



**DOTTORATO DI RICERCA IN
MEDIUM E MEDIALITÀ**

CICLO XXXVIII

COORDINATORE Prof. Francesco Napoli

NEL LABORATORIO CREATIVO DI ARIOSTO: DALL'INCHIOSTRO ALL'EDIZIONE GENETICA DIGITALE DEI
FRAMMENTI AUTOGRAFI DELL'*ORLANDO FURIOSO* COME *VIRTUAL RESEARCH ENVIRONMENT*

Settore Scientifico Disciplinare L-FIL-LET/13

Dottorando

Dott. Ambra Torregrossa

A handwritten signature in black ink, reading 'Ambra Torregrossa'. Below the signature is a horizontal line.

Tutor

Prof. Dario Mantovani

A handwritten signature in black ink, reading 'Dario Mantovani'. Below the signature is a horizontal line.

A mia madre,
alle donne di grande intelletto che cercano,
attraverso lo studio, di rendere il mondo un luogo
migliore, ma soprattutto a quelle cui non è stato
concesso di farlo.

Le antique donne hanno mirabil cose
fatto ne l'arme e ne le sacre muse;
e di lor opre belle e gloriose
gran lume in tutto il mondo si diffuse.
Arpalice e Camilla son famose,
perché in battaglia erano esperte et use;
Safo e Corinna, perché furon dotte,
splendono illustri, e mai non veggon notte.

Le donne son venute in eccellenza
di ciascun'arte ove hanno posto cura;
e qualunque all'istorie abbia avvertenza,
ne sente ancor la fama non oscura.
Se 'l mondo n'è gran tempo stato senza,
non però sempre il mal influsso dura;
e forse ascosi han lor debiti onori
l'invidia o il non saper degli scrittori.

(*Orlando furioso*, Canto XX, ottave 1-2)

«L'informatica è troppo importante per essere lasciata solo agli uomini!»
(Karen Spärck Jones)

INDICE

INTRODUZIONE

I CAPITOLO

- 1.1 La genesi del poema: l'*editio princeps* del 1516
- 1.2 La storia redazionale del poema: i *Cinque Canti* e l'edizione del 1521.
- 1.3 La storia redazionale del poema: la terza edizione del 1532 e i frammenti autografi
 - 1.3.1 Storia, fisionomia materiale ed ordinamento dei frammenti autografi
 - 1.3.2 I nuovi episodi aggiunti: trama, collocazione nell'impianto del poema e attestazione negli autografi ariosteschi
 - 1.3.3 L'edizione debenedettiana de *I Frammenti autografi dell'Orlando Furioso*: contesto storico, premesse teoriche e innovazioni metodologiche
 - 1.3.4 La ricezione critica e le implicazioni metodologiche dell'edizione debenedettiana

II CAPITOLO

- 2.1 Dall'edizione cartacea all'edizione digitale dei frammenti autografi
- 2.2 Metodologia dell'edizione scientifica digitale: il modello "sorgente-output"
 - 2.2.1 Modellizzazione del processo editoriale
 - 2.2.2 Digitalizzazione delle fonti documentarie come immagini ad alta definizione
 - 2.2.3 Digitalizzazione del contenuto testuale dei documenti in MRF
 - 2.2.4 Codifica del testo secondo lo standard XML-TEI
 - 2.2.5 Progettazione dell'interfaccia utente e pubblicazione dell'edizione digitale

III CAPITOLO

- 3.1 Digitalizzazione e trascrizione automatica dei frammenti autografi
- 3.2 La piattaforma *Transkribus* come caso esemplare di *Virtual Research Environment*
 - 3.2.1 Genesi, modelli di business e di governance della piattaforma
 - 3.2.2 Evoluzione tecnologica della piattaforma: *TranskribusX*, *TranskribusLite* e *Transkribus web app*
 - 3.2.3 Il workflow del riconoscimento del testo manoscritto (HTR) sulla piattaforma *Transkribus*

3.2.4 Dalla marcatura semantica del documento all'addestramento di *field model* personalizzati

3.3 Addestramento di modelli di IA personalizzati per i frammenti autografi

3.3.1 Creazione del campione *ground truth*

3.3.2 Marcatura semantica del documento e addestramento di un *field model* per il riconoscimento automatico della *mise en page* e dei *marginalia* ariosteschi

3.3.3 Addestramento di un modello di HTR personalizzato per i frammenti autografi

IV CAPITOLO

4.1 Prototipo di modello di codifica XML-TEI per l'edizione genetica digitale dei frammenti autografi

4.1.1 Il modulo *TEI Header*

4.1.6 Il modulo *Critical Apparatus*

4.1.2 Il modulo *Manuscript Description*

4.1.3 Il modulo *Facsimile*

4.1.4 I moduli *Default Text Structure* e *Verse*

4.1.5 Il modulo *Representation of Primary Sources*

4.2 Pubblicazione online dell'edizione genetica digitale dei frammenti autografi

4.2.1 EVT 1 (2010–2019)

4.2.2 EVT 2 (2016–2020)

4.2.3 EVT 3 (2021–oggi)

4.3 Disseminazione online dei risultati della ricerca: il sito web dedicato all'edizione digitale

CONCLUSIONE

BIBLIOGRAFIA

ABSTRACT

La presente ricerca si colloca all'intersezione tra filologia d'autore e informatica umanistica, proponendo la realizzazione di un prototipo di edizione critico-genetica digitale dei frammenti autografi dell'*Orlando furioso* di Ludovico Ariosto. L'obiettivo è coniugare il rigore metodologico della tradizione filologica con le potenzialità offerte dalle tecnologie digitali, al fine di restituire non soltanto la forma compiuta del poema, ma anche le tracce materiali e testuali del suo processo di elaborazione. Dalla ricostruzione della storia redazionale e della tradizione autografa del *Furioso* – a partire dal lavoro pionieristico di Santorre Debenedetti fino alle interpretazioni di Gianfranco Contini e Cesare Segre – il progetto approda alla definizione di un modello metodologico integrato, capace di unire la prospettiva genetica della filologia d'autore ai principi dell'open science. In tale quadro, la digitalizzazione, la codifica XML-TEI e l'adozione di standard internazionali (TEI P5, IIIF, METS, FITS) garantiscono interoperabilità, tracciabilità e sostenibilità dei dati digitali, in conformità ai principi FAIR (Findable, Accessible, Interoperable, Reusable).

Un elemento di particolare innovatività è rappresentato dall'applicazione di tecniche di intelligenza artificiale per l'analisi dei manoscritti ariosteschi, condotta attraverso la piattaforma *Transkribus*. Le sperimentazioni nel campo del riconoscimento automatico della scrittura (*Handwritten Text Recognition*, HTR) e della *mise en page* hanno evidenziato elevati livelli di accuratezza, dimostrando la possibilità di integrare modelli di *machine learning* nei flussi di lavoro filologici, senza rinunciare al controllo interpretativo dello studioso. Il modello di codifica XML-TEI sviluppato per l'edizione genetica digitale sarà pubblicato tramite la piattaforma EVT (*Edition Visualization Technology*), che consente di rappresentare in ambiente ipertestuale la complessità stratigrafica dei frammenti e di navigare tra le diverse fasi del processo creativo. Il progetto si configura pertanto come una infrastruttura di ricerca dinamica e collaborativa, aperta a futuri sviluppi e implementazioni. In conclusione, la tesi dimostra come la filologia digitale rappresenti non solo un ambito applicativo, ma un autentico paradigma epistemologico, capace di rinnovare le modalità di analisi, rappresentazione e trasmissione del testo letterario, restituendo la voce di Ariosto nella dimensione processuale e vitale della sua scrittura.

Parole chiave: filologia digitale; edizione genetica; Ludovico Ariosto; informatica umanistica; intelligenza artificiale.

This research is situated at the intersection of authorial philology and digital humanities, and presents the development of a digital critical-genetic prototype edition of the autograph fragments of Ludovico Ariosto's *Orlando furioso*. Its aim is to combine the methodological rigor of traditional philology with the potential of digital technologies, in order to reconstruct not only the final form of the poem, but also the material and textual traces of its creative process. From the reconstruction of the editorial history and the autograph tradition of the *Furioso* – beginning with the pioneering work of Santorre Debenedetti and extending to the interpretations of Gianfranco Contini and Cesare Segre – the project leads to the definition of an integrated methodological model that brings together the genetic perspective of authorial philology and the principles of open science. Within this framework, digitization, XML-TEI encoding, and the adoption of international standards (TEI P5, IIIF, METS, FITS) ensure the interoperability, traceability, and sustainability of digital data, in accordance with the FAIR principles (Findable, Accessible, Interoperable, Reusable).

A key innovative aspect lies in the application of artificial intelligence techniques to the analysis of Ariosto's manuscripts through the *Transkribus* platform. Experiments in *Handwritten Text Recognition* (HTR) and layout analysis have achieved high levels of accuracy, demonstrating the potential for integrating machine learning models into philological workflows without compromising the interpretative control of the scholar. The XML-TEI encoding model designed for the digital genetic edition will be published through the *Edition Visualization Technology* (EVT) platform, which enables a hypertextual representation of the fragments' stratigraphic complexity and facilitates navigation across the various phases of the creative process. The project thus takes the form of a dynamic and collaborative research infrastructure, open to future developments and implementations. In conclusion, this study demonstrates that digital philology constitutes not merely an area of technological application but a genuine epistemological paradigm, capable of renewing the modes of analysis, representation, and transmission of literary texts, while restoring Ariosto's voice within the living and processual dimension of his writing.

Keywords: digital philology; genetic edition; Ludovico Ariosto; digital humanities; artificial intelligence.

INTRODUZIONE

Il presente lavoro di ricerca nasce all'incrocio fra filologia d'autore e informatica umanistica, e si propone di realizzare un prototipo di edizione critico-genetica digitale dei frammenti autografi dell'*Orlando furioso* di Ludovico Ariosto. Tale progetto si colloca nel solco delle più recenti sperimentazioni nel campo delle *digital humanities*, orientate alla creazione di ambienti di ricerca e di edizione capaci di coniugare il rigore della metodologia filologica con le potenzialità offerte dalle tecnologie digitali di analisi, rappresentazione e diffusione del testo. La vicenda redazionale del poema ariostesco rappresenta, per complessità e stratificazione, un banco di prova ideale per esplorare le possibilità di un'edizione digitale di tipo genetico, volta non soltanto a restituire la forma finale del testo, ma anche le tracce materiali e testuali del suo farsi. L'*Orlando furioso*, infatti, non è un'opera statica: la sua storia redazionale, che si snoda dalle prime redazioni manoscritte fino alla terza edizione a stampa del 1532, documenta un processo di continua revisione, ampliamento e riscrittura, nel quale i frammenti autografi costituiscono una testimonianza privilegiata. Essi rivelano non solo le modalità di lavoro dell'autore, ma anche le tensioni interne tra tradizione manoscritta e stampa, tra gesto autoriale e mediazione editoriale, tra la materia viva del testo e la sua fissazione tipografica. La tesi è articolata in quattro capitoli, che corrispondono ad altrettanti livelli di approfondimento, dal piano storico-filologico a quello tecnico-metodologico.

Nel primo capitolo si offre una ricostruzione della genesi del poema e della sua evoluzione redazionale, con particolare attenzione alla tradizione autografa e alla ricezione critica dei frammenti ariosteschi dal Cinquecento ad oggi. Un momento di svolta fondamentale è rappresentato dall'edizione dei *Frammenti autografi dell'Orlando furioso* curata da Santorre Debenedetti, la prima impresa sistematica volta a restituire e ordinare i materiali manoscritti ariosteschi, rivelandone il valore come documenti del processo creativo. Il lavoro di Debenedetti non si limita al recupero e alla trascrizione dei testimoni, ma inaugura un metodo filologico innovativo, attento alla dimensione genetica del testo e alla sua stratificazione temporale. La ristampa dell'edizione con la premessa di Cesare Segre ne rilanciò la portata metodologica,

interpretandola alla luce delle riflessioni più avanzate della filologia del secondo Novecento; mentre la ricezione critica di Gianfranco Contini, che ne riconobbe il valore esemplare, contribuì a fissarne l'importanza nella storia della filologia d'autore italiana, consolidando il nesso tra critica delle varianti e interpretazione del processo creativo.

Il secondo capitolo illustra i principi teorici e pratici che guidano la realizzazione di un'edizione scientifica digitale. Si descrive il paradigma metodologico *sorgente–output*, soffermandosi dapprima sulla preparazione del sorgente – dalle fasi di digitalizzazione delle fonti documentarie, alla loro metadattazione, fino alla trascrizione e alla codifica XML-TEI – e approfondendo poi l'output, ovvero la progettazione dell'interfaccia utente. A questo riguardo, si sottolinea la conformità del progetto alle linee guida del *Piano Nazionale di Digitalizzazione* (PND) e alle disposizioni del *Codice dei beni culturali e del paesaggio* (D.Lgs. 42/2004), che stabiliscono i criteri di digitalizzazione, conservazione e accesso ai beni bibliografici e archivistici, garantendo al tempo stesso la tutela, l'autenticità e la fruibilità dei materiali digitali. Particolare attenzione è riservata all'adozione di protocolli e standard di metadattazione interoperabili, in coerenza con i principi FAIR (*Findable, Accessible, Interoperable, Reusable*), al fine di assicurare la tracciabilità, la riusabilità e la sostenibilità dei dati digitali nel lungo periodo. In tale prospettiva, il progetto si ispira ai principi dell'*open science*, promuovendo l'accesso aperto ai dati, la trasparenza metodologica e la condivisione dei risultati della ricerca in ambienti digitali cooperativi. L'uso di formati aperti, di schemi di descrizione conformi alle specifiche TEI P5, armonizzati con le raccomandazioni dello standard IIF (*International Image Interoperability Framework*), nonché l'impiego combinato di FITS (*File Information Tool Set*) – per l'estrazione e la normalizzazione dei metadati tecnici – e di METS (*Metadata Encoding and Transmission Standard*) – per la strutturazione e la trasmissione dei pacchetti informativi – garantiscono l'interoperabilità tecnica e semantica, la compatibilità e la sostenibilità dei diversi livelli di rappresentazione e conservazione del patrimonio digitale.

Il terzo capitolo introduce il lettore nel campo dell'intelligenza artificiale applicata alla filologia, con particolare riferimento alla piattaforma *Transkribus*, qui analizzata come caso esemplare di *Virtual Research Environment*. Viene esaminato il modello di governance della piattaforma, la sua evoluzione tecnologica e illustrata l'infrastruttura informatica che attualmente ne supporta il funzionamento. Si descrivono i diversi servizi offerti dalla piattaforma, soffermandosi in particolare sulle procedure di riconoscimento automatico della scrittura manoscritta (HTR), di marcatura semantica e di addestramento di modelli personalizzati di intelligenza artificiale per il

riconoscimento della *mise en page* e dei marginalia ariosteschi. Il quarto capitolo presenta il modello di codifica XML-TEI messo a punto per l'edizione genetica digitale, illustrandone l'architettura modulare e le prospettive di pubblicazione online tramite la piattaforma EVT (*Edition Visualization Technology*). Si descrivono le principali componenti informatiche del software *EVT* nelle sue tre versioni e gli elementi che ne fanno, ad oggi, uno degli strumenti più versatili per la pubblicazione digitale di edizioni testuali. Il capitolo si chiude con una riflessione sulla disseminazione dei risultati della ricerca attraverso la presentazione di un sito web dedicato, concepito come spazio di condivisione e collaborazione scientifica. In questa prospettiva, si evidenzia la necessità di definire strategie integrate per la gestione e la salvaguardia della proprietà intellettuale e dei prodotti della ricerca, nel rispetto dei principi dell'*open source* e della *open science*.

Essendo il progetto per sua natura interdisciplinare, la presente tesi è stata redatta con un linguaggio divulgativo e inclusivo, volto a rendere accessibili i concetti e le metodologie di entrambe le aree coinvolte. Si è cercato, da un lato, di fornire agli studiosi di filologia e letteratura gli strumenti essenziali per comprendere i principi e le tecniche dell'edizione digitale; dall'altro, di offrire agli informatici e agli specialisti delle tecnologie del linguaggio una chiara esposizione dei fondamenti teorici e storici della filologia d'autore. L'obiettivo è quello di creare un terreno comune, in cui la precisione terminologica e il rigore scientifico si accompagnino alla chiarezza espositiva e alla volontà di dialogo tra discipline. Il lavoro intende dunque porsi non solo come contributo alla conoscenza del laboratorio creativo di Ludovico Ariosto, ma anche come esperimento metodologico nel campo della filologia digitale, in grado di suggerire nuovi paradigmi per la rappresentazione e la fruizione dei testi letterari in ambiente ipertestuale. In questa prospettiva, l'edizione genetica digitale dei frammenti autografi dell'*Orlando furioso* non è pensata come un punto d'arrivo, ma come un prototipo dinamico, aperto a futuri sviluppi, aggiornamenti e interazioni tra studiosi di ambiti differenti. Essa si propone di restituire, attraverso le risorse del digitale, non solo la voce di Ariosto, ma anche la vitalità del suo processo creativo: un *work in progress* che, a distanza di cinque secoli, continua a interrogarci sul senso della scrittura, della revisione e della trasmissione dei testi.

I

CAPITOLO

1.1 La genesi del poema: l'editio princeps del 1516

La disamina dei frammenti autografi richiede, in via preliminare, una sintetica ricostruzione dell'iter redazionale del poema ariostesco. La composizione del *Furioso* fu un processo lungo e complesso, che possiamo ricostruire nelle sue tappe principali grazie alla ricca documentazione epistolare degli Este e dei Gonzaga – destinatari ideali nonché primi lettori privilegiati dell'opera *in progress* - e dello stesso autore.¹ Un primo indizio significativo proviene da una lettera del 3 febbraio 1507, in cui Isabella d'Este, moglie del marchese di Mantova Francesco II Gonzaga, scrive al fratello Ippolito per esprimere la propria gratitudine verso Ludovico Ariosto, inviato del cardinale. La visita, racconta, le aveva portato grande conforto, poiché l'autore le aveva narrato la trama dell'opera a cui stava lavorando, rendendole più lievi i giorni seguiti al parto.² Da questa testimonianza emergono due aspetti rilevanti. In primo luogo, si evince che l'*Orlando furioso*

¹ Per approfondire l'argomento, cfr.: G. Sangirardi, *Orlando furioso*, in *Ludovico Ariosto*, Firenze, Le Monnier, 2006, pp. 95-202; S. Zatti, *Il lavoro di tutta una vita. L'elaborazione del Furioso e i Cinque Canti*, in L. Ariosto, *Orlando furioso e Cinque Canti*, a cura di R. Ceserani e S. Zatti, Torino, Utet, 1997, vol. I, pp. 27-38; W. Binni, *Il capolavoro*, in *Metodo e poesia di Ludovico Ariosto ed altri studi ariosteschi*, a cura di Rosanna Alhaique Pettinelli, Firenze, La Nuova Italia, 1996, pp. 185-274; A. Casadei, *La tradizione delle opere di Ariosto: Orlando Furioso, Frammenti, Cinque Canti*, in *Storia della letteratura italiana, vol. X: La tradizione dei testi*, a cura di E. Malato e C. Ciociola, Roma, Salerno Editrice, 2001, pp. 827-831; Id., *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993; C. Segre, *Storia interna dell'Orlando furioso*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 29-41; S. Debenedetti e C. Segre, *Nota al testo*, in L. Ariosto, *Orlando furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, Bologna, Commissione per i testi di lingua, 1960, vol. II, pp. 1649-1697.

² La lettera è riprodotta in M. Catalano, *Vita di L. Ariosto ricostruita su nuovi documenti*, Genève, L.S. Olschki, 1930-31, vol. II, pp. 78-79. A questo proposito, si ricorda che il volume di Catalano ha rappresentato uno dei contributi più rilevanti alla biografia ariostesca. L'autore, basandosi su un ampio corpus di materiali d'archivio in larga parte inediti, è riuscito a delineare con grande precisione il profilo umano e intellettuale di Ludovico Ariosto, inserendolo nel vivo del contesto politico, culturale e sociale della Ferrara del primo Cinquecento. La ricerca di Catalano non si è limitata a integrare i dati già noti, ma ha rivisto criticamente molte acquisizioni tradizionali, ha corretto cronologie, ha chiarito rapporti familiari e professionali e ha restituito un'immagine più concreta e documentata della vita del poeta. In tal modo, l'opera è diventata un punto di riferimento imprescindibile per gli studi ariosteschi, poiché ha mostrato come l'attività letteraria dell'autore dell'*Orlando Furioso* sia inscindibile dalle condizioni storiche e biografiche in cui essa è maturata.

doveva già avere una struttura narrativa sufficientemente definita da poter essere esposta. In secondo luogo, l'opera appare fin dall'inizio legata a una funzione sociale di intrattenimento, in linea con la tradizione del romanzo cavalleresco. Ne deriva l'immagine di un poema concepito non solo come progetto letterario, ma anche come dono cortigiano destinato a dilettere i signori estensi. Un'ulteriore conferma, in questo senso, arriva da una nota del 5 luglio 1509, in cui il duca Alfonso chiede al fratello Ippolito rinforzi da Modena, insieme ad alcune aggiunte di Ariosto al poema boiardo.³ È difficile stabilire a quale punto della stesura fosse allora giunto il *Furioso*, ma è chiaro che già si riconosceva in esso un duplice statuto: da un lato, continuazione dell'*Orlando innamorato*; dall'altro, opera encomiastica al servizio della corte estense. Tuttavia, qualche mese più tardi, in una lettera datata 25 dicembre 1509, Ariosto si congratula con Alfonso per la vittoria riportata nella battaglia della Polesella e accenna al “padaglione” di Ruggiero, descritto in alcune ottave destinate ad essere collocate nella parte conclusiva del poema.⁴ Questa testimonianza non consente di stabilire se il poeta avesse già concluso l'opera o se, invece, componesse episodi isolati da inserire successivamente nell'intreccio narrativo. Ad ogni modo, una lettera del 14 luglio 1512 indirizzata al marchese di Mantova mostra che il poema, ormai giunto alle dimensioni di un “libro”, non era ancora compiuto:⁵

Ill^{mo} et ex^{mo} principi et domino, domino meo obser^{mo} Domino Marchioni Mantuae. Mantuae.

Ill^{mo} et ex^{mo} Signor mio. Prima per il Molina e poi per Ieronimo me è stato fatto intendere che vostra ex^{tia} haveria piacere de vedere un mio libro al quale già molti di (continuando la invention del conte Matheo Maria Boiardo) io dedi principio. Io, bono e deditissimo servitore de Vostra Signoria, alla prima richiesta le haverei satisfatto et hauto de gratia che quella si fusse degnata legere le cose mie, s'el libro fusse stato in termine da poterlo mandare in man sua. Ma, oltra ch'el libro non sia limato né fornito anchora, come quello che è grande et ha bisogno de grande opera, è anchora scritto per modo, con infinite chiose e liture e trasportato di qua e de là, che fòra impossibile che altro che io lo legessi: e de questo la Ill^{ma} Signora Marchesana sua consorte me ne pò far fede, alla quale (quando fu a questi giorni) a Ferrara i ne lessi un poco. Ma pur, dispostissimo alli servitij de vostra ex^{tia}, cercarò el più

³ Cfr. Ivi., vol. II, pp. 92-93.

⁴ Cfr. L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 139.

Si ricorda che l'edizione delle *Lettere* a cura di Angelo Stella, insieme alla voce “Ariosto” del progetto *Autografi dei letterati italiani*, curata da Simone Albonico (Roma, Salerno editrice, 2022, vol. III) – che per la prima volta fornisce le segnature dei manoscritti autografi – costituisce la base per un nuovo censimento delle lettere di Ludovico Ariosto, disponibile nell'omonima sezione della piattaforma *Epistulae* a cura di Chiara De Cesare (disponibile al seguente link: <https://epistulae.unil.ch/projects/ariosto>).

⁵ Cfr. Ivi., p. 151.

presto che mi serà possibile de far che ne veda almeno parte: e ne farò transcrivere, cominciando al principio, quelli quinterni che mi pareranno star manco male, e scritti che siano li mandarò a V. Sig^{ria} III^{ma}, alla quale humilmente mi racomando.

Ferrariae, XIV Iulij MDXII.

Deditissimo Servo de V. S.

Ludovico Ariosto

Ariosto scrive al marchese spiegando di aver ricevuto, tramite intermediari, la notizia del suo desiderio di leggere l'opera a cui sta lavorando, presentata ancora una volta come continuazione dell'*Orlando innamorato* di Matteo Maria Boiardo. L'autore dichiara la propria disponibilità, ma precisa che al momento il poema non è rifinito né completo: il testo è voluminoso, richiede ancora molto lavoro ed è pieno di correzioni e annotazioni tali da renderlo leggibile soltanto a lui stesso. A riprova, ricorda che pochi giorni prima ne aveva letto un brano a Isabella d'Este, moglie del marchese, la quale poteva confermarne lo stato ancora provvisorio. Tuttavia, assicurando la sua devozione, Ariosto promette di trascrivere e inviare al marchese almeno alcune parti, quelle che riterrà più accettabili, iniziando dall'inizio dell'opera.⁶ La risposta del marchese esprimeva un desiderio ancora più forte di conoscerla e rivela come attorno al poema si fosse già creata un'attesa crescente. Questo episodio può essere interpretato come una raffinata operazione di "marketing" ante litteram. Ariosto seppe infatti alimentare la curiosità con due strategie complementari: offrendo letture selettive in ambienti prestigiosi, soprattutto nella cerchia estense, e legando costantemente il proprio lavoro al successo consolidato dell'*Orlando innamorato*.

Nel frattempo, per tutelare la propria opera da edizioni abusive e garantirsi i possibili profitti, ma anche per preservarne l'integrità fino ai minimi dettagli grafici, Ariosto iniziò a richiedere privilegi ancora prima che la stampa del *Furioso* fosse intrapresa. La prassi editoriale dell'epoca, infatti, era caratterizzata da scarsa attenzione nella preparazione dei testi, spesso soggetti a modifiche di contenuto o di forma linguistica. In tale contesto il ricorso agli strumenti legislativi disponibili, come il suddetto privilegio librario, rappresentò per l'autore un mezzo indispensabile per salvaguardare al tempo stesso i propri interessi economici e la propria reputazione letteraria.⁷

⁶ I materiali preparatori sui quali Ariosto lavorò per giungere alla prima edizione del *Furioso* non si sono conservati, il che rende complessa la ricostruzione del processo di elaborazione dell'opera. In questo senso risulta particolarmente rilevante il contributo di Marco Dorigatti, che ha indagato la fase genetica del poema pubblicando *Il manoscritto dell'Orlando Furioso (1505-1515)*, in *L'uno e l'altro Ariosto. In corte e nelle delizie* (a cura di G. Venturi, Firenze, Olschki, 2011, pp. 1-44), offrendo nuove prospettive sul laboratorio creativo ariostesco e sul percorso che condusse alla princeps del 1516.

⁷ I privilegi librari sono una fonte fondamentale per la storia del libro, in quanto mostrano i rapporti tra autori e stampatori e il valore attribuito alle opere in un'epoca di grandi trasformazioni dovute alla stampa. In assenza di

Il 20 giugno 1515, su sua richiesta, papa Leone X concesse il privilegio di stampa, redatto dal segretario ai brevi Pietro Bembo. L'autorizzazione parlava genericamente di "libri in volgare e in verso" sulle imprese di "soldati erranti", composti sì per diletto, ma frutto di lungo studio e fatica. La formula non dovette soddisfare Ariosto, che infatti non la incluse nell'edizione a stampa. Molti anni dopo, nel 1535, lo stesso Bembo ripubblicò il documento con una piccola ma significativa modifica: i protagonisti non erano più "soldati erranti", bensì "cavalieri erranti". Un cambiamento minimo in apparenza, che riflette però un riconoscimento più pieno della dignità del genere cavalleresco e, insieme, la stima che Bembo nutriva per l'opera a distanza di vent'anni. La questione si risolse definitivamente il 27 marzo 1516, con un nuovo privilegio pontificio redatto da Jacopo Sadoletto, successore del Bembo. Questa volta l'autorizzazione si riferiva in modo esplicito a un'opera intitolata *Orlando furioso*, composta per diletto ma con grande dedizione e molte veglie. Il passaggio da una definizione generica a un titolo preciso segna un punto decisivo: il poema aveva ormai acquisito un'identità riconosciuta. Resta aperto, però, un interrogativo: se fu Bembo a non voler attribuire al testo una fisionomia compiuta, oppure se fu Ariosto stesso a scegliere il titolo definitivo solo a ridosso della pubblicazione. Alcuni mesi dopo, il 25 ottobre 1515, Ariosto ritenne opportuno rivolgersi anche al Doge di Venezia per ottenere un privilegio analogo: ⁸

[AL DOGE DI VENEZIA. A VENEZIA]

Ill^{mo} e Ser^{mo} Principe e Signore mio obser^{mo}. Supplico alla Sub.^{ità} V.^a io, divoto et affectionatissimo servo suo Ludovico Ariosto nobile ferrarese e familiare del R.^{mo} S.^{re} Car.^{le} Estense, come, havendo cum mie longe vigilie e fatiche, per spasso e recreatione de' S.^{ri} e persone di animi gentilli e madone, composta una opera in la quale si tratta di cose piacevoli e delectabeli de arme e de amore, e desiderando ponerla in luce per solazo e piacere di qualunque vorà e che se delecterà de legerla, et anche cum quello più beneficio e remuneratione delle fatiche mie duratoli più anni in componerla che consequire posso, ho deliberato de farla stampire dove meglio a me parerà. Ma dubitando che qualche altro, a concorrentia della stampa ch'io ne farò, sùbito che tal mia opera e stampa sia fuori, non se

leggi sul diritto d'autore, i privilegi costituivano l'unico strumento legale di tutela: garantivano agli scrittori la possibilità di scegliere l'editore, impedendo ristampe non autorizzate che avrebbero danneggiato prestigio e guadagni. Di conseguenza, molti autori li utilizzarono consapevolmente come protezione economica e reputazionale. Lo studio dei privilegi consente di chiarire le dinamiche commerciali che interessarono Ludovico Ariosto nella produzione, promozione e diffusione dell'*Orlando Furioso*, offrendo nuove prospettive sul percorso editoriale del poema e spunti per ulteriori ricerche.

Per approfondire, cfr. E. Squassina, *La protezione del Furioso: Ariosto e il sistema dei privilegi in Italia*, «Bibliothecae.it», 6(1), 2017, pp. 9–38, consultabile al seguente link: <<https://doi.org/10.6092/issn.2283-9364/7024>> (ultimo accesso 17 ottobre 2025).

⁸ Cfr. Ivi., pp. 157-58.

intrometta a ristampare o farne ristampare una altra, e che non pigli il bene et utile de le fatiche, che doveriano venir a me: pertanto prego e supplico la prefata V.^{ra} Sub.^{tà}, che quella sia contenta per suo decreto e privilegio concedermi de gratia che per tucto el tempo della vita mia non sia licito a persona alcuna, cossì terrera come forestiera e di qualunque grado se voglia esser o sia, che ardisca né presuma in le terre e loci e dominio de V.^{ra} S.^{tà} presumere di stampare né di fare stampare in forma alcuna de letttera né di foglio grande, piccolo, né picolino, nè possa vender o far vender ditta mia opera sença expressa licentia e concessione de mì Ludovico Ariosto auctore di epsa, sotto pena de perder tal opere tucte che si trovassino stampate, o venderse, e de ducati mille per cadauno che presumerà stamparla o farla stampare, o venderla o farla vender: la qual pena per la mità se applichi a cui piacerà alla Sub.^{tà} V.^{ra}, e l'altra mità e libri stampati o venduti a mì Ludovico pre nominato servitore di quella. Cuius gratiae etc.

Die 25 octobris 1515.

Il poeta, dopo aver composto con lunga fatica e veglie un'opera "di armi e d'amore" destinata al diletto dei signori e delle dame, desidera pubblicarla a stampa sia per offrire piacere ai lettori, sia per ottenere un giusto riconoscimento delle sue fatiche. Tuttavia, teme che altri, approfittando della sua edizione, possano ristampare l'opera e trarne profitto a suo danno. Per questo, prega l'autorità veneziana di concedergli un privilegio esclusivo che vieti, per tutta la durata della sua vita, a chiunque – suddito o forestiero – di ristampare o vendere il poema senza la sua autorizzazione. In caso di infrazione, chiede che siano confiscate le copie illegittime e comminata una pena pecuniaria di 1000 ducati, metà da destinare a discrezione del doge e metà a lui stesso.

In una lettera del 27 febbraio 1516, Ariosto informa il duca di Genova Ottaviano Fregoso, che ha iniziato la stampa del suo *Orlando furioso*, opera che il destinatario aveva già manifestato interesse a leggere. Riportiamo di seguito il testo integrale della missiva:⁹

Ill^{mo} et ex^{mo} domino domino meo obs^{mo} Domino Octaviano Fulgosio Duci Ianuae. Ianuae.

Ill^{mo} Sig^{re} mio obs^{mo}. Havendo io cominciato a porre a stampa quel mio libro, el quale già vostra ex^{tia}, nel tempo che in Urbino et in Roma la vidi, per sua humanità dimostrava desiderare di veder fornito, e dubitando io che non si faccia per colpa de li impressori incorretto, ho ottenuto da la S.^{tà} di nostro S. e da la Maestà Chris.^{ma} e da Sig^{ri} Venetiani, e spero de ottenere da altre potentie anchora, che fin ch'io viva a nessuno sia licito, se non a chi parrà a me, di stampirlo né di venderlo. Hor, essendo io deditissimo servitore de vostra ex^{tia} et havendo gran fede in essa, piglio seco questa sicurtà di suplicarla, che essa ancora mi voglia concedere similmente questa gratia. Io non ho appresso a quella persona ch'io conosca, alla quale io possa di questo dar cura; sì che vostra ex^{tia}, per sua solita gentileza, si degnarà di commettere ad alcun suo che lo le ricordi, e che sia mio sollicitatore, e che con più prestezza che può mi mandi expedito in buona forma, o per privilegio o per lettera

⁹ L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, pp. 159-60.

patente o per rescritto o in qual modo parra a vostra ex^{tia}, che a nessuno sia licito, vivendo io, né in Genua né in terra alcuna al dominio di Genua soggetta, stampir né far stampire, né vendere né far vendere questo mio libro ch'io nomino *Orlando furioso* senza mia expressa licentia, sotto pena di perdere tutte l'opere e di pagar per ciascuna quattro o cinque ducati, o più o manco, e da essere applicati secondo che parrà a vostra ex^{tia}, offerendole che di questa gratia io le haverò grandissima e perpetua obligatione; alla quale raccordo ch'io sono affectionatissimo servitore e senza fine me le racomando.

Ferrariae, xxvii februarij MDXVI.

de Vostra Ill^{ma} Sig^{ria}

deditissimo servitore Ludovico Ariosto.

Il poeta teme che gli stampatori possano produrne versioni scorrette e, per proteggersi da ristampe abusive, ha già ottenuto privilegi di stampa dal papa, dal re di Francia e dalla Repubblica di Venezia, e conta di ottenerne da altre potenze. Chiede quindi al duca di Genova di concedergli la stessa protezione: un privilegio che vieti, per tutta la sua vita, a chiunque nei territori genovesi di stampare o vendere il poema senza la sua autorizzazione, sotto pena di confisca delle copie e di una multa (stabilita a discrezione del duca).

La stampa venne affidata a Giovanni Mazzocco dal Bondeno, editore ferrarese che, pur non molto prolifico, aveva dimostrato grande coerenza e qualità nelle sue scelte. Attivo dal 1509, egli aveva pubblicato testi che riflettevano la varietà culturale del primo Cinquecento: dagli scritti medici di Niccolò Leonicensino al *Mambriano* di Francesco Cieco da Ferrara, dalle poesie erotiche di Guarino Veronese alle sentenze greche di Giovanni Crastone, dalla grammatica di Costantino Lascais ai testi religiosi di Savonarola, fino ai proverbi di Erasmo da Rotterdam. Un catalogo che univa scienza, riscoperta dei classici, poesia cavalleresca, spiritualità e cultura umanistica. In questo contesto l'*Orlando Furioso* trovò una collocazione naturale, a testimonianza della piena consonanza culturale tra autore ed editore. La decisione di Ariosto di affidare il suo poema alla tipografia ferrarese del Mazzocco, anziché alle stamperie veneziane – allora leader indiscusse per quantità e qualità della produzione libraria – si spiega con il desiderio di seguire da vicino l'allestimento dell'edizione. La bottega del Mazzocco si trovava infatti in via dei Sabbioni, nella contrada di San Romano, non lontano dalla casa del poeta, il che agevolava i contatti diretti. Mazzocco ebbe un ruolo determinante nella pubblicazione del *Furioso*: contribuì alle spese di produzione, partecipò alla distribuzione e, in quanto libraio e cartolaio del cardinale Ippolito d'Este, poté contare sul sostegno del suo illustre cliente, che si adoperò personalmente per la fornitura della carta. A Ferrara, infatti, la carta era scarsa, costosa e difficile da reperire a causa dei dazi d'importazione. Per ridurre le spese, Ariosto ottenne, tramite Ippolito d'Este, il libero

transito di mille risme provenienti dalla riviera di Salò attraverso il territorio mantovano.¹⁰ Il 22 aprile 1516 il poema uscì in circa 1.300 copie.¹¹ L'accoglienza dell'*Orlando Furioso* fu immediata. Il 7 maggio 1516, da Mantova, il segretario del marchese Ippolito Calandra scrisse a Federico Gonzaga, che in quel momento si trovava in Francia, per informarlo dell'arrivo di Ariosto con una cassa di libri. Si trattava di un volume imponente, quasi pari per estensione all'*Innamoramento* di Boiardo, ma intitolato *Orlando Furioso* e giudicato da Calandra persino più bello del modello. Tre copie furono donate ai membri più illustri della famiglia Gonzaga, mentre le altre vennero destinate alla vendita. Nella stessa lettera Calandra domandava al marchese se desiderasse riceverne un esemplare: in tal caso, egli si sarebbe occupato di acquistarlo, farlo legare e spedirlo al suo signore. L'impresa aveva dunque una doppia finalità: da un lato, l'omaggio cortigiano, che consolidava i rapporti di protezione e radicava l'opera nel contesto estense; dall'altro, la volontà di raggiungere un pubblico più ampio, oltre i confini della corte.¹² La diffusione fu rapida: già nel dicembre 1517 il poema era noto a Firenze, dove Machiavelli, scrivendo a Luigi Alamanni, lo giudicò "bello tutto", pur con il rammarico di non figurare tra le personalità celebrate nell'ultimo canto.¹³ L'edizione del 1516 in 40 canti (siglata

¹⁰ La notizia ci giunge tramite una lettera del 17 settembre 1515, inviata dal cardinale Ippolito al cognato Francesco Gonzaga. Per approfondire, cfr. L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 156.

¹¹ L'edizione si presentava come «Orlando Furioso de Ludovico Ariosto da Ferrara. (In fine.) Impresso a Ferrara per Maestro Giovanni Mazzocco dal Bondeno. adi XXI. de Aprile. M.D. XVI.». Per approfondire, cfr. U. Guidi, *Annali delle edizioni e delle versioni dell'Orlando Furioso e d'altri lavori al poema relativi*, Bologna 1861, pp. 3-5, consultabile al seguente link: <https://ia600203.us.archive.org/9/items/annalidelleedizi00guiduoft/annalidelleedizi00guiduoft.pdf> (ultimo accesso 17 ottobre 2025). Sul verso della seconda carta compare un'incisione raffigurante un alveare, da cui le api fuggono spinte dal fuoco appiccato al di sotto. L'immagine è racchiusa entro una cornice decorata con più gruppi composti da una mazza e una scure, attorno ai quali si avvolge una serpe. Nei quattro angoli della cornice è ripetuto il motto *Pro bono malum*.

Sulla scorta del censimento effettuato da Marco Dorigatti in preparazione dell'edizione critica dell'*Orlando Furioso* secondo la princeps del 1516 (Firenze, Olschki, 2016), sono sopravvissuti circa dodici esemplari di questa edizione, oggi conservati in biblioteche istituzionali quali quelle di Treviso, Ferrara, Rovigo e Firenze. L'edizione del 1516, inoltre, è stata oggetto di un'importante analisi anche da parte di C. Fahy, che in *L'autore in tipografia: le edizioni ferraresi dell'Orlando Furioso* (in *I libri di 'Orlando Innamorato'*, Modena, Panini, 1987) ha messo in luce il ruolo diretto di Ariosto nel processo tipografico, evidenziando le complesse dinamiche tra autore, stampatore e committenza estense.

¹² La lettera è riprodotta in M. Catalano, *Vita di L. Ariosto ricostruita su nuovi documenti*, Genève, L.S. Olschki, 1930-31, vol. II, pp. 157-158.

¹³ Cfr. N. Machiavelli, *Lettere*, a cura di Franco Gaeta, in *Tutte le opere*, vol. III, Milano, Mondadori («I Meridiani»), 1971, pp. 413-415.

A) fu quindi il *Furioso* che i contemporanei conobbero e ammirarono.¹⁴ Non sorprende che, nel 1520, Ariosto scrivesse a Mario Equicola di non sapere più in quale parte d'Italia fossero rimaste copie invendute. Riportiamo di seguito il testo integrale della missiva:¹⁵

Mag^{co} tanquam fratri honorando D. Mario Equicolo Secretario. Mantuae.

Mag^{co} m. Mario mio hn^{mo}. Per m. Gian Iacomo Bardelone ho havuto sei lire di nostra moneta, li quali vostra M^{tia} mi ha rimessi, credo, per parte de li denari che si hanno d'havere dal venditore de li miei *Orlandi* a Verona. Di che ringratio quella, ma mi paron pochi a quelli ch'io aspettava; e non posso credere che quel libraro non li habbia expediti tutti, perché in nessuno altro luogo di Italia non so dove ne restino più da vendere: e se fin qui non gli ha venduti, non credo che più li venda. Per questo seria meglio che il libraro li rimettesse qui, perché subito troverei di expedirli, perché me ne son dimandati ogni dì. Vostra Mag^{tia}, essen<do> risanata come spero che ella sia, la prego che si sfor<z> di saper la cosa, che troverà che li libri son vendut<i> e che> quel libraro vuole rivalersi de quelli denari. La <quale> si ricordi ch'io son suo, e sempre me gli raccomando.

<Ferrariae>, 8 Novembris MDXX.

Vostro Lud. Ariosto.

Ariosto informa Equicola di aver ricevuto tramite un intermediario sei lire, che suppone provenire dalla vendita dell'*Orlando furioso* a Verona. Tuttavia, la somma è molto inferiore a quanto si aspettava e dubita che il libraio non abbia già venduto tutte le copie, poiché non gli risulta che ne siano rimaste invendute altrove in Italia. Se il libraio non le ha vendute finora, difficilmente potrà venderle in futuro; per questo Ariosto suggerisce che i volumi invenduti vengano rimandati a Ferrara, dove potrebbero essere subito collocati, dato che vi è ancora forte richiesta. Il diretto coinvolgimento nelle fasi di produzione e distribuzione mostra come Ariosto avesse piena consapevolezza del proprio ruolo di autore e considerasse imprescindibile esercitare il massimo controllo sulla stampa del *Furioso*. Per oltre un quindicennio, fino alla redazione del 1532, questo testo rappresentò il fondamento della fama ariostesca e fu accolto come un nuovo capolavoro della letteratura volgare.¹⁶

¹⁴ La siglatura tradizionale delle tre edizioni dell'*Orlando furioso* – A (1516), B (1521) e C (1532) – fu introdotta da Santorre Debenedetti nell'articolo *Per la storia dell'Orlando furioso* («Giornale storico della letteratura italiana», 1922) e ripresa nell'edizione critica da lui approntata nel 1928, pubblicata tuttavia priva dell'apparato variantistico, benché già predisposto. Tale apparato fu reso disponibile solo più tardi, con l'edizione curata da Cesare Segre: Ludovico Ariosto, *Orlando Furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, a cura di S. Debenedetti e C. Segre, Bologna, Commissione per i testi di lingua, 1960. Tale nomenclatura, recepita e consolidata dalla critica, è divenuta lo standard di riferimento negli studi ariosteschi.

¹⁵ Cfr. L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 177.

¹⁶ Per approfondire la ricezione del primo *Furioso*, cfr. i seguenti testi: AA.VV., *Orlando furioso: Rezeptionsgeschichte und Interpretationsansätze. Studi sulla ricezione e sull'interpretazione del testo*, a cura di

1.2 La storia redazionale del poema: i Cinque Canti e l'edizione del 1521.

La composizione dell'Orlando furioso non si arrestò con la stampa del poema. In una lettera del 15 ottobre 1519, indirizzata a Mario Equicola, Ariosto scriveva:¹⁷

<Ma>^{g^{co}} et Doctissimo Viro Domino Mario Equicolae mihi hn^{mo} etc. Mantuae.

M. Mario mio hn^{mo}. Io ringratio molto vostra S. de la offerta ch'ella mi fa di prestarmi l'opera sua, accadendomi, ne li miei litigij; la quale accetto di buono animo, e credo di usarla: ma non mi basteria il scrivere quello ch'io dimandassi. Ho pensiero di trasferirmi un giorno a Mantua, et informarvi bene di quel ch'io voglio: ma non è il tempo anchora. [...] È vero ch'io faccio un poco di giunta al mio *Orlando furioso*, cioè io l'ho cominciata; ma poi da l'un lato il Duca, da l'altro il card^{le}, havendomi l'un tolto una possessione che già più di trecent'anni era di casa nostra, l'altro un'altra possessione di valore appresso dece mila ducati, de facto e senza pur citarmi a mostrare le ragion mie, m'hanno messo altra voglia che di pensare a favole. Pur non resto per questo ch'io non segua facendo spesso qualche cosetta; s'io seguirò, mon mi uscirà di mente di fare il debito mio; e tanto meglio che non ho fatto pel passato, quanto questo debito da quel tempo in qua è cresciuto in infinito. [...]

Ferrariae, xv octobris 1519.

Vostro Ludovico Ariosto.

Il poeta ringrazia Mario Equicola per avergli offerto il proprio aiuto nei contenziosi legali e rimanda a un futuro incontro a Mantova per spiegargli meglio la situazione. Inoltre, ammette di aver iniziato a scrivere “un poco di giunta” al *Furioso*, ma confessa che le ingiustizie subite dal duca e dal cardinale, che gli hanno sottratto due importanti proprietà di famiglia senza permettergli di difendersi, lo hanno distolto dalla poesia e gli hanno fatto perdere la voglia di scrivere “favole”. Nonostante ciò, continua di tanto in tanto a comporre e promette che, se proseguirà, non dimenticherà il proprio dovere di poeta, anzi lo adempirà con più impegno di quanto non abbia fatto in passato, poiché sente questo dovere ormai accresciuto in misura grandissima.

In questa “giunta” la critica ha solitamente identificato i cosiddetti *Cinque Canti*, rimasti inediti fino a dodici anni dopo la morte dell'autore, quando furono pubblicati dapprima dal figlio Virginio Ariosto presso gli eredi di Aldo Manuzio come appendice al poema in 46 canti (cioè secondo la versione fissata dall'autore nel 1532) e successivamente in edizione autonoma dagli

Christian Rivoletti e Kai Nonnenmacher, *Romanische Studien, Beihefte 102*, München, AVM, 2020; Ludovico Ariosto, *Orlando Furioso secondo la princeps del 1516*, a cura di M. Dorigatti, Firenze, Olschki, 2006.

¹⁷ Cfr. L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere, vol. III: Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 172-73.

eredi di Bernardo Giunti (Firenze, 1546). Non si conservano materiali autografi relativi ai *Cinque Canti*; il testo è noto grazie a tre fonti principali:¹⁸

- l'*editio princeps* sopra menzionata, in cui manca l'inizio del primo canto e svariate ottave (Venezia, 1545);¹⁹
- l'edizione pubblicata da Gabriele Giolito, che sanò parzialmente lacune dell'edizione aldina (Venezia, 1548);²⁰
- il cosiddetto *Codice Taddei*, trascritto da Giulio di Giammaria Ariosto (parente del poeta e copista di fiducia di Virginio Ariosto) e oggi custodito presso la Biblioteca Comunale Ariostea di Ferrara (ms. Cl. I 706).²¹

¹⁸ Una trattazione esaustiva dell'argomento non rientra negli obiettivi del presente lavoro. Per approfondire la tradizione dei *Cinque Canti*, cfr. i seguenti saggi: V. Gritti, *Per l'edizione critica dei 'Cinque Canti' di Ariosto*, «Filologia italiana», 13, 2016, pp. 139-191, consultabile al seguente link: <https://sfera.unife.it/retrieve/e309ade3-2399-3969-e053-3a05fe0a2c94/113922360214.pdf?utm_source=chatgpt.com> (ultimo accesso 17 ottobre 2025); S. Larosa, *Poesia e cronologia nei Cinque Canti. Una nuova ipotesi*, Rende, Centro Editoriale e Librario, Università della Calabria, 2001; A. Casadei, *I 'Cinque Canti' o l'ultima eredità di Boiardo*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 113-128; Id. *Sulla situazione testuale e sulla cronologia dei 'Cinque Canti'*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 179-192; P. Trovato, *Con ogni diligenza corretto. La stampa e le revisioni editoriali dei testi letterari italiani (1470-1570)*, Bologna, Il Mulino, 1991; C. Segre, *Studi sui Cinque Canti*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 121-177.

¹⁹ Nell'edizione si legge: «Orlando Furioso di Messer Lodovico Ariosto, et di piu aggiuntovi in fine piu di cinquecento stanze del medesimo autore non piu vedute. Riveduto & corretto con somma diligenza. In Venegia, del M.D. XLV. (In fine.) In Vinegia, nell'anno M.D. XLV. In casa de' figliuoli d'Aldo». I *Cinque Canti* hanno un titolo separato: «Cinque Canti di un nuovo libro di M. Lodovico Ariosto, i quali seguono la materia del Furioso. Di nuovo mandati in luce. Con privilegio del Sommo Pontefice, et della Illustrissima Signoria di Vinegia, M.D. XXXXV.» Per approfondire, cfr. U. Guidi, *Annali delle edizioni e delle versioni dell'Orlando Furioso e d'altri lavori al poema relativi*, Bologna 1861, pp. 47-49, consultabile al seguente link: <<https://ia600203.us.archive.org/9/items/annalidelleedizi00guiduoft/annalidelleedizi00guiduoft.pdf>> (ultimo accesso 17 ottobre 2025).

²⁰ Nell'edizione si legge: «Orlando Furioso di M. Lodovico Ariosto...Ornato di varie figure, con alcune stanze, e cinque canti d'un nuovo libro del medesimo nuovamente aggiunti, e ricorretti con alcune allegorie, e nel fine una breve tavola di tutto quello, che nell'opera si contiene. In vinegia appresso Gabriel Giolito de Ferrari, M.D. XLVIII, in 4to fig. in legno». Per approfondire, cfr. Ivi., pp. 51- 52.

²¹ Cesare Segre, nel saggio *Studi sui Cinque Canti* (ed. orig. 1954), confluito successivamente nel volume *Esperienze ariostesche* (Pisa, Nistri-Lischi, 1966, p. 121), non fornisce una datazione precisa del manoscritto ma ricorda che esso fu acquisito dalla biblioteca ferrarese agli inizi del Novecento come dono degli eredi di Antonio Taddei. La proposta di datarlo alla metà del XVI sec. è invece di Alberto Casadei, sulla base dell'individuazione delle filigrane, in *Il percorso del Furioso. Ricerche intorno alle redazioni del 1516 e del 1521* (Bologna, il Mulino, 1993, p. 179).

La maggior parte delle edizioni cinquecentesche dell'*Orlando furioso* posteriori al 1545 accolse i *Cinque Canti*, inaugurando un vivace dibattito critico circa la loro destinazione e funzione.²² Una tappa decisiva di tale dibattito fu la pubblicazione, nel 1556, di una nuova edizione del *Furioso* approntata da Girolamo Ruscelli (1518–1566), letterato e poligrafo attivo nella Venezia di metà secolo, noto per la cura di antologie poetiche e per i suoi contributi in materia di lingua e metrica.²³ Nel frontespizio, l'edizione si presentava nei seguenti termini:²⁴

Orlando furioso di m. Lodovico Ariosto, tutto ricorretto, et di nuove figure adornato. Al quale di nuovo sono aggiunte le Annotationi, gli Avvertimenti, & le Dichiarationi di Girolamo Ruscelli, la Vita dell'Autore descritta dal Signor Giovambattista Pigna, gli Scontri de' luoghi mutati dall'Autore doppo la sua prima impressione, La Dichiaratione di tutte le favole, il Vocabolario di tutte le parole oscure, & altre cose utili & necessarie, In Venetia, Appresso Vincenzo Valgrisi, nella bottega d'Erasmus, 1556.

L'iniziativa editoriale ebbe straordinario successo, conoscendo numerose ristampe, arricchite talora di ulteriori apparati esegetici rispetto al nucleo originario. Sul piano tipografico-editoriale,

²² Per approfondire le tappe fondamentali della filologia ariostesca nel Cinquecento – vale a dire i diversi momenti di intervento sul testo dell'*Orlando Furioso* attraverso revisioni, interpretazioni e operazioni editoriali, considerate nel più ampio quadro culturale e librario dell'epoca – cfr. M. Dorigatti, *Momenti della filologia ariostesca nel Cinquecento*, in *La filologia in Italia nel Rinascimento*, a cura di C. Caruso e E. Russo, Roma, Edizioni di Storia e Letteratura, 2018, pp. 193-215; C. Segre, *Petrarca, Ariosto e la critica delle varianti nel Cinquecento*, in *Storia della letteratura italiana*, vol. XI: *La critica letteraria dal Due al Novecento*, a cura di E. Malato e P. Orvieto, Roma, Salerno Editrice, 2003, pp. 353-67.

²³ Per approfondire, cfr. P. Trovato, *Un correttore 'libraro': Ruscelli e compagni (1553-1555)*, in *Con ogni diligenza corretto. La stampa e le revisioni editoriali dei testi letterari italiani (1470-1570)*, Bologna, Il Mulino, 1991, pp. 241-268.

²⁴ Per approfondire, cfr. U. Guidi, *Annali delle edizioni e delle versioni dell'Orlando Furioso e d'altri lavori al poema relativi*, Bologna 1861, pp. 59-61, consultabile al seguente link: <https://ia600203.us.archive.org/9/items/annalidelleedizi00guiduoft/annalidelleedizi00guiduoft.pdf> (ultimo accesso 17 ottobre 2025); Cfr. L. Morini, *Ruscelli e le pretese varianti ariostesche al Furioso del '32*, in *In ricordo di Cesare Angelini. Studi di letteratura e filologia*, a cura di F. Alessio e A. Stella, Milano, Il Saggiatore, 1979, pp. 160-84; P. Trovato, *Filologia e linguistica del Ruscelli (1552-1556)*, in *Con ogni diligenza corretto. La stampa e le revisioni editoriali dei testi letterari italiani (1470-1570)*, Bologna, Il Mulino, 1991, pp. 269-298. Online è possibile trovare diversi esemplari digitalizzati in alta definizione dell'edizione Valgrisi (1556), tra cui possiamo menzionare: l'esemplare conservato dalla Biblioteca Nazionale Centrale di Firenze (consultabile sulla piattaforma del Museo Galileo al seguente link: <https://bibdig.museogalileo.it/tecanew/opera?bid=19740&seq=2>); l'esemplare custodito dalla Bnf - Bibliothèque nationale de France (consultabile sulla piattaforma *Gallica* al seguente link: <https://gallica.bnf.fr/ark:/12148/bd6t5772384v>) e l'esemplare posseduto dalla Duke University Libraries (consultabile sul portale *Internet Archive* al seguente link: <https://archive.org/details/orlandofurios00ario/page/n1/mode/2up>).

Ruscelli sottopose il poema ariostesco a un intervento di regolarizzazione ortografica (il “ricorretto” dichiarato nel frontespizio, cui egli stesso dedica un paragrafo nella lettera *Ai lettori*), corredandolo, oltre che di illustrazioni, di un commento essenziale, limitato ai luoghi più oscuri. Inoltre, subito dopo la dedica ad Alfonso II d’Este, inserì la *Vita dell’Ariosto* composta da Giovan Battista Pigna (1529-1575), già pubblicata due anni prima all’interno dell’opera *I romanzi [...] divisi in tre libri, ne’ quali della poesia, & della vita dell’Ariosto con nuovo modo si tratta, In Vinegia, nella bottega d’Erasmus, appresso Vincenzo Valgrisi, 1554*. Infine, gli *Scontri de’ luoghi mutati dall’Autore doppo la sua prima impressione* costituiscono un’altra sezione dei suddetti *Romanzi*, anche se non risulta chiaro, dal solo frontespizio, a chi si dovesse la loro elaborazione.²⁵

La decisione di ripubblicare la *Vita* del Pigna rispondeva non soltanto a un intento di arricchimento editoriale – rendendo così più appetibile il volume sul mercato librario – ma anche a un preciso disegno polemico. Nella *Vita*, infatti, Pigna sosteneva:²⁶

Cominciò l’Ariosto un altro poema, che dalla invention del Furioso non si partiva, del quale (forse contra sua voglia per non esser egli stato il publicator d’essi) cinque Canti si leggono, che il palagio del Signor delle fate, hanno nel primo aspetto. Egli dicea, che questa era un’orditura, & che deliberato havea di traporvi abbattimenti, & viaggi, & altre somiglianti cose, che compimento le dessero. Dal che comprender si può qual fosse la via del comporre da lui usata. Primieramente molti Episodii atti a essere allargati, raccoglieva in uno, & le attioni poi vi frammetteva, che gli paressero a dare spirito al rimanente bastevoli. Questi cinque Canti fanno un poema tale, quale è l’Odissea, che seguita l’Iliade in Ulisse, & seguono la materia del Furioso con nuovo & diverso soggetto, che dai proposti principii non si scosta. Giudicano alcuni, ch’essi da lui sarebbono stati sparsi qua & là per varii luoghi del suo Orlando, il che egli non disse già mai. Anzi pur per contrario lasciossi intendere, ch’egli di fare un’altra opera intendea, che dovesse star da per sé.

Secondo la lettura del Pigna, dunque, i *Cinque Canti* costituivano l’ossatura di un nuovo poema, destinato a proseguire la materia del *Furioso* mediante l’inserzione di ulteriori azioni ed episodi. La sua interpretazione, tuttavia, non era univoca: lo stesso autore, quasi *en passant*, ricorda

²⁵ L’edizione critica de *I romanzi* di Giovan Battista Pigna è stata curata da Salvatore Ritrovato per la Commissione per i testi di lingua (Bologna, 1997). Online è possibile trovare diverse copie digitalizzate dell’opera (Valgrisi, 1554), tra cui possiamo menzionare quella conservata dalla Bnf (consultabile sulla piattaforma *Gallica* al seguente link: <https://gallica.bnf.fr/ark%3A/12148/bpt6k58600k.image?utm_source=chatgpt.com#>>).

²⁶ G. Battista Pigna, *La vita di M. Lodovico Ariosto*, in L. Ariosto, *Orlando furioso*, a cura di G. Ruscelli, Venezia, Valgrisi, 1556, Gal. 29 - II, Galileo. I.19, Galilei Galileo. 19, Orlando furioso postillato, BNFC, (antica segnatura: Div. 2a - P. I, t. 19), BNFC, cc. 4 v. - 6 r., consultabile al seguente link: <<https://bibdig.museogalileo.it/tecanew/opera?bid=19740&seq=8>>.

l'esistenza di altre ipotesi circolanti negli ambienti eruditi, segno della vitalità del dibattito coevo. Di segno radicalmente opposto è invece l'esegesi proposta da Ruscelli nella lettera *Ai lettori*, nel paragrafo intitolato *Dei cinque canti nuovamente aggiunti negli altri Furiosi stampati*. Egli riferisce, sulla base di testimonianze orali provenienti da coloro che avevano conosciuto Ariosto, che l'intento originario del poeta era stato quello di raggiungere un numero tondo di cinquanta canti; un'idea poi abbandonata per ragioni di misura e di opportunità. Secondo il progetto iniziale, il poema non si sarebbe chiuso con la vittoria dei paladini cristiani e con le nozze di Ruggero e Bradamante, destinate a fondare la casata estense, ma con la tragica rotta di Roncisvalle, lasciando i lettori senza alcuna consolazione di fronte alla sorte luttuosa dei protagonisti. I *Cinque Canti*, secondo Ruscelli, non erano altro che il relitto di questa prima intenzione, un abbozzo scartato e in seguito definitivamente rigettato:²⁷

Perché dunque dappoi ch'egli hebbe finito di comporre tutto questo bellissimo libro suo, lo tenne molti anni per polirlo, & adornarlo in quella perfettione, nella quale lo veggiamo, egli havendo già risecati, & come gittati via quegli altri Canti di soverchio [cioè i Cinque Canti], attese a polire, correggere, & coltivar solamente gli eletti, senza pensar più in alcun modo ai reprobati. Et di qui si vede, che di questi che chiamano cinque Canti sono tutti tronchi, rotti, confusi, & quello che più importa, tanto pieni di errori di lingua, & d'ogn'altra cosa, che ben si potrebbe credere, che l'Autore se tornasse vivo, si sdegnerebbe grandemente, che doppo la morte sua, si sien dati fuori per lui, contra ogni sua voglia et intentione.

La posizione di Ruscelli, in contrapposizione a quella di Pigna, interpreta i *Cinque Canti* non come progetto embrionale di un nuovo poema, ma come materiale di scarto, “tronco, rotto, confuso”, e soprattutto indegno di rappresentare la volontà ultima dell'autore. Il dibattito tra i due critici esemplifica bene il nodo centrale posto dai *Cinque Canti*: la loro collocazione all'interno del corpus ariostesco. Se per alcuni essi costituivano la struttura portante di un'opera distinta, destinata a proseguire la materia del *Furioso*, per altri rappresentavano invece un residuo privo di legittimità letteraria. Il contrasto tra queste interpretazioni mette in luce la duplice natura del frammento: da un lato come estensione del poema maggiore, dall'altro come nucleo autonomo e incompiuto, caratterizzato da peculiarità tematiche e stilistiche proprie. Questa ambiguità si inserisce in un quadro più ampio: non tutte le narrazioni in ottave concepite da Ariosto confluirono nella redazione definitiva del *Furioso*. Restano infatti a testimonianza 84 stanze note come *Per la storia d'Italia* e 15 ottave denominate *Scudo della regina Elisa*, originariamente pensate per l'episodio della rocca di Tristano, poi confluito, con modifiche

²⁷ G. Ruscelli, *Ai lettori*, in L. Ariosto, *Orlando furioso*, Venezia, Valgrisi, 1556, Gal. 29 - II, Galileo. I.19, Galilei Galileo. 19, *Orlando furioso postillato*, BNFC, (antica segnatura: Div. 2a - P. I, t. 19), c. 6 v.- c. 7r., consultabile al seguente link: <<https://bibdig.museogalileo.it/tecanew/opera?bid=19740&seq=12>>.

sostanziali, nel canto XXXII dell'edizione del 1532.²⁸ Tale dato dimostra come il metodo compositivo ariostesco si fondasse su un complesso lavoro di selezione, rielaborazione e montaggio, del quale la tradizione manoscritta e a stampa non conserva tracce sistematiche. A questo riguardo, l'ipotesi oggi più accreditata è quella avanzata da Casadei, secondo cui i *Cinque Canti* devono essere intesi come continuazione diretta del *Furioso* e non come opera autonoma, in virtù del fitto intreccio di rimandi tematici e narrativi con il poema. Essi assumono dunque la funzione di una giunta coerente, fondata su una consequenzialità interna che ne rivela la dipendenza dal testo maggiore. Pur distinguendosi per toni più cupi e accenti pessimistici che ne giustificano una lettura autonoma, i *Cinque Canti* vanno considerati parte di un progetto unitario con cui Ariosto intendeva prolungare e completare la vicenda del *Furioso*. La posizione di Casadei, maturata a partire dal saggio del 1988 (*Alcune considerazioni sui 'Cinque Canti'*, «Giornale storico della letteratura italiana», CLXV, pp. 161-179) e sviluppata in *Il percorso del Furioso* (1993), si distingue per la proposta di una diversa collocazione rispetto alla tesi tradizionale.²⁹ Secondo l'interpretazione più diffusa, i *Cinque Canti* si sarebbero dovuti inserire nel canto XL del *Furioso* (redazioni A e B), subito dopo l'ottava 45. Casadei osserva tuttavia come tale soluzione generi incongruenze narrative e contraddizioni con il resto del poema. Per superare queste difficoltà, mette in evidenza il rapporto strutturale fra l'ottava I, 59 dei *Cinque Canti* e le ottave XL, 46-48 del *Furioso*, dedicate alle feste per le nozze di Ruggiero e Bradamante: la ripresa lessicale e situazionale rivela infatti una continuità narrativa che non può essere casuale. Ariosto non avrebbe potuto scrivere l'incipit di I, 59 senza presupporre la conoscenza, da parte del lettore, delle ottave immediatamente precedenti del *Furioso*. Da qui la conclusione di Casadei: i *Cinque Canti* non si riallacciano a XL, 45 mantenuto all'interno del

²⁸ Per la storia d'Italia è un componimento in ottave attribuito a Ludovico Ariosto, noto anche come le "ottantaquattro stanze per la storia d'Italia". Si tratta di un frammento poetico incompiuto, di carattere storico-celebrativo, che Ariosto avrebbe iniziato a comporre tra la fine del primo e l'inizio del secondo decennio del Cinquecento (circa 1510-1513), nel periodo in cui era legato alla corte estense di Ferrara. Il poemetto ha come tema principale gli eventi politico-militari contemporanei all'autore, in particolare le guerre d'Italia, viste dal punto di vista della casata d'Este. Le ottave narrano e commentano fatti reali – battaglie, alleanze, rovesci di fortuna – con un tono che alterna l'elogio cortigiano e la riflessione storica e morale. Ariosto intendeva probabilmente tracciare una sorta di "storia poetica" dell'Italia moderna, simile per impianto all'epos cavalleresco, ma saldamente ancorata alla realtà politica del suo tempo. Di questo frammento non è noto alcun originale autografo superstite: ciò che si conserva è una copia custodita presso la Biblioteca Ambrosiana di Milano (ms. H 55 inf., in passato indicato anche come H 66 inf.). Si ritiene che questa copia sia stata preparata per Gian Vincenzo Pinelli, in un periodo in cui gli autografi si trovavano presso Cornelio Pigna, che li aveva ricevuti a sua volta da Agostino Mosti. Nel ms. H 55 inf., oltre alle prime venti stanze del suddetto frammento (cc. 280 r–283 v), si trovano anche altre carte apografe (cc. 283 v–286 v) che tramandano le stanze 76–81 e 97–109 del Canto XXXII dell'*Orlando furioso*, relative all'episodio della Rocca di Tristano.

²⁹ Per approfondire, cfr. A. Casadei, *I 'Cinque Canti' o l'ultima eredità di Boiardo*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 113-128; Id. *Sulla situazione testuale e sulla cronologia dei 'Cinque Canti'*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 179-192.

poema, ma a quella stessa ottava se collocata all'inizio di un nuovo canto, il primo dei *Cinque Canti*. La stanza 45, infatti, può essere spostata senza compromettere lo sviluppo narrativo, poiché svolge la funzione di raccordo e di annuncio della continuazione rappresentata dal nuovo materiale (il tradimento di Gano e gli episodi connessi). L'ipotesi alternativa – collocare la parte finale del canto XL del *Furioso* (dalla stanza 46 in avanti) dopo *Cinque Canti* I, 58 – appare invece meno convincente, poiché richiederebbe eccessivi interventi di ricomposizione e altererebbe l'unità interna tanto dell'ultimo canto quanto del frammento aggiuntivo.

L'altro nodo problematico posto dai *Cinque Canti* riguarda l'epoca della loro composizione. In assenza di prove documentarie certe, negli ultimi decenni sono state avanzate ipotesi anche molto divergenti, a partire dall'edizione critica di Segre (1954).³⁰ La maggior parte degli studiosi concorda oggi con Dionisotti nel fissare al 1519 il termine iniziale, e con Segre nell'individuare il triennio 1526-1528 come quello dell'ultimo intervento linguistico prima della rinuncia definitiva al progetto. Alla luce di queste considerazioni, sia di ordine storico sia di ordine linguistico, i *Cinque Canti* sarebbero stati composti tra il 1519 e il 1521, negli anni del fallimento delle ambizioni curiali di Ariosto e della rottura con il cardinale Ippolito, per poi essere ripresi e infine abbandonati intorno al 1528.³¹ Il tono di cupo pessimismo e la diversità delle tecniche narrative – si tratta infatti di un blocco unitario, privo di *entrelacement* e di registro ironico – resero la loro integrazione sempre più inopportuna agli occhi dell'autore, che tuttavia sfruttò parte di quel materiale per l'elaborazione dei nuovi episodi (i cosiddetti “frammenti autografi”) confluiti nell'edizione del 1532 (siglata C). L'opera si colloca dunque in una fase di passaggio: da un lato rimane radicata nell'arcaismo di matrice boiardesca, dall'altro riflette l'urgenza di un'attualità politica segnata dal conflitto tra Francia e Spagna e dall'irrompere della Riforma. Questo duplice registro produce un ibrido che Ariosto non riuscì a

³⁰ L. Ariosto, *Cinque Canti*, in *Orlando Furioso. Opere minori*, a cura di C. Segre e L. Caretti, 2 voll., Milano-Napoli, Ricciardi, 1954.

³¹ Una trattazione esaustiva dell'argomento non rientra negli obiettivi del presente lavoro; la questione è stata efficacemente sintetizzata da Ida Campeggiani nel saggio *Una nuova datazione per i Cinque Canti*, pubblicato in «Storie e linguaggi. Rivista di studi umanistici», 2 (1), 2016, pp. 71-94. Per approfondire, cfr. i seguenti saggi: A. Casadei, *I 'Cinque Canti' o l'ultima eredità di Boiardo*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 113-128; Id. *Sulla situazione testuale e sulla cronologia dei 'Cinque Canti'*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 179-192; L. Caretti, *Storia dei Cinque Canti e Note*, in L. Ariosto, *Cinque Canti*, a cura di L. Caretti, Torino, Einaudi, 1977 (prima ed. Venezia, Corbo & Fiore, 1974); C. Segre, *Studi sui Cinque Canti*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 121-177; C. Dionisotti, *Appunti sui Cinque Canti e sugli studi ariosteschi*, in *Studi e problemi di critica testuale*, a cura di Raffaele Spongano, Bologna, Commissione per i testi di lingua, 1961, pp. 369-382; Id., *Per la data dei Cinque Canti*, «Giornale Storico della letteratura italiana», CXXXVII, 1960, pp. 1-40.

portare a compimento, probabilmente per la difficoltà di conciliare l'eredità epico-cavalleresca con le nuove tensioni storico-politiche. I *Cinque Canti* si presentano così come testimonianza di una transizione irrisolta: un'opera che rivela la volontà del poeta di misurarsi con il proprio tempo, ma che al contempo mette in luce l'impossibilità di integrare pienamente quella tensione con i modelli letterari tradizionali.

La seconda edizione del poema fu stampata a Ferrara, presso l'officina tipografica del milanese Giovan Battista Pigna, e pubblicata il 13 febbraio 1521 sotto la diretta supervisione dell'autore, che si fece carico anche dei costi di produzione e dello smercio.³² L'edizione del 1521 (siglata B) si colloca in un contesto editoriale complesso: il successo immediato dell'opera aveva favorito la diffusione di copie pirata, spesso scorrette, che resero urgente un intervento correttivo e una nuova stampa. Ariosto stesso, come si evince dalla lettera inviata a Mario Equicola l'8 novembre 1520, riconosceva la necessità di una seconda edizione per sopperire all'esaurimento della prima tiratura, pur non avendo ancora avviato un progetto di revisione sistematica.³³ Il poeta lavorò dunque su una copia della *princeps*, depositandovi progressivamente il proprio meticoloso lavoro di correzione linguistica. La seconda edizione fu preparata con una certa rapidità, tanto che non rispecchiò appieno l'intenzione – dichiarata ma rimasta inattuata – di una revisione organica e di un ampliamento sostanziale (“quasi tutto formato di nuovo e ampliato”, si legge nel frontespizio). Nondimeno, essa rappresenta una tappa di rilievo nel processo di definizione del

³² L'edizione si presenta come «Orlando Furioso di Ludovico Ariosto nobile Ferrarese ristampato et con molta diligentia da lui corretto et quasi tutto formato di nuovo et ampliato con gratie e privilegi. (In fine.) Finisce Orlando Furioso de Lodovico Ariosto: Stampato in Ferrara per Giovanni Battista da la Pigna Milanese (sic) A di. XII. de febraro. M.D.XX». Per approfondire, cfr. U. Guidi, *Annali delle edizioni e delle versioni dell'Orlando Furioso e d'altri lavori al poema relativi*, Bologna 1861, pp. 5-6, consultabile al seguente link: <<https://ia600203.us.archive.org/9/items/annalidelleedizi00guiduoft/annalidelleedizi00guiduoft.pdf>> (ultimo accesso 17 ottobre 2025). Il frontespizio è racchiuso entro una cornice decorata con più gruppi di mazza e scure avvolti da una serpe; agli angoli della cornice ricorre, ripetuto, il motto *Pro bono malum*.

M. Catalano (*Vita di L. Ariosto ricostruita su nuovi documenti*, Genève, L.S. Olschki, 1930-31, vol. I, pp. 530-33) e C. Fahy (*L'autore in tipografia: le edizioni ferraresi dell'Orlando Furioso*, in AA.VV., *I libri di 'Orlando Innamorato'*, Modena, Panini, 1987, p. 108) ricordano che la tiratura del 1521 fu molto limitata (probabilmente 500 copie), il che spiega la rarità degli esemplari superstiti. Secondo il censimento nazionale EDIT16 - Edizioni Italiane del XVI sec. (scheda CNCE 2542, consultabile al seguente link: <<https://edit16.iccu.sbn.it/resultset-titoli/-/titoli/detail/2542>>), in Italia si conservano solamente due esemplari noti, rispettivamente presso la Biblioteca Angelica e la Biblioteca dell'Accademia nazionale dei Lincei e Corsiniana, Roma. A questi esemplari, bisogna aggiungere la copia digitalizzata (segn. S.B.699), posseduta dalla SLUB di Dresda (consultabile al seguente link: <https://digital.slub-dresden.de/werkansicht/dlf/92224/1?utm_source=chatgpt.com>).

³³ Cfr. L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 177.

poema: pur mantenendo i 40 canti, introduce un numero relativamente limitato di varianti rispetto al testo del 1516, ma di significativo valore storico-letterario (cfr. Tab. 1.1). Sul piano strutturale si registrano l'aggiunta e la soppressione di undici ottave ciascuna, nonché lo spostamento di alcune strofe; sul piano linguistico, si nota invece un più marcato adeguamento al codice petrarchesco, in linea con il gusto cortigiano del tempo. Le correzioni linguistiche riflettono l'attenzione di Ariosto alle istanze bembiane, già diffuse in forma di linee guida prima della pubblicazione delle *Prose della volgar lingua* (1525), e testimoniano la sua volontà di affinare la lingua del poema secondo le nuove norme della letteratura volgare.³⁴

Tabella 1.1 – Varianti strutturali fra le edizioni del 1516 e del 1521

Tipo di intervento	Descrizione	Significato critico
Aggiunte	XVII 62 = C XIX 62 XXIII 5-6 = C XXV 5-6 XXIII 81-82 = C XXV 81-82 XXIX 41 = C XXXI 41 XXX82 = C XXXIII 106 XXXV 1 = C XXXIX 1 XXXVIII 20-22 = C XLII 20-22	Rispondono ad esigenze di chiarificazione narrativa e teorica.
Soppressioni	XXXV 1-10 e XL 71	Motivate dalla nuova situazione storica e da possibili critiche dei lettori contemporanei.
Spostamenti	XXXIII 16-19 → XXX 6-9 = C XXXII 6-9 XXXV 11-12 → XXXIV 89-90 = C XXXVIII 89-90	Testimoniano l'attenzione di Ariosto alla coerenza strutturale e alla ricezione del pubblico.

³⁴ Una trattazione esaustiva dell'argomento non rientra negli obiettivi del presente lavoro. Per approfondire lo studio dell'edizione del 1521, cfr. i seguenti testi: G. Sangirardi, *Orlando furioso*, in Ludovico Ariosto, Firenze, Le Monnier, 2006, pp. 98-100; S. Zatti, *Il lavoro di tutta una vita. L'elaborazione del Furioso e i Cinque Canti*, in L. Ariosto, *Orlando furioso e Cinque Canti*, a cura di R. Ceserani e S. Zatti, Torino, Utet, 1997, vol. I, pp. 27-38; A. Casadei, *I mutamenti del 1521*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 91-112; Id., *La strategia delle varianti. Le correzioni storiche al terzo Furioso*, Lucca, Pacini Fazzi, 1988; C. Segre, *Storia interna dell'Orlando furioso*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 29-41; S. Debenedetti e C. Segre, *Nota al testo*, in L. Ariosto, *Orlando furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, Bologna, Commissione per i testi di lingua, 1960, vol. II, pp. 1649-1697.

Queste innovazioni, che a un primo sguardo potrebbero apparire marginali, rivelano invece la piena consapevolezza di Ariosto di doversi misurare non solo con esigenze di perfezionamento stilistico, ma anche con le reazioni dei lettori e con il mutato quadro politico-culturale. In quest'ottica, la soppressione delle dieci ottave d'apertura del canto XXXV e dell'ottava 71 dell'ultimo canto risulta particolarmente eloquente. Nel primo caso, Ariosto aveva proposto una riflessione ironica sulle alleanze di Ferrara, arrivando a evocare un immaginario "comunismo delle donne", paradossalmente fonte di felicità; nel secondo, Ferrara appariva, sotto la guida del cardinale Ippolito, come l'unica città italiana davvero libera. La cancellazione di entrambi i passi riflette verosimilmente il mutare della situazione storica, ma potrebbe anche risentire di osservazioni critiche precoci. Lodovico Dolce, ad esempio, giudicava inammissibile la presenza di una stanza che parlasse delle donne in termini così disinvolti, sostenendo che un poema epico dovesse mantenere una condotta "casta" in ogni sua parte. Ariosto sembra dunque tener conto delle prime reazioni al testo, modulando le sue scelte in funzione del pubblico e delle convenzioni del genere. Le aggiunte, d'altro canto, rispondono a esigenze di chiarificazione e di difesa. Nel canto XXIII Ariosto spiega perché Ruggiero non si converta subito al cristianesimo dopo l'incontro con Ricciardetto: l'eroe resta legato all'onore militare, che gli impone di non abbandonare Agramante in difficoltà, e teme di essere accusato di viltà se scegliesse proprio allora la fede cristiana. È evidente la volontà di rafforzare la coerenza del personaggio e prevenire accuse di contraddizione, come quelle che anni dopo muoverà Sperone Speroni, rimproverando ad Ariosto di aver reso la conversione di Ruggiero ora eccessivamente pronta, ora troppo esitante. Analoga funzione ha l'inserimento nel canto XXXVIII della replica ad alcune critiche di Federico Fulgoso circa l'episodio del duello di Lipadusa (cfr. B XXXVIII 20-22 = C XLII 20-22). Fulgoso sosteneva l'inverosimiglianza della scena, poiché l'isola, aspra e montuosa, non poteva accogliere un combattimento equestre. Ariosto risponde ribadendo la piena autonomia della letteratura rispetto alla realtà empirica: l'isola, al tempo della narrazione, avrebbe avuto una spianata adatta al combattimento; solo in seguito un terremoto fece precipitare un masso che la coprì, rendendo il terreno accidentato. In tal modo egli trasforma un'obiezione in spunto per rivendicare la libertà dell'immaginazione, affermando che la veridicità della poesia non si misura sui fatti storici ma sulla coerenza interna della narrazione. La poesia, pertanto, non può mai mentire, poiché non appartiene al medesimo orizzonte della realtà fattuale.

Da queste correzioni emergono tre nuclei critici che orientano la seconda edizione: il rapporto con la storia contemporanea e le sue mutazioni; la condotta esemplare di Ruggiero, figura che Ariosto intende rendere coerente con il proprio ruolo di eroe; e, infine, la riflessione sul senso

della poesia come costruzione autonoma, distinta dalla verità materiale. È in questa direzione che si devono leggere anche le giunte dell'edizione del 1532, che ampliano ulteriormente il progetto poetico. Parallelamente, la revisione ariostesca si colloca in un contesto letterario dinamico, caratterizzato dall'emergere di poemi derivati dal *Furioso*. Già nel 1521 Cassio da Narni pubblica a Ferrara la *Morte del Danese*; seguono, a breve distanza, il *Belisardo* di Marco Guazzo (1525) e il *Rinaldo furioso* di Marco Cavallo (1526). È plausibile che Ariosto abbia sentito la necessità di differenziare la propria opera da questi tentativi di imitazione, riaffermando, attraverso le sue revisioni, l'originalità del proprio poema. Non meno decisivo è il dato della ricezione: dal 1524 in avanti, il *Furioso* conosce numerose ristampe in diverse città italiane, segno di un successo crescente. Proprio in questi anni il poema diventa punto di riferimento per autori giovani e sperimentali, che lo assumono insieme come modello e come termine di confronto polemico. Teofilo Folengo, con l'*Orlandino* (1520), Pietro Aretino, con la *Marfisa* (1532), e Luigi Tansillo, con il *Vendemmiatore* (1532), oscillano tra rispetto e trasgressione, testimoniando l'immediata centralità dell'opera ariostesca nel panorama letterario del tempo. In conclusione, l'edizione del 1521 segna un momento cruciale nell'iter redazionale del poema non tanto per la quantità delle varianti, quanto per il loro significato: essa documenta il dialogo di Ariosto con il proprio pubblico, con la tradizione epica e con la storia contemporanea. Le correzioni linguistiche, le soppressioni e le aggiunte rivelano una riflessione che va oltre l'estetica, investendo la funzione della poesia, il ruolo dell'eroe e il rapporto tra letteratura e verità. In questo senso, il secondo *Furioso* non è una semplice revisione, ma l'avvio di una nuova fase del laboratorio ariostesco, che troverà pieno compimento nell'edizione del 1532.

1.3 La storia redazionale del poema: la terza edizione del 1532 e i frammenti autografi

Il 19 giugno 1531 Ariosto manifestava al conte Nicolò Tassone da Este, oratore ducale, l'intenzione di ristampare il *Furioso*. Riportiamo di seguito il testo integrale della missiva:³⁵

All'III^{mo} e M^{co} Conte Nicolò Tassone da Este duc^{le} Oratore ápresso l'III^{mo} Duca di Milano maggiore honorando.

Signor conte mio hn^{mo}. Vostra S^{ria} non si graverà s'io le darò fatica, ché l'humanità sua verso di me mi daria ardire di affaticarlo in molto maggior cosa di questa, con fidutia che non meno la faria volentieri, che io havria piacere ch'ella lo facesse. Io vorrei stampare di novo il mio *Orlando furioso*, acciò che io gli emendassi molti errori che, oltra quelli che per poca diligentia vi ho fatti io, hanno fatto anchora li stampatori; et ancho vi ho fatto alcune aggiunte che spiero che non spiacerranno a chi le leggerà. E perché vorrei essere sicuro che li stampatori non l'havessino a stampare contra mia volontà, prima ch'io lo stampisca ho ottenuto da quasi tutte le potentie d'Italia che finché viva io nessuno lo possa stampare senza mia licentia. Io vorrei ancho ottenere il medesimo da l'III^{mo} Signor Duca di Milano, e così prego Vostra S^{ria} che sia contenta d'impetrarmi questa gratia da sua ex^{tia}; et acciò sia informata di quello che vorrei, li mando qui annessa una coppia de la lettera che circa questo mi ha fatto il S^{re} Duca di Mantova. Di questo io la priego strettamente, alla quale mi offero e raccomando sempre: e la prego che mi raccomandi in bona gratia del S^{re} Conte Maximiano.

Ferrariae, xix Iunij 1531.

di Vostra S^{ria} Ludovico Ariosto.

Il poeta si rivolge dunque a Nicolò Tassone da Este affinché interceda presso il Duca di Milano per ottenere un privilegio di stampa che tuteli il *Furioso* da edizioni non autorizzate. Ariosto motiva la richiesta con l'intenzione di pubblicare una nuova edizione, rivista e ampliata, resa necessaria tanto dagli errori a lui imputabili quanto da quelli introdotti dai tipografi. Sottolinea inoltre come analoghi privilegi gli siano già stati concessi da altri principi italiani, rafforzando così la legittimità e l'urgenza della sua istanza.

La terza edizione del poema (siglata C) uscì a Ferrara il 1° ottobre 1532, in circa 3.000 copie, per i tipi di Francesco Rosso da Valenza.³⁶ Figlio di Lorenzo, principale editore ferrarese dei primi

³⁵ L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 458.

³⁶ L'edizione si presenta come «Orlando Furioso di Messer Ludovico Ariosto nobile Ferrarese nuovamente da lui proprio corretto e d'altri canti nuovi ampliato con gratie e privilegi. (In fine.) Impresso in Ferrara per maestro Francesco Rosso da Valenza, a di primo d'ottobre.. M.D.XXXII». Per approfondire, cfr. U. Guidi, *Annali delle edizioni e delle versioni dell'Orlando Furioso e d'altri lavori al poema relativi*, Bologna 1861, pp. 18-23, consultabile [al](#) [seguinte](#) [link](#):

decenni del Cinquecento, Rosso esercitava in proprio da meno di dieci anni e fino ad allora aveva pubblicato pochissimi testi, talvolta guardati con sospetto in ambito religioso: una traduzione latina di un trattato sulla peste del medico Antonio Manardi, alcuni scritti di Pellegrino Morato sul Natale e sul *Padre nostro*, gli oroscopi dell'astrologo Zacuth, un trattato di logica e un commento a Porfirio di Antonio Musa Brasavola, una lezione sul matrimonio di Camillo Vistarini e una preghiera di Bartolomeo Ricci da Lugo. Una scelta editoriale così singolare lascia riflettere: il tipografo, infatti, rappresentava spesso anche un preciso orientamento culturale. Forse si tratta di un indizio di inquietudini intellettuali e religiose più profonde di quanto si pensi, anche considerando che lo stesso Ariosto, già in gioventù, aveva mostrato interesse per testi ermetici e neoplatonici, frequentando studiosi come Celio Calcagnini.

Anche in questa occasione il poeta seguì da vicino il lavoro tipografico, intervenendo con correzioni significative quando il testo era già in composizione.³⁷ Come nel 1521, consegnò al tipografo una copia della redazione precedente, fittamente annotata con varianti autografe, accompagnata da quaderni separati contenenti oltre settecento ottave nuove, che portarono il

<https://ia600203.us.archive.org/9/items/annalidelleedizi00guiduoft/annalidelleedizi00guiduoft.pdf> (ultimo accesso 17 ottobre 2025).

³⁷ Secondo la *recensio* condotta da Connor Fahy e confluita nel volume *L'Orlando Furioso del 1532. Profilo di un'edizione* (Milano, Vita e Pensiero, 1989), gli esemplari superstiti della terza stampa ferrarese sarebbero ventiquattro. L'indagine di Fahy si colloca nel solco degli studi di Santorre Debenedetti e Michele Catalano, integrandone i risultati con un'attenzione particolare alla storia interna della stampa del Furioso del 1532. L'organizzazione del volume riflette un preciso orientamento metodologico: la descrizione dell'"esemplare ideale", nozione cardine della bibliografia testuale. Il concetto di *ideal copy*, come formulato da G. T. Tanselle, non si limita a designare una copia concreta ma consiste nella ricostruzione storica della forma o delle forme che l'edizione assunse al momento della sua distribuzione. Tale ricostruzione comprende necessariamente tutte le varianti, intenzionali o accidentali, che caratterizzarono gli esemplari prodotti. L'"esemplare ideale" non ha dunque una consistenza materiale, bensì costituisce una costruzione teorica, indispensabile per orientarsi nella complessità delle testimonianze bibliografiche e per distinguere ciò che appartiene alla progettualità editoriale da ciò che è il risultato di contingenze tipografiche.

In questa prospettiva, l'analisi di Fahy si lega strettamente alla textual bibliography, ovvero alla filologia dei testi a stampa. Essa, lungi dal considerare il libro come un semplice veicolo neutro di contenuti letterari, indaga invece la materialità dell'oggetto tipografico – mise en page, apparato paratestuale, errori e correzioni di stampa, sequenza dei fascicoli – quale parte integrante della trasmissione testuale. Tale approccio non soltanto consente di chiarire le dinamiche produttive che hanno dato forma all'edizione ariostesca del 1532, ma apre anche la strada a riflessioni metodologiche di più ampio respiro, investendo la storia della stampa italiana e la stessa pratica della filologia sui testi a stampa. Il lavoro di Fahy assume così un valore duplice: da un lato, documenta e interpreta in modo più completo la complessa tradizione dell'edizione definitiva dell'*Orlando Furioso*; dall'altro, mostra come la descrizione e la collazione degli esemplari superstiti siano condizioni preliminari e imprescindibili di ogni indagine filologica che voglia fondarsi su basi storicamente solide.

Per approfondire, cfr. i seguenti testi: L. Spagnolo, *Variantistica ed ecdotica dell'Orlando furioso*, «Tipo-filologia», 1, 2008, pp. 61-87; H. Harris, *Filologia e bibliologia a confronto nell'Orlando furioso del 1532*, in AA.VV., *Libri, tipografi, biblioteche. Ricerche storiche dedicate a Luigi Balsamo*, Firenze, Olshki, 1997, pp. 105-122.

poema all'assetto definitivo di 46 canti.³⁸ L'edizione, curata dall'autore fin nei minimi dettagli, si distingue per una veste tipografica raffinata: per la prima volta vi compare un suo ritratto tratto da un disegno di Tiziano; il colophon raffigura due vipere – una già con la lingua recisa, l'altra in procinto di subirne la stessa sorte da una mano armata di forbici – accompagnate dal motto *Dilexisti malitiam super benignitatem*. Il precedente motto *Pro bono malum* non scompare del tutto, ma in alcuni esemplari è spostato in coda, dopo l'explicit, insieme a una piccola incisione che raffigura una lupa nell'atto di allattare un lupicino. Le prime cinque copie furono donate rispettivamente al duca di Ferrara Alfonso, al marchese di Mantova Federico Gonzaga, al marchese del Vasto, a Isabella d'Este e a Margherita Paleologa Gonzaga. La copia destinata a Isabella era accompagnata da una lettera dell'autore, datata 9 ottobre 1532, riportata qui integralmente:³⁹

All'III^{ma} et ex^{ma} Signora mia obser^{ma} la S^{ra} Marchesana di Mantova etc. In Venetia.

III^{ma} et ex^{ma} Signora mia obser^{ma}. Io mando a vostra ex^{ma} uno de li miei *Orlandi furiosi*, che, havendoli meglio corretti et ampliati di sei canti e di molte stanze sparse chi qua chi là pel libro, mi parrebbe molto uscire del debito mio, s'io inanzi a tutti gli altri non ne facessi copia a vostra ex^{tia}, come a quella che riverisco e adoro, et alla quale so che le mie compositioni (sieno come si vogliono) essere gratissime sogliono. Quella si degnerà di accettarlo insieme col buono animo col quale io le fo questo picciol dono: in buona gratia de la quale mi ricomando sempre.

³⁸ La maggior parte delle varianti dell'*Orlando Furioso* si può ricavare dal confronto diretto fra le tre edizioni a stampa, poichè Ariosto, in vista della pubblicazione della seconda e poi della terza redazione, consegnava in tipografia una copia dell'edizione precedente annotata a mano con le correzioni da introdurre. Questo metodo di lavoro è dimostrato dalla presenza di refusi comuni: alcuni compaiono sia nella *princeps* del 1516 (A) sia nell'edizione del 1521 (B), altri invece si ritrovano in B e nella definitiva del 1532 (C). Si tratta di un chiaro indizio, leggibile con gli strumenti della filologia lachmanniana, della dipendenza diretta di B da A e di C da B (cfr. S. Debenedetti e C. Segre, *Nota al testo*, in L. Ariosto, *Orlando furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, Bologna, Commissione per i testi di lingua, 1960, vol. 2, pp. 1649-1697). Questa procedura, tuttavia, poteva servire solo per interventi limitati e puntuali. Per modifiche più estese, per rielaborazioni strutturali o per l'aggiunta di interi passi, Ariosto dovette necessariamente servirsi di fogli sciolti o materiali separati. È probabile che tali supporti, in quanto strumenti di lavoro quotidiano e di natura effimera, non siano mai stati archiviati con cura e siano andati dispersi molto presto. Lo stesso vale per le copie annotate delle edizioni precedenti: essendo esemplari "di servizio", utilizzati in tipografia e privi di valore autonomo, una volta svolta la loro funzione, difficilmente venivano conservati. La dispersione dei materiali autografi di Ariosto rientra in una dinamica più generale della tradizione letteraria rinascimentale: raramente i laboratori tipografici mantenevano traccia sistematica delle fasi preparatorie, e le carte di lavoro degli autori andavano quasi sempre perdute. Ragion per cui, la tradizione delle varianti ariostesche può essere ricostruita quasi soltanto attraverso le tre edizioni a stampa, che costituiscono – con l'unica eccezione dei cosiddetti "frammenti autografi" – gli unici testimoni superstiti del processo di revisione.

³⁹ L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 488.

Ferrariae, 9 octobris 1532.
di vostra ex^{tia}
Servitore dedit^{mo} Ludovico Ariosto.

Questa lettera conferma che la terza edizione segna un ampliamento considerevole del poema, che passa da 40 a 46 canti. Ariosto vi inserisce quattro ampi episodi: la storia di Olimpia (innestata sul IX canto di *AB* e collocata tra il IX e l'XI di *C*), comprensiva della celebre invettiva contro le armi da fuoco; l'episodio di Bradamante e dei tre re nordici ospiti alla Rocca di Tristano (originariamente nel XXX di *AB*, collocato tra il XXXII e il XXXIII di *C*), con la rilevante ekphrasis politica delle pitture di Merlino, profetiche delle disfatte francesi in Italia; la vicenda del tiranno Marganorre, che occupa un canto interamente nuovo, il XXXVII di *C*, inserito dopo l'antico XXXIII di *AB*; e infine la più ampia, la storia del principe greco Leone, che si colloca tra il penultimo e l'ultimo canto di *AB* e occupa i canti XLIV, XLV e gran parte del XLVI di *C*. Tra questi, il XLV è del tutto nuovo e rappresenta l'ultima peripezia romanzesca prima della conclusione epica (con il duello tra Ruggiero e Rodomonte) e dinastica (con le nozze dei due capostipiti estensi). Accanto a questi ampi inserimenti, l'edizione del 1532 presenta anche aggiunte e soppressioni di singole ottave o interi gruppi, spesso in relazione a personaggi ed eventi storici contemporanei: vengono eliminate le ottave XVI 75, XXXII 34, XXXIX 112, XL 10 e 64, mentre si aggiungono le ottave XIII 71-72, XV 18-36, XXVI 50-52, XXVII 124, XXXIII 65-76, XXXV 32, XXXVI 84, XLIII 56-59 e XLVI 102-103. Inoltre, tutto il poema è sottoposto a un'accurata revisione linguistica e stilistica nel segno dei dettami di Pietro Bembo, come testimonia la seguente lettera:⁴⁰

Al Rev^{mo} Mons. Pietro Bembo.

Mag^{co} e Rev^{mo} m. Pietro patron mio hon^{mo}. Virginio mio figliuolo viene a Padova per studiare. Io gli ho commesso che, la prima cosa che faccia, venga a far riverentia a V. Sig^{ria}, e si faccia da lei conoscere per suo servitore. Io priego V. Sig^{ria} che dove gli sarà bisogno il suo favore sia contenta di prestarglielo, e sempre che lo vederà lo admonisca et esorti a non gittare il tempo: alla quale mi offero e raccomando sempre.

Io sono per finir di rivedere il mio Furioso; poi verrò a Padova per conferire con V. Sig^{ria}, et imparare da lei quello che per me non sono atto a conoscere. Che Dio conservi sempre.

Ferrara, alli xxiii Febraro MDXXXI.

Di Vostra Signoria

Servitore Ludovico Ariosto.

⁴⁰ L. Ariosto, *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere*, vol. III: *Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984, p. 457.

In questa lettera Ariosto raccomanda a Bembo il figlio Virginio, che sta per recarsi a Padova a studiare, chiedendogli sostegno e consiglio. In chiusura, annuncia di essere in procinto di terminare la revisione del *Furioso* e di voler poi confrontarsi direttamente con lui.

La terza redazione segna una trasformazione profonda del poema, che da racconto cavalleresco di tono romanzesco si orienta verso forme più classicizzanti ed epiche.⁴¹ Fin dal primo verso il cambiamento è evidente: non più «*Di donne e cavallier li antiqui amori*», con l'accento sul confronto tra antico e moderno, ma «*Le donne, i cavallier, l'arme, gli amori*». Il proemio, costruito secondo un impianto a chiasmo – che associa donne e amori, cavalieri e armi, senza tuttavia escludere donne guerriere e cavalieri innamorati – mette in primo piano due nuclei tematici centrali: i rapporti tra uomini e donne e l'intreccio fra amore e guerra. Ne risulta un poema meno rivolto al passato e più radicato nella contemporaneità: da un lato la vita di corte, dall'altro la possibilità di contrapporre l'amore alle armi e la letteratura alla guerra. Questa impostazione accentua anche l'impronta virgiliana del *Furioso*, grazie al richiamo all'incipit dell'*Eneide* («*Arma virumque cano*»), introdotto solo nella terza edizione e volto a rafforzare la parentela poetica e ideologica con il modello classico. In tale prospettiva si comprendono le nuove aggiunte – dall'invettiva contro le armi da fuoco (canto IX), ai riferimenti alle invasioni francesi (canto XXXIII), fino alle allusioni alle scoperte geografiche (canto XV) – che testimoniano il superamento dell'orizzonte municipale delle prime versioni in favore di una prospettiva nazionale ed europea. L'intensificarsi dei richiami alla storia contemporanea riflette sia il passaggio degli Estensi dall'alleanza francese al sostegno di Carlo V, sia le tensioni religiose legate alla Riforma, particolarmente sentite a Ferrara dopo l'arrivo di Renata di Francia nel 1528. La revisione va dunque collocata nel contesto della crisi politica italiana seguita alla battaglia di Pavia (1525), al trattato di Madrid (1526) e al Sacco di Roma (1527), eventi che segnarono la fine delle aspirazioni di autonomia degli stati italiani. Sul piano culturale, essa corrisponde alla piena affermazione del toscano come lingua letteraria nazionale, in linea con i dettami bembiani. Non bisogna dimenticare, infatti, che il *Furioso* è il primo grande capolavoro europeo concepito per la tipografia, circostanza che ne favorì la diffusione sovraregionale e che contribuì a consacrare il toscano come lingua nazionale.

La ricostruzione del processo di riscrittura del poema è stata resa possibile grazie alla pubblicazione de *I frammenti autografi* a cura di Santorre Debenedetti (1937) e all'edizione critica di Cesare Segre (1960), delle quali si darà conto più diffusamente nel prosieguo del

⁴¹ Per approfondire l'argomento, cfr. L. Caretti, *Ariosto e Tasso*, Torino, Einaudi, 1961.

capitolo. La critica ha individuato tre direttrici fondamentali che orientano il lavoro ariostesco: l'italianizzazione linguistica, la classicizzazione strutturale e l'epicizzazione del racconto. Nella redazione definitiva del *Furioso* confluiscono due momenti distinti: un primo, ancora di respiro municipale, e un secondo, maturato in un contesto storico segnato dalle guerre franco-spagnole e dall'ideologia imperiale, con il conseguente rafforzamento dei riferimenti storici e la progressiva attenuazione dell'ironia. Le innovazioni introdotte rispondono anche a esigenze di ordine strutturale. Così, il salvataggio di Olimpia da parte di Orlando, in parallelo a quello di Angelica da parte di Ruggiero, crea un contrappunto che riequilibra i ruoli eroici; la disposizione dei nuovi episodi, inoltre, colloca le fasi della follia di Orlando esattamente al centro del poema, conferendole una funzione strutturale cruciale. Alcuni episodi, come quello di Marganorre, correggono lo squilibrio narrativo della vicenda delle "femmine omicide"; altri, invece, rilanciano la tematica amorosa, indagando le distorsioni dell'etica sentimentale e intrecciandole con valori quali la fedeltà alla parola data e la lotta contro l'arbitrio, che trovano espressione paradigmatica nelle vicende di Olimpia, Leone e Ruggiero.

Infine, resta la questione, a lungo discussa, circa l'eventualità che Ariosto abbia continuato a rielaborare il poema anche dopo la stampa. Girolamo Ruscelli, nelle *Annotationi, gli Avvertimenti, & le Dichiarationi* contenute nell'edizione dell'*Orlando furioso* da lui curata, trattando le «mutazioni et miglioramenti che m. Lodovico Ariosto havea fatti per mettere nell'ultima impressione del Furioso», riferisce che Galasso Ariosto, fratello di Ludovico, durante un incontro avvenuto a Reggio Emilia nel 1543, gli aveva mostrato «un Furioso degli ultimi stampati in Ferrara, il quale era solamente legato in un cartone rozo, et non era tagliato in torcolo o uguagliate le carte altramente per non restringere il margine, da potervi scrivere sopra. Et questo libro era per tutto notato et postillato di mano dell'autore stesso; dicendomi M. Galasso (come da me stesso io potei ancor riconoscere) che M. Ludovico era in animo di farlo ristampare ultimamente così tutto ricorretto et migliorato da lui medesimo».⁴² La testimonianza suscitò però dubbi già fra i contemporanei. Lodovico Dolce Lodovico Dolce, nell'edizione del 1568 curata per Giovanni Varisco, giudicava inverosimile l'esistenza di un «esemplare ricorretto di mano dell'Ariosto», poiché «l'Ariosto si morì pochi mesi da poi che questo tale esemplare o libro suo

⁴² G. Ruscelli, *Le Annotationi, gli Avvertimenti, & le Dichiarationi di Girolamo Ruscelli*, in *Orlando furioso di m. Lodovico Ariosto...*, In Venezia, Appresso Valgrisi, 1556, citato in I. Andreoli, *Girolamo Ruscelli nella bottega d'Erasmus*, in *Dall'accademia alla corte alla tipografia. Atti del Convegno internazionale di studi* (Viterbo, 6-8 ottobre 2011), a cura di P. Marini e P. Procaccioli, Roma, Vecchiarelli Editore, 2012, pp. 682-83.

ultimo uscì fuori, e non ebbe il tempo di rivederlo».⁴³ A sostegno di questa posizione si cita anche una lettera che lo stesso Galasso scrisse a Bembo nel luglio 1533, poi confluita nella raccolta sansoviniana delle lettere di Bembo (1560), in cui affermava che Ludovico non aveva potuto ristampare il poema «come havea in animo di fare», poiché subito dopo la pubblicazione della *princeps* «s'ammalò, et dopo l'essere stato otto mesi infermo finalmente s'è morto».⁴⁴ Più recentemente, Morini ha definitivamente escluso l'esistenza del postillato.⁴⁵ Al di là della controversia, l'episodio conferma la meticolosa attenzione con cui Ariosto seguì sempre la propria opera. Il *Furioso* del 1532 si presenta infatti come un testo aperto, costantemente suscettibile di revisione e ampliamento, un autentico *work in progress* consegnato dall'autore ai lettori e alla storia. Se i contemporanei apprezzarono soprattutto le edizioni del 1516 e del 1521, è tuttavia la versione del 1532 quella che, da cinque secoli, continua ad alimentare letture, entusiasmi e discussioni.

1.3.1 Storia, fisionomia materiale ed ordinamento dei frammenti autografi

Per lo studio della tradizione e della fisionomia materiale dei frammenti ariosteschi ci siamo basati sulla ricostruzione storico-filologica e codicologica proposta da Santorre Debenedetti nell'introduzione alla sua edizione de *I frammenti autografi dell'Orlando furioso*.⁴⁶ Come ricordato nei paragrafi precedenti, non ci sono pervenuti testimoni autografi anteriori all'edizione del 1516. Con ogni probabilità, la loro scomparsa fu determinata dalle operazioni di stampa, circostanza piuttosto frequente all'epoca. Alcuni fogli, tuttavia, sopravvissero più a lungo, conservati con cura dalla famiglia. Secondo la testimonianza di Saverio Bettinelli e

⁴³ L. Ariosto, *Orlando furioso, corretto e dichiarato da M. Lodovico Dolce*, In Venezia, Appresso Giovanni Varisco, 1568, citato in I. Andreoli, *Girolamo Ruscelli nella bottega d'Erasmus*, in *Dall'accademia alla corte alla tipografia. Atti del Convegno internazionale di studi* (Viterbo, 6-8 ottobre 2011), a cura di P. Marini e P. Procaccioli, Roma, Vecchiarelli Editore, 2012, p. 683.

⁴⁴ P. Bembo, *Delle lettere di M. Pietro Bembo*, In Venezia, Appresso F. Sansovino, 1560, citato in P. Trovato, *Con ogni diligenza corretto. La stampa e le revisioni editoriali dei testi letterari italiani (1470-1570)*, Bologna, Il Mulino, 1991, p. 295 (nota 71).

⁴⁵ Per approfondire, cfr. L. Morini, *Ruscelli e le pretese varianti ariostesche al Furioso del '32*, in *In ricordo di Cesare Angelini. Studi di letteratura e filologia*, a cura di F. Alessio e A. Stella, Milano, Il Saggiatore, 1979, pp. 160-84.

⁴⁶ Cfr. S. Debenedetti, *Introduzione*, in *I frammenti autografi dell'Orlando furioso*, a cura di S. Debenedetti, rist. con prefazione di Cesare Segre, Roma, Edizioni di Storia e Letteratura, 2010, pp. XIII-XVII (prima ed. Torino, Chiantorre, 1937).

Giovannandrea Barotti, erano ancora visibili nel Settecento, prima di scomparire definitivamente. Insieme ad essi andarono dispersi esemplari di straordinario valore, tra cui la copia del 1516 rivista per la ristampa del 1521 e quella del 1521 fittamente corretta dall'autore, che – secondo Giovan Battista Pigna nei *Romanzi* – si trovava presso Virginio Ariosto e che, stando a Girolamo Ruscelli, egli ebbe modo di vedere a Reggio nel 1543 presso Galasso Ariosto; si tratta con ogni probabilità del medesimo esemplare che il poeta aveva consegnato al tipografo Francesco Rosso da Valenza nel 1532.⁴⁷ Nonostante queste perdite, diversi quaderni e frammenti di quaderni autografi sono giunti fino a noi e documentano, in vari stadi di elaborazione e conservazione, la prima, la terza e la quarta giunta. Della seconda giunta non resta invece alcun originale, ma disponiamo di un apografo che ne restituisce, sia pure in modo approssimativo, il testo e l'ordinamento. A questo corpus, destinato in gran parte a confluire poi nel *Furioso*, si aggiunge un foglio autografo di particolare interesse: un frammento narrativo legato alla materia del poema, noto come *Lo scudo della regina Elisa*.

Alla morte di Ariosto, nel luglio del 1533, l'eredità passò al figlio Virginio. È probabile che tra i beni paterni fossero compresi anche gli autografi del *Furioso*, rimasti nella casa di famiglia a Ferrara per circa due secoli. La prima notizia a riguardo ci viene fornita da Ruscelli, il quale racconta di averli visti a Reggio presso Galasso Ariosto nel 1543. Nelle *Annotationi, Avvertimenti & Dichiarationi*, pubblicate in appendice alla sua edizione dell'*Orlando furioso*, trattando le “Mutazioni et miglioramenti” introdotti dal poeta per la terza edizione del poema, Ruscelli afferma:⁴⁸

Il qual M. Galasso mi mostrò molti quaderni di carta scritti parte di mano di detto M. Lodovico, come egli me la dichiarò, et parte d'altre diverse mani. Et questi erano quei fogli o quaderni ov'egli veniva scrivendo et componendo quel libro suo. Et vi erano delle stanze et de' versi molto cassati et postillati per sopra et nei margini, et altre più nette che dovevano essere state rescritte per aventura più d'una volta, et in molti luoghi vi erano diece, XV, XX, et cinquanta stanze tutte tagliate via, che non l'havea voluto mettere; et per certo ve ne erano delle bellissime, ch'io non seppi considerar perché le havesse risecate via, se non per non

⁴⁷ Cfr. G. Battista Pigna, *Romanzi*, In Venezia, Appresso Valgrisi, 1554, p. 121 citato in S. Debenedetti, *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XIV (nota 2); G. Ruscelli, *Le Annotationi, gli Avvertimenti, & le Dichiarationi di Girolamo Ruscelli*, in *Orlando furioso di m. Lodovico Ariosto...*, In Venezia, Appresso Valgrisi, 1556, citato in I. Andreoli, *Girolamo Ruscelli nella bottega d'Erasmus*, in *Dall'accademia alla corte alla tipografia. Atti del Convegno internazionale di studi* (Viterbo, 6-8 ottobre 2011), a cura di P. Marini e P. Procaccioli, Roma, Vecchiarelli Editore, 2012, pp. 682-83.

⁴⁸ G. Ruscelli, *Le Annotationi, gli Avvertimenti, & le Dichiarationi di Girolamo Ruscelli*, in *Orlando furioso di m. Lodovico Ariosto...*, In Venezia, Appresso Valgrisi, 1556, citato in S. Debenedetti, *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XIV (nota 4).

essere soverchiamente lungo, et queste erano la più parte in cose dell'istoria d'Italia. Benché in molti altri soggetti, che si leggono negli stampati, ve n'erano di molte cassate o tagliate via, come ho detto. Et fra l'altre cose, erano quelle, che doppo la morte sua hanno stampate et batezzate cinque canti d'un nuovo libro.

Secondo quanto riferisce Ruscelli, Galasso mostrò alcuni quaderni contenenti parti dell'*Orlando Furioso*, in parte autografi di Ariosto e in parte scritti da altre mani. In quei fogli si trovavano numerose correzioni, cancellature e annotazioni a margine, accanto a passi ricopiati più volte in forma più ordinata. Ariosto vi aveva soppresso molte stanze, anche di notevole bellezza, probabilmente per evitare che il poema risultasse troppo esteso, soprattutto nei brani di argomento storico legati all'Italia. Fra i materiali scartati figuravano anche quelli che, dopo la morte del poeta, furono dati alle stampe con il titolo di *Cinque Canti*. Debenedetti invita comunque a non fidarsi ciecamente della testimonianza di Ruscelli, pur ammettendo che l'accenno alla presenza di carte d'altra mano possa essere verosimile. Si ricorda infatti che il codice delle *Satire*, a lungo creduto autografo, si è poi rivelato copiato da un amanuense e successivamente corretto dall'autore in persona.

Alcuni frammenti, tuttavia, si dispersero già molto prima del Settecento. Nel Cinquecento un pezzo dell'episodio della Rocca di Tristano entrò nella collezione di Agostino Mosti, che lo donò a Cornelio Pigna a Ferrara: oggi ne resta soltanto una copia di modesto valore (cioè il suddetto apografo), conservata alla Biblioteca Ambrosiana (ms. H. 55 inf.). Nel Seicento un foglio e una carta uscirono dalla dimora ferrarese per raggiungere Napoli. Nel 1640 il predicatore Lodovico Antinori li ricevette dagli eredi di Ariosto e l'anno seguente – come attesta una scrittura sul verso della carta di guardia – li depositò nella biblioteca del convento dei Santi Apostoli. In seguito passarono alla biblioteca del convento di San Martino e infine alla Biblioteca Nazionale, dove furono legati insieme e si conservano ancora oggi (ms. San Martino 353). Il foglio, ben conservato, tramanda le stanze 50-71 del canto XLVI; la carta, invece, in pessime condizioni, contiene le ottave relative allo scudo della regina Elisa sopra menzionate. Un altro foglio, ceduto in data imprecisata, giunse all'Ambrosiana nel 1840 grazie al lascito di Giberto Borromeo. Le due carte, oggi divise e collocate nella cosiddetta “cartella ariostesca”, presentano due brevi estratti dell'episodio di Olimpia: rispettivamente IX 1-11; 92-94 e X 1-7, copiati con grande precisione. Il nucleo principale degli autografi rimase comunque presso la famiglia Ariosto fino alla metà del Settecento. Gli ultimi discendenti, ormai ridotti in gravi difficoltà economiche, custodirono i manoscritti fino all'estinzione della casata con Claudio Ariosto, nel 1747. Poco dopo il bibliofilo Giuseppe De Carli riuscì ad acquisirne una parte consistente e, nel 1758, la lasciò in eredità alla nascente Biblioteca di Ferrara. L'anno seguente il primo direttore della

biblioteca, Giovannandrea Barotti, completò la raccolta donando anche l'ultimo fascicolo in suo possesso. Così, tra dispersioni, passaggi di mano e fortunati recuperi, i frammenti superstiti del *Furioso* sono giunti fino a noi: documenti fragili ma di straordinario valore, che offrono uno sguardo diretto sul laboratorio creativo di Ariosto.

Il manoscritto conservato presso la Biblioteca Comunale di Ferrara (F) costituisce, insieme alle due carte Borromeo della Biblioteca Ambrosiana di Milano (M) e al foglio di Napoli (N I-2), un complesso unitario. F è un codice cartaceo in formato folio (mm 314 × 225), originariamente composto da 54 carte, numerate in epoca moderna da 1 a 53, poiché non è stata conteggiata una carta interamente bianca. La maggior parte delle carte reca filigrane riconducibili a due tipologie principali: più frequentemente la balestra e, più raramente, la bilancia, attestata in due varianti solo lievemente differenziate. Alla filigrana a bilancia si accompagna talvolta una contromarca (cfr. Fig. 1.1).

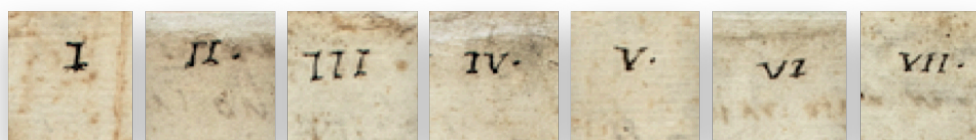
L'analisi delle filigrane richiederebbe oggi un approfondimento specifico, indispensabile per identificare con maggiore precisione l'area di produzione della carta e per una più puntuale datazione dei manoscritti. Una simile indagine, tuttavia, esula dagli obiettivi del presente lavoro, tanto più che presupporrebbe la consultazione diretta dei documenti originali, circostanza che non è stata possibile. Ci si limita pertanto, anche in questo caso, a riportare sinteticamente le osservazioni di Debenedetti. Per quanto riguarda la provenienza della carta, egli sottolinea l'impossibilità di formulare ipotesi certe. La filigrana della bilancia risulta infatti assai comune, ma anche l'altra tipologia era largamente diffusa, come attesta il repertorio classico di Briquet, *Les filigranes* (Genève, 1907, I, nn. 2364-2616, 701-770; cfr. anche pp. 178 ss., 49 ss.). Entrambi i segni compaiono di frequente nelle carte italiane: sebbene Debenedetti non sia riuscito a rintracciare esemplari perfettamente identici, rileva che nell'area veneta la bilancia costituisce il motivo più diffuso (cfr. D. Urbani, *Segni di cartiere antiche*, Venezia, 1870, p. 19: «Le balance: è questa la filigrana che s'incontra più numerosa di ogni altra negli stampati in Venezia...»; per la balestra Urbani menziona soltanto due esemplari, p. 30). Ulteriori riscontri sono forniti anche da F. Ongania, *L'art de l'imprimerie pendant la Renaissance à Venise* (Venezia, 1895-96).

Figura 1.1 – Esempi di filigrane e contromarca



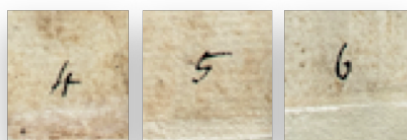
Il manoscritto ferrarese si compone di sette fascicoli, ciascuno provvisto di una camicia aggiunta da Barotti, che numerò sia i quaderni sia le rispettive camicie con numeri romani da I a VIII (cfr. Fig. 1.2).

Figura 1.2 – Numerazione dei fascicoli I-VIII apposta da Barotti sul margine superiore destro



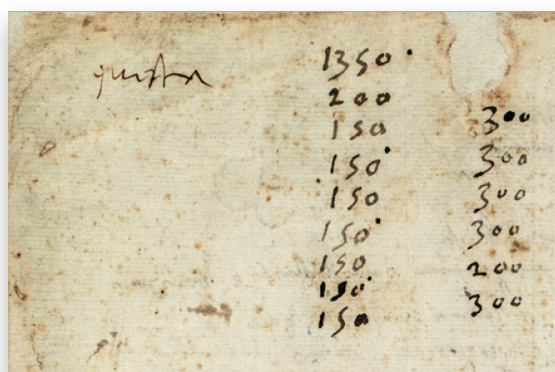
Non si sa a chi attribuirli, ma sono sicuramente posteriori, i numeri 4-6 sulla prima facciata, nel margine inferiore dei corrispondenti fascicoli (cfr. Fig. 1.3).

Figura 1.3 – Numerazione apocrafa dei fascicoli 4-6



L'insieme è protetto da una semplice coperta di cartone. Le carte sono scritte su entrambe le facciate, con due eccezioni: il verso della c. 35, lasciato in bianco (dove appare la parola “questa” e una serie di altri numeri, scritti da una mano diversa) e una carta non numerata, interamente bianca, collocata dopo la c. 25 (cfr. Fig. 1.4).

Figura 1.4 – c. 35 v.



Le pagine attribuibili direttamente all'Ariosto sono in tutto 105; vi si trovano tuttavia inserzioni estranee – appunti, numeri e segni – dovute a Barotti e ad altri bibliotecari, le quali si rivelano in genere superflue, talvolta inesatte e fuorvianti per il lettore. Ad esempio, Debenedetti fa notare che a c. 19 v., sul margine sinistro, in sostituzione dei versi «E con la coda alzar fa in modo l'onde / Che 'l ciel ne sferza e spesso il sol nasconde» (XI 35), il poeta aveva scritto:

«Ribombano al rumor che di lei s'ode / Le selve, i scogli e le lontane prode»; versi che poi – sebbene ciò non risulti direttamente da questa carta – collocò altrove, come chiusa della stanza 43. L'annotatore, tuttavia, non seppe resistere alla tentazione di collegare, con segni convenzionali, la stanza 35 con un abbozzo della 43, presente sulla stessa facciata, cadendo così in un evidente anacronismo (cfr. Fig. 1.5). Così come Ariosto contrassegnava alcuni versi con una lineetta, anche l'annotatore aggiungeva segni propri, con esiti fuorvianti (c. 32 v.: XXXVII 90, 1; c. 35 v.: 122, 1) (cfr. Figg. 1.6 e 1.7). Altre note, di data molto più tarda e di tono quasi scherzoso, compaiono sul margine inferiore di cc. 22 v. e 43 v.: «Quel che d'Orlando ecc. (segue dopo questi due fogli bianchi mal intromessi)» (cfr. Fig. 1.8), «Segue dopo i due fogli bianchi» (cfr. Fig. 1.9). Si tratta delle camicie barottiane. Due sole note meritano di essere ricordate: la prima, dello stesso Barotti, apposta in fondo alla prima faccia dell'ultimo fascicolo (c. 44 r.), recita «Dr. Barotti donò 1759» (cfr. Fig. 1.10). La seconda è di un illustre visitatore della biblioteca, Vittorio Alfieri, che nel maggio del 1783, dopo essersi recato nei sepolcri di Dante e Petrarca, giunse a Ferrara per visitarvi la tomba e i manoscritti dell'Ariosto. Sul margine inferiore dell'ultima carta del manoscritto (c. 53 v.) rimane la sua nota a lapis: «Vittorio Alfieri vide e venerò 18 giugno 1783» (cfr. Fig. 1.11).

Figura 1.5 – c. 19 v

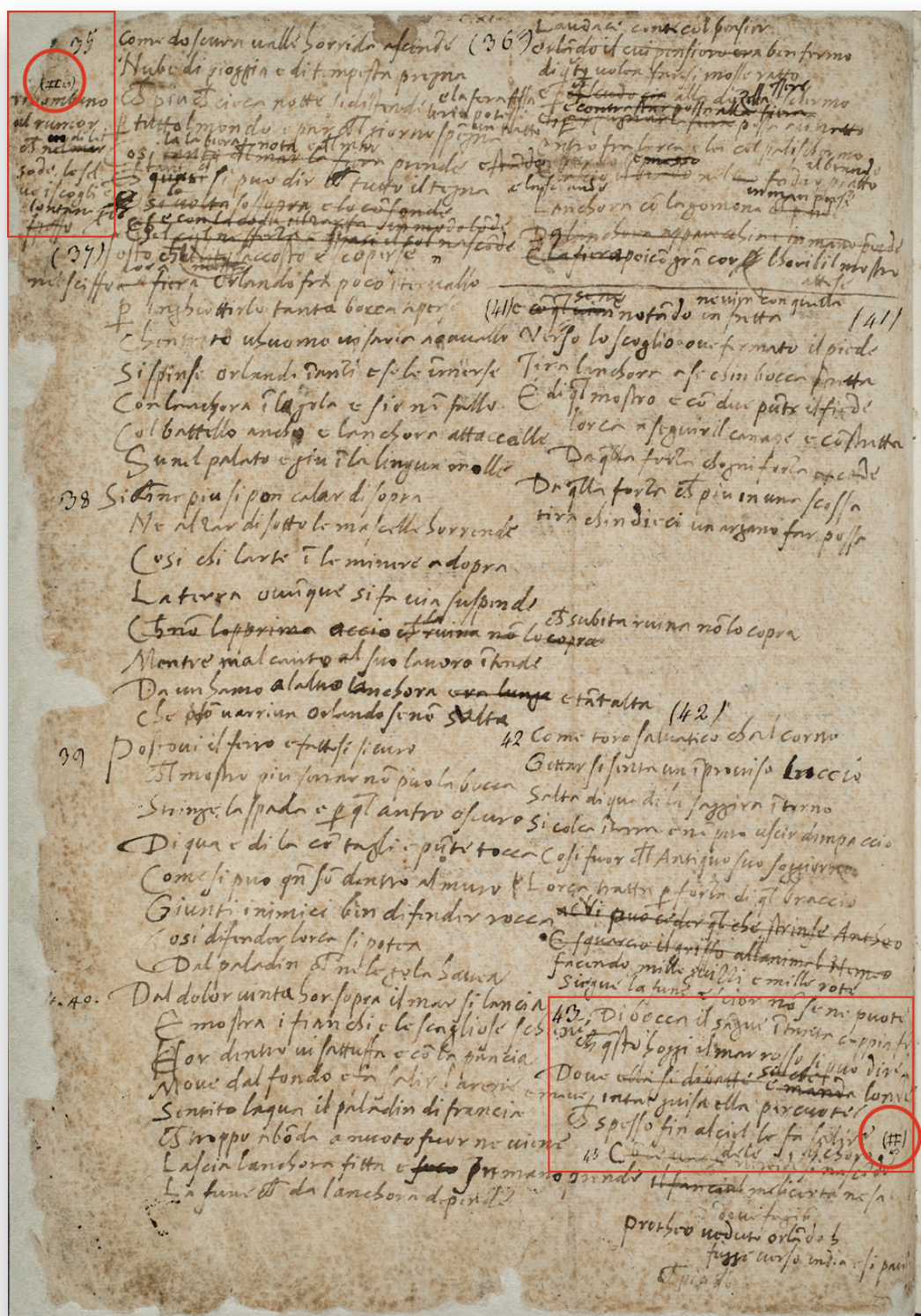


Figura 1.6 – c. 32 v.: XXXVII 90, 1

32
89
La cunivione di colui fu onta
Dal rapace Tanagro. come lo detto
Et a colui di poi data l'impresa
Di gl'acqua et de il crudel offeso
E no era ita come l'altre i chista
Et di gl'et loqui spino i sospetto
Anzi i quel tempo de la porta usata
Dove esser credon salui era fuggita
Hauuto nunquam poi di lei spira
Ch'esse ~~venuta~~ ^{venuta} ~~non~~ ^{non} ~~era~~ ^{era} ~~fin~~ ^{fin} ~~ostentata~~ ^{ostentata} ~~lo~~ ^{lo} ~~qual~~ ^{qual} ~~era~~ ^{era} ~~ridotta~~ ^{ridotta} ~~i~~ ⁱ ~~ostentata~~ ^{ostentata}
(92) ~~come il grā fiume et di~~ ^{Non ha cospato mai di cercar via}
~~refuso esser~~ ^{Come i man Babilin accorrendo}
~~quanto piu innanzi e verso il~~ ^{Et finalmente l'annata via}
~~mar di fondo~~ ^{Alta da doni e da proferte ricche}
~~et colui l'ombra et ricin~~ ^{Hafuto ch'ha barto et afficuro}
~~si misse~~ ^{La donna i sua terra a Margano}
~~gl'adri e gli altri unde~~ ^{Sanctificata e non ha potestà}
~~in l'ora prende~~ ^{Dona parola e fura cossa dirota}
~~quanto piu altiro e impetoso~~ ^{Sopra il somer portar fin a cossina}
~~misse~~ ^{Latte e la ben come la murea fissa}
Onde ~~se~~ ^{poi} ~~sta~~ ^{sta} ~~giu~~ ^{giu} ~~ad~~ ^{ad} ~~passata~~ ^{passata}
cosi ~~per~~ ^{per} ~~quella~~ ^{quella} ~~piu~~ ^{piu} ~~colpe~~ ^{colpe} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
~~nessa intanto~~ ^{Quasi condotta e del grā d'ombra}
di ~~marginor~~ ^{marginor} ~~colpe~~ ^{colpe} ~~due~~ ^{due} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
~~giu~~ ^{giu} ~~per~~ ^{per} ~~quella~~ ^{quella} ~~piu~~ ^{piu} ~~colpe~~ ^{colpe} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
~~si~~ ^{si} ~~gl'adri~~ ^{gl'adri} ~~e~~ ^e ~~gli~~ ^{gli} ~~altri~~ ^{altri} ~~unde~~ ^{unde} ~~in~~ ⁱⁿ ~~l'ora~~ ^{l'ora} ~~prende~~ ^{prende}
~~piu~~ ^{piu} ~~di~~ ^{di} ~~giu~~ ^{giu} ~~colpe~~ ^{colpe} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
90
La cunivione di colui fu onta
Dal rapace Tanagro. come lo detto
Et a colui di poi data l'impresa
Di gl'acqua et de il crudel offeso
E no era ita come l'altre i chista
Et di gl'et loqui spino i sospetto
Anzi i quel tempo de la porta usata
Dove esser credon salui era fuggita
Hauuto nunquam poi di lei spira
Ch'esse ~~venuta~~ ^{venuta} ~~non~~ ^{non} ~~era~~ ^{era} ~~fin~~ ^{fin} ~~ostentata~~ ^{ostentata} ~~lo~~ ^{lo} ~~qual~~ ^{qual} ~~era~~ ^{era} ~~ridotta~~ ^{ridotta} ~~i~~ ⁱ ~~ostentata~~ ^{ostentata}
(93) ~~come il grā fiume et di~~ ^{Non ha cospato mai di cercar via}
~~refuso esser~~ ^{Come i man Babilin accorrendo}
~~quanto piu innanzi e verso il~~ ^{Et finalmente l'annata via}
~~mar di fondo~~ ^{Alta da doni e da proferte ricche}
~~et colui l'ombra et ricin~~ ^{Hafuto ch'ha barto et afficuro}
~~si misse~~ ^{La donna i sua terra a Margano}
~~gl'adri e gli altri unde~~ ^{Sanctificata e non ha potestà}
~~in l'ora prende~~ ^{Dona parola e fura cossa dirota}
~~quanto piu altiro e impetoso~~ ^{Sopra il somer portar fin a cossina}
~~misse~~ ^{Latte e la ben come la murea fissa}
Onde ~~se~~ ^{poi} ~~sta~~ ^{sta} ~~giu~~ ^{giu} ~~ad~~ ^{ad} ~~passata~~ ^{passata}
cosi ~~per~~ ^{per} ~~quella~~ ^{quella} ~~piu~~ ^{piu} ~~colpe~~ ^{colpe} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
~~nessa intanto~~ ^{Quasi condotta e del grā d'ombra}
di ~~marginor~~ ^{marginor} ~~colpe~~ ^{colpe} ~~due~~ ^{due} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
~~giu~~ ^{giu} ~~per~~ ^{per} ~~quella~~ ^{quella} ~~piu~~ ^{piu} ~~colpe~~ ^{colpe} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
~~si~~ ^{si} ~~gl'adri~~ ^{gl'adri} ~~e~~ ^e ~~gli~~ ^{gli} ~~altri~~ ^{altri} ~~unde~~ ^{unde} ~~in~~ ⁱⁿ ~~l'ora~~ ^{l'ora} ~~prende~~ ^{prende}
~~piu~~ ^{piu} ~~di~~ ^{di} ~~giu~~ ^{giu} ~~colpe~~ ^{colpe} ~~che~~ ^{che} ~~di~~ ^{di} ~~se~~ ^{se} ~~ip~~ ^{ip} ~~quella~~ ^{quella}
94 Ma prima librar la donna e somsto Carco alla rana. e qm piu si unde
che fin condotta da quei bini a morte de l'v. sicur dal cacciar la p.
L'atar di briglia e s'pone all'finco presto e da li cam a truar per si no
fice a posti di stitir far la uir corse getta la soma e doue appar m.
La mesnada assalita mai di questo
No s'bbe incontru i uita sua piu forte la maccata i anti
Siet sa di grata di lasciar i piedi
E la donna e l'armis e fuggir nudi
no furimmo adach spodi
giu m'psti no fur gli a
et li fuffo q'balni ad afflar

Figura 1.7 – c. 35 v.: 122, 1

35.

120 Tanto qui s'indugia et marphisa
 faccia per la sua legge la colera
 Contraria a quella et in uita incisa
 Amore et ignominia dopoi dopa
 Da questa amara compagnia diuola
 Resta quella che dopo s'ama di pona quella d'gloriosa p'riser la zona
 Che copre i corte obubrio prima
 Se no s'adonna e riputazione i primos seno / p'riser e adna come prima

121 Quin li male ^{allora} e Mayanorre
 Di li n'ro i potere et esse poi
 p'et no s'bruin i qual'et modo a liore
 E done e cunthor di nouo annoi
 Lo fex u' ^{di giorno} alzar piu dura torre
 Che no il se i ^{maggior} salto a giorni suoi
 Ma no li dica di li p'iusi parti ^{ella sua amara et la sua e s'p'ntis} No p'us d'li n'ro d'li suoi si parli

122 Ma Della tua compagnia et da uoso Arli
 Tutto el giorno i fumi e labro appresso
 Hora di ter la andaro: e poi et furo
 Giunti d'ore i due strada e il camin fesso
 L'un a un al campo e l'altra d'arli al muro
 Tornar gli ananpi ad abbracciarsi i fesso
 E tor ~~l'argento~~ e sempre acerbo e duro ^{combinato}
 Dur se ne andaro le due Al fin le due al campo e puggir ito
 I son al te furo ma il canto a qui finito
 Per le pontelle al campo e puggir ito
 Et in Arli al fine ~~il canto~~ qui finito
^{p'us d'li n'ro d'li suoi si parli}
~~p'us d'li n'ro d'li suoi si parli~~
 al suo re in Arli /

Al fin le done

Figura 1.8 – c. 22 v. «Quel che d'Orlando ecc. (segue dopo questi due fogli bianchi mal intromessi)»

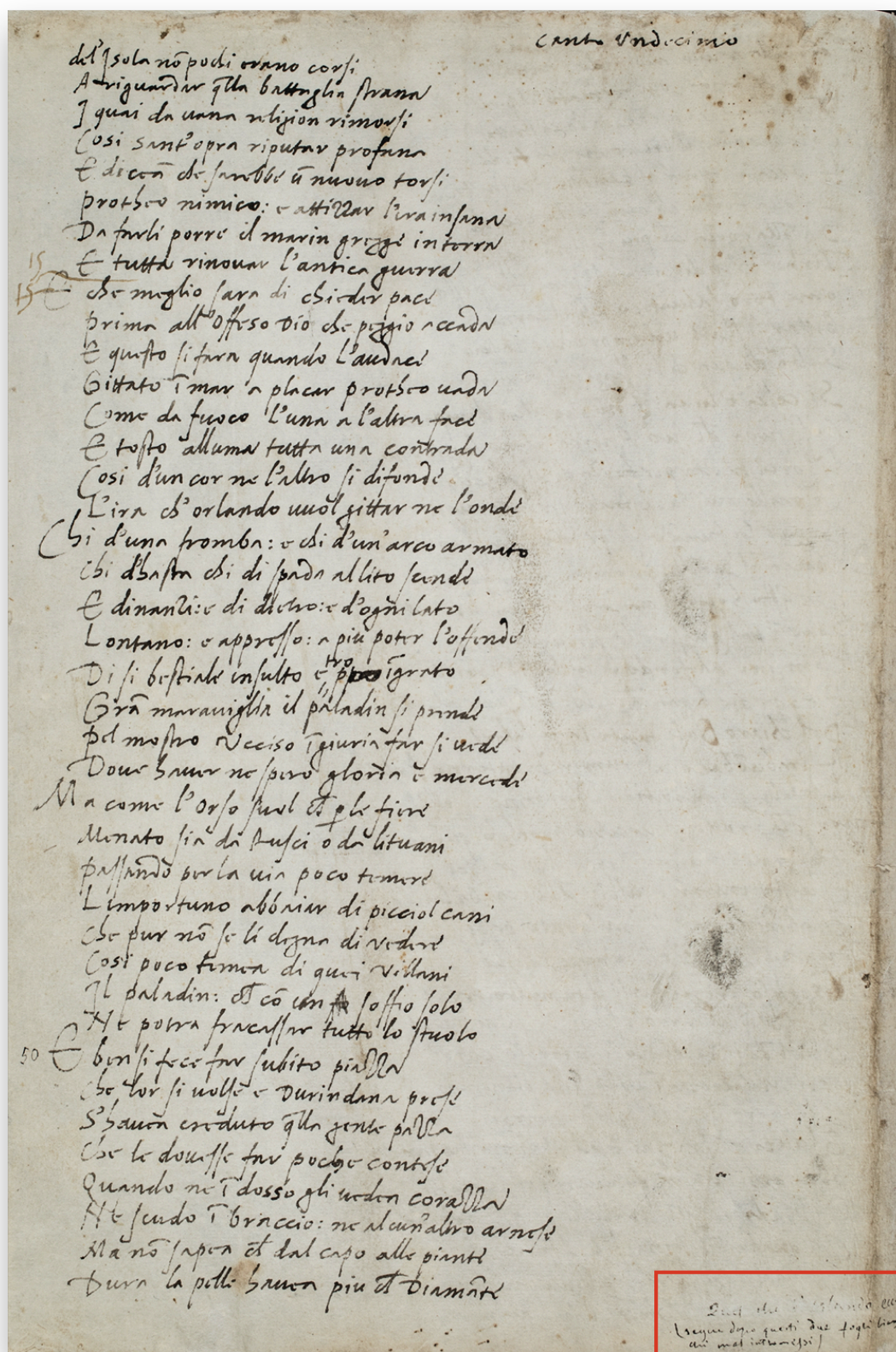


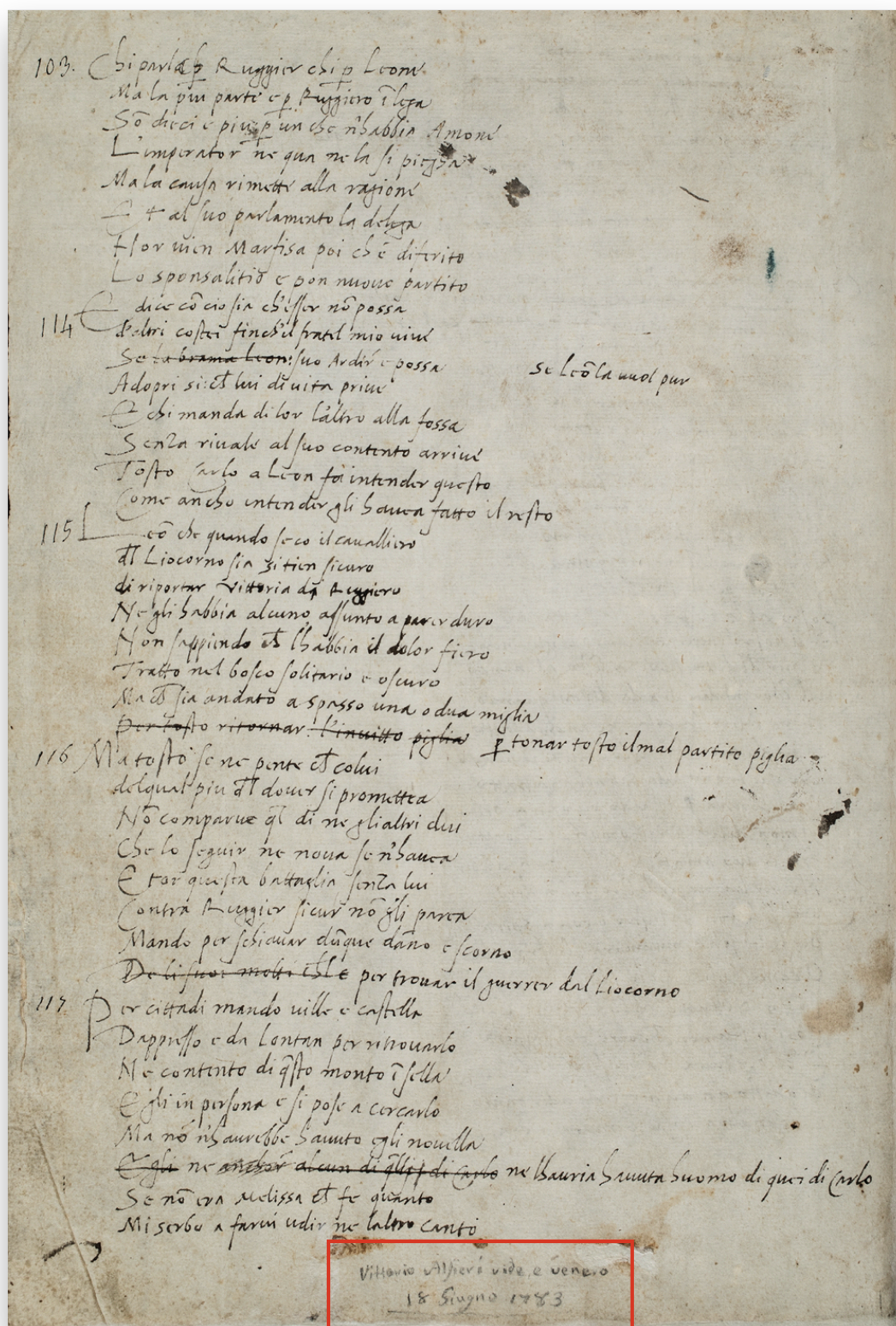
Figura 1.9 – c. 43 v. «Segue dopo i due fogli bianchi»

[illegible]

Figura 1.10 – c. 44 r. «Dr. Barotti donò 1759»

canto XLV
 44 16 J. no mi leuaro da questi piedi
 disse ella signor mio: se ti fello
 E' uccise il mio figliuol no mi co'icai
 di uindicar lor et habbia prigione
 Olte che stato ti. n'ipote uedi
 quanto t'amo ue di quito opre buone
 Ha porte tutto e uedi slaurai torro
 Se uendetta no fai di chi l'ha morto
 di no lo uindicar di chi l'ha morto
 17 Vidi E' co' fortuna ch'ella anebax si dolse
 uist'lonarix) ti do da poter di farla et ti suadi a farlo (17)
 E' seguio di conto: benet co'le
 tre uolte e piu l'imperator le uolte
 Et la prima das pie no fegli tolse
 E' promesso le fu di contentarla
 Comando constantin ch'el si facesse
 Colui condurre e in mand'li si disse
 18 E' osi ben diange e colsi ben si duole
 e colsi bene et efficece parla
 Ne d'ali piedi tor mai se li muole
 E' co'le tre uolte e quattro di leuarla
 V'ss'li constantino atti e parole
 Che gli e sforzato al fin di contentarla
 E' cosi comando et si facessi
 Colui condurre e in ma' di lei si disse
 19 E' per no' far i'pacio lunga dimora
 fu' condotto il guerrer dal Liocorno Sine
 e dato i' mano alla crudel Theodora
 Che no' ui fu intervallo piu di giorno
 il far che sia squartato uiu e morto
 publicamente co' obbro e scorno
 poca pena a lei par e studin e pensa
 altra trouar' alcuna i'usitata e imensa
 20 La femina crual lo fece porre
 In catenato e mani e piedi e collo
 nel tenebroso fondo duna torre
 doue mai no' e' uero raggio d'Apello
 fuor ch' un poco di per omofino terre
 gli fe ogni cibo e senza ancor l'usello
 Duo di talhora e lo die i' guardin a tal
 E' era di lei piu pre'to a farli male
 o se d'Amon la ualerosa figlia A.
 D. Barotti
 Dono. 1759

Figura 1.11 – c. 53 v. «Vittorio Alfieri vide e venerò 18 giugno 1783»



Debenedetti osserva che la scrittura ariostesca si presenta chiara, ordinata e priva di tratti fortemente individuali. Le pagine sono divise in due colonne: quella di sinistra accoglie il testo,

solitamente cinque stanze; quella di destra, invece, raccoglie le correzioni che non trovarono spazio nell'interlinea o nel margine sinistro della carta (anche quando riguardavano l'incipit del verso), oltre alle aggiunte. Raramente il testo si discosta da questa disposizione, e solo per esigenze particolari. Le carte non recano di norma numerazione autografa, salvo una serie – interrotta da due lacune – in cui ai numeri moderni (apposti in inchiostro blu, sul margine superiore sinistro del recto di ogni carta) corrispondono numeri apposti dallo stesso autore in inchiostro marrone sul margine superiore destro. In questa sequenza risultano tuttavia assenti le cc. 18-19, sulle quali si tornerà più avanti (cfr. Fig. 1.12).

Figura 1.12 – Numerazione moderna e autografa delle carte



È presente inoltre una seconda numerazione, anch'essa autografa, che si rinnova di dieci in dieci stanze, conferendo a questa sezione del manoscritto una particolare rilevanza. Di seguito, se ne riportano due occorrenze a scopo esemplificativo (cfr. Figg. 1.13 e 1.14).

Figura 1.13 – c. 7 (num. autogr. 2) v.: «7 E» (canto IX, st. 18)

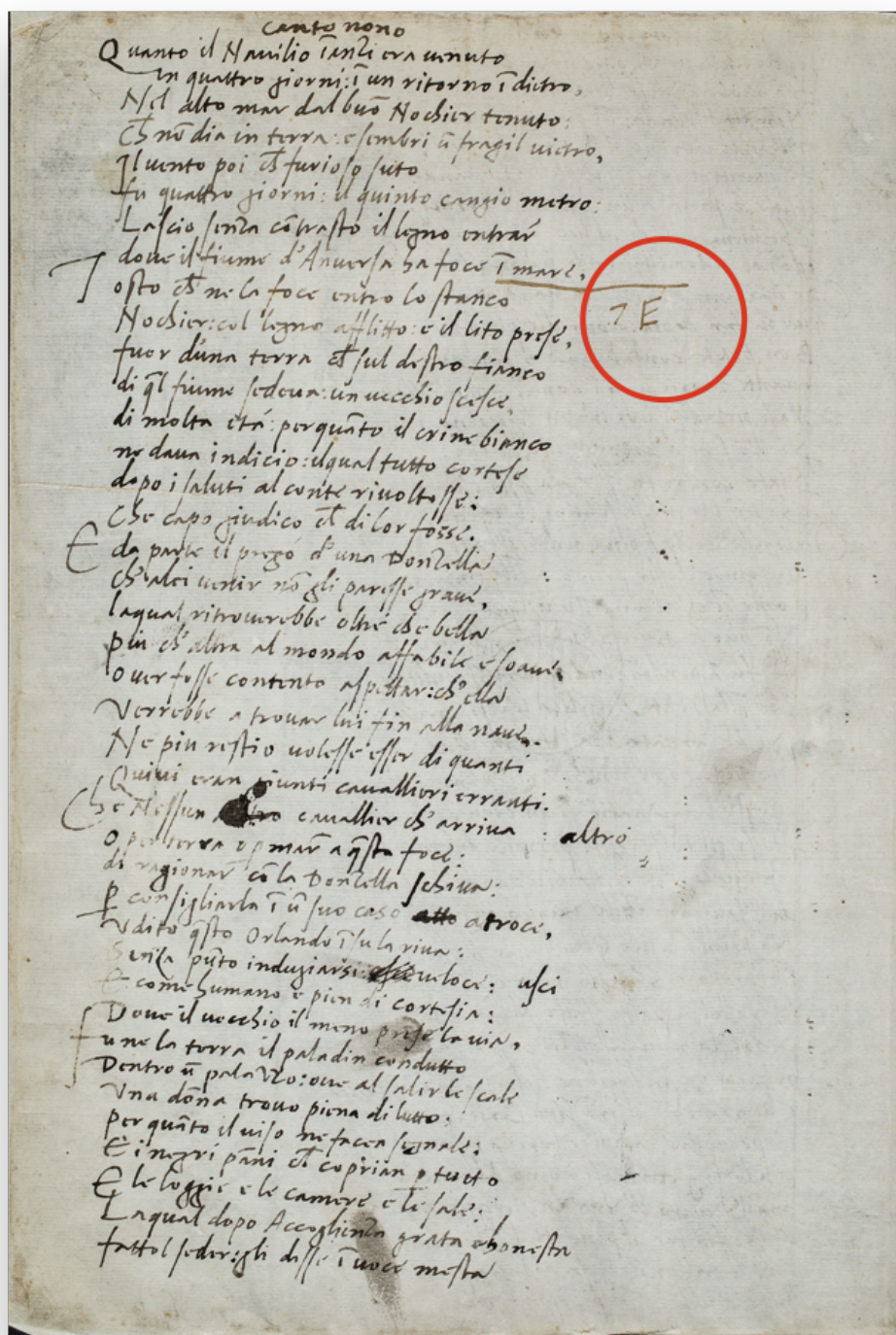


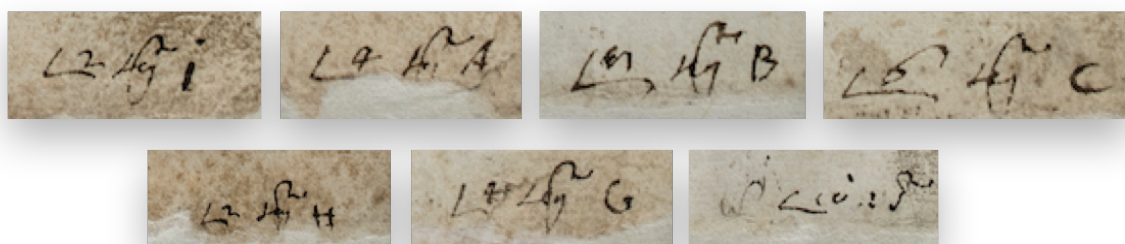
Figura 1.14 – c. 8 (num. autogr. 3) v.: «8 E» (canto IX, st. 28)

Canto nono

Il mio bñ padre: al qual sol piaceva questo
 Iamè pincin: ne mai turbar mi uolse:
 per consolar mi: e far cessar il pianto
 chio ne tra con la pratica di ciolse,
 di et il supbo Re di frisa tanto
 idigno prese: e a tanto odio si uolse:
 Et curò in olanda: e cominciò la guerra
 Et tutto il sangue mio caccia sotto terra.
 O l'or et si robusto: e si possente
 Et poëbi pari a no s'ha ita ritenute.
 E si astuto in mal far: et albrui niente
 La possanza: alardir: l'ingegno: gioua
 porta alcun arme et l'antica gente
 N'io uide mai: ne fuor et alui la noua:
 Un ferro bugio: lungo da dua braccia:
 Dentro a cui polue et una palla caccia.
 Col fuoco: dietro oue la cana e chiusa:
 Poco a spiraglio et si uide a uita,
 a quila et toccò il medico uita
 doue e bisognò d'alleciar la uita.
 Onde non co tal fuo la palla esclusa:
 Et si pur dir et tuona e et balena.
 Ne men et seglia il fulmine oue passa
 cio et tocca arde: abatte: arde: e fracassa
 Pos due uolte il nostro campo i rotta
 Et questo ingegno: e i miei fratelli uccise.
 Nel primo assalto il primo: et la botta
 rotta l'isburgo in mille il cor gli mise:
 Ne l'altra uolte a l'altra et el re in preta
 fugga: dal corpo l'anima diuise.
 E lo feri l'onta dietro la spalla:
 E fuor et posto uscir fece la palla,
 Difendendosi poi mio padre a giorno
 Dentro a castel et sol gli era rimaso:
 Et tutto il resto hanno perduto ritorno:
 Lo fe co simil colpo ire all'ocaso.
 Che mentre andaua et et fecea ritorno
 promouendo sur a qsto sur a qto caso,
 Dal traditor fu i mille gli occhi calto
 Et hanno di loro di mira tolto,

Infine, si incontrano alcune numerazioni di difficile interpretazione, riportate a piè di pagina sul recto delle cc. 3, 5, 7, 15, 19, 23 e 44: Debenedetti le ritiene antiche, ma probabilmente non attribuibili ad Ariosto (cfr. Fig. 1.15).

Figura 1.15 – Numerazione ignota

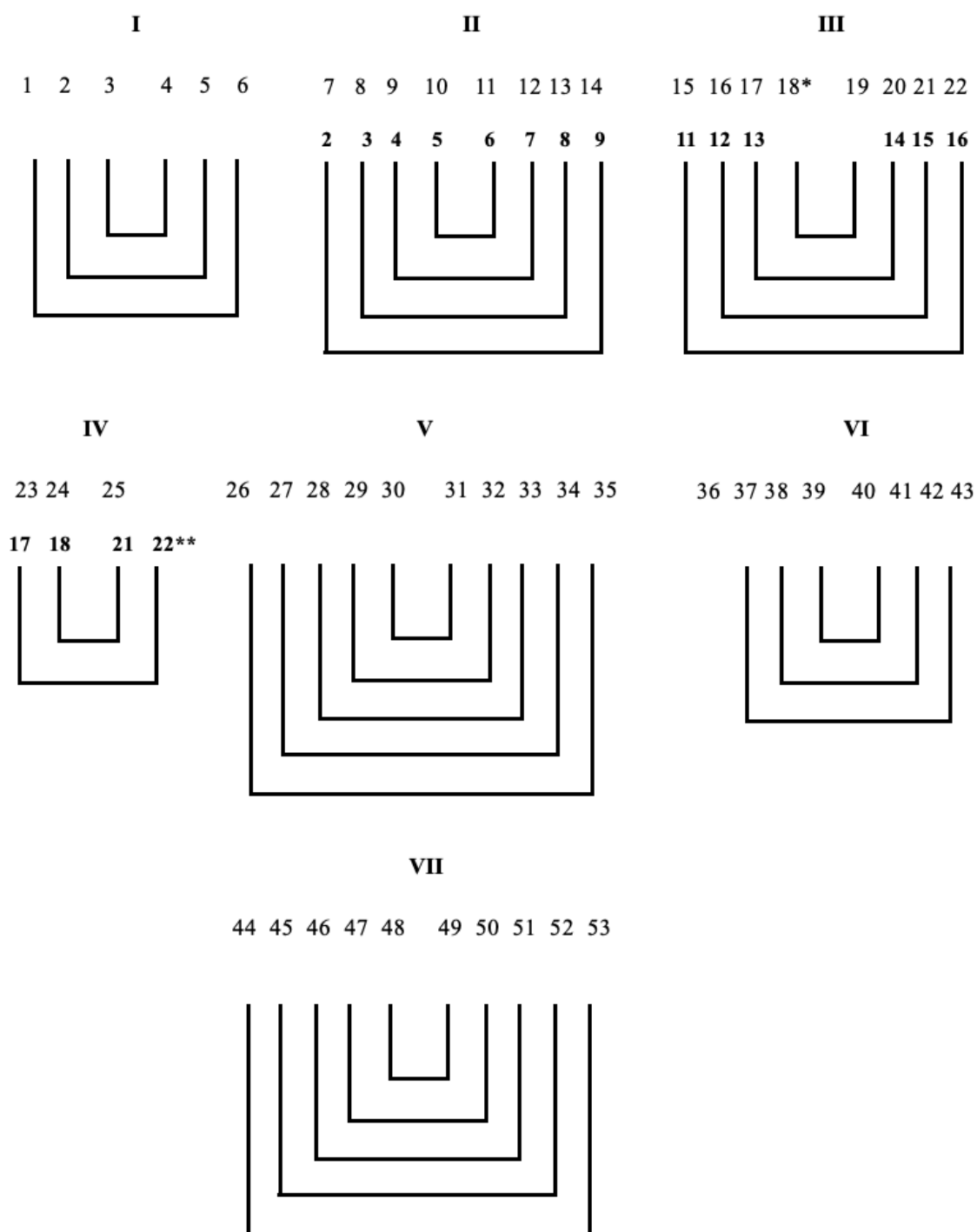


Le 54 carte superstiti risultano ripartite nei fascicoli nel modo seguente:

- Fascicolo I: cc. 1-6;
- Fascicolo II: cc. 7-14;
- Fascicolo III: cc. 15-22;
- Fascicolo IV: cc. 23-25 e c. bianca non numerata;
- Fascicolo V: cc. 26-35;
- Fascicolo VI: cc. 36-43;
- Fascicolo VII: cc. 44-53;

L'ordinamento dei fascicoli è quello originario, fatta eccezione per una discrepanza individuata e segnalata da Debenedetti. Nell'infografica che segue vengono riportate entrambe le cartolature: quella autografa, parziale e resa in grassetto, e quella moderna (cfr. Fig. 1.16).

Figura 1.16 – Ordinamento dei fascicoli



* Le cc.18 e 19, che ora sono staccate, formavano un solo foglio, sciolto dal fascicolo.

** Questa carta è bianca.

Figura 1.17 – c. 18 recto e verso

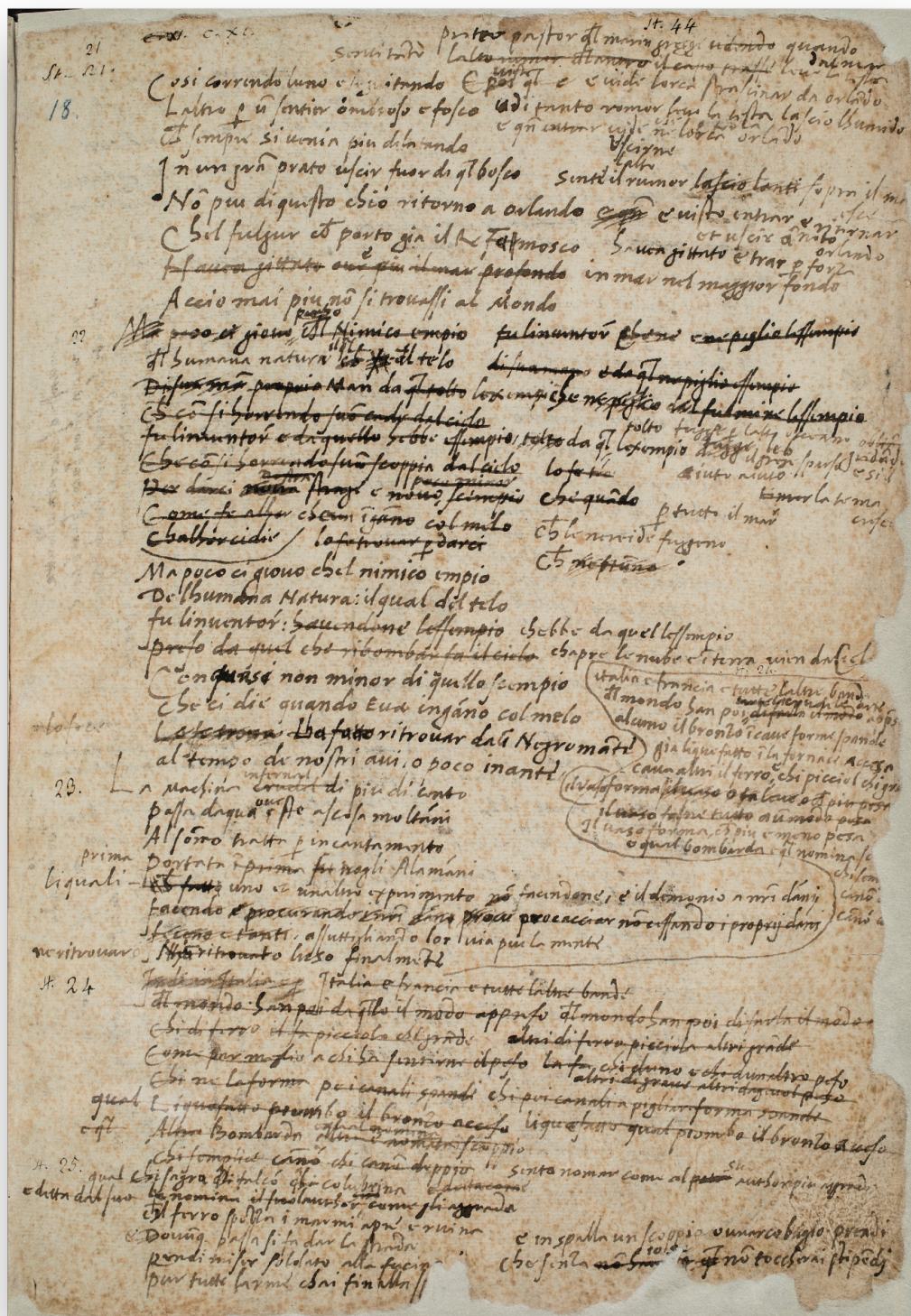
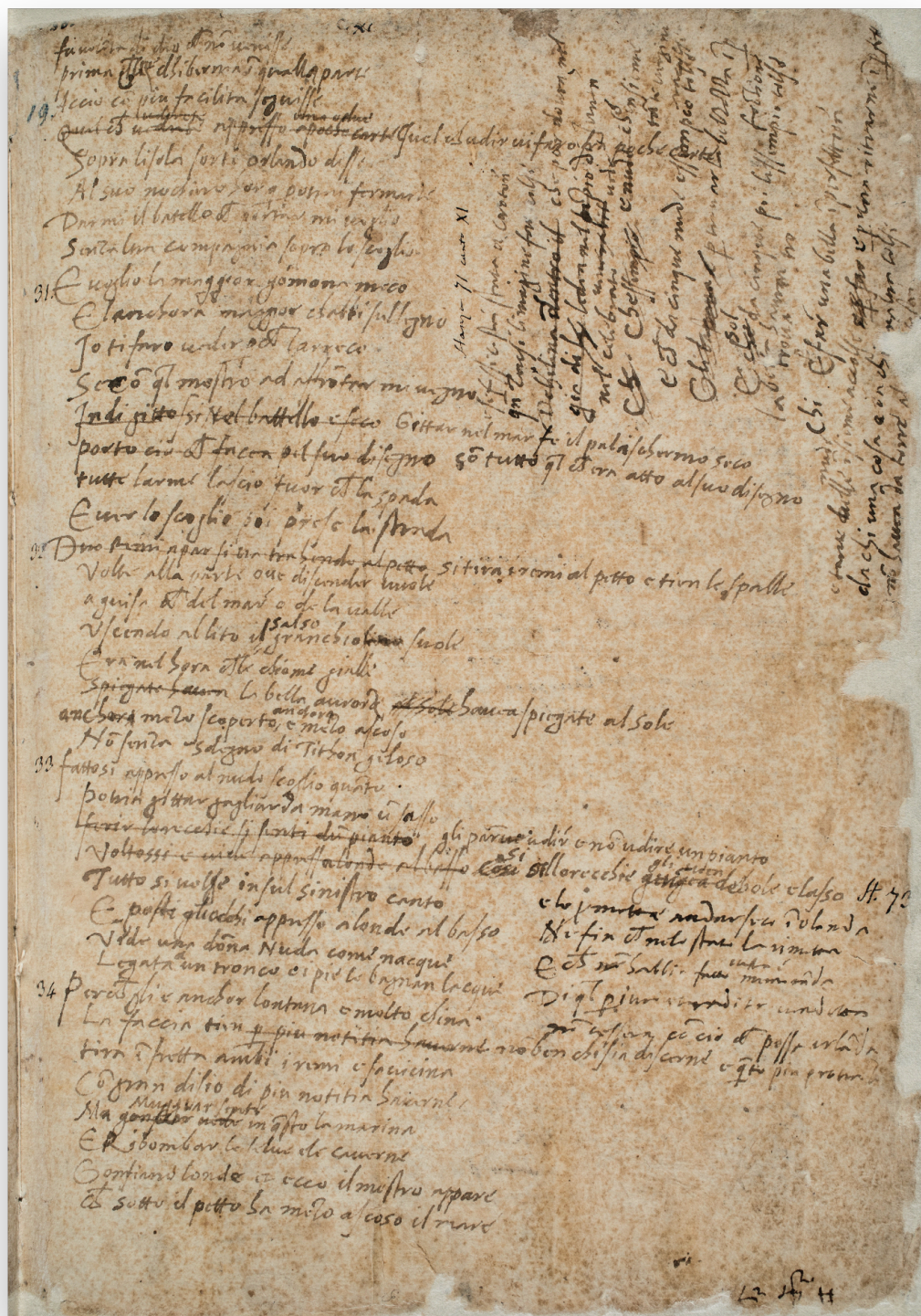


Figura 1.18 – c. 19 recto e verso



35 Com' de pueri ualli barida ascende (36) L'andrea concul possem
 Nubi di pioggia e di neppa p'grina
 36 L'andrea concul possem
 37 L'andrea concul possem
 38 L'andrea concul possem
 39 L'andrea concul possem
 40 L'andrea concul possem
 41 L'andrea concul possem
 42 L'andrea concul possem
 43 L'andrea concul possem
 44 L'andrea concul possem
 45 L'andrea concul possem
 46 L'andrea concul possem
 47 L'andrea concul possem
 48 L'andrea concul possem
 49 L'andrea concul possem
 50 L'andrea concul possem
 51 L'andrea concul possem
 52 L'andrea concul possem
 53 L'andrea concul possem
 54 L'andrea concul possem
 55 L'andrea concul possem
 56 L'andrea concul possem
 57 L'andrea concul possem
 58 L'andrea concul possem
 59 L'andrea concul possem
 60 L'andrea concul possem
 61 L'andrea concul possem
 62 L'andrea concul possem
 63 L'andrea concul possem
 64 L'andrea concul possem
 65 L'andrea concul possem
 66 L'andrea concul possem
 67 L'andrea concul possem
 68 L'andrea concul possem
 69 L'andrea concul possem
 70 L'andrea concul possem
 71 L'andrea concul possem
 72 L'andrea concul possem
 73 L'andrea concul possem
 74 L'andrea concul possem
 75 L'andrea concul possem
 76 L'andrea concul possem
 77 L'andrea concul possem
 78 L'andrea concul possem
 79 L'andrea concul possem
 80 L'andrea concul possem
 81 L'andrea concul possem
 82 L'andrea concul possem
 83 L'andrea concul possem
 84 L'andrea concul possem
 85 L'andrea concul possem
 86 L'andrea concul possem
 87 L'andrea concul possem
 88 L'andrea concul possem
 89 L'andrea concul possem
 90 L'andrea concul possem
 91 L'andrea concul possem
 92 L'andrea concul possem
 93 L'andrea concul possem
 94 L'andrea concul possem
 95 L'andrea concul possem
 96 L'andrea concul possem
 97 L'andrea concul possem
 98 L'andrea concul possem
 99 L'andrea concul possem
 100 L'andrea concul possem

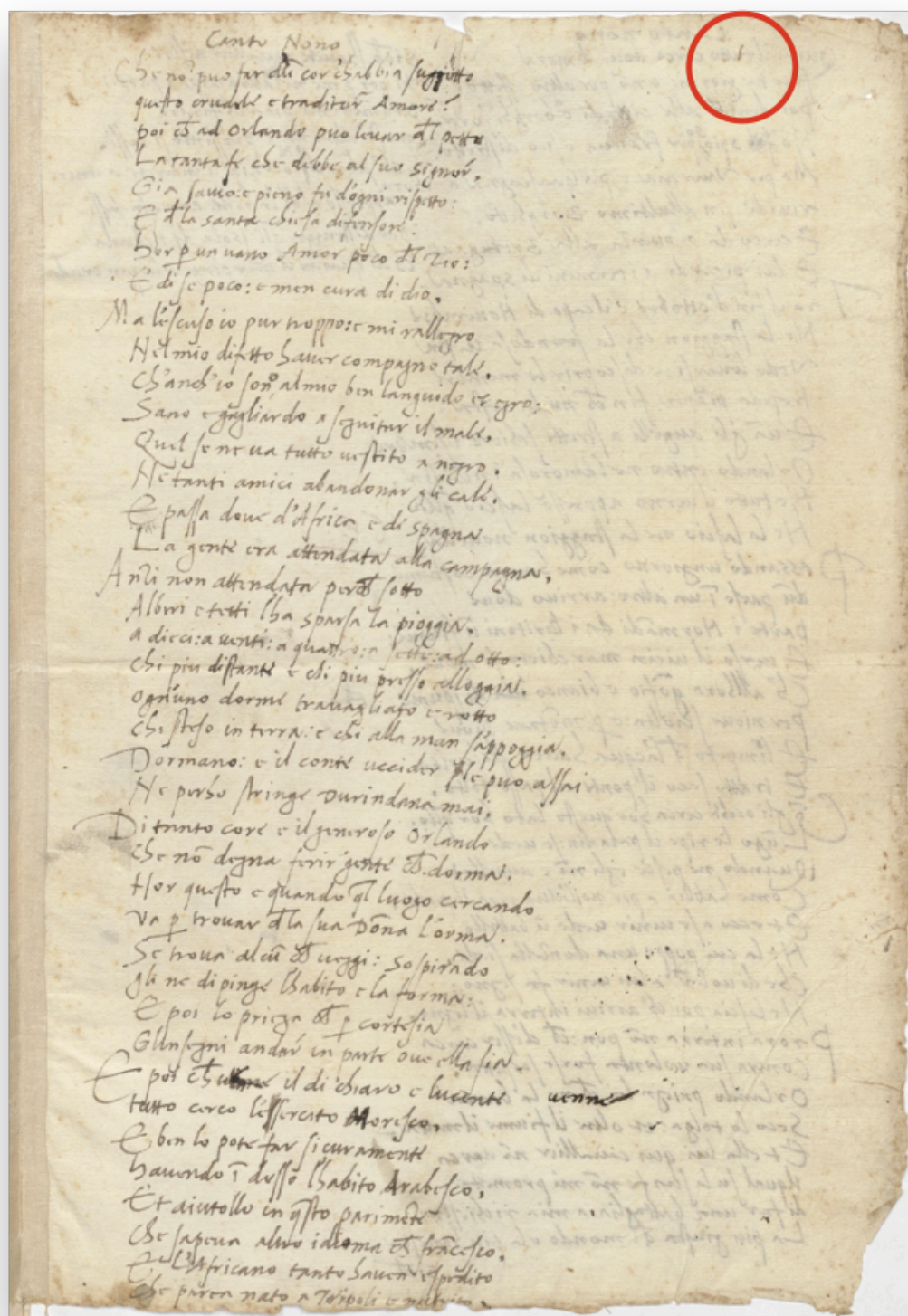
La suddetta discrepanza emerge nel fascicolo III: le carte centrali (numerazione moderna 18 e 19) non presentano numerazione autografa (cfr. Figg. 1.17 e 1.18), mentre questa procede senza interruzioni dalle prime tre carte alle ultime tre. È evidente quindi che le carte furono erroneamente inserite nel fascicolo. A confermarlo concorrono diversi elementi, tra cui lo stadio di elaborazione del testo e un particolare significativo: lo sgorbio visibile sul margine superiore di c. 17 v. corrisponde infatti all'impressione a rovescio dell'inchiostro ancora fresco delle ultime lettere della parola "Undecimo" (indicante il canto XI), vergata da Ariosto sul margine superiore di c. 20 r. Si tratta di una prova certa che le due carte erano originariamente contigue. Debenedetti ricorda che anche in altri luoghi del manoscritto (cc. 13 r., 25 r.) vi sono tracce dell'abitudine di Ariosto di voltare velocemente le carte, senza attendere che l'inchiostro fosse asciutto, dopo aver apposto l'intestazione o qualche altro segno. Ne consegue che le cc. 18-19 debbano essere rimosse dall'attuale collocazione e reinserite tra il primo e il secondo fascicolo.

Dopo aver riordinato il manoscritto ferrarese, Debenedetti mise in evidenza la necessità di ricostruirne, per quanto possibile, la composizione originaria, reintegrando le carte mancanti. Le due carte Borromeo conservate presso l'Ambrosiana (M), dello stesso formato di F, recano i numeri autografi 1 e 10, oltre alla numerazione supplementare già ricordata (cfr. Figg. 1.19 e 1.20):

- c. 1 v.: «6 E» (canto IX, st. 8);
- c. 10 v.: «15 E» (canto X, st. 3).

Esse colmano dunque la lacuna individuata nel fascicolo II del manoscritto ferrarese. Il foglio conservato a Napoli (N I-2), anch'esso del medesimo formato, rappresenta invece ciò che resta di un fascicolo successivo all'ultimo di F. Uno degli obiettivi della nostra edizione digitale era quello di ricostruire, almeno in forma virtuale, la composizione originaria del manoscritto ferrarese, integrandolo secondo le indicazioni di Debenedetti; tale intento, tuttavia, non è stato finora pienamente raggiunto a causa della mancata disponibilità delle carte napoletane.

Figura 1.19 – c. 1 recto e verso



Canto nono
Quini il tutto ceros: doue dimora
fecce ho giorni: e nò per altro offito.

poi dentro alle cittadi e borghi forn

No fte spio, pòr francia e suo distretto.

Ma per Vuerria e per Guascogna andorn

riuide in allultimo Borgsetto.

E ceros da prouincia alla Bretagna:

E dai picur di a i termini di spagna.

Tra il fin d'ottobre e il capo di Homprebre

Ne la spaggion che la profonda iusta

Vede leuarsi: e di coprir le monbre

tropida pienta fin es nuda iusta:

E un gli augelli a frotte schiere i fombre

Orlando entro ne lamorosa in cista

Ne tutto il uerno appresso lascio glio:

Ne la lascio ne la spaggion nouella.

Passando un giorno come sauer costumi

du parte i Hermadi da i Britani u fiumi:

E uerso il uicin mar chieto si moue.

Es alloro gofio e bianco uia di spume

per nicur sciolta: e p motane piodu.

E limpeto d lacqua sauer di ciolto

E tratto seco il ponte cil passo tolto.

Cogli occhi circa sor questo lato sor glio.

Lago le ripe il paladin: se uide

Quando ne pise e gli nò e augello)

Come labbia a per ne l'ultima ripe il piede.

Et ecco a se uenir uide u battello

Ne la cui puppe una donella fide.

Che di uolere a lui uenir fa segno:

Ne lascia poi es arrui interra il legno.

Prore interra nò pun: et disse carca

Contra sua uolenta forse sospetto.

Orlando prigo lei et oue la baren

Seco lo tolga: et olu il fiume il metta.

Et ella lui qui cauallier nò uaren

il qual su la fure nò mi prometta

di far una battaglia a mia richiesta:

La piu giusta d mondo e la piu sanesta

Siet haute cauallier di fire

di per p me nialtra ripa i passi:

promettermi pma et finire

questo altro uide prossimo si lassì.

Es alu d'indernia uandarete a uini

appresso alfl la bella armata fessi

per distruggere gli Isola d'hibuda

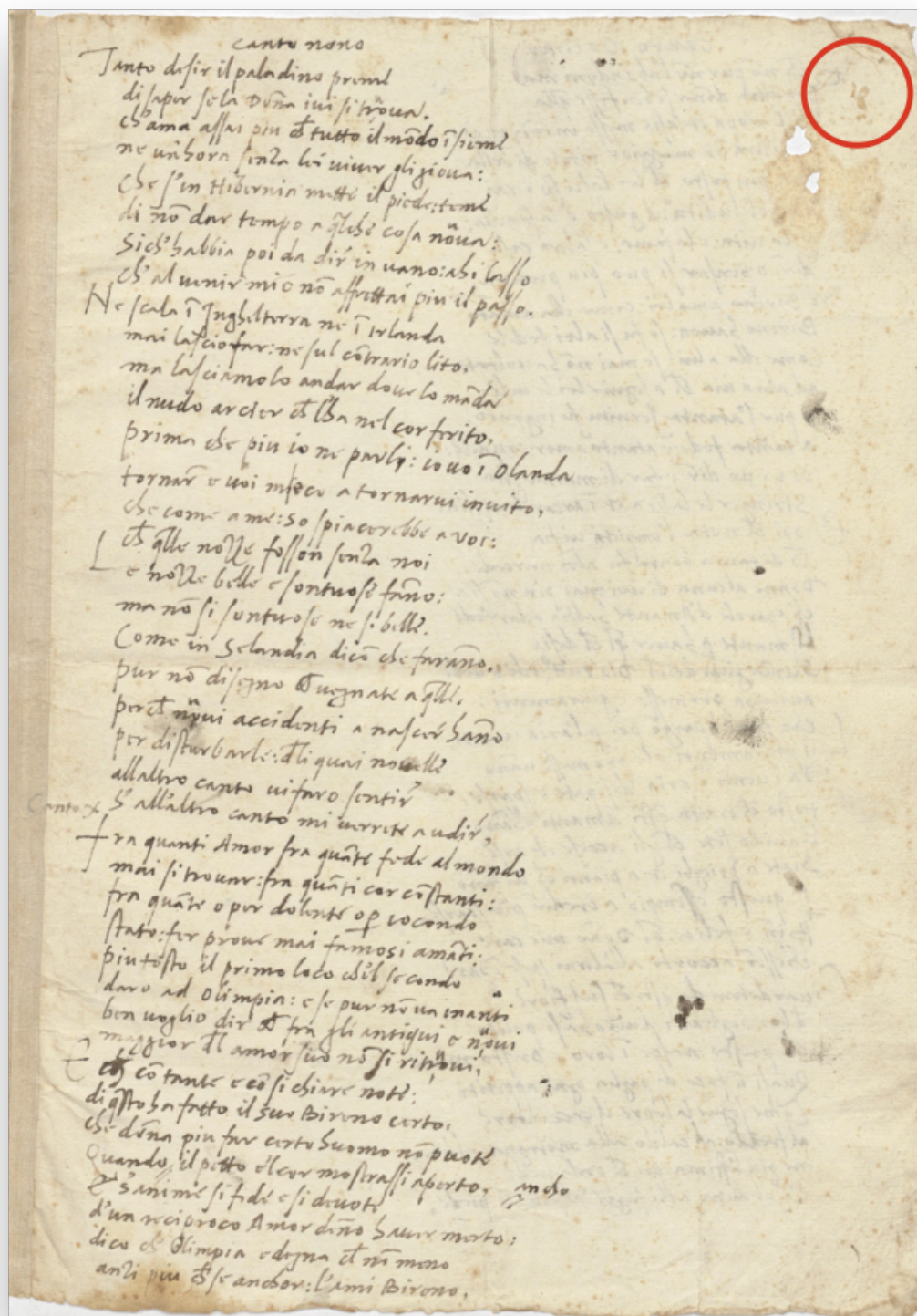
Es di quante il mar cinge: e la piu cruda.

CE

cta

Siet

Figura 1.20 – c. 10 recto e verso



canto De como
E ch'io pur n'abandoni mai
per altri d'anni: se ben fosse q'la
di Europa et alia misse in tante guai:
altra ha maggior titolo di bella:
ma piu fusto et lei lasci so i rai:
al sol: ludita: il gusto: et la fucilla:
et la vita: et la fama: et l'altra cola
dir o pensar si puo piu p'curarla:
Se Bireno ama lei: come ella amato
Bireno hauea: se fu si alui fedele
Come ella a lui: se mai no ha uoltrato
ad altra uia et a seguir lei le uelle:
o pur l'atanta seruita fu ingrato:
a tanta fede: et tanto amor: crudele:
io ui uo dir: et far di marauiglia
Stringer le labra et tancar le ciglia.
E poi et resta l'impista uia fin
et di tanta bonta fu alui mercede:
Donne alcuna di uoi mai piu no tin
et a parole d'Amante babbia adorfede
LAmante p' bauer q' et desia
S'entra giurda et Dio tuu' odo et uidi
auulupa promisse: et giuramenti:
[Che tutti soango poi p' laria iurati,
i giuramenti et le promisse: unno
da i uenti et aria dissipate et sparse:
tosto et cratta q'li amanti banno
L'uida se et gli accese et aspi:
Siate a prieghi et a pianti et u'anno
p' questo s'empio a cridar piu scarsi:
Ben e felice et done mic care
E' offer accorto all'altrui spio: et pare
Guarda: tui da q'li et sul fiori
et lor buglianti il uiso ha si polito:
et p'usto nasce i loro et p'usto more
Quasi a face di puglia ogni apportito.
Come segue la lore il cacciatore
al preda: al caldo alla mangiara al lito:
ma piu l'istima poi et porta uide
E' sol dietro adli fuggi affrettu i piedi,

1.3.2 I nuovi episodi aggiunti: trama, collocazione nell'impianto del poema e attestazione negli autografi ariosteschi

L'episodio di Olimpia si sviluppa nelle seguenti ottave del poema: IX, 8-94 (fine); X, 1-34; XI, 21 (vv. 5-8)-80; XII, 1-4 (vv. 1-6). Olimpia, figlia del conte d'Olanda e innamorata perdutamente del duca di Selandia Bireno, è chiesta in sposa da Cimosco, re di Frisa, per il figlio Arbante. Al suo rifiuto, Cimosco devasta l'Olanda grazie alla sua nuova arma, l'archibugio; Orlando interviene allora a liberare la terra e ristabilire la pace. La giovane sembra finalmente destinata a coronare il suo amore, ma Bireno si invaghisce della figlia di Cimosco e abbandona Olimpia su un'isola deserta, dove viene catturata dagli abitanti ed esposta a un'orca. Ancora una volta l'arrivo di Orlando si rivela provvidenziale e garantisce il lieto fine. L'episodio assolve a una duplice funzione: da un lato, riequilibrare la distribuzione dello spazio narrativo tra Ruggiero e Orlando, originariamente sbilanciata a favore del primo; dall'altro, posticipare l'avvio della ricerca amorosa di Orlando, che giungerà al castello di Atlante soltanto dopo aver portato a compimento due imprese eroiche. Ariosto, in tal modo, arricchisce uno dei principali fili narrativi del poema, facendo emergere come nucleo centrale il percorso di iniziazione dell'eroe eponimo.

L'episodio della rocca di Tristano si colloca nelle seguenti ottave: XXXII, 50-110 (fine); XXXIII, 1-59 e 65-76. La rocca costituisce uno spazio regolato da una norma particolarmente crudele, istituita da Tristano per punire l'eccessiva gelosia di Clodione nei confronti della propria compagna: essa prevede che il castello possa accogliere un solo ospite per volta e di un solo sesso, salvo il caso in cui giungano più persone contemporaneamente, nel qual caso tutti sono ammessi. Se l'ospite è un cavaliere, deve affrontare in duello il precedente occupante; se invece è una donna, la contesa assume la forma di una gara di bellezza. La tensione narrativa raggiunge il suo apice con l'arrivo di Bradamante: presentatasi dapprima in abiti maschili, ella vince il duello contro il cavaliere di guardia; ma, una volta rivelatasi donna, mette in discussione le convenzioni legate all'identità di genere.

L'episodio di Marganorre si sviluppa a partire dall'ultima ottava del canto XXXVI (v. 84) e prosegue per l'intero canto XXXVII (ottave 1-122). Marganorre vi è rappresentato come un tiranno spietato che, per vendicare la morte dei propri figli – entrambi caduti a causa di passioni amorose nutrite per donne già maritate – ha imposto un editto disumano: tutte le donne della sua città devono essere bandite, non prima però di subire l'umiliazione dello “scorciamento” della gonna, così da esporre pubblicamente ciò che dovrebbe restare celato. L'intervento di

Bradamante e Marfisa si rivela decisivo per porre fine a tale barbaro costume e ristabilire l'ordine.

L'episodio di Ruggiero e Leone si sviluppa nelle seguenti ottave: XLIV, 12-14 e 36-104 (fine); XLV, 1-117 (intero canto); XLVI, 19 (vv. 7-8)-66, 69-72. Quando ormai tutto sembra pronto per il matrimonio tra Bradamante e Ruggiero, sopraggiunge un nuovo ostacolo: Amone, padre della giovane, decide di concederla in sposa a Leone, figlio dell'imperatore d'Oriente Costantino, ritenendo che, pur essendo Ruggiero un guerriero di straordinario valore, egli non disponga né di titoli nobiliari né di possedimenti. Bradamante, angosciata, dichiara che accetterà come marito soltanto colui che riuscirà a vincerla in duello. Ruggiero, deciso a colmare la propria mancanza di prestigio, si reca allora in Oriente e si schiera al fianco dei Bulgari contro l'imperatore, conducendoli alla vittoria con imprese straordinarie che suscitano l'ammirazione di Leone, ignaro della sua identità. Catturato in un'imboscata e successivamente liberato proprio da Leone, Ruggiero, per riconoscenza, si impegna a combattere in sua vece contro Bradamante, affinché il principe possa ottenerne la mano. Dopo averle resistito in duello, la consegna a Leone e si ritira in solitudine. Nel frattempo, Marfisa interviene presso Amone per rivendicare la legittimità della promessa di nozze già esistente tra Ruggiero e Bradamante; tuttavia Amone resta irremovibile, stabilendo che Ruggiero potrà sposare la figlia soltanto se riuscirà a sconfiggere Leone in combattimento. Ritrovatolo, Leone induce Ruggiero a confidarsi e viene così a conoscenza della sua vera identità: comprende che il cavaliere da lui liberato e colui che aveva combattuto in sua difesa erano la stessa persona. Riconosciuto il sincero amore di Ruggiero per Bradamante, Leone decide infine di ricambiare la cortesia ricevuta, rinunciando al matrimonio.⁴⁹

E importante sottolineare che le giunte non vanno intese come blocchi narrativi autonomi, giustapposti a una trama già compiuta, poiché la loro collocazione risponde a una precisa strategia compositiva dell'autore. Esse instaurano, da un lato, una continuità diegetica con i

⁴⁹ Per lo studio delle fonti dell'*Orlando Furioso* è imprescindibile il volume di Pio Rajna, *Le fonti dell'Orlando Furioso* (Firenze, Sansoni, 1876), che ha segnato una tappa fondamentale negli studi ariosteschi. L'autore ricostruisce con grande rigore filologico la trama delle tradizioni letterarie, storiche e cavalleresche che hanno nutrito l'immaginazione di Ariosto, delineando un quadro complesso che va dalle chansons de geste francesi ai romanzi cavallereschi italiani, includendo opere come il *Morgante* di Pulci e l'*Innamorato* di Boiardo, oltre a numerosi altri testi della tradizione europea. L'indagine di Rajna, tuttavia, non si riduce a una mera ricognizione erudita: il valore duraturo del suo studio risiede nella capacità di mostrare come Ariosto non si sia limitato a recepire i motivi ereditati, ma li abbia trasformati, adattati e rielaborati, innestandoli nella sua opera con un'originalità poetica capace di rinnovare dall'interno la materia cavalleresca. In questo senso, il lavoro di Rajna non costituisce soltanto un repertorio di fonti, ma anche un modello metodologico: esso coniuga comparazione filologica, attenzione alle strutture narrative e ricostruzione storica, offrendo agli studi successivi uno strumento critico di riferimento imprescindibile per comprendere la genesi e l'originalità del poema ariostesco.

segmenti che le precedono e seguono; dall'altro, una rete di rimandi interni che contribuisce a complicare e, al tempo stesso, ad arricchire sia la macrostruttura del poema sia le sue possibilità interpretative. La giunta di Olimpia, in particolare, sospende la ricerca amorosa di Orlando, appena avviata nel canto VIII, e colloca la liberazione di Angelica da parte di Ruggiero tra le due liberazioni di Olimpia compiute da Orlando.⁵⁰ La duplicazione dell'episodio produce un effetto di specularità che incide non solo sui rapporti di forza tra i due eroi principali, ma anche sulle modalità di lettura del discorso amoroso, in cui all'infedeltà di Ruggiero si contrappone la fedeltà di Olimpia. Analogamente, le giunte della rocca di Tristano e di Marganorre si innestano, sul piano diegetico, nelle vicende di Bradamante, Ruggiero e Marfisa, approfondendo la riflessione sui rapporti tra i sessi e sulle dinamiche fra amore e gelosia; sul piano tematico, esse si configurano come un vero e proprio controcanto all'episodio delle femmine omicide nei canti XVII e XVIII delle prime due redazioni (corrispondenti ai canti XIX-XX).⁵¹ L'ultima giunta, a sua volta, complica e differisce, in termini narrativi, il destino matrimoniale di Bradamante e Ruggiero, mentre sul piano tematico istituisce un legame con il canto XIX delle redazioni A e B (poi XXI), che racconta la vicenda di Filandro e Gabrina.⁵²

⁵⁰ In questi episodi Ariosto rielabora il motivo mitologico della fanciulla offerta in sacrificio al mostro marino, ispirato al celebre mito di Andromeda e Perseo. Nel canto X, Angelica, legata nuda a uno scoglio e destinata a essere divorata da un'orca, viene salvata da Ruggiero che, in groppa all'ippogrifo e armato della lancia d'oro, ferisce la creatura e sottrae la giovane al suo destino. La scena, di forte impatto spettacolare, intreccia l'eroismo cavalleresco con la dimensione fiabesca della meraviglia, consolidando l'immagine di Ruggiero come cavaliere salvatore. In modo analogo, Olimpia, tradita da Bireno e abbandonata su un'isola deserta, viene catturata dagli abitanti e offerta anch'essa in sacrificio a un'orca; questa volta è Orlando a intervenire, sconfiggendo il mostro e liberando la giovane dalla condanna. Ariosto, dunque, attinge a un medesimo archetipo mitico, ma lo declina secondo funzioni narrative e simboliche diverse, adattandolo di volta in volta alle esigenze della trama e alla costruzione dei personaggi.

⁵¹ L'episodio racconta di una comunità di donne che, ribellatesi al dominio maschile, hanno promulgato una legge crudele: ogni uomo che giunga nel loro territorio viene messo a morte dopo essere stato utilizzato per la procreazione. Le amazzoni giustificano tale pratica come forma di vendetta per le violenze subite in passato dagli uomini. Nel corso della narrazione, Orlando e altri cavalieri si imbattono in questa società femminile rovesciata, che rappresenta una radicale sovversione dell'ordine naturale e sociale. L'episodio assume così un valore esemplare e simbolico: da un lato, denuncia gli eccessi di un potere femminile degenerato in tirannide; dall'altro, mette in scena, in forma paradossale, i conflitti tra i sessi e i rischi connessi a una distorta gestione dell'amore e della gelosia.

⁵² L'episodio narra la vicenda di Filandro, giovane cavaliere, che si innamora della vecchia Gabrina, donna astuta e malvagia, e ne diventa schiavo per passione cieca. Per compiacerla, egli arriva perfino a tradire la propria lealtà e a macchiarsi di colpe, ma la donna lo ripaga con inganni e crudeltà. Alla fine, Filandro, vittima del suo stesso accecamento amoroso, viene ucciso, mentre Gabrina sopravvive, incarnando il modello della donna ingannatrice e corrottrice. L'episodio, di tono cupo ed esemplare, si inserisce nel poema come riflessione sulle degenerazioni dell'amore e sulla potenza distruttiva dell'inganno femminile.

Dunque, l'ultimo *Furioso* non si configura come un semplice superamento delle redazioni precedenti, ma come un organismo narrativo che le ingloba e rielabora, precisando e approfondendo alcuni nuclei tematici che, nella prima stesura, risultavano meno centrali o meno pressanti. Tutte e quattro le giunte appaiono infine legate da una fitta rete di corrispondenze interne, articolata attorno a tre nuclei tematici fondamentali: la contestazione dell'autorità costituita, la questione dell'identità sessuale e la fedeltà reciproca. La figura di Marganorre richiama quella di Cimosco nell'episodio di Olimpia, entrambi accomunati dal tratto tirannico e crudele; l'amore illecito e ossessivo dei figli di Marganorre per donne già sposate rinvia alla gelosia patologica di Clodione; l'episodio di Leone, Bradamante e Ruggiero riprende, da un lato, il tema della lotta all'autorità già presente nella vicenda di Marganorre e, dall'altro, quello dell'identità sessuale, centrale nella rocca di Tristano. Risulta evidente, quindi, come le giunte rispondano a un progetto compositivo consapevole, fondato su un calibrato montaggio dei singoli episodi all'interno della trama complessiva. Nonostante ciò, esse non furono accolte con favore dall'unanimità dei lettori contemporanei, segno che il confronto tra la prima e la terza redazione del *Furioso* rimase aperto ancora per diversi anni. Lodovico Dolce, nell'*Apologia contra ai detrattori dell'Ariosto*, testimonia infatti come molti lettori giudicassero i nuovi canti aggiunti superflui e inferiori agli originali, sia per lo stile sia per la materia trattata. La replica dello stesso Dolce non è tuttavia priva di riserve: egli ammette che il poema sarebbe potuto restare nella sua forma originaria, senza aggiunte, ma sottolinea che Ariosto ritenne necessario intervenire per colmare alcune lacune e offrire una rappresentazione più compiuta e coerente dei temi trattati.⁵³

Dopo aver esaminato la trama delle quattro giunte e la loro collocazione nell'architettura complessiva del poema, è opportuno soffermarsi sulla distribuzione di questi nuovi materiali negli autografi superstiti. Come ricordato nel paragrafo precedente, sono pervenute fino a noi – con l'esclusione del testimone che tramanda le ottave su *Lo scudo della regina Elisa* – 58 carte originali, delle quali 54 (di cui una completamente bianca) conservate presso la Biblioteca Comunale di Ferrara (F), due presso la Biblioteca Ambrosiana di Milano (M) e due presso la Biblioteca Nazionale di Napoli (N I-2). Questo corpus, costituito da sette fascicoli e due fogli sciolti, presenta i contenuti riportati nella tabella seguente, ove per ciascun fascicolo o foglio si indica anche la relativa numerazione delle carte. Si precisa, inoltre, che le serie delle stanze, così

⁵³ Per approfondire, cfr. L. Dolce, *Apologia contra ai detrattori dell'Ariosto*, in L. Ariosto, *Orlando furioso*, In Venezia, Appresso Bindoni-Pasini, 1535, citato da S. Jossa, *Ariosto*, Bologna, il Mulino, 2009, pp. 109-110.

come quelle delle carte nei fascicoli, saranno interrotte mediante puntini di sospensione nei luoghi in cui il testo presenta delle lacune (cfr. Tab. 1.2).

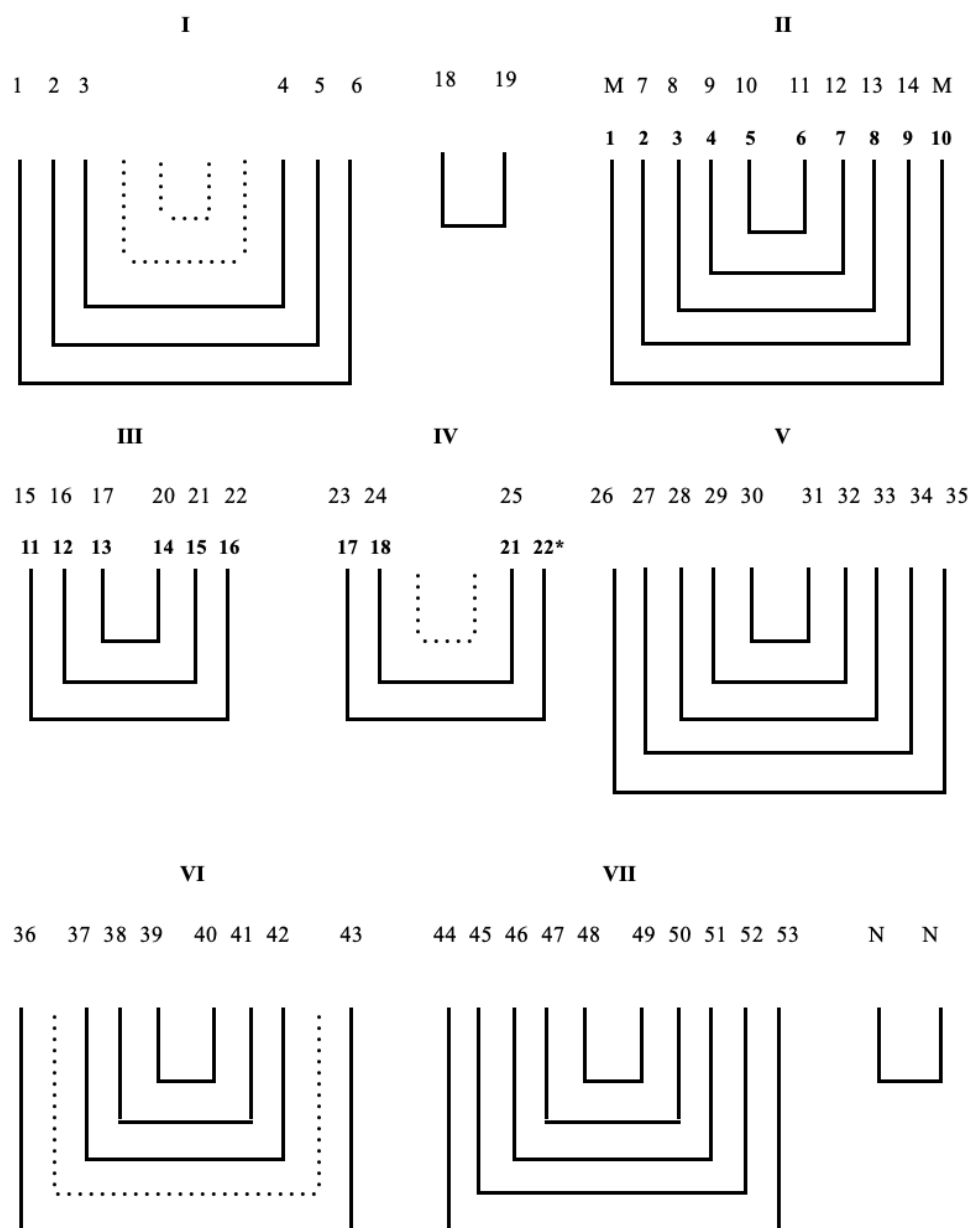
Tabella 1.2 – Distribuzione delle quattro giunte negli autografi superstiti

Fascicoli e Fogli	Testo	Carte
	EPISODIO DI OLIMPIA	
I	IX 8-39...83-94; X 1-18	F 1-3...4-6
Un foglio	XI 21-42, 48 (abbozzo), 44 (c.s.), 71 (c.s.), 78 (vv. 1-5); XII 1 (abbozzo)	F 18-19
II	IX 1-94 (fine); X 1-7	M 1, F 7-14, M 2
III	X 8-35 (vv. 1-3); XI 21-50	F 15-17, 20-22
IV	XI 51-70... XII 8-17	F 23-24...25-c. non num.
	EPISODIO DI MARGANORRE	
V	XXXVII 25-122 (fine)	F 26-35
	EPISODIO DI RUGGERO E LEONE	
VI	XLIV 11-20...81-92...104 (fine); XLV 1-8, 7-15	F 36...37-42...43
VII	XLV 16-117 (fine)	F 44-53
Un foglio	XLVI 50-71	N 1-2

Tenendo presente la consuetudine ariostesca di scrivere cinque ottave per pagina – prassi cui Ludovico rinunciava solo di rado – Debenedetti ha ricostruito la configurazione originaria dei fascicoli che presentano lacune interne (I, IV, VI).⁵⁴ Nel fascicolo I, tra le carte 3 e 4, si registra un'interruzione di 43 stanze (IX 40-82); dunque, appare evidente che dal quinterno siano stati asportati il quarto e il quinto foglio. Nel fascicolo IV, dopo due carte scritte, si osserva una lacuna, tra le carte 24 e 25, di 20 stanze (XI 71-83; XII 1-7). Il terno, composto dalle carte oggi numerate 23 e 24, dal foglio perduto contenente le 20 stanze, e dalle carte 25 più una (bianca e non computata), risulta privo del foglio centrale, ossia della parte più vulnerabile a strappi e dispersioni. Nel fascicolo VI, infine, si rileva una prima lacuna di 10 stanze (XLIV 21-30) dopo la carta iniziale, e una seconda di 11 stanze (XLIV 93-103) dopo la settima carta; pertanto, il quinterno risulta mancante del secondo foglio. Alla luce di queste evidenze, è possibile rappresentare con maggiore precisione la struttura della documentazione superstite, mediante un'infografica in cui una linea tratteggiata segnala i fogli perduti nei fascicoli I, IV e VI. Le carte prive di sigla, salvo diversa indicazione, appartengono al manoscritto di Ferrara (cfr. Fig. 1. 21).

⁵⁴ S. Debenedetti, *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XIX.

Figura 1. 21 – Ricostruzione dei fascicoli che presentano lacune interne



Dopo aver delineato la trama dei nuovi episodi, illustrato la loro collocazione nell'impianto definitivo del poema e la loro distribuzione nei documenti superstiti, si ritiene opportuno, in chiusura di questo paragrafo, considerare tali materiali dal punto di vista del grado di elaborazione letteraria. Data la complessità e l'ampiezza del corpus, l'esame di ciascun fascicolo e l'analisi dettagliata dei singoli loci critici – cancellature, correzioni, aggiunte e revisioni di interi versi o stanze – saranno rinviati alla successiva fase di codifica e di pubblicazione dell'edizione digitale. Come si vedrà più avanti, la codifica del testo all'interno di un'edizione scientifica digitale non rappresenta una mera operazione tecnica, ma un vero e proprio atto critico-interpretativo, fondato su un coerente ragionamento filologico. In questa sede ci si limiterà pertanto a osservare, seguendo le indicazioni di Debenedetti, che il materiale può essere suddiviso in sette sezioni, delle quali le prime cinque risultano costituite dai fogli e dai fascicoli del cosiddetto manoscritto ferrarese (F), opportunamente ordinato e integrato con le carte dell'Ambrosiana (M 1–10) e della Nazionale di Napoli (N 1–2), secondo il seguente schema.

1. Fasc. I: *Olimpia*, prima parte
2. Un foglio sciolto: *Olimpia*, seconda parte
3. Fasc. II–III–IV: *Olimpia*, prima e seconda parte
4. Fasc. V: *Marganorre*
5. Fasc. VI–VII + un foglio: *Ruggiero e Leone*

Le due unità restanti sono invece costituite, rispettivamente, da:

6. l'apografo dell'Ambrosiana, che conserva un frammento dell'episodio della Rocca di Tristano;
7. la carta 3 dell'autografo di Napoli, che tramanda le ottave relative allo *Scudo della regina Elisa*.

A un primo esame, le sette unità si distinguono nettamente per lo stato di redazione. Il foglio interposto tra il fascicolo I e il fascicolo II costituisce una mala copia particolarmente tormentata – la più travagliata dell'intero insieme – che si risolve in una serie di primi getti e abbozzi. Anche il fascicolo I e il fascicolo V possono essere considerati brutte copie, fortemente corrette, sebbene più ordinate della precedente. I fascicoli VI–VII, insieme al foglio superstite del fascicolo successivo oggi perduto, offrono invece un testo di qualità decisamente superiore, pur

recando tracce di significative innovazioni. I fascicoli II–IV si distinguono su tutti per cura formale e nettezza grafica, presentandosi come copie definitive, non soltanto corrette ma anche stilisticamente compiute. Dei due pezzi conservati separatamente, il mezzo foglio napoletano appare in gran parte di primo getto, ma presenta una fisionomia diversa da quella del foglio più tormentato, priva di ottave semplicemente abbozzate. L'apografo milanese, invece, riflette un originale già in pulito e con pochi ritocchi, senza che ciò implichi l'identificazione con l'ultima redazione del testo. L'insieme delle carte testimonia dunque una notevole varietà di stati redazionali, che spaziano dagli abbozzi più lontani alle trascrizioni definitive, passando per una gamma di copie intermedie, più o meno curate. Come osserva Debenedetti:⁵⁵

Ci sta dunque innanzi l'Ariosto mentre compone, e nei momenti più diversi. Così, quelle che potremmo chiamare le sue abitudini di lavoro ci diventano, alla lettura, quasi familiari, e persino qualche volta riusciamo a sorprendere i suoi processi fantastici, ora nel momento germinativo, ora nel loro svolgersi fino alla piena attuazione.

È opportuno, tuttavia, precisare che questi autografi non presentano elementi di carattere memoriale o biografico: non vi si trovano date, annotazioni o ricordi personali. Le uniche note marginali consistono in lettere ordinatrici, lineette di richiamo accanto ai versi che richiedono revisione, rimandi interni e appunti funzionali alla composizione

1.3.3 L'edizione debenedettiana de 'I Frammenti autografi dell'Orlando Furioso': contesto storico, premesse teoriche e innovazioni metodologiche

L'edizione a stampa degli autografi ariosteschi, finora richiamata soltanto marginalmente, merita un esame approfondito, al fine di comprenderne il contesto storico, le premesse teoriche e i criteri editoriali che ne hanno guidato l'allestimento, nonché le implicazioni metodologiche e pratiche per gli sviluppi successivi della disciplina filologica. Il volume de *I frammenti autografi dell'Orlando Furioso*, curato e pubblicato da Santorre Debenedetti (1878-1948) nel 1937, rappresentò l'unico numero della collana *Testi inediti e rari* del «Giornale storico della letteratura italiana».⁵⁶ Pochi anni prima, lo stesso Debenedetti, massimo interprete della tradizione ariostesca, aveva dato alle stampe l'edizione critica del *Furioso* (Bari, Laterza, 1928,

⁵⁵ Cfr. S. Debenedetti, Introduzione, in *I frammenti Autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XXII.

⁵⁶ L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, a cura di S. Debenedetti, Torino, Chiantorre, 1937.

nella serie «Scrittori d'Italia»), destinata a imporsi immediatamente come testo di riferimento e, ancora oggi, assunta come base imprescindibile per studi e commenti.⁵⁷ Nelle intenzioni del curatore, quell'edizione avrebbe dovuto essere corredata da un apparato critico che registrasse le varianti tra le tre redazioni autorizzate dal poeta (1516, 1521, 1532), secondo l'auspicio già formulato da Carducci.⁵⁸ Tuttavia, la collana laterziana non prevedeva apparati critici, e Benedetto Croce, direttore e ispiratore della serie, non concesse deroghe: venne così meno l'occasione di seguire passo a passo l'evoluzione testuale del poema verso la sua forma definitiva. Solo trentadue anni più tardi, Cesare Segre avrebbe realizzato, sulla base dei materiali di Debenedetti, un'edizione dotata di apparato critico.⁵⁹ Nonostante tali limitazioni, Debenedetti proseguì le sue ricerche con la consueta tenacia. Nel 1929 Egidio Piermarini, bibliotecario della Nazionale di Napoli, pubblicò alcune pagine autografe dell'Ariosto conservate in quell'istituto: si trattava della bella copia di alcune ottave poi confluite nella terza edizione del 1532.⁶⁰ Debenedetti accolse con favore la scoperta, non senza segnalare necessarie correzioni, e ricordò che altri frammenti relativi alla medesima redazione erano custoditi presso la Biblioteca Comunale Ariostea di Ferrara e la Biblioteca Ambrosiana di Milano.

Già in epoca precedente, del resto, gli studiosi che avrebbero potuto valorizzare i materiali autografi non se ne servirono in modo adeguato: Giannandrea Barotti, che in occasione della terza ristampa delle *Opere dell'Ariosto* (1766) ebbe occasione di consultarli, li menzionò soltanto sporadicamente, senza trarne alcun reale beneficio; né più incisivo fu il contributo di Girolamo Baruffaldi il giovane, i cui accenni rimasero superficiali e puramente esteriori.⁶¹ I materiali ferraresi furono resi più accessibili grazie alla riproduzione fototipica curata da Giuseppe Agnelli nel 1904.⁶² Tuttavia, come osserva lo stesso Debenedetti, pochi si cimentarono con quella lettura ardua e insidiosa, che richiedeva competenze paleografiche raffinate: tra i

⁵⁷ L. Ariosto, *Orlando Furioso*, a cura di S. Debenedetti, Bari, Laterza («Scrittori d'Italia»), 1928.

⁵⁸ G. Carducci, *Su l'Orlando Furioso*, in L. Ariosto, *Orlando Furioso*, Milano, Fratelli Treves, 1881.

⁵⁹ L. Ariosto, *Orlando Furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, a cura di S. Debenedetti e C. Segre, Bologna, Commissione per i Testi di Lingua, 1960.

⁶⁰ E. Piermarini, *Un episodio inedito dell'Orlando Furioso*, «Pegaso», I (2), febbraio 1929, pp. 168-18.

⁶¹ S. Debenedetti, *Introduzione*, a L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, con una premessa di C. Segre, Roma, Edizioni di Storia e Letteratura, 2010, p. XXI-XXII.

⁶² G. Agnelli, *I frammenti autografi dell'Orlando Furioso*, Roma, 1904.

pochi studiosi che vi si dedicarono, egli ricorda Giuseppe Lisio.⁶³ Quest'ultimo fu l'unico ad affrontare con autentica serietà l'impresa, vagheggiata già da Carducci, di una ristampa del *Furioso* del 1532 accompagnata dalle varianti delle edizioni precedenti e degli autografi, il tutto in forma diplomatica. Lisio pubblicò un saggio preliminare (i primi due canti, a tiratura limitata, nel 1909) e riuscì a licenziare per la stampa l'opera fino all'ottava 58 del canto IX. La morte improvvisa, sopraggiunta nel 1912, lo colse prima di poter rivedere la composizione successiva, già avanzata sino a circa due canti ulteriori: l'edizione rimase pertanto incompiuta, inedita e pressoché sconosciuta. L'iniziativa di Lisio, pur meritoria, presentava tuttavia limiti evidenti. Come sottolinea Debenedetti, non è metodologicamente corretto porre sullo stesso piano le varianti delle edizioni a stampa – che ottennero pieno riconoscimento nella tradizione del poema – e quelle dei manoscritti, spesso nate e scomparse nell'immediato, senza altra funzione che per l'autore stesso. Analogamente, la scelta di una trascrizione diplomatica delle ottave, fitte di correzioni, ne comprometteva inevitabilmente la leggibilità, rendendole quasi inaccessibili. Ciò nonostante, l'impresa avrebbe potuto fornire un contributo significativo alla conoscenza della storia del *Furioso*, soprattutto se si considera la ricchezza dei materiali impiegati e la qualità delle note predisposte da Lisio, già chiaramente visibili nel saggio pubblicato.⁶⁴

In questo quadro si colloca la decisione di Debenedetti di offrire agli studiosi e agli appassionati una trascrizione integrale e leggibile di tutti i frammenti autografi relativi alle aggiunte del 1532. Il progetto era radicalmente innovativo: la rappresentazione tipografica delle varianti e delle correzioni, disposte graficamente in modo da riflettere l'assetto degli autografi, consentiva di restituire non soltanto il testo ma anche la dinamica del lavoro creativo ariostesco. Un simile approccio non aveva precedenti nella storia della filologia, fatta eccezione per un caso pressoché dimenticato: la trascrizione del cosiddetto "codice degli abbozzi" (cioè l'autografo petrarchesco, Vat. Lat. 3196, che raccoglie le prime stesure e le revisioni di vari componimenti poi confluiti nel *Canzoniere*), realizzata nel XVII secolo da Federico Ubaldini. Quest'ultimo, erudito di eccezionale competenza e autore anche di un importante commento ai *Documenti d'Amore* di Francesco da Barberino, aveva infatti distinto con accorgimenti grafici i primi abbozzi, le correzioni e i rifacimenti del poeta, aprendo inconsapevolmente la strada a una modalità di edizione che soltanto nel Novecento avrebbe trovato nuova legittimazione.⁶⁵

⁶³ G. Lisio, *Il canto primo e il canto secondo dell'Orlando Furioso. Testo critico comparato*, Milano, La Gutenberg, 1909.

⁶⁴ S. Debenedetti, *Introduzione*, a L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, con una premessa di C. Segre, Roma, Edizioni di Storia e Letteratura, 2010, p. XXI-XXII.

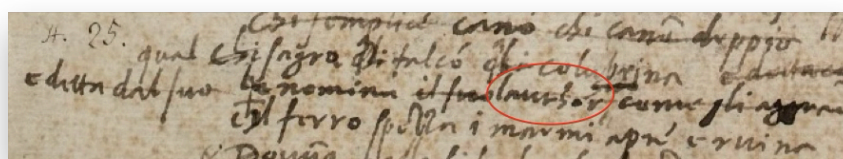
⁶⁵ Per approfondire, cfr. P. Italia e G. Raboni, *Che cos'è la filologia d'autore*, Roma, Carocci, 2010, pp. 19-22.

Nel paragrafo conclusivo della sua introduzione ai *Frammenti*, Debenedetti descrive gli sforzi compiuti per scovare anche le correzioni più sottili operate dal poeta:

La scrittura dell'Ariosto è, lo sappiamo, chiara e uniforme. Ma d'altro lato, con tutto quel mutare e rimutare, è ben naturale che le brutte e le pessime copie, e non di rado anche le buone, quando sono corrette a punta di penna perché la pagina non faccia un brutto vedere, debbano presentare delle difficoltà. E ne presentano, e son molte, e talora gravi. Fu necessario ricercare pazientemente tra parola e parola, e sotto le parole, e tra le righe i segni a volte quasi impercettibili di quella mano, e molte prove occorsero per giungere a renderli in modo perspicuo.⁶⁶

Il filologo individua quindi alcune modalità correttorie ricorrenti, per ciascuna delle quali offre una ricca casistica che qui si ripropone, corredata dalle immagini delle carte corrispondenti; le correzioni sono cerchiare in rosso. Gli esempi che seguono testimoniano la finezza del suo acume filologico: laddove non è possibile basarsi sulla coerenza semantica o metrica del verso, egli ricorre ad altri indizi materiali – come la gestione dello spazio grafico, la forma e l'andamento del tratto, o la densità dell'inchiostro – offrendo in tal modo una preziosa lezione di metodo. Tra le molte correzioni presenti, particolare attenzione richiedono quelle ottenute anteponendo, sulla stessa linea, una o più lettere alla parola già scritta. Ad esempio, osservando attentamente la *l* premessa ad *author*, si riconosce che si tratta di un'aggiunta, poiché annulla quasi del tutto la distanza abituale tra la parola e quella che la precede (cfr. Fig. 1.22).

Figura 1.22 – Correzione in rigo, cc. F 18 r. (XI 25, 2)

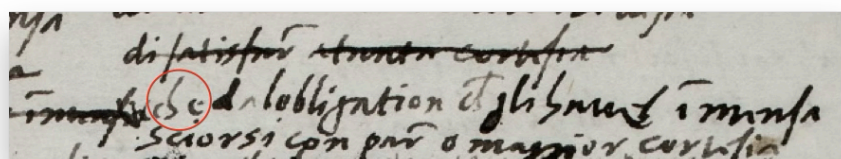


Un altro caso esemplare è offerto dal verso: «Che da l'obligation che gli have immensa». A una prima lettura sembrerebbe redatto in questa forma; tuttavia, si osserva che le lettere *ch* risultano strettamente unite fra loro, ma eccessivamente distanziate dalla *e*; inoltre, appaiono troppo vicine all'ultima parola (*immensa*) del verso cancellato nella colonna di sinistra. Per fugare ogni

⁶⁶ S. Debenedetti, *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XLI.

possibile dubbio, il poeta rimosse con maggiore decisione anche la *a* finale di *immensa*. Tutti questi indizi dimostrano che le due lettere sono un'aggiunta successiva. Il verso era stato in origine scritto come «E da l'obligation che gli have immensa», appartenente a una redazione intermedia del periodo, ora agevolmente ricostruibile (cfr. Fig. 1.23).

Figura 1.23 – Correzione in rigo, c. F 47 v. (XLV 52, 3)



Le correzioni più insidiose e frequenti, però, sono quelle in cui Ariosto sovrascrive la parola eliminata con una nuova, oppure ne modifica i tratti con sottile abilità grafica. Per riconoscere le lezioni celate, Debenedetti invita a soffermarsi su ogni minima deviazione rispetto alle abitudini scritte dell'autore: un'interlinea insolitamente ridotta, un inchiostro più denso, una forma anomala di alcune lettere (non di rado, nei ritocchi, Ariosto ricorre alla maiuscola) oppure un legamento inconsueto (cfr. Fig. 1.24). Un esempio significativo è costituito dalla parola *Per*, che copre l'articolo *La*: poiché in altre occorrenze la lettera *P* presenta l'asta prolungata oltre l'occhiello, la presenza della *L* sottostante si riconosce soltanto grazie alla linea orizzontale, ancora percepibile nel punto di attacco tra la *P* e la *e* (cfr. Fig. 1.25).

Figura 1.24 – Correzione sovrascritta con lettera maiuscola, c. M 1 r. (IX 3, 7)

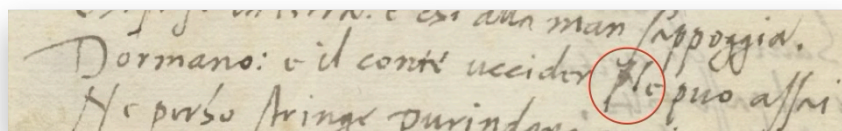
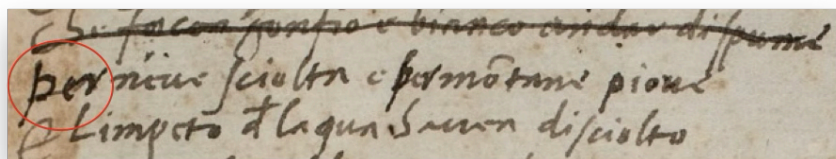


Figura 1.25 – Correzione sovrascritta, c. F 1 r. (IX 8, 6)



Un altro caso emblematico è rappresentato dalla parola *poppe*, in cui tutte le lettere – ad eccezione della prima, rimasta intatta – appaiono più scure e con tratti forzati; a un esame ravvicinato si distinguono infatti le tracce delle lettere originarie *r*; *o*; *r*; *a*, corrispondenti a *prora* (cfr. Fig. 1.26). In modo analogo, la *i* di *gli* in un verso della stessa carta, più spessa del consueto, lascia intuire la presenza di un'altra lettera sottostante, mentre la *g* risulta eccessivamente vicina alla parola precedente *si*: si tratta in realtà di un *lo* corretto successivamente in *gli* (cfr. Fig. 1.27).

Figura 1.26 – Correzione sovrascritta, c. F 1 r. (IX 9, 6)

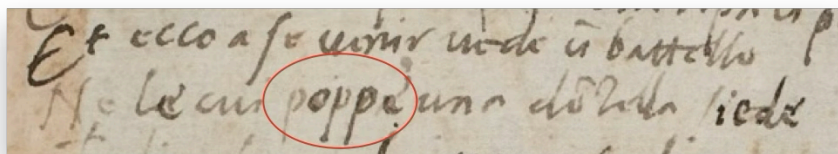
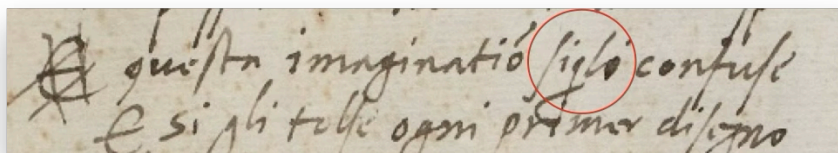


Figura 1.27 – Correzione sovrascritta, c. F 1 v. (IX 15, 1)



Anche la lezione *in quello stato* appare chiara; tuttavia, un piccolo ricciolo sull'asta della *s* lascia intravedere una *I* maiuscola sottostante, che consente di ricostruire la parola *Isola* (cfr. Fig. 1.28).

Poco oltre, sotto la *M* di *esser Morta*, si distingue l'inizio di una parola, *uc*, che si completa senza difficoltà in *uccisa* (cfr. Fig. 1.29).

Figura 1.28 – Correzione sovrascritta, c. F 3 r. (IX 32, 4)

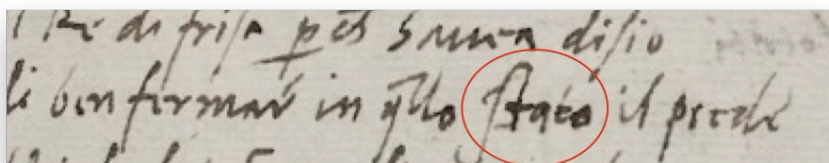
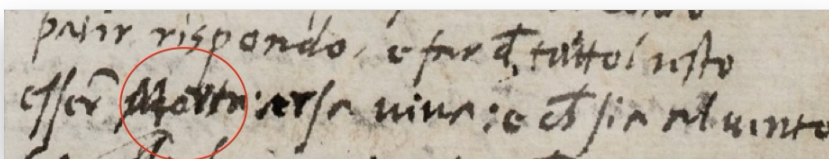
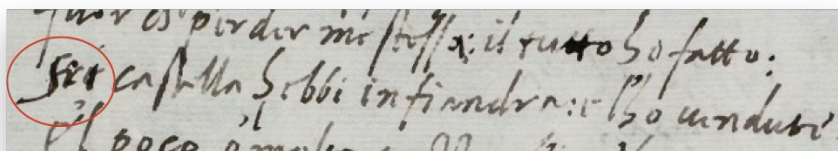


Figura 1.29 – Correzione sovrascritta, c. F 3 r. (IX 34, 3)



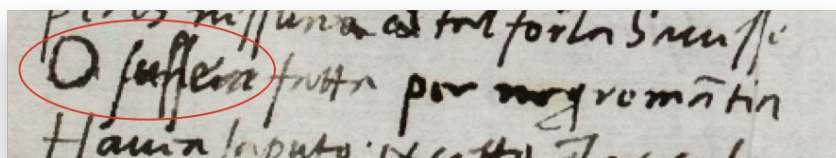
In un altro passo, si legge *sei castella*: le lettere sono insolitamente marcate; la *S* si lega alla *e* in modo anomalo; sotto la *e* si intravede una *r* e la *i* conserva ancora l'apice di una *e*. La lezione originaria era quindi *tre castella* (cfr. Fig. 1.30).

Figura 1.30 – Correzione sovrascritta, c. F 10 v. (IX 48, 3)



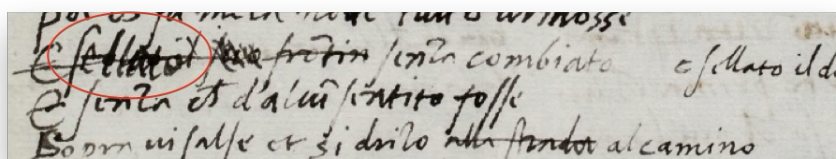
Ulteriori esempi confermano questa dinamica correttoria. Debenedetti osserva che, nel verso «O fusse fatta per negromantia», la scrittura appare a prima vista regolare; tuttavia la forma rigonfia e l'inchiostro particolarmente denso della *O* richiedono un'analisi più ravvicinata: sotto di essa affiorano infatti i resti di una *L* e di una *a*. Anche *fusse* nasconde un'altra parola: lo indicano il tratto incerto, l'irregolarità dell'inchiostro e la persistenza della desinenza *-ia*. Le lettere iniziali, pur abilmente dissimulate, si riconoscono con chiarezza non appena si intuisce la parola *lancia* (cfr. Fig. 1.31).

Figura 1.31 – Correzione sovrascritta, c. F 48 v. (XLV 65, 6)



Analogo è il caso di *sellato*: l'aspetto ingombrante e confuso della parola suggerisce una correzione. Tralasciando la *s* – verosimilmente aggiunta in un secondo momento, come attesta l'esiguità dello spazio – e osservando la *e*, vistosa e anomala, si riconosce al di sotto una *t*. Un ulteriore esame conferma che la lezione originaria è *tolto* (cfr. Fig. 1.32).

Figura 1.32 – Correzione sovrascritta, c. F 50 v. (XLV 85, 5)



Seguendo questi indizi, Debenedetti recupera, persino in carte che a prima vista sembrano prive d'insidie, numerose parole celate: nella prima carta, apparentemente chiarissima, individua sei varianti, tutte sfuggite a Giuseppe Lisio. Ciò nondimeno, il lavoro del filologo non si esaurisce nel mero recupero di lettere o parole; la parte più ardua consiste piuttosto nel ricostruire la corretta sequenza cronologica delle varianti d'autore. Nella maggior parte dei casi, Debenedetti

riesce a stabilire l'ordine delle correzioni, sia perché l'autore impiega modalità diverse di cancellatura (cfr. Fig. 1.33), sia perché lo deduce dalle sue abitudini scrittorie o dalla forma definitiva assunta dal passo nel manoscritto o nell'edizione del 1532. Non mancano, peraltro, passaggi spinosi, capaci di mettere a dura prova anche il suo acume filologico. La difficoltà si avverte soprattutto negli abbozzi, come mostra, ad esempio, la stanza *Tutto a un tempo Ruger*, in cui Debenedetti ricostruisce la successione delle stesure dei vv. 5-8 basandosi sulla loro disposizione spaziale, ovvero seguendo il progressivo allontanamento dalla redazione iniziale fino a quella definitiva, collocata addirittura più in alto dell'incipit della stanza (cfr. Fig. 1.34).

Figura 1.33 – Modalità diverse di cancellatura, c. F 48 v. (XLV 62, 6)

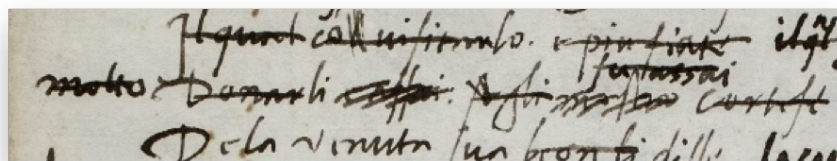
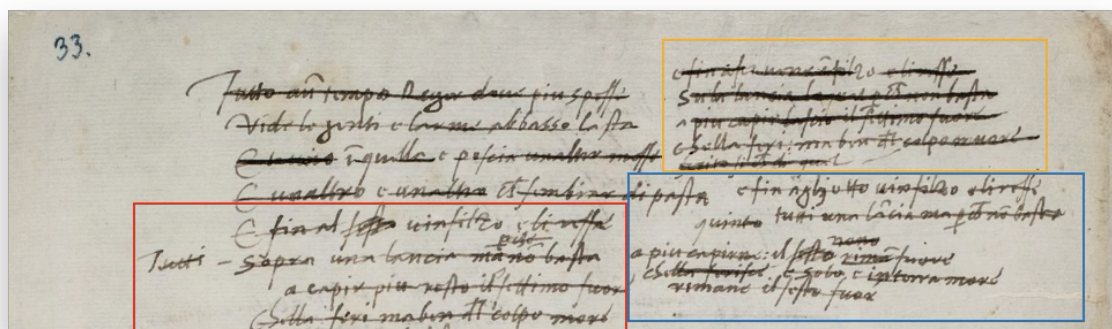


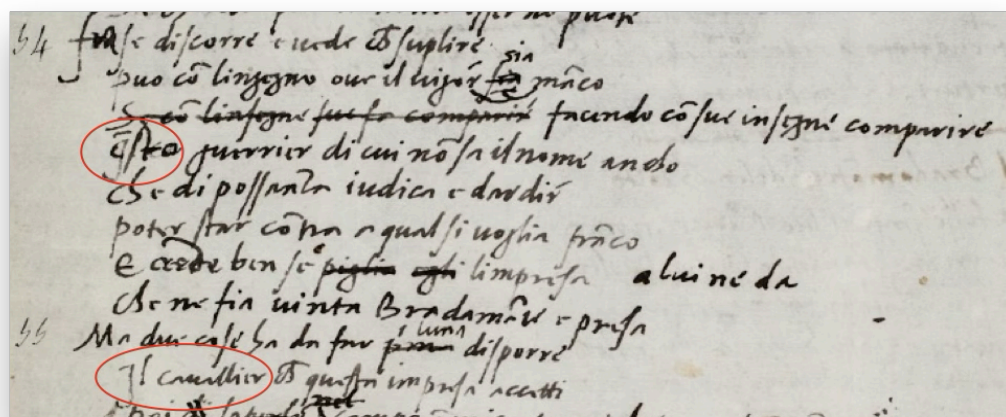
Figura 1.34 – Disposizione spaziale delle diverse stesure, c. F 33 r.



Se in questi casi le difficoltà risultano tutto sommato ancora sormontabili, ben più complesso diventa il lavoro ermeneutico quando Ariosto, appena abbozzata una parola – talvolta una sola lettera – la cancella o vi sovrascrive un'altra, proseguendo oltre: in simili circostanze Debenedetti procede per via ipotetica. Nel verso «Questo guerrier di cui non sa il nome ancho», sotto le prime lettere di *Questo* si riconosce un *Il*, e sotto le ultime due la sillaba *ca*.

Evidentemente, iniziata la parola, Ariosto si rammentò del verso dell'ottava successiva («Il cavallier che questa impresa accetti») e corresse all'istante, lasciando la parola incompiuta (cfr. Fig. 1.35).

Figura 1.35 – Correzione sovrascritta, c. F 47 v. (XLV 54, 4)



A quasi un secolo di distanza, queste osservazioni possono essere rilette alla luce delle nuove tecnologie spettroscopiche, che hanno rivoluzionato lo studio dei manoscritti. Si tratta di un insieme di metodologie ottiche capaci di analizzare la radiazione emessa o assorbita dai materiali e di restituire informazioni sulla loro composizione chimica e fisica. Tecniche come Raman, NIR (vicino infrarosso), fluorescenza a raggi X (XRF) o FTIR (Fourier Transform Infrared Spectroscopy) consentono di distinguere con estrema precisione gli strati di scrittura, rilevando differenze nella densità dell'inchiostro, nella granulometria dei pigmenti o nella presenza di composti metallici. Questi strumenti, non invasivi e privi di contatto diretto, penetrano sotto la superficie visibile e permettono di recuperare tracce ormai invisibili all'occhio umano. Il loro impiego si rivela particolarmente significativo per la filologia d'autore, poiché consente di verificare la successione cronologica delle varianti. In questo modo è possibile stabilire se due correzioni siano state eseguite con inchiostri diversi e distinguere interventi coevi da rielaborazioni posteriori. Inoltre, l'imaging multispettrale permette di isolare stratificazioni successive, separando il tratto originario da quello correttivo e restituendo con maggiore chiarezza il processo compositivo.⁶⁷

⁶⁷ Per approfondire, cfr. P. Italia, *Filologia d'autore digitale e multidisciplinare. Dall'Authorship alla fotonica*, in *Saperi Umanistici nella Contemporaneità. Atti del Convegno Internazionale dei dottorandi* (17-18 settembre 2015,

Applicata ai frammenti autografi, la spettroscopia potrebbe rappresentare un potente strumento di verifica e di scoperta. Da un lato, consentirebbe di convalidare le letture di Debenedetti, corroborando con dati oggettivi le sue intuizioni basate sull'osservazione paleografica: ad esempio, riconoscere inchiostri sovrapposti o confermare che una lettera apparentemente uniforme nasconde in realtà un segno sottostante. Dall'altro, offrirebbe l'opportunità di individuare nuove tracce correttorie, rimaste invisibili ai metodi tradizionali, come abbozzi di parole cancellate, residui chimici di inchiostro o differenze microscopiche nella pressione del tratto. In tal senso, l'applicazione delle tecniche spettroscopiche ai documenti ariosteschi non soltanto garantirebbe una verifica scientifica delle ipotesi di Debenedetti, ma permetterebbe anche di spingersi oltre il suo lavoro pionieristico, restituendo una mappa ancora più dettagliata della stratificazione autoriale. In definitiva, si configurerebbe come un ampliamento della filologia d'autore nel segno dell'interdisciplinarietà, dove la sensibilità critica del filologo e la precisione degli strumenti ottici collaborano per rendere visibile l'invisibile e dare nuovo spessore alla comprensione del laboratorio creativo di Ariosto.

Dopo aver delineato, attraverso un ampio repertorio di esempi, le abitudini correttorie di Ariosto, Debenedetti si sofferma a illustrare i criteri editoriali adottati nella trascrizione del testo e nella resa grafica delle varianti d'autore. Tali criteri, che qui si ripropongono corredati dalle relative immagini delle carte e dell'edizione cartacea, costituiscono un imprescindibile termine di confronto – talora anche in senso negativo – per la nostra edizione digitale. In apertura, lo studioso precisa:

Letti e decifrati i manoscritti, ho cercato il modo di renderli con tutta evidenza e semplicità, di far sì che la pagina dica subito quello che deve dire. La presente edizione non è strettamente diplomatica, ma interpretativa; e perciò, mentre rispetta sostanzialmente l'originale, non esita a scostarsi dalla fedeltà esteriore, quando il farlo possa giovare alla chiarezza.⁶⁸

L'edizione da lui approntata si configura quindi come interpretativa più che diplomatica, poiché mira a rispettare la sostanza dell'originale, pur concedendosi alcune deroghe formali laddove queste risultino funzionali alla comprensione del testo. Da questa premessa discendono alcune scelte rilevanti: la disposizione del testo in due colonne, che tenta di riprodurre, per quanto possibile, l'assetto dell'autografo; la suddivisione del testo in ottave numerate – caratteristica

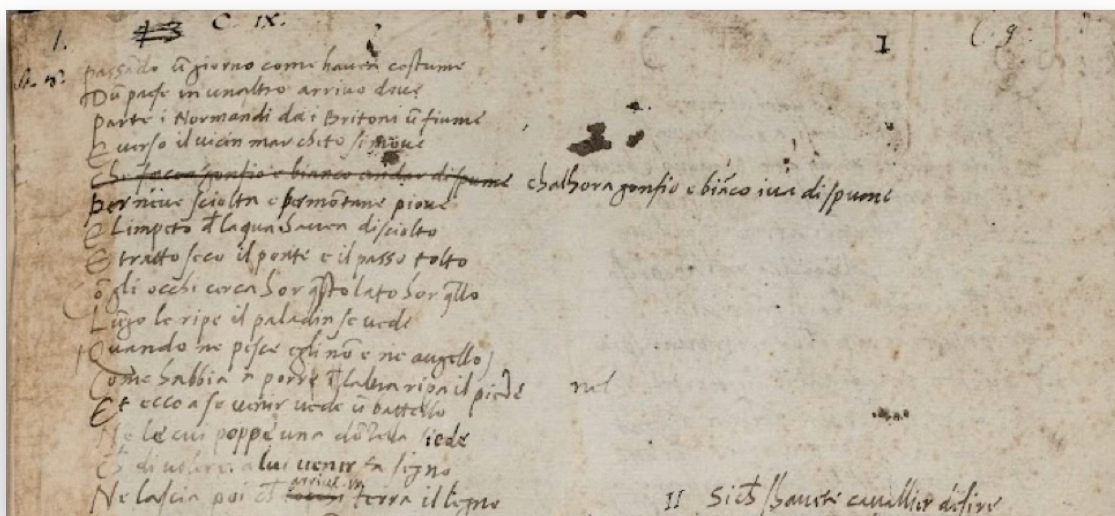
Università degli Studi di Palermo), Palermo, Palumbo, 2018, pp. 322-334; P. Italia, G. Raboni et al, *What is Authorial Philology?* Cambridge (UK), Open Book Publishers, 2021.

⁶⁸ S. Debenedetti, *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XLIII.

pressoché assente nel manoscritto, dove l'incipit di ciascuna è segnalato dalla maiuscola iniziale leggermente sporgente; la distinzione sistematica tra le lettere *u* e *v*; lo scioglimento delle abbreviazioni; la separazione delle parole; la normalizzazione delle maiuscole e l'adeguamento della punteggiatura all'uso moderno.

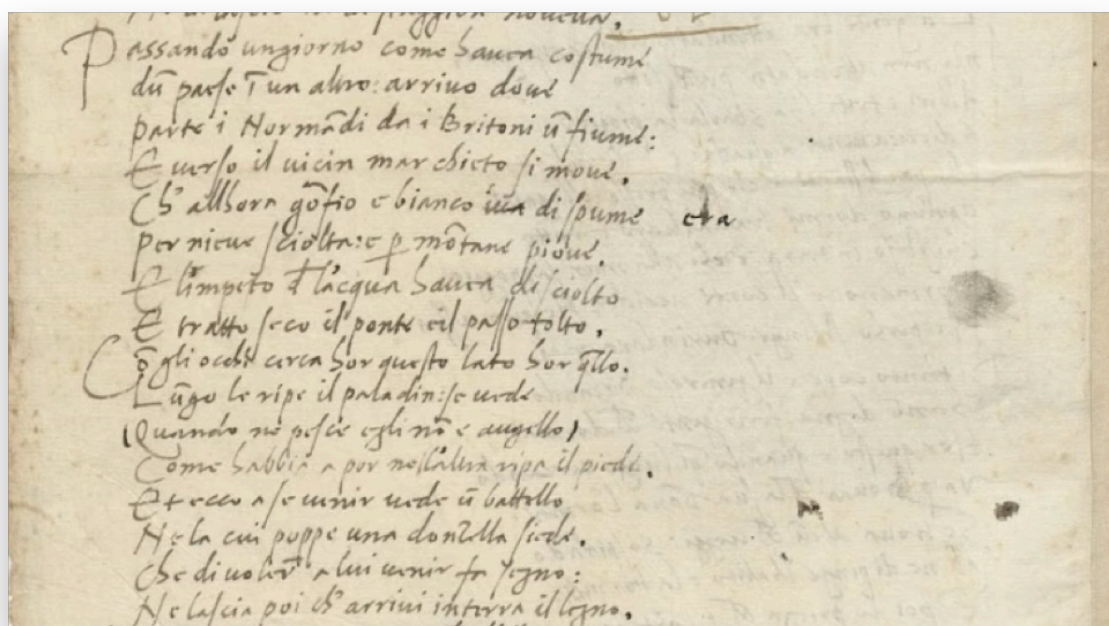
Per quanto riguarda l'uso delle maiuscole, Ariosto appare disordinato in tutte le copie, applicandole senza distinzione a nomi comuni, aggettivi e persino verbi. Diversa è la situazione della punteggiatura, che varia sensibilmente da un manoscritto all'altro in rapporto al grado di elaborazione del testo. L'autore, infatti, conosce i principali segni d'interpunzione – il punto (adoperato talvolta anche in funzione di virgola o di punto e virgola, quest'ultimo mai attestato negli autografi), la virgola e i due punti (usati con valore equivalente), il punto interrogativo, la parentesi, l'accento e l'apostrofo – ma non li impiega con regolarità nella sua pratica scrittoria. In particolare, le parole che normalmente richiederebbero l'apostrofo (per esempio l'articolo *un* davanti a un sostantivo maschile) compaiono spesso unite al termine successivo, mentre l'accento ricorre raramente e quasi sempre con funzione di disambiguazione semantica. A differenza degli abbozzi e delle brutte copie, le redazioni definitive presentano una punteggiatura più regolare e frequente. Un esempio significativo si trova nelle due stanze che aprono la “mala copia” di *Olimpia* (IX, 8-9), dove compare un solo segno di interpunzione: una parentesi al v. 3 della seconda stanza (cfr. Fig. 1.36).

Figura 1.36 – c. F 1 r.



Nel testo definitivo, invece, oltre alla parentesi ricorrono i due punti (dopo *altro* v. 2, *fiume* v. 3, *sciolta* v. 6; *paladin* v. 2, *segno* v. 7) e il punto (dopo *move* v. 4, *piove* v. 6, *tolto* v. 8; *quello* v. 1, *piede* v. 4, *siede* v. 6, *legno* v. 8). L'apostrofo è originario in *Ch'allhora* v. 5 e *ch'arrivi* v. 8, mentre sembra aggiunto in un secondo momento in *limpeto*, *lacqua* v. 7 e *laltra* v. 4 (cfr. Fig. 1.37).

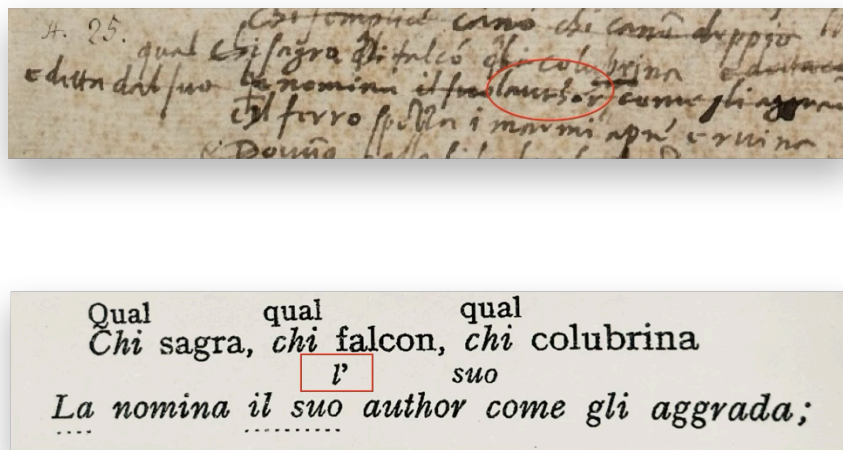
Figura 1.37 – c. M 1 v.



Debenedetti sceglie di non riprodurre gran parte delle maiuscole presenti nei manoscritti, siano esse originarie o introdotte in fase di correzione, destinate a dare rilievo a una parola o semplicemente a rendere più evidenti i ritocchi. Quando però la maiuscola assume un preciso valore semantico, viene mantenuta nella trascrizione (ad esempio, la maiuscola nella personificazione *Antri* nel verso «Rispondean gli Antri che pietà n'havieno», X 22, 8). Un analogo criterio di normalizzazione guida l'intervento sulla punteggiatura sia nelle brutte copie, dove i segni sono sporadici o del tutto assenti, sia in quelle “in pulito”, dove risultano spesso aggiunti in un secondo momento e collocati in modo improprio, con il rischio di generare fraintendimenti. Per tale ragione il filologo avverte che chi intenda approfondire nel dettaglio le questioni ortografiche del Cinquecento dovrà fare riferimento all'edizione fototipica di Agnelli (Roma, 1904) o ad altri testimoni.

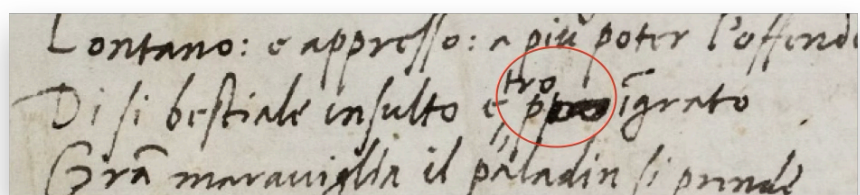
Sulla scorta delle soluzioni adottate da Federico Ubaldini nella sua edizione del codice degli abbozzi, Debenedetti ricorre a espedienti tipografici tanto semplici quanto efficaci per rappresentare le varianti d'autore. Per quanto riguarda le lettere aggiunte, queste vengono riportate in interlinea sia quando nel manoscritto si trovano effettivamente in quella posizione, sia quando risultano inserite direttamente nella riga del verso (per esempio, il suddetto *author* corretto in *lauthor*, con l'aggiunta dell'articolo davanti). In tali casi, la modifica dell'editore è segnalata in nota (cfr. Fig. 1.38).

Figura 1.38 – c. F 18 r. (XI 25, 2), ed. Debenedetti, p. 22



Le parole cancellate sono invece trascritte in corsivo. Debenedetti avverte tuttavia che non sempre conviene riprodurre alla lettera gli ingegnosi adattamenti di Ariosto, poiché la stampa, più che restituirli, rischierebbe talvolta di deformarli. Alcuni ritocchi, infatti, sono eseguiti con tale scrupolo da conservare persino il minimo segno della parola originaria. Così, per trasformare *poco* in *troppo*, Ariosto aggiunge *tro* in interlinea, modifica la *o* in *p* allungandone il tratto e racchiude le ultime due lettere in un cerchietto, pur avendo a disposizione spazio libero sul foglio (cfr. Fig. 1.39).

Figura 1.39 – c. F 22 v. (XI 48, 5), ed. Debenedetti, p. 64.

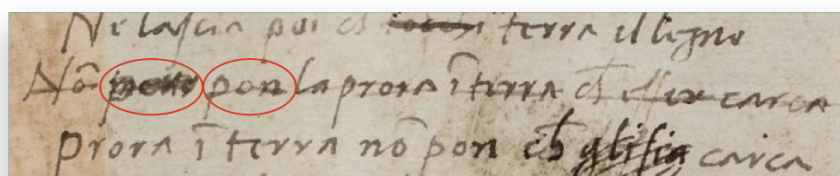


Lontano: e appresso: a più poter l'offende
 Di sí bestiale insulto e ~~poco~~ ^{troppo} ingrato
 Gran maraviglia il paladin si prende

Lontano e appresso, a più poter l'offende.
 Di sí bestiale insulto e ^{troppo in} poco grato
 Gran maraviglia il paladin si prende:

Quando una parola cancellata è sostituita da un'altra lezione (in rigo, in interlinea o comunque nella stessa colonna), poi anch'essa rifiutata, la prima correzione scartata è segnalata con dei puntini.

Figura 1.40 – c. F 1 r. (IX 10, 1), ed Debenedetti, p. 5

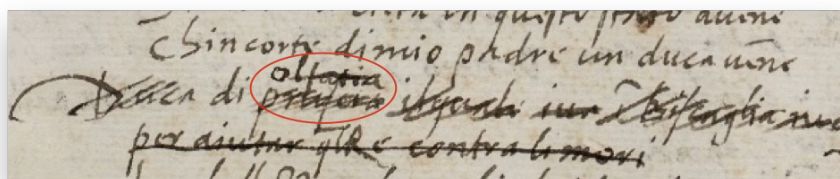


Ne lascia poi es ~~mette~~ terra il legno
 Nò ~~mette~~ ^{pon} la prova in terra es ~~esser~~ carica
 prova in terra nò pon es ~~glifia~~ carica

[10]
 Non mette pon la prova in terra; che esser carica

In un primo momento Ariosto tenta di trasformare la parola *mette* in *pon*, ma, non riuscendovi, la cancella e la riscrive a seguire (cfr. Fig. 1.40).

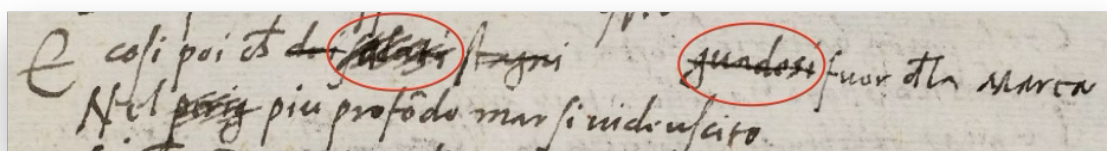
Figura 1.41 – c. F 2 r. (IX 23, 1), ed Debenedetti, p. 7



[23]
Olsatia
 Duca di Pxxxa, il quale iva in Biscaglia iva

In questo caso, la parola *Pxxxa* allude verosimilmente a *Prussia*, mentre *Olsatia* risulta sovrascritta a un'altra voce preesistente (cfr. Fig. 1.41).

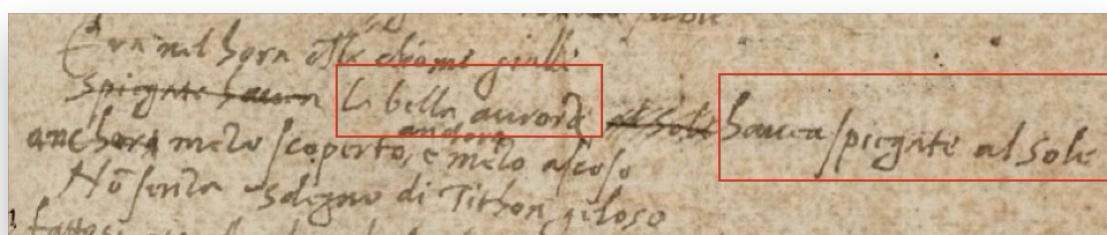
Figura 1.42 – c. F 4 v. (IX 90, 1) ed Debenedetti, p. 13



[90]
 E così, poi che *dei* ^{guadosi} salati stagni

In questo caso, Ariosto elimina la lezione *salati* e, nella seconda redazione a margine, la sostituisce con *guadosi* (cfr. Fig. 1.42). In alcune correzioni, per agevolare la lettura, tra parentesi quadre si ripetono parole o persino interi versi della redazione precedente (cfr. Fig. 1.43).

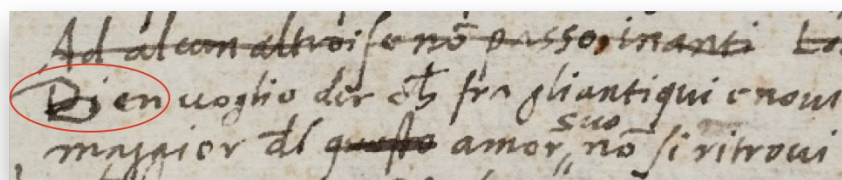
Figura 1.43 – c. F 19 r. (XI 32, 6), ed. Debenedetti, p. 25



Era ne l'ora ~~che le chiome gialle~~
 Spiegate havea la bella Aurora al Sole,
^{Anchor} ~~Ch'era~~ ^{anchora} mezo scoperto e mezo ascoso,
 Non senza sdegno di Tithon geloso.

Dunque, nei documenti la lezione definitiva può collocarsi immediatamente dopo quella espunta, oppure in interlinea o nel margine. Debenedetti rileva come, in tali circostanze, l'interpretazione della dinamica correttoria non presenti grandi difficoltà. Diverso è il caso in cui la nuova variante ricopra parzialmente o totalmente la lezione precedente, situazione che richiede ulteriori precisazioni. Qualora l'intervento correttorio sia avvenuto durante la stessa fase di scrittura del verso, la nuova parola viene apposta subito dopo quella cassata: si pensi, ad esempio, al *Di* coperto dalla prima lettera di *Ben* nel passo «Ben voglio dir...», dove Ariosto aveva iniziato a scrivere *Di*[r], salvo poi sostituirlo con *Ben* (cfr. Fig. 1.44).

Figura 1.44 – c. F 5 r. (X 1, 7), ed. Debenedetti p. 14



Ad alcun altro, se non passò inanti.

Di Ben voglio dir che fra gli antiqui e novi
de l' suo
Maggior *di questo* amor non si ritrovi;

Nei casi in cui non sia possibile stabilire con certezza se si tratti di una variante immediata o tardiva, il verso viene interrotto al termine della parola rifiutata e riproposto integralmente nella forma corretta, con le parole già trascritte racchiuse fra parentesi quadre. Si veda l'esempio del verso «De l'isola guardian lascia il cugino» in cui la *g* (maiuscola) è stata sovrascritta su una preesistente *a* (cfr. Fig. 1.45).

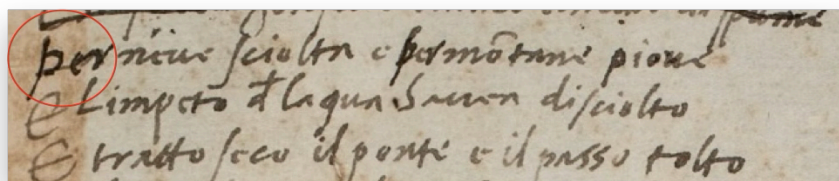
Figura 1.45 – c. F 4 r. (IX 86, 8), ed. Debenedetti, p. 12

Et egli tratto poi da un'altra cura
De le fortezze e di tutto il domino
De l'isola *Guardian* lascia il cugino.

Et egli, tratto poi da un'altra cura,
De le fortezze e di tutto il domino
De l'isola *a* [De l'isola] guardian lasciò il cugino;

Altrimenti, se si fosse stampato «De l'isola *a* guardian lasciò il cugino» evidenziando la semplice cassatura della *a* con il corsivo, il lettore avrebbe potuto credere che le due parole fossero state scritte di seguito, mentre una copre l'altra. Quando, invece, la correzione è intervenuta in un momento successivo alla stesura compiuta del verso, la nuova lezione viene registrata in interlinea o nella colonna di destra, a seconda delle necessità. Si veda l'esempio del verso «Per neve sciolta...» dove *Per* è stato sovrascritto su *La* (cfr. Fig. 1.46).

Figura 1.46 – c. F 1 r. (IX 8, 6), ed. Debenedetti, p. 5



Che facea gonfio e bianco andar di spume
 Per i per
 La neve sciolta e le montane piove:
 E l'impeto de l'aqua havea disciolto
 E tratto seco il ponte, e il passo tolto.

Debenedetti interviene anche sugli errori di scrittura inconsapevoli, ossia dovuti a mera disattenzione, rilevando che una trascrizione alla lettera di forme come *ora* al posto di *orca* o *barma* per *barba* finirebbe per compromettere interi versi a causa di un eccesso di pedanteria. Il lettore, tuttavia, viene sempre informato della grafia originaria. Diversamente, Debenedetti mantiene inalterati gli errori che riflettono un intento autoriale: essi possono dipendere da difetti di pronuncia (come nel caso di *Basilarda*) o da altre particolarità che meritano di essere registrate. Debenedetti conclude affermando:

Degli autografi, insomma, la presente è una edizione fatta con intenti critici; [...]. Ho tentato, con quei mezzi che la stampa mi consentiva, di rendere facili all'occhio e familiari i manoscritti dell'Ariosto. Grande semplicità di procedimenti e insieme uno sforzo continuo di chiarezza. Ho fatto sí che quasi non si vedesse la fatica da me durata, e non soltanto da me (i miei amici lo sanno), per sorprendere ogni più piccolo segno di quella mano e disporre le varianti nel loro vero ordine, badando alla sostanza delle cose. Nessun commento, se pur talora la tentazione era forte, per lasciare tutto solo il lettore con queste carte.⁶⁹

Sono parole significative, che sintetizzano il senso complessivo dell'impresa del filologo. Se da un lato è vero – come ha osservato la critica successiva – che l'edizione anticipa di alcuni decenni la *critique génétique* della scuola filologica francese, poiché si presenta come

⁶⁹ S. Debenedetti, *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, Roma, Edizioni di Storia e Letteratura, 2010, p. XLI.

pubblicazione integrale di tutto l'avantesto del *Furioso*;⁷⁰ dall'altro è altrettanto evidente che essa non ha un mero intento riproduttivo (lo stesso Debenedetti, d'altronde, rinvia all'edizione fototipica di Agnelli per chi sia interessato a conoscere la fisionomia dell'originale). Al contrario, si tratta di un'edizione concepita con finalità critico-interpretative, in quanto il filologo descrive la fenomenologia delle correzioni, formula ipotesi sulla corretta successione cronologica delle varianti d'autore e propone soluzioni tipografiche mirate a rappresentarle sulla pagina. In questo senso, il lavoro di Debenedetti non costituisce soltanto l'atto fondativo della moderna filologia d'autore, ma pone anche le basi di quella sua specifica diramazione disciplinare che è la critica delle varianti.

⁷⁰ La *critique génétique* è una disciplina filologica e critica che si è affermata in Francia negli anni '70, in netta reazione allo strutturalismo e alle teorie che consideravano il testo letterario come un'entità chiusa e statica. Il suo obiettivo primario è studiare il processo di creazione e di elaborazione di un'opera, focalizzandosi non sul testo finito e pubblicato, ma sulla sua genesi, ovvero su tutte le fasi che hanno condotto alla sua forma finale. I principi chiave della *critique génétique* sono elencati di seguito.

- La centralità del manoscritto e degli *avanttextes*: a differenza della filologia tradizionale che si concentra sull'edizione "ideale" o "definitiva" di un testo, la *critique génétique* pone al centro della sua indagine i materiali preparatori dell'opera. Gli *avant-textes* (o *dossier genetico*) possono includere appunti, abbozzi, schemi, prime stesure, rielaborazioni, ma anche, corrispondenze epistolari, disegni, ecc. L'analisi di questi materiali permette di ricostruire il "laboratorio" dell'autore e di comprendere le sue strategie di scrittura, le esitazioni, le scelte e i ripensamenti.
- Il testo come processo dinamico: la *critique génétique* concepisce il testo non come un prodotto statico, ma come un processo in divenire, una "scrittura in atto"; l'opera è vista come il risultato di una serie di operazioni, spesso complesse e non lineari, che l'autore compie nel tempo.
- L'analisi delle varianti d'autore: un aspetto fondamentale è lo studio delle varianti, ovvero le modifiche apportate dall'autore al proprio testo; queste varianti non sono considerate "errori" da eliminare, ma tracce preziose del pensiero e del lavoro creativo dell'autore.
- La ricostruzione dell'intenzione autoriale (e la sua problematicità): sebbene l'obiettivo sia comprendere come l'opera sia nata e si sia sviluppata, la *critique génétique* è consapevole della difficoltà e della parzialità della ricostruzione dell'intenzione autoriale; l'analisi dei materiali genetici permette di formulare ipotesi sul processo creativo, ma non sempre di penetrare completamente la mente dell'autore.
- L'influenza dell'informatica umanistica: con l'avvento del digitale, la *critique génétique* ha trovato nuovi strumenti per la sua applicazione. Le edizioni genetiche digitali permettono di rappresentare in modo più esaustivo e interattivo la complessità dell'elaborazione d'autore, offrendo al lettore la possibilità di navigare tra le diverse fasi redazionali e di visualizzare le varianti.

La *critique génétique* è stata sviluppata soprattutto dall'*Institut des Textes et Manuscrits Modernes* (ITEM) del CNRS di Parigi, grazie al contributo di figure come Almuth Grésillon (cfr. *Eléments de critique génétique*, a cura di L. Hay, Paris, Pouf, 1994) e Louis Hay (cfr. *La naissance du texte*, a cura di L. Hay, Paris, Corti, 1989). Per approfondire questo argomento, cfr. A. Stussi, *Filologia d'autore*, in *Introduzione agli studi di filologia italiana*, Bologna, Il Mulino, 2006, pp. 158-170; P. Italia e G. Raboni, *La filologia d'autore e la critique génétique*, in *Che cos'è la filologia d'autore*, Roma, Carocci, 2010, pp. 26-28.

1.3.4 La ricezione critica e le implicazioni metodologiche dell'edizione debenedettiana

Questo paragrafo non pretende di offrire una trattazione esaustiva, ma soltanto alcuni rapidi cenni sulla ricezione critica dei *Frammenti* di Debenedetti, rinviando a prossimi lavori un'analisi più dettagliata della fortuna degli autografi ariosteschi presso la critica contemporanea.⁷¹ Gianfranco Contini, principale allievo di Debenedetti, rimase subito colpito dall'originalità di questa edizione e ne scrisse una recensione, *Come lavorava l'Ariosto*, apparsa sul «Meridiano di Roma», con la quale inaugurò quella “critica delle varianti” destinata a diventare il nucleo della sua attività.⁷² Sebbene non si fosse mai occupato prima della poesia ariostesca, il volume gli offrì l'occasione di sistematizzare intuizioni fino ad allora isolate sul modo di accostarsi alla poesia e di valutarla criticamente. Contini si chiede:

Che significato hanno, per il critico, i manoscritti corretti degli autori? Vi sono essenzialmente due modi di considerare un'opera di poesia: vi è un modo, per dir così statico, che vi ragiona attorno come un oggetto o risultato, e in definitiva riesce ad una descrizione caratterizzante; e vi è un modo dinamico, che la vede quale opera umana o lavoro in fieri, e tende a rappresentarne drammaticamente la vita dialettica. Il primo stima l'opera poetica un “valore”; il secondo, una perenne approssimazione al “valore”; e potrebbe definirsi, rispetto a quel primo e assoluto, un modo in senso altissimo, “pedagogico”.⁷³

Il passo mette a fuoco il diverso significato che i manoscritti corretti possono assumere per il critico letterario. Contini distingue due modi di concepire l'opera poetica. Da un lato vi è un approccio “statico”, che considera il testo come un prodotto compiuto, già definito, suscettibile quindi di una “descrizione caratterizzante” e di una valutazione estetica. In questa prospettiva, l'opera appare come un “valore” acquisito e fissato, una realtà conclusa. Dall'altro lato emerge un approccio “dinamico”, che guarda all'opera come a un processo, a un lavoro in continuo

⁷¹ Per approfondire la fortuna degli autografi ariosteschi presso la critica contemporanea, cfr. S. Albonico, *Pour un retour aux manuscrits de l'Arioste*, «Génésis», n. 49, 2019, pp. 31-46; poi Id., *Dalla parte di Ariosto*, in *Gli 'scartafacci' degli scrittori. I sentieri della creazione letteraria in Italia*, a cura di C. Del Vento e P. Musitelli, Roma, Carocci, 2022, pp. 143-62; Id., *Ariosto*, in *Autografi dei letterati italiani. Il Cinquecento*, a cura di M. Motolese, P. Procaccioli e E. Russo, Roma, Salerno editrice, 2022, vol. III, pp. 3-35. I. Campeggiani, *L'ultimo Ariosto. Dalle Satire ai Frammenti autografi*, Pisa, Scuola Normale Superiore, 2017; Id., *I frammenti autografi dell'Orlando Furioso: un'ipotesi per lo 'scrittoio' di Ariosto*, in *Dreaming again on things already dreamed: 500 years of Orlando furioso (1516-2016)*, Atti del convegno tenuto a Oxford nel 2016 in occasione del 5° centenario dell'Orlando furioso, a cura di M. Dorigatti e M. Pavlova, Oxford, Peter Lang, 2019, pp. 133-143; V. Gritti, *Intorno a una serie di correzioni nell'“Orlando Furioso”: i metaplasmi e le forme eteroclite*, «Storie e linguaggi», 2020, pp. 63-92.

⁷² G. Contini, *Come lavorava l'Ariosto*, «Meridiano di Roma», 18 luglio 1937; poi in Id., *Esercizi di lettura sopra autori contemporanei, con un'appendice su testi non contemporanei*, Torino, Einaudi, 1974, pp. 232-41.

⁷³ Ibid., p. 233.

divenire. In questo senso, i manoscritti con correzioni e varianti assumono una rilevanza particolare, poiché testimoniano il percorso creativo dell'autore, le sue esitazioni e le sue scelte progressive. L'opera non è più soltanto un risultato, ma una tensione costante verso il "valore", una ricerca che si sviluppa dialetticamente. Questo secondo atteggiamento, definito "pedagogico" nell'accezione più alta del termine, non si limita a contemplare il prodotto finale, ma intende rappresentare la vitalità stessa della creazione poetica, rendendo visibile la dimensione formativa e umana del testo letterario. In tal modo, il critico non coglie soltanto l'opera nella sua staticità, ma anche nella sua dimensione vitale e processuale. Contini prosegue:

L'idea direttiva del poeta dovrà coincidere col mondo poetico quale risulta dalla descrizione caratterizzante; il secondo metodo costituirà una verifica o controprova del primo; o per dir meglio, i due metodi non saranno che due aspetti empiricamente differenziati della medesima lettura dei testi e giudizio sulla poesia.⁷⁴

L'“idea direttiva” del poeta – cioè il nucleo ispiratore che orienta la creazione – deve trovare riscontro nel “mondo poetico” ricostruito dal critico attraverso la descrizione caratterizzante. In altri termini, l'analisi deve cogliere l'unità tra progetto creativo e risultato estetico. Il metodo dinamico non si pone come alternativo, ma funge da verifica: mette alla prova la coerenza tra intenzione e sviluppo concreto dell'opera, illuminandone tensioni e passaggi. I due approcci non vanno quindi concepiti come separati o antitetici, ma come prospettive complementari, due modalità empiricamente distinte di un'unica operazione critica: leggere i testi e formulare un giudizio sulla poesia. Nel suo commento a Contini, Cesare Segre prendendo in prestito la terminologia della linguistica strutturale, osserva che «se il testo costituisce un'unità sincronica, le varianti elaborative ci portano invece a quel susseguirsi di sincronie di cui consiste la diacronia della stesura: sono due modi di leggere il testo che debbono necessariamente integrarsi, tramite una visione temporale». ⁷⁵ In questo modo Contini mirava a ricondurre il metodo dinamico nell'alveo dell'idealismo crociano, presentandolo come esercizio di valenza “pedagogica” volto a confermare per via sperimentale la critica estetica. È chiaro che la funzione pedagogica cui egli alludeva non coincide più con quella cinquecentesca, intesa come

⁷⁴ Ibid., p. 234.

⁷⁵ C. Segre, *Premessa*, a L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, a cura di S. Debenedetti, Roma, Edizioni di Storia e Letteratura, 2010, p. VII.

apprendimento diretto dell'arte poetica, bensì con la comprensione critica del processo creativo.⁷⁶ Muovendo dalla dicotomia crociana poesia/non-poesia, Contini afferma:

Queste successive redazioni e varianti possono offrire due stati ben distinti; in un caso, i rapporti dell'essere al non-essere poetico, l'inventio delle vecchie arti retoriche, la scoperta o rivelazione del fantasma in relazione allo stato di attesa, la progressiva identificazione di esso; in un altro, le vere e proprie "correzioni", cioè la rinuncia a elementi frammentariamente validi per altri organicamente validi, l'espunzione di quelli e l'inserzione di questi.⁷⁷

La funzione pedagogica della critica delle varianti si sostanzia nell'individuazione di due processi variantistici distinti ma interdipendenti: da un lato, lo "stato di attesa", ossia la "scoperta o rivelazione del fantasma" che segna l'instaurarsi del valore poetico; dall'altro, la sostituzione di valori "frammentariamente validi" con altri "organicamente validi", in vista della coesione stilistica. Sulla scorta di questa distinzione, in seguito Contini classificherà le varianti d'autore in due categorie: le "varianti instaurative", che segnano la conquista del valore poetico aprendo nuove possibilità espressive, e le "varianti sostitutive", in cui un valore puntuale viene rielaborato per assicurare coerenza stilistica.⁷⁸ Ogni variante modifica un equilibrio preesistente e ne impone uno nuovo, poiché ogni elemento del testo è in relazione con gli altri e con l'intero sistema culturale che lo sostiene.

Contini conclude il suo primo esercizio di "critica delle varianti" osservando che «la direzione costante che si individua nel lavoro correttivo dell'Ariosto si trova a coincidere perfettamente con la migliore descrizione caratterizzante che sia data fin qui della sua poesia» – allusione evidente al saggio di Croce sull'Ariosto (1920), in cui il *Furioso* era interpretato come luogo in

⁷⁶ Nel Cinquecento la raccolta e l'analisi delle varianti d'autore avevano un obiettivo eminentemente pratico e didattico: non si trattava di studiare il processo creativo in sé, ma di fornire ai poeti contemporanei strumenti per imparare a scrivere meglio, affinando lo stile sul modello dei classici o dei moderni illustri.

Un primo esempio è Giovanni Pontano, umanista napoletano, che nelle sue *Actius* e in altri dialoghi aveva posto come modello Jacopo Sannazzaro, esaltandone la capacità di mediare tra tradizione latina e volgare. Il commento di Pontano non si limitava alla celebrazione, ma insisteva sull'utilità formativa dello studio delle sue scelte linguistiche e stilistiche: il testo del maestro diventava una sorta di laboratorio per i discepoli, che potevano esercitarsi nell'imitazione e nell'emulazione.

Un altro caso significativo è quello di Bernardino Daniello, commentatore di Petrarca, che nel 1541 pubblicò un'edizione delle *Rime* con particolare attenzione alle varianti testuali. La sua operazione non aveva un intento filologico moderno, ma piuttosto pedagogico: mostrare ai giovani poeti come il modello petrarchesco fosse stato progressivamente affinato, così da offrire esempi concreti di "correzione" e perfezionamento stilistico. Le varianti erano dunque intese come "lezioni di scrittura", strumenti per apprendere l'arte poetica attraverso l'osservazione delle scelte di un autore canonico.

⁷⁷ Ibid., p. 234.

⁷⁸ Cfr. G. Contini, *Varianti e altra linguistica*, Torino, Einaudi, 1970.

cui una sovrana legge dell'armonia riconduceva a unità temi, episodi e sentimenti particolari.⁷⁹ Su tali basi, Contini sviluppò il proprio metodo di analisi delle varianti, applicandolo a Petrarca (cfr. *Saggio d'un commento alle correzioni del Petrarca volgare*, Firenze, Le Monnier, 1943), Leopardi (cfr. *Implicazioni leopardiane*, Firenze, Sansoni, 1949), Mallarmè (cfr. *Sulla trasformazione dell' "Après-midi d'un faune"*, in «Letteratura», 1948) e Proust (cfr. *Introduzione alle "paperoles"*, in «Letteratura», 1947; *"Jean Santeuil", ossia l'infanzia della "Recherche"*, in «Paragone», 1953). L'articolo di Contini ebbe larga circolazione, mentre i *Frammenti* furono presto dimenticati a causa delle leggi razziali del 1938, che privarono Debenedetti della cattedra di filologia romanza all'università di Torino, della direzione editoriale del «Giornale storico della letteratura italiana» e della possibilità di pubblicare. Le copie superstiti furono poche e l'opera non poté essere confrontata con il lavoro di Contini, che la definì «la cosa più eccellente elaborata in questo secolo dalla critica ariostesca».⁸⁰

Nella premessa alla ristampa dei *Frammenti* (Roma, Edizioni di Storia e Letteratura, 2010), Cesare Segre osserva che il metodo di Debenedetti non ebbe fortuna immediata per le suddette ragioni storiche. In un contesto diverso, più stabile, è plausibile che il suo approccio si sarebbe sviluppato parallelamente a quello di Contini, anticipando anche la *critique génétique* francese. Il concetto di genesi è infatti già presente in Debenedetti, che parla di “momento germinativo” dei processi creativi; la sua attenzione ad autografi, trascrizioni a stampa e correzioni rivela un'impostazione chiaramente orientata in senso genetico. Diversamente, il metodo di Contini trovò subito accoglienza, a partire dal 1937, nella “critica delle varianti” della scuola italiana. La divergenza tra i due metodi riflette la diversa postura critica dei due studiosi. Sempre Segre, citando Émile Benveniste, osserva: «un'analisi del fare poetico, con i suoi pentimenti e sviluppi, insomma un'analisi attuata, diacronicamente, sulla corrente stessa dell'enunciazione (Debenedetti), e un'analisi del passaggio del testo da una sincronia ad un'altra, da un enunciato all'altro (Contini). [...] Debenedetti si abbandona alla genesi, la rivive nelle pagine dell'Ariosto; Contini osserva il succedersi degli enunciati cercando di afferrare quel tanto di determinismo che essi permettono di svelare».⁸¹ Debenedetti mostra un coinvolgimento appassionato nello studio della scrittura ariostesca, seguendone da vicino difficoltà e progressi; Contini, al contrario,

⁷⁹ B. Croce, *Ariosto*, Torino, Laterza, 1920 (originariamente apparso in «La Critica»).

⁸⁰ G. Contini, *Come lavorava l'Ariosto*, «Meridiano di Roma», 18 luglio 1937.

⁸¹ C. Segre, *Premessa*, a L. Ariosto, *I frammenti autografi dell'Orlando Furioso*, a cura di S. Debenedetti, Roma, Edizioni di Storia e Letteratura, 2010, p. XI.

mantiene un distacco scientifico, volto alla generalizzazione teorica. Debenedetti immagina Ariosto “mentre compone”, lo coglie nell’atto di interrompere il lavoro per rileggere appunti o conquistare con fatica un risultato; la scrittura, a tratti fluida e immediata, a tratti tormentata, rivela attraverso le varianti le fasi del processo creativo. La sua impostazione si fonda anche sulla competenza filologica: partendo dalla ricostruzione dell’unico manoscritto completo, Debenedetti mostra come Ariosto segnasse i passi problematici, ricorresse a cancellature, organizzasse la pagina con spazi destinati ai rifacimenti. Dalle prime stesure alle copie definitive, egli mette in luce tanto le esitazioni quanto i momenti di invenzione rapida. Il contributo più originale della sua impresa consiste nell’aver inserito lo studio delle correzioni nel quadro complessivo dell’ampliamento del *Furioso*, condotto da Ariosto sia riutilizzando personaggi delle redazioni precedenti, sia introducendone di nuovi, talora in sostituzione di progetti abbandonati come i *Cinque Canti*. Grazie agli autografi, emerge la perizia con cui il poeta perfezionò il tessuto narrativo. Debenedetti permette così al lettore di entrare, con rigore filologico e sensibilità critica, nel laboratorio creativo ariostesco.

In conclusione di questo capitolo, intendiamo fare nostro l’auspicio formulato da Simone Albonico, il quale, in un recente contributo, ha proposto di ripensare la tradizione critica e filologica relativa ai manoscritti di Ludovico Ariosto, con particolare attenzione all’*Orlando furioso* e alle *Satire*.⁸² Lo studioso muove dal riconoscimento di una condizione peculiare: l’archivio ariostesco si presenta frammentario e lacunoso, ma conserva due nuclei di eccezionale rilievo – i *Frammenti* autografi dell’*Orlando furioso* e il manoscritto delle *Satire*. In particolare, il caso del *Furioso* costituisce un momento emblematico della tradizione ecdotica italiana, poiché l’edizione dei *Frammenti* curata da Santorre Debenedetti nel 1937 e la successiva recensione di Gianfranco Contini segnarono l’avvio della cosiddetta “critica delle varianti”. Tuttavia, la lettura continiana – condizionata dall’estetica crociana – privilegiò soltanto alcune varianti ritenute “organicamente valide”, trascurando la complessità del processo creativo documentato dai manoscritti. Come già osservato da Cesare Segre, ne derivò una ricezione squilibrata: mentre l’edizione di Debenedetti, rigorosa nella rappresentazione mimetica dei materiali e pionieristica dal punto di vista genetico, cadde rapidamente nell’ombra, la recensione di Contini si impose come testo fondativo della filologia italiana del Novecento. Albonico invita pertanto a rivalutare i *Frammenti* non solo come semplici testimoni di varianti, ma come veri e propri documenti genetici, capaci di rivelare le strategie di riscrittura, le oscillazioni linguistiche

⁸² S. Albonico, *Dalla parte di Ariosto*, in *Gli “scartafacci” degli scrittori. I sentieri della creazione letteraria in Italia*, a cura di C. Del Vento e P. Musitelli, Roma, Carocci, 2022, pp. 143-62.

e le tensioni stilistiche dell'Ariosto, anche in relazione al dibattito coevo sulla lingua, in particolare alle *Prose* di Bembo. La trascuratezza critica nei loro confronti ha infatti impedito, fino ad oggi, una piena comprensione dei meccanismi di invenzione e di elaborazione poetica ariostesca.

Nella seconda parte del suo contributo, Albonico prende in esame le *Satire*, la cui tradizione manoscritta si fonda soprattutto sul cosiddetto manoscritto F di Ferrara. L'analisi si concentra sulle scelte editoriali di Cesare Segre, che attribuì un peso determinante a criteri linguistici per distinguere le mani (copista/autore), adottando talora soluzioni ricostruttive suscettibili di compromettere l'oggettività documentaria. Anche in questo caso, lo studioso richiama la necessità di un approccio che, senza sacrificare il rigore filologico, sappia valorizzare la stratificazione materiale e redazionale del testo. Sulla scorta di queste considerazioni, Albonico sottolinea l'urgenza di un "ritorno ai manoscritti" ariosteschi, al fine di superare la prospettiva riduttiva imposta dall'estetica crociana e dalle letture semplificanti del Novecento. Solo un rinnovato interesse per la fenomenologia dei testi autografi, intesi come spazi dinamici di scrittura e riscrittura, può consentire di restituire appieno la portata letteraria, storica e linguistica dell'opera ariostesca, aprendo la via a nuove prospettive critiche.

II

CAPITOLO

2.1 Dall'edizione cartacea all'edizione digitale dei frammenti autografi

A quasi un secolo di distanza dalla sua pubblicazione, l'edizione de *I frammenti autografi dell'Orlando Furioso*, a cura di Santorre Debenedetti, rimane ancora il testo di riferimento per tutti gli studiosi che desiderano approfondire la conoscenza della materia, nonché un'opera pionieristica che ha anticipato di qualche decennio la *critique génétique* della scuola francese. Pertanto, il confronto con questa edizione è stato imprescindibile; se, da una parte, il confronto ha permesso di apprezzare le innovative acquisizioni filologiche e metodologiche di Debenedetti, dall'altra, ha messo in evidenza anche i limiti dell'edizione a stampa quando si tratta di rappresentare sulla pagina la complessità dell'elaborazione d'autore. Questi limiti dipendono da alcune caratteristiche intrinseche al supporto cartaceo, quali la staticità, la linearità e la scarsità di spazio a disposizione dell'editore, sia a livello di singolo foglio che di volume. Come si è visto nel capitolo precedente, a monte della sua impresa filologica, Debenedetti ha dovuto compiere una scelta editoriale cruciale tra due forme di rappresentazione testuale mutuamente esclusive a causa della limitatezza del formato libro. Posto di fronte all'alternativa tra una trascrizione strettamente diplomatica e una critico-interpretativa, l'editore ha optato per quest'ultima modalità: ha suddiviso in ottave il testo che negli originali si presenta come un unico blocco continuo, senza spazi fra una ottava e l'altra; ha normalizzato il testo secondo le convenzioni ortografiche e interpuntive moderne; ha corretto gli errori, sacrificando, di fatto, l'originario dettato ariostesco in favore di una maggiore leggibilità del contenuto. Inoltre, Debenedetti ha dovuto ricorrere alle soluzioni tipografiche di cui si è parlato sopra per rappresentare sulla pagina la successione delle varianti d'autore, con risultati che spesso si discostano dalla fisionomia degli originali ariosteschi. L'editore ha infatti optato per una resa tipografica che privilegia la cronologia delle varianti e la fedeltà alle abitudini scritte del poeta, piuttosto che la loro effettiva topografia sulle carte.

A titolo esemplificativo, si consideri il cosiddetto “foglio sciolto” (cc. 18-19), che conserva una copia particolarmente tormentata dell'episodio di Olimpia (Canto XI). Il foglio è già stato

descritto nel Capitolo I e corredato dai relativi facsimili (cfr. Figg. 1.17 e 1.18); lo confrontiamo ora con la resa tipografica dell'edizione debenedettiana:

- Il testo della c. F 18 r. è stato distribuito su tre pagine:
 - p. 21: nella colonna sinistra è riportato il testo delle st. 21, 22 (prima e seconda stesura) e 23; nella colonna destra, invece, sono collocate le correzioni più articolate;
 - p. 22: nella colonna sinistra compare il testo delle st. 24 (prima stesura) e 25, mentre nella colonna destra sono riportate le revisioni della st. 24 (seconda e terza stesura, che nell'originale risultano invertite nella loro collocazione) e le correzioni alla st. 25;
 - p. 28: nella colonna sinistra compare il testo della st. 44, che nell'originale appare nell'angolo superiore destro.

- Il testo della c. F 18 v. è stato distribuito su tre pagine:
 - p. 23: nella colonna sinistra è riportato il testo delle st. 26, 27 e 28; nella colonna destra, invece, sono collocate le revisioni della st. 26 (seconda e terza stesura, che nell'originale appaiono rispettivamente in senso orizzontale e verticale nell'angolo superiore destro della carta) e le correzioni più articolate;
 - p. 24: nella colonna sinistra compare il testo della st. 29 (prima e seconda stesura), mentre nella colonna destra sono riportate le correzioni alla prima stesura;
 - p. 30: nella colonna destra ricorre la st. I del Canto XII con le relative correzioni nell'angolo inferiore destro.

- Il testo della c. F 19 r. è stato distribuito su tre pagine:
 - p. 24: nella colonna sinistra è riportato il testo delle st. 30 e 31; nella colonna destra, le relative correzioni;
 - p. 25: nella colonna sinistra compare il testo delle st. 32, 33 e 34; nella colonna destra, le relative correzioni;

 - p. 29: nella colonna sinistra compare il testo della st. 71 con la relativa revisione (che nell'originale appare nell'angolo superiore destro in senso verticale) e

l'abbozzo della st. 73, collocato nella parte inferiore destra dell'originale.

- Il testo della c. F 19 v. è stato distribuito su tre pagine:
 - p. 25: nella colonna sinistra è riportato il testo della st. 35 e, nella colonna destra, le relative correzioni;
 - p. 26: nella colonna sinistra compaiono le st. 37, 38 e 39, mentre nella colonna destra è collocata l'aggiunta della st. 36;
 - p. 27: nella colonna sinistra compaiono le st. 40 (margine inferiore sinistro dell'originale), 41, 42 e l'abbozzo della st. 43, che nell'originale si trovavano nella colonna destra, benché non fossero aggiunte.

In questo esempio si può osservare come il contenuto delle carte sia stato trascritto senza soluzione di continuità – ad esempio, la p. 24 riunisce stanze appartenenti alle cc. 18 v. e 19 r., così come la p. 25 riunisce stanze appartenenti alle cc. 19 r. e 19 v. – e distribuito su più pagine a causa della scarsità di spazio e della linearità del supporto tipografico. Inoltre, le ottave scritte nella colonna di destra degli originali, quando non costituivano aggiunte al testo di base, nell'edizione debenedettiana sono state spostate nella colonna di sinistra e collocate in sequenza, al fine di agevolare la lettura del processo elaborativo dell'autore. Ciò nondimeno, tale difformità nella resa tipografica impedisce ai lettori – ai quali è preclusa la possibilità di osservare *de visu* gli originali ariosteschi – di farsi un'idea precisa della loro fenomenologia. Di conseguenza, l'edizione digitale dei frammenti autografi è stata concepita con l'intento precipuo di superare, se non del tutto almeno in parte, le criticità sopra evidenziate, coniugando una fedele resa topografica delle varianti con un'accurata ricostruzione dei loro rapporti cronologici. Ciò, pur riconoscendo le sfide che le complesse e spesso inesplicabili stratigrafie correttorie degli autografi ariosteschi pongono all'editore nell'applicazione degli strumenti digitali attualmente disponibili. A questo proposito, è bene sottolineare che la tecnologia digitale non è di per sé una panacea di tutti i problemi che la tecnologia analogica non è stata in grado di risolvere finora; tuttavia, essa può diventare un valido e prezioso ausilio quando è supportata da una solida base filologica.

Prima di passare alla descrizione particolareggiata della metodologia editoriale adottata, sono necessari alcuni chiarimenti preliminari. In primo luogo, la nostra edizione si qualifica come “scientifica” (trad. dall'inglese *scholarly edition*), una locuzione con la quale, in ambito

accademico, si è soliti fare riferimento alla pubblicazione di un testo del passato, allestita secondo criteri scientifici rigorosi e pienamente documentati; infatti, la caratteristica distintiva che eleva un'edizione al rango di "scientifica" è la verificabilità del lavoro editoriale. Questo principio cardine, mutuato dalla ricerca scientifica in generale, stabilisce che qualsiasi risultato debba essere replicabile qualora vengano ripristinate le medesime condizioni di partenza. Nel contesto di un'edizione scientifica, ciò significa che il lettore deve essere in grado di ripercorrere le scelte dell'editore, comprenderne le motivazioni e, potenzialmente, convalidarne o confutarne le conclusioni. Dunque, è questa trasparenza e la possibilità di scrutinio metodologico che conferiscono all'edizione il suo statuto di scientificità.⁸³

Nelle edizioni critiche a stampa tradizionali, la verificabilità dei risultati (laddove per "risultati" si intendono le lezioni testuali che l'editore ha selezionato per la pubblicazione nel testo critico, dopo aver vagliato attentamente una pluralità di varianti, siano esse d'autore e/o di copista) si basa primariamente sulla nota al testo (il cosiddetto "philological statement") e sull'apparato critico. Questi elementi, generalmente collocati a piè di pagina o in appendice al volume, assolvono la funzione di documentare le scelte editoriali; infatti, mentre la nota al testo esplicita in modo dettagliato le decisioni prese riguardo a passaggi specifici, l'apparato critico presenta in modo sistematico le varianti scartate e le loro fonti. Ciò consente al lettore di ripercorrere il processo ecdotico, di comprendere le motivazioni sottese a ogni lezione adottata e, in ultima analisi, di validare criticamente l'affidabilità dell'edizione.⁸⁴

Nelle edizioni digitali, invece, la verificabilità dei risultati presenta un'intrinseca complessità che si articola su due livelli distinti ma strettamente correlati: il *front-end* e il *back-end* del sito web dell'edizione. Sul versante *front-end*, l'attendibilità è garantita dalla presenza di una nota al testo esaustiva e di un apparato critico rigoroso che, come i loro equivalenti cartacei, consentono all'utente di risalire alle fonti, comprendere le scelte editoriali e valutare la coerenza dell'edizione rispetto alla tradizione testuale.⁸⁵ Parallelamente, il *back-end* introduce un ulteriore

⁸³ Per approfondire, cfr. T. Mancinelli e E. Pierazzo, *Che cos'è un'edizione scientifica digitale*, Roma, Carocci Editore, 2020 pp. 9-18; P. Sahle, What is a Scholarly Digital Edition?, in AA.VV., *Digital Scholarly Editing: Theory and Practice*, a cura di M. J. Driscoll e E. Pierazzo, Cambridge, Open Book Publisher, 2016, pp. 19-40.

⁸⁴ Per approfondire questo argomento, cfr. P. Maas, *La critica del testo*, trad. a cura di Giorgio Ziffer, Roma, Edizioni di Storia e Letteratura, 2021; A. Stussi, *L'edizione*, in *Introduzione agli studi di filologia italiana*, Bologna, Il Mulino, 2006, pp. 115-146.

⁸⁵ A questo proposito, è bene sottolineare che - malgrado la divergenza di opinioni - la nota al testo e l'apparato critico continuano ad assolvere una funzione cruciale nelle edizioni digitali, sia che si tratti di edizioni di filologia della copia che di edizioni di filologia d'autore; nel primo caso l'apparato riporterà le varianti di copista tratte dai diversi testimoni manoscritti di uno stesso testo, nel secondo, invece, l'apparato riporterà le varianti d'autore che intercorrono tra le diverse redazioni, manoscritte o a stampa, di una medesima opera letteraria. Ciò che cambia sono gli strumenti e le modalità adottate di volta in volta dall'editore nella progettazione della GUI (*Graphic User Interface*) dell'edizione digitale, in funzione degli obiettivi scientifici perseguiti. Ad esempio, la nota al testo e

strato di trasparenza e verificabilità. Su questo versante, la possibilità per l'utente di accedere al codice sorgente dell'edizione e osservare la concreta applicazione del modello di codifica adottato dall'editore, assume un'importanza cruciale; ciò consente di estendere la verifica a livello infrastrutturale, poiché l'utente può analizzare le logiche di rappresentazione dei dati e validare l'aderenza dell'implementazione digitale ai principi filologici dichiarati (cfr. § 2.4).⁸⁶

Dunque, la combinazione di questi due approcci – gli elementi ecdotici tradizionali visibili all'utente nella GUI (*Graphic User Interface*) dell'edizione e la trasparenza del codice sottostante – è fondamentale per assicurare l'affidabilità delle edizioni digitali nell'ambito della ricerca accademica. Ration per cui, nella realizzazione del nostro progetto editoriale, abbiamo adottato due soluzioni complementari. Da una parte, abbiamo predisposto un sito web dedicato che ha “rimediato” la classica nota al testo dell'editore.⁸⁷ In questo spazio digitale, gli utenti possono consultare i riferimenti bibliografici e tutte le informazioni necessarie relative alla storia e alla fenomenologia degli autografi ariosteschi, alle abitudini compositive e correttorie del poeta, nonché alla metodologia digitale adottata, comprensiva di una descrizione circostanziata del modello di codifica. Una sezione specifica del sito rimanda alla piattaforma dell'edizione digitale propriamente detta, nella quale il testo di base – trascritto in forma diplomatica e, auspicabilmente, anche in versione normalizzata – sarà corredato da un apparato critico verticale e fotografico, la cui configurazione varierà in funzione del software di visualizzazione adottato. L'apparato critico verticale costituisce una modalità di rappresentazione delle varianti che dispone il testo base e le sue lezioni alternative in colonne o livelli paralleli, sovrapposti l'uno all'altro, anziché in calce o a piè di pagina. Le varianti non vengono semplicemente

l'apparato critico potrebbero configurarsi come collegamenti ipertestuali a pagine web dedicate oppure come finestre di dialogo a comparsa, in sovrapposizione al testo critico, per focalizzare l'attenzione dell'utente su una determinata lezione testuale.

Per approfondire questo argomento, cfr. P. Robinson, *What is a critical digital edition?*, «Variants: The Journal of the European Society for Textual Scholarship», 1, 2002, pp. 43–62; E. Pierazzo, *Modelling Digital Scholarly Editing: From Plato to Heraclitus*, in AA.VV., *Digital Scholarly Editing: Theory and Practice*, a cura di M. J. Driscoll e E. Pierazzo, Cambridge, Open Book Publisher, 2016, pp. 41–58; M. Buzzoni, *A Protocol for Scholarly Digital Editions? The Italian Point of View*, in AA.VV., *Digital Scholarly Editing: Theory and Practice*, a cura di M. J. Driscoll e E. Pierazzo, Cambridge, Open Book Publisher, 2016, pp. 59–82; P. Italia, G. Raboni et al., *Authorial philology in the digital era*, in *What is Authorial Philology*, Cambridge, Open Book Publishers, 2021, pp. 18–28.

⁸⁶ Per approfondire questo argomento, cfr. P. Italia, *Editing Duemila. Per una filologia dei testi digitali*, Roma, Salerno Editrice, 2020, pp. 171–185.

⁸⁷ Con il termine “rimediazione” (trad. dall'inglese *remediation*) si intende il processo attraverso il quale un medium ne assorbe, ne rappresenta o ne trasforma un altro. Questo concetto, sviluppato principalmente da Jay David Bolter e Richard Grusin prendendo le mosse dalle teorie di Marshall McLuhan, descrive come i nuovi media non emergano dal nulla, ma piuttosto rielaborino e riconfigurino le forme e le convenzioni dei media precedenti. In conclusione, la rimediazione è un meccanismo costante di evoluzione mediatica, in cui ogni nuovo medium commenta, critica e riorganizza i suoi predecessori, creando un ciclo continuo di innovazione e riferimento reciproco. Per approfondire l'argomento, cfr. M. McLuhan, *Gli strumenti del comunicare*, Milano, Il Saggiatore, 2011; J.D. Bolter e R. Grusin, *Remediation. Competizione e integrazione fra media vecchi e nuovi*, Milano, Guerini e Associati, 2003.

elencate, ma collocate in corrispondenza diretta con i segmenti testuali di riferimento, favorendo una lettura comparata e analitica. Nell'ambito delle edizioni digitali, tale sistema si traduce in una visualizzazione stratificata o multilivello, che consente al lettore di confrontare dinamicamente il testo di base con le sue varianti. Questo tipo di rappresentazione può essere implementato tramite codifica XML-TEI, poiché il modello TEI consente di strutturare il testo e le sue varianti in maniera gerarchica e marcata, mentre gli ambienti di visualizzazione dedicati, come *EVT*, sfruttano proprio la struttura TEI multilivello per permettere all'utente di espandere o contrarre informazioni testuali o apparati critici, in funzione delle esigenze di lettura o di analisi (cfr. § 2.24 e 2.2.5).

L'apparato fotografico, invece, accompagna il testo con riproduzioni digitali dei testimoni – manoscritti o stampe – che ne costituiscono la base documentaria. Esso ha la funzione di documentare visivamente le lezioni e le varianti, consentendo la verifica diretta della trascrizione sul facsimile. Nelle edizioni digitali contemporanee, tale apparato è frequentemente implementato attraverso viewer IIIF (come quelli integrati in *EVT* o *Transkribus*), che permettono di visualizzare, ingrandire e confrontare le immagini dei testimoni con la trascrizione o con l'apparato critico stesso. L'apparato fotografico risulta, pertanto, complementare a quello verticale, fornendo la necessaria base visiva e diplomatica per la verifica scientifica del testo.⁸⁸ Dall'altro lato, è prevista la pubblicazione del codice sorgente dell'edizione digitale in modalità *open source*, per favorire la trasparenza metodologica e la collaborazione scientifica. Il codice potrà essere reso disponibile su piattaforme di versionamento e condivisione come *GitHub* o

⁸⁸ Secondo P. Italia e G. Raboni, l'apparato verticale «consiste nel mettere in colonna e rappresentare verticalmente tutte le correzioni di un verso, dalla prima all'ultima, indicando con marcatori tipografici opportunamente scelti (corsivi, neretti) le cassature che hanno portato alla scrittura di una nuova lezione o gli inserimenti di nuove porzioni di testo. Il testo di riferimento può essere l'ultimo ricostruibile dal manoscritto o il primo, e può essere evidenziato da ulteriori marcatori tipografici, oppure riprodotto integralmente nella medesima pagina del testo incolonnato o all'inizio dell'edizione, come punto di partenza o d'arrivo, da leggere distesamente. La verticalità coincide con la temporalità; i versi o le porzioni di verso non soggetti a variante non sono ripetuti, in modo da rendere immediatamente evidente dove e come si impianta la correzione. È evidente come questo tipo di apparato sia adottabile esclusivamente per testi in poesia perché il verso non eccede la dimensione di una riga tipografica e consente l'incolonnamento delle varianti nelle righe seguenti, soluzione impossibile per un testo in prosa. Questo tipo di apparato è stato utilizzato per la prima volta per i *Canti* di Leopardi nell'edizione di Moroncini del 1927, ed è stato seguito in altre importanti edizioni di testi poetici». Le note filologiche, invece, «hanno il compito di spiegare più analiticamente la fenomenologia dei manoscritti e la dinamica delle correzioni, proponendo varie interpretazioni del testo, oppure indicando i casi di dubbia lettura. Nelle note filologiche si può inserire, sotto forma di commento, tutto ciò che si vorrebbe dire a giustificazione di determinate scelte di apparato, ma che non è possibile costringere nel breve spazio al piede del testo».

Cfr. P. Italia, G. Raboni, *Che cos'è la filologia d'autore*, Roma, Carocci, 2010, pp. 50-52; ivi, p. 68.

GitLab.⁸⁹ Le convenzioni editoriali saranno documentate all'interno del <teiHeader>, secondo le specifiche TEI, in particolare:

- nell'elemento <encodingDesc>, destinato a descrivere le scelte di codifica (criteri TEI, politiche di normalizzazione, rappresentazione delle varianti, ecc.);
- nel sotto-elemento <editorialDecl>, che specifica le politiche editoriali seguite, come la normalizzazione ortografica, la gestione delle abbreviazioni e l'uso della punteggiatura (cfr. § 4.1).

Tale impostazione garantirà la tracciabilità delle decisioni editoriali e la riproducibilità del modello di codifica all'interno della comunità scientifica.

In secondo luogo, la nostra edizione rappresenta un vero e proprio esempio di filologia digitale e computazionale per la scelta e l'applicazione di strumenti e metodi di cui si parlerà più avanti. A tal proposito, è utile richiamare la distinzione – elaborata da Elena Pierazzo sulla scorta della classificazione proposta da Patrick Sahle – tra filologia digitalizzata, filologia digitale e filologia computazionale.⁹⁰ La filologia digitalizzata integra strumenti e tecniche digitali all'interno di un flusso di lavoro che mantiene un impianto metodologico prevalentemente tradizionale. Le edizioni prodotte in questo ambito somigliano spesso a edizioni a stampa, sia nel loro funzionamento che nella presentazione dei risultati, come nel caso di versioni PDF stampabili. In sostanza, la tecnologia digitale è usata per migliorare o accelerare metodi editoriali preesistenti, ma non per ridefinire la natura stessa dell'edizione. Al contrario, la filologia digitale comprende pratiche editoriali innovative che sono nate e modellate dal digitale. Le edizioni autenticamente digitali stravolgono il tradizionale concetto di “pagina” e, per loro stessa natura,

⁸⁹ Tra le piattaforme per la pubblicazione e la gestione collaborativa del codice sorgente, *GitHub* rappresenta la soluzione più diffusa e consolidata nel campo delle edizioni digitali e della filologia computazionale. Gratuita e stabile, è adottata da numerosi enti di ricerca e biblioteche digitali per le sue funzionalità di version control, gestione delle revisioni e tracciamento delle modifiche (*issue tracking*). Consente inoltre di archiviare e condividere codice, file XML-TEI, metadati e documentazione, nonché di realizzare siti web o dimostrazioni dell'edizione tramite *GitHub Pages*. L'integrazione con *Zenodo* permette infine di attribuire un DOI alle release, garantendo la citabilità accademica dei contenuti. Un'alternativa equivalente è costituita da *GitLab*, piattaforma open source con funzioni analoghe, particolarmente indicata per progetti interni o collaborazioni su scala istituzionale. In presenza di un'installazione locale (ad esempio presso un'università o un centro di ricerca), *GitLab* offre un controllo più rigoroso sui dati e una gestione avanzata dei permessi, pur risultando generalmente meno visibile pubblicamente rispetto a *GitHub*.

⁹⁰ Per approfondire questo argomento, cfr. T. Mancinelli e E. Pierazzo, *Che cos'è un'edizione scientifica digitale*, Roma, Carocci Editore, 2020 pp. 9-18; P. Sahle, *What is a Scholarly Digital Edition?*, in AA.VV., *Digital Scholarly Editing: Theory and Practice*, a cura di M. J. Driscoll e E. Pierazzo, Cambridge, Open Book Publisher, 2016, pp. 19-40.

non possono essere stampate senza una significativa perdita di informazione e funzionalità; la multimedialità, l'interattività, la stratificazione dei dati e la possibilità di navigazione complessa – ovvero la loro intrinseca ipertestualità – sono caratteristiche essenziali di queste edizioni, che non trovano corrispondenza nel formato cartaceo.⁹¹

A queste due categorie, è possibile aggiungerne una terza, la cosiddetta “filologia computazionale”, cui afferiscono pratiche editoriali che sfruttano metodi informatici avanzati, come algoritmi di intelligenza artificiale (IA) applicati alla *computer vision* e al *natural language processing* (NLP).⁹² A scopo esemplificativo, possiamo menzionare: la collazione assistita da computer, che consente di automatizzare il confronto tra diversi testimoni di uno stesso testo per identificare le varianti; l'*Optical Character Recognition* (OCR) e la *Handwritten Text Recognition* (HTR), che consente la lettura e la trascrizione automatica dei testi manoscritti e a stampa.⁹³ In seguito, i dati ottenuti con questi sistemi di apprendimento automatico (trad.

⁹¹ Parafrasando Patrick Sahle, Paola Italia ricorda che «le edizioni scientifiche digitalizzate, tuttavia, non devono essere confuse con le edizioni scientifiche digitali. Mentre l'edizione digitalizzata non differisce da quella cartacea che per il supporto, e una può essere facilmente convertita nell'altra senza perdita di dati fondamentali (si può stampare il PDF di un'edizione scientifica, ottenendo la sua forma cartacea, o pubblicare una edizione critica cartacea in un sito digitale), l'edizione critica digitale non può essere convertita in un'edizione critica cartacea perché perderebbe quegli elementi di interattività e dinamicità che costituiscono la peculiarità dell'edizione stessa». Cfr. P. Italia, *Editing Duemila. Per una filologia dei testi digitali*, Roma, Salerno Editrice, 2020, p. 61.

⁹² La *computer vision* è una branca dell'intelligenza artificiale (IA) che si propone di dotare i sistemi informatici della capacità di “vedere” e “comprendere” il mondo visivo in modo simile all'organo biologico della vista. La visione artificiale non si limita ad acquisire immagini e video digitali, ma elabora, analizza e interpreta questi dati per estrarre informazioni significative e utili a prendere decisioni o compiere azioni. A tal fine, si avvale ampiamente di tecniche di *machine learning*, in particolare del *deep learning* (come le Reti Neurali Convoluzionali, o CNNs), che vengono addestrate su vasti *dataset* di immagini etichettate per riconoscere pattern e caratteristiche visive complesse.

Il *natural language processing* (NLP) è un campo di applicazione dell'intelligenza artificiale (IA), che si propone di dotare i computer della capacità di comprendere, interpretare e generare linguaggio umano, sia esso scritto o parlato. In pratica, il suo obiettivo fondamentale è superare il divario tra la comunicazione umana e quella computazionale, trasformando il linguaggio in dati strutturati. Questo campo abbraccia due macro-aree principali: la *natural language understanding* (NLU) che concerne la capacità del computer di decifrare il significato del testo o del parlato umano, superando le ambiguità e cogliendo le sfumature semantiche e contestuali; e la *natural language generation* (NLG) che attiene alla capacità del computer di produrre testo o parlato coerente, grammaticalmente corretto e rilevante per il contesto, spesso a partire da dati strutturati o da input specifici. Il NLP attinge conoscenze e tecniche da diverse discipline, tra cui la linguistica computazionale, l'informatica, la statistica e il *machine learning* (in particolare il *deep learning*). I modelli di NLP vengono addestrati su vasti *dataset* di testo e voce per riconoscere pattern, regole grammaticali, relazioni semantiche e contesti d'uso.

Per approfondire l'argomento, cfr. E. Jezek e R. Sprugnoli, *Linguistica computazionale. Introduzione all'analisi automatica dei testi*, Bologna, Il Mulino, 2023.

⁹³ Sia la *computer vision* che il NLP trovano un'applicazione sempre più cruciale ed integrata nell'ambito dell'OCR/HTR. Mentre la *computer vision* interviene nella fase iniziale dell'OCR/HTR analizzando visivamente le immagini del testo scritto per rilevare i caratteri, il NLP entra in gioco successivamente per affinare la trascrizione. Infatti, una volta che la *computer vision* ha estratto sequenze di caratteri o potenziali parole, il NLP utilizza modelli linguistici per migliorare l'affidabilità della trascrizione. Ad esempio, se l'algoritmo di visione non è sicuro se un carattere sia una “o” o una “a” in una parola, il modello linguistico può analizzare il contesto e le probabilità delle parole nella lingua in questione per suggerire la lettera più probabile. Se “c*sa” è l'output provvisorio, un modello linguistico capirà che “casa” o “cosa” sono più probabili di “cxa”.

dall'inglese *machine learning*) possono essere estrapolati e sottoposti ad ulteriori analisi computazionali per mezzo di algoritmi di *data mining*, al fine di generare ipotesi e supportare decisioni filologiche su larga scala.⁹⁴

A corollario della distinzione tra le diverse accezioni di filologia nell'era digitale, possiamo affermare, parafrasando Elena Pierazzo, che la filologia digitale e quella computazionale non si limitano a produrre risultati innovativi solo sul piano ecdotico. A differenza della filologia digitalizzata che - come abbiamo detto sopra - si concentra principalmente sull'applicazione di strumenti digitali a pratiche editoriali tradizionali e consolidate, queste discipline generano progresso anche sul piano informatico; questo progresso si sostanzia nella creazione di nuovi software (ad esempio, per la trascrizione e la collazione automatica, per la visualizzazione e pubblicazione online di edizioni digitali, ecc.), nello sviluppo di modelli di codifica all'avanguardia e nell'implementazione di sistemi di meta-datazione innovativi, per rispondere alle esigenze specifiche dei vari progetti editoriali. Questo duplice impatto, sia sul versante ecdotico che su quello informatico, fa sì che tali discipline si impongono come catalizzatori di una trasformazione più profonda nella gestione e nell'analisi dei testi storico-letterari.

Infine, la nostra edizione scientifica digitale si qualifica come “genetica”, sullo stampo dell'*édition génétique* francese; infatti, l'edizione pubblica in blocco tutti i materiali preparatori che costituiscono il cosiddetto “avantesto” dell'*Orlando furioso*. Questi materiali presentano diversi gradi di elaborazione letteraria e commensurabilità del testo con la redazione definitiva del poema del 1532; il contenuto dell'edizione spazia dalle carte autografe che documentano la gestazione dei nuovi episodi introdotti dal poeta, alle carte che testimoniano le successive rielaborazioni degli stessi episodi, fino alle copie in pulito pronte per essere inviate in tipografia. Inoltre, prendendo spunto dalle ipotesi interpretative di Debenedetti sull'iter elaborativo dei testi, l'edizione cerca di districare e organizzare le complesse stratigrafie correttorie degli autografi ariosteschi, formulando congetture sulla corretta successione cronologica delle varianti d'autore.

⁹⁴ Il *data mining* è un processo computazionale che, applicando tecniche statistiche e di *machine learning*, estrae pattern, tendenze e conoscenze significative da grandi volumi di dati. Il suo obiettivo è trasformare i dati grezzi in informazioni utili e azionabili, supportando la previsione di comportamenti, l'ottimizzazione di processi e la rilevazione di anomalie.

2.2 Metodologia dell'edizione scientifica digitale: il modello sorgente-output ⁹⁵

Dal punto di vista metodologico, l'edizione genetica digitale dei frammenti autografi aderisce al modello cosiddetto "sorgente-output" che accomuna tutte le edizioni scientifiche digitali (ESD) in circolazione. Infatti, pur manifestando una notevole eterogeneità in termini di resa testuale e strumenti analitici, le ESD rivelano una certa omogeneità metodologica riconducibile proprio alla predominanza di questo paradigma concettuale e tecnologico.⁹⁶

Secondo il modello sorgente-output, il testo codificato (cfr. § 2.4 e Cap. IV) è archiviato in uno o più file che nell'insieme costituiscono il codice sorgente. Nella maggior parte dei contesti, questi file sorgente sono codificati con il linguaggio di marcatura XML (*Extensible Markup Language*), conformemente alle direttive della TEI (*Text Encoding Initiative*). La TEI, riconosciuta come standard *de facto* nel campo delle edizioni digitali, funge anche da fondamentale riferimento intellettuale. Un insieme di programmi informatici elabora il sorgente come input, generando una pluralità di oggetti digitali in uscita (gli output). A titolo esemplificativo, da un singolo sorgente XML è possibile derivare più versioni HTML per la pubblicazione sul web, documenti PDF stampabili, eBook ottimizzati per tablet, indici interattivi e molto altro. Inoltre, l'efficacia di questo approccio risiede nella sua capacità di produrre, a partire dal medesimo sorgente, diverse tipologie di edizioni, quali: edizioni diplomatiche che riproducono fedelmente la fisionomia dei documenti e riportano il testo così come si legge negli originali; edizioni interpretative che normalizzano il testo secondo le convenzioni ortografiche e interpuntive moderne; edizioni critiche che, attraverso un'analisi comparativa di tutte le fonti disponibili (manoscritte, stampe antiche, ecc.), punta a ricostruire la forma più vicina possibile all'originale voluto dall'autore o a stabilire la versione più autorevole e storicamente significativa; edizioni dei testimoni che incamerano l'intera tradizione testuale di un'opera letteraria; descrizioni codicologiche e ricostruzioni dello *stemma codicum*.

L'editore scientifico gode di una libertà pressoché illimitata nell'incorporare informazioni nel sorgente, sottostando unicamente a due vincoli: la leggibilità intrinseca del file sorgente e il tempo disponibile per la sua creazione. Questi vincoli, sebbene possano apparire esclusivamente di natura tecnica, assumono una rilevanza strategica, nella misura in cui le risorse temporali ed

⁹⁵ Per approfondire questo argomento, cfr. T. Mancinelli e E. Pierazzo, *Che cos'è un'edizione scientifica digitale*, Roma, Carocci Editore, 2020.

⁹⁶ Per avere una cognizione generale del panorama delle edizioni digitali, cfr. AA.VV., *Edizioni critiche digitali. Edizioni a confronto*, a cura di P. Italia e C. Bonsi, Roma, Sapienza Università Editrice, 2016, consultabile al seguente link: <<https://doi.org/10.13133/9788893770033>> (ultimo accesso 17 ottobre 2025).

economiche a disposizione dell'editore scientifico sono limitate ed impongono, pertanto, una gestione efficiente e razionale del lavoro. Praticamente, ciò si traduce nella selezione accurata degli elementi da valorizzare attraverso la codifica, prediligendo quelli ritenuti fondamentali per il conseguimento degli obiettivi scientifici del progetto editoriale. Di conseguenza, la fase preliminare e cruciale del processo editoriale consiste nella modellizzazione concettuale dell'edizione da produrre.⁹⁷

2.2.1 Modellizzazione del processo editoriale

Nel contesto delle edizioni scientifiche, la modellizzazione rappresenta la fase preliminare e fondativa del processo editoriale, in cui si analizza l'oggetto di studio prescelto e si definiscono le strategie di rappresentazione più adeguate. Essa consiste nell'identificare, selezionare e formalizzare le caratteristiche ritenute essenziali ai fini dell'indagine, escludendo al contempo gli elementi considerati marginali rispetto agli obiettivi di ricerca. Tale attività, già centrale nelle edizioni tradizionali, assume un ruolo ancora più rilevante in ambito digitale, dove ogni processo computazionale – dalla riproduzione fotografica digitale delle fonti documentarie, alla trascrizione automatica del contenuto testuale mediante sistemi di OCR/HTR, alla codifica XML-TEI, fino alla pubblicazione su piattaforme interattive – presuppone l'esistenza di modelli precisi e ripetibili. Tali modelli, una volta formalizzati, si traducono in algoritmi, ovvero sequenze strutturate di operazioni basate su regole esplicite, che possono essere eseguite dal calcolatore. È importante sottolineare che un modello non mira a riprodurre integralmente la complessità dell'oggetto che rappresenta, bensì a semplificarla secondo un punto di vista selettivo e funzionale al raggiungimento di un determinato scopo. Questa riduzione intenzionale della complessità è ciò che consente al modello di essere efficace. Una carta geografica costituisce un modello del territorio nella misura in cui ne rappresenta selettivamente alcune caratteristiche, in funzione dell'uso per cui è stata realizzata. Ad esempio, una carta fisica mostra elementi naturali come montagne, fiumi, pianure ecc., ma, a differenza di una carta politica, non include confini, divisioni amministrative o insediamenti umani. Questa mancanza non è un limite, bensì una scelta funzionale al tipo di analisi territoriale per cui la carta è concepita. Allo stesso modo, un'edizione scientifica rappresenta un modello del testo per diversi motivi: si basa

⁹⁷ E. Pierazzo, *Digital Scholarly Editing: Theories, Models and Methods*, Farnham, Ashgate Publishing, 2015, consultabile al seguente link: <<https://hal.univ-grenoble-alpes.fr/hal-01182162/document>> (ultimo accesso 17 ottobre 2025).

su una selezione ragionata dei testimoni disponibili; riflette l'approccio critico dell'editore; adotta convenzioni editoriali consolidate (come l'edizione critica, diplomatica o semidiplomatica); e propone una selezione mirata dei fenomeni testuali, come la scelta delle lezioni da inserire nel testo principale e delle varianti da riportare nell'apparato, oltre alla definizione di una specifica *facies* linguistica.

Nel caso specifico dell'edizione digitale dei frammenti autografi dell'*Orlando Furioso* di Ludovico Ariosto, la modellizzazione del testo richiede una serie di decisioni puntuali e tra loro strettamente connesse, tra cui:

- la selezione dei frammenti autografi rilevanti, con l'esclusione di copie non autografe (come, ad esempio, il manoscritto apografo dell'Ambrosiana che contiene alcuni passi dell'episodio della Rocca di Tristano) e delle redazioni a stampa successive, come quella del 1532;
- l'identificazione dei fenomeni testuali da annotare, quali varianti, correzioni, segni di scrittura e cancellature;
- la definizione delle categorie testuali oggetto di trascrizione, come il testo principale, le annotazioni marginali o interlineari, e i simboli grafici;
- la selezione dei marcatori XML-TEI per la codifica strutturata dei dati linguistici, grafici, paleografici e codicologici;
- la scelta tra una trascrizione diplomatica (o semidiplomatica), orientata a finalità filologiche, e una trascrizione interpretativa, con normalizzazione linguistica e apparati esplicativi, più adatta a scopi didattici;
- la progettazione dell'interfaccia di visualizzazione, che determina le modalità di interazione tra utente e contenuto (ad esempio, funzionalità di zoom, sincronizzazione testo-immagine, apparati dinamici, ecc.).

L'insieme di queste operazioni concorre alla costruzione di un modello concettuale e formale dell'edizione: un sistema coerente di principi e criteri che consente di rappresentare il testo in ambiente digitale in modo sistematico, analitico e compatibile con l'elaborazione computazionale.

Come ha evidenziato Willard McCarty, i modelli impiegati nell'ambito delle scienze umane computazionali si distinguono in due categorie: il “modello di qualcosa” (*model-of*) e il “modello per qualcosa” (*model-for*). Il primo ha finalità descrittive e mira a rappresentare fedelmente un oggetto o un insieme di conoscenze; il secondo ha finalità operative e progettuali e serve a costruire un nuovo artefatto, digitale o concettuale. Nel caso della nostra edizione digitale, il *model-of* consiste nella descrizione del manoscritto autografo in termini linguistici, paleografici e codicologici; il *model-for* guida invece la progettazione dell'edizione, determinando formati, strumenti e modalità di fruizione.⁹⁸ Infine, è fondamentale sottolineare il carattere iterativo del processo di modellizzazione. Raramente un modello iniziale riesce a cogliere tutte le complessità del materiale; nuove acquisizioni – come la scoperta di ulteriori frammenti o la revisione di ipotesi filologiche – possono rendere necessarie modifiche sostanziali al modello stesso. In ambito digitale, tali revisioni comportano la rielaborazione di procedure già eseguite (riproduzione fotografica digitale, trascrizione, codifica, riconfigurazione del software di visualizzazione), con inevitabili costi in termini di tempo e risorse. Per questo motivo, è essenziale una pianificazione iniziale accurata, volta ad anticipare il maggior numero possibile di variabili e casi limite, così da contenere la necessità di rielaborazioni successive e garantire la sostenibilità del progetto editoriale. Nel seguito del capitolo, offriremo una breve panoramica delle principali fasi che caratterizzano la metodologia delle edizioni digitali, sottolineando come la modellizzazione costituisca un'attività trasversale a ogni fase del processo.

2.2.2 Digitalizzazione delle fonti documentarie come immagini ad alta definizione

La genesi della nostra edizione ha preso le mosse dalla digitalizzazione degli autografi ariosteschi. In termini generali, la digitalizzazione può essere definita come un processo che consiste in una serie di operazioni volte a convertire un documento analogico (sia esso un foglio manoscritto, un rotolo, un volume rilegato o persino una superficie muraria) in un oggetto digitale. Questa conversione può avvenire in due modi distinti e complementari: mediante la riproduzione facsimilare del documento originale sotto forma di immagine fotografica ad alta risoluzione, oppure attraverso la codifica del testo come sequenza di caratteri alfanumerici riconoscibili dalla macchina (la cosiddetta “Machine Readable Form” o “MRF”). Ogni processo

⁹⁸ Per approfondire l'argomento, cfr. W. McCarty, *Humanities Computing*, Basingstoke-New York, Palgrave Macmillan, 2005.

di digitalizzazione solleva particolari problemi ed è sempre un'operazione intrinsecamente selettiva, che pone in risalto determinate caratteristiche dell'oggetto di studio a scapito di altre. In buona sostanza, ogni digitalizzazione è la concreta espressione di un modello esplicito del testo stesso.

In questo paragrafo, ci soffermeremo sulla prima modalità, contestualizzando la riproduzione facsimilare dei frammenti autografi nel quadro più ampio della trasformazione digitale del patrimonio culturale testuale italiano, avviata dal Ministero della Cultura (MiC). L'avvento della fotografia digitale ha rappresentato un punto di svolta cruciale: in un lasso di tempo sorprendentemente breve, essa ha soppiantato le altre tecniche di riproduzione per via delle sue prestazioni tecnologiche qualitativamente superiori, la sua semplicità di utilizzo e accessibilità economica. Congiuntamente alla disponibilità di sistemi di *hosting* online per l'archiviazione e la gestione dei dati, la fotografia digitale ha rivoluzionato non solo la gestione del patrimonio culturale testuale da parte di biblioteche ed archivi, ma ha anche trasformato radicalmente le metodologie di lavoro dei ricercatori. Inoltre, le sue ripercussioni si sono estese profondamente alla pedagogia di discipline umanistiche fondamentali come la filologia, la paleografia, la codicologia e la diplomatica, aprendo scenari precedentemente inimmaginabili.

Nell'ultimo ventennio, MiC ha posto un'enfasi crescente sulla digitalizzazione dei fondi librari e archivistici conservati negli istituti italiani, riconoscendo l'importanza di preservare e rendere accessibile la memoria storica e culturale del Paese. Questa spinta verso la transizione digitale è stata dettata dalla necessità di salvaguardare materiali fragili, ampliare l'accesso ai contenuti, favorire la ricerca e la fruizione da parte di un pubblico più ampio.⁹⁹

Le direttrici sulle quali il MiC ha operato per incentivare la digitalizzazione sono state molteplici. In primo luogo, il MiC ha promosso diversi piani strategici nazionali volti a coordinare progetti di digitalizzazione su larga scala. Tra le iniziative più recenti e rilevanti, merita un'attenzione particolare il *Piano Nazionale di Digitalizzazione del Patrimonio Culturale* (PND) che rappresenta la visione strategica con la quale il MiC intende promuovere e organizzare il processo di transizione digitale del patrimonio culturale italiano nel quinquennio

⁹⁹ Per completezza, è doveroso menzionare la Biblioteca Apostolica Vaticana (BAV), sebbene non sia un'istituzione statale italiana. Notoriamente, la BAV detiene un patrimonio di importanza universale e in anni recenti ha condotto una delle più ambiziose e sofisticate campagne di digitalizzazione di manoscritti e incunaboli. L'obiettivo della BAV è quello di preservare e rendere accessibili i suoi inestimabili fondi attraverso l'uso di tecnologie avanzate (come gli scanner *Hasselblad* ad alta risoluzione), protocolli di conservazione digitale a lungo termine (come i file FITS) e framework di condivisione (come l'*International Image Interoperability Framework* - IIF). A tal fine, ha collaborato con partner tecnologici come EMC e Dedagroup, e più recentemente con NTT DATA per esplorare nuove tecnologie come il *Web3/NFT* per la promozione e la raccolta fondi. La Digital Vatican Library (DVL) è consultabile al seguente link: <<https://digi.vatlib.it/>> (ultimo accesso 17 ottobre 2025).

2022-2026. Il PND è stato elaborato, coordinato e promosso dall'Istituto Centrale per la Digitalizzazione del Patrimonio Culturale – Digital Library (comunemente chiamato “Digital Library”), che riveste un ruolo centrale nell’indirizzare e vigilare sulle iniziative di digitalizzazione del Ministero. L’obiettivo primario del *Piano* è quello di superare la frammentazione che ha caratterizzato in passato i progetti di digitalizzazione, creando un ecosistema culturale, tecnologico e scientifico unitario e coerente per l’attuazione della transizione digitale. I destinatari del PND sono principalmente i musei, gli archivi, le biblioteche, gli istituti centrali, le soprintendenze e tutti i luoghi della cultura pubblici che conservano, gestiscono e valorizzano i beni culturali (cfr. Cap. III).¹⁰⁰

Per l’implementazione dei piani strategici sopra menzionati, il MiC ha stanziato finanziamenti specificamente destinati a progetti di digitalizzazione. I finanziamenti sono stati assegnati sia tramite bandi pubblici rivolti a biblioteche, archivi e altri istituti di conservazione, sia attraverso contributi diretti a istituzioni di particolare importanza. Essi hanno permesso di sostenere i costi connessi all’acquisto di attrezzature tecnologiche, al reclutamento di personale specializzato e allo sviluppo di piattaforme digitali. A tal proposito, ricordiamo che il suddetto *Piano Nazionale di Digitalizzazione del Patrimonio Culturale* (PND) si colloca all'interno del più ampio *Piano Nazionale di Ripresa e Resilienza* (PNRR), il programma di riforme e investimenti che è stato approvato dalla Commissione Europea nel giugno 2021 per favorire la modernizzazione del Paese; infatti, il PNRR fornisce le risorse economiche e le strategie operative per supportare e accelerare i processi di digitalizzazione delineati dal PND nel settore culturale.¹⁰¹

¹⁰⁰ Il PND è articolato in tre sezioni interconnesse, che concorrono a definire il percorso di digitalizzazione.

- La *Visione* prefigura la trasformazione e le opportunità offerte dal digitale, indicando gli obiettivi a lungo termine; si concentra sul valore del patrimonio culturale come generatore di identità e sviluppo, e sulla necessità di renderlo pienamente fruibile nell’era digitale.
- La *Strategia* definisce il percorso concreto per implementare e conseguire gli obiettivi della *Visione*. In questa sezione, sono descritte le azioni e le iniziative necessarie per la digitalizzazione, la gestione delle risorse digitali, l’aggiornamento delle competenze e lo sviluppo di servizi innovativi.
- Le *Linee Guida* rappresentano gli strumenti operativi che supportano la pianificazione e l’esecuzione delle attività legate alla digitalizzazione del patrimonio e alla trasformazione digitale degli istituti culturali. In particolare, esse forniscono indicazioni metodologiche e tecniche precise per la riproduzione digitale (formati, risoluzioni), la descrizione e l’indicizzazione dei contenuti (metadati), e la conservazione a lungo termine degli oggetti digitali.

Per approfondire l’argomento, cfr. i seguenti link: <<https://digitallibrary.cultura.gov.it/>> (ultimo accesso 17 ottobre 2025).

¹⁰¹ Il PNRR, acronimo di *Piano Nazionale di Ripresa e Resilienza*, è il programma di investimenti e riforme che l’Italia ha predisposto per accedere ai fondi del *Next Generation EU* (NGEU), lo strumento di ripresa economica lanciato dall’Unione Europea per fronteggiare le conseguenze della pandemia di COVID-19. Il PNRR persegue una serie di obiettivi ambiziosi, raggruppati in sei “missioni” principali, tra le quali figura la missione *Digitalizzazione, Innovazione, Competitività, Cultura e Turismo* che mira a promuovere la trasformazione digitale della Pubblica Amministrazione, del sistema produttivo e dei servizi ai cittadini, con investimenti significativi in infrastrutture digitali e innovazione. Per quanto concerne il settore culturale, questa missione include misure per la

In secondo luogo, il MiC ha incentivato l'instaurazione di partnership con enti di ricerca, università e società tecnologiche private, sia a livello nazionale che internazionale, per lo sviluppo di soluzioni innovative e la condivisione di *expertise* e risorse. Uno degli accordi di collaborazione più noti e controversi degli ultimi anni è quello che ha coinvolto, da una parte, la Biblioteca Nazionale Centrale di Firenze (BNCF) e la Biblioteca Nazionale Centrale di Roma (BNCR) e, dall'altra, il colosso dell'hi-tech Google, per la digitalizzazione dei fondi librari (in particolare, opere fuori diritto d'autore) delle suddette biblioteche e la loro pubblicazione sulla piattaforma *Google Books*. Lanciata nei primi anni 2000, *Google Books* si proponeva di digitalizzare milioni di libri appartenenti a biblioteche di tutto il mondo, rendendoli ricercabili e in parte consultabili online. Per le biblioteche, la proposta era molto allettante: digitalizzare grandi volumi di opere con costi ridotti o nulli, aumentando la visibilità del proprio patrimonio – specialmente quella di opere rare o difficilmente consultabili – rendendolo accessibile a un pubblico globale. L'adesione delle istituzioni italiane a *Google Books* è stata una vicenda complessa e a lungo dibattuta, che ha visto alternarsi momenti di entusiasmo e forte scetticismo. Alla promessa di rendere disponibili vastissimi *corpora* di testi storico-letterari, si contrapponevano diverse criticità riguardanti: la gestione dei diritti sulle opere digitalizzate (specialmente quelle ancora sotto *copyright*); la proprietà e l'uso dei file digitali creati, con il timore di una “privatizzazione” del patrimonio pubblico; la qualità della digitalizzazione di massa e la difficoltà di integrare i dati di *Google Books* con altre piattaforme nazionali o europee.

Tuttavia, nel tempo, l'attenzione del MiC si è spostata verso progetti di digitalizzazione autonomi e lo sviluppo di piattaforme digitali centralizzate, tra cui spiccano: *Internet Culturale*, in cui convergono numerosi progetti di digitalizzazione intrapresi dagli istituti di conservazione italiani, sia statali che privati;¹⁰² e *Manus Online*, dedicato alla catalogazione e alla digitalizzazione dei manoscritti medievali e moderni conservati nelle biblioteche italiane.¹⁰³ Queste piattaforme sono confluite, a loro volta, all'interno della biblioteca digitale europea

digitalizzazione dell'intero patrimonio, incluso quello bibliografico e archivistico. Per approfondire l'argomento, cfr. il seguente link: <<https://temi.camera.it/leg19/pnrr.html>> (ultimo accesso 17 ottobre 2025).

¹⁰² *Internet Culturale* è un portale online gestito dall'ICCU, che aggrega risorse digitali (quali, manoscritti, edizioni antiche, periodici storici, mappe, materiale grafico, fondi musicali, ecc.) provenienti da migliaia di biblioteche, archivi e altre istituzioni culturali italiane. Il portale nasce con l'intento precipuo di offrire agli utenti un unico punto di accesso al patrimonio culturale nazionale digitalizzato, promuovendo l'interoperabilità e la visibilità delle collezioni digitali. È possibile accedere a *Internet Culturale*, tramite il seguente link: <<https://www.internetculturale.it/>> (ultimo accesso 17 ottobre 2025).

¹⁰³ È possibile accedere alla piattaforma *Manus Online* (anche essa gestita dall'ICCU), tramite il seguente link: <<https://manus.iccu.sbn.it/>> (ultimo accesso 17 ottobre 2025).

Europeana, un'iniziativa della Commissione Europea che aggrega milioni di oggetti digitali (immagini, testi, suoni, video) provenienti da migliaia di istituzioni culturali di tutta Europa (biblioteche, archivi, musei, gallerie). L'obiettivo sotteso alla creazione di *Europeana* è fornire un unico punto di accesso al patrimonio culturale europeo e promuovere la sua valorizzazione. L'Italia rappresenta uno dei principali *contributor* di *Europeana*. La partecipazione italiana, promossa e coordinata dal MiC (attraverso la Digital Library, l'Istituto Centrale per il Catalogo Unico – ICCU per le biblioteche e la Direzione Generale Archivi), si sostanzia nella condivisione di contenuti; nella collaborazione a progetti tematici (ad esempio, sulla Prima Guerra Mondiale, sulle migrazioni, ecc.); e nella definizione e adozione di standard comuni per i metadati (ad esempio, l'*Europeana Data Model* – EDM) e per la qualità della digitalizzazione, garantendo così che i dati siano ben descritti e interoperabili.¹⁰⁴

In terzo luogo, il MiC ha elaborato e promosso l'adozione di linee guida e standard tecnici, sia per la riproduzione digitale (formati e risoluzioni) sia per la descrizione, l'indicizzazione e la distribuzione delle immagini digitali del patrimonio culturale testuale (metadati). Questa iniziativa, che si allinea ai principi FAIR (*Findable, Accessible, Interoperable, Reusable*) fortemente caldeggiati dalla comunità scientifica internazionale, si prefigge lo scopo di garantire omogeneità e qualità nella digitalizzazione e, al contempo, di favorire l'interoperabilità tra diverse banche dati e sistemi informativi (cfr. Cap. III).¹⁰⁵ A tal proposito, è bene ricordare che

¹⁰⁴ È possibile accedere a *Europeana*, tramite il seguente link: <<https://www.europeana.eu/it>> (ultimo accesso 17 ottobre 2025). Per maggiori informazioni, sul framework EDM, è possibile consultare la seguente pagina web: <<https://pro.europeana.eu/page/metadata>> (ultimo accesso 17 ottobre 2025).

¹⁰⁵ I principi FAIR sono stati formalizzati nel 2014, in occasione del workshop *Jointly Designing a Data Fairport*, tenutosi al Lorentz Center nei Paesi Bassi; ma la loro diffusione su larga scala si deve, però, ad un articolo pubblicato nel 2016 su *Scientific Data*, intitolato *The FAIR Guiding Principles for scientific data management and stewardship*, firmato da Mark D. Wilkinson e da un nutrito gruppo di coautori. È importante sottolineare che i principi FAIR non sono uno standard tecnico, ma piuttosto un *framework* concettuale, elaborato da un'equipe internazionale di ricercatori provenienti da diversi ambiti disciplinari, con l'intento di definire un approccio condiviso per la gestione e il riutilizzo dei dati di ricerca nell'era dell'*open science*. Secondo questo quadro concettuale, i dati devono essere:

- *findable*, ovvero facilmente trovabili da chiunque ne abbia bisogno, il che implica che i dati debbano essere pubblicati e registrati in archivi persistenti e ricercabili, con identificatori unici e globalmente risolvibili (come i DOI); inoltre, è fondamentale che i metadati siano ricchi e dettagliati per garantire una ricerca efficace;
- *accessible*, cioè accessibili tramite protocolli standard e aperti (possono esserci restrizioni etiche o legali all'accesso, purché in questi casi il processo per ottenere l'accesso sia chiaro e trasparente);
- *interoperable*, nel senso che i dati e i metadati devono poter essere facilmente integrati e combinati con altri dati; ciò richiede l'uso di formati e linguaggi standardizzati, vocabolari controllati e ontologie che permettano ai sistemi informatici di comprendere e interpretare i dati in modo coerente;
- *reusable*, al fine di massimizzare l'utilizzo dei dati stessi. Ciò presuppone che i dati debbano essere ben descritti (con metadati ricchi e precisi), avere una chiara licenza d'uso che ne specifichi i termini di

nel campo della gestione e dell'accesso ai dati digitali, soprattutto per quanto concerne il patrimonio culturale, i due standard più importanti sono il FITS e il IIIF. Questi standard operano a livelli differenti nella catena di produzione e distribuzione delle risorse digitali.

Il FITS (*Flexible Image Transport System*) è un formato di file a standard aperto usato principalmente per l'archiviazione, la trasmissione e l'elaborazione di dati scientifici, in particolare immagini. Sebbene sia nato e sia più diffuso in ambito astronomico (dov'è il formato standard per lo scambio di dati tra osservatori), il FITS trova sempre più applicazione anche in altri settori che richiedono la conservazione a lungo termine e l'integrità di dati complessi e ad alta risoluzione, come appunto la digitalizzazione del patrimonio culturale. In Italia, la norma UNI 11845 2022 (*Processi di gestione della conservazione a lungo termine di immagini digitali con l'uso del formato FITS*) ha confermato l'importanza e l'affidabilità di questo formato per la conservazione digitale del patrimonio culturale.¹⁰⁶

Il IIIF (*International Image Interoperability Framework*) è un insieme di standard API (*Application Programming Interface*) che definiscono come le immagini digitali e i media associati (come audio e video) possono essere richiesti, distribuiti e visualizzati sul web in modo standardizzato e interoperabile. È bene sottolineare che il IIIF non è un formato di file per la

utilizzo, ed essere conformi agli standard di dominio pertinenti; inoltre, devono essere sufficientemente documentati per permettere a futuri ricercatori (o macchine) di replicare, validare o estendere i risultati.

In conclusione, i principi FAIR promuovono una cultura di gestione dei dati che valorizza l'apertura, la trasparenza e la riproducibilità della ricerca, facilitando la scoperta di nuove conoscenze e il progresso scientifico.

¹⁰⁶ I file FITS si sono affermati come soluzione privilegiata per la conservazione digitale a lungo termine grazie ad una serie di caratteristiche chiave che si trovano elencate di seguito.

- Standard aperto: essendo uno standard aperto e ben documentato, il formato FITS assicura che i dati siano accessibili nel tempo; infatti, può essere letto da un'ampia gamma di software, indipendentemente dall'hardware o dal software specifici usati per la sua creazione.
- Integrità dei dati: i file FITS sono concepiti per archiviare dati "raw" (grezzi) e non compressi (o con compressione *lossless*), evitando la perdita di informazioni che potrebbe verificarsi con formati compressi come JPEG. Ciò li rende ideali per la conservazione di immagini master di altissima qualità.
- Incorporazione di metadati ricchi: ogni file FITS è strutturato in "Header + Data Units" (HDU). L'Header contiene metadati leggibili dall'uomo in formato ASCII, organizzati in coppie chiave/valore. Questi metadati descrivono le caratteristiche dell'immagine (come dimensioni e calibrazioni) e, in particolare per il patrimonio culturale, dettagli sull'opera originale, il processo di digitalizzazione, la provenienza e i diritti. Questa ricchezza di metadati è essenziale per la descrizione e la futura ricercabilità dei beni digitalizzati.
- Struttura flessibile: possono contenere non solo immagini 2D, ma anche array multidimensionali (ad esempio, cubi di dati) o tabelle, rendendoli versatili per diversi tipi di dati scientifici.
- Progettati per la conservazione a lungo termine: la filosofia "once FITS, always FITS" (una volta FITS, sempre FITS) sottolinea l'impegno a mantenere la compatibilità retroattiva con le versioni precedenti, assicurando che i file creati oggi possano essere letti e interpretati anche in futuro, un aspetto vitale per gli archivi digitali.

Per approfondire l'argomento, cfr. il seguente link: <https://fits.gsfc.nasa.gov/fits_standard.html> (ultimo accesso 17 ottobre 2025).

conservazione (come il FITS sopra menzionato), né uno strumento per la digitalizzazione stessa; piuttosto, è un “linguaggio comune” che permette a piattaforme e applicazioni diverse di dialogare fra loro per accedere alle stesse risorse digitali e riutilizzarle. Prima della diffusione di IIIF, le istituzioni culturali che digitalizzavano i loro patrimoni tendevano a implementare soluzioni software e formati di visualizzazione proprietari. Questa frammentazione complicava notevolmente il lavoro degli studiosi, impedendogli di: visualizzare in modo efficiente immagini ad alta risoluzione senza sovraccaricare i server o richiedere download massivi; confrontare immagini provenienti da collezioni digitali diverse; riutilizzare risorse digitali in contesti differenti da quelli offerti dall’istituzione originaria (ad esempio, per integrarle all’interno di edizioni scientifiche digitali oppure per creare annotazioni su immagini afferenti a collezioni diverse).

I vantaggi dello standard IIIF, per le istituzioni culturali che lo adottano, sono i seguenti:

- l’interoperabilità fa sì che le istituzioni possano condividere e visualizzare le proprie collezioni digitali in modo standardizzato, superando le barriere tecnologiche rappresentate dai formati di archiviazione interni;
- la riusabilità fa sì che i patrimoni culturali siano facilmente aggregabili, consultabili e riutilizzabili all’interno di nuove applicazioni o ambienti di ricerca virtuali;
- il miglioramento della *user experience* attiene alla possibilità per l’utente di usufruire di funzionalità avanzate (quali, lo zoom, la rotazione, l’annotazione collaborativa e la visualizzazione *side-by-side* di immagini provenienti da diverse collezioni) nei visualizzatori IIIF (come *OpenSeadragon*, *Mirador*, *Universal Viewer*, ecc.);
- la riduzione dei costi di gestione è una diretta conseguenza del fatto che le istituzioni non devono creare e mantenere molteplici versioni delle stesse immagini per scopi diversi; infatti, il server IIIF le genera "on demand", il che riduce il carico di dati gravante sui server, ottimizza l’uso della banda e i costi connessi;
- la conservazione digitale a lungo termine deriva dal fatto che essendo IIIF uno standard aperto e gestito da una comunità internazionale, garantisce la longevità dell’accesso ai dati digitali (ad esempio, evitando la dipendenza da soluzioni proprietarie);

- la collaborazione tra istituzioni a livello nazionale e internazionale per progetti congiunti che coinvolgono il patrimonio digitale.¹⁰⁷

Dunque, una volta che le istituzioni culturali hanno digitalizzato e archiviato le proprie collezioni in formato FITS, possono renderle fruibili tramite server conformi a IIIF; gli utenti possono usare un visualizzatore IIIF, sul proprio browser, per accedere alle immagini – ingrandendole, confrontandole con documenti di altre istituzioni e annotandole – il tutto senza il bisogno di scaricare i pesanti file FITS originali. A titolo esemplificativo, consideriamo un prezioso manoscritto come quello che raccoglie i frammenti autografi dell’*Orlando Furioso*, conservato presso la Biblioteca Comunale Ariostea di Ferrara. La Biblioteca ha digitalizzato e archiviato il manoscritto in formato JPEG. Supponendo che la Biblioteca adotti gli standard sopra menzionati, il FITS sarebbe il “caveau blindato” in cui i documenti originali sarebbero conservati, protetti e catalogati nei minimi dettagli, per garantirne la sopravvivenza nei secoli; invece, il IIIF sarebbe il “museo virtuale” che consentirebbe agli utenti di visualizzare su un monitor riproduzioni fedeli di quei documenti, ingrandirli per apprezzarne meglio i dettagli, confrontarli con altri documenti, provenienti anche da altre biblioteche (ad esempio, quelli

¹⁰⁷ Lo standard IIIF si compone di diverse API che operano in sinergia. Di seguito, si trovano elencate le API principali.

- La *IIIF Image API* (API Immagine) costituisce la vera e propria spina dorsale di IIIF, poichè consente a qualsiasi visualizzatore o applicazione compatibile con IIIF di richiedere a un server conforme a IIIF un’immagine ad alta definizione o una specifica versione di essa. Ciò include la possibilità di: richiedere porzioni specifiche di un’immagine, invece di doverla scaricare integralmente; ridimensionare l’immagine (miniatura o vista a piena risoluzione); ruotare l’immagine con specifiche angolazioni; specificare la qualità dell’immagine (ad esempio, bianco e nero, scala di grigi, colore), il formato (JPEG, PNG, GIF) e la sua compressione; identificare l’immagine mediante il cosiddetto “Image URI” (*Uniform Resource Identifier*), ovvero un URL strutturato che incapsula tutte queste opzioni di manipolazione, permettendo a qualsiasi visualizzatore conforme a IIIF di interpretare e visualizzare l’immagine come desiderato. Pertanto, grazie all’*Image API*, le istituzioni culturali non sono più obbligate a creare e archiviare molteplici versioni (miniatura, media e alta risoluzione) di una stessa immagine per diverse finalità, poichè il server IIIF le genera “on demand”.
- La *IIIF Presentation API* (API Presentazione) descrive la struttura e i metadati di un oggetto digitale complesso (come un libro, un manoscritto, una mappa o un’opera d’arte multidimensionale) per la sua presentazione sul web. A differenza della *Image API*, che gestisce la richiesta e la consegna dei pixel dell’immagine, la *Presentation API* si concentra sulla logica organizzativa e contestuale delle risorse digitali. Il fulcro della *Presentation API* è il *Manifest*, un documento in formato JSON (*JavaScript Object Notation*), che funge da “indice” o “scheda descrittiva” completa dell’oggetto digitale e contiene tutte le informazioni necessarie a un visualizzatore IIIF per presentare l’opera in modo coerente e significativo.
- La *IIIF Authentication API* (API Autenticazione) consente alle istituzioni di controllare l’accesso a contenuti specifici (ad esempio, immagini protette da copyright o a risoluzione molto alta) attraverso meccanismi di autenticazione standardizzati.
- La *IIIF Search API* (API Ricerca) consente di effettuare ricerche testuali all’interno di un’opera (ad esempio il testo OCR di un manoscritto) e di evidenziare i risultati direttamente sull’immagine.

Per approfondire l’argomento, cfr. i seguenti link: <<https://iiif.io/>>; <<https://www.culturaitalia.it/lo-standard-iiif/>> (ultimo accesso 17 ottobre 2025).

conservati presso l'Ambrosiana di Milano e la Nazionale di Napoli), senza mai toccare gli originali nel *caveau*. Se questo scenario si concretizzasse, le implicazioni per il nostro progetto di ricerca sarebbero considerevoli. Potremmo integrare le immagini digitali direttamente nel sito web dell'edizione digitale, eliminando la necessità di acquisire e gestire i file stessi, che verrebbero invece caricati da remoto in modo semplice e trasparente. Di conseguenza, la produzione e la manutenzione delle immagini ricadrebbero direttamente sulle biblioteche custodi dei documenti originali, riducendo così in modo significativo i costi complessivi dell'edizione. In conclusione, mentre il FITS garantisce la conservazione a lungo termine e l'integrità del dato sorgente, il IIF incentiva la distribuzione e l'accessibilità web di quel dato, facilitando l'interoperabilità tra diverse collezioni e l'esperienza utente avanzata. Molte biblioteche nazionali, archivi di stato e musei hanno già reso o stanno rendendo disponibili le proprie collezioni attraverso IIF (ad esempio, la BAV, la Biblioteca Ambrosiana e la Biblioteca Estense Universitaria).

Infine, il MiC ha progressivamente incoraggiato l'adozione di politiche di *open access* ai contenuti digitalizzati, in modo da favorire la fruizione pubblica e lo sviluppo di servizi innovativi basati sul patrimonio culturale. A tal proposito, si ricorda che la *Legge 4 agosto 2017, n. 124*, (nota anche come “Legge annuale per il mercato e la concorrenza”) ha introdotto importanti innovazioni nel panorama normativo italiano, con ricadute significative anche in materia di digitalizzazione, in particolare per quanto concerne la liberalizzazione della riproduzione e diffusione dei beni culturali. Infatti, l'articolo 1, comma 171 della suddetta *Legge* ha modificato l'articolo 108 del *Codice dei beni culturali e del paesaggio* (D. Lgs. 42/2004), introducendo novità sostanziali (cfr. Cap. III). In precedenza, il *Codice* prevedeva restrizioni significative sulla riproduzione dei beni culturali, soprattutto per i beni archivistici e librari, e limitava fortemente la diffusione delle riproduzioni, anche per finalità non direttamente lucrative (ad esempio, il cosiddetto “lucro indiretto” era spesso interpretato in modo molto restrittivo). Inoltre, vi era incertezza sulla possibilità di divulgare liberamente riproduzioni online. La *Legge 124/2017* introduce le seguenti innovazioni:¹⁰⁸

- libera riproducibilità dei beni culturali per scopi non lucrativi (ovvero, studio e ricerca, libera manifestazione del pensiero o espressione creativa, promozione della conoscenza del patrimonio culturale);

¹⁰⁸ Il testo integrale della *Legge 124/2017* può essere consultato al seguente link: <https://www.gazzettaufficiale.it/eli/id/2017/08/14/17G00140/s> (ultimo accesso 17 ottobre 2025).

- la libertà di riproduzione viene espressamente estesa anche ai beni librari e archivistici, superando le precedenti restrizioni che li escludevano o li limitavano fortemente;
- libera divulgazione, con qualsiasi mezzo, delle immagini di beni culturali, legittimamente acquisite, a condizione che ciò avvenga in modo da non poter essere ulteriormente riprodotte a scopo di lucro, il che favorisce la condivisione sui social media e su piattaforme personali per scopi privati e non commerciali;
- chiarimento del concetto di “non lucro”: la semplice pubblicazione online o l’uso per fini culturali e didattici (anche se su piattaforme che generano pubblicità, purché il contenuto culturale non sia direttamente monetizzato) non rientra nell’ambito del “lucro”.

In linea di principio, queste innovazioni dovrebbero avere un impatto positivo sulla digitalizzazione, poiché:

- incentivano la produzione di contenuti digitali da parte delle istituzioni culturali che dovrebbero sentirsi rassicurate dal fatto che le riproduzioni potranno essere liberamente utilizzate per scopi culturali e di studio;
- promuovono una maggiore circolazione del patrimonio digitale, rendendolo più accessibile a cittadini, ricercatori e operatori culturali, stimolando la creatività e la conoscenza;
- allineano l’Italia a standard internazionali, in quanto questa apertura si avvicina a filosofie di *open access* e *public domain enhancement* adottate in altri paesi e promosse da iniziative come *Europeana* e *Creative Commons* (anche se con alcune differenze significative sul riutilizzo commerciale).

Quanto detto evidenzia la sussistenza delle precondizioni necessarie – istituzionali, finanziarie, legislative e tecnologiche – affinché il Paese possa avviarsi verso una digitalizzazione profonda e capillare; nel prossimo futuro, sarà compito degli istituti di conservazione cogliere questa opportunità recependo le misure predisposte dal MiC, di modo che la promessa della digitalizzazione non rimanga un mero discorso retorico, ma diventi una realtà tangibile e trasformativa, in grado di valorizzare il nostro immenso patrimonio culturale testuale e renderlo accessibile a tutti.

A chiusura di questo paragrafo, è necessario formulare alcune brevi riflessioni critiche riguardo ai metodi di digitalizzazione e alla presentazione delle immagini sui siti istituzionali delle biblioteche e all'interno di molte edizioni scientifiche digitali. Sebbene sia innegabile che la possibilità di accedere a riproduzioni facsimilari di documenti antichi abbia rivoluzionato e migliorato considerevolmente il lavoro del filologo, non possiamo limitarci ad assumere una visione acritica o meramente celebrativa. Questa pratica, infatti, ha introdotto problematiche specifiche che non possono essere trascurate. In particolare, non possiamo sottovalutare quanto le fotografie dei manoscritti tendano a presentare le informazioni in modo selettivo e, a volte, distorto. Quando osserviamo un facsimile digitale, le nostre percezioni rispetto all'oggetto riprodotto risultano alterate; ad esempio, non abbiamo alcuna idea della dimensione reale del manoscritto né della posizione che la pagina riprodotta occupa al suo interno; inoltre, qualsiasi informazione relativa alla struttura codicologica è praticamente assente; spesso, le pagine vengono presentate in modo scorporato, senza la loro controparte destra o sinistra, con la conseguente perdita di informazioni fondamentali sulla *mise en page*, sull'iconografia e, più in generale, sulla materialità del documento.

La scelta di privilegiare la pagina come unità di rappresentazione nella digitalizzazione di documenti analogici, anziché la doppia pagina, deriva da una duplice motivazione. Da un lato, è una conseguenza tecnica diretta del processo di digitalizzazione: fotografare una pagina alla volta minimizza ombre e distorsioni causate dalla rilegatura. Dall'altro, questa scelta valorizza implicitamente la prospettiva di chi è interessato principalmente alla lettura del testo, piuttosto che quella di chi intende studiare il documento nella sua integrità materiale e strutturale. Di conseguenza – parafrasando Elena Pierazzo e Tiziana Mancinelli – la pagina si è affermata come un modello (*model-of*) per la digitalizzazione dei testi antichi e, al tempo stesso, come il modello di presentazione privilegiato per i prodotti digitali, influenzando ogni ricerca condotta mediante queste riproduzioni; in sostanza, la pagina è diventata un modello per la creazione di nuova conoscenza (*model-for*). Tuttavia, questo stato di cose non giustifica la posizione dominante della pagina come modello unico per i libri digitalizzati, a scapito di altre unità strutturali altrettanto valide e significative, come fogli, bifolii o fascicoli.¹⁰⁹

¹⁰⁹ Cfr. T. Mancinelli e E. Pierazzo, *Che cos'è un'edizione scientifica digitale*, Roma, Carocci Editore, 2020, pp. 55-56.

2.2.3 Digitalizzazione del contenuto testuale dei documenti in MRF

Da quanto detto nel paragrafo precedente, si evince che i facsimili digitali degli autografi ariosteschi costituiscono una componente essenziale e ineliminabile dell'edizione digitale; inoltre, l'acquisizione dei facsimili digitali rappresenta la preconditione necessaria per effettuare la trascrizione del loro contenuto testuale in un formato di testo leggibile dalle macchine (*machine-readable form* – MRF). La trascrizione del testo è un'operazione preliminare fondamentale per la produzione di qualsiasi edizione, sia essa cartacea o digitale. Tuttavia, mentre nella prassi editoriale tradizionale, la trascrizione è spesso considerata un mero punto di partenza per ulteriori analisi filologiche (ad esempio, per la collazione dei testimoni di un'opera letteraria e per la ricostruzione del testo critico), nella prassi editoriale digitale – specialmente quella basata sull'uso dello standard di codifica XML-TEI – essa assume un ruolo centrale e di primo piano. Questo cambiamento di status deriva dalle potenzialità intrinseche del processo di codifica che consente non solo di trascrivere fedelmente i documenti originali, ma anche di normalizzare l'ortografia e la punteggiatura secondo le convenzioni moderne; in questo modo, il sistema linguistico originale non è compromesso e contemporaneamente la leggibilità complessiva del testo risulta agevolata. Conseguentemente, la trascrizione evolve da semplice attività propedeutica e “di servizio” a vero e proprio prodotto editoriale; infatti, la trascrizione può diventare direttamente un'edizione diplomatica o interpretativa, pronta per essere distribuita e fruita come prodotto finale autonomo.

La trascrizione del testo (diplomatica o interpretativa), giustapposta al facsimile digitale del documento originale da cui deriva, concorre a determinare un modello editoriale molto diffuso nel web: si tratta dell'edizione documentaria digitale (trad. dall'inglese *Digital Documentary Edition* – DDE) che Elena Pierazzo ha definito come “la registrazione del maggior numero possibile di caratteristiche del documento originale considerate significative dagli editori, visualizzate in tutti i modi che gli editori ritengono utili per i lettori, inclusi tutti gli strumenti necessari per raggiungere tale scopo.”¹¹⁰ Questa sintetica definizione mette in risalto alcuni aspetti fondamentali che meritano di essere approfonditi. In primo luogo, quando parla della “registrazione del maggior numero possibile di caratteristiche del documento originale considerate significative dagli editori”, la Pierazzo richiama l'attenzione sul livello di dettaglio a cui l'editore deve attenersi nella riproduzione della fonte documentaria; infatti, durante la trascrizione di un documento, una delle decisioni più rilevanti per l'editore è determinare quanto

¹¹⁰ E. Pierazzo, *A Rationale of Digital Documentary Editions*, «Literary and Linguistic Computing», 26 (4), 2011, pp. 463-477, p. 475.

fedelmente riprodurre la fisionomia dell'originale. In passato, questa decisione era fortemente condizionata dai limiti tecnici della tipografia. L'avvento del digitale ha notevolmente ampliato le possibilità "mimetiche" della trascrizione, grazie a due fattori chiave: l'introduzione di set di caratteri molto ampi che includono persino abbreviazioni e lettere ormai in disuso (come quelli descritti dal MUFI – *Medieval Unicode Font Initiative*);¹¹¹ e la capacità dei software di videoscrittura e del web di rappresentare con precisione i minimi dettagli e la *mise en page* originale del documento. Di conseguenza, oggi gli studiosi si trovano nella condizione di dover definire nuovi criteri per stabilire fino a che punto spingersi nella riproduzione della fonte documentaria. Paradossalmente, l'ampia disponibilità di immagini digitali, anziché sollevare lo studioso dalla necessità di una riproduzione fedele (dato che l'immagine è spesso parte integrante dell'edizione digitale), ha spinto verso una fedeltà ancora maggiore al documento originale.¹¹² In epoca pre-digitale, l'edizione diplomatica (a cui il modello delle DDE è strettamente legato) era considerata un surrogato del documento originale ed era usata per riprodurre testi preziosi e unici in modo da renderli accessibili a studiosi che non potevano consultare l'originale. Oggi, grazie all'ampia disponibilità di immagini digitali, questa funzione "vicaria" è venuta meno. Pertanto, le DDE assumono un ruolo nuovo e ben più importante: costituiscono un'opportunità per esplorare, conoscere e studiare un documento da molteplici prospettive (testuale, paleografica, codicologica, culturale, iconografica, ecc.). Parafrasando la Pierazzo, possiamo affermare che, in un certo senso, l'edizione documentaria sta al documento fonte come una carta geografica sta al territorio: ne costituisce un modello selettivo, ma diventa anche uno strumento estremamente utile per l'utente che desidera esplorare un territorio sconosciuto.¹¹³

¹¹¹ Per maggiori informazioni sulla MUFI –Medieval Unicode Font Initiative, è possibile consultare la seguente pagina web: <<https://mufi.info/q.php?p=mufi>> (ultimo accesso 17 ottobre 2025).

¹¹² Elena Pierazzo afferma che: «lunghi dall'essere inglobate dalle edizioni basate su immagini, le DDE le hanno di fatto incorporate, perché il formato più comune per le edizioni documentarie digitali è il layout affiancato, in cui l'edizione diplomatica (o ultra-diplomatica) è giustapposta al suo facsimile digitale. In questo formato l'editore si impegna in una sorta di competizione con il facsimile dove, come notato da Kathryn Sutherland, 'l'editore è continuamente sotto esame, aperto al resoconto e alla correzione'. La relazione dialettica tra l'edizione diplomatica e la rappresentazione in facsimile, pur richiedendo un rigore editoriale estremo, coinvolge gli utenti in un'attenta ispezione delle trascrizioni/traduzioni messe in atto dall'editore in una sorta di competizione immaginaria». Cfr. E. Pierazzo, *Digital Documentary Editions and the Others*, «Scholarly Editing: the Annual of the Association for Documentary Editing», 35, 2014, p. 4, consultabile al seguente link: <<https://scholarlyediting.org/2014/essays/essay.pierazzo.html>> (ultimo accesso 17 ottobre 2025).

¹¹³ Secondo Elena Pierazzo, «l'edizione diplomatica affiancata al facsimile fornisce al lettore una versione semplificata, mediata e interpretata di 'ciò che è sulla pagina'; pertanto, l'edizione diplomatica, anziché essere resa ridondante dalla presenza dell'immagine, rappresenta una sorta di mappa e chiave per la comprensione e la navigazione di quell'immagine. Proprio come una mappa di una data area della Terra fornisce all'utente un modo per navigare in un luogo sconosciuto, l'edizione diplomatica fornisce al lettore gli strumenti per decifrare i codici linguistici e bibliografici che informano il documento. Pertanto, mentre la produzione di facsimili digitali di alta qualità ha effettivamente reso superflua la creazione di edizioni diplomatiche come sostituto delle edizioni

In secondo luogo, quando afferma che le caratteristiche dei documenti possono essere “visualizzate in tutti i modi che gli editori ritengono utili per i lettori”, la Pierazzo allude evidentemente alla flessibilità delle DDE, in cui le trascrizioni possono assumere “on demand” la forma di edizioni diplomatiche, ultra-diplomatiche, semi-diplomatiche o interpretative. Tutto ciò è reso possibile dalla codifica digitale del testo e dalla pubblicazione online per mezzo di software che consentono di modificare la fisionomia e la presentazione del testo a seconda delle specifiche esigenze di ricerca.¹¹⁴ Sulla scorta di questa peculiarità, la Pierazzo sostiene che le DDE si qualificano come “edizioni paradigmatiche”, poiché sono in grado di incamerare nel file sorgente una pluralità di opzioni di visualizzazione alternative per la stessa stringa di testo in modo non lineare; successivamente, sarà possibile attivare l’una o l’altra opzione di visualizzazione, grazie alla configurazione dell’interfaccia utente del software di pubblicazione. Viceversa, le edizioni che visualizzano il testo in un solo modo alla volta (ad esempio, quelle cartacee), possono essere definite “edizioni sintagmatiche”.¹¹⁵ Infine, facendo riferimento a “tutti

facsimilari – dal momento che le edizioni facsimilari ora possono essere prodotte facilmente – ha anche rivelato che le edizioni diplomatiche (e in particolare le DDE) sono strumenti accademici e non surrogati, la cui creazione è utile perché sono strumenti interpretativi indispensabili».

Cfr. E. Pierazzo, *Digital Documentary Editions and the Others*, «Scholarly Editing: the Annual of the Association for Documentary Editing», 35, 2014, p. 4, consultabile al seguente link: <<https://scholarlyediting.org/2014/essays/essay.pierazzo.html>> (ultimo accesso 17 ottobre 2025).

¹¹⁴ Elena Pierazzo ha definito la DDE come «un’edizione di un testo, basata su un singolo documento, che cerca di riprodurre le peculiarità del documento stesso, anche se ciò può causare un’interruzione nel normale flusso del testo presentato dal documento. Essa può assumere diversi formati, presentando il contenuto testuale del documento come edizione semi-diplomatica, diplomatica, ultra-diplomatica, o persino facsimilare, che si differenziano per il livello di intervento editoriale, variando dalla massima alla minima concessione alle abitudini di lettura del pubblico di riferimento. L’edizione documentaria si distingue dalla semplice trascrizione nel senso che un’edizione documentaria è destinata alla pubblicazione, mentre una trascrizione rappresenta la prima fase privata del lavoro editoriale. Una trascrizione può diventare un’edizione documentaria nel momento in cui il suo creatore (l’editore) decide che è sufficientemente accurata e segue le convenzioni accademiche abbastanza bene da essere pubblicata».

Cfr. E. Pierazzo, *Digital Documentary Editions and the Others*, «Scholarly Editing: the Annual of the Association for Documentary Editing», 35, 2014, pp. 4-5, consultabile al seguente link: <<https://scholarlyediting.org/2014/essays/essay.pierazzo.html>> (ultimo accesso 17 ottobre 2025).

¹¹⁵ A proposito delle DDE, Elena Pierazzo afferma che «forse queste edizioni potrebbero essere chiamate “edizioni paradigmatiche”, poiché incorporano molteplici opzioni alternative per la stessa sequenza di testo in modo non lineare. Questo si contrappone alle edizioni che possono visualizzare il testo in un solo formato (come, tra le altre, le edizioni a stampa), che potrebbero invece essere definite “edizioni sintagmatiche”. La codifica del testo ha infatti permesso agli editori di avere la botte piena e la moglie ubriaca: caratteristiche che un tempo venivano normalizzate senza indugio per produrre edizioni di lettura o critiche (o sintagmatiche), ora possono essere mantenute e semplicemente attivate o disattivate a piacimento per soddisfare pubblici diversi, aprendo così la strada a nuove ricerche e letture. Infatti, è possibile memorizzare quante informazioni o punti di vista alternativi si desiderano per ogni sequenza di testo, seguendo l’asse paradigmatico, in modo non lineare, ed elaborati diversamente a seconda della visione editoriale. Queste funzionalità hanno attratto l’attenzione degli studiosi per i quali i documenti e le fonti documentarie rivestono un ruolo centrale, in particolare quelli impegnati nella *New o Material Philology* e nella *Critique Génétique*. Gli studiosi che si riconoscono in questi orientamenti editoriali sono stati infatti tra i primi ad abbracciare il mezzo digitale. Poiché le edizioni paradigmatiche sono in grado di immagazzinare una quantità indefinita di informazioni, è possibile rimanere fedeli al documento (mantenendo, ad esempio, grafie antiche e

gli strumenti necessari”, la Pierazzo sottolinea che il concetto di “edizione documentaria digitale” abbraccia sia il contenuto testuale del documento fonte, sia l’infrastruttura digitale sottostante (sia essa visibile o meno all’utente finale), indispensabile per la pubblicazione e l’analisi del contenuto testuale stesso.

L’edizione digitale dei frammenti autografi di Ludovico Ariosto rientra a tutti gli effetti nella categoria delle DDE. D’altra parte, è importante puntualizzare che questa classificazione non ci impedisce di definire l’edizione anche come “genetica” e “critico-interpretativa”. Questo perché, se da un lato l’edizione raccoglie il dossier genetico dell’*Orlando furioso* (il che ne fa una antesignana dell’edizione genetica di stampo francese, intesa come l’edizione integrale di tutto avantesto), dall’altro lato, essa si propone anche di formulare ipotesi sulla corretta successione cronologica delle varianti depositate dall’autore sulle pagine del manoscritto, seguendo la lezione della filologia d’autore italiana.¹¹⁶

La trascrizione del contenuto delle immagini digitali in un formato di testo leggibile dalle macchine può avvenire in due modi: mediante la digitazione manuale o automaticamente. Per la trascrizione dei facsimili digitali degli autografi ariosteschi, abbiamo deciso di ricorrere alla seconda modalità. Le ragioni e le implicazioni pratiche di questa scelta metodologica saranno discusse in dettaglio più avanti (cfr. Cap. III). Per il momento, ci limitiamo a fare il punto sullo stato dell’arte dell’ATR (*Automatic Text Recognition*) e a illustrare le caratteristiche distintive delle due principali implementazioni tecniche di questo metodo, ovvero l’OCR (*Optical Character Recognition*) e l’HTR (*Handwritten Text Recognition*). Recentemente, il riconoscimento automatico del testo ha destato grande interesse all’interno della comunità scientifica internazionale, trovando ampia applicazione in numerosi progetti digitali volti a rendere il patrimonio bibliografico e archivistico non solo pienamente accessibile, ma anche

insolite, abbreviazioni, errori di copista, ecc.), e allo stesso tempo fare concessioni al lettore, che ha così l’opportunità di scegliere quale versione leggere».

Cfr. E. Pierazzo, *Digital Documentary Editions and the Others*, «Scholarly Editing: the Annual of the Association for Documentary Editing», 35, 2014, pp. 4-5, consultabile al seguente link: <<https://scholarlyediting.org/2014/essays/essay.pierazzo.html>> (ultimo accesso 17 ottobre 2025).

¹¹⁶ Sulla complessità del panorama editoriale digitale odierno, si sono espresse anche Tiziana Mancinelli e Elena Pierazzo, le quali osservano che «l’edizione documentaria digitale è il modello predominante del mondo digitale; o almeno questo è quanto sostiene Peter Robinson in un articolo del 2013. In realtà un’indagine a tappeto sulle edizioni digitali effettivamente esistenti ci presenta una situazione alquanto variegata e meno chiara di quanto Robinson sembra suggerire. [...] Questo breve excursus ci porta a valutare da una parte la complessità delle edizioni offerte e l’impossibilità di categorizzare chiaramente i prodotti digitali che si presentano quindi come edizioni critiche, genetiche, documentarie allo stesso tempo, dall’altra ribadisce la centralità della trascrizione e dell’edizione documentaria digitale, sia essa indipendente o parte di un’edizione critica (che comunque è tutt’altro che assente)». Cfr. T. Mancinelli e E. Pierazzo, *Che cos’è un’edizione scientifica digitale*, Roma, Carocci Editore, 2020, pp. 57-58.

facilmente ricercabile e processabile.¹¹⁷ Questo accresciuto interesse è strettamente correlato all'evoluzione dei sistemi di HTR che rappresentano attualmente una delle aree di ricerca più promettenti nell'ambito delle *digital humanities*.¹¹⁸

I progressi della tecnologia di HTR sono stati fortemente stimolati dalle campagne di digitalizzazione di massa intraprese dalle principali biblioteche e archivi a livello mondiale. Nel paragrafo precedente, abbiamo trattato diffusamente l'argomento in relazione al contesto italiano, osservando come il fattore trainante di queste campagne sia stata la diffusione di tecnologie di *imaging* digitale che coniugano costi ridotti e prestazioni elevate, unitamente alla disponibilità di sistemi online per l'*hosting* delle immagini digitali. Ciò nondimeno, bisogna constatare che le iniziali aspettative ottimistiche sulle campagne di digitalizzazione sono state presto ridimensionate dalla difficoltà di fruire e analizzare il contenuto di gran parte dei documenti digitalizzati, a causa della mancanza di trascrizioni *machine-readable* accurate.¹¹⁹

Finora, la trascrizione manuale delle immagini digitali, pur essendo tecnicamente possibile, è stata un'opzione praticabile solo per una minoranza di istituzioni che disponevano delle risorse economiche necessarie per sostenere i costi; infatti, si tratta di un'attività estremamente onerosa in termini di tempo e forza lavoro, sia che si ricorra all'impiego di professionisti, sia che si opti per il coinvolgimento di volontari in iniziative di *crowdsourcing*. D'altro canto, le tecniche di OCR disponibili, adottate da molte biblioteche e archivi digitali (ad esempio, *Google Books* o *Internet Archive*), si sono dimostrate efficaci principalmente per la trascrizione di documenti moderni, stampati e redatti in lingue diffuse, come l'inglese; ma faticano con i manoscritti e i libri a stampa antichi. Un problema che permane malgrado i progressi conseguiti nel campo dell'OCR, grazie all'introduzione del *machine learning* che consente di decifrare anche i testi manoscritti purché i caratteri siano ben definiti e uniformi. Infatti, le tecniche di OCR più

¹¹⁷ Per approfondire le *best practice* adottate dalle biblioteche per l'archiviazione, metadattazione e condivisione dei dati generati mediante sistemi di HTR, cfr. Melissa Terras, *The Role of the Library When Computers Can Read: Critically Adopting Handwritten Text Recognition (HTR) Technologies to Support Research*, in *The Rise of AI: Implications and Applications of Artificial Intelligence in Academic Libraries*, a cura di Amanda Wheatley e Sandy Hervieux, ACRL - Association of College & Research Libraries, Atlanta, pp. 137-148.

¹¹⁸ Per approfondire l'argomento, cfr. AA.VV., *Handwriting: recognition, development and analysis*, a cura di Byron Leite Dantas Bezerra et al., New York, Nova Science Publishers, 2017.

¹¹⁹ A questo proposito Nikolina Milioni afferma: «biblioteche digitali e archivi sono portali essenziali per avere accesso ad un ricco patrimonio informativo. Questi enti svolgono un lavoro imponente di digitalizzazione per ampliare le loro collezioni digitali e mettere a disposizione degli utenti dati testuali preziosi. Tuttavia, i documenti manoscritti sono spesso disponibili solo come immagini, senza le necessarie trascrizioni. Questa mancanza di testo leggibile automaticamente ostacola i ricercatori moderni, limitandone la capacità di condurre studi avanzati, specialmente nell'ambito delle *Digital Humanities* con approcci come il *distant reading* e il *data mining*».

Per approfondire l'argomento, cfr. N. Miloni, *Automatic Transcription of Historical Documents: Transkribus as a Tool for Libraries, Archives and Scholars* (Master's Degree Project), Department of ALM, Uppsala Universitet, 2020, pp. 7-8.

avanzate combinano algoritmi di *computer vision*, che funzionano meglio con caratteri nitidi e quando vi è un contrasto elevato tra sfondo e caratteri, e algoritmi di NLP (*Natural Language Processing*), che richiedono lingue standardizzate e la disponibilità di vocabolari informatizzati per quelle stesse lingue.¹²⁰ Di conseguenza, l'OCR risulta quasi del tutto inadeguato alla maggior parte dei documenti di interesse per gli studiosi di discipline umanistiche, in particolare, filologi e storici che si trovano a lavorare prevalentemente con testi difficilmente leggibili, danneggiati, scritti in lingue ibride o con ortografie non standardizzate.

Tuttavia, nell'ultimo turno di anni, l'impressionante incremento della potenza di calcolo dei computer e il conseguente avvento di sistemi di *machine learning* avanzati che simulano l'attività cognitiva umana, come le reti neurali profonde (*Deep Neural Network* – DNN), hanno contribuito allo sviluppo di tecnologie di HTR altamente performanti.¹²¹ I nuovi e sofisticati sistemi di HTR offrono risultati di trascrizione decisamente più accurati rispetto al passato, sia per i testi manoscritti che per quelli a stampa; ragion per cui, essi hanno rapidamente soppiantato i più noti, ma ormai obsoleti, sistemi di OCR. Questa transizione tecnologica ha, in qualche misura, alimentato l'erronea convinzione che l'HTR fosse la naturale evoluzione dell'OCR. Pertanto, è fondamentale chiarire, in questa sede, che si tratta di due tecnologie distinte, anche se effettivamente, fino ad un certo punto, le rispettive storie si intrecciano.

¹²⁰ Per approfondire l'argomento, cfr. R. Cordell e D. Smith, *Report: A Research Agenda for Historical and Multilingual Optical Character Recognition*, Northeastern University Library, 2018, consultabile al seguente link: <<http://hdl.handle.net/2047/D20297452>> (ultimo accesso 17 ottobre 2025); Id., *'Q i-jtb the Raven': Taking Dirty OCR Seriously*, «Book History», 20, 2017, pp. 188-225, consultabile al seguente link: <<https://scispace.com/pdf/q-i-jtb-the-raven-taking-dirty-ocr-seriously-4joz3t45nt.pdf>> (ultimo accesso 17 ottobre 2025); AA.VV., *Knowledge-Based Intelligent Techniques in Character Recognition*, cura di Lakhmi Jain e Beatrice Lazzerini, Boca Raton, CRC Press, 1999.

¹²¹ Le reti neurali artificiali (*Artificial Neural Networks* – ANN) sono sistemi computazionali avanzati, modellati sulla struttura e il funzionamento del cervello umano, per elaborare informazioni e apprendere da insiemi di dati complessi. Le ANN sono costituite da unità di elaborazione dette “neuroni artificiali” o “nodi”, collegate tra loro tramite connessioni (analoghe alle sinapsi biologiche), che trasmettono segnali. A ciascuna connessione è associato un peso, ovvero un valore numerico che ne determina l'influenza sull'attivazione del nodo successivo. I nodi sono organizzati in una struttura gerarchica suddivisa in tre sezioni principali: uno strato di input, uno o più strati intermedi (detti anche “strati nascosti”) e uno strato di output. Gli strati intermedi svolgono la maggior parte dell'elaborazione del segnale.

Una variante particolarmente potente delle reti neurali è rappresentata dalle reti neurali profonde (*Deep Neural Networks* – DNN). Le DNN funzionano secondo gli stessi principi delle ANN tradizionali, ma si distinguono per la presenza di un numero significativamente maggiore di strati nascosti. La profondità consente alle DNN di apprendere rappresentazioni di dati molto più complesse: ogni strato successivo è in grado di riconoscere *pattern* sempre più astratti e sofisticati, basandosi sulle informazioni elaborate dagli strati precedenti. Questo processo di apprendimento gerarchico fa sì che le DNN siano particolarmente efficaci nello svolgimento di compiti complessi come il riconoscimento di immagini, la traduzione automatica e l'elaborazione del linguaggio naturale (*Natural Language Processing* – NLP).

Per approfondire l'argomento, cfr. G. Roncaglia, *L'architetto e l'oracolo. Forme digitali del sapere da Wikipedia a ChatGPT*, Bari, Laterza, 2023; F. Ciotti, *Minerva e Il Pappagallo. IA generativa e modelli linguistici nel laboratorio dell'umanista digitale*, «Testo E Senso», 26, 2023, pp. 289-315, consultabile al seguente link: <<https://doi.org/10.58015/2036-2293/671>> (ultimo accesso 17 ottobre 2025).

Originariamente, l'acronimo HTR è stato adottato nel campo dell'informatica per indicare strumenti capaci di trascrivere il contenuto di manoscritti digitalizzati in un formato di testo elettronico; di conseguenza, questi strumenti erano anche definiti "OCR per manoscritti", un'espressione ancora invalsa che però risulta imprecisa e fuorviante poichè se, da una parte, accomuna due sistemi di riconoscimento distinti, dall'altra, ne circoscrive l'applicabilità a campi predefiniti. La principale differenza fra i due sistemi di ATR risiede nelle modalità di interpretazione dei segni grafici: l'OCR funziona riconoscendo le realizzazioni grafiche dei singoli caratteri (come suggerisce la locuzione "Optical Character Recognition") e in seguito sfrutta il confronto con un vocabolario informatizzato per migliorare la precisione della trascrizione e minimizzare gli errori di lettura che possono essere causati da difetti di stampa o di scansione; al contrario, l'HTR individua schemi ricorrenti di segni grafici all'interno di una riga di testo (come si intuisce dalla locuzione "Handwritten Text Recognition"), perfezionando la predizione dei risultati di trascrizione tramite l'analisi del contesto e un addestramento preliminare fornito dall'utente.¹²² Pertanto, oggi si tende a credere che l'OCR sia ideale per l'interpretazione di testi a stampa in cui i caratteri sono ben definiti e separati, mentre l'HTR sia adatto esclusivamente ai manoscritti in cui lettere e parole possono essere collegate in modo irregolare a causa delle peculiarità della grafia di ciascun autore o per via delle circostanze di scrittura.¹²³ Ma a ben vedere, questa distinzione risulta troppo rigida e semplicistica, in quanto i sistemi di HTR sono usati con successo anche per la trascrizione automatica di testi a stampa

¹²² I moderni sistemi di HTR si basano su tecniche di *machine learning*, quali le reti neurali profonde (DNN), per analizzare le caratteristiche visive di un testo e identificare caratteri e parole all'interno di una riga di testo precedentemente segmentata. Questo processo avviene tramite un "calcolo di probabilità sovrapposte" che può essere schematizzato nel modo seguente:

1. il sistema analizza una piccola porzione di immagine di testo per estrarre caratteristiche visive (*feature extraction*), quali curve, linee, angoli, ecc;
2. basandosi su queste caratteristiche visive, il sistema genera diverse ipotesi su quali caratteri o parole potrebbero essere presenti;
3. per ciascuna ipotesi, il sistema calcola una probabilità che essa sia corretta, tenendo conto sia delle caratteristiche visive sia del contesto (ad esempio, la probabilità che una certa lettera segua un'altra in una specifica lingua);
4. infine, il sistema confronta le probabilità di tutte le ipotesi sovrapposte e sceglie quella con la probabilità più alta come la trascrizione più probabile per quel segmento di testo.

Dunque, il concetto di "probabilità sovrapposte" si riferisce al fatto che il sistema non fa una scelta binaria (è "A" o non è "A"), ma valuta un intervallo di probabilità per diverse interpretazioni possibili che si "sovrappongono" sulla stessa porzione di immagine.

¹²³ Per approfondire, cfr. P. Stokes e B. Kiessling, *Sharing Data for Handwritten Text Recognition (HTR)*, «Digital Humanities in Practice», 2024, consultabile al seguente link: <<https://hal.science/hal-04444641v1/document>> (ultimo accesso 17 ottobre 2025).

antichi (ad esempio, libri in caratteri gotici) o di difficile interpretazione (ad esempio, i giornali tedeschi dei primi del Novecento in *Fraktur*), raggiungendo un livello di affidabilità impensabile per i coevi sistemi OCR.

Inoltre, sebbene le tecniche sottese al funzionamento dell'OCR e dell'HTR presentino spesso delle somiglianze, l'evoluzione dell'HTR ha seguito un percorso parzialmente indipendente. A cavallo tra la metà degli anni '60 e i primi anni 1970, mentre i sistemi hardware OCR di seconda generazione si andavano affermando, la trascrizione automatica dei testi manoscritti cominciava a muovere i passi, soprattutto grazie alle intuizioni pionieristiche del filosofo dell'informazione Kenneth M. Sayre (1928-2022).¹²⁴ Quest'ultimo mise in evidenza le difficoltà insite nell'interpretazione di scritture continue. La questione affrontata da Sayre è nota come “paradosso di Sayre”: per decifrare una parola manoscritta, è necessario scomporla nei caratteri che la compongono; tuttavia, per essere segmentata, la parola deve essere preliminarmente interpretata. Per uscire da questa *impasse*, Sayre suggerì di regolarizzare i caratteri che compongono le parole e di adottare tecniche di *template matching* (confronto di modelli), al fine di semplificare la segmentazione e il successivo riconoscimento dei caratteri. Nello specifico, l'approccio proposto da Sayre consisteva nel suddividere una linea di testo continua in frammenti discreti (operando una sorta di “segmentazione implicita”) e di cercare una corrispondenza probabilistica tra questi frammenti e un set di parole regolarizzato.¹²⁵

Tuttavia, le teorie di Seyre ebbero scarsa risonanza scientifica, rimanendo di fatto inapplicate. Dunque, per circa una decina d'anni, non vi furono progressi tecnologici significativi nel campo dell'ATR. Frattanto, la ricerca scientifica si orientò verso lo sviluppo di dispositivi hardware per il riconoscimento dei gesti e della grafia, con l'ambizioso obiettivo di rimpiazzare le periferiche di input tradizionali (tastiera e mouse) nell'interazione uomo-macchina. Quest'area di ricerca

¹²⁴ I sistemi OCR di seconda generazione rappresentarono un significativo passo avanti rispetto ai loro predecessori che dipendevano da font specifici. Questi sistemi operavano tramite la tecnica del *template matching* che consisteva nel confrontare i caratteri scansionati con un catalogo di modelli predefiniti: dopo un'iniziale fase di *preprocessing* in cui i caratteri venivano isolati e il rumore visivo dall'immagine ridotto, il carattere segmentato veniva confrontato con una libreria di immagini di riferimento per trovare la corrispondenza più vicina, mediante metodi di correlazione o similarità statistica. Alcuni sistemi più avanzati introdussero anche la tecnica della *feature extraction* (estrazione delle caratteristiche) che consentiva di riconoscere le caratteristiche strutturali dei caratteri (come linee e incroci), così da ridurre la dipendenza da font specifici. L'hardware di questi sistemi era spesso specializzato e dedicato, con circuiti appositi per accelerare il processo di *matching* e riconoscimento. Ne sono un esempio l'IBM 1287, il primo sistema misto analogico-digitale, e l'IBM 1418, entrambi capaci di supportare *template* multipli e la lettura automatica su formati standardizzati (ad esempio, l'OCR-A e l'OCR-B).

In conclusione, sebbene più flessibili della generazione precedente, questi sistemi erano ancora vincolati a font e layout altamente strutturati, il che ne limitava l'applicabilità in presenza di layout complessi o di variazioni nella scrittura. Dunque, a differenza dei sistemi di ATR attuali, mancavano di capacità di apprendimento automatico e adattamento dinamico ai font e, in generale, alle altre caratteristiche dei documenti.

¹²⁵ Per approfondire, cfr. A. Vinciarelli, *Offline Cursive Handwriting: From Word To Text Recognition*, IDIAP Research Record, 2003, consultabile al seguente link: <<https://infoscience.epfl.ch/record/82879>> (ultimo accesso 17 ottobre 2025).

(nota come “pen computing”), che sfrutta penne digitali e tavolette grafiche per elaborare segnali analogici, prese il via nel 1983, quando l’azienda Pencept realizzò prototipi innovativi come il PenPad 200 e il PenPad 320 (quest’ultimo già compatibile con le applicazioni MS-DOS dei primi personal computer IBM).

I primi software di HTR videro la luce fra la fine degli anni ‘80 e l’inizio degli anni ‘90 e si diffusero prevalentemente in ambito postale e commerciale per identificare brevi scritture manoscritte (quali, codici postali, cifre su assegni bancari, ecc.) in cui l’interpretazione del testo era facilitata dalla disposizione chiara e ben separata dei caratteri. Questi software continuavano ad essere basati su un approccio *segmentation-based*: l’interpretazione del testo avveniva isolando i singoli caratteri e riconoscendoli per sovrapposizione con un set predefinito di segni grafici, tramite tecniche di *feature extraction* e *template matching*. L’approccio *segmentation-based* presentava però un grosso limite legato all’adozione di modelli euristici e algoritmi *ad hoc* che compromettevano la scalabilità dei software di HTR. Infatti, essendo dipendenti dai dati originali di addestramento, questi risultavano meno robusti e performanti nell’elaborare nuovi materiali.

All’inizio del nuovo millennio, i progressi conseguiti nel campo del *machine learning* hanno segnato un superamento del vecchio approccio *segmentation-based*, rivoluzionando profondamente le tecnologie di ATR. I ricercatori hanno iniziato a sfruttare la capacità di memoria predittiva delle reti neurali ricorrenti (RNN), per migliorare l’affidabilità dei software di ATR.¹²⁶ In particolare, i sistemi di HTR hanno beneficiato notevolmente di questa innovazione, dato che la loro efficacia risiede nella capacità di trascrivere accuratamente testi sconosciuti, ovvero testi che non sono mai stati processati prima dalla macchina. A questo scopo, si rende preliminarmente necessario un addestramento manuale del sistema di HTR che si basa su dati forniti dall’utente, presi come modello di riferimento ideale.¹²⁷ La preparazione di

¹²⁶ Una Rete Neurale Ricorrente (RNN) è un tipo di rete neurale artificiale, progettata per elaborare sequenze di dati, come testo, serie temporali o audio. La caratteristica distintiva delle RNN è la capacità di mantenere una “memoria” degli input precedenti nella sequenza, grazie a connessioni che formano cicli. In altre parole, l’output di un nodo in un dato momento dipende non solo dall’input corrente, ma anche dagli input precedenti nella sequenza. Ciò consente alle RNN di apprendere dipendenze e schemi all’interno di sequenze, rendendole ideali per compiti come la traduzione automatica, il riconoscimento vocale, la generazione di testo e l’analisi di serie temporali.

¹²⁷ L’apprendimento delle reti neurali artificiali si basa su due processi computazionali opposti e complementari: la *forward propagation* e la *backpropagation*.

La *forward propagation* (propagazione in avanti) è il processo mediante il quale una rete neurale genera un’inferenza o una previsione, basata sui dati che ha “visto” e imparato a processare. Questo processo, caratterizzato da un flusso unidirezionale di informazioni che si propaga dallo strato di input verso lo strato di output, può essere schematizzato come segue:

1. i dati di input (rappresentati, ad esempio, da vettori numerici corrispondenti a immagini o sequenze di testo) sono immessi nel primo strato della rete neurale;

questi dati di addestramento, noti come “ground truth” (cioè “verità di base”), rappresenta il fulcro dell’intero processo di addestramento, poichè determina direttamente l’efficienza e l’adattabilità della tecnologia HTR a contesti diversi. Ne consegue che la trascrizione manuale dei documenti deve rispettare criteri che assicurino un certo grado di robustezza del sistema (*consistency*), il quale può essere occasionalmente sottoposto a operazioni di messa a punto (*fine tuning*). Analogamente a quanto avviene per i *Large Language Model* (LLM), che sotto questo aspetto assomigliano molto ai sistemi di HTR, l’addestramento della macchina è tanto più efficiente quanto maggiori sono la quantità e la qualità dei materiali processati. Ciò nondimeno, è fondamentale sottolineare che, per evitare l’iper-specializzazione del sistema, a volte è consigliabile preferire una maggiore variabilità interna dei dati piuttosto che una quantità massiccia di dati dello stesso tipo.¹²⁸

A margine di questa breve analisi, è opportuno fare una considerazione che si ricollega direttamente a quanto detto sopra in merito al concetto di modellizzazione. Nel caso dei LLM, come *ChatGPT*, la creazione della *ground truth* è un’operazione onerosa e laboriosa, che richiede la revisione di miliardi di parole; nel caso dell’ATR, invece, i dati immessi dall’utente

2. ogni neurone in uno strato riceve input dai neuroni dello strato precedente; questi input sono moltiplicati per i rispettivi “pesi sinaptici” (cioè valori numerici che la rete ha “imparato” durante l’addestramento e che quantificano la forza delle connessioni) e sommati a un “bias”;
3. il risultato di questa somma ponderata è successivamente trasformato da una “funzione di attivazione” (cioè una funzione matematica) che introduce la “non-linearità” necessaria affinché la rete sia in grado di apprendere relazioni complesse nei dati.
4. l’output di ciascun neurone diviene l’input per i neuroni dello strato successivo; questo schema di propagazione si ripete per tutti gli strati nascosti, fino a raggiungere lo strato di output finale.
5. lo strato di output finale produce la previsione della rete, che potrebbe consistere, ad esempio, nella probabilità che una sequenza di caratteri corrisponda alla parola “casa” o “cosa”.

In pratica, la *forward propagation* rappresenta il modo in cui la rete “pensa” e produce una risposta basandosi sulla sua attuale conoscenza (i pesi e i bias). La *backpropagation* (retropropagazione), invece, è il processo che consente alla rete neurale di imparare dai propri errori e aggiustare i suoi pesi e bias. Questo processo si muove dall’output verso l’input, in direzione opposta alla *forward propagation*.

In estrema sintesi, mentre la *forward propagation* calcola l’output, la *backpropagation* calcola come l’errore si distribuisce all’indietro e come i parametri interni della rete (pesi e *bias*) devono essere modificati per migliorare le prestazioni. Questi due passaggi vengono ripetuti milioni di volte durante l’addestramento, permettendo alla rete neurale di apprendere e migliorare continuamente.

¹²⁸ Un *Large Language Model* (LLM) è un modello di intelligenza artificiale addestrato su grandi quantità di testo per comprendere, generare e manipolare il linguaggio naturale. Utilizza architetture neurali avanzate, come i *Transformer*, per prevedere parole, frasi o interi testi, ed è in grado di svolgere compiti complessi come traduzione, sintesi, scrittura automatica e risposte a domande.

Un *Transformer* è un’architettura di rete neurale, introdotta nel 2017 con l’articolo *Attention Is All You Need*, che ha rivoluzionato il campo dell’elaborazione del linguaggio naturale (NLP) e, successivamente, quello della visione artificiale. La sua caratteristica distintiva e innovativa risiede nell’uso intensivo del meccanismo di auto-attenzione (*self-attention*), che gli permette di pesare l’importanza di diverse parti di una sequenza di input quando ne elabora un’altra parte. A differenza delle Reti Neurali Ricorrenti (RNN), che elaborano le sequenze in modo seriale, i *Transformer* possono elaborare tutte le parti della sequenza in parallelo, migliorando notevolmente l’efficienza computazionale e la capacità di cogliere dipendenze a lungo raggio.

possono essere quantitativamente minori, ma ciò non toglie che i costi e le implicazioni ideologiche ed etiche dell'operazione siano comunque rilevanti. In entrambi i casi, infatti, i dati di addestramento (una sorta di *big data* testuale) derivano da una selezione operata dall'utente, che influenza direttamente lo sviluppo della tecnologia (*model-of*) e conseguentemente le sue future applicazioni pratiche (*model-for*). Questa selezione è spesso orientata dalle dinamiche culturali dominanti, che tendono a privilegiare contenuti già centrali per la ricerca e la divulgazione scientifica. Ne deriva il rischio concreto che patrimoni culturali marginali, già poco visibili, siano ulteriormente esclusi, mentre quelli già ampiamente accessibili accrescano la propria centralità, il che può condurre al consolidamento di assetti culturali egemonici. In sintesi, quindi, la creazione della *ground truth* non è un'operazione culturalmente neutra, specialmente quando ad occuparsene sono aziende private che perseguono principalmente finalità economiche.

A tal proposito, bisogna ricordare che inizialmente sono stati proprio i grandi gruppi editoriali e le *tech company* che, attratti dalle straordinarie potenzialità dei moderni sistemi di HTR, li hanno integrati nei loro servizi commerciali.¹²⁹ Ad esempio, Adam Matthew Digital ha introdotto l'HTR nella sua piattaforma *Quartex* per consentire agli utenti di effettuare ricerche avanzate all'interno delle proprie collezioni digitalizzate. Parallelamente, Google ha sviluppato l'applicazione *Fabricius*, che sfrutta l'intelligenza artificiale per la trascrizione automatica dei geroglifici.¹³⁰ Tuttavia, è importante sottolineare che in questi contesti specifici, gli algoritmi utilizzati e il codice sorgente non sono resi pubblici. La mancanza di trasparenza circoscrive l'implementazione e l'accessibilità di tali tecnologie a una ristretta cerchia di utenti con competenze specialistiche. Fortunatamente, in anni recenti si è andata profilando una valida alternativa a questi servizi commerciali: si tratta di piattaforme di HTR, sviluppate all'interno di una più ampia comunità scientifica internazionale, con l'intento di garantire i principi di equità e

¹²⁹ Per approfondire l'argomento, cfr. Adam Matthew Digital, *Artificial Intelligence Transforms Discoverability of Handwritten Manuscripts*, 2020, consultabile al seguente link: <<https://www.amdigital.co.uk/products/handwritten-text-recognition>> (ultimo accesso 17 ottobre 2025); Adam Matthew Digital, *Historic 17th & 18th Century Manuscript Documents Made Fully Searchable Using Artificial Intelligence*, 2017, consultabile al seguente link: <<https://www.amdigital.co.uk/about>> (ultimo accesso 17 ottobre 2025); Google Arts and Culture, *What is Fabricius?*, 2020, consultabile al seguente link: <<https://experiments.withgoogle.com/fabricius>> (ultimo accesso 17 ottobre 2025); Google Cloud, *Detect Handwriting in Images, AI and Machine Learning Products*, 2021, consultabile al seguente link: <<https://cloud.google.com/vision/docs/handwriting>> (ultimo accesso 17 ottobre 2025).

¹³⁰ È possibile accedere alla piattaforma *Quartex* tramite il seguente link: <<https://www.amdigital.co.uk/transform-your-digital-archives/am-quartex>> (ultimo accesso 17 ottobre 2025). L'applicazione *Fabricius* è disponibile al seguente link: <<https://artsandculture.google.com/experiment/fabricius/gwHX41Sm0N7-Dw?hl=en>> (ultimo accesso 17 ottobre 2025).

accessibilità della conoscenza, in linea con le direttive europee sulla digitalizzazione del patrimonio culturale. Ne sono esempi virtuosi le piattaforme *eScriptorium* e *Transkribus* che si configurano come veri e propri *Virtual Research Environment* (VRE), ovvero applicazioni collaborative progettate per supportare la ricerca scientifica di un'utenza differenziata, con molteplici interessi di studio. Questo approccio promuove la condivisione delle conoscenze e facilita l'adozione diffusa delle tecnologie di HTR, democratizzando l'accesso a strumenti di ricerca avanzati.

Nel panorama italiano delle *digital humanities*, due iniziative si contraddistinguono per l'efficace applicazione delle suddette piattaforme di HTR alla valorizzazione del patrimonio bibliografico. Innanzitutto, segnaliamo la biblioteca digitale creata nell'ambito del progetto *HumanitiesConnect*, avviato nel 2020 dalla Bibliotheca Hertziana - Istituto Max Planck per la storia dell'arte, con l'obiettivo di costituire un repertorio digitale di fonti a stampa rare in ambito storico-artistico e scientifico. Il progetto *HumanitiesConnect*, che risponde ai principi FAIR dell'*open science*, si articola su tre direttrici principali. In primo luogo, punta sull'accessibilità del patrimonio storico-letterario, digitalizzando migliaia di volumi preziosi provenienti dalle collezioni della Bibliotheca Hertziana (Roma), del Kunsthistorisches Institut (Firenze) e del Max Planck Institute for the History of Science (Berlino). In secondo luogo, incentiva la ricerca *full-text*, integrando scansioni ad alta risoluzione con trascrizioni testuali al fine di consentire una ricerca completa nei contenuti, piuttosto che limitarsi alla mera consultazione di immagini. Infine, promuove l'interoperabilità e il riuso dei dati, attraverso l'adozione di standard aperti come IIIF e TEI-XML.

Il fulcro della *HumanitiesConnect Digital Library* è rappresentato dalla piattaforma *Read&Search* che affianca le scansioni ad alta risoluzione dei volumi alle loro trascrizioni testuali. Le trascrizioni sono state generate mediante l'applicazione *Transkribus* che sfrutta motori neurali per automatizzare e ottimizzare la trascrizione di documenti storico-letterari, manoscritti o stampati. *Read&Search* non solo rende i testi leggibili e ricercabili, ma supporta anche il perfezionamento continuo delle trascrizioni, l'annotazione dei contenuti e l'accesso ai dati tramite API, facilitando così il *data mining* e l'analisi computazionale per la ricerca umanistica.¹³¹

Ricordiamo poi il gruppo di ricerca *Progetto Mambrino*, istituito nel 2003 presso l'Università di Verona, che ha intrapreso un'importante iniziativa di studio sulla ricezione italiana del romanzo

¹³¹ Per maggiori informazioni sul progetto *Humanities Connect*, cfr. E. Bastianello, *Digitalizzazione, trascrizione, citazione: le fonti testuali per le pubblicazioni digitali*, «DigItalia», 18(2), 2023, pp. 162-172, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00084>> (ultimo accesso 17 ottobre 2025).

cavalleresco spagnolo, concentrandosi in particolare sui romanzi cavallereschi pubblicati a Venezia nel Cinquecento (1546-1568). Questi testi, che ebbero una vasta diffusione nell'Europa dell'Ancien Régime, costituiscono un corpus ampio e ancora poco esplorato. Attualmente non esistono edizioni moderne di tali opere, che sono state per lungo tempo ignorate dalla cultura accademica a causa della loro natura commerciale e della loro lunghezza (alcuni testi contano quasi mille pagine). Le edizioni antiche e gli esemplari sopravvissuti sono rari e difficilmente accessibili, rendendo arduo un confronto sistematico tra di essi.

Per ovviare a queste criticità, il *Progetto Mambrino* si propone di sviluppare una biblioteca digitale che consenta la consultazione interattiva delle suddette opere. La piattaforma include trascrizioni delle opere affiancate dalle immagini delle fonti originali, nonché collegamenti ipertestuali a risorse complementari, tra cui un database bibliografico degli esemplari conservati. Il raggiungimento di questo obiettivo è stato reso possibile grazie al finanziamento del PRIN 2017, con il progetto *Mapping Chivalry: Spanish Romances of Chivalry from Renaissance to XXI Century: a Digital Approach*, nonché grazie al supporto del Progetto di Eccellenza *Le digital humanities applicate alle lingue e letterature straniere* (2018-2022) e del successivo progetto *Inclusive Humanities* (2023-2027). Questi finanziamenti hanno permesso di affidare lo sviluppo della *Progetto Mambrino Digital Library* all'impresa Net7 di Pisa, specializzata nella creazione di biblioteche digitali umanistiche, avvalendosi anche del supporto tecnico del dipartimento universitario.

Una delle principali sfide nell'implementazione del progetto è stata la trascrizione dei testi, data l'enorme estensione del corpus. La trascrizione manuale si è rivelata un processo troppo lungo e gravoso, ostacolando il progresso di altre attività di ricerca. Per accelerare questa fase, quindi, si è fatto ricorso a strumenti di riconoscimento automatico dei caratteri (OCR). Tuttavia, la trascrizione automatica di testi antichi, come le cinquecentine, ha presentato diverse difficoltà: la stampa utilizzava segni grafici particolari (come segni tironiani, legature e abbreviazioni), e i testi erano spesso danneggiati da problemi di conservazione, quali lacune, macchie e rotture. Inoltre, le distorsioni derivanti dalle scansioni manuali, come la trasparenza dell'inchiostro e l'oscillazione nelle inquadrature, complicavano ulteriormente il processo. Per superare questi ostacoli, è stato adottato il software *Transkribus* che ha consentito di ottenere trascrizioni di alta precisione, con un margine d'errore inferiore al 2%, grazie all'addestramento del programma su un set iniziale di circa venti pagine trascritte manualmente. Una volta trascritti e corretti secondo rigorosi criteri editoriali, i testi sono stati codificati in XML-TEI per la visualizzazione nella biblioteca digitale. La struttura del testo è stata accuratamente marcata, con l'indicazione delle partizioni principali, come il titolo del libro, i titoli dei capitoli e le interruzioni di pagina.

Inoltre, per agevolare l'interrogazione dei contenuti, sono stati isolati elementi chiave, quali i nomi dei personaggi, i luoghi, i temi e i motivi narrativi trattati nelle opere. La struttura del front-end della *Progetto Mambrino Digital Library* si articola in tre sezioni principali.

Il *Database bibliografico* è organizzato secondo una gerarchia a tre livelli – *Opere*, *Edizioni* ed *Esemplari* – collegati fra loro da un sistema di filtri di ricerca che include parametri quali lingua, ciclo narrativo, autore, traduttore e altri criteri specifici. Tale struttura consente di:

- individuare un'opera a partire dall'anno di prima edizione e da ulteriori dati editoriali;
- accedere a un elenco delle edizioni note, ordinato cronologicamente, corredato da informazioni sulle biblioteche che custodiscono i diversi esemplari;
- consultare la descrizione di ciascuna copia, corredata da eventuali riproduzioni fotografiche e da collegamenti diretti alle versioni digitali.

Il nucleo centrale della biblioteca è costituito dalle *Edizioni scientifiche digitali*, incentrate soprattutto sulla pubblicazione di edizioni critiche di testi privi di una moderna versione a stampa. Il primo titolo pubblicato in questo formato è lo *Sferamundi di Grecia*, prosecuzione del ciclo di *Amadis*. Le edizioni, concepite come strumenti interattivi, permettono la consultazione simultanea delle immagini digitalizzate e delle trascrizioni testuali. Ogni capitolo è corredato da un riassunto accessibile tramite collegamenti ipertestuali, che facilita l'esplorazione dei personaggi, dei luoghi e dei motivi ricorrenti. L'indice dei personaggi annovera oltre 1.300 voci e la piattaforma integra ulteriori strumenti di consultazione, quali mappe, alberi genealogici e altre risorse di supporto. Infine, i *Database semantici* sono dedicati all'indicizzazione e all'interpretazione di personaggi, luoghi e motivi narrativi, con un'integrazione dinamica ai riassunti dei capitoli.¹³²

In definitiva, l'adozione della tecnologia HTR si profila come una vera e propria svolta per lo studio dei documenti storico-letterari. Questa innovazione non solo amplifica l'accesso a vaste collezioni documentarie, rendendo interrogabili archivi finora inaccessibili, ma apre anche la

¹³² Per maggiori informazioni sul Progetto Mambrino e le iniziative da esso intraprese nel campo delle Digital Humanities, cfr. i seguenti saggi: S. Bazzaco, *La trascrizione automatica di documenti a stampa antichi. Appunti per un modello di riconoscimento della tipografia in corsivo*, «DigItalia», 1, 2024, pp. 63-86, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00094>> (ultimo accesso 17 ottobre 2025); A. Bognolo, *Le Digital Humanities e l'accessibilità del patrimonio letterario, il caso del 'Progetto Mambrino'*, in *Informatica umanistica, digital humanities: verso quale modernità?*, a cura di M. Gatto et al., Bari, Cacucci Editore, 2024, pp. 105-118; T. Mancinelli, *Early printed edition and OCR techniques: what is the state-of-art? Strategies to be developed from the working-progress Mambrino project work*, «Historias Fingidas», 4, 2016, pp. 255-260.

strada a nuove domande di ricerca, più ambiziose e interdisciplinari. Le trascrizioni automatiche ottenute grazie all'HTR permettono infatti di estrarre, visualizzare e analizzare contenuti testuali attraverso strumenti avanzati delle *digital humanities*, superando le barriere tradizionali dell'analisi manuale. Come sottolineano Günter Mühlberger e Melissa Terras, "l'impatto di queste tecnologie sarà dirompente, gettando le basi per analisi computazionali sofisticate su scala semantica, linguistica e geospaziale delle fonti storiche".¹³³ Con il rapido perfezionamento dei modelli HTR, il dibattito si sta spostando dalle potenzialità generali agli aspetti operativi e strategici legati alla loro implementazione nei progetti di digitalizzazione del patrimonio culturale. Tra le sfide emergenti, si evidenziano la minimizzazione dell'intervento umano nella preparazione dei dati di addestramento,¹³⁴ il riuso di modelli esistenti, lo sviluppo di modelli HTR ibridi (i cosiddetti "general model" o "super model") e il *fine-tuning* mirato per migliorare le prestazioni su specifici *corpora*.¹³⁵ In questo scenario, l'adesione ai principi FAIR non è solo una buona pratica, ma una condizione necessaria per garantire la sostenibilità e la scalabilità delle soluzioni adottate. L'HTR, dunque, non è un mero strumento tecnico: è un catalizzatore di trasformazione metodologica nella filologia digitale, nell'archivistica storica e, più in generale, nello studio e nella valorizzazione del patrimonio testuale.

¹³³ Günter Mühlberger et al., *Transforming scholarship in the archives through Handwritten Text Recognition. Transkribus as a case study*, «Journal of Documentation», 75(5), 2019, p. 956, consultabile al seguente link: <<https://doi.org/10.1108/JD-07-2018-0114>> (ultimo accesso 17 ottobre 2025).

¹³⁴ Per approfondire l'argomento, cfr. E. Perdiki, *Preparing Big Manuscript Data for Hierarchical Clustering with Minimal HTR Training*, «Journal of Data Mining and Digital Humanities», 2023, consultabile al seguente link: <<https://hal.science/hal-03880102v4>> (ultimo accesso 17 ottobre 2025).

¹³⁵ Gli attuali motori di HTR consentono lo sviluppo di modelli generalizzati (i cosiddetti "super model"). Tali modelli sono in grado di processare non una singola grafia specifica, bensì una molteplicità di scritture storiche all'interno di un medesimo sistema grafico, estendendosi su un intervallo temporale considerevole (superiore a un secolo). Per approfondire l'argomento, cfr. T. Hodel et al., *General Models for Handwritten Text Recognition: Feasibility and State-of-the Art. German Kurrent as an Example*, «Journal of Open Humanities Data», 7, 2021, consultabile al seguente link: <<https://doi.org/10.5334/johd.46>> (ultimo accesso 17 ottobre 2025); REED-COOP, *Should I use a base model when training a custom model?*, «Transkribus blog», 7 febbraio 2025, consultabile al seguente link: <<https://blog.transkribus.org/en/using-base-model-when-training-a-custom-model>> (ultimo accesso 17 ottobre 2025); REED-COOP, *Introducing Transkribus Super Models - Get Access to 'The Text Titan I'*, «Transkribus blog», 9 luglio 2023, consultabile al seguente link: <<https://blog.transkribus.org/en/introducing-transkribus-super-models-get-access-to-the-text-titan-i>> (ultimo accesso 17 ottobre 2025).

2.2.4 Codifica del testo secondo lo standard XML-TEI

Come anticipato nel paragrafo precedente, la trascrizione del testo presente nelle immagini digitali dei frammenti ariosteschi è un'operazione propedeutica alla successiva fase di codifica del testo. A questo proposito, si ricorda che la *Text Encoding Initiative* (TEI) costituisce oggi uno degli standard più consolidati e autorevoli nell'ambito delle *digital humanities*, configurandosi come un'infrastruttura metodologica e tecnologica di riferimento per la codifica testuale. Avviata nel 1987 in risposta alla crescente informatizzazione della ricerca umanistica, la TEI persegue l'obiettivo di uniformare e razionalizzare le pratiche di rappresentazione digitale dei testi, spesso caratterizzati da strutture complesse e da una stratificazione di significati linguistici, storici, materiali e ideologici. In questo contesto, la TEI si propone sin dall'inizio come una soluzione standardizzata ma flessibile, capace di adattarsi a una vasta tipologia di materiali – dai testi a stampa ai manoscritti, dalle epigrafi ai corpora lessicografici – offrendo un modello descrittivo in grado di restituire in modo formale tanto la struttura quanto il contenuto semantico del testo. Uno degli aspetti distintivi dell'iniziativa risiede nella sua governance partecipativa. Le *TEI Guidelines* non sono emanazione di un'autorità normativa centralizzata, ma il frutto di un processo collaborativo che coinvolge studiosi, bibliotecari, archivisti, informatici e altri attori del mondo culturale. Dal 2000, il coordinamento delle attività è affidato al TEI Consortium, che promuove lo sviluppo dello standard, organizza incontri annuali, gestisce strumenti di comunicazione (come la mailing list *TEI-L*) e favorisce il dialogo tra progetti e comunità scientifiche eterogenee. Questa struttura aperta e inclusiva ha favorito la diffusione della TEI anche tra i non specialisti, grazie a un'ampia documentazione, esempi pratici, tutorial e a una *community* attiva nel fornire supporto. L'architettura teorico-operativa della TEI si articola attorno a cinque principi fondamentali, esemplificati di seguito.

1. Formalizzazione della struttura testuale

Il primo principio riguarda la necessità di descrivere formalmente la struttura logica, semantica e materiale dei testi. La TEI si basa su XML (*eXtensible Markup Language*), un linguaggio gerarchico e leggibile sia dall'uomo sia dalla macchina, che consente la definizione precisa di “elementi” (ad esempio, <p>, <div>, <name>, <date>, <placeName>) e dei relativi “attributi” (ad esempio, @type, @n, @lang), nonché delle regole di annidamento e contesto che ne disciplinano l'uso all'interno del documento. L'adozione di uno schema XML-TEI permette di validare automaticamente i documenti,

assicurandone la coerenza strutturale e riducendo il rischio di ambiguità interpretative.¹³⁶ Strumenti come *Roma*, sviluppato dal TEI Consortium, facilitano la creazione di schemi personalizzati, selezionando esclusivamente i moduli pertinenti a uno specifico progetto.¹³⁷

2. Flessibilità e modularità

Le linee guida TEI comprendono centinaia di elementi codificabili, ma non impongono l'adozione di un set rigido: ciascun progetto può configurare un sottoinsieme adatto alle proprie esigenze scientifiche. Questa modularità consente di mantenere un equilibrio tra standardizzazione e adattabilità, favorendo la personalizzazione senza compromettere la compatibilità tra diversi progetti.¹³⁸

3. Centralità del contenuto semantico

Diversamente dai sistemi WYSIWYG (*What You See Is What You Get*), in cui l'informazione è veicolata principalmente attraverso la formattazione visiva (font, stile, colore), la TEI pone l'accento sulla funzione semantica degli elementi testuali. Ogni unità significativa – che si tratti di un nome proprio, un toponimo, una citazione o un termine tecnico – è marcata con un *tag* esplicito, consentendo di distinguere, ad esempio,

¹³⁶ Un documento XML-TEI può essere considerato valido solo quando rispetta due condizioni fondamentali. La prima riguarda la correttezza formale (*well-formedness*). In pratica, il file deve obbedire alle regole di base dell'XML: ogni tag aperto deve essere anche chiuso, la struttura dei tag deve essere correttamente annidata senza sovrapposizioni improprie, e l'intero documento deve avere un unico elemento radice (ad esempio, `<TEI>...</TEI>`). Anche gli attributi devono essere sempre racchiusi tra virgolette e, se compaiono caratteri speciali come il simbolo `&`, questi vanno codificati in modo corretto (ad esempio, `&`). Questa è la prima verifica da superare, e vale per qualunque file XML, non solo per quelli conformi allo standard TEI.

La seconda condizione riguarda invece la conformità allo schema TEI. Significa che il documento deve rispettare le regole strutturali e semantiche stabilite dalla comunità TEI, descritte in un documento di riferimento come una DTD (*Document Type Definition*), uno schema XML o uno schema RELAX NG. Questi schemi definiscono quali elementi e attributi sono consentiti, dove possono comparire e in che ordine. Per esempio, un elemento `<title>` è previsto all'interno di `<titleStm>`, ma non in un paragrafo `<p>`, a meno che il modello usato non dica diversamente. Per controllare la validità rispetto a queste regole si possono usare diversi strumenti: programmi da riga di comando come *xmllint* (con l'opzione `--valid` o `--schema`), validatori come *Jing* per RELAX NG, oppure editor XML professionali come *Oxygen XML Editor* o *XMLSpy*, che integrano funzioni di validazione automatica.

¹³⁷ L'applicazione *Roma* è disponibile al seguente link: <https://roma.tei-c.org/> (ultimo accesso 17 ottobre 2025).

¹³⁸ I moduli TEI sono insiemi tematici di elementi e attributi XML che consentono di personalizzare la codifica digitale di testi umanistici in base alle specifiche esigenze di un determinato progetto. Tra i moduli più comuni figurano: *tei.core* che fornisce gli elementi fondamentali comuni a tutti i documenti TEI; *tei.header*, dedicato alla codifica delle informazioni descrittive e bibliografiche contenute nel tag `<teiHeader>`; e *tei.textstructure* che definisce la struttura generale del testo come sezioni e paragrafi. Altri moduli importanti sono: *tei.namesdates*, utile per il riconoscimento e la marcatura di nomi di persona, luoghi e date; *tei.verse*, pensato per la struttura della poesia; *tei.msdescription*, specifico per la descrizione codicologica di manoscritti; e *tei.transcription* che offre strumenti per la trascrizione e l'analisi paleografica dei testi.

tra “Londra” come città britannica, città canadese o cognome. Tale granularità semantica abilita analisi sofisticate e trasformazioni automatizzate, garantendo una rappresentazione più ricca e funzionalmente utile del testo.

4. *Interoperabilità e indipendenza tecnologica*

Il formato XML garantisce una neutralità tecnologica: i documenti TEI possono essere letti, trasformati o manipolati con qualsiasi software compatibile, su qualsiasi sistema operativo. A differenza dei formati proprietari (ad esempio, *.docx* o *.pages*), la TEI non vincola a un ambiente di lavoro specifico, assicurando durabilità, portabilità e riusabilità dei dati nel lungo periodo.

5. *Accessibilità e sostenibilità della comunità*

La struttura aperta e collaborativa della TEI favorisce non solo l’evoluzione dello standard, ma anche la sua accessibilità. Il supporto fornito dalla comunità e la ricca documentazione rendono lo standard praticabile anche da chi si avvicina per la prima volta alla codifica testuale, facilitando l’inclusione e la diffusione delle buone pratiche.

A titolo esemplificativo, si prendono in esame le stanze 22 e 23 del *Canto IX* dell’*Orlando Furioso*, conservate nella c. 8 r. del fascicolo II del manoscritto custodito presso la Biblioteca Comunale Ariostea di Ferrara (ms. Cl. I. A., It. 1.). Come si evince dalla riproduzione digitale dell’autografo, si tratta di una copia in bella, probabilmente destinata alla stampa (cfr. Fig. 2.1). Le due stanze sono tratte dall’episodio in cui Olimpia racconta in prima persona, con tono diretto e confidenziale, le circostanze della propria vita familiare e l’inizio della relazione amorosa con Bireno, duca di Selandia. Le due ottave passano con naturalezza dalla descrizione biografica (“figliuola fui del conte d’Olanda, a lui sì grata”) alla dimensione emotiva, che si apre con “la bellezza e l’età ch’in lui fioriva” e culmina nella certezza dell’amore (“ch’amassi et ami me con cor sincero”). La fluidità del passaggio è favorita dagli enjambement (ad esempio, tra “stato avvenne / che ne la nostra terra un duca venne” o tra “captiva / tanto più che per quel ch’apparea fuori”). L’anafora della congiunzione *che* e le ripetizioni (ad esempio, “Credea e credo e creder credo il vero”) scandiscono il ritmo e sottolineano la progressione narrativa, conferendo enfasi al racconto. La paratassi (“Io voglio che sappiate che figliuola... / fui del conte d’Olanda...”) e gli iperbati (“Duca era di Selandia”) contribuiscono a un tono narrativo vivo, che rispecchia la dimensione orale del racconto. Infine, l’alternanza tra lessico affettivo (“grata”, “captiva”, “cor

sincero”) e toponimi storici (“Selandia”, “Biscaglia”, “Mori”) unisce la sfera privata all’orizzonte cavalleresco.¹³⁹

[22]

Io voglio ch sappiate: ch figliuola
fui dl conte d'olanda: a lui si grata:
Quantuq prole io no gli fossi sola:
Ch'era da lui fratelli accompagnata:
Ch'a quato io gli chiedea: da lui parola
Contraria no mi fu mai replicata,
Standomi lieta i questo stato: avene
Che ne la nostra terra un duca venne,

[23]

Duc^a era di Selandia: e se ne giva
verso Biscaglia a guerreggiar co i mori.
La bellezza e l'eta ch'in lui fioriva
E li no piu da me sentiti amori:
Co poca guerra me gli fer captiva:
tanto piu ch p quel ch'apparea fuori
Io Credea: e credo: e creder credo il vero:
ch'amassi et ami me con cor sincero.

Il testo è stato trascritto in forma diplomatica, mantenendo integralmente l’ortografia originale, la punteggiatura e l’uso di maiuscole e minuscole così come presenti nel manoscritto. La

¹³⁹ Nella st. 22, Olimpia si rivolge all’interlocutore con l’incipit “Io voglio che sappiate”, dichiarando apertamente le proprie origini regali: è figlia del re d’Olanda (qui indicato come conte), amata profondamente dal padre (“a lui si grata”), anche se non unica erede (“Quantuque prole io non gli fossi sola”). Il rapporto d’intesa con il genitore è sottolineato dal fatto che questi non le oppone mai parola contraria alle sue richieste (“Contraria non mi fu mai replicata”). L’equilibrio e la serenità di questa vita familiare sono però destinati a cambiare con l’arrivo, “nella nostra terra”, di un duca straniero.

Nella st. 23, il nuovo venuto è presentato come duca di Selandia, in viaggio verso la Biscaglia per intraprendere una campagna contro i Mori. Olimpia, che ne è subito colpita, ne descrive la bellezza e il vigore giovanile (“La bellezza e l’età ch’in lui fioriva”), qualità che si intrecciano alla nascita di un sentimento inedito e profondo (“E li non più da me sentiti amori”). L’attrazione si trasforma rapidamente in amore (“Con poca guerra me gli fer captiva”), alimentata non soltanto dalle doti esteriori, ma dalla persuasione interiore che Bireno l’ami sinceramente e con costanza (“ch’amassi et ami me con cor sincero”). Queste due ottave disegnano così il passaggio di Olimpia da una condizione di sicurezza domestica e filiale a una nuova dimensione, segnata dall’irrompere dell’amore e dal destino legato a un uomo straniero, preludio agli sviluppi drammatici dell’episodio.

codifica è stata realizzata con un metodo semi-automatico, secondo lo standard XML-TEI, applicando criteri volti a rendere fedelmente sia gli aspetti testuali sia quelli materiali.¹⁴⁰

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Orlando Furioso, Canto IX, Stanze 22-23</title>
        <respStmt>
          <resp>Trascrizione e codifica</resp>
          <name>Ambra Torregrossa</name>
        </respStmt>
      </titleStmt>
      <publicationStmt>
        <p>Non pubblicato</p>
      </publicationStmt>
      <sourceDesc>
        <msDesc>
          <msIdentifier>
            <repository>Biblioteca Comunale Ariostea di Ferrara</repository>
            <idno>Cl. I. A., It. 1.</idno>
          </msIdentifier>
          <msPart>
            <msContents>
              <msItem>
                <title>Orlando Furioso</title>
                <locus>carta 8r, fascicolo II, foglio 3</locus>
                <note>Le stanze 22 e 23 narrano l'episodio di Olimpia</note>
              </msItem>
            </msContents>
            <physDesc>
              <objectDesc>
                <supportDesc>
                  <support>Frammento autografo in bella copia di Ludovico Ariosto</support>
                </supportDesc>
              </objectDesc>
            </physDesc>
          </msPart>
        </msDesc>
      </sourceDesc>
      <bibl>
        <author>Ludovico Ariosto</author>
      </bibl>
    </fileDesc>
  </teiHeader>
</TEI>
```

¹⁴⁰ La codifica in XML-TEI è stata realizzata seguendo un approccio misto. Inizialmente il testo è stato elaborato con *Transkribus*, che ha effettuato il riconoscimento automatico del contenuto e ha generato direttamente un file di output in formato TEI, fornendo così una struttura di base già allineata agli standard di riferimento. Successivamente, il file generato è stato sottoposto a un validatore automatico, utilizzato per individuare eventuali errori formali o discrepanze rispetto allo schema TEI adottato. Infine, il documento è stato sottoposto a una revisