo è disponibile all'URL <https://dig-ed-cat.acdh.oeaw.ac.at/>.

di libertà nella definizione di un proprio modello di codifica, ha come conseguenza la possibilità di offrire molte soluzioni alternative per lo stesso problema di codifica. La combinazione di elementi, attributi e – soprattutto – valori per questi ultimi scelti da ciascun utente dà spesso origine a documenti formalmente compatibili, ma che risultano interoperabili soltanto a seguito di un processo di verifica e normalizzazione, condotto in base ai requisiti specifici del progetto che importa i dati. In pratica, un riutilizzo diretto è possibile quasi esclusivamente nei casi di marcatura molto leggera; la situazione si complica ulteriormente nel momento in cui gli schemi adottati sono stati oggetto di personalizzazioni richieste dal progetto. Infine, i documenti TEI prodotti nell'ambito di una DSE vengono quasi sempre convertiti in formato HTML5 per la pubblicazione e la visualizzazione online. Questo passaggio introduce un'ulteriore variabilità, dovuta tanto alle eventuali difficoltà di conversione tra i due formati – essendo la TEI concepita per una annotazione semantica mentre l'HTML è destinato alla presentazione – quanto alla diversità degli algoritmi di *rendering* implementati dai vari software di visualizzazione. A ciò si aggiunge l'evoluzione continua degli standard web, che impone agli strumenti di visualizzazione un costante aggiornamento per mantenere la compatibilità.

Per dare forma concreta alle osservazioni precedenti, si può considerare un caso specifico che riveste, tra l'altro, un interesse diretto per il progetto *LiMINA*: quello dei *marginalia* e, più in generale, delle forme di contenuto paratestuale. Questi fenomeni risultano particolarmente complessi da codificare mediante gli schemi della TEI, poiché costituiscono fenomeni estremamente eterogenei, spesso difficili da classificare in modo univoco, e caratterizzati da una notevole sovrapposizione tra categorie diverse (Fig. 1).

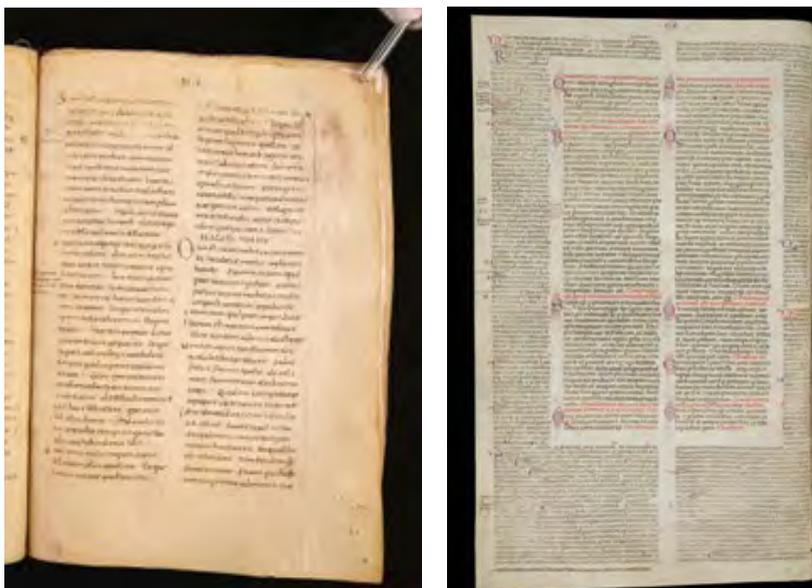


Fig. 1. Brevi note a margine nel ms. Vercelli, Biblioteca Capitolare, Codex Vercellensis LVIII, f. 43r a confronto con l'esteso testo di commento al ms. Cologny, Fondation Martin Bodmer, Cod. Bodmer, f. 59v.

Come accennato in precedenza, gli schemi TEI offrono un notevole grado di libertà nella codifica, tuttavia richiedono che gli elementi a disposizione del filologo siano combinati in modo sistematico, tenendo conto delle specifiche funzionalità e dei vincoli associati a ciascuno di essi. Nel caso dei *marginalia*, esistono diverse soluzioni possibili, la cui scelta dipende principalmente dall'elemento TEI selezionato come punto di partenza per rappresentare il fenomeno. Un'opzione relativamente semplice sul piano trascrizionale è l'uso dell'elemento `<add>`, che consente di indicare l'inserimento di testo nei margini del documento specificandone la posizione mediante l'attributo `@place`. Secondo la definizione sintetica contenuta nelle *Guidelines TEI*,⁶ `<add>` “contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector”.⁷

6 TEI Consortium 2025.

7 <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-add.html>.

Da un punto di vista semantico, tuttavia, l'elemento `<add>` risulta piuttosto generico: esso descrive infatti qualsiasi aggiunta scribale, anche all'interno dello specchio di scrittura, come nel caso di correzioni o glosse interlineari. Inoltre `<add>` si riferisce a un fenomeno di natura eminentemente documentaria, connesso più al processo materiale della scrittura che alla tipologia del contenuto introdotto. Sarebbe possibile specificare, tramite l'attributo `@type`, che si tratta ad esempio di una glossa marginale – come in `<add type="gloss" place="margin right">` – ma perché ricorrere a tale costruzione quando la TEI mette a disposizione un elemento specifico come `<gloss>`,⁸ semanticamente molto più preciso e compatto? Una rapida analisi delle caratteristiche di `<gloss>` rivela, purtroppo, che questo elemento non contempla l'attributo `@place`, rendendolo di fatto inadeguato alla rappresentazione dei *marginalia*. Un'indagine più approfondita consente però di individuare un'alternativa valida nell'elemento `<note>`,⁹ che risulta particolarmente flessibile e potente: non solo offre gli attributi necessari (`@place in primis`), ma è anche molto versatile dal punto di vista del contenuto. Essendo pensato anche per la codifica di note editoriali e di altro tipo è consigliabile l'impiego degli attributi `@type` e `@resp`, rispettivamente per indicare la tipologia dell'annotazione (glossa, commento, nota filologica, etc.) e l'autore materiale dell'intervento (scriba, autore, possessore del manoscritto, etc.).

Tuttavia, l'elemento `<note>` potrebbe non essere la soluzione più adeguata nei casi in cui si debbano rappresentare paratesti di una certa estensione e complessità strutturale (Fig. 1). In simili circostanze, risulta più opportuno ricorrere a elementi strutturali come `<div>`¹⁰ oppure `<ab>`¹¹. L'impiego di questi elementi richiede però un'attenta analisi della sintassi del documento TEI, al fine di garantirne un'integrazione coerente all'interno della gerarchia testuale complessiva.¹² Tra i due elementi considerati, `<ab>` presenta alcune limitazioni, in particolare

8 <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-gloss.html>.

9 <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-note.html>.

10 <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-div.html>.

11 <https://tei-c.org/release/doc/tei-p5-doc/en/html/ref-ab.html>.

12 Fortunatamente, la natura stessa dei *marginalia*, tendenzialmente accessoria e periferica rispetto al testo principale, consente in molti casi di evitare il rischio di sovrapposizioni strutturali (*overlap*) a livello gerarchico.

l'assenza dell'attributo `@place`, mentre `<div>` risulta più versatile, anche grazie all'introduzione recente proprio di tale attributo. In virtù di queste caratteristiche, `<div>` appare come la soluzione più adatta per rappresentare *scholia*, glosse estese o altre forme di annotazione paratestuale lunga e articolata. Resta comunque necessario stabilire un legame esplicito tra l'elemento annotativo e il passaggio del testo principale cui esso si riferisce, operazione che può essere effettuata, ad esempio, tramite l'attributo `@corresp`.

Come emerge da questa rapida ricognizione delle risorse offerte da XML/TEI per la codifica di glosse e altri elementi paratestuali, non esiste una soluzione preconfezionata che risulti valida in ogni contesto. Al contrario, è indispensabile procedere a una valutazione accurata delle caratteristiche del documento originale, nonché definire con chiarezza gli obiettivi specifici del progetto di edizione, affinché la scelta degli elementi e degli attributi TEI risulti coerente e funzionale. Va inoltre sottolineato che la codifica costituisce solo il primo passo di un processo più ampio: spetta infatti al software di visualizzazione dell'edizione digitale il compito di restituire in maniera efficace tali componenti testuali, eventualmente stabilendo un collegamento interattivo con il testo principale, rendendole accessibili alla ricerca, alla consultazione, e per altri scopi in base agli obiettivi del progetto.

Nel contesto dello sviluppo di *EVT* (Edition Visualization Technology),¹³ un software per la consultazione di edizioni digitali scientifiche,¹⁴ su richiesta di alcuni progetti – in particolare il progetto relativo all'edizione digitale delle *Leges Langobardorum*¹⁵ e del *Canzoniere* 1919 di Umberto Saba (progetto *Saba 1919*)¹⁶ – stiamo proprio adesso avviando la progettazione di funzionalità specifiche per la gestione di contenuti paratestuali come i *marginalia*. Come di consueto, il punto di partenza imprescindibile è rappresentato dalla codifica TEI; al tempo stesso, uno dei principi fondamentali nello sviluppo di tale strumento resta quello di non imporre agli editori l'adozione di un particolare stile di codifica. Il nostro obiettivo ideale è offrire un supporto efficace per tutti i documenti TEI che ricadono nell'ambito dei moduli relativi alla

13 Home page: <http://evt.labcd.unipi.it/>.

14 Rosselli Del Turco 2019.

15 PRIN PNRR 2022 “Leges Langobardorum”; Buzzoni-Rosselli Del Turco 2024.

16 Buzzoni 2024, Buzzoni *et al.* 2024.

filologia digitale attualmente gestiti da *EVT*. Tuttavia, la potenziale variabilità nelle combinazioni di elementi e attributi rende impraticabile una compatibilità perfetta con ogni possibile implementazione. Nel caso specifico dei *marginalia*, ciò comporta la necessità di privilegiare un approccio ben definito per quanto riguarda le scelte di codifica – nel nostro caso, l'utilizzo di `<note>` per glosse e annotazioni, e di `<div>` per commenti lunghi e strutturati – garantendo comunque un livello di supporto accettabile anche per quei progetti che optano per soluzioni alternative.

Situazioni analoghe, in cui si rende necessario operare scelte talvolta complesse, si presentano con regolarità nello sviluppo di *EVT*. Un esempio emblematico è costituito dalla codifica dell'apparato critico, come nel caso del progetto *RETI*,¹⁷ dedicato all'edizione di testi classici in latino e greco, per il quale è richiesto un livello di dettaglio estremamente elevato. Un ulteriore caso significativo riguarda la rappresentazione degli epistolari: sebbene la TEI abbia compiuto importanti progressi nel mettere a disposizione strumenti specifici per la codifica di questo genere testuale, permane una notevole varietà di approcci, legata alla possibilità di combinare in modi molto diversi gli elementi disponibili. Analoghe complessità emergono anche in relazione alla trascrizione di documenti manoscritti, dove le caratteristiche materiali e storiche dei testimoni impongono spesso soluzioni di codifica personalizzate, alla ricerca di un ideale equilibrio tra precisione descrittiva e sostenibilità tecnica del markup utilizzato.

In conclusione, assistiamo oggi a un notevole incremento sia quantitativo sia qualitativo nelle edizioni scientifiche digitali pubblicate online, grazie anche alla crescente disponibilità di strumenti avanzati per la pubblicazione e la consultazione di tali risorse – una situazione ben diversa rispetto anche solo a un decennio fa. Tuttavia, a questa evoluzione positiva si accompagna un insieme di criticità strutturali, spesso concatenate, sulle quali è possibile intervenire *a valle* – mediante scelte di codifica consapevoli basate sugli schemi TEI e adeguamenti mirati degli strumenti di visualizzazione – ma difficilmente *a monte*, poiché le caratteristiche fondamentali del linguaggio XML, su cui la TEI si fonda, restano fuori dal controllo diretto degli sviluppatori e degli editori attivi nel campo della filologia digitale. A complicare ulteriormente il quadro, si aggiunge

17 PRIN 2022 “RETI - REnDerIng Texts and Images. Digital Scholarly Editions with Edition Visualization Technology”.

l'emergere di nuove tecnologie e paradigmi, come l'impiego delle ontologie e di modelli di rappresentazione basati sul web semantico, che stanno progressivamente affiancando – e in alcuni ambiti potenzialmente integrando – la codifica TEI. È quindi ragionevole ipotizzare che lo stesso ecosistema TEI possa subire, in un futuro più o meno prossimo, trasformazioni significative nella struttura, nelle finalità e nei meccanismi di implementazione. Una innovazione importante, che potrebbe portare a una riduzione dell'eterogeneità delle codifiche, potrebbe essere l'introduzione di un vocabolario controllato (ingl. *controlled vocabulary*) condiviso che renda omogenei, e quindi più facilmente elaborabili dai software di navigazione, i valori degli attributi utilizzati nel markup TEI.

In questo scenario in evoluzione, appare essenziale adottare un atteggiamento progettuale aperto e flessibile, che sappia non solo valorizzare al meglio gli strumenti attualmente disponibili, ma anche predisporre le basi per un adattamento consapevole ai mutamenti futuri. La capacità di integrare nuove tecnologie mantenendo al centro le esigenze filologiche e le pratiche critiche rimane, e rimarrà, una delle sfide decisive per chi opera nel campo delle edizioni digitali.

Bibliografia

- Buzzoni 2024 = Marina Buzzoni, *Il "Progetto Saba": dare voce a un manoscritto inedito del Canzoniere*, in *Informatica umanistica, Digital Humanities: verso quale modernità?*, a cura di Maristella Gatto, Alessandra Squeo e Silvia Silvestri, Bari, Cacucci Editore, 2024, pp. 153-166. <https://iris.unive.it/handle/10278/5016061?mode=complete>
- Buzzoni-Rosselli Del Turco 2024 = Marina Buzzoni, Roberto Rosselli Del Turco, *Towards an integrated digital edition of the Leges Langobardorum*, in *Atti del XIII Convegno Annuale AIUCD 2024*, a cura di Antonio Di Silvestro e Daria Spampinato, 2024, pp. 226-231. <https://doi.org/10.6092/unibo/amsacta/7927>
- Buzzoni et al. 2024 = Marina Buzzoni, Davide Cucurnia, Cristina Fenu, Roberto Rosselli Del Turco e Giulia Tancredi, *Progetto di edizione genetica digitale del Canzoniere manoscritto di U. Saba (1919-20)*, in *Atti del XIII Convegno Annuale AIUCD 2024*, a cura di Antonio Di Silvestro e Daria Spampinato, 2024, pp. 215-220. <https://doi.org/10.6092/unibo/amsacta/7927>
- Franzini 2017 = Greta Franzini, Melissa Terras e Simon Mahony, *A Catalogue of Digital Editions*, in *Digital Scholarly Editing: Theories*

- and Practices*, a cura di Matthew James Driscoll e Elena Pierazzo, Cambridge, Open Book Publishers, 2017, pp. 161-82. <http://books.openedition.org/obp/3416>
- Renear *et al.* 1996 = Allen H. Renear, Elli Mylonas e David Durand., *Refining our Notion of What Text Really Is: The Problem of Overlapping Hierarchies*, in *Research in Humanities Computing*, a cura di Nancy Ide e Susan Hockey, Oxford, Oxford University Press, 1996, pp. 263-280. <https://hdl.handle.net/2142/9407>
- Rosselli Del Turco 2019 = Roberto Rosselli Del Turco, *Designing an Advanced Software Tool for Digital Scholarly Editions: The Inception and Development of EVT (Edition Visualization Technology)*, «Textual Cultures», 12 (2019), pp. 91.111. <https://doi.org/10.14434/textual.v12i2.27690>
- Sperberg-McQueen 2018 = C. Michael Sperberg-McQueen, *Representing concurrent document structures using Trojan Horse markup*, in *Balisage: The Markup Conference 2018*, Washington DC, 2018. <https://doi.org/10.4242/BalisageVol21.Sperberg-McQueen01>



10.69118/2025_L4

Work in progress: the palimpsest Comedy of the “Dante Poggiali” (BNCF, ms. Pal. 313)

Ciro Perna, Elisabetta Tonello

Università degli Studi della Campania “Luigi Vanvitelli”, Università di Salerno
ciro.perna@unicampania.it, etonello@unisa.it

Lavori in corso: il palinsesto della *Commedia* del “Dante Poggiali”

Abstract (ITA)

La nota anticipa un lavoro in corso sul codice Palatino 313 della Biblioteca Nazionale Centrale di Firenze, volto alla decrittazione e trascrizione della *scriptio inferior* di alcune carte del manufatto. Grazie alla scansione attraverso il microdome, prototipo della KU Leuven, sono stati svelati i versi delle carte palinseste, su cui si tornerà in un saggio a più mani ospitato dal prossimo fascicolo di questa rivista.

Abstract (ENG)

The article anticipates ongoing work on Palatino codex 313 of the National Central Library of Florence, aimed at decrypting and transcribing the *scriptio inferior* of some of the artifact’s folios. Thanks to scanning through the microdome, a prototype of the KU Leuven, the verses of the palimpsest sections have been revealed, on which a multi-handed essay hosted by the next issue of this journal will focus.

Parole chiave / Keywords

Dante Alighieri, Commedia, Palatino 313, palinsesto, microdome / Dante Alighieri, Commedia, Palatino 313, palimpsest, microdome

The synergy between the *DanteLimina.it* project and the National Central Library of Florence¹ allowed for the holding of an international workshop on Feb., 28th 2025 (<https://limina.dantelimina.it/incontro/firenze2502>), in which paleographic, historical and philological skills profitably met different declinations of the so-called “new technologies”: at the center an extraordinary example of the very early Florentine suc-

1 The partnership was formalized by an agreement signed on September, 17th 2024.

cess of the *Commedia*: the Palatino 313 manuscript. Bearer not only of the Dante Poem, but, integrated in a truly unique layout, also of an extraordinary figurative cycle (unfinished, but planned for all the hundred cantos) and of an exegetical vernacular anonymous apparatus, substantially limited to the first cantica (known as *Chiose Palatine*), the codex has never gone beyond the chronological limits of the “antica vulgata” in the various datings proposed, to configure itself as a true milestone of Dantean codicological, historical-artistic, philological studies.²

On the occasion of the Florentine meeting, a selection of folios from the “Dante Poggiali” was submitted to multispectral scans by microdome technology, patented by the KU of Leuven,³ which, in addition to giving valuable insights into the use of color and preparatory drawings of the 37 tabulae pictae analyzed, made it possible to clearly read the text of the *scriptio inferior* of the palimpsest folios 156r-159v and 162rv. Already marked in lapis by a modern hand directly on the pages of the manuscript (Fig. 1), the verses below, derived from cantos XXII and XXIII of *Purgatorio* and I of *Paradiso*, are now legible (example in Fig. 2) and could for the first time be transcribed and stemmatically placed.

The results of this work, conducted by Elisabetta Tonello, Beatrice Mosca and Serena Malatesta, will be published in the next issue of this journal (2 2025) and promise to be not only a useful acquisition for the definition of the dating of the artifact, but also an additional piece in the genealogical relations between witnesses of the second and third canticles, as part of the critical edition of the *Commedia* of the so-called “Ferrara group.”

Bibliografia

- Abardo 2005 = *Chiose Palatine (Ms. Pal. 313 della Biblioteca Nazionale di Firenze)*, a cura di Rudy Abardo, Roma, Salerno Editrice, 2005.
- Azzetta et al. 2021 = «Onorevole e antico cittadino di Firenze» *Il Bargello per Dante. Catalogo della mostra. Firenze, Museo Nazionale del Bargello, 21 aprile-31 luglio 2021*, a cura di Luca Azzetta, Sonia Chiodo, Teresa De Robertis, Firenze, Mandragora, 2021.

2 The bibliography on the manuscript is endless: we simply mention Abardo 2005, Azzetta et al. 2021, pp. 232-34 and the description in Manus On Line: <https://manus.iccu.sbn.it/risultati-ricerca-manoscritti/-/manus-search/cnmd/0000300055?>

3 See Watteeuw-Hameeuw 2022.

Watteeuw-Hameeuw 2022 = Lieve Watteeuw, Hendrik Hameeuw, *Art technical research on manuscript heritage with the white light and multispectral portable light dome system (multi-light reflectance imaging)*, in *Immaginare la Commedia*, a cura di Ciro Perna, Roma, Salerno Editrice, 2022, pp. 225-235.

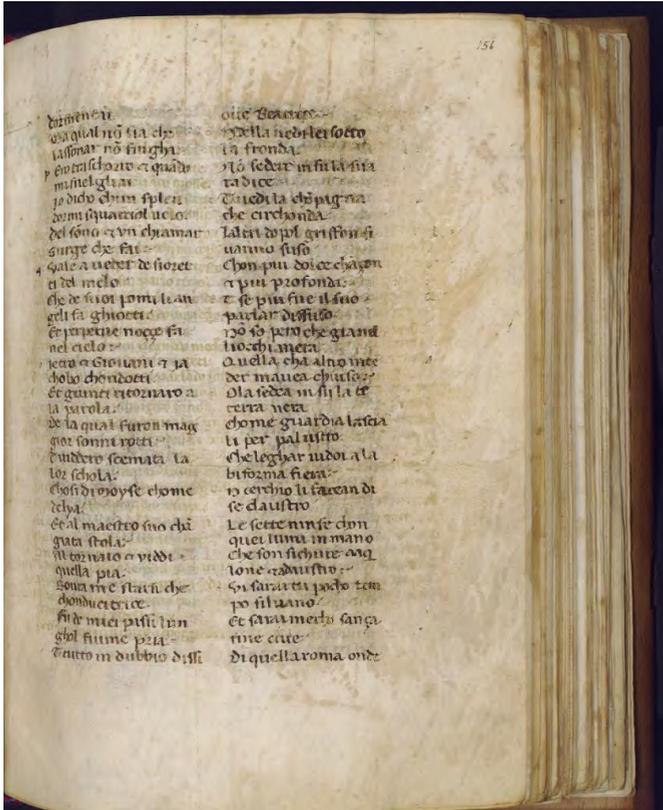


Fig. 1, c. 156r



10.69118/2025_L5

*Divine (virtual) Reality: un mondo virtuale per immergersi nella Commedia**

Serena Malatesta

Università degli studi di Padova - Université de Tours
serena.malatesta@phd.unipd.it

Divine (virtual) Reality: a virtual world to delve into the Comedy

Abstract (ITA)

Il contributo presenta il progetto *Divine Reality*, sviluppato all'interno dell'infrastruttura D.A.N.T.E., come esempio di applicazione avanzata delle tecnologie immersive alla valorizzazione didattica e critica del patrimonio dantesco. A partire da una riflessione teorica sui concetti di realtà virtuale, *gamification* e rimedializzazione il lavoro esplora le potenzialità della VR come spazio epistemico per la costruzione di modelli storici e iconografici interattivi. Dopo una prima fase di sperimentazione, è stato progettato un ambiente virtuale, che consente di esplorare una stamperia cinquecentesca e una biblioteca digitale immersiva contenente edizioni illustrate della *Commedia*. Il progetto permette di testare l'efficacia della *gamification* come stimolo cognitivo e di rilevare la convergenza di competenze interdisciplinari – filologiche, storiche, tecnologiche – nella progettazione di ambienti educativi immersivi.

Abstract (ENG)

This paper presents *Divine Reality*, a project developed within the D.A.N.T.E. infrastructure, as an example of advanced application of immersive technologies to the didactic and critical valorisation of Dante's heritage. Starting from a theoretical reflection on the concepts of virtual reality, *gamification* and remedialisation, the work explores the potential of VR as an epistemic space for the construction of interactive historical and iconographic models. After an initial experimentation phase, a virtual environment was designed to explore a 16th century print shop and an immersive digital library containing illustrated editions of the *Divine Comedy*. The project allows to test the effectiveness of *gamification* as a cognitive stimulus and to detect the convergence of interdisciplinary skills – philological, historical, technological – in the design of immersive educational environments.

* I responsabili della programmazione e della realizzazione del secondo prototipo in realtà virtuale e delle immagini da Fig. 1 a Fig. 9 sono Alessandro Calamo (alessandro-calamo.info@gmail.com) e Marco Serra (marcoserra89@gmail.com) per Remedia srl (agenzia specializzata in cultura, tecnologia e narrazione digitale).

Parole chiave / Keywords

Divine Reality, Realtà Virtuale, Edizioni a stampa, media studies, educational studies / Divine Reality, Virtual Reality, Early Modern Print, media studies, educational studies

Negli ultimi anni, l'evoluzione delle tecnologie digitali e la loro crescente integrazione nel mondo materiale hanno riportato al centro del dibattito le tecnologie immersive – e in particolare la realtà virtuale (VR, *virtual reality*), – le quali stanno profondamente trasformando le modalità di percezione e interazione con l'ambiente circostante. Nel contesto odierno, dove i confini tra il mondo reale e il mondo digitale si fanno sempre più labili, la realtà virtuale non è relegata a esperienze marginali o settoriali, ma si afferma come componente strutturale della quotidianità, con implicazioni rilevanti nei campi della cultura, della comunicazione e dell'educazione. Le Digital Humanities (DH), in quanto snodo interdisciplinare tra sapere umanistico e sistemi computazionali, sono chiamate a interrogarsi criticamente su questo scenario, riconoscendo nelle tecnologie immersive uno strumento epistemologico in grado di modellare sia la produzione che la fruizione della conoscenza. Tale trasformazione si inserisce in un più ampio ripensamento delle metodologie formative, volto a superare il paradigma tradizionale in favore di pratiche di apprendimento attive, multisensoriali e interattive, capaci di coinvolgere trasversalmente tutte le categorie di utenti e di favorire una più inclusiva ed efficace circolazione del sapere.

Tuttavia, l'introduzione della VR in ambito accademico e formativo si confronta ancora con il persistente *stigma* dell'intrattenimento di massa. Associata perlopiù al mondo del *gaming* e del consumo audiovisivo, la realtà virtuale fatica a essere riconosciuta come uno strumento pienamente legittimo di produzione e trasmissione della conoscenza. È in questo spazio di tensione che i progetti educativi e di ricerca che sfruttano le tecnologie immersive, in particolare nel contesto delle Digital Humanities,¹

1 I videogiochi sono tra le forme di new media più diffuse e influenti oggi e sono stati centrali nella storia dell'informatica stessa, fin dagli albori dei computer. Nonostante il profondo interesse nelle Digital Humanities, per diversi decenni i giochi sono stati in gran parte esclusi dalla ricerca. Nel corso del tempo, i cambiamenti nelle pratiche e nelle piattaforme, insieme all'ascesa dei *media studies*, hanno portato a un maggiore focus sui *media born digital*, inclusi i videogiochi. In anni recenti, i giochi, gli ambienti videoludici e

si collocano come sperimentazioni critiche e potenzialmente trasformative sul metodo della ricerca.

Un nodo teorico cruciale di questa evoluzione riguarda la modellazione della realtà attraverso l'ambiente videoludico. Lungi dall'essere un mero contenitore narrativo o un prodotto d'intrattenimento, il videogioco si propone come dispositivo epistemologico: uno spazio strutturato da regole, interazioni e dinamiche sistemiche che modellano ambienti simbolici e concettuali. In tale prospettiva, la realtà virtuale non si presenta come una "seconda realtà" o una realtà alternativa, ma come una soglia permeabile, uno spazio ibrido in cui mondo fisico e rappresentazione digitale si intrecciano e si contribuiscono l'uno alla definizione dell'altro. Reale e digitale si intrecciano sempre più, dando forma a quello che Steven E. Jones, riprendendo un'intuizione di William Gibson,² definisce l'*everting* del cyberspazio, cioè il suo rovesciamento nel mondo fisico, che «leaks out into the physical world».³ Questo processo si è intensificato nell'attuale quadro sociale, caratterizzato dalla sovraesposizione ai contenuti digitali e dalla crescente confusione tra esperienze mediate ed esperienze fattuali. I social network, i flussi continui di immagini e l'interconnessione tra sfera privata e pubblica hanno contribuito a smarginare la distinzione tra reale e virtuale. Le recenti implementazioni dell'ecosistema Meta, promosso da Mark Zuckerberg, rappresentano un'ulteriore accelerazione in questa direzione: la cosiddetta realtà mista (MR, *mixed reality*) – ovvero la sovrapposizione e integrazione di elementi digitali nel mondo fisico – delinea oggi un nuovo paradigma, in cui lo spazio virtuale diviene luogo possibile dell'azione e del pensiero.

In questo scenario, la realtà virtuale si configura non più come un "altrove" simulato, ma come un laboratorio di sperimentazione della realtà, in cui vengono messi alla prova modelli concettuali, dinamiche di sistema e strutture narrative. I videogiochi, in particolare, possono

la cultura del gioco sono diventati una parte della ricerca nell'ambito DH. Per una storia di integrazione tra videogiochi e Digital Humanities, si veda Jones 2016, 84-97.

2 William Gibson è noto per aver coniato il termine cyberspace nel racconto *Burning Chrome* (1982) e successivamente reso celebre nel romanzo *Neuromancer* (1984), considerato uno dei testi fondativi del cyberpunk. In *Neuromancer*, il cyberspazio è definito come una «consensual hallucination experienced daily by billions of legitimate operators». Come è noto, molte delle tecnologie più sofisticate sono state anticipate, sotto forma di immaginazione narrativa, dalla letteratura fantascientifica.

3 Jones 2012.

essere letti come sistemi algoritmici di simulazione, nei quali la libertà del giocatore si esercita entro vincoli formalizzati. In essi, il *gameplay* diventa una forma di modellazione ricorsiva, in cui ogni azione è un atto di interrogazione del sistema rappresentato e conseguente apprendimento. A differenza dei media tradizionali, che sono prevalentemente rappresentazionali, le simulazioni videoludiche incorporano modelli di comportamento che reagiscono agli input secondo condizioni programmate, trasformando l'interazione in una forma di partecipazione strutturata e critica. Questa trasformazione del digitale in spazio di esercizio cognitivo e operativo trova una sua cornice teorica nel concetto di *remediation*, così come formulato da Bolter e Grusin (1999): ogni nuovo medium, lungi dal nascere in modo isolato, rimedia e rielabora le forme espressive e le strutture percettive dei media precedenti. La realtà virtuale, in particolare, si situa in una tensione costante tra immediatezza e ipermediazione: da un lato, mira a creare un'illusione di trasparenza, cancellando le tracce della mediazione (*immediacy*), dall'altro, rende visibile e articolata la propria interfaccia, moltiplicando i livelli informativi e i canali di accesso (*hypermediacy*).⁴ Questo tipo di ambienti apre scenari inediti anche nelle scienze umane, in particolare per la costruzione di modelli storici, testuali e iconografici. Attraverso la visualizzazione e la manipolazione di ambienti narrativi, si genera una nuova relazione tra conoscenza, spazio e percezione: il digitale cessa di essere un mero supporto o archivio e si afferma come estensione interpretativa del reale. Nel contesto educativo e delle scienze umane, la rimedializzazione ipermediatica non è un limite, bensì una risorsa: consente di visualizzare le logiche della mediazione, di renderle oggetto di riflessione critica e di utilizzarle come strumenti di accesso stratificato alla conoscenza. L'apprendente non è solo utente o spettatore, ma agente consapevole all'interno di una struttura che modella – e a sua volta interroga – la realtà. Un ruolo centrale è svolto dalla *gamification*, ovvero l'impiego intenzionale di elementi ludici in contesti non ludici,⁵ vale a dire il trasferimento di principi e modelli (*patterns*) di design dei giochi a diverse applicazioni interattive e persino configurazioni sociali. In ambito edu-

4 Si veda Bolter-Grusin 1999, pp. 21-50.

5 Deterding *et al.* 2011 (p. 2425) definiscono *gamification* «an informal umbrella term for the use of video game elements in non-gaming systems to improve user experience and user engagement».

cativo e scientifico, la *gamification* è oggi riconosciuta come strategia efficace per aumentare la motivazione, l'attenzione e il coinvolgimento attivo degli utenti, attivando processi cognitivi profondi, sostenuti dalla ripetizione, dalla scoperta, dalla partecipazione e dalla critica dei contenuti, in linea con i principi della didattica costruttivista⁶ e del *learning by doing*.⁷ Quando integrata in ambienti immersivi, essa non si limita a “rendere il contenuto più accessibile”, ma trasforma la struttura dell'apprendimento stesso, rendendola interattiva, situata e dinamica.

In questo ambito si colloca *Divine Reality*, sviluppato all'interno del più ampio progetto D.A.N.T.E. (*Digital Archive and New Technologies for E-content*), finanziato dal fondo V:ALERE 2022 dell'Università della Campania “Luigi Vanvitelli”. D.A.N.T.E. rappresenta un'iniziativa interdisciplinare, che interseca la storia della stampa, l'iconografia, le Digital Humanities e le tecnologie immersive, nutrendo del suo portato il portale *DanteLimina.it* (www.dantelimina.it). Uno degli obiettivi del progetto D.A.N.T.E. è quello di catalogare, analizzare e interpretare il patrimonio iconografico della *Commedia*, sia in forma cartacea, sia attraverso un'infrastruttura digitale che consenta una fruizione interattiva e dinamica dei contenuti. Su un altro versante, *Divine Reality* rappresenta un'ancora digitale volta a sperimentare metodologie educative che, attraverso la *gamification*, mirano a valorizzare e diffondere le illustrazioni quattrocentesche alla *Commedia*. Questo approccio consente di coniugare lo studio dell'iconografia, della storia del libro e della stampa con l'innovazione tecnologica, offrendo un'esperienza di apprendimento immersiva e partecipativa. La traduzione dell'universo dantesco in ambienti digitali implica una rimediazione in chiave educativa complessa tra fedeltà letteraria e licenza interpretativa. Nel caso di *Divine Reality*, il team *DanteLimina* insieme al responsabile scientifico del progetto, Ciro Perna, ha declinato l'esperienza da proporre al pubblico attraverso due modalità partecipative, entrambe in un contesto di realtà virtuale.

La prima prova è costituita da un filmato immersivo a 360° della durata di 6 minuti, sviluppato dal team di Re:LAB. Il trailer è stato costruito utilizzando asset 3D modellati in Blender e integrati all'interno del motore grafico Unreal Engine. L'obiettivo è coinvolgere lo spettatore attraverso un'esplorazione in terza persona del percorso dei

6 Piaget 1969.

7 Dewey 1938.

due *viatores* all'interno di un ambiente che ricostruisce fedelmente la celebre mappa dell'*Inferno* di Sandro Botticelli.⁸ L'utente segue quindi Dante e Virgilio nel loro viaggio, partendo dalla selva oscura, passando per la porta dell'*Inferno* fino alla riva dell'Acheronte, dove incontrano Caronte. La scelta di un video a 360° consente di coniugare la narrazione testuale con la dimensione visiva legata al capolavoro Botticelliano, traducendo il percorso poetico in un'esperienza immersiva. Al termine di questa prima sperimentazione, sono emerse due criticità rilevanti che hanno orientato le successive fasi di sviluppo. In primo luogo, il file non risultava pienamente ottimizzato per una fruizione diretta attraverso il visore Meta Quest 3, il che ha compromesso l'immediatezza dell'accesso e l'autonomia dell'esperienza da parte dell'utente. In secondo luogo, l'elevata richiesta di risorse hardware imponeva l'uso di dispositivi costosi e poco accessibili, limitando fortemente la replicabilità in contesti educativi o museali. A queste difficoltà tecniche si è aggiunta una limitata dimensione interattiva, che ha reso l'esperienza principalmente esplorativa e visiva, finendo per non attivare pienamente quei meccanismi di coinvolgimento cognitivo e partecipazione che, come appena anticipato, risultano essenziali per il consolidamento dei contenuti e il mantenimento dell'attenzione nel tempo. Nonostante ciò, il trailer mantiene una notevole forza comunicativa e un alto valore formativo, grazie all'aderenza topografica al disegno botticelliano, che consente allo spettatore di orientarsi visivamente lungo il percorso danteresco con coerenza e suggestione visiva.

Per superare questi ostacoli, è stato avviato un secondo sviluppo, progettato da chi scrive in collaborazione con il gruppo Remedia srl,⁹ con l'obiettivo di superare le criticità del primo trailer e incrementare in modo significativo la varietà delle interazioni e l'immediatezza operativa. Questo nuovo ambiente virtuale è interamente programmato in Unreal Engine, utilizzando un'architettura modulare ottimizzata per visori Meta Quest 3. L'esperienza è stata costruita su una logica di interazione diretta con oggetti fisici e interfacce contestuali semplificate [Fig. 7, Fig. 8] per estendere l'accessibilità anche a un pubblico non specializzato. È stato concepito come un ambiente didattico immersivo ad

8 Ms. Città del Vaticano, Biblioteca Apostolica Vaticana, Reg.lat.1896.pt.A, c. 101r.

9 Remedia Srl è una realtà consolidata nel panorama nazionale della produzione culturale e multimediale con una solida esperienza nell'ideazione di contenuti ad alto valore artistico.

alto grado di coinvolgimento, volto a riprodurre, con rigore filologico e accuratezza storica, l'esperienza di una stamperia cinquecentesca, restituendo la dimensione visiva e soprattutto quella gestuale e operativa della stampa. All'interno di questo spazio virtuale, l'utente può esplorare visivamente gli ambienti e gli strumenti tipici della stampa antica e interagire attivamente con essi mediante un sistema di feedback *hands-based* in realtà virtuale. Tale modalità consente, ad esempio, di azionare una pressa tipografica [Fig. 3, Fig. 4], manipolare matrici xilografiche, sollevare fogli virtuali e riprodurre gesti tipici dell'officina editoriale rinascimentale, come l'inchiostratura e la disposizione dei fogli sui telai. L'intero sistema è implementato in Unreal Engine attraverso un *Blueprint Event Graph* [Fig. 5], che permette di gestire interazioni e gesti, simulando il comportamento degli oggetti in risposta al movimento e alla pressione delle mani dell'utente. In parallelo, è possibile navigare virtualmente una biblioteca [Fig. 6] dove sono visualizzabili le scansioni digitali di importanti edizioni illustrate della *Commedia* – in particolare [*La Commedia col commento di Cristoforo Landino*] (Brescia, Bonino, 1487) e *La comedia di Dante Alighieri con la nuova esposizione di Vellutello* (Venezia, Marcolini, 1544), entrambe conservate alla Biblioteca Nazionale di Napoli, digitalizzate nell'ambito del progetto.¹⁰ La biblioteca è concepita come una galleria che consente di mettere a confronto e approfondire i dettagli delle soglie iconografiche dei due testi. A incorniciare l'intera esperienza, un sistema di *storytelling* collega la dimensione storica della stampa all'attualità della conservazione digitale, traducendo la narrazione in ambienti immersivi ricostruiti con tecniche di modellazione 3D. In particolare, la scenografia virtuale – che include pavimentazioni originali, scaffalature, banchi da stampa e ambienti bibliotecari – è stata generata a partire da rilievi fotogrammetrici ad alta definizione e da scannerizzazioni LiDAR (*Light Detection and Ranging*), successivamente rielaborati e texturizzati in ambienti 3D. [Fig. 9] Questa combinazione di rilievo reale e modellazione digitale permette di restituire topografia e consistenza materica degli spazi, favorendo un'esperienza immersiva coerente e fondata. Questa doppia traiettoria

10 I rispettivi esemplari sono: Biblioteca Nazionale di Napoli, S.Q. 23. I. 20 = *La Commedia col commento di Cristoforo Landino*, Brescia, Bonino de' Bonini, 31 maggio 1487 e Biblioteca Nazionale di Napoli, BIBL. BRANC Sez. Dant. D 0032 = *La Comedia di Dante Alighieri con la nova esposizione di Alessandro Vellutello*, Venezia, Francesco Marcolini, giugno 1544.

– ambientale e documentaria – rende lo *storytelling* non solo un elemento di cornice ma un dispositivo strutturante dell'intera esperienza, capace di connettere passato e presente, dimensione materiale e codifica digitale, memoria culturale e pratiche innovative di accesso alla conoscenza: un nuovo modulo che integra la dimensione storica artistica e materiale del libro antico con una riflessione critica sulle tecnologie digitali di conservazione e valorizzazione, in linea con gli obiettivi generali del progetto *Limina*. L'ambiente virtuale descritto è attualmente in una fase di programmazione avanzata e verrà rilasciato a settembre 2025 all'interno dell'infrastruttura D.A.N.T.E., completamente fruibile tramite visore Meta Quest 3.

La progettazione di questo secondo prototipo, sviluppata sulla base degli ostacoli emersi nella prima fase sperimentale, ha rappresentato un'opportunità per interrogarsi su questioni teoriche e concrete poste in apertura: l'importanza dell'interattività, non come valore aggiunto ma come struttura portante del processo educativo; la coerenza storica e iconografica, necessaria per fondare la credibilità e la solidità interpretativa delle ricostruzioni e infine, la rimedializzazione, che supera il concetto di trasposizione da un *medium* all'altro, ma diventa un processo consapevole di trasformazione e rinegoziazione del senso e della forma.

Questo progetto ha permesso di sperimentare sul campo come una narrazione immersiva, sostenuta da un impianto scientifico solido e da un design interattivo accessibile, possa rendere fruibili contenuti complessi come la storia editoriale e iconografica della *Commedia*, e soprattutto stimolare una riflessione più ampia su come le tecnologie digitali ridefiniscano i modi dell'apprendere, del ricercare e del ricordare. In questo senso, *Divine Reality* non è soltanto un prodotto didattico o una dimostrazione tecnica: è un dispositivo critico, un campo di prova in cui si mettono in discussione i confini tra esperienza e conoscenza, tra spazio museale e spazio virtuale, tra interpretazione e manipolazione. Nel vasto ambito della ricerca, necessariamente interdisciplinare, è stato sperimentato come, all'interno della costruzione di uno strumento immersivo, convergano competenze provenienti da ambiti differenti: storia del libro, filologia, ingegneria del software, interaction design e pedagogia. Inoltre, è stato dimostrato una volta di più che le DH rappresentano un contesto fertile per la sperimentazione di modelli ibridi di trasmissione culturale. Ogni fase del lavoro ha permesso di generare nuove questioni e nuove forme di accesso alla conoscenza, non limitandosi solo alla fase di realizzazione tecnica dell'ambiente. Guar-

dando al futuro, naturalmente non sarebbe saggio considerare questo progetto, né gli altri sviluppati all'interno dell'universo *DanteLimina.it*, come iniziative concluse e impermeabili all'evoluzione dei linguaggi e delle tecnologie. Al contrario, si intende concepirli come un osservatorio permanente, un'infrastruttura critica e operativa da cui analizzare, interpretare e, auspicabilmente, orientare le traiettorie emergenti dell'interazione tra media digitali e patrimonio culturale. In quest'ottica, *Divine Reality* è un modello evolutivo, destinato ad arricchirsi di nuove funzioni, contenuti e livelli di complessità e soprattutto aperto alla co-creazione con i suoi utenti finali, siano essi studenti, studiosi, insegnanti o semplici visitatori. Solo mantenendo questa tensione costante tra sperimentazione e rigore, tra innovazione e riflessione, sarà possibile creare dei veri strumenti di conoscenza trasformativa, in grado di ridefinire il nostro modo di leggere, insegnare e custodire conoscenze e testi del passato.



Fig. 1. Screenshot del trailer con Dante e Virgilio che si muovono verso la riva dell'Acheronte.

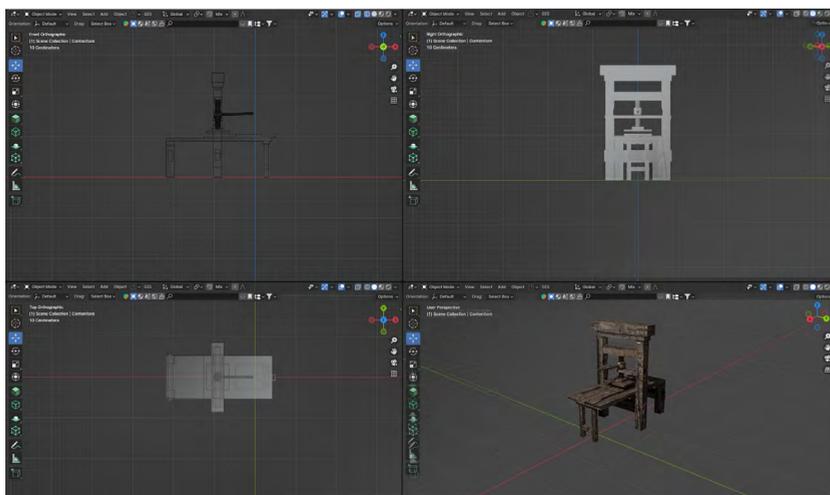


Fig. 2. Screenshot del trailer con Caronte che si avvicina allo spettatore.



Fig. 3. Ricostruzione di una pressa tipografica Cinquecentesca in Blender; out-line e modellazione solida con e senza texture.

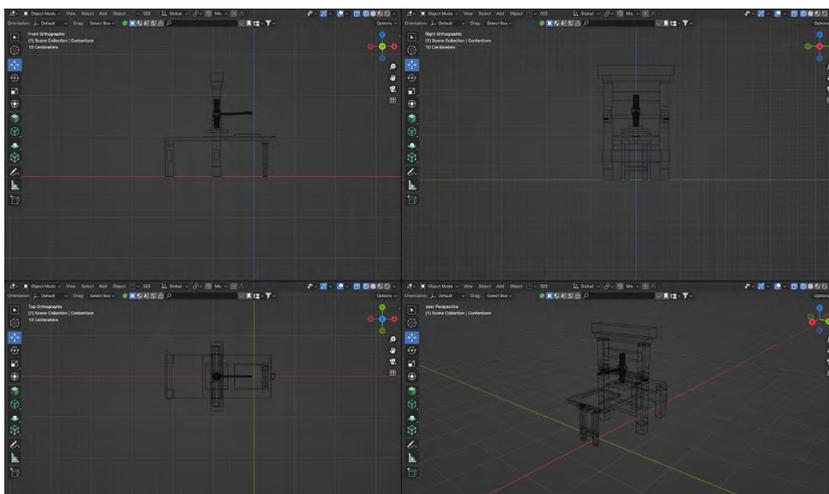


Fig. 4. Outline della pressa tipografica in 4 viste differenti.

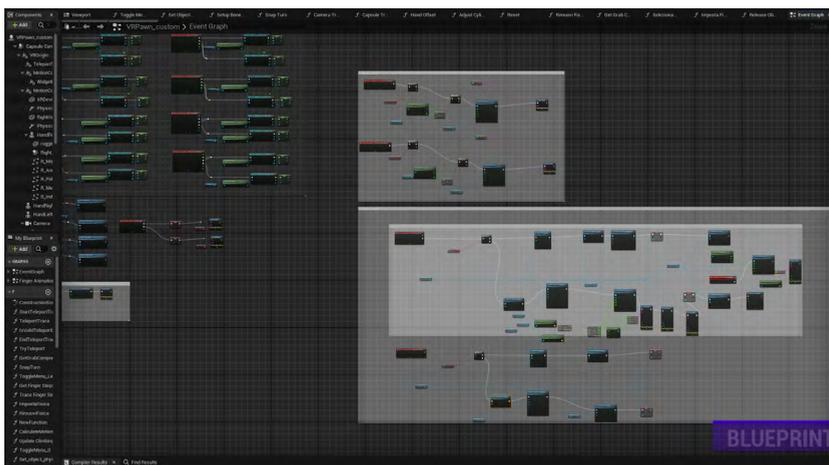


Fig. 5. Event Graph del Blueprint in Unreal Engine, con focalizzazione sulle funzioni VR hands-based.

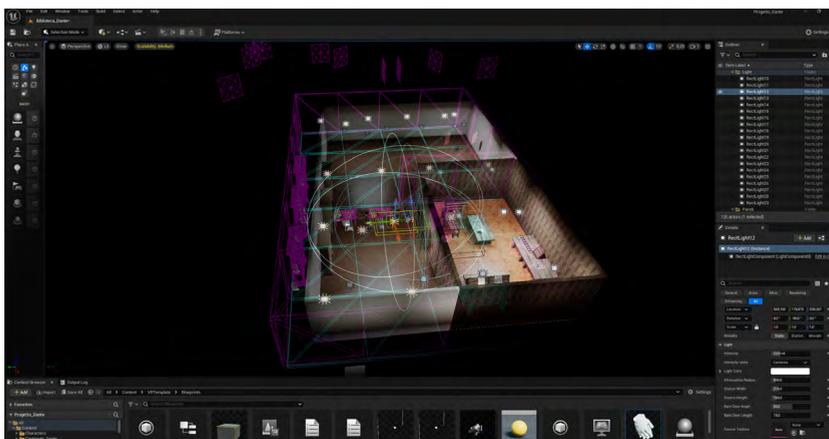


Fig. 6. Vista dall'alto del modello tridimensionale della stanza usata per i test sulle funzionalità di movimento e delle interazioni con l'applicazione dell'outline che evidenzia i punti luce e i bordi degli oggetti interagibili.



Fig. 7. Screenshot catturato direttamente dal visore di un dettaglio della pressa con matrice e stampa poggiati sopra (materiali di prova).

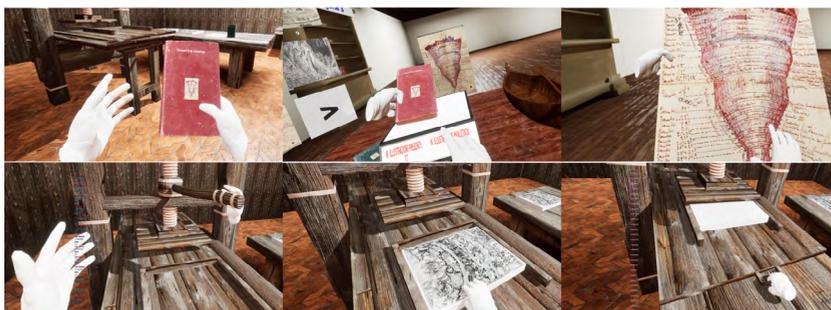


Fig. 8. Screenshot catturato direttamente dal visore di un dettaglio della pressa con matrice e stampa poggiati sopra (materiali di prova).



Fig. 9. Scansioni 3D con la tecnica del Gaussian Splatting, in queste immagini 3 fasi del procedimento che partono dall'acquisizione di una nuvola di punti all'avere un oggetto tridimensionale con tanto di texture.

Bibliografia

- Bolter-Grusin 1999 = Jay David Bolter, Richard Grusin, *Remediation. Understanding New Media*, Cambridge (Mass), MIT Press, 1999.
Gibson 1984 = William Gibson, *Neuromancer*, New York, Ace Books, 1984.
Gibson 1986 (1982) = William Gibson, *Burning Chrome*, New York, Ace Books, 1986.

- Deterding *et al.* 2011 = Sebastian Deterding, Miguel Sicart, Lennart Nacke, Kenton O'Hara, Dan Dicon, *Gamification: Using Game Design Elements in Non-Gaming Contexts*, in *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems: proceeding, Vancouver, BC, Canada - May 07-12, 2011*, New York, ACM Press, pp. 2425-2428.
- Dewey 1938 = John Dewey, *Experience and education*, Indianapolis (US), Kappa Delta Pi, 1938.
- Jones 2012 = Steven E. Jones, *The Emergence of the Digital Humanities (as the Network Is Everting)*, in Matthew K. Gold (ed.), *Debates in the Digital Humanities*, University of Minnesota Press, 2012. <https://dhdebates.gc.cuny.edu/read/untitled/section/09efe573-98e0-4a10-aaa3-e4b222d018fe>
- Jones 2016 = Steven E. Jones, *New Media and Modeling Games and the Digital Humanities*, in *A New Companion to Digital Humanities*, Hoboken, John Wiley & Sons, Ltd, 2016, pp. 84-97.
- Piaget 1969 = Jean Piaget, *Psychologie et pédagogie*, Paris, Denoël, 1969.



10.69118/2025_L6

Building and maintaining the crowdsourced archive *Dante Today**

Elizabeth Coggeshall, Arielle Saiber

Florida State University, Johns Hopkins University
ecoggeshall@fsu.edu, asaiber@jhu.edu

Creazione e mantenimento dell'archivio crowdsourced *Dante Today*

Abstract (ITA)

L'archivio digitale curato e realizzato in crowdsourcing *Dante Today* cataloga gli adattamenti e le rivisitazioni delle opere di Dante nelle culture contemporanee di tutto il mondo. Nel loro insieme, le oltre 2000 voci mostrano i modi in cui la gente di oggi intrattiene un dialogo transstorico con Dante. Questo dialogo prende forma attraverso diversi media: dai romanzi e fumetti alle arti visive e performative, dalla pubblicità ai meme e alla street art. Oltre a presentare alcuni reperti provenienti dal fondo dell'archivio, questo saggio illustra la storia e l'organizzazione dell'archivio e discute i piani futuri per il suo sviluppo.

Abstract (ENG)

The curated, crowdsourced digital archive *Dante Today* catalogues adaptations and appropriations of Dante's works across contemporary global cultures. Collectively, the 2000+ artifacts showcase the ways that people today engage in a transhistorical dialogue with Dante. This dialogue takes shape across multiple media: from novels and comics to visual and performing arts, advertisements to memes and street art. In addition to presenting a few artifacts from the archive's holdings, this essay presents the archive's history and organization, and discusses future plans for development.

Parole chiave / Keywords

crowdsourcing; archivi digitali; intermedialità; storia della ricezione / crowdsourcing; digital archives; intermediality; reception history

* Some of the material printed here has appeared previously in other formats. In particular, see Coggeshall 2022, pp. 253-272, and Coggeshall 2021, pp. 324-337. We thank the editors for the permission to reprint.

Arielle Saiber had long kept a cabinet-drawer full of “Dante sightings”, either given to her by students or friends, or that she had stumbled upon herself. These marvelous artifacts were just sitting there, quietly, in the dark. She would often refer to them in lectures and conversations, but never had them on hand to show people. As the internet developed in the early 2000s, she realized that she could place these items online as digitized scans or as links, organize them, think about the patterns that emerged when grouped together, and open the “cabinet” to contributors around the world. This is how *Dante Today: Citings and Sightings of Dante’s Works in Contemporary Culture* was born at Bowdoin College in Brunswick, Maine, in 2006.

Working with Bowdoin’s senior producer of academic media, David Israel, and a team of undergraduate students, Saiber migrated the newspaper and magazine clippings, photographs, scribbled lists of consumer products, and website URLs onto WordPress. From the outset, the archive – which seeks to capture and promote the effects of Dante’s works as they echo within, across, and even counter to contemporary cultures – was conceived as a crowdsourced repository to which anyone could contribute (the term was, coincidentally, coined the same year *Dante Today* went online).¹ But Saiber, then the sole editor of the site, decided *Dante Today* would also curate submissions as they rolled in, maintaining the archive’s commitment to the rigorous examination of Dante’s resonance around the world while also taking stock of where Dante’s different publics hear that resonance in their daily lives.²

In 2012, Elizabeth Coggeshall joined Saiber as co-editor of *Dante Today*. They – we – have worked together for over a decade with the support of IT staff, web designers, Digital Humanities specialists, and students from our home institutions and beyond. After nineteen years, many grants, and several major redesigns, *Dante Today* has holdings of nearly 2,200 artifacts, and the archive has secured its place as one of

1 On the complicated role of crowdsourcing as the mechanism for the growth of *Dante Today*, see Coggeshall 2022. For additional research directly addressing or stemming from our work on the archive, see Coggeshall 2023, Coggeshall 2020, Saiber in press, Saiber 2023, and Saiber 2020.

2 On the various ways that public humanities initiatives like *Dante Today* conceive of and engage with their respective “publics,” see Coggeshall 2025, pp. 145-153.

the most respected resources for understanding the impact of Dante's works on cultures around the world today.³

Our goals as archivists are twofold: we collect and preserve an array of ephemeral cultural artifacts – both material and digital – that may otherwise pass unnoticed or even disappear; and we host a platform for an unspoken dialogue between contemporary artists, writers, and other makers from around the world who utilize Dante's works as a lens through which to describe, analyze, and critique the world as they see it. Our archive offers a space for our users to listen to the varied ways Dante's voice has resonated in and across global cultures in the last century and beyond.

Some of the *Commedia's* stamp on twentieth- and twenty-first-century culture has been well-documented: a growing body of research testifies to the broad (but not always deep) imprint that the *Commedia* has left on Anglo-European literatures, cinema, visual art and architecture, music, and even comic books and graphic novels.⁴ But Dante's reach extends far beyond the borders of the Anglo-European world, just as it extends beyond what we think of as sanctioned genres or categories of reference. We see Dante in advertisements, in street art, in internet memes, on social media, on restaurant menus, in place names, and more. *Dante Today* serves as a single reference point for all these engagements, from high culture to mass media.

One of our project's cornerstones is its expanding bibliography of Dante reception (“[Bibliography](#)”). We have created a public Zotero library currently featuring over 400 scholarly resources examining the translation, adaptation, appropriation, and reimagining of Dante's works across media and cultures. This bibliography, however, can only scratch the surface of available materials, as scholarship on Dante's reception continues to grow. Like the posts themselves, the bibliography depends on crowdsourced research contributions.

The archive and the bibliography receive submissions (via the “[Submit a Citing](#)” tab) from students and teachers as well as artists, writers,

3 See, for example, the following essays on Dante reception that praise *Dante Today* as a central resource: Ciabattini-Marchesi (*ed.*) 2023; Lazzarin-Dutel (*ed.*) 2018; De Rooy 2017; Marchesi 2020, pp. 216-222; Moscardi 2016.

4 For an extensive, ever-growing bibliography of critical scholarship on Dante's resonance across media and cultures, see *Dante Today's* public Zotero library. https://www.zotero.org/groups/2604094/dante_today/library (accessed 30 April 2025).

scholars, performers, gallerists, editors, and other “Dante fans” from around the world. We then read through and approve these submissions, tagging, mapping, and categorizing them. We avoid imposing value judgments onto the references, as our primary objective for the archive has been to *document* the global reception of Dante’s works. We do not provide “expert” commentary on the artifacts in the collection, nor enable the participation of users through upvoting, “liking,” or commenting on individual posts. Rather, we think of our work strictly as that of archivists, who carefully select the objects included in the collection, thoughtfully devise their descriptions, and then step aside to let our audience engage with them as they will – through browsing media formats, tag lists, and the map, or using the search box. Each visitor employs their own metrics to explore the ways these artifacts use Dante’s works to comment on contemporary culture.

We have chosen not to classify references according to the depth of their engagement, nor do we assess the fidelity of a reference to the original poem. Fidelity and depth are not the only measures that matter in the analysis of these data. To use a well-known and much discussed example, the *Dante’s Inferno* video game (Fig. 1) might engage deeply with the original poem, but it is hardly an accurate retelling of the plot, what with a crusader pilgrim fighting to save his damsel-in-distress Beatrice.⁵ Does this matter to the way we evaluate the game’s value to research or in the classroom? Consider another example: the fifth episode of season one of the HBO television series *Westworld*, titled “Contrapasso” (Fig. 2). The episode makes no other explicit mention of Dante. Is the video game, for all its detail and popularity, any more interesting or relevant to the question of Dante’s resonance in contemporary culture than the title of the TV episode? We leave this question up to our users to answer.

Yet we do make decisions as to whether an artifact “makes the cut” and is verifiably related to Dante and his works in some way. To give an example: the *Simpsons* episode *Treehouse of Horror IV*, which aired in 1993 (S05E05) (Fig. 3). In the episode, Homer bargains his soul for a donut; when he caves to temptation and eats the donut, a devilish Ned Flanders swoops him off to Hell, where he is sent to the “Ironic Punish-

5 On its inaccuracy, see Barolini 2010, p. 79 and Saiber 2010. Essary and Powlesland argue for the value of the video game as an (albeit imperfect) analytical and/or pedagogical tool, opening a door to understanding the participatory experience of the poem: Essary 2019; Powlesland 2022.

ment Division” of Hell Labs. In Dantesque retribution, Homer is forced by the devil to binge-eat thousands of pink-glazed donuts, which he does, gleefully, much to the devil’s consternation. The episode makes no direct reference to Dante, but the very notion of an “ironic punishment” that reveals something of the nature of the crime itself (like Homer’s gluttony), comes to American popular culture directly through the influence and legacy of Dante’s poem.

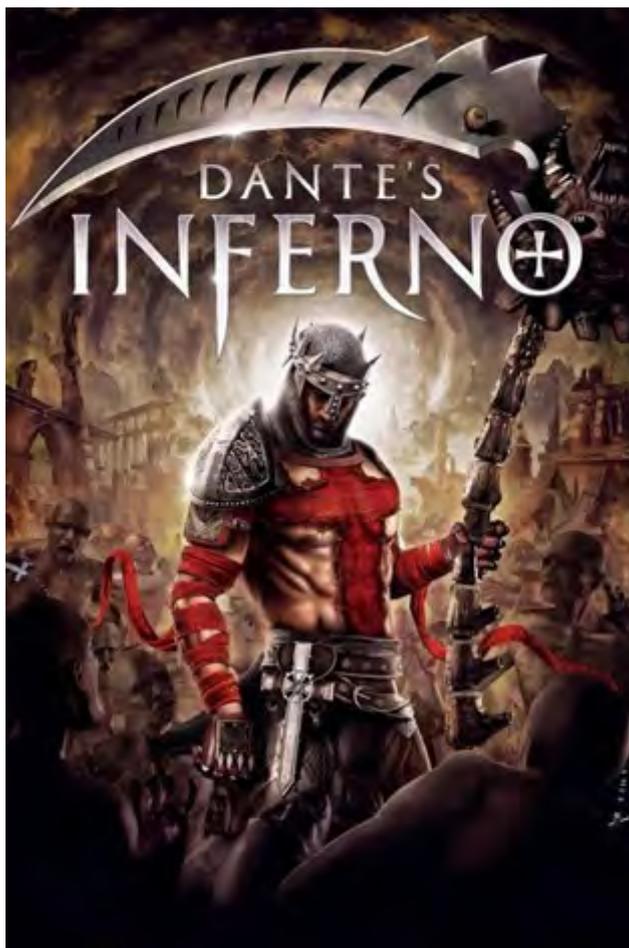


Fig. 1. Promotional image for EA Games’ *Dante’s Inferno* video game (2010).
Image from IMDB. https://www.imdb.com/title/tt1409575/mediaviewer/rm4293343489/?ref=tt_ov_i. Accessed 1 May 2025.



Fig. 2. Still from HBO's *Westworld*, season 1, episode 5, *Contrapasso* (2016). Image from IMDB. <https://www.imdb.com/title/tt4625866/mediaviewer/rm3773432064/>. Accessed 1 May 2025.



Fig. 3. Image from *The Simpsons*, season 5, episode 5, *Treehouse of Horror IV* (October 28, 1993). Image from IMDB. https://www.imdb.com/title/tt0701279/mediaviewer/rm3848421888/?ref=tt_ph_more. Accessed 1 May 2025.

1. *Dante Today*, Present and Future

After nearly twenty years of growth, *Dante Today* is entering a new phase. We have recently launched a complete redesign of the site. We worked with the award-winning team at Studio Rainwater and migrated the archive to the URL [dantetoday.org](https://www.dantetoday.org) in November of 2024. While the site is now publicly available, it is still *in allestimento*, as the data are still being transferred and refined. Our data have been reorganized to facilitate the search-and-filter process, either through a search box or a structured system of tags. We continue to repopulate our map (“[Map](#)”) as the data migration continues, but one can immediately see the massive worldwide circulation of Dante’s works (Fig. 4). Because we are in the United States and the majority of sightings we receive are from America, we explain to users why the USA seems, and is, overrepresented (“[How to Use this Website](#)”). We also have a take out page dedicated to resources for classroom use (“[Teaching Resources](#)”) where we have gathered suggestions from our own pedagogical practices and those of other teachers to compile prompts for discussion, writing, and research related to the archive. We are in the process of updating this page.



Fig. 4. *Dante Today* Map. <https://www.dantetoday.org/map/>. Author’s screenshot. Accessed 1 May 2025.

As mentioned above, the 2024 revision of *Dante Today* features a completely overhauled search-and-filter system. Previously, users could locate artifacts on the site by typing queries into a search box, by sifting through the nine media categories (Consumer Goods, Digital Media, Dining and

Leisure, Music, Odds and Ends, Performing Arts, Places, Visual Arts and Architecture, and Written Word), or by selecting from thousands of available tags. The redesign involved reverse-engineering the site's existing data architecture by organizing the thousands of existing tags into a coherent taxonomy. During data migration, our team methodically reclassified each entry according to this new taxonomy, substantially expanding filtering options and improving navigability. The user can now filter for much more granularized topics, such as performance pieces featuring the character Beatrice, or street art that adapts Gustave Doré's lithographs to political ends.

In addition to the great number and variety of references one can discover on *Dante Today*, users are also privy to narrative threads hidden therein. We conceive of our archive much like a museum collection in which various stories are concealed: the proliferation of self-help books that use Dante's poem as a model for spiritual and psychological betterment; a Jesuit priest and antiwar activist, whose rewriting of *Purgatorio* condemns the nuclear arms race, the military-industrial complex, and interreligious conflict; the myriad ways the poem has been revisited in China or East Africa. As it stands, each of these artifacts can only be discovered through brief posts, tags, or targeted search. Thus far, we have left the work of uncovering those narratives squarely in the hands of the user, each going on their own mission to find what connects artifacts across times, spaces, media formats, or themes. In the near future, however, we plan to feature some of those embedded narratives in a "gallery" that would parse and analyze select materials held in the collection, organized into digital exhibits. We hope these "exhibits" will inspire users to do more digging.

One exhibit we have in mind is a collation and discussion of the #Dante2018 phenomenon, a virtual collective reading that went viral in the first months of 2018. In mid-November of the previous year, Pablo Maurette, an Argentinian essayist and then professor of literature at a small college in the American Midwest, tweeted his intention to reread Dante's *Divine Comedy*, one canto per day, for the first hundred days of the new year. Within a few weeks, Maurette's tweet had mobilized a mass of fellow readers across the Spanish-speaking world. With the hashtag #Dante2018, Twitter users shared a diverse array of content – from quotations and commentary to graffiti, illustrations, and selfies – documenting their collective journey through the poem. The reading stimulated illustrators, designers, and performers to create and distribute

original works responding to their reading. The initiative attracted an impressive following: an informal poll partway through found some five thousand participants spread across five continents. The initiative culminated in performances and live readings sponsored by the Comune di Ravenna, featuring Maurette and other writers, artists, and contributors, who came together to conclude their collective reading experience. We were made aware of #Dante2018 near the end of its run, but we collaborated with Maurette to document some of the art that was produced during the #Dante2018 moment. These can be found scattered through the *Dante Today* archive. But the gallery exhibit would collect them all in a single virtual space, allowing the user to learn about this mass reading initiative that brought together a mixed community of active participants for an enriching collective experience.

One of the great privileges of working behind-the-scenes on the site over these years has been the opportunity for us to speak with – and even meet – a diverse group of creatives, whose engagement with Dante always reinvigorates our teaching and our scholarship. We have done on-site interviews with the master goldsmith Donna Distefano in New York City and visionary artists such as Kazumasa Chiba in Tokyo and Paul Laffoley in Boston. We have met with renowned makers of all genres, such as Mary Jo Bang (poet, USA), Sandow Birk (painter, USA), Kimberly Campanello (poet, USA/UK), Marina Ceballos (graphic artist, Argentina), Lorna Goodison (poet, Jamaica), Kat Mustatea (performance and digital media artist, USA), Matthew Pearl (novelist, USA), John Sayles (filmmaker, USA), Timothy Schmalz (sculptor, Canada), and BL Shirelle (musician and producer, USA). Our students have also followed our lead and corresponded with makers, such as with the rapper GM Grimm, who was in prison at the time, on his album *Digital Tears: E-mail from Purgatory* (2004). The gallery will open a space for these kinds of conversations to be shared.

2. Discussing the *Divine Comedy* with Dante

The dialogues catalyzed by *Dante Today* are part of a larger global conversation among artists, writers, performers, filmmakers, advertisers, and other culture-makers who think through Dante's poem to speak to and about their world. This great conversation is powerfully figured in the 2006 oil painting *Discussing the Divine Comedy with Dante* by the Chinese artists Dai Dudu, Li Tiezi, and Zhang Anjun (Fig. 5).

The painting features an ahistorical (or, better, antihistorical) mashup of 103 world figures, from Ramses II and Genghis Khan to Luciano Pavarotti and Mike Tyson. Some are gathered around tables, some are seated on the floor, some are milling about. Some look at the viewer, others scan the room, and still others seemingly fix their eyes on a specific figure in their gathering space. Some hold tools, either of their own trade or of someone else's: Tolstoy plays an accordion, George W. Bush gazes through a spyglass, Mao Zedong holds a cigarette, Hemingway a whiskey bottle, and Che Guevara a rifle. Meanwhile, in the upper right-hand corner of the image, elevated above the fray, are four individuals: the three artists and the *sommo poeta*, who holds a copy of his poem open in his hand (Fig. 6). The four figures gaze down silently on the spectacle of history unfolding below.

When we redesigned the *Dante Today* site in 2024, we selected *Discussing the Divine Comedy with Dante* as the new backdrop for the landing page (Fig. 7). The image celebrates how many people have “talked” with Dante in their work about the “divine comedy” of human experience in all its facets, expressions, histories, cultures, and languages. Our work as archivists has been to find, curate, and listen in to these dialogues, and to create space for our users to do the same.

Over the past nineteen years, we have consistently aimed to welcome the widest possible userbase to the expanding collection. Although our visitor metrics have always been strong, extending and diversifying our reach remains a key priority. Once all the data is fully loaded onto the site, we will launch a publicity and promotion campaign – which we have never done – and cultivate a presence on social media. We will continue to seek out new features for the site, new collaborations, and new ways of helping users both contribute and engage with the data. *Dante Today* is a living, dynamic archive that will continue to evolve, as technology and culture do, and as our human collective keeps discussing the *Divine Comedy* with Dante.



Fig. 5. Dai Dudu, Li Tiezi, Zhang Anjun. *Discussing the Divine Comedy with Dante*. Author's screenshot; image captured from <http://cliptank.com/PeopleofInfluencePainting.htm>. Last accessed 1 May 2025.



Fig. 6. Dai Dudu, Li Tiezi, Zhang Anjun. *Discussing the Divine Comedy with Dante* (detail). Author's screenshot; image captured from <http://cliptank.com/PeopleofInfluencePainting.htm>. Last accessed 1 May 2025.

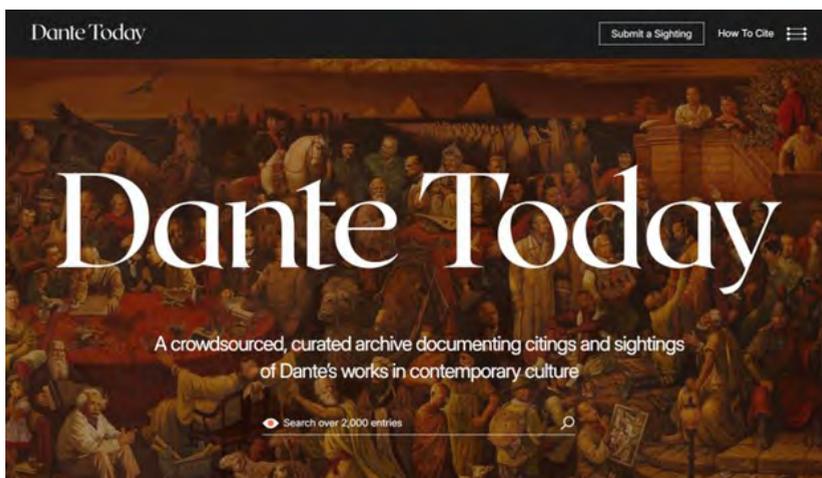


Fig. 7. *Dante Today* Homepage. <https://www.dantetoday.org/>. Author's screenshot. Accessed 1 May 2025.

Bibliografia

- Barolini 2010 = Teodolinda Barolini, *Expert View*, *Entertainment Weekly* (26 February 2010).
- Ciabattioni-Marchesi (ed.) 2023 = *Dante Alive: Essays on a Cultural Icon*, ed. Francesco Ciabattioni and Simone Marchesi, London, Routledge, 2023.
- Coggeshall 2020 = Elizabeth Coggeshall, *Dante's Afterlife in Popular Culture*, in *Approaches to Teaching Dante's Divine Comedy*, 2nd edition, ed. Christopher Kleinhenz and Kristina Olson, New York, Modern Language Association, 2020, pp. 185-191.
- Coggeshall 2021 = Elizabeth Coggeshall. *Dante Today: Tracking the Global Resonance of the Commedia*, in *Dante Beyond Borders: Context and Reception*, ed. Nick Havely and Jonathan Katz, Legenda, 2021, pp. 324-337
- Coggeshall 2022 = Elizabeth Coggeshall, *Discussing the Divine Comedy with Dante: On Crowdsourcing and Transcultural Resonance*, «Bibliotheca Dantesca» 5, (2022), pp. 253-272. <https://repository.upenn.edu/entities/publication/8c51d223-120b-44c2-96a8-193d5ddb7a48>
- Coggeshall 2023 = Elizabeth Coggeshall, *The Hell Franchise: Dante's Commedia in American Marketing* in *Dante Alive: Essays on a*

- Cultural Icon*, ed. Francesco Ciabattoni and Simone Marchesi, London, Routledge, 2023, pp. 195-212.
- Coggeshall 2025 = Elizabeth Coggeshall, *Resonant Texts in Noisy Spaces: On the 'Publics' of the Public Humanities*, PMLA 140.1, January 2025, pp. 145-153.
- De Rooy 2017 = Ronald De Rooy, *Divine Comics*, «European Comic Art», 10.1, Spring 2017, pp. 94-109.
- Essary 2019 = Brandon Essary, *Dante's 'Inferno', Video Games, and Pop Pedagogy*, «Parole rubate/Purloined Letters», 20 (2019), pp. 59-82.
- Lazzarin-Dutel (ed.) 2018 = *Dante pop: la Divina Commedia nella letteratura e nella cultura popolare contemporanea*, ed. Stefano Lazzarin and Jérôme Dutel, Manziana, Vecchiarelli, 2018.
- Marchesi 2020 = Simone Marchesi, *Dante's Comedy as First-Year Seminar: From Early Engagement to Self Reliance*, in *Approaches to Teaching Dante's Divine Comedy*, 2^a ed., ed. Christopher Kleinhenz and Kristina Olson, New York, Modern Language Association, 2020, pp. 216-222.
- Moscardi 2016 = Iuri Moscardi, *Dal Maine un super sito su Dante*, *La Stampa* (5 June 2016).
- Powlesland 2022 = Katherine Powlesland, *Dante and Video Games: The Unrealised Potential of the Virtual Commedia*, «Italian Studies», 77.2 (2022), pp. 146-156.
- Saiber 2010 = Arielle Saiber, in Benjamin Popper, *Dante Alighieri: Epic Poet, Ass Kicker*, *The Atlantic* (February 2010). <https://www.theatlantic.com/magazine/archive/2010/02/dante-alighieri-epic-poet-ass-kicker/307936/>.
- Saiber 2020 = Arielle Saiber, *Hell, yes! Dante in Contemporary American Satire* in *Dante Satiro: The Concept of Satire in the Middle Ages and in Dante's Works*, ed. Fabian Alfie and Nicolino Applauso, Lanham, Rowman & Littlefield, 2020, pp. 171-186.
- Saiber 2023 = Arielle Saiber, *Dante and American Science Fiction*, in *Dante Alive: Essays on a Cultural Icon*, ed. Francesco Ciabattoni and Simone Marchesi, London, Routledge, 2023, pp. 148-160.
- Saiber *in press* = Arielle Saiber, *L'amor che move: Dante's Paradiso in Contemporary Culture* in *Dante's Paradiso: A Reader's Guide*, ed. Filippo Gianferrari and Ronald Herzman, London, Routledge, forthcoming.



10.69118/2025_L7

Custodire il passato per il futuro: la transizione digitale della Biblioteca Apostolica Vaticana

Paola Manoni

*Biblioteca Apostolica Vaticana
manoni@vatlib.it*

Preserving the Past for the Future: the digital transition of the Vatican Apostolic Library

Abstract (ITA)

Le collezioni digitali della Biblioteca Apostolica Vaticana (BAV) costituiscono sistemi complessi, in cui elementi interconnessi determinano la struttura e l'accessibilità delle risorse. La gestione di tali sistemi richiede l'integrazione di piattaforme distinte, capaci di operare in sinergia per garantire un'efficace interoperabilità. La piattaforma digitale DVL (Digital Vatican Library) della BAV adotta API conformi allo standard IIIF (International Image Interoperability Protocol), assicurando la fruizione avanzata dei contenuti. La seconda fase di implementazione del IIIF, realizzata in BAV, ha portato alla creazione della piattaforma Thematic Pathways on the Web, dedicata alle mostre virtuali, che arricchisce l'esperienza utente mediante l'applicazione di annotazioni ai manoscritti, migliorando la valorizzazione del patrimonio culturale e l'accesso alle risorse digitalizzate.

Abstract (ENG)

The digital collections of the Vatican Apostolic Library (BAV) constitute complex systems where interconnected elements define their structure and accessibility. Managing these systems requires the seamless integration of distinct platforms designed to operate in synergy, ensuring robust interoperability. The BAV's digital platform, DVL (Digital Vatican Library), employs APIs compliant with the IIIF (International Image Interoperability Protocol) standard, facilitating advanced access to digitized content. In its second phase of IIIF implementation, the BAV developed Thematic Pathways on the Web, a platform dedicated to virtual exhibitions. Through the application of manuscript annotations, this initiative enhances user engagement, enriches the cultural heritage experience, and broadens access to digital resources.

Parole chiave / Keywords

*Collezioni digitali, Biblioteca Apostolica Vaticana, Digital Vatican Library /
Digital collections, Vatican Apostolic Library, Digital Vatican Library*

1. Organizzazione delle collezioni digitali

Le collezioni digitali della Biblioteca Apostolica Vaticana (BAV) rappresentano entità articolate, composte da elementi interconnessi che ne determinano struttura e accessibilità. In primo luogo, comprendono le immagini che costituiscono l'oggetto digitale, come i fogli di un codice manoscritto, le pagine di un incunabolo, stampe, disegni, monete e medaglie: i gemelli digitali delle collezioni storicamente conservate nei magazzini dei Dipartimenti Manoscritti, Stampati e Numismatico della Biblioteca. Gli oggetti digitali includono i metadati descrittivi, fondamentali per identificare e comprendere la risorsa digitale pubblicata. Questi metadati contengono tutte le caratteristiche rilevanti, oggetto di catalogazione. I metadati strutturali si affiancano ad essi, risultando essenziali per creare connessioni logiche tra le immagini e altre funzioni. Ad esempio, rendono possibile l'associazione di un indice dei contenuti, facilitando così la navigazione e l'interazione con i materiali digitali.

La gestione di sistemi interconnessi richiede l'impiego di piattaforme distinte, concepite per operare in sinergia. I metadati descrittivi, tradizionalmente legati ai cataloghi (OPAC), sono integrati con software per la gestione delle immagini digitali. Queste informazioni sono tipicamente gestite nei record dei cataloghi e rese disponibili attraverso applicazioni con interfacce web per la ricerca. Nel contesto delle biblioteche digitali l'interazione di tali sistemi si svolge mediante API (Application Programming Interface), che consentono lo scambio di dati e funzionalità tra le piattaforme, senza necessità di un'integrazione diretta. L'introduzione delle API rappresenta un pilastro fondamentale per le biblioteche digitali, in quanto consente di armonizzare i dati provenienti da fonti eterogenee, offrendo agli utenti un'esperienza di consultazione più efficace e integrata. Grazie all'impiego delle tecnologie *open data*, le API favoriscono una maggiore interoperabilità e migliorano significativamente l'integrazione tra i sistemi, superando le barriere tradizionali legate alla frammentazione dei database.

La piattaforma digitale della Biblioteca Apostolica Vaticana (BAV), nota con l'acronimo DVL (Digital Vatican Library),¹ utilizza API conformi allo standard IIIF (International Image Interoperability Protocol).² Il DVL è stato progettato seguendo le specifiche funzionali richieste dalla

1 Cfr. <https://digi.vatlib.it> (data di consultazione 28/03/2025).

2 Cfr. <https://iiif.io> (data di consultazione 28/03/2025).

Biblioteca. Implementata come un sistema integrato, la piattaforma consente di combinare strumenti di ricerca avanzati con meccanismi innovativi per l'accesso ai dati. Ogni collezione è aggiornata in tempo reale, ivi incluse le informazioni sul numero di esemplari disponibili, e organizza i materiali digitalizzati all'interno di gallerie collegate ai fondi documentari di appartenenza.

Il processo di consultazione è strutturato secondo diversi criteri organizzativi. È possibile effettuare ricerche sfruttando tutti gli elementi distintivi delle sintassi catalografiche oppure scegliere interrogazioni più mirate, grazie a filtri specifici che migliorano la precisione dei risultati. Inoltre, un'altra modalità di consultazione si basa sull'organizzazione dei dati mediante schemi di metadati descrittivi, i quali raggruppano le risorse secondo categorie documentali e tipologie, permettendo una navigazione più sistematica.

Un elemento essenziale della piattaforma è la conformità allo standard OAIS,³ che assicura una gestione avanzata dei cosiddetti *Dissemination Information Packages* (DIPs): pacchetti informativi strutturati che integrano dati e immagini digitali. I DIPs comprendono non solo le rappresentazioni digitali degli oggetti, come le immagini delle pagine di manoscritti e documenti, ma anche i metadati descrittivi e strutturali che arricchiscono e contestualizzano ciascuna risorsa. Questo approccio consente una consultazione immediata degli oggetti digitalizzati, anche in assenza di una catalogazione completa nell'OPAC, semplificando e ampliando l'accesso al patrimonio culturale.

Il sistema offre una ricerca avanzata che consente l'utilizzo di termini singoli o frasi e permette di identificare con precisione diversi componenti testuali, come nel caso dei manoscritti, distinguendo ad esempio tra *incipit* di testo, di prefazione o dedica, grazie agli attributi associati a ciascun elemento. La ricerca per scorrimento si basa su indici alfabetici arricchiti da riferimenti incrociati e annotazioni relative ai record di autorità, quali nomi, titoli e soggetti. Questa combinazione di strumenti rende più semplice il recupero delle informazioni e offre un'esperienza di navigazione più fluida. Ogni oggetto digitale è corredato da un *permalink* specifico, che garantisce un riferimento permanente e stabile all'oggetto stesso. Inoltre, il sistema assegna un indirizzo unico per ciascuna pagina visualizzata, offrendo un accesso preciso e diretto alle singole compo-

3 Cfr. <http://www.oais.info/> (data di consultazione 28/03/2025).

menti dell'oggetto. In parallelo, viene generato dinamicamente un URI (*Uniform Resource Identifier*) per il *manifest* IIIF⁴, che descrive in dettaglio la struttura, i metadati e i contenuti dell'oggetto digitale. Questo URI favorisce l'interoperabilità e l'integrazione delle risorse all'interno di ecosistemi digitali compatibili con il *framework* IIIF, migliorando significativamente sia l'accessibilità che la gestione scientifica del patrimonio digitale reperibile nel Web.

2. Adozione del IIIF in BAV

La Biblioteca Apostolica Vaticana ha adottato il protocollo IIIF oltre dieci anni fa, avviando un progetto pilota nel 2014 con il sostegno della Mellon Foundation. Questo progetto ha permesso di sperimentare, in particolare, la gestione del modello ontologico *Shared Canvas*,⁵ ponendo le basi per l'annotazione dei fogli di manoscritti digitalizzati. Alla base del modello *Shared Canvas* vi sono i principi dei Linked Open Data (LOD), che consentono di collegare e integrare dati provenienti da diverse fonti attraverso un'architettura distribuita conforme al web. Questo significa che ogni elemento di una risorsa digitale (come le pagine di un manoscritto, le miniature o le annotazioni) è rappresentato come un'entità separata, identificata da un URI univoco, che ne facilita l'accesso, l'organizzazione e la condivisione. *Shared Canvas* permette di 'ricomporre' digitalmente opere complesse, integrando immagini e metadati provenienti da archivi diversi, con l'obiettivo di rendere le risorse utilizzabili all'interno di sistemi compatibili con lo standard IIIF. Questo approccio facilita la collaborazione tra istituzioni culturali, archivi e biblioteche, contribuendo a migliorare l'accessibilità e l'interoperabilità delle risorse digitali.

La Biblioteca Apostolica Vaticana, in seguito al progetto pilota, ha adottato due strategie principali per la gestione del protocollo IIIF. La

4 Un *manifest* IIIF include informazioni essenziali per descrivere e organizzare un oggetto digitale. Tra queste ci sono i metadati descrittivi sull'oggetto, la struttura delle immagini distinte in sequenze e *canvas*. Queste ultime rappresentano elementi concettuali che fungono da 'spazio (tela) virtuale' su cui vengono posizionati e organizzati i contenuti digitali, come immagini, testi o annotazioni. Il *manifest* comprende inoltre gli URI per l'accesso diretto alle immagini, le annotazioni aggiuntive come note o trascrizioni, e i dettagli sulla presentazione, inclusi i parametri di visualizzazione come dimensioni e qualità delle immagini.

5 Cfr. <https://iiif.io/api/model/shared-canvas/1.0/> (data di consultazione 28/03/2025).

prima ha riguardato l'adeguamento dell'intera collezione digitale agli standard delle API del IIIF, tramite l'implementazione del DVL, garantendo un'organizzazione uniforme delle risorse. La seconda si è concentrata sulla produzione di annotazioni basate sui principi del IIIF, utilizzate come livello avanzato di analisi su una selezione di manoscritti. Queste annotazioni si conformano al modello *Shared Canvas*, ampliandone l'applicazione.

Il protocollo IIIF si struttura principalmente attorno a due API, essenziali per assicurare la conformità a tali specifiche delle piattaforme di biblioteca digitale, come il DVL: l'*Image API* e la *Presentation API*. L'*Image API* stabilisce una struttura standard per la gestione delle immagini digitali tramite URI, definendo parametri come area, dimensioni, rotazione e qualità. La *Presentation API*, invece, fornisce una rappresentazione strutturata degli oggetti digitali, arricchita da metadati descrittivi e da tutte le altre informazioni eventualmente disponibili. Attraverso un *manifest* IIIF, organizzato in sequenze e *canvas*, vengono descritti contenuti visivi e informazioni associate. Il livello *canvas* costituisce un componente fondamentale, poiché permette di distinguere tra diversi 'strati' informativi relativi al contesto originario. Questa organizzazione consente di integrare annotazioni e approfondimenti, offrendo una rappresentazione dettagliata e stratificata delle risorse digitali.

L'integrazione di questa tecnologia nelle piattaforme digitali della Biblioteca non solo ha promosso una gestione distribuita e accessibile delle risorse, ma ha anche permesso un significativo avanzamento nei processi di condivisione e collaborazione, facilitando lo studio dei manoscritti provenienti da collezioni diverse. Questo approccio contribuisce attivamente a valorizzare il patrimonio documentale, rendendolo fruibile a un pubblico più ampio, e allo stesso tempo promuove l'interoperabilità, facilitando lo scambio di dati e contenuti digitali su scala globale. Il visualizzatore integrato nel DVL, sviluppato in conformità con lo standard IIIF, permette di confrontare oggetti digitali provenienti da diversi archivi, purché compatibili col protocollo. Lo strumento offre la possibilità di interagire con immagini digitali in formato JPEG2000, consentendo operazioni come ingrandimento, scorrimento, rotazioni di immagini e la navigazione sequenziale delle pagine, simulando cioè l'esperienza dello sfogliare un libro fisico. Inoltre, supporta la comparazione di risorse di collezioni digitali appartenenti ad altre istituzioni che adottano lo standard IIIF, rendendo possibile un'analisi congiunta di materiali provenienti da fonti diverse.

3. Realizzazione dei *Thematic Pathways on the Web* della BAV

La seconda fase di implementazione del protocollo IIF presso la Biblioteca Apostolica Vaticana ha portato allo sviluppo di una piattaforma per mostre virtuali, denominata *Thematic Pathways on the Web*.⁶ Questo progetto rappresenta un'integrazione allo sviluppo del IIF, concepito per valorizzare il patrimonio culturale della Biblioteca e potenziare le possibilità di fruizione delle risorse digitalizzate, attraverso l'applicazione delle annotazioni ai contenuti dei manoscritti.

La piattaforma, sviluppata con il supporto della Mellon Foundation e in collaborazione con le Stanford University Libraries, si basa sull'uso delle annotazioni IIF per arricchire le immagini digitalizzate dei manoscritti. Queste annotazioni includono dati contestuali quali trascrizioni, note interpretative, approfondimenti storici e collegamenti a risorse esterne, offrendo così una comprensione più approfondita dei contenuti e consentendo un'esplorazione interattiva e multilivello.

All'interno della piattaforma, le mostre virtuali sono concepite come narrazioni tematiche, che guidano il visitatore attraverso percorsi di analisi e scoperta. Ogni mostra si compone di una selezione accurata di manoscritti e altri materiali digitalizzati, consultabili tramite il visualizzatore IIF integrato. Grazie alle annotazioni, è possibile mettere in risalto dettagli specifici, come decorazioni marginali, miniature o testi di *incipit* ed *explicit*, consentendo uno studio più mirato e approfondito.

Le annotazioni stesse costituiscono una base dati condivisibile, che può essere utilizzata per ulteriori ricerche o integrata in altre piattaforme compatibili con il protocollo IIF.

Tra i percorsi tematici sviluppati all'interno dei *Thematic Pathways on the Web*, quello dedicato alla Paleografia Latina⁷ offre un'analisi approfondita dell'evoluzione della scrittura latina, tracciando il suo sviluppo dall'antichità fino al Rinascimento. Il percorso utilizza annotazioni dettagliate per esaminare le caratteristiche stilistiche, i sistemi grafici e gli elementi decorativi presenti nei manoscritti. Questo consente di indagare aspetti specifici della scrittura nella sua evoluzione. In parallelo, il percorso dedicato alla Paleografia Greca⁸ si concentra sulla tradizione scrittoria greca, analizzan-

6 Cfr. <https://spotlight.vatlib.it> (data di consultazione 28/03/2025).

7 Cfr. <https://spotlight.vatlib.it/latin-paleography> (data di consultazione 28/03/2025).

8 Cfr. <https://spotlight.vatlib.it/greek-paleography> (data di consultazione 28/03/2025).

done l'origine, la diffusione e le trasformazioni nel corso dei secoli. In entrambi i due percorsi paleografici, le annotazioni includono traduzioni di testi rilevanti e commenti critici che guidano il lettore nella comprensione delle peculiarità linguistiche, semantiche e stilistiche dei manoscritti, con particolare attenzione ai contesti culturali in cui questi furono prodotti.

Un'altra rimarchevole 'mostra' è rappresentata dal percorso sui *Classici Latini*,⁹ che analizza approfonditamente la trasmissione e le trasformazioni che le opere latine hanno attraversato nel corso della storia. Questo percorso esamina come i manoscritti abbiano svolto un ruolo cruciale nel preservare e reinterpretare la produzione letteraria latina attraverso le epoche, ponendo enfasi sulla circolazione dei testi e sulle dinamiche culturali che ne hanno influenzato la ricezione.

Di grande interesse è anche il percorso dedicato ai *Palinsesti Vaticani*,¹⁰ che si avvale di tecnologie digitali all'avanguardia, come l'*i-maging* multispettrale, per il recupero di testi anticamente cancellati o riscritti. Il percorso non solo permette di accedere a contenuti altrimenti invisibili, ma apre nuove prospettive sulla comprensione delle identità culturali nascoste e dimenticate.

Infine, il percorso intitolato *Biblioteca di un 'Principe Umanista'*¹¹ esplora i manoscritti appartenuti a Federico da Montefeltro, offrendo uno sguardo sull'importanza della cultura umanistica nella formazione delle biblioteche rinascimentali. Attraverso annotazioni IIF e approfondimenti mirati, il percorso illustra il valore simbolico e intellettuale delle opere collezionate, mettendo in luce le interconnessioni tra cultura, potere e mecenatismo.

Tra gli altri percorsi tematici presenti, si distinguono inoltre le esposizioni virtuali relative al *Progetto Polonsky*¹² e la mostra *Viaggiare con Dante*¹³. Quest'ultima è stata progettata in occasione delle celebrazioni

9 Cfr. <https://spotlight.vatlib.it/latin-classics> (data di consultazione 28/03/2025).

10 Cfr. <https://spotlight.vatlib.it/palimpsests> (data di consultazione 28/03/2025).

11 Cfr. <https://spotlight.vatlib.it/humanist-library> (data di consultazione 28/03/2025).

12 Collaborazione tra la Biblioteca Apostolica Vaticana (BAV) e le Bodleian Libraries dell'Università di Oxford, sostenuta dalla Polonsky Foundation. Questo progetto, avviato tra il 2012 e il 2017, ha avuto l'obiettivo di digitalizzare e rendere accessibili online oltre 1,5 milioni di pagine di manoscritti e incunaboli provenienti dalle collezioni delle due istituzioni. Cfr. <https://spotlight.vatlib.it/selected-manuscripts-from-the-polonsky-foundation-digitization-project> (data di consultazione 28/03/2025).

13 Cfr. <https://spotlight.vatlib.it/dante> (data di consultazione 28/03/2025).

per il settimo centenario della morte di Dante Alighieri, configurandosi come un'indagine approfondita sulla vita e sulle opere del Poeta. La mostra virtuale si avvale di un'ampia selezione di manoscritti e materiali digitalizzati. Il percorso tematico segue una struttura narrativa che mette in luce gli aspetti fondamentali della produzione dantesca. I manoscritti medievali e gli incunaboli selezionati sono arricchiti da annotazioni che approfondiscono elementi testuali, iconografici e storici, consentendo una lettura stratificata e contestualizzata. Tra i punti centrali dell'esposizione, spicca lo studio dettagliato della *Divina Commedia*, con particolare attenzione ai manoscritti miniati, i cui decorazioni e illustrazioni arricchiscono l'interpretazione visiva dell'opera. Un altro nucleo rilevante riguarda l'analisi del contributo di Giovanni Boccaccio. Le annotazioni evidenziano il ruolo del Boccaccio nella diffusione dell'opera dantesca, approfondendo il suo metodo esegetico e il valore dei suoi interventi critici. Questo approccio consente di esplorare le interconnessioni tra i testi danteschi e le tradizioni culturali, nonché l'impatto che le opere hanno avuto nel contesto medievale e rinascimentale.

I manoscritti annotati nei *Thematic Pathways on the Web* esaminano sia le peculiarità testuali sia le caratteristiche iconografiche delle opere, creando un'esperienza fruibile su diversi livelli di approfondimento. Le annotazioni contribuiscono a contestualizzare i documenti all'interno delle loro dimensioni storiche, sociali e culturali, rendendo la piattaforma uno strumento significativo per l'analisi critica e lo studio interdisciplinare.

Questa iniziativa rappresenta un passo significativo nel processo di trasformazione digitale della Biblioteca Apostolica Vaticana, dimostrando come l'adozione di nuove tecnologie possa coniugare la conservazione del patrimonio storico con le esigenze della ricerca contemporanea.

Conclusioni

Le piattaforme digitali descritte non si limitano a essere strumenti tecnologici, ma rappresentano un ponte innovativo tra passato e futuro, capace di preservare e reinterpretare la ricchezza culturale delle collezioni vaticane in una dimensione globale e partecipativa.

Attraverso il patrimonio digitalizzato, arricchito da annotazioni dettagliate e metadati, viene aperta la strada a nuove prospettive di ricerca e analisi. La struttura interoperabile degli identificatori univoci (URI) e la ricchezza dei dati contenuti nei manoscritti, dai testi alle decorazioni

iconografiche, forniscono una base solida da cui partire per applicazioni innovative.

Infatti questi *dataset* rappresentano un legame concreto tra la tradizione e l'innovazione, offrendo nuove opportunità di indagine per le Digital Humanities, attraverso strumenti avanzati come il *machine learning*. Ad esempio, possono essere utilizzati per addestrare modelli di riconoscimento ottico dei caratteri (OCR) e di testi manoscritti (HTR) che decifrano scritture antiche o per sviluppare algoritmi di *computer vision* capaci di analizzare e catalogare elementi artistici e decorativi.

La complessa rete di connessioni tra dati visivi e descrittivi consente lo sviluppo di modelli avanzati in grado di correlare elementi testuali e iconografici alle rispettive funzioni culturali e storiche, offrendo un importante contributo alla ricerca in ambito umanistico. Inoltre, la scalabilità e l'uniformità del *framework* IIF facilitano l'integrazione e l'analisi di dati provenienti da collezioni eterogenee, incrementando il valore delle collezioni digitali e favorendo la collaborazione tra istituzioni culturali quali archivi, musei e biblioteche.

L'adozione di tali tecnologie consente di sviluppare strumenti per l'analisi automatizzata delle risorse culturali, favorendo l'accessibilità globale e promuovendo il sapere attraverso una prospettiva interdisciplinare. Tale combinazione, di archiviazione digitale e metodologie avanzate, trasforma il patrimonio culturale in un elemento fondamentale per lo studio e l'approfondimento accademico. Significa, di conseguenza, che le tecnologie digitali non sono solo strumenti, ma linguaggi che ridefiniscono il modo in cui interpretiamo e condividiamo il sapere. Questo concetto sottolinea un cambiamento significativo: il passaggio da una concezione del digitale come semplice strumento tecnico a una visione in cui esso diventa un elemento culturale e semantico. Questo approccio mette in evidenza il potenziale del digitale nel trasformare le scienze umane e nell'aprire nuove prospettive di ricerca interdisciplinare, ridefinendo il modo in cui interpretiamo e condividiamo la conoscenza.

La prospettiva qui delineata invita a considerare il digitale come una componente viva dell'evoluzione culturale contemporanea, un agente che non solo conserva il passato, ma lo proietta nel futuro, arricchendo le possibilità di interazione e reinterpretazione. Attraverso questa sintesi tra tradizione e innovazione, si creano nuovi paradigmi per il sapere, ponendo le basi per un dialogo continuo tra cultura e tecnologia.

Bibliografia

- Benanti 2020 = Paolo Benanti, *Digital Age. Teoria del cambio d'epoca. Persona, famiglia e società*, Milano, San Paolo, 2020.
- Manoni 2019 = Paola Manoni, *The Mellon Project at the Vatican Library: WEB thematic pathways of Medieval manuscripts from the Vatican collections using International Image Interoperability Framework*, in *Miscellanea Bibliothecae apostolicae Vaticanae XXV*, Città del Vaticano, Biblioteca Apostolica Vaticana, 2019, pp. 256-281.
- Manoni 2020a = Paola Manoni, *L'adozione del IIF nell'ecosistema digitale della Biblioteca Apostolica Vaticana*, «Digitalia», 15/2 (2022), pp. 96-105. ([DOI:10.36181/digitalia-00017](https://doi.org/10.36181/digitalia-00017)).
- Manoni 2020b = Paola Manoni, *Thematic Pathways on the Web: IIF Annotations of Manuscripts from the Vatican Collections: il "Progetto Mellon" della Biblioteca Vaticana*, «Rivista di storia della miniatura», 24 (2020), pp. 211-222.
- Manoni 2022 = Paola Manoni, *Discoverability in the IIF Digital Ecosystem*, «JLIS», 1 (2022), pp. 312-320 (DOI: 10.4403/jlis.it-12770).
- Manoni 2024 = Paola Manoni, *The DigiVatLib: Descriptive and Structural Metadata, IIF Interoperability Protocol in The process for the digitization of manuscripts in the Vatican Library*, Città del Vaticano, Biblioteca Apostolica Vaticana, 2024, pp. 126-140.
- Ponzi 2023 = Eva Ponzi, *Dante cibernetico. L'esperienza vaticana in una mostra online*, «Dante e l'arte», 10 (2023), pp. 125-136.



10.69118/2025_L8

Transforming manuscripts studies through technology: methods, challenges and future directions

Arianna Traviglia

Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia
arianna.traviglia@iit.it

Trasformare lo studio sui manoscritti attraverso la tecnologia: metodi, sfide e direzioni future

Abstract (ITA)

Il contributo esplora il valore dello studio dei manoscritti non solo come testi, ma come oggetti materiali. Analizzarne la dimensione ‘fisica’ – materiali, struttura, segni d’uso – consente infatti di ricostruire la storia della loro produzione e fruizione nel tempo. Le più recenti innovazioni tecnologiche, dall’imaging avanzato alla digitalizzazione, stanno trasformando profondamente il campo della paleografia e della filologia. La combinazione di tecniche di analisi materiale e strumenti digitali, come i sistemi di riconoscimento e trascrizione automatiche della scrittura, apre nuove prospettive nella comprensione dei manoscritti, ampliando l’accesso e accelerando i processi di studio.

Abstract (ENG)

This paper examines the study of manuscripts beyond their textual content, highlighting the importance of their material dimension. Investigating physical features – materials, construction, usage traces – offers crucial insights into their historical and cultural contexts. Recent technological advances, from imaging tools to digital platforms, are reshaping manuscript research. The integration of material analysis and digital methods, including handwriting recognition and transcription systems, is revolutionising access and interpretation, allowing scholars to engage with manuscripts in deeper, more meaningful ways.

Parole chiave / Keywords

Materiality of manuscripts, Digital transformation, Handwriting recognition / Materialità dei manoscritti, Trasformazione digitale, Riconoscimento automatico della scrittura

When reflecting on a manuscript, the focus of a humanities scholar's thinking is typically drawn first to its textual content, as if the physical aspect of the manuscript – as an object – were somehow secondary, reduced to a mere writing support. However, approaching the manuscript as a physical artifact reveals a richer field of inquiry, one that encourages close examination of its material form, construction, and history. This perspective offers critical insights into the circumstances of its production, the history of its preservation, and the ways in which it was used and understood over time.

Every manuscript (whether a scroll, codex, single sheet, booklet, or fragment) is crafted from specific materials such as paper or parchment and bears physical features that tell stories of writing, cultural and social context, and artistic practices. Its dimensions, the quality of the material, the type of ink used, any decorative elements, signs of handling and preservation, bindings, and even folds or abrasions can reveal valuable information about the place and time of its production, the craftsmanship of the scribe, and the manuscript's interaction with its historical environment. It is precisely the manuscript's 'material dimension' that enables it to transcend its function as a mere 'container' of text. This aspect deepens understanding of the document and its historical trajectory, offering an additional layer of meaning and interpretation.

Modern technology has profoundly reshaped the study and analysis of manuscripts, offering tools that extend well beyond traditional textual reading or examination. Advances in imaging techniques, such as infrared spectroscopy, multi and hyperspectral analysis, and even X-ray tomography now allow for detailed examination of a manuscript's material features. These instruments can reveal hidden layers of ink, corrections, additions, or earlier writings that would otherwise be difficult or impossible to detect. In doing so, they enable scholars to explore the manuscript's physical history in unprecedented depth. At the same time, digital technologies have greatly expanded access to manuscript collections. Thanks to high-resolution photography, occasionally coupled with other sensors, it is now possible to create detailed digital copies of manuscripts, making them globally accessible to scholars and researchers. This development addresses one of the major challenges in manuscript studies, namely the access to documents, especially rare or fragile items that cannot be handled frequently. The process of converting 'physical' information recorded on paper into digital format offers a valuable solution to this problem, preserving the physical integrity of manuscripts

while ensuring secure and immediate access to their contents. Digital technologies further enhance manuscript usability by enabling precise searches, facilitating cross-text comparisons, and automating transcription through optical character recognition (OCR) algorithms. In this way, the combination of technological innovations in material sciences and digital techniques is opening new frontiers in manuscript research, leading to a deeper understanding not only of the content, but also of the material and historical context surrounding these documents.

The process of digital transformation encompasses a range of activities. In the case of manuscripts, it goes beyond the simple conversion of physical documents into digital formats through the production of high-resolution images, which constitutes their 'digitisation': while this step is crucial for accurate preservation and documentation, it represents only part of a much broader spectrum. Digital transformation also includes the use of digital data – thus evolving into 'digitalisation' – extracted from manuscripts to support analysis, interpretation, and transcription. Among these, transcription has seen particularly significant advancements lately, with the automatic recognition of handwriting emerging as one of the most promising developments for the discipline. Systems based on transcription algorithms are becoming increasingly capable of reading and transcribing handwritten texts with remarkable accuracy. Some deep learning models have already demonstrated the ability to drastically reduce transcription time and to improve the legibility of texts that would otherwise be difficult to transcribe. Naturally, the development of these systems does not and cannot replace the work of scholars (palaeographers, philologists, historians), but must be steered by them in order for computer scientists to design high-quality models. These tools should always be understood as support for scholars, enabling them to focus on analysis and interpretation rather than on the task of transcribing the text. This acceleration facilitates broader access to knowledge, making a significantly larger number of manuscripts and documents available, many of which would otherwise remain confined to archives.

OCR transcription technology is now widely known and used in everyday contexts, in a variety of applications. However, applying Artificial Intelligence techniques to historical manuscript transcription, while similar in principle, is generally more complex than working with modern printed or handwritten documents, due to a number of factors that make historical text recognition significantly more challenging. First and most obviously, while transcription can still be relatively straightforward for

standardised scripts, such as Gothic hands, it becomes far more difficult with cursive scripts and personal handwriting of the following centuries. Unlike the rigid and uniform letterforms used before, personal scripts often lack consistent structure. This makes it extremely challenging to develop AI models capable of generalising across a single letter that may be written in many different ways by different individuals, each with their own style, slant, and habits. The stylistic variation inherent in personal writing means that each scribe may alter letterforms depending on their handwriting, writing speed, or even the condition of the materials and tools used. As a result, AI models must be trained to identify and compare a wide range of variants, requiring a high degree of adaptability. In addition, layout variation presents another layer of complexity. Manuscripts often feature diverse writing patterns and page layouts, even within a single document, whereas printed books tend to follow more standardised formatting conventions. AI-based transcription systems must therefore be trained not only to recognise handwriting, but also at the same time to interpret and adapt to varying layouts and structural features of the manuscript page. Additional challenges stem from the visual characteristics of manuscripts, particularly the high presence of what is commonly referred to as ‘noise.’ In this context, ‘noise’ refers to any element in the document image that hinders the character recognition process. In manuscripts, noise can result from various factors, including age-related deterioration, suboptimal storage conditions, or environmental exposure. Another significant obstacle involves physical damage to the manuscript, such as holes, stains, and other imperfections that compromise legibility and further complicate character recognition.

It is well known that Machine Learning models require data, that is, examples from which to learn: in the case of automatic transcription of historical documents, these models rely on training datasets composed of manuscript images paired with accurate transcriptions produced by subject matter experts. Thus, the preliminary creation of high-quality datasets is a necessary step in the development of effective transcription models. The collaboration of paleographers, philologists and historians is crucial in this process, as they are responsible for producing the reliable transcriptions that serve as ground truth data for training. However, experts with such specialised skills are increasingly scarce, and even fewer are willing or able to dedicate time to the creation of datasets for AI training purposes. As a result, in the absence of high-quality annotated data, computer scientists are often left to produce their own

training sets, frequently relying on lower-quality transcriptions or limited sample sizes. This, in turn, negatively affects the performance and accuracy of the Machine Learning models. In addition, it must be emphasised once again that Machine Learning models are data-driven, meaning their performance depends directly on the quantity and quality of the data they are given: the more data available during training, the better the model can learn and make accurate predictions. Conversely, when only limited data are available, the model's performance is similarly constrained. In other words, with insufficient training data, the model is less likely to produce reliable or effective results, outcomes that are unlikely to meet the standards or needs of scholars relying on these systems.

The creation of large, high-quality datasets that align with the expectations and requirements of domain experts is therefore a *conditio sine qua non* for developing Machine Learning models that are both reliable and effective, and that respond to scholarly needs. With this in mind, the Center for Cultural Heritage Technology (CCHT), the unit of the Italian Institute of Technology dedicated to technological applications in the field of cultural heritage, and directed by the present author, has, in recent years, produced datasets aligned with these principles. This work has been carried out in close collaboration with paleographers and has drawn on the *Mariegole*, a rich documentary tradition preserved in the historical archives of Venice (Italy). Building on this approach, the CCHT has adopted a streamlined method in which each training sample consists of a single line of text paired with its corresponding transcription. These line-level segments form the core of the datasets used to train the models. While alternative approaches, such as character-level datasets, are possible, the line-based format offers a practical balance between transcription accuracy and the time required for dataset preparation. Creating character-level datasets would demand significantly more annotation effort, making them impractical within the scope of the research.

One of the main features of the *Mariegole* manuscripts used in model development is the frequent presence of abbreviations. While in manual transcription performed by paleographers, the expansion of abbreviations is a standard practice, current automatic transcription models often do not follow this convention, instead adopting a diplomatic approach that preserves abbreviations as they appear in the source. Only a few studies have focused on the automatic transcription of texts containing historical abbreviations, and CCHT was among the first research groups

worldwide to address this specific challenge. Notably, it was the first to propose models capable of performing directly the expansion without the need for a preliminary, separate diplomatic transcription step. This innovation significantly reduces the labeling effort, which would otherwise be duplicated for both diplomatic and normalised transcriptions.

This effort to enable and enhance the performance of automated normalised transcription is due to several important advantages it offers. First, it ensures greater legibility and accessibility, making the text understandable to a broader audience, including specialists who may not be familiar with regional or local script conventions. Additionally, normalisation helps reduce ambiguities inherent in original manuscripts, clarifying the intended meaning and facilitating more consistent analysis across different texts. In particular, abbreviations often carry significant linguistic and contextual information that, if left unresolved, can obscure meaning and hinder interpretation. In historical documents like the *Mariegole*, where abbreviations were frequently used to save space or reflect contemporary writing conventions, the expansion process helps clarify the original intent of the scribe and facilitates a more precise understanding of the text. Above all, adopting a technological perspective and considering future possibilities with automated methods, fully expanded transcriptions enhance interoperability with digital tools, thereby improving searchability and enabling more effective comparative analysis across different manuscripts.

The experience gained through this research activity makes it evident that achieving meaningful progress in this field will require ever-stronger cooperation between paleographers, philologists, historians, and computer scientists, the only ones that can fully grasp the significance of key features such as specific scribal marks, textual organisation, graphic conventions, and stylistic choices that characterise each manuscript. These elements go well beyond the mere recognition of words and are essential to developing high-quality transcription models. The real challenge lies in transferring this expertise into Machine Learning models, so that the system not only “reads” the text, but begins to interpret it within its historical, cultural, and formal context, ultimately ensuring a more faithful transcription.

Digital technologies extend beyond the mere conversion of analogue materials into digital formats or the processing of extracted data. They also encompass the potential to derive deeper insights from manuscripts through physico-chemical analysis, which produce data and images suit-

able for further digital interpretation and elaboration. These imaging techniques allow for the examination of the material composition of manuscripts, such as ink and paper, and can reveal details invisible to the naked eye, including portions of text that are hidden or damaged – and thus unreadable – providing a deeper understanding of the manuscript's textual content, its production processes, and its historical interactions.

Among the most rapidly advancing techniques is hyperspectral imaging, a completely non-invasive method that captures both spectral and spatial information simultaneously, allowing for the acquisition of hundreds of images across the electromagnetic spectrum in a single session, collecting data at multiple wavelengths. The resulting product is a 'datacube', an image with two spatial and one spectral dimension, where each pixel contains a spectrum that provides detailed information on the constituent materials. Hyperspectral imaging is primarily used to identify pigments and dyes in manuscripts, thanks to its ability to detect specific spectral bands that correspond to the chemical composition of materials. By applying specialised algorithms to the data extracted from the instruments, it is possible to generate distribution maps of the identified materials. These maps illustrate, for example, how pigments are used across different areas of the manuscript, allowing for their precise localisation. Another example of the usefulness of hyperspectral imaging lies in cases where two inks – one used for the main text and another for a correction or later addition – appear visually identical to the naked eye. Hyperspectral imaging in the short-wave infrared (SWIR) range can clearly reveal differences between them, enabling the identification of distinct chemical formulations. This is particularly valuable for studying changes made to a document over time, including corrections, restorations, or later additions. Such analysis allows to reconstruct a detailed history of the manuscript, shedding light on how it was written, altered, and preserved throughout its existence.

Another emerging technique in manuscripts analysis is Terahertz imaging, which operates in the electromagnetic spectrum between far-infrared and microwaves. Due to its non-invasive nature and ability to penetrate opaque materials, Terahertz imaging is proving highly valuable for the study and preservation of handwritten texts: it detects hidden or overwritten texts, revealing annotations, corrections, or earlier versions concealed beneath layers of ink or miniature. The technique also allows for the identification of multiple writing layers and material stratigraphy, such as ink, paper, and restoration layers, without damaging

the original. Furthermore, Terahertz radiation interacts differently with various substances, aiding in the identification of materials and the detection of physical damage or prior restoration efforts invisible to the naked eye. Since Terahertz imaging captures both spatial and spectral data, it can be considered a form of hyperspectral imaging, though operating in a different frequency range. Unlike standard spectral techniques, however, Terahertz data require more complex processing and advanced algorithms to accurately interpret and reconstruct usable images. A key challenge lies in the image quality: as frequency increases, images tend to become blurrier and more affected by noise. These two distortions are inversely related and difficult to correct simultaneously, often resulting in imperfect, hard-to-interpret data. Addressing this issue requires the development of advanced algorithms capable of mitigating both distortions concurrently, thereby ensuring consistent and reliable results across a range of wavelengths. Artificial intelligence can assist – once more – in this process through models specifically designed for image enhancement and spectral reconstruction, improving data quality and facilitating more accurate analysis of manuscript features.

This brief overview of emerging technologies has demonstrated how the integration of digital tools, computational methods, and advanced imaging techniques is profoundly transforming the study and understanding of historical manuscripts. Far from serving merely as tools for visual reproduction or preservation, these technologies are increasingly embedded in the processes of scholarly interpretation, shaping how manuscripts are studied, understood, and contextualised. In particular, the growing ubiquitous application of artificial intelligence across various domains is opening up unprecedented opportunities for the integrated analysis of both the textual content and material characteristics of manuscripts. While many of these methods are still under development and require further refinement, especially in terms of standardisation, interoperability, and usability, their impact on historical, philological, palaeographic and codicological research is already evident. The key challenge going forward lies in strengthening interdisciplinary collaboration to ensure that technological innovation aligns with the needs and values of humanistic inquiry. Only by doing so can the content of manuscripts be not merely accessed, but interpreted with greater depth and precision, revealing layers that extend beyond the textual surface to include their historical, cultural, and material dimensions.



10.69118/2025_L9

Limina from the virtual unwrapping of the Herculaneum Papyri*

Gianluca Del Mastro

Università degli Studi della Campania “Luigi Vanvitelli”

gianluca.delmastro@unicampania.it

Limina dallo srotolamento virtuale dei papiri di Ercolano

Abstract (ITA)

La nuova tecnologia che consente attraverso l’IA di leggere le immagini di papiri ercolanesi prodotte da una tomografia (Micro-CT) rivela non solo nuovi testi, ma anche una fitta serie di elementi paratestuali. I segni che si leggono nel primo papiro svolto con questo procedimento (*PHerc.Paris.4*) sono quelli che conosciamo dagli altri papiri ercolanesi e greco-egizi. L’unica particolarità si osserva sotto l’ultima linea, dove, in luogo della più frequente coronide, si trova una *paragraphos*.

Abstract (ENG)

The new technology that employs AI to interpret the images of Herculaneum papyri produced through – computed tomography (Micro-CT) reveals not only previously unknown texts but also a rich array of paratextual elements. The signs visible in the first papyrus examined using this method (*PHerc.Paris. 4*) correspond to those already known from other Herculaneum and Greco-Egyptian papyri. The only notable peculiarity appears beneath the final line, where a *paragraphos* is present instead of the more common *coronis*.

Parole chiave / Keywords

Herculaneum Papyri, carbonized rolls, Epicurean philosophy, titles, punctuation marks / Papiri di Ercolano, rotoli carbonizzati, filosofia epicurea, titoli, segni di punteggiatura

* This article is produced by the PRIN PNRR 2022 project, *Digital Papyrology. New Approaches to Preservation, Edition and Dissemination of Papyrus Collections in Southern Italy* (P2022J8CAJ, PI, Lucio Del Corso; Head of the Local Unit of the Vanvitelli University, Gianluca Del Mastro).

In recent years, a new technology, based on X-ray tomography and machine learning, has made it possible to virtually unroll ancient carbonized scrolls that could not be opened mechanically.

After an initial successful attempt on a carbonized scroll that preserves traces of the Book of *Leviticus*, found in Israel,¹ and then on some Herculaneum papyri,² the team led by Brent Seales at University of Kentucky within the framework of the *Vesuvius Challenge* managed to virtually open the final 16 columns of a Herculaneum papyrus (*PHerc.Paris. 4*). The scroll had been donated by King Ferdinand IV to Napoleon Bonaparte in 1802 and is now preserved at the *Institut de France* in Paris.³

However, the ‘simple’ scan – obtained at the Diamond Light Source synchrotron in Didcot, Oxfordshire – and the segmentation, flattening, and texturing phases carried out by the computer scientists, were not sufficient to make the ink on the carbonized Herculaneum papyrus legible. In fact, the Herculaneum ink contains a small amount of metal that can only be partially seen from the scan.⁴ To address this, Nat Friedman, Daniel Gross, and Brent Seales launched the *Vesuvius Challenge* in 2023,⁵ an international competition aimed at making the digitized papyri available to scholars worldwide and enabling the detection of ink traces.

AI was used by Luke Farritor, Youssef Nader and Julian Schilliger - the three winners of the first phase of the Challenge - to ‘teach’ the machine to detect ink signal (not to read letters!) based initially on pieces of papyri previously opened with mechanical means, and then on ink signal partially visible in the scan data.⁶

Only at this point were the papyrologists able to identify the letters and read the text.

The first papyrus opened using this system is an Epicurean text dealing with the theory of perceptions.⁷ In particular, the final part of the scroll

1 Seales *et al.* 2016.

2 Seales-Delattre 2013. At the same time, other groups of scientists were conducting research using similar technologies (cfr. Del Mastro-Delattre-Mocella 2015).

3 Delattre 2004.

4 On the metal in the ink of the Herculaneum Papyri, cfr. Bonnerot *et al.* 2020.

5 www.scrollprize.org.

6 “The identification of letterforms in the output images is, therefore, the result of human interpretation and not of algorithmic hallucination” Nicolardi *et al.* 2024, p. 5.

7 Τὸ σπάνιον (the scarcity) and τὸ ἀφιλές (the abundance).

discusses taste and perception of food through the contrast between its scarcity and abundance.⁸

One initial ‘paratextual’ issue that has not yet been resolved by scholars (both computer scientists and papyrologists) is that the final title has not been identified.⁹ In such titles, as is common in all papyri, including those of Herculaneum, we could find the name of the author, the title of the work, possibly the book number and, if we are fortunate, other information, such as the number of *stichoi*, or, more rarely, the number of columns (*selides*) and sheets (*kollemata*) making up the roll.¹⁰

In the Herculaneum papyri, the title is usually located at the end of the roll, either beneath the last column or to the right of it: as I have noted,¹¹ among the papyri already opened using traditional methods, the title is found beneath the final column in only eight cases (8 out of approximately 130).¹² Today it looks unlikely that the title of the *PHerc. Paris. 4* was not placed beneath the final column: the “segmenters”,¹³ in fact, are still searching, but it is clear that the space under the last column appears to be blank. The innermost portion of the roll remains to be analyzed. However, this section is particularly complex and detecting traces of letters there is difficult.

In an effort to identify the title of the papyrus, another possible route would be to search for an initial title. In fact, fewer than ten Herculaneum papyri preserve such initial titles. The actual number must have been much higher, but as I have pointed out elsewhere, in many cases the initial portion of the roll – the outermost section – was the most likely to be lost during the eruption or during early mechanical unwrapping attempts in previous centuries.¹⁴

8 The edition of this text has been published by Nicolardi *et al.* 2024b.

9 Cfr. Nicolardi *et al.* 2024b, p. 12.

10 In the Herculaneum papyri all three pieces of information were found in *PHerc.* 1385 (Epicurus, *On nature*) and *PHerc.* 1414 (Philodemus, *On gratitude*). See Del Mastro 2014, p. 23.

11 Del Mastro 2014, p. 7.

12 Del Mastro 2014, p. 11.

13 About the segmentation after the tomography (Micro-CT), cfr. Parker *et al.* 2019 and Nicolardi *et al.* 2024, p. 2 f.

14 Del Mastro 2014, p. 10.

It is quite likely that a ‘reverse’ reading (i.e., starting from the final columns) of *PHerc.Paris. 4* will eventually lead to the identification of the initial title, placed – as in other cases – on the *recto* of the roll.

The AI underlying the reading of papyri virtually unrolled through X-ray tomography, as mentioned earlier, does not read graphic signs in the traditional sense. Instead, it detects the physical characteristics of the ink as it spreads across the texture of the papyrus fibers. This provides clear evidence that the machine is not interpreting what it sees.

For this reason, even the non-alphabetic signs detected on the papyrus during the various phases of the *Challenge* are ‘real’ – that is, they do not result from any interpretive process carried out by the AI. In all cases, these signs correspond to those already known from the Herculaneum papyri, the Greco-Egyptian papyri, and those described in ancient treatises.¹⁵

In *PHerc.Paris.4*, various punctuation marks are present, including the *paraglyphos* – a stroke-shaped sign equivalent to our modern full stop – found at columns -15,4; -14,2; -7,13 (cfr. Fig. 1); -6,14; -1,7¹⁶ and the *diple*, an angular sign indicating a stronger pause, located at -15,2 (cfr. Fig. 2). Additionally, an *asteriskos* appears at the end of a line (col. -12,7; cfr. Fig. 3), likely serving as a filler.¹⁷ This symbol resembles an X, slightly larger than the surrounding letters, and is encircled by four small dots that are barely visible in the virtual reconstruction.

At the end of the text, beneath the final column, Herculaneum papyri almost always feature a *coronis*¹⁸ – a sign that, in many cases, consists of a central element (often spiral-shaped) accompanied by a series of curved and straight strokes above and below the spiral.¹⁹

In many instances, confirming the presence of this sign is difficult, as the final column – or the precise spot where the *coronis* would typ-

15 As regards the Herculaneum papyri, cfr. Del Mastro 2017.

16 See *infra* and Fig. 4.

17 This sign is quite used in the Herculaneum papyri. Cfr. e.g. *PHerc. 1050* (Philodemus, *On Death* IV, col. 70,19 Delattre 2022).

18 In the Herculaneum papyri (as in the Greco-Egyptian papyri), the *coronis* is also found within the text to delimit specific sections or to clearly separate portions of text.

19 The name derives from the crow (in Greek κορώνη) and, in some cases, the shape of the sign is zoomorphic (cfr. Stephen 1959). In *PHerc. 208* (Colotes, *Against Plato's Lysis*), an intermediate *coronis* is clearly bird-shaped (see Fig. 4).

ically appear – has been lost in a lacuna (cfr. *PHerc.* 182, Philodemus, *On Anger*; *PHerc.* 1497, Philodemus, *On Music* 4).²⁰ In the case of *PHerc.Paris.* 4, only a *paragraphos* is present,²¹ a simple horizontal stroke placed beneath the first letter of the line to indicate the conclusion of the treatise (Fig. 5).²² Francesca Schironi has pointed out that, among the papyri that preserve hexametric texts «there are only two that present a *paragraphos* without a *coronis*, and both are Ptolemaic or very early Roman papyri».²³



Fig. 1. The *paragraphos* in the *PHerc.Paris.*4.

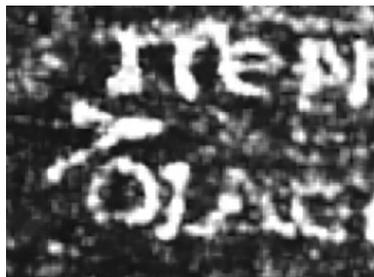


Fig. 2. The *diple* in the *PHerc.Paris.*4.

20 The same problem, for the Greco-Egyptian papyri, has been highlighted by Schironi 2010, p. 8 f.

21 See Schironi 2010, p. 71, for examples from Greco-Egyptian papyri.

22 Some scholars (Tanzi-Mira 1920) believed that the *coronis* was only an ornate form of *paragraphos*. In reality, the existence of a specific name for this sign in antiquity testifies to the will of the scribes to clearly distinguish the two signs. An example is given by *PLond.Lit.* 11 (LDAB 1957, 1st century AD) which preserves fragments of the books 3-4 of the *Iliad*, in which, parallel to the last column, one can read an epigram which begins with Ἐγ[ὼ κο]ρωνίς εἰμι | γραμμάτων φύλαξ ... (“I am the coronis, guardian of the letters...”). See also Meleager, *AP* 12.257, which begins with: Ἄ πόματον καμπτήρα καταγγέλλουσα κορωνίς, ἐρκοῦρος γραπταῖς πιστοτάτα τελίειν ... (“I, the coronis, which announces the final lap, the most trustworthy guardian of the end of the written sheets...”). These quotations (on which see Stephen 1959, p. 12-14) represent the desire to clearly name the *coronis* and to distinguish it from other signs.

23 Schironi 2010, p. 78. In our case, the *PHerc.Paris.* 4 on a paleographic basis (as far as we can see from the images produced by Micro-CT) can be dated between the 1st century BC and the 1st century AD (see Nicolardi *et al.* 2024b, p. 10).



Fig. 3. The space-filler in the *PHerc. Paris. 4*.



Fig. 4. The *coronis* in the *PHerc. 208*

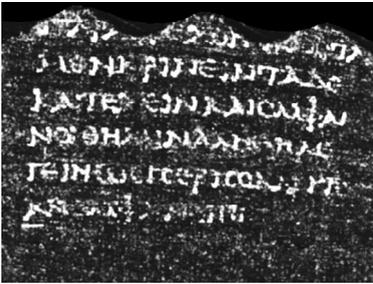


Fig. 5. The *paragraphos* at the end of *PHerc. Paris. 4*.

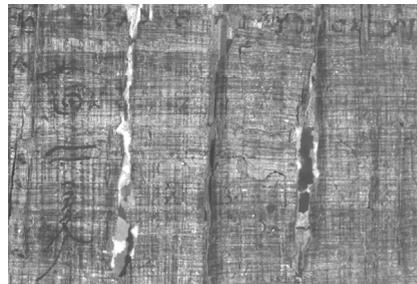


Fig. 6. The *coronis* at the end of *PHerc. 1426*.

At the end of an entire book, beneath the first letter of the final line, the *coronis* often takes on an elaborate and elegant form. Its complexity can vary greatly – from simpler designs, such as in *PHerc. 832/1015* (Philodemus, *On Rhetoric* 8), where an S-shaped stroke drawn backward ends in a loop and is decorated below with a few additional lines, to highly intricate compositions such as in *PHerc. 1426* (Philodemus, *On Rhetoric* 3; Cfr. Fig. 6). In this latter case, the upper part consists of a series of intersecting rays (corresponding to the final line), a spiral-shaped central body, and a lower section with wavy strokes of varying lengths.

In many cases, the *paragraphos* is incorporated into the *coronis*, marking the final line of the treatise (*PHerc. 1427*, Philodemus, *On Rhetoric* 1). However, there are rarer instances – such as in *PHerc. Paris. 4* –

where only a *paragraphos* appears, without a *coronis*. This also occurs in *PHerc.* 1050 (Philodemus, *On Death* 4), where, as Daniel Delattre has recently noted: «à la fin de la dernière colonne (118) du *volumen*, devant la l. 25 et dernière, on notera que ne figure qu’une *paragraphos* renforcée en L, au lieu de la *coronis* qu’on attendrait en cet endroit». ²⁴ The same scholar, conversely, has pointed out the use of *coronides* «à l’intérieur du livre... afin de souligner un passage particulier de l’écrit». Even in the Roman era (the *De morte* papyrus dates to the second half of the 1st century BC),²⁵ a *paragraphos* – in this case, the “forked” version – could be used to mark the end of an entire book.

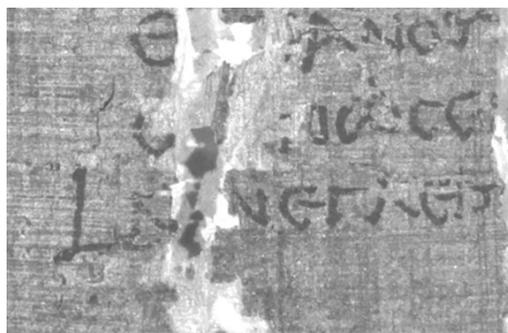


Fig. 7. The forked *paragraphos* at the end of *PHerc.* 1050.

As has been demonstrated on several occasions,²⁶ these are discretionary choices made by the scribes, only partially influenced by what they found in the exemplar (*antigraphon*). The Herculaneum papyri confirm that even when we possess two copies of the same work (*e.g.*, a draft and a final version), the system of punctuation marks may vary.

In the case of *PHerc.Paris.* 4, we observe a more ‘basic’ editorial choice: a simple *paragraphos* marking the end of the text. We now await the reading of the book’s title to determine whether it follows this same minimalist trend or features some form of decorative flourish.

24 Delattre 2022, p. LVI.

25 Cavallo 1983, p. 52.

26 Cfr. Turner 1987, p. 9 f.

Bibliography

- Bonnerot *et al.* 2020 = Olivier Bonnerot, Gianluca Del Mastro, Jürgen Hammerstaedt, Vito Mocella, Ira Rabin, *XRF Ink Analysis of some Herculaneum Papyri*, «Zeitschrift für Papyrologie und Epigraphik», 216 (2020), pp. 50-52.
- Cavallo 1983 = Guglielmo Cavallo, *Libri Scritture Scribi a Ercolano*, «Cronache Ercolanesi» Suppl. 1, Naples, Macchiaroli, 1983.
- Delattre 2004 = Daniel Delattre, *Le retour du PHerc. Paris 2 à l'Institut de France: un rouleau épicurien inédit en 279 fragments*, «Comptes Rendus des séances de l'Académie des Inscriptions et Belles-Lettres», 3 (2004), pp. 1351-1391.
- Delattre 2022 = Daniel Delattre, *Philodème de Gadara, Sur la mort, Livre IV*, Paris, Le Belles Lettres, 2022.
- Del Mastro 2014 = Gianluca Del Mastro, *Titoli e annotazioni bibliologiche nei papiri greci di Ercolano*, «Cronache Ercolanesi», suppl. 5, Napoli, CISPE, 2014.
- Del Mastro 2017 = Gianluca Del Mastro, *Les σημεία dans les papyrus grecs d'Herculaneum*, in *Signes dans les textes, textes sur les signes*, Actes du colloque international Liège 2013, ed. Gabriel Nocchi Macedo e Maria Chiara Scappaticcio, Liège, Presses Universitaires, 2017, pp. 77-96.
- Del Mastro-Delattre-Mocella 2015 = Gianluca Del Mastro, Daniel Delattre, Vito Mocella, *Una nuova tecnologia per la lettura non invasiva dei papiri ercolanesi*, «Cronache Ercolanesi», 45 (2015), pp. 227-230.
- Nicolardi *et al.* 2024 = Federica Nicolardi, Stephen Parsons, Daniel Delattre, Gianluca Del Mastro, Robert L. Fowler, Richard Janko, Tobias Reinhardt, Clifford Seth Parker, Christy Chapman, William Brent Seales, *Revealing ink from a still-rolled Herculaneum papyrus scroll (PHerc.Paris. 4)*, «Zeitschrift für Papyrologie und Epigraphik», 229 (2024), pp. 1-13.
- Nicolardi *et al.* 2024b = Federica Nicolardi, Daniel Delattre, Gianluca Del Mastro, Robert L. Fowler, Richard Janko, *The final columns of PHerc. Paris. 4 revealed through virtual unwrapping*, «Cronache Ercolanesi», 54 (2024), pp. 9-27.
- Parker *et al.* 2019 = Clifford Seth Parker, Stephen Parsons, Jack Bandy, Christy Chapman, Frederik Coppens, William Brent Seales, *From Invisibility to Readability: Recovering the Ink of Herculaneum*, PLoS

ONE 14.5, May 2019, pp. 1-17 (<https://doi.org/10.1371/journal.pone.0215775>).

Schironi 2010 = Francesca Schironi, TO MEΓA BIBAION: *Book-Ends, End-Titles and Coronides in Papyri with Hexametric Poetry*, «The American Society of Papyrologists», 48, Durham, The American Society of Papyrologists, 2010.

Seales-Delattre 2013 = William Brent Seales-Daniel Delattre, *Virtual Unrolling of Carbonized Herculaneum Scrolls: Research Status (2007-2012)*, «Cronache Ercolanesi», 43 (2013), pp. 191-208.

Seales *et al.* 2016 = William Brent Seales- Clifford Seth Parker-Michael Segal-Emanuel Tov-Pnina Shor-Yosef Porath, *From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi*, «Science Advances», 2/9 (2016) (<https://doi.org/10.1126/sciadv.1601247>).

Stephen 1959 = Gwendolen M. Stephen, *The Coronis*, «Scriptorium» 13 (1959), pp. 3-14.

Tanzi-Mira 1920 = Giuseppina Tanzi-Mira, *Paragraphoi ornate in papyri letterari greco-egizi*, «Aegyptus», 1 (1920), pp. 224-227.

Turner 1987 = Eric G. Turner, *Greek Manuscripts of the Ancient World*, Oxford. Rev. ed. P.J. Parsons, «BICS», suppl. 46, London, 1987.



10.69118/2025_L10

“Voce dei Margini”: le postille di Giorgio Bassani in un’edizione scientifica digitale*

Angelo Mario Del Grosso, Angela Siciliano

CNR-Istituto di Linguistica Computazionale “A. Zampolli”, Università di Parma
angelomario.delgrosso@cnr.it, angela.siciliano@unipr.it

“Voice from the Margins”: Giorgio Bassani’s marginalia in a Digital Scholarly Edition

Abstract (ITA)

L'articolo presenta il lavoro di edizione scientifica digitale che è in corso di svolgimento sulle postille di Giorgio Bassani ai libri della sua biblioteca personale. Dopo aver esposto le ragioni per cui l'edizione digitale rappresenta la soluzione più adatta a rappresentare la complessa e strutturata testualità della postilla, si illustrano i criteri adottati nella codifica XML-TEI delle postille di Bassani al volume *La scuola dell'uomo* di Guido Calogero, centrale nella sua formazione umana e culturale. Infine, si descrivono le funzionalità di fruizione implementate dall'applicazione web tramite l'utilizzo del framework HTML templating della piattaforma eXist-db.

Abstract (ENG)

This article presents the ongoing work on the digital scholarly edition of Giorgio Bassani’s marginalia in the books from his personal library. After outlining the reasons why a digital edition is the most suitable approach for representing the complex and structured textual nature of marginalia, the paper illustrates the criteria adopted in the XML-TEI encoding of Bassani’s annotations in the book *La scuola dell'uomo* by Guido Calogero, a work central to his human and cultural development. Finally, the contribution describes the functionalities implemented in the web application through the use of the HTML templating framework of the eXist-db platform.

Parole chiave / Keywords

XML-TEI, postille, Giorgio Bassani, eXist-db, archivi digitali /
XML-TEI, marginalia, Giorgio Bassani, eXist-db, digital archives

* Sebbene i contenuti siano il risultato di un lavoro di confronto continuo, la redazione dei paragrafi è così suddivisa: il paragrafo 1 è stato scritto da Angela Siciliano; il paragrafo 2 da Angelo Mario Del Grosso; il paragrafo 3 è frutto di una riflessione e stesura congiunta.

1. Un modello di edizione digitale per i postillati di Giorgio Bassani

Negli ultimi decenni, la comunità scientifica ha mostrato un interesse sempre crescente nei confronti delle biblioteche d'autore,¹ riconoscendone l'importanza e le potenzialità sul piano interpretativo: anche solo scorrendo i titoli posseduti da un autore (scrittore, filosofo, scienziato, ecc.), possiamo delinearne il profilo culturale, la formazione e gli studi, le preferenze e le idiosincrasie. La nostra analisi può tuttavia spingersi a un livello più profondo, sottile e puntuale, se i volumi sono impreziositi da postille: seguendo e decodificando le annotazioni è possibile ricostruire il metodo di lettura di un autore, scoprire le fonti delle sue opere, o ancora della sua poetica e del suo pensiero.

È in tal senso altamente significativo il caso della biblioteca di Giorgio Bassani, conservata a Ferrara presso la sede dell'omonima Fondazione: l'esame dei 3395 volumi che la costituiscono, e in particolare dei 297 postillati,² ha consentito di indagare la personalità e la produzione dello scrittore ferrarese da un punto di vista inedito, contribuendo a mettere a fuoco periodi poco noti della sua biografia (come la militanza nell'antifascismo, tra il 1937 e il 1945)³ e a inquadrare correttamente il rapporto con i modelli letterari (primo tra tutti, Marcel Proust).⁴ Insieme alla conferma del valore critico dei postillati, lo studio della biblioteca di Bassani ha però indotto a una constatazione: i libri annotati sono documenti materialmente complessi, perché complessa e sfaccettata è la forma testuale della postilla. La postilla è, infatti, un testo che generalmente deriva da un altro testo (il brano a cui si riferisce) per commentarlo, spiegarlo,

1 All'interno della ricca bibliografia, segnaliamo almeno: Caproni 2004, Del Vento 2019, D'Iorio-Ferrer 2001, Pink 2018.

2 La biblioteca di Bassani ha rappresentato l'argomento della tesi di dottorato dell'autrice del paragrafo (Siciliano 2022), La tesi presenta, insieme al catalogo della biblioteca (pubblicato in Siciliano 2023), l'edizione integrale delle postille di Bassani a suoi libri, la loro discussione critica e un prototipo di edizione digitale delle annotazioni (vd. *infra*). Per quanto riguarda il prosieguo del progetto sul fronte digitale, si ringrazia la dottoressa Vanessa Bianconi per il lavoro di sviluppo completato nel contesto della sua tesi di laurea in Informatica Umanistica (presso l'Università di Pisa).

3 Siciliano 2023a.

4 Siciliano 2022a.

smentirlo.⁵ La sua complessità si rivela, poi, in altri due aspetti: da un lato, sfugge a una rigida e univoca categorizzazione, perché possiamo classificare le note per temi, per forma (postille verbali, di solo testo; postille non verbali, come barre laterali, sottolineature, asterischi; postille miste, formate da elementi verbali e non verbali), per funzione, per anni, ecc., ovvero categorie che non sono mutuamente esclusive; dall’altro, è al centro di una fitta rete intratestuale e intertestuale, perché può dialogare con annotazioni allo stesso volume, con note vergate dall’autore su altri libri, con passi dei manoscritti, delle lettere e dei taccuini, ecc.

Chi intenda realizzare l’edizione di un postillato dovrà porsi, allora, tre interrogativi:

- ◇ come rappresentare correttamente e adeguatamente il rapporto tra la postilla e il testo annotato?
- ◇ come classificare un oggetto plurimo e sfuggente?
- ◇ come evidenziare, in modo chiaro e dinamico, le relazioni concettuali tra la postilla e le altre note o quelle con i materiali d’archivio?

Tali interrogativi si sono trasformati in criticità nell’allestimento dell’edizione cartacea dei postillati bassaniani. Abbiamo riscontrato, infatti, che questo tipo di formalizzazione può confondere il lettore e, soprattutto, disperdere e/o dislocare l’informazione perché:

- ◇ ricorre a molteplici simboli e corpi tipografici per rendere il legame tra postilla e testo annotato (punto 1);
- ◇ indica solo implicitamente la categoria formale e, tra le altre potenziali classi, seleziona quella che si ritiene criticamente più pertinente e rilevante (tematica), per non congestionare la rappresentazione (punto 2);
- ◇ relega in una nota a piè di pagina l’illustrazione dei rapporti con le altre note o con altri documenti (punto 3).

Le tre criticità si manifestano nell’edizione di una postilla all’*Orlando Furioso* dell’Ariosto (dal volume *Opere*, Trieste, Sezione letterario-artistica del Lloyd austriaco, 1857; Fig. 1). Innanzitutto, i triangoli rovesciati (►◄) rappresentano la barra laterale in cui il passo è racchiuso da Bassani per evidenziarlo (qui in corpo minore); la postilla (in corpo mag-

5 Un riferimento importante per la riflessione sulle testualità (nel nostro caso applicata alla postilla) è Boot 2009.

giore), è tuttavia specificamente riferita al sintagma «fronte lieta», per cui il vero punto di rappiccio della nota è indicato dal simbolo della losanga (◆): per rendere correttamente la struttura della postilla e il suo rapporto con il testo annotato è stato necessario, pertanto, ricorrere a un complesso di segni che non è immediatamente comprensibile per l'utente. In più, la classificazione formale (postilla mista) non è dichiarata, perché il dato si ricava dalla rappresentazione stessa. Tra le altre possibili categorie si segnala solo quella tematica (qui letteraria) e l'intreccio tra questa postilla al *Furioso* e un'altra postilla dello stesso volume (riferita alla sezione delle *Canzoni*) è compresso, perdendo visibilità, nella nota in calce alla pagina.

Di persona [Alcina] era tanto ben formata, / Quanto me' finger san pittori industri, / Con bionda chioma lunga ed annodata; / ► Oro non è che più risplenda e lustri. / Spargesi per la guancia delicata / Misto color di rose e di ligustri; / Di terso avorio era la ◆ fronte lieta, ◀ / Che lo spazio finia con giusta meta. [*Orlando Furioso*, VII, 11, vv. 1-8; p. 44; postilla di letteratura]

vedi | Canz.[one] I²¹

¹Bassani intende evidenziare l'analogia tra il sintagma «fronte lieta», qui riferito alla maga Alcina, e l'espressione «fronte serena», utilizzato da Ariosto nella *Canzone I*: con rimando circolare, Bassani annota anche la *Canzone I*, sottolineando in particolare l'attributo «serena» e commentandolo, significativamente, con la postilla «Aggettivo proprio di Alcina».

Fig. 1. Edizione critica di una postilla di Giorgio Bassani all'*Orlando Furioso* di Ludovico Ariosto.

Da queste considerazioni, non possono che sorgere altre domande: come classifichiamo in un'edizione cartacea una postilla che può appartenere a più categorie tematiche (una nota di letteratura e arte, di storia e politica, di filosofia e scienza, ecc.)? E, ancora, come rappresentare le postille esclusivamente non verbali?

Partendo dall'osservazione dei limiti dell'edizione cartacea, è stato quindi elaborato un modello di edizione digitale dei postillati di Bassani, che ovviasse alle criticità appena rilevate: un prototipo esemplato sul volume *La scuola dell'uomo* di Guido Calogero (Sansoni, 1939), tra i libri cardine nell'educazione antifascista dello scrittore, letto e annotato tra il 1941 e il 1942.⁶

⁶ Una più ampia illustrazione del prototipo è in Siciliano-Del Grosso 2022, Siciliano-Del Grosso 2023 e Siciliano *et al.* 2023. Per una panoramica su progetti digitali dedicati ai postillati si rimanda al primo contributo (da integrare con Milone 2022). Sul processo di edizione digitale un utile riferimento è Bozzi *et al.* 2019.

Il primo passo è consistito nella scelta del linguaggio di codifica: l’XML-TEI.⁷ La rappresentazione digitale multidimensionale, multimodale e gerarchica di un oggetto testuale suggerita dalle linee guida del markup XML-TEI è sembrata fin da subito, e si è puntualmente dimostrata, la più adatta a riprodurre la testualità strutturata e composita della postilla, con un’ottima resa in termini di espressività e completezza dell’informazione. Vediamo come, presentando ordinatamente la struttura del modello: prima i metadati e poi l’edizione vera e propria. Prenderemo come caso di studio la p. 96 della *Scuola dell’uomo* (Figura in appendice), in cui figurano due annotazioni di Bassani.

A livello dei metadati, ogni postilla è stata considerata e descritta come un `<msItem>` nel `<msDesc>` (Fig. 2):

```
<msItem n="99" xml:id="p99" class="#postilla_filosofica_di_commento #sottolineatura_interlineare"
facs="#pag96.p1">
  <msItem n="99.1" class="#postilla_filosofica_di_commento">
    <locus target="#t96.1" facs="#pag96.p1det"> p. 96 </locus>
    <textLang> italiano </textLang>
    <note resp="#AS"> postilla verbale a lapis; sul margine sinistro; presenta due parole
sottolineate </note>
  </msItem>
  <msItem n="99.2" class="#sottolineatura_interlineare">
    <locus facs="#pag96.p1det1"> p. 96 </locus>
    <note resp="#AS"> postilla muta, costituita sottolineatura interlineare; a lapis </note>
  </msItem>
</msItem>
<msItem n="100" xml:id="p100" class="#postilla_filosofica_di_commento #barra_laterale_ondulata"
facs="#pag96.p2">
  <msItem n="100.1" class="#postilla_filosofica_di_commento">
```

Fig. 2. Frammento di codice XML relativo alla sezione `<msItem>`.

È questa una soluzione innovativa, non prevista dalla TEI, e rispondente a una diversa interpretazione della postilla. Nel modulo 11.7.2 delle linee guida,⁸ la TEI prescrive che le annotazioni siano presentate solo a livello descrittivo, con il tag `<additions>` in `<msDesc>` (e in particolare nella sezione `<PhysDesc>`): le considera cioè semplici aggiunte, appendici marginali al testo postillato (manoscritto o a stampa), che è l’oggetto primario di studio e rappresentazione. Il nostro punto di vista è invece opposto: in quanto testo d’autore, è la postilla ad essere gerarchicamente superiore al testo annotato, e come tale va rappresentata.

7 Burnard 2014; Buzzoni 2016; Pierazzo 2015.

8 <https://www.tei-c.org/release/doc/tei-p5-doc/it/html/MS.html> (maggio 2025).

La descrizione della postilla comprende, tra i vari aspetti distintivi, anche le categorie in cui essa è classificata: ogni `<msItem>` è qualificato dall'attributo `@class`, che punta alle tassonomie codificate, con l'elemento `<taxonomy>`, in un'altra sezione dei metadati, il `<classDecl>`. Per le note alla *Scuola dell'uomo*, sono state individuate tre `<taxonomy>` (Fig. 3, Fig. 4, Fig. 5): formale, funzionale e tematica (queste due sono valide soltanto per le postille contenenti testo, cioè le verbali o le miste).

```

<taxonomy>
  <category xml:id="postilla_verbale"/>
  <category xml:id="postilla_non_verbale">
    <category xml:id="sottolineatura_interlineare"/>
    <category xml:id="barra_laterale_ondulata"/>
    <category xml:id="doppia_barra_laterale_ondulata"/>
    <category xml:id="barra_laterale"/>
    <category xml:id="doppia_barra_laterale"/>
    <category xml:id="tripla_barra_laterale"/>
    <category xml:id="parentesi_quadra"/>
    <category xml:id="doppia_parentesi_quadra"/>
    <category xml:id="parentesi_quadra_spezzata"/>
    <category xml:id="parentesi_quadra_e_barra_laterale"/>
    <category xml:id="parentesi_quadra_e_sottolineatura_interlineare"/>
    <category xml:id="parentesi_quadra_e_ics_e_sottolineatura_interlineare"/>
    <!-- ... -->
  </category>
  <category xml:id="postilla_mista">
    <category xml:id="postille_verbale_legata_a_segno_non_verbale"/>
    <category xml:id="segno_non_verbale">
      <category xml:id="sottolineatura_interlineare_snv"/>
      <category xml:id="barra_laterale_ondulata_snv"/>
      <category xml:id="doppia_barra_laterale_ondulata_snv"/>
      <category xml:id="barra_laterale_snv"/>
      <category xml:id="doppia_barra_laterale_snv"/>
      <category xml:id="tripla_barra_laterale_snv"/>
      <category xml:id="parentesi_quadra_snv"/>
      <category xml:id="doppia_parentesi_quadra_snv"/>
      <category xml:id="parentesi_quadra_spezzata_snv"/>
      <category xml:id="parentesi_quadra_e_barra_laterale_snv"/>
      <category xml:id="parentesi_quadra_e_sottolineatura_interlineare_snv"/>
      <!-- ... -->
    </category>
  </category>
</taxonomy>

```

Fig. 3. Frammento di codice XML con definizione della tassonomia formale.

```
<taxonomy>
  <category xml:id="postilla_di_commento">
    <category xml:id="postilla_esegetica"/>
    <category xml:id="postilla_critico-dialogica"/>
    <category xml:id="postilla_polemica"/>
    <category xml:id="postilla_elogiativa"/>
    <category xml:id="postilla_personale"/>
  </category>
  <category xml:id="postille_di_correzione_e_integrazione">
    <category xml:id="correzione_di_refuso"/>
    <category xml:id="integrazione_testuale"/>
  </category>
  <category xml:id="postille_di_soglia">
    <category xml:id="postilla_su_frontespizio"/>
    <category
      xml:id="postilla_riferita_a_titolo_o_intertitolo"/>
  </category>

```

Fig. 4. Frammento di codice XML con definizione della tassonomia funzionale.

```
<taxonomy>
  <category xml:id="postilla_di_filosofia"/>
  <category xml:id="postilla_di_letteratura"/>
  <category xml:id="postilla_di_storia"/>
  <category xml:id="postilla_di_pedagogia"/>
  <category xml:id="postilla_di_lingua"/>
  <category xml:id="postilla_di_poetica"/>
  <category xml:id="postilla_di_stile"/>
  <category xml:id="postilla_di_arte"/>

```

Fig. 5. Frammento di codice XML con definizione della tassonomia tematica.

I vantaggi della rappresentazione digitale, legati alla struttura gerarchica ed elastica della TEI, sono evidenti:

- ◇ è possibile una triplice marcatura tipologica delle postille, di contro alla doppia dell’edizione cartacea, formale-tematica (che per il primo aspetto è per di più implicita);
- ◇ la classificazione è più strutturata, granulare e graduabile, perché ogni tassonomia può contenere più elementi, in cui si annidano altri elementi di livello inferiore (ma più precisi). In `<msItem>` la nota

di Bassani sarà perciò marcata riferendosi all'elemento più specifico, perché la grammatica gerarchica della TEI consente di ricavare l'informazione generale dalla proprietà particolare: la prima postilla di p. 96,⁹ ad esempio, sul piano tematico è codificata come postilla personale (<category> inferiore) e non come postilla critico-dialogica o di commento (<category> superiore).¹⁰ Sul versante formale, invece, la nota è marcata non come postilla mista, ma con l'elemento subordinato "postilla verbale legata a segno non verbale": questa soluzione consente, significativamente, di scomporre la postilla in tutte le sue componenti verbali e segniche, che sono rappresentate come <msItem> annidati nel <msItem> genitore.

- ◇ una nota può appartenere a più <category> di una tassonomia: la stessa di nota di p. 96 è contemporaneamente classificata come postilla di storia, di filosofia, di politica, risolvendo il problema della sovrapposizione di sottocategorie.

Passiamo ora all'edizione delle postille, che si presenta in una duplice forma: critica e diplomatica.¹¹ L'articolazione della prima rispecchia, coerentemente, la prospettiva adottata nel <msDesc>: il rovesciamento della gerarchia postilla-testo annotato. Mentre la TEI prevede che in <text> le postille siano codificate con i tag <note> o <label> all'interno del testo annotato (sempre come appendice o aggiunta),¹² nella nostra edizione critica ogni postilla è trascritta come elemento <div> in un primo <text> (Fig. 6) e in un secondo <text>, in posizione chiaramente subalterna, si riportano le pericopi di testo annotato, anch'esse considerate <div> (Fig. 7):

9 La riproduzione digitale fototipica è allegata in appendice, in calce al contributo.

10 Secondo classificazione delle note di Bassani (che riprende, adattandola, quella elaborata per Alfieri in Del Vento 2019) le postille di commento sono quelle che illustrano il testo secondo varie modalità: puntualmente (esegetiche), formulando un giudizio e/o dialogando con esso (critico-dialogiche), confrontandolo con altri testi (intertestuali), analizzandone gli aspetti tecnico-strutturali (metriche e prosodiche), facendone il punto di partenza per nuovo testo d'autore (di genesi). Le postille personali sono una sottocategoria delle critico-dialogiche: si tratta, nello specifico, delle note con cui Bassani non prende posizione sul testo (esprimendo il suo consenso o dissenso), ma se appropriata commentandolo attraverso il filtro delle sue passioni e delle conoscenze pregresse.

11 Boschetti-Del Grosso 2020; Nava 2022.

12 Estill 2016.

```
<div n="99" corresp="#p99" xml:id="t96.1" type="postilla" subtype="marginalia">
  <p> ma la storia è sempre prodotto, risultato, della volontà umana? La storia non travolge spesso la
  volontà? L'atto dell'uomo è sempre documento di una sufficiente consapevolezza? la storia non è qualche
  volta il segno di una irrazionalità irrefrenabile, il regno del dio cattivo? </p>
</div>
<div n="100" corresp="#p100" xml:id="t96.2" type="postilla" subtype="marginalia">
  <p> la storia come educatrice <persName ref="#Croce"> Croce </persName> </p>
//144...
```

Fig. 6. Frammento di codice XML relativo alla sezione <text> (trascrizione delle postille).

```
<div n="114" xml:id="idtext114" type="printedtext" facs="#pag96.p1t">
  <ab xml:id="ab114">
    ultime azioni dovettero pure esser sentite come preferibili ad ogni
    altra dalla presente coscienza di chi le compì, giacché, altrimenti, esse non
    sarebbero state compiute.
  </ab>
</div>
<div n="115" xml:id="idtext115" type="printedtext" facs="#pag96.p2t">
  <ab xml:id="ab115">
    ...
```

Fig. 7. Frammento di codice XML relativo alla sezione <text> (trascrizione del testo annotato).

L'edizione diplomatica è invece accolta nel <sourceDoc>, il modulo che riproduce la dimensione documentale della postilla, esplicitandone le caratteristiche genetiche (cassature, correzioni, sostituzioni) e materiali (articolazione spezzata su più righe, abbreviazioni, lacune, ecc.). Sofferiamoci ancora sulla prima postilla di p. 96 (Fig. 8):

```
<surface xml:id="pag96">
  <graphic url="pag96.jpg" width="1668px" height="2500px"/>
  <zone xml:id="pag96.p1" points = ' 1035,445 1030,489 164,491 165,1231 511,1228 513,671
  1097,661 1097,608 1483,608 1478,429 '>
    <zone xml:id="pag96.p1det" ulx="164" uly="533" lrx="511" lry="1230" hand="#m1">
      <line> ma la storia è </line>
      <line> <hi rend="italic"> sempre </hi> prodotto, risul </line>
      <line> tato, della volontà </line>
      <line> umana? La storia </line>
      <line> non travolge spesso </line>
      <line> la volontà? L'atto </line>
      <line> dell'uomo è <hi rend="italic"> sempre </hi> </line>
      <line> documento di una </line>
      <line> sufficiente consapevo </line>
      <line> lezza? La storia </line>
      <line> non è qualche volta </line>
      <line> il segno di una </line>
      <line> irrazionalità irrefre </line>
      <line> nabile, il regno del </line>
      <line> dio cattivo? </line>
```

Fig. 8. Frammento di codice XML relativo al modulo <sourceDoc>.

Contestualmente, il `<sourceDoc>` consente di descrivere – con il metodo dell'*embedded-transcription* – la postilla come una superficie (`<surface>`) formata da più aree (`<zone>`), corrispondente al `<msItem>` genitore composto da più `<msItem>`: la corrispondenza è segnalata attraverso l'attributo `@facs` in `<msItem>` (nell'elemento `<locus>`), che punta all'`@xml:id` della relativa `<zone>` nel `<sourceDoc>`. Inoltre, il legame tra la nota di Bassani («ma la storia [...] dio cattivo?») e il testo a stampa è espresso agganciando la `<zone>` della postilla a quella della sottolineatura interlineare con l'attributo `@target` e, parallelamente, a quella della porzione di testo che la sottolineatura evidenzia, trascritta come `<div>` nel secondo `<text>` e qui ripresa con l'attributo `@start`. Attraverso questo articolato sistema di rimandi è possibile, quindi, rappresentare in modo chiaro, razionale, sistematico il rapporto testo annotato-postilla e tra le componenti della nota.

Il modello non registra, però, soltanto le postille contenenti testo (miste, come quella di p. 96, o semplicemente verbali): si applica anche, con una leggera modifica, alle postille esclusivamente non verbali. Ciascuna di esse è, infatti, codificata come un `<msItem>` nel `<msDesc>` (Fig. 9), come `<metamark>` nel `<sourceDoc>` (Fig. 10) e tipologicamente in `<classDecl>` (Fig. 3).

```
<msItem n="106" xml:id="p106" class="#sottolineatura_interlineare">
  <locus target="#t106.1" facs="#pag106.p1"> p. 106 </locus>
  <note resp="#AS"> sottolineatura interlineare a lapis </note>
</msItem>
<msItem n="107" xml:id="p107" class="#barra_laterale_ondulata">
  <locus target="#t106.2" facs="#pag106.p2det"> p. 106 </locus>
  <note resp="#AS"> barra laterale ondulata a lapis, sul margine sinistro </note>
</msItem>
```

Fig. 9. Frammento di codice XML relativo alla sezione `<msItem>` (per le postille esclusivamente non verbali).

```
<surface xml:id="pag106">
  <graphic url="pag106.jpg" width="1668px" height="2500px"/>
  <zone xml:id="pag106.p1" points = ' 825,1254 829,1298 597,1306 599,1402 1317,1394 1315,1346 1469,1348 1463,1238 ' hand="#m1" corresp="#pag106.p1t">
    <metamark function="highlighting" rend="underlining" place="inline" target="#pag106.p1"/>
  </zone>
  <zone xml:id="pag106.p1t" ulx="590" uly="1247" lrx="1467" lry="1386" type="printedtext" start="idtext122"/>
  <zone xml:id="pag106.p2" points = ' 565,1770 565,1922 727,1918 727,1874 1469,1860 1465,1766 '>
  <zone xml:id="pag106.p2det" ulx="569" uly="1784" lrx="611" lry="1918" hand="#m1" corresp="#pag106.p2t">
    <metamark function="highlighting" rend="wavy" place="vertical_line" target="#pag106.p2"/>
  </zone>
</surface>
```

Fig.10. Frammento di codice XML relativo al modulo `<sourceDoc>` (per le postille esclusivamente non verbali).

Abbiamo visto, dunque, come il modello di edizione digitale offra una risposta, adeguata e concreta, ai problemi di classificazione e rappresentazione delle postille. Resta da valutare la sua efficacia in termini di sistematizzazione dei dati. Da questo punto di vista, il prototipo è in grado di esprimere in forma ipertestuale, sintetica e dinamica, la relazione concettuale tra una postilla e le altre presenti nello stesso volume.

La nota di p. 96 è, in tal senso, nuovamente esemplare. Si tratta, in fondo, di una domanda che Bassani rivolge a sé stesso (pensando all’infuriare della guerra e della violenza fascista): «ma la storia è *sempre* prodotto, risultato della volontà umana? La storia non travolge spesso la volontà? L’atto dell’uomo è *sempre* documento di una sufficiente consapevolezza? La storia non è qualche volta il segno di una irrazionalità irrefrenabile, il regno del dio cattivo?» La risposta sembra arrivare da una postilla di p. 98, secca e laconica: «ma la nostra è in gran parte l’epoca delle tirannie irresponsabili, al di là della psicologia e della volontà». Il rapporto tra le due note è codificato nell’elemento `<listRelation>` (Fig. 11), incluso in `<standOff>` (che segue la chiusura del `<teiHeader>`): `<listRelation>` contiene a sua volta l’elemento `<relation>`, in cui è possibile qualificare la relazione tra due entità identificandone la tipologia (“richiamo tra postille”, con l’attributo `@type`), denominandola (“domanda-risposta”, con l’attributo `@name`) e assegnando un ruolo attivo o passivo agli agenti coinvolti (rispettivamente, con gli attributi `@active` per la nota di p. 98 e `@passive` per quella di p. 96).

```
<standOff>
  <listRelation>
    <relation type="richiamo_tra_postille" name="domanda-risposta" passive="#p99" active="#p101" />
  </listRelation>
</standOff>
```

Fig. 11. Frammento di codice XML relativo alla sezione `<standOff>`.

2. L’ambiente web di interrogazione e fruizione dell’edizione delle postille di Bassani

Dal punto di vista strettamente tecnico, la realizzazione dell’edizione scientifica digitale delle postille si è sviluppata attraverso tre fasi principali. La prima si è concentrata sulla definizione del modello di rappresentazione del testo postillato, ovvero sulla formalizzazione della tipologia di testualità da adottare, che ha orientato l’intero processo di codifica. La seconda fase, quella di codifica vera e propria, ha compor-

tato l'identificazione e la marcatura dei fenomeni testuali definiti nel modello e presenti nella fonte primaria; tale operazione necessita di un sistema formale quale il linguaggio di markup XML¹³ vincolato dalle prassi definite nelle linee guida del consorzio TEI.¹⁴ La terza fase è stata dedicata all'elaborazione e alla restituzione del testo digitale all'interno di un ambiente per la conservazione, la consultazione e la fruizione da parte dell'utente.

Questa fase ha previsto la progettazione e l'implementazione di uno strumento open-source accessibile via web, pensato per l'interrogazione e la visualizzazione dei dati codificati. Lo sviluppo dell'applicazione è stato inizialmente orientato a soddisfare le esigenze del filologo interessato allo studio del pensiero e dell'opera di Bassani. Tuttavia, i principi che ne hanno guidato la realizzazione risultano pienamente generalizzabili e applicabili all'analisi di qualsiasi modalità di rappresentazione e fruizione di opere postillate e/o commentate, sia nell'ambito della filologia d'autore, sia in altri contesti di analisi filologica.



Fig. 12. Elenco dei volumi postillati presentati allo studioso dall'applicazione "Voce dei Margini".

13 <https://www.w3.org/TR/xml11/>.

14 <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>.



Fig. 13. Indice delle pagine per il volume selezionato.

Per quel che concerne la progettazione dell’ambiente digitale, il primo compito è stato quello di raccogliere i requisiti funzionali, ovvero quelle specifiche che definiscono le attività che il sistema deve essere in grado di svolgere per soddisfare gli obiettivi per cui è stato progettato. In questa fase sono state individuate le principali funzionalità e le modalità di interazione per l’utente finale. In questa fase sono state disegnate anche le principali interfacce grafiche di fruizione dei dati (GUI, *Graphical User Interfaces*), divise in tre macro categorie:

- ◇ L’unità testuale da approfondire a diversi livelli di granularità, quali ad esempio la scelta del volume da consultare e/o la pagina da consultare (Fig. 12, Fig. 13).
- ◇ Il dettaglio dell’attività postillatoria rispetto alla fonte primaria e alla modalità di codifica (Fig. 14).
- ◇ L’indagine, sia libera che guidata, consente di esplorare il testo postillato; ad esempio, è possibile utilizzare una preselezione di termini rilevanti per accedere alle loro occorrenze e visualizzare le concordanze con il testo annotato (Fig. 15).

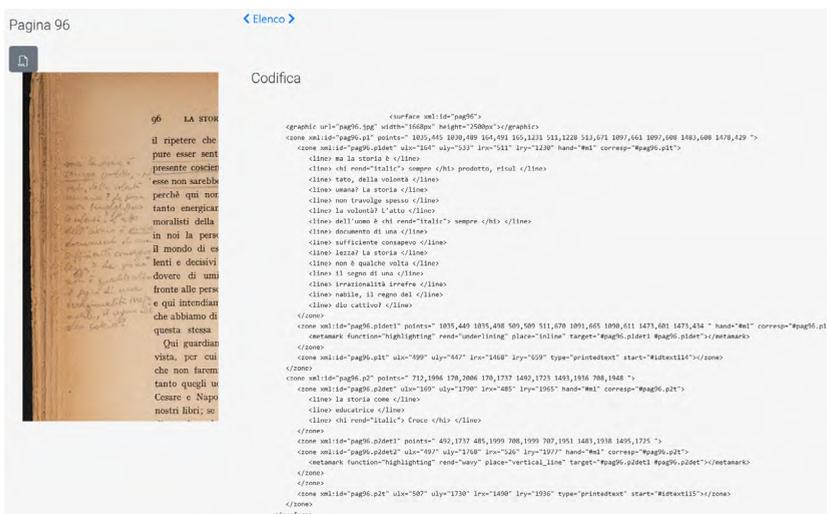


Fig.14. Fonte primaria pagina 96 del volume *La scuola dell'uomo* e rappresentazione digitale mediante XML-TEI.

Dalle postille verbali, annotate sul volume *La scuola dell'uomo*, è stato possibile creare un vocabolario dell'antifascismo che contiene le parole, annotate dal giovane Bassani, che rimandano alla sua attività antifascista risalente agli anni 1937-44

Selezione una parola:

Croce Educazione Etica Liberalismo Libertà Morale Politica Spirito Storia Storicismo

Parola selezionata: Croce

Occorrenze: 2

Pagina	Risultato	Testo a stampa
Pagina 34	Croce	Ciò può fare difficoltà solo a chi, abituato alle logiche e alle dialettiche del sì e del no, dell'affermazione e della negazione, creda che il mondo o lo spirito proceda per antitesi e per salti,
Pagina 96	la storia come educatrice Croce	Dal punto di vista di chi non ha più come agire, la storia non può presentarsi che nell'aspetto della pura comprensione, quale semplice ricostituzione sgustificatrice dell'esperienza di chi agì

Fig.15. Esempio di selezione di termini notevoli e occorrenze nel volume *La scuola dell'uomo*.

Una volta definiti anche gli aspetti operazionali del sistema, cioè dopo aver concordato le condizioni tecniche che devono essere soddisfatte affinché il sistema possa essere eseguito in modo efficace/efficiente, è stato possibile procedere con l'implementazione del prototipo dell'applica-

zione web, denominata “Voce dei Margini”.¹⁵ L’edizione digitale si è quindi concretizzata in uno strumento digitale destinato alla fruizione e all’interrogazione della collezione di testi postillati, precedentemente codificati in XML-TEI secondo le linee guida stabilite in una specifica ODD (One Document Does it all).¹⁶

Il prototipo è stato sviluppato interamente a partire dalle tecnologie messe a disposizione dalla piattaforma XML-nativa eXist-db.¹⁷ In particolare, la pubblicazione dell’edizione è resa possibile da un’applet web costruita mediante il modulo di “HTML templating framework”¹⁸ integrato nella piattaforma eXist-db. “Voce dei Margini” consente la lettura in modalità open access dell’intero *corpus* annotato della biblioteca di Bassani e ne supporta l’interrogazione avanzata attraverso strumenti adatti ad una indagine filologica.

L’impiego di un ambiente XML-nativo consente di gestire l’intero flusso di lavoro editoriale attraverso un unico modello dati efficace e coerente, sfruttando tecnologie perfettamente adatte agli obiettivi scientifici del progetto. Dalla gestione e codifica dei dati testuali, alla loro conservazione e indicizzazione, fino alla visualizzazione finale, ogni fase del processo è integrata in un ecosistema omogeneo e progettualmente sostenibile. La scelta dell’architettura applicativa riprende lo spirito – seppur con uno scopo meno ampio – del progetto TEIPublisher,¹⁹ che ha sviluppato un’applet in ambiente eXist-db per la realizzazione di edizioni digitali filologicamente curate e dotate di un elevato livello di complessità e di flessibilità nella gestione dell’informazione testuale.²⁰ Il nucleo tecnologico dell’applet è costituito dal linguaggio di interrogazione XQuery²¹ e dal framework HTML Templating integrato in eXist-db. Questa combinazione permette il recupero, l’elaborazione e l’organizzazione dinamica dei contenuti codificati, garantendo al tempo stesso un’elevata coerenza

15 Il prototipo in versione demo è accessibile al link: <https://48002.omega.ilc.cnr.it/exist/apps/postille/index.html>.

16 <https://tei-c.org/guidelines/customization/getting-started-with-p5-odds/>.

17 <https://exist-db.org/exist/apps/homepage/index.html>.

18 <https://github.com/eXist-db/templating>.

19 <https://teipublisher.com/exist/apps/tei-publisher-home/index.html>.

20 Il codice sorgente dell’applicazione “Voce dei Margini” è disponibile sulla piattaforma GitHub all’indirizzo: <https://github.com/CoPhi/postille>.

21 <https://www.w3.org/TR/xquery-31/>.

semantica dei dati e grande flessibilità nella progettazione delle interfacce e delle modalità di interazione con l'utente.

Più nel dettaglio, il modulo eXist-db *HTML Templating Framework* si articola in tre componenti principali, che implementano un noto *pattern* architeturale denominato in gergo tecnico *Model-View-Controller* (MVC):²²

- ◇ Model: costituito dallo schema adottato per la realizzazione dei documenti codificati in XML;
- ◇ View: un modulo responsabile della generazione dinamica delle pagine HTML a partire da specifici template ;
- ◇ Controller: implementato mediante istruzioni XQuery che gestiscono l'integrazione tra la componente di visualizzazione e il modello dati sottostante.

Le immagini che seguono illustrano un tipico caso d'uso implementato nel progetto di edizione delle postille illustrando nel dettaglio le diverse componenti dell'architettura del framework HTML Templating coinvolte nel processo di sviluppo. In particolare, la Fig. 16 mostra un dettaglio della pagina web accessibile allo studioso a partire dalla selezione di uno dei volumi della biblioteca. Tra le informazioni visualizzate si distinguono vari metadati bibliografici e descrittivi, segnatamente: il titolo del volume (*La scuola dell'uomo*), l'autore (Guido Calogero), l'edizione (prima edizione), la casa editrice (Sansoni, Firenze), l'anno di pubblicazione (1939), la collana (Biblioteca Italiana, 17), il luogo di conservazione (Biblioteca di Giorgio Bassani, Fondazione Giorgio Bassani, Ferrara), e la datazione delle postille (1941–1942).

La Fig. 17 mostra un frammento di codice HTML impiegato per la generazione dinamica della pagina web illustrata in precedenza all'interno dell'applicazione (Fig. 16). Si evidenziano alcune peculiarità nell'uso del framework, in particolare attraverso gli attributi `@data-template` e `@data-template-with`, conformi alla sintassi standard del linguaggio HTML5. Gli attributi introdotti permettono di integrare direttamente nel markup le logiche di accesso, di analisi e di elaborazione dei dati codificati. In particolare, i valori specificati all'interno degli attributi `@data-template` degli elementi `<p>`, annidati nel contenitore `<div>` con attributo `@id="librodet"`, corrispondono a vere e proprie invo-

22 <https://martinfowler.com/eaCatalog/modelViewController.html>.

cazioni di procedure computazionali che accedono ai documenti XML. Tali dati costituiscono la base informativa dell’edizione e vengono archiviati e indicizzati nativamente nella piattaforma eXist-db, che ne garantisce l’accessibilità, la persistenza e l’interrogabilità. Tra i valori richiamati si annoverano, a titolo esemplificativo, `app:stampaautore`, `app:stampaedizione`, `app:stampaeditore`, `app:annedizione`, `app:serie`, `app:biblioteca` e `app:biblioluogo`: ciascuno di essi corrisponde a uno specifico metadato bibliografico, coerente con quelli visualizzati in Fig. 16.

```
<div xmlns="http://www.w3.org/1999/xhtml" data-template="templates:surround"
data-template-with="templates/page.html"
  data-template-at="content">
  <div class="contenuto1"> ...
  </div>
  <div class="infolibro">
    <button type="button" class="btn btn-danger" id="dettagliibro">
      <i class="bi bi-info-circle" />
    </button>
    <div class="info" id="librodet" style="display: none;">
      <p data-template="app:stampaautore" />
      <p data-template="app:stampaedizione" />
      <p data-template="app:stampaeditore" />
      <p data-template="app:annedizione" />
      <p data-template="app:serie" />
      <p data-template="app:biblioteca" />
      <p data-template="app:biblioluogo" />
    </div>
  </div>
  <div> ...
  </div>
```

Fig. 16. Dettaglio applicazione “Voce dei Margini” con i dati relativi al volume.

```
<div xmlns="http://www.w3.org/1999/xhtml" data-template="templates:surround"
data-template-with="templates/page.html"
  data-template-at="content">
  <div class="contenuto1"> ...
  </div>
  <div class="infolibro">
    <button type="button" class="btn btn-danger" id="dettagliibro">
      <i class="bi bi-info-circle" />
    </button>
    <div class="info" id="librodet" style="display: none;">
      <p data-template="app:stampaautore" />
      <p data-template="app:stampaedizione" />
      <p data-template="app:stampaeditore" />
      <p data-template="app:annedizione" />
      <p data-template="app:serie" />
      <p data-template="app:biblioteca" />
      <p data-template="app:biblioluogo" />
    </div>
  </div>
```

Fig. 17. Frammento di codice HTML con la definizione del template per la pagina Web relativa al dettaglio dei dati di descrizione bibliografica dei volumi postillati.

Le procedure richiamate sono implementate attraverso istruzioni scritte nel linguaggio XQuery, specificamente concepito per la manipolazione di dati XML. Esse contengono la logica applicativa necessaria per accedere ai dati (mediante espressioni XPath),²³ elaborarli secondo necessità (attraverso funzioni XPath e XQuery)²⁴ e generare i frammenti HTML che andranno a popolare dinamicamente il documento di template nel punto esatto in cui è presente la chiamata alla funzione.

La Fig. 18 mostra il dettaglio di due funzioni, `app:annoedizione` e `app:serie`, le quali definiscono un insieme di istruzioni atte a estrarre, processare e restituire i dati da visualizzare all'interno della pagina web (Fig. 16), tramite l'inserimento dinamico delle informazioni all'interno del codice HTML illustrato in Fig. 17. Senza addentrarsi in dettagli eccessivamente tecnici, si può osservare come entrambe le funzioni contengano: un'istruzione per l'accesso al documento XML dell'edizione (`postille.xml`); un'espressione XPath per selezionare l'anno dell'edizione del volume corrente (nella funzione `app:annoedizione`); e, nella funzione `app:serie`, ulteriori espressioni XPath per l'estrazione delle informazioni relative al titolo e al numero della collana. Infine, l'istruzione di restituzione (`return`) produce il frammento HTML, completato con i dati selezionati ed eventualmente manipolati.

```

declare function app:annoedizione($node as node(), $model as map(*)){
  let $doc := doc("/db/apps/postille/filexml/postille.xml")
  let $anno := $doc//tei:biblStruct[@xml:id=$app:volume]//tei:imprint/tei:date
  return
  <p id="annoedizione"><b>Anno:</b>{data($anno)}</p>
};

declare function app:serie($node as node(), $model as map(*)){
  let $doc := doc("/db/apps/postille/filexml/postille.xml")
  let $serietit := $doc//tei:biblStruct[@xml:id=$app:volume]//tei:series/tei:title
  let $serienum := $doc//tei:biblStruct[@xml:id=$app:volume]//tei:series/tei:biblScope
  return
  <p id="serie"><b>Collana:</b>{data($serietit)}, {data($serienum)}</p>
};

```

Fig. 18. Frammento di codice XQuery con la definizione di due funzioni `app:annoedizione` e `app:serie`, che accedono alla collezione di dati XML e restituiscono rispettivamente l'anno di edizione del volume selezionato e il titolo della collana.

23 <https://www.w3.org/TR/xpath-31/>.

24 <https://www.w3.org/TR/xpath-functions-31/>.

L’applicazione di filologia computazionale offre una presentazione integrale del *corpus* delle postille, affiancando le annotazioni autografe al corrispondente testo a stampa (Fig. 19). Come anticipato, una volta selezionato un volume, l’utente può accedere alla riproduzione digitale fototipica delle singole pagine annotate, selezionandole una ad una.

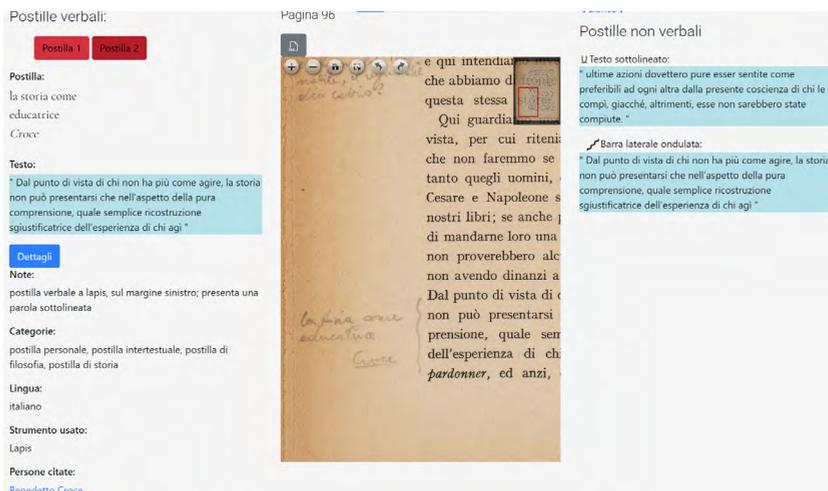


Fig. 19. Pagina di consultazione delle postille presenti su pagina 96 della *Scuola dell'uomo*.

Ogni elemento postillato è rappresentato all’interno di una pagina web che offre una vista completa e strutturata dei contenuti. Al centro della schermata viene visualizzata la scansione della pagina a stampa, resa possibile grazie al componente *JavaScript OpenSeadragon*.²⁵ Sulla sinistra sono presentate le trascrizioni delle postille verbali, selezionabili eventualmente tramite un apposito pulsante. Più in basso, nella stessa sezione, vengono segnalati gli interventi d’autore sul testo originale trascritto. Sul lato destro della schermata si trovano invece le descrizioni delle annotazioni non verbali e i collegamenti tra le postille e il testo annotato. Infine, nella parte inferiore della schermata, a sinistra, sono visualizzati i dati descrittivi e catalografici relativi al volume e alle singole postille (Fig. 19). Nel caso specifico della postilla esemplificativa «la storia come | educatrice | Croce»,

25 <https://openseadragon.github.io/>.

la sezione “Note” riporta il testo: «postilla verbale a lapis, sul margine sinistro; presenta una parola sottolineata». Nella sezione “Categorie” sono indicate: postilla personale, postilla intertestuale, postilla di filosofia, postilla di storia. La “Lingua” è specificata come italiano, mentre lo “Strumento usato” è il lapis. Infine, nella sezione “Persone citate”, compare Benedetto Croce: il nome è attivo sulla pagina web tramite un collegamento ipertestuale che rimanda alla voce corrispondente su Wikipedia²⁶ dedicata al celebre filosofo e critico letterario italiano.²⁷

L'applicazione integra anche un modulo di analisi linguistica automatica che associa a ciascuna parola del testo trascritto della postilla una serie di annotazioni tramite il modulo di Natural Language Processing sviluppato dalla Stanford University,²⁸ disponibile come componente per la piattaforma eXist-db.²⁹ Nella versione attuale, il modulo consente di estrarre automaticamente informazioni linguistiche di base, come ad esempio l'identificazione dei nomi propri di persona (Fig. 20). Ogni parola viene rappresentata all'interno di un elemento <word>, accompagnato da un elemento <POS> che specifica la parte del discorso (*Part of Speech*). Nel caso mostrato in figura, la parola “educatrice” riceve l'annotazione POS uguale alla lettera “S”, a indicare un sostantivo, mentre per la parola “Croce”, la POS è “SP”, a segnalare un nome proprio di persona. L'elemento <NER> (*Named Entity Recognition*) specifica se la parola è identificata come entità nominata.

Il modulo viene importato nell'applicazione tramite l'istruzione XQuery `import module namespace nlp30 e successivamente utilizzato con la funzione nlp:parse($testo, $opzioni). Questa funzione riceve in input il testo da annotare ($testo) e alcune opzioni di configurazione ($opzioni), restituendo una struttura dati XML gerarchica. La struttura è descritta dalla radice <StanfordNLP>, che contiene, a livelli successivi, gli elementi <sentences>, <sentence>, <tokens>, e <token>.`

26 Si veda la pagina Wikipedia all'indirizzo web https://it.wikipedia.org/wiki/Benedetto_Croce.

27 Il riferimento biografico attualmente rimanda a Wikipedia, ma dal punto di vista tecnico è possibile integrare anche altre risorse online che raccolgono e pubblicano profili di personaggi illustri, come ad esempio le voci del *Dizionario Biografico degli Italiani* o dell'*Enciclopedia Treccani*.

28 Palmero-Moretti 2016.

29 <https://github.com/eXist-db/exist-stanford-nlp>.

30 Con valore del modulo uguale a “<http://exist-db.org/xquery/stanford-nlp>”.

L’elemento `<token>` include a sua volta gli elementi `<word>`, `<CharacterOffsetBegin>`, `<CharacterOffsetEnd>`, `<POS>` e `<NER>`, che tra loro condividono lo stesso livello gerarchico (*sibling*).

The screenshot displays a user interface for linguistic analysis. At the top, under the heading "Postilla:", there is a list of items: "la storia come", "educatrice", and "Croce". Below this, under the heading "Testo:", a light blue highlighted text block contains the quote: " Dal punto di vista di chi non ha più come agire, la storia non può presentarsi che nell'aspetto della pura comprensione, quale semplice ricostruzione sgiustificatrice dell'esperienza di chi agì ". A dark button labeled "Analisi linguistica" is positioned below the text. A window titled "Analisi linguistica" is open, showing the following XML-like structure for the word "educatrice":

```
<word>educatrice</word>  
  
<POS>S</POS>  
  
<NER>O</NER>  
  
</StanfordNLP>  
  
<StanfordNLP>
```

Fig. 20. Analisi linguistica apportata automaticamente al testo della postilla.

Tra le funzionalità implementate nell’applicazione “Voce dei Margini” figura un sistema di interrogazione del *corpus* particolarmente flessibile. Il componente è accessibile tramite un’interfaccia minimale ma altamente efficace mostrata in Fig. 21. Attualmente sono supportate quattro modalità di ricerca: 1) Ricerca con *wildcard*, che consente di individuare parole parziali o forme variabili. Ad esempio, inserendo nella maschera di ricerca la stringa “lib*”, il sistema restituirà tutte le postille contenenti parole come “liberalismo”, “libertà” o “liberale”; 2) Ricerca fuzzy, utile per tollerare errori di digitazione o varianti ortografiche. Ad esempio, digitando “rilce”, il sistema sarà in grado di restituire le postille che contengono la parola “Rilke”; 3) Ricerca di prossimità, finalizzata a individuare la co-occorrenza di due o più termini all’interno di una finestra di

contesto personalizzabile. Un esempio concreto è riportato in Fig. 22, dove vengono ricercate le postille in cui la parola “storia” compare a una distanza massima di cinque parole dalla parola “Croce”; 4) Ricerca per entità nominate, che permette di esplorare il *corpus* attraverso nomi propri, toponimi o concetti storici e culturali. Per esempio, è possibile individuare le postille in cui l’autore menziona “Guicciardini” oppure lo stesso “Croce”. Il motore di ricerca è implementato configurando opportunamente il sistema di indicizzazione full-text integrato nella piattaforma eXist-db, basato sulla tecnologia *Lucene*.³¹ Grazie a questa ulteriore tecnologia, come visto, è possibile formulare interrogazioni complesse ottenendo risultati ad alta precisione.

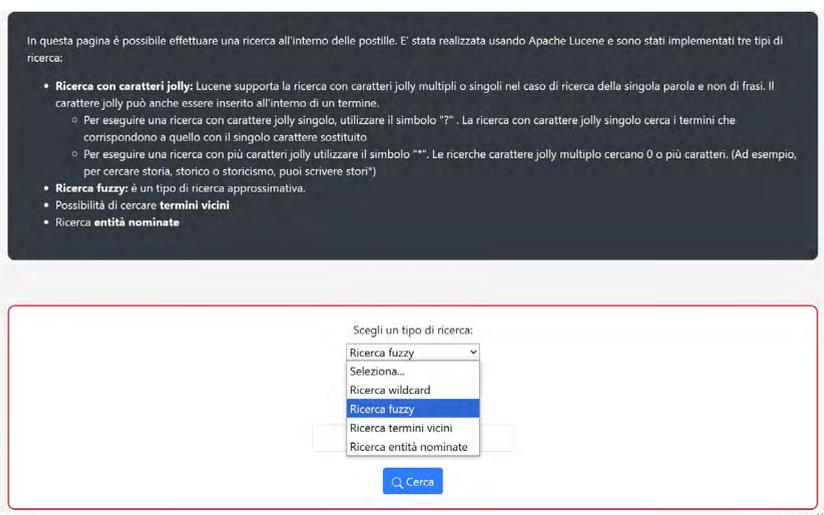


Fig. 21. Maschera di Ricerca implementata per l’applicazione “Voce dei Margini”.

Un’ulteriore funzionalità, complementare alla ricerca libera, è rappresentata dalla presenza di una pagina web dedicata con la presenza di una lista predefinita di termini “notevoli”. Come mostrato in Fig. 15, questa lista è composta da parole selezionate in precedenza dai curatori dell’edizione per la loro particolare rilevanza tematica in relazione al lessico del giovane Bassani antifascista (Croce, Educazione, Etica, Liberalismo, Libertà, Morale, Politica, Spirito, Storia, Storicismo).

31 <https://lucene.apache.org/>

Scegli un tipo di ricerca:
Ricerca entità nominate ▾

Tipo di ricerca: termini vicini
Termini cercati: storia Croce
Distanza selezionata: 5
Occorrenze: 1

Pagina	Risultato
Pagina 96	la storia come educatrice Croce

Fig. 22. Esempio di ricerca di prossimità con termini “storia” e “Croce” a distanza massima di 5 parole.

Alle funzionalità già descritte si affiancheranno, attualmente in fase di progettazione, ulteriori estensioni, tra cui: un collegamento ottimizzato tra le porzioni del testo a stampa e le relative immagini annotate; un motore di ricerca avanzato, capace di interrogare anche tutti i metadati descrittivi e catalografici dei volumi e delle singole postille; funzionalità sperimentali di text mining, orientate al raggruppamento automatico di postille simili secondo criteri linguistici, stilistici o tematici.

3. Una proposta di modello generale

Il nostro prototipo è stato concepito e organizzato sulle specificità strutturali, materiali e concettuali della forma testuale della postilla. Solo in un secondo tempo, abbiamo proceduto ad adattarlo al caso particolare delle annotazioni di Bassani alla *Scuola dell'uomo*. Il risultato è, perciò, un modello generale e flessibile che potrebbe essere – di volta in volta – personalizzato e specializzato sul caso di studio in esame, indipendentemente dall’epoca della postillatura e dall’autore delle note.

Per quanto riguarda la classificazione, ad esempio, nell’edizione digitale delle postille di Bassani abbiamo inserito solo le tre tassonomie fondamentali (formale, funzionale e tematica), tendenzialmente comuni a tutte le annotazioni. Tuttavia, nella struttura dell’elemento `<classDecl>` è possibile inserire ulteriori elementi `<taxonomy>`, per rappresentare ad esempio la data della postillatura, che non abbiamo considerato perché – come si è detto – il prototipo riguarda un unico volume, con annotazioni risalenti a un arco temporale omogeneo, 1941-1942. La tassonomia cronologica è risultata, quindi, non necessaria. Lo è, invece, se includiamo

nell'edizione un numero cospicuo di libri annotati dallo scrittore ferrarese o, ancora, se studiamo un singolo postillato con note databili a periodi diversi, vergate dalla stessa mano o da più mani.

Oltre a integrare nuove <taxonomy>, si può rimodulare l'articolazione interna di quelle di base. Prendiamo la <taxonomy> formale, che per Bassani prevede tre elementi <category>, ovvero postilla verbale, postilla non verbale e postilla mista, ciascuna strutturata per mezzo di successivi elementi <category> di livello inferiore. Se scorriamo le <category> incluse tra le postille non verbali notiamo, in particolare, l'assenza delle *maniculae*: un'assenza giustificata, perché questo tipo di annotazione – molto diffuso nei manoscritti medioevali³² e rinascimentali, e impiegato fino (almeno) a Manzoni³³ – scompare dalle abitudini dei postillatori novecenteschi. Se il nostro postillato presenta delle *maniculae*, potremmo allora aggiungerle tra le <category> delle postille non verbali; diversamente, nel caso in cui siano associate a una nota testuale, le codificheremmo tra le postille miste con la <category> “postilla verbale e manicola”.

L'elasticità del modello si rileva anche nella struttura dell'elemento <sourceDoc>. In questa sezione è possibile, infatti, rappresentare fenomeni materiali tipici dei postillati più antichi, che nelle note alla *Scuola dell'uomo* (e in generale nei libri annotati da Bassani) sono:

- ◇ assenti, come la presenza di più mani che postillano il testo (manoscritto o a stampa), distinguibili con l'attributo @hand nelle <zone> della postilla; o ancora le lacune, potenzialmente caratteristiche di volumi che hanno avuto una lunga circolazione o usurati dall'uso, che indicheremmo con l'elemento <gap> (se la porzione di testo mancante non è ricostruibile) o <damage> (se la lezione è congetturabile, e dunque proposta nell'elemento <supplied>).
- ◇ presenti, ma in una forma differente. È il caso delle abbreviazioni, che un postillatore novecentesco come Bassani (o anche un lettore di oggi) usa senza riferirsi a una precisa grammatica, e perciò con oscillazioni: il cognome dell'autore della *Scuola dell'uomo*, per esempio, potrebbe essere abbreviato con “C.” o “Cal.”. Al contrario, nei postillati più antichi le abbreviazioni partecipano di un sistema più standardizzato (anche se, talvolta, con piccole variazioni). La differenza principale,

32 Fiorilla 2015.

33 Martinelli 2018.

però, è che quest’ultime – a differenza di quelle di Bassani – presentano segni e caratteri tachigrafici speciali (si pensi, tra tutti, al *titulus*) che non sono sempre registrati nel *character set* di un documento XML: se non sono disponibili, occorre prima descriverli mediante l’espressività dell’elemento `<charDecl>` (figlio dell’elemento `<encodingDesc>`) con l’elemento `<char>` o `<glyph>`, e poi scioglierli nell’elemento `<g>` (nella struttura di rappresentazione del testo presente nel blocco `<text>`), senza che sia necessario ricorrere all’elemento `<abbr>`. Per Bassani, ricorremo invece al classico metodo di scioglimento delle abbreviazioni, con l’uso dell’elemento `<choice>` strutturato con gli elementi `<abbr>` ed `<expan>`.

Per valutare l’effettiva applicabilità del nostro modello, abbiamo deciso di sperimentarlo su un postillato rappresentativo di gran parte dei fenomeni sopra descritti (Fig. 23, Fig. 24): un esemplare della seconda edizione delle *Prose della volgar lingua* di Pietro Bembo (Venezia, Marcolini, 1538), di proprietà di un privato, e che è stato postillato da tre mani anonime, databili tra il XVI e il XVII secolo.³⁴ Lavori successivi documenteranno i dettagli della sperimentazione poc’anzi introdotta.

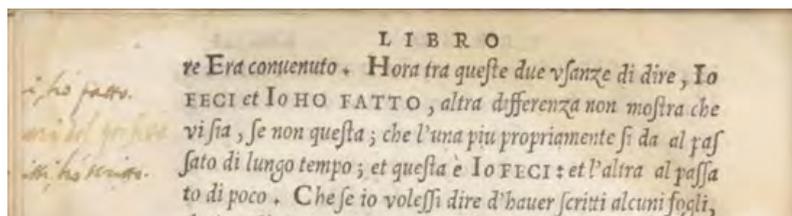


Fig. 23. Esempio di intreccio materiale di due mani e di lacuna (dovuta a rifilatura dell’esemplare).

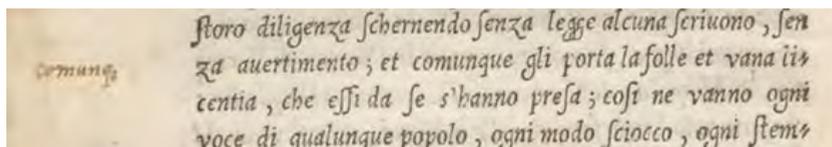


Fig. 24. Esempio di abbreviazione.

34 L’esemplare è stato utilizzato in un percorso di didattica digitale sulle *Prose*, curato da Matilde Passafaro: <https://pdd.fclit.unibo.it/pietro-bembo-prose/page/esplora-edizione>. Si ringrazia Paola Italia per averci consentito lo studio del postillato.

Bibliografia

- Boot 2009 = Peter Boot, *Mesotext: Digitised Emblems, Modelled Annotations and Humanities Scholarship*, Amsterdam, Pallas Publications, 2009.
- Boschetti-Del Grosso 2020 = Federico Boschetti, Angelo Mario Del Grosso, *L'annotazione di testi storico-letterari al tempo dei social media*, «*Italica Wratislaviensia*», 1 (2020), pp. 65-99.
- Bozzi *et al.* 2019 = Andrea Bozzi, Maria Sofia Corradini, Giacomo Ferrari, *Percorsi Di Linguistica e Di Filologia Computazionali*, Pisa, Edizioni ETS, 2019.
- Burnard 2014 = Lou Burnard, *What Is the Text Encoding Initiative?*, Marseille, OpenEdition Press, 2014.
- Buzzoni 2016 = Marina Buzzoni, *A Protocol for Scholarly Digital Editions? The Italian Point of View*, in *Digital Scholarly Editing: Theories and Practices*, ed. Matthew James Driscoll and Elena Pierazzo, Cambridge, Open Book Publishers, 2016, pp. 59-82.
- Caproni 2004 = Attilio Mauro Caproni, *Le biblioteche d'autore: definizione, caratteristiche e specificità: alcuni appunti*, «*Bibliotheca*», 1 (2004), pp. 15-22.
- Del Vento 2019 = Christian Del Vento, *La biblioteca ritrovata. La prima biblioteca di Vittorio Alfieri*, Alessandria, Edizioni dell'Orso, 2019.
- D'Iorio-Ferrer 2001 = *Bibliothèques d'écrivains*, sous la direction de Paolo D'Iorio et Daniel Ferrer, Paris, CNRS Éds, 2001.
- Estill 2016 = Laura Estill, *Encoding the Edge: Manuscript Marginalia and the TEI*, «*Digital Literary Studies*», 1 (2016).
- Fiorilla 2005 = Maurizio Fiorilla, *Marginalia figurati nei codici di Petrarca*, Firenze, Olschki, 2005.
- Martinelli 2018 = Donatella Martinelli, *Dalle orecchie di lettura ai collettori: nel cantiere manzoniano delle postille di lingua*, «*Prassi ecdotiche della modernità letteraria*», 3 (2108), pp. 233-263.
- Milone 2022 = Federico Milone, *Postille "amargine": problemi e metodi per l'edizione digitale dei postillati di autori del secondo Novecento*, «*Linguistica e Letteratura*», 1-2 (2022), pp. 195-209.
- Nava 2022 = Beatrice Nava, *Are We All Bédierian? Perspectives for Digital Genetic Editions*, «*Umanistica Digitale*», 14 (2002), pp. 19-40. <https://doi.org/10.6092/issn.2532-8816/14949>
- Palmero-Moretti 2016 = Alessio Palmero Aprosio, Giovanni Moretti, *Italy Goes to Stanford: A Collection of CoreNLP Modules for Italian*, *ArXiv E-Prints*, 2016. <http://arxiv.org/abs/1609.06204>

- Pierazzo 2015 = Elena Pierazzo, *Digital Scholarly Editing: Theories, Models and Methods. Digital Research in the Arts and Humanities*, Farnham Surrey, Ashgate, 2015.
- Pink 2018 = Gillian Pink, *Voltaire à l’ouvrage. Une étude de ses traces de lecture et de ses notes marginales*, Paris, CNRS Éds, 2018.
- Siciliano 2022 = Angela Siciliano, *La biblioteca di Giorgio Bassani: origini, trasformazioni, usi, intertestualità*, tesi di dottorato in Studi Italianistici, discussa presso l’Università di Pisa (cotutela con l’Université Grenoble Alpes), 2022.
- Siciliano 2022a = *In margine a Proust: le postille di Bassani alla «Recherche», tra memoria e stile*, in *Il ‘tono’ di Proust. Dagli avantesti alla ricezione*, a cura di Anna Dolfi, Firenze, Firenze University Press, 2022, pp. 412-426.
- Siciliano 2023 = Angela Siciliano, *Catalogo della biblioteca di Giorgio Bassani*, Ravenna, Giorgio Pozzi Editore, 2023.
- Siciliano 2023a = *Per una storia intellettuale dell’antifascismo di Bassani: letture, scritture, strategie della resistenza*, in *«Strapparsi di dosso il fascismo»: l’educazione di regime nella «generazione degli anni difficili»*, a cura di Rosanna Morace, Napoli, La scuola di Pitagora editrice, 2023, pp. 415-438.
- Siciliano-Del Grosso 2022 = Angela Siciliano, Angelo Mario Del Grosso *From Print to Digital: An Encoding Model for the Scholarly Edition of Giorgio Bassani’s Notes*, «Umanistica Digitale», 13 (2022), pp. 21-46. <https://doi.org/10.6092/issn.2532-8816/13688>
- Siciliano et al. 2023 = Angela Siciliano, Vanessa Bianconi, Angelo Mario Del Grosso, “Voce dei Margini”: *Le annotazioni di Giorgio Bassani*, in *La memoria digitale. Forme del testo e organizzazione della conoscenza*, a cura di Emmanuela Carbé, Gabriele Lo Piccolo, Valenti Alessia e Francesco Stella, Siena, AIUCD, 2023, <https://doi.org/10.6092/unibo/amsacta/7721>
- Siciliano-Del Grosso 2023 = Angela Siciliano, Angelo Mario Del Grosso, *Giorgio Bassani’s notes between tradition and innovation*, in *Digital Humanities 2023. Collaboration as Opportunity*, a cura di Walter Scholger, Georg Vogeler, Toma Tasovac, Anne Baillet, 2023. <https://doi.org/10.5281/zenodo.7961822>

II. SCHEDE BIBLIOGRAFICHE



LORELLA VIOLA, *The Humanities in the Digital: Beyond Critical Digital Humanities*, Cham, Palgrave Macmillan, 2023, XXVI + 173 pp.

In un contesto in cui il processo di digitalizzazione della società e dei saperi, accelerato dalla pandemia da COVID-19, è ormai divenuto strutturale, le Digital Humanities si configurano non più come un ambito collaterale o sperimentale, ma come uno snodo cruciale nella ridefinizione delle pratiche di produzione del sapere. Il volume di Lorella Viola, *The Humanities in the Digital: Beyond Critical Digital Humanities* (2023), si inserisce nel dibattito contemporaneo sull'evoluzione epistemologica del rapporto tra scienze umane e tecnologie digitali. Muovendo dal presupposto che il cambiamento nella produzione della conoscenza è stato *nel* (e non *verso il*) digitale, l'A. denuncia l'inadeguatezza della dicotomia fra digitale/non digitale, *humanities*/scienze dure nel descrivere l'incremento di sapere nel XXI secolo. Lungi dal proporre una semplice integrazione tecnologica, Viola auspica un cambiamento di paradigma: l'aumento delle conoscenze deve fondarsi su un modello collaborativo e simmetrico tra competenze, in cui la responsabilità epistemica sia condivisa.

Il volume si apre con una *premessa* di Johanna Drucker a cui seguono la *Prefazione* (pp. XIII-XVI), i *Crediti* (pp. XVII-XVIII), una breve biografia dell'autrice (pp. XXI), una lista di acronimi (pp. XXIII-XXIV), un *Indice delle figure* (pp. XXV-XXVI) e si articola quindi in cinque sezioni: 1) *The Humanities in the digital* (pp. 1-35); 2) *The Importance of Being Digital* (pp. 37-56); 3) *The Opposite of Unsupervised* (pp. 57-80); 4) *How Discrete* (pp. 81-106); 5) *What the Graph* (pp. 107-135); 6) *Conclusion* (pp.137-146). A conclusione del volume troviamo le note bibliografiche (pp. 147-167) e l'indice analitico (pp. 169-173).

Il primo capitolo parte da una panoramica sulle trasformazioni imposte dalla digitalizzazione prima e dopo la pandemia da COVID-19. L'A. sottolinea come la crisi sanitaria globale abbia agito da catalizzatore, imponendo un'accelerazione al processo di digitalizzazione in ambito universitario e accademico, mettendo in crisi le precedenti strategie 'evolutive' di digitalizzazione basate su un'integrazione graduale e parziale delle tecnologie digitali e il modello stesso di produzione del sapere. In proposito scrive:

I therefore maintain that universities and higher-education institutions can no longer afford to consider the digital as something

that is happening to knowledge creation. It is time to recognise that knowledge creation is happening in the digital. As digital vs non-digital positions have entirely lost relevance, we must recognise that the current model of knowledge, grounded in rigid divisions, is at best irrelevant and unhelpful and at worst artificial and harmful. Scholars, researchers, universities and institutions have therefore a central role to play in assessing how digital knowledge is created, not just today, but also for the purpose of future generations, and clear responsibilities to shoulder, those that come from being in the digital. (pp. 10-11)

Successivamente l'A. introduce la diffusa convinzione secondo cui gli algoritmi opererebbero in modo imparziale e oggettivo. Servendosi della metafora del *Mechanical Turk* del XVIII secolo – un automa che celava al suo interno un operatore umano – la studiosa mette in luce come la supposta neutralità dei processi computazionali celi, in realtà, una serie di interventi umani e l'influenza di categorie culturali preesistenti. In tale contesto, diventa imprescindibile adottare un atteggiamento di vigilanza critica nei confronti dei criteri di selezione, classificazione e interpretazione incorporati negli algoritmi: essi sono modellati da valori, ideologie e rapporti di potere umani e dunque culturalmente e storicamente orientati. Procedendo nell'analisi l'A. richiama il saggio di C. P. Snow (*The two cultures and the scientific revolution*, 2013) sul divario fra scienze e lettere, evidenziando come tale frattura persista nel dibattito contemporaneo, anche all'interno di contesti che si dichiarano interdisciplinari. Le stesse etichette *Digital Humanities* e *Critical Digital Humanities*, anziché rappresentare un reale superamento delle barriere epistemologiche, finiscono per reiterare e consolidare la divisione tra ambiti del sapere, perpetuando l'illusione di una loro conciliazione. Riguardo le DH in particolare, l'A. commenta:

This model, I contend, is highly problematic because besides promoting intense schism, it inevitably leads disciplines to operate within a hierarchical, competitive structure in which they are far from equal. For example, Liu's critique mirrors the persistent dichotomy of science vs humanities: due to the lack of cultural criticism – typical of the sciences but not of the humanities – DH is not humanities at all. DH may be instrumental to the humanities (i.e., the humanities is superior to DH but inferior to the sciences), but it is reduced to a servant role. (pp. 24-25)

Viola conclude il primo capitolo sottolineando la necessità di un approccio autenticamente ‘rivoluzionario’, che vada oltre la semplice integrazione del digitale a supporto delle pratiche tradizionali. Per l’A., è fondamentale riconoscere che la produzione del sapere avviene ormai integralmente *nel* digitale e non *con* il digitale. Questo mutamento di paradigma implica l’abbandono di una logica competitiva e gerarchica tra discipline a favore di un modello di conoscenza fluido, fondato sulla collaborazione paritetica tra saperi e competenze.

Nel secondo capitolo, significativamente intitolato *The Importance of Being Digital*, l’A. riflette sulle implicazioni della produzione e della manipolazione di oggetti digitali nel contesto delle discipline umanistiche. A partire da questa riflessione, propone una nuova cornice metodologica incentrata sul concetto di *post-authentic digital object*, ossia un oggetto che prescinde dai concetti tradizionali di autenticità e copia, non essendo né l’una né l’altra. Si tratta di una costruzione dinamica e instabile, frutto di processi di mediazione multipla, in cui intervengono elementi umani (ricercatori, utenti), tecnici (software e hardware) e storici (politiche istituzionali e contesti di produzione e ricezione). Ne deriva l’inadeguatezza di ogni approccio che si limiti a considerare la digitalizzazione come un mero processo di ‘semplice’ trasposizione da un supporto all’altro. L’A. insiste sulla necessità di riconoscere la natura trasformativa dell’intervento digitale, che modifica tanto l’oggetto quanto il contesto epistemico in cui si inserisce. A supporto di ciò, introduce due categorie mutuete dalla biologia, che trasferisce efficacemente nel campo delle DH per descrivere i rapporti interdipendenti tra uomo e tecnologia nel *framework* epistemico: *symbiosis* e *mutualism*.

Symbiosis and mutualism embed in themselves the principle of knowledge as fluid and inseparable, in which areas of knowledge do not compete against each other, but benefit from a mutually compensating relationship. When asking ourselves the questions ‘How do we produce knowledge today?’, ‘How do we want our next generation of students to be trained?’, the concepts of symbiosis and mutualism may guide the new reconfiguration of our understanding of knowledge in the digital. (p. 32)

Applicate al dibattito sul rapporto tra scienze umane e tecnologie digitali, queste categorie contribuiscono a superare la tradizionale gerarchia disciplinare, delineando un modello cooperativo e non più subordinato, in cui umanisti e tecnologi co-producono l’oggetto conoscitivo. I casi

studio a supporto dell'efficacia dell'approccio *post-authentic* sono rappresentati da DeXTER-DeepteXTminER4 e ChroniclItaly 3.0. DeXTER è un flusso di lavoro che implementa tecniche di deep learning per aumentare la contestualità del materiale testuale digitale e ChroniclItaly 3.0 è una collezione di giornali italoamericani pubblicati negli Stati Uniti tra il 1898 e il 1936. In particolare, per quest'ultimo, ogni fase – dalla selezione dei testi alla loro annotazione – conferisce al *corpus* lo statuto di *artefatto critico*, ovvero un prodotto modellato da scelte intenzionali e da mediazioni culturali, piuttosto che un mero contenitore imparziale di dati.

Nel terzo capitolo (*The Opposite of Unsupervised*), l'A. si concentra sull'analisi di tre funzioni computazionali: il *Named Entity Recognition* (NER), la geolocalizzazione e la *sentiment analysis*. Ciascuna di esse viene discussa non solo nei suoi aspetti tecnici, ma soprattutto alla luce delle implicazioni metodologiche e interpretative che emergono dalla rispettiva applicazione a testi storici e linguisticamente complessi, come quelli che compongono il corpus *ChroniclItaly 3.0*. Attraverso il caso studio, l'A. mostra come operazioni apparentemente “neutrali” – la normalizzazione ortografica, l'identificazione delle entità nominate e l'attribuzione di emozioni – siano in realtà profondamente influenzate dal punto di vista epistemologico. Esse riflettono modelli culturali, linguistici e ideologici spesso non adeguati al contesto storico del testo analizzato.

The creation and enrichment of ChroniclItaly 3.0 show how the conjuncture of the implicated structural forces and factors cannot be envisioned as a network of linear relations and as such, cannot be predicted. The acknowledgement of the limitations and biases of specific tools and choices adopted in the curation of ChroniclItaly 3.0 takes the form of a thorough documentation of the steps and actions undertaken during the process of creation of the digital object. (p. 59)

Il quarto capitolo del volume (*How Discrete*) affronta le implicazioni teorico-metodologiche legate all'impiego del *topic modelling* nei progetti di Digital Humanities. L'A. decostruisce la fiducia, spesso ingenua, riposta nei modelli tematici probabilistici, evidenziandone i limiti strutturali e interpretativi. Al centro di tale critica vi è la constatazione che l'efficacia computazionale del *topic modelling* si fonda su un'operazione preliminare concettualmente problematica: la scomposizione del testo in unità linguistiche discrete, ovvero la trasformazione del testo continuo e semanticamente denso in un insieme di unità indipendenti, statisticamente trattabili, ma private della loro profondità contestuale. L'A. puntualizza:

The problem with topic modelling is that it provides a probabilistic representation of words' distributions in the ingested documents, but it is completely agnostic of the underlying continuous structure of such documents, such as the ambiguity of words' use in each document and across texts as well as the documents' coherent substructure, let alone their wider historical, social and cultural entrenchment. (pp. 92-93)

Ciò implica che elementi cruciali come l'ambiguità, la coesione argomentativa e l'intenzionalità discorsiva vengano sistematicamente elisi a favore di configurazioni semplificate e, talvolta, fuorvianti. Tale distorsione si manifesta con particolare evidenza nel trattamento di testi storici e plurilingue come quelli contenuti in *ChronicleItaly 3.0*, in cui convergono idiomi differenti, registri ibridi e retoriche diasporiche. Lungi dal proporre un rifiuto del *topic modelling*, l'A. ne promuove un uso consapevole e metodologicamente fondato, che riconosca i vincoli epistemici dei modelli probabilistici e ne valorizzi il potenziale solo all'interno di un quadro critico esplicito. In questa prospettiva, il capitolo si inserisce coerentemente nel più ampio framework *post-authentic*, richiamando la necessità di interfacce critiche interpretative capaci di restituire, anziché annullare, la complessità e l'instabilità semantica dei dati.

Il quinto capitolo, *What the Graph*, si concentra sull'analisi critica delle pratiche di visualizzazione dei dati. L'A. esamina in dettaglio come le scelte grafiche nella rappresentazione dei dati – siano esse mappe, reti o distribuzioni tematiche – costituiscano vere e proprie operazioni interpretative, portatrici di specifiche visioni epistemologiche e narrative. In linea con la prospettiva *post-authentic*, la visualizzazione viene intesa non come uno strumento di mera esposizione, bensì come un atto critico, che stabilisce quali aspetti del dato debbano essere resi visibili, accessibili o leggibili, e quali, invece, nascosti o omessi. Le interfacce grafiche, dunque, assumono il ruolo di dispositivi di mediazione e selezione, in grado di orientare attivamente la ricezione e l'interpretazione dei contenuti. L'A. ribadisce ancora una volta la necessità di un approccio critico e consapevole alla progettazione e alla performatività dei dispositivi visuali, fondata su principi di trasparenza, collaborazione interdisciplinare e responsabilità condivisa.

Nel capitolo conclusivo, l'A. riannoda i fili concettuali e metodologici tracciati nei capitoli precedenti. La riflessione finale sottolinea la consapevolezza che ogni oggetto computazionale è frutto di scelte culturali, tecniche, politiche e infrastrutturali ma insiste soprattutto sul fatto che tali scelte non sono mai neutre, bensì implicano responsabilità epistemiche

specifiche, di cui deve farsi carico chi progetta, sviluppa, interpreta e difonde contenuti digitali. In questo scenario, il ruolo delle istituzioni accademiche è centrale. L'A. si rivolge direttamente a queste ultime scrivendo:

Universities also have a responsibility to train in ethical digital management the next generation of thinkers, scholars and academics as well as of digital citizens at large. Equally, research funding agencies must specifically require that the issue of digital ethics is explicitly addressed by researchers in their projects, for instance, by demanding a digital critical component in their proposals. (p. 145)

In ultima analisi, la prospettiva *post-authentic* si configura allora come una vera e propria etica della ricerca digitale, fondata su trasparenza metodologica, simmetria epistemica tra saperi e valorizzazione del ruolo attivo dell'interprete. Le *humanities*, conclude l'A., non devono più assumere una postura difensiva nei confronti della tecnologia, ma al contrario rivendicare un ruolo guida nella progettazione critica delle infrastrutture conoscitive del presente. Il volume di L. Viola rappresenta un contributo di rilievo per le istituzioni accademiche impegnate nella definizione di pratiche digitali eticamente responsabili e metodologicamente fondate. La prospettiva *post-authentic* fornisce un quadro teorico solido per affrontare le sfide poste dalla mediazione computazionale nella produzione del sapere e si configura inoltre come una cornice operativa per orientare la progettazione, la documentazione e la valutazione della ricerca digitale, promuovendo trasparenza, corresponsabilità e collaborazione. La lettura di questo saggio rappresenta uno stimolo a superare dicotomie ormai anacronistiche e poco efficaci, e utili per indirizzare l'orientamento culturale e metodologico delle Digital Humanities in un contesto accademico contemporaneo consapevole e realmente interdisciplinare.

DENISE TIRONE

Università degli studi di Napoli L'Orientale
d.tirone@studenti.unior.it



10.69118/2025_L12

JULIANNE NYHAN, GEOFFREY ROCKWELL, STÉFAN SINCLAIR, ALEXANDRA ORTOLJA-BAIRD (a cura di), *On Making in the digital humanities: the scholarship of digital humanities development in honour of John Bradley*, London, UCL Press, 2023, XX + 279 pp.

Nel panorama contemporaneo delle *digital humanities* (o DH), il concetto di *making* ha assunto una centralità crescente, cui si attribuisce, oltre alla dimensione operativa dello sviluppo tecnico, una forma epistemica autonoma. *On Making in the digital humanities*, curato da Julianne Nyhan, Geoffrey Rockwell, Stéfan Sinclair e Alexandra Ortolja-Baird, si inserisce in questo dibattito, proponendo un'articolata esplorazione del 'fare' nelle DH attraverso una struttura corale e multidimensionale. Si tratta di un volume che mira a ridefinire il campo delle DH partendo dal basso, dalle pratiche, dagli strumenti e dalle persone che le animano. Il volume si apre con una sezione dedicata alle biografie dei *Contributors* (pp. VII-X), segue una *List of figures and table* (XI-XII). Di particolare rilievo è *Introduction: On making in the digital humanities* (pp. 1-13) dei curatori Nyhan e Rockwell, i quali enunciano sei assunti fondanti che guideranno l'intero volume: il primato del progetto come unità epistemica, la centralità delle persone, la riflessività del 'fare', la collaborazione interdisciplinare, la manutenzione come prassi intellettuale e l'importanza dell'infrastruttura. A questi principi si collega organicamente il primo capitolo, *Four Corners of the Big Tent* (pp. 14-47), firmato da John Bradley, figura chiave nella storia delle DH. Assumendo un tono insieme autobiografico e tecnico-scientifico, Bradley propone una riflessione stratificata su quattro ruoli da lui ricoperti nell'arco della sua carriera (i quattro angoli della tenda, per l'appunto): lo sviluppatore, il modellatore concettuale, il collaboratore interdisciplinare e il pensatore critico del 'fare'. Attraverso il racconto della propria esperienza nei progetti TACT (*Text Analysis Computing Tools*), PASE (*Prosopography of Anglo-Saxon England*) e PBW (*Prosopography of the Byzantine World*), Bradley mette in discussione la distinzione tra teoria e tecnica, rivendicando il valore del *maker* come co-autore del sapere. In questa prospettiva, chiarisce che un «tool is, so to speak, an objectified idea [...] It is thought in action» e che persino un software come il suo COGS (*Concordance Generating System*), «with all its limitations, could be considered at least a product of thought – an 'objectified idea' – about some scholarly concerns» (p.

18). L'intreccio tra teoria e prassi, tra soggetto e infrastruttura, si sviluppa pienamente nella prima delle tre macroaree tematiche in cui è diviso il volume, *Making Projects* (pp. 49-126), e prosegue nelle successive *People Making* (pp.127-176) e *Making Praxis* (pp. 177-276), ciascuna delle quali approfondisce un diverso aspetto del “fare” nelle DH: rispettivamente il progetto come forma epistemica, la soggettività come nodo relazionale e politico, la prassi come generatrice intrinseca di teoria. Il volume, nel suo insieme, può essere considerato una riflessione critica sulle forme di sapere che emergono nel momento stesso in cui si costruisce, si collabora, si modella e si riflette. Chiude il volume una quarta sezione *In Memoriam* (pp. 277-287) emblematicamente collocata come conclusione dell’opera. La prima parte, *Making Projects*, si apre con una riflessione condivisa e profondamente articolata sul valore epistemologico dei progetti nelle Digital Humanities. I quattro saggi che la compongono – *Prosopography meets the digital: PBW and PASE* (pp. 51-65) , *Braving the new world: REED at the digital crossroads* (pp. 66-77), *Sustainability and modelling at King’s Digital Lab: between tradition and innovation* (pp. 78-104) , *People of Medieval Scotland database as history* (pp. 105-126) – convergono su una visione ormai matura della progettualità digitale secondo cui il progetto DH non è un semplice contenitore operativo né un supporto alla ricerca, ma si configura come ambiente epistemico, laboratorio ermeneutico e dispositivo teorico. I casi studio presentati dimostrano come ogni fase della progettazione digitale implichi decisioni concettuali complesse che trasformano le modalità tradizionali di produrre, organizzare e interrogare il sapere umanistico. Il saggio inaugurale, *Prosopography Meets the Digital*, firmato da Averil Cameron, Charlotte Roueché e Janet L. Nelson, affronta l’evoluzione della prosopografia da disciplina storica e tradizionale a metodologia computazionale.

The researchers worked systematically through each source, recording all statements about individuals, and attaching such statements to person records. The information for an individual is not reconciled, as in an article-based prosopography, and in the PmbZ: though some sources are clearly more accurate than others, all available testimony is recorded. The PBE data set is a guide to what is said in the sources; it has not set itself the task of source criticism, establishing which sources are more ‘valuable’, ‘accurate’ or ‘true’. This was a radical new approach, made possible by working in a digital environment. (p. 59)

I progetti PASE (*Prosopography of Anglo-Saxon England*) e PBW (*Prosopography of the Byzantine World*) dimostrano come la digitalizzazione delle fonti comporti una riconcettualizzazione delle stesse. Tali processi superano la dimensione strumentale della narrazione storica, assumendo il ruolo di dispositivi in grado di generare nuovi quesiti. Centrale in entrambi è l'adozione del modello *factoid-based*, sviluppato da John Bradley, che consente di trattare ogni affermazione testuale come un'unità minima e modulare, permettendo analisi flessibili, trasversali e ricombinabili. Il *factoid* non solo riflette un'epistemologia della frammentazione strutturata, ma impone anche una profonda riflessione sulla rappresentazione della conoscenza e sulle modalità del suo attraversamento. Questa logica ritorna con forza nel secondo saggio, *Braving the New World: REED at the Digital Crossroads*, in cui Sally-Beth MacLean documenta la trasformazione del progetto REED (*Records of Early English Drama*) in un'infrastruttura digitale. L'integrazione tra archivio documentario, sistema prosopografico e piattaforme interattive evidenzia come l'ambiente computazionale ridefinisca l'oggetto stesso della ricerca – il teatro popolare premoderno – analizzato non più attraverso singole fonti, ma come rete di relazioni multifattoriali. Il contributo di Bradley ha permesso lo sviluppo di una piattaforma capace di rappresentare l'intreccio tra eventi, attori, spazi e strutture di potere, trasformando REED Online da risorsa di consultazione a un ambiente generativo e di interpretazione. Un'altra prospettiva sul valore teorico del progetto digitale è proposta da Arianna Ciula e James Smithies in *Sustainability and Modelling at King's Digital Lab*. L'attenzione si sposta qui sull'infrastruttura, sulla manutenzione e sulla modellazione come pratiche intellettuali. L'approccio del King's Digital Lab alla sostenibilità oltrepassa la conservazione passiva, e va inteso piuttosto come esercizio critico di ri-progettazione continua. Il saggio conclusivo della sezione, *The People of Medieval Scotland Database as History*, a firma di Dauvit Broun e Joanna Tucker, rappresenta l'esito più compiuto dell'intersezione tra database e storiografia. PoMS (*People of Medieval Scotland*) è qui presentato non come strumento di supporto, ma come forma stessa della narrazione storica. Il *factoid* consente di articolare una rete complessa di relazioni, attraverso cui si ricostruisce una 'biografia collettiva' della Scozia medievale non secondo la linearità narrativa tradizionale, ma mediante un'esplorazione relazionale e modulare dei dati.

[...] this is based on a conception of 'infrastructure' that moves beyond material technical necessities [...] towards one that acknowledges the centrality of people, funding, ethics, technology strategy,

software engineering method, and data management to the long term health of our research infrastructures. (p. 85)

In questo contesto, il database si configura come uno spazio ermeneutico aperto, in cui lo storico assume simultaneamente i ruoli di modellatore concettuale, architetto dei percorsi interpretativi e agente critico della rappresentazione storica. Oltre a presentare esempi di *making*, la prima sezione del volume contribuisce a definire una teoria delle DH fondata sul progetto come pratica in cui il codice, la modellazione e la collaborazione sono strumenti critici. Il concetto di *factoid* emerge come modello paradigmatico di una conoscenza storica modulare e interrogabile. Parallelamente, la figura del collaboratore tecnico si afferma come attore intellettuale centrale, capace di superare le tradizionali e fuorvianti separazioni tra teoria e prassi. La sezione *People Making* interroga in modo radicale la soggettività che istruisce i progetti delle Digital Humanities. Se nella prima parte del volume il ‘progetto’ è apparso come motore concettuale e infrastrutturale della conoscenza, in questa seconda sezione è invece la figura del *maker* – inteso al di là della sua funzione di sviluppatore tecnico, come soggetto incarnato, posizionato e agente – a divenire il fulcro dell’epistemologia digitale. I tre saggi che compongono la sezione – *The history of the ‘techie’ in the history of digital humanities* (pp. 129-147), *Jobs, roles and tools in digital humanities* (pp. 148-161), *The politics of digital repatriation and its relationship to Rongowhakaata cultural data sovereignty* (pp. 162-176) – pur profondamente differenti per oggetto, approccio e scala analitica, condividono una comune tensione critica: restituire visibilità e legittimità ai soggetti, ai ruoli e alle lotte che ne costituiscono la trama operativa e politica. Nel suo contributo, *The History of the ‘Techie’ in the History of digital humanities*, Julianne Nyhan smonta la dicotomia tra ‘intellettuale’ e ‘tecnico’. Infatti scrive:

Digital humanities often categorises itself as an interdisciplinary and collaborative field that has been built by a wide range of actors, including technical experts, information professionals, curators, members of the general public and academics... Nevertheless, the contributions of some individuals, like the technical experts referred to in the quote from Bradley above, have tended to be overlooked and have generally been held in lower esteem than the contributions of academics. (p. 129)

Il saggio rappresenta un intervento critico di natura storico ricostruttiva volto a comprendere e problematizzare la figura del *techie* – il tecnico, il programmatore, lo sviluppatore – all'interno della genealogia delle Digital Humanities. Attraverso un approccio metodologico che coniuga oralità e ricerca d'archivio, l'autrice mette in discussione la narrazione dominante della disciplina, che ha spesso collocato i contributi tecnici ai margini del discorso teorico, relegandoli a ruoli ancillari, e invisibili. La dicotomia tra pensiero umanistico e implementazione tecnica, ancora largamente operante nel sistema accademico e nelle pratiche di attribuzione del merito, viene letta come un costrutto ideologico che riflette e riproduce gerarchie di potere culturale e disciplinare. Nyhan fa dunque emergere una contro-narrazione delle DH, centrata su soggetti la cui *agency* è stata sistematicamente marginalizzata, ma il cui lavoro ha strutturato, in modo determinante, l'ontologia computazionale del sapere umanistico. A questa prospettiva fa eco Julia Flanders, che in *Jobs, Roles and Tools in digital humanities* propone un'analisi particolareggiata delle pratiche lavorative all'interno dei progetti DH, distinguendo con precisione tra impieghi formali (*jobs*), funzioni effettivamente svolte (*roles*) e strumenti operativi (*tools*). Flanders mostra come l'organizzazione del lavoro sia caratterizzata da una costante tensione tra aspettative istituzionali e necessità operative: figure formalmente 'tecniche' assumono ruoli strategici e concettuali, mentre studiosi umanisti si trovano a svolgere compiti progettuali o gestionali. Come esemplifica Flanders:

The position of 'librarian' possesses relevant subject expertise [...] but is not typically accorded the agency to participate in project leadership [...] The 'programmer' position is designed around technical expertise but has no provision for subject expertise. The 'faculty' job has subject knowledge but typically little or no project management expertise. (p. 149)

Questa fluidità dei ruoli, lungi dall'essere un'anomalia, viene presentata come cifra distintiva – ma anche criticamente ambigua – della prassi delle DH. Flanders invita a riconoscere tale complessità come elemento costitutivo della produzione di conoscenza digitale, suggerendo la necessità di un ripensamento radicale dei criteri di valutazione, merito e riconoscimento accademico. Il terzo saggio, firmato da Arapata Hakiwai, Karl Johnstone e Brinker Ferguson, allarga lo sguardo oltre l'accademia, introducendo una prospettiva decoloniale che interroga i fondamenti etici e politici del *making* nelle DH. In *The Politics of Digital Repatria-*

tion and Its Relationship to Rongowhakaata Cultural Data Sovereignty, gli autori raccontano il caso della restituzione digitale della *whareniui Te Hau-ki-Tūranga* alla comunità *Rongowhakaata* della Nuova Zelanda, ponendo al centro il concetto di *cultural data sovereignty*. Gli autori pongono al centro dell'analisi la necessità di riconoscere che ogni atto di digitalizzazione non è mai neutrale: l'atto del 'fare' – realizzare un modello 3D, archiviare immagini, strutturare un database – diventa un'operazione carica di significato politico, capace di riaffermare, oppure negare, la sovranità di una comunità sul proprio patrimonio. Infatti,

while indigenous peoples have long claimed sovereign status over their lands and territories, debates around 'data sovereignty' and especially 'cultural data sovereignty' have largely been missing... Indigenous cultural data sovereignty calls into question the multifaceted legal and ethical dimensions of data storage, ownership, access and consent to intellectual property rights and practical considerations regarding how data are used in the context of research, policy and practice. (p. 175)

La restituzione digitale, in questa prospettiva, è possibile solo se accompagnata da processi di consultazione, consenso informato e co-progettazione, che riconoscano le epistemologie e i protocolli delle comunità coinvolte. Il contributo amplia in modo decisivo l'orizzonte delle Digital Humanities, richiamando alla necessità di assumere una postura etica e politicamente consapevole nei confronti delle tecnologie digitali applicate alla memoria e all'identità. Il digitale, lontano dall'essere un dominio disimpegnato, si configura come un territorio di potere, nel quale si giocano oggi partite decisive di giustizia epistemica e riparazione storica. Da questi tre contributi emerge una visione del 'fare' – nel senso più ampio e profondo – che si configura qui come atto etico e politico: non semplice costruzione di strumenti, ma costruzione di relazioni, responsabilità e memorie. La sezione *Making Praxis* rappresenta il momento di più alta sofisticazione teorica dell'intero volume. I quattro saggi che la compongono – *Towards an operational approach to computational text analysis* (pp. 179-212), *From TACT to CATMA; or, a mindful approach to text annotation and analysis* (pp. 213-250), *Pursuing a combinatorial habit of mind and machine* (pp. 251-266), *Historians, texts and factoids* (pp. 267-274) – delineano un quadro plurale ma fortemente coerente, in cui il fare nelle DH si emancipa definitivamente dalla sua immagine strumentale e viene riconosciuto come atto conoscitivo a pieno titolo. Dino Buzzetti,

Jan Christoph Meister, Willard McCarty e Manfred Thaller propongono, ciascuno da prospettive diverse, una concezione delle DH come *epistemologia incarnata*: una disciplina che elabora teoria attraverso il codice, la manipolazione dei dati, l'annotazione dei testi, la modellazione di database e l'interazione con la macchina. Nel saggio *Towards an Operational Approach to Computational Text Analysis*, Dino Buzzetti propone una concezione innovativa dell'analisi testuale computazionale, fondata su un approccio che rifiuta l'intervento del digitale come irrilevante. Ogni trasformazione del testo in ambiente computazionale è interpretata come un atto mediato. L'analisi computazionale si configura così come un processo ricorsivo, fatto di lettura, astrazione e ritorno al testo, che mette in gioco una dinamica riflessiva tra soggetto umano e macchina.

Computational practice is indeed one essential element of the digital humanities, a domain that can be thought of as a meta discipline comprising an operational component, consisting in the application of computational methods, and a theoretical component, resulting from an epistemological reflection on the adequacy of the methods applied in a given context to achieve relevant results. (p. 180)

Elemento centrale della sua argomentazione è la distinzione tra testo e operazione: ciò che la macchina elabora non è più il testo originario, ma una sua riformulazione modellata attraverso strutture formali e categorizzazioni. Questa trasformazione implica una ridefinizione del rapporto con il testo, aprendo nuove possibilità di indagine, ma anche introducendo vincoli interpretativi e assunti teorici spesso impliciti. Per Buzzetti, diventa quindi cruciale rendere espliciti e interrogabili i presupposti che guidano la costruzione dei modelli computazionali che non sono mai innocui, ma veicolano precise concezioni del significato, del testo e della conoscenza. Un analogo spostamento dal tecnico al teorico si riscontra nel saggio di Jan Christoph Meister, *From TACT to CATMA; or, a mindful approach to text annotation and analysis*, che ripercorre l'evoluzione degli strumenti di annotazione e analisi testuale attraverso il confronto tra due ambienti chiave: quello del TACT (*Text Analysis Computing Tools*), sviluppato negli anni Ottanta, e la piattaforma CATMA (*Computer Assisted Text Markup and Analysis*), da lui stesso co-progettata. Questa genealogia strumentale diventa per l'autore l'occasione per una riflessione più ampia sulla natura epistemica dell'annotazione computazionale, intesa come atto interpretativo consapevole, fondato sulla pluralità e sulla negoziazione del significato. In questo quadro, Meister introduce il concetto

di markup ermeneutico, ovvero una forma di annotazione che non si propone di rappresentare oggettivamente la struttura del testo, ma di esplicitare le operazioni interpretative compiute dal lettore. Questo approccio consente la coesistenza di annotazioni multiple, sovrapposte e anche contraddittorie, rispettando così la complessità dell'interpretazione letteraria e la soggettività del processo esegetico. L'autore pone inoltre particolare enfasi sulla valenza pedagogica della piattaforma CATMA, progettata per incoraggiare la riflessività degli utenti e promuovere una concezione della lettura come pratica interpretativa aperta, anziché come esercizio normativo. Inoltre, scrive:

Inspired by TACT and my own fledgling attempts at modelling narrated events in prolog in terms of what I might nowadays call 'latent actions' I argued that we should rather strive for a consensus *qua machina*: digital approaches in the humanities presented us with a unique opportunity to make explicit and transparent, via annotation as well as formal modelling, many of the premises and assumptions that the traditional humanities had hitherto been able to avoid addressing. (p. 242)

Se liberata da vincoli ontologici rigidi, sostiene Meister, l'annotazione testuale può diventare una pratica formativa ad alta intensità teorica, capace di integrare metodo, tecnologia e pensiero critico in un'unica esperienza ermeneutica. Willard McCarty, nel suo contributo intellettualmente più meditativo, propone l'idea di una abitudine combinatoria che unisce mente e macchina. Il cuore del saggio ospita un'argomentazione a favore di una computazione 'critica' e 'immaginativa', in cui l'atto di programmare, annotare o modellare lungi dall'essere ridotto a gesto tecnico, è riconosciuto come parte integrante del processo interpretativo. La macchina, per McCarty, ha valore non in quanto simula la mente umana, ma perché opera secondo una logica differente, che costringe il pensiero a ripensarsi, a ridefinire le proprie categorie e ad affrontare le ambiguità e i paradossi del significato. A tal proposito, McCarty invita a un'ampia esplorazione delle strategie di pensiero combinatorio:

My third, last and most adventurous suggestion is to seek analogical help from uses of combinatorial enquiry wherever it is to be found: close at hand culturally, in the natural sciences and mathematics, and further afield, from other people, in other times, places and cultures. What have people done with counting and sorting when, unassisted by computers, they were at a loss to proceed? (p. 264)

A chiudere la sezione, il saggio di Manfred Thaller affronta il tema della modellazione della conoscenza storica attraverso la lente della computazione, concentrandosi in particolare sul concetto di *factoid*, già ampiamente discusso in ambito prosopografico. Thaller sottolinea come il database storico non possa essere inteso come un semplice archivio neutro di informazioni, ma debba piuttosto essere considerato una narrazione computazionale, ovvero una forma strutturata di racconto che impone allo storico di prendere posizione – in modo esplicito e consapevole – su una serie di scelte teoriche: quali entità inserire, come definirle, quali relazioni rappresentare e quali trascurare. Ogni struttura dati è, in questo senso, una presa di posizione che determina ciò che può essere detto, collegato, interrogato. Thaller mostra come la progettazione di un database comporti una serie di vincoli e di possibilità che trasformano l'interpretazione storica: il lavoro dello storico diventa anche quello di architetto concettuale, chiamato a mediare tra la complessità della fonte e le esigenze di formalizzazione computazionale. Per questa ragione, per far sì che i *factoid* possano davvero fiorire nella ricerca storica, è essenziale un supporto tecnologico evoluto:

That is, factoids would have to be supported as 'objects' in the full sense of software technology – where an object is defined not as a bundle of data in some format, but as a collection of data, together with methods supporting their processing. In other words: for factoids to flourish among historians, we need a technical implementation which does for them what TACT (Bradley 1989) did to text retrieval/ concordances. (p. 273)

Ancora una volta la digitalizzazione viene presentata come trasposizione, ma anche rielaborazione di dati, in questo caso storiografici, frutto di scelte tutt'altro che meccanizzate. Il volume si chiude con l'ultima sezione, *In memoriam*, che consta di un solo saggio che ne riassume, in forma affettiva e teorica insieme, l'intera traiettoria concettuale. In *If Voyant then Spyril: Remembering Stéfan Sinclair* (pp. 277-287), Geoffrey Rockwell non si limita a commemorare una figura centrale delle DH contemporanee, ma costruisce un testo che è, esso stesso, un atto di *making*: un'opera in cui codice, memoria e riflessione si fondono in un esercizio di scrittura teorica computazionale. La spirale che dà nome al progetto *Spyril*, prende forma nel codice, si articola nella collaborazione, si dispiega nell'interazione tra macchina e mente. Questa immagine – il pensare come spirale, come pratica ricorsiva e aperta – offre una metafora

potente per leggere l'intera raccolta. *On Making in the digital humanities* non è solo un manuale di esperienze, strumenti e pratiche ma invita a ripensare radicalmente il senso del 'fare' nella ricerca umanistica. Il volume costruisce un discorso coerente senza divenire rigido, in cui la progettualità, la soggettività e la prassi emergono come dimensioni intrecciate del lavoro nelle DH. Il codice, l'annotazione, il database, l'interfaccia sono potenziali veicoli di interrogazione, in cui l'infrastruttura è pensata come ambiente epistemico e ogni scelta progettuale come decisione teorica. Isolata nel saggio conclusivo, la figura di Stéfan Sinclair assume un valore emblematico. Sia come pioniere di strumenti innovativi, che come esempio di intellettuale capace di unire immaginazione computazionale, sensibilità umanistica ed etica della collaborazione. Il gesto di Rockwell è l'atto che chiude il cerchio e rilancia, al tempo stesso, l'urgenza di fare delle DH un campo interdisciplinare, soggettivo e trasformativo. A riguardo Rockwell scrive:

Like Spyril notebooks the staircase, as a representation of thinking-through, is braided from these two sources of light: the light of thinking from above and the warmth from below of technologies also used for thinking. (p. 286)

Il volume, in definitiva, si presenta come un invito a esplorare il digitale come spazio critico, relazionale e progettuale, dove *fare* teoria significa costruire architetture di senso, immaginare futuri condivisi. Il *making*, qui, non è mai solo esecuzione: è un modo di concepire e abitare la conoscenza.

DENISE TIRONE
Università degli studi di Napoli L'Orientale
d.tirone@studenti.unior.it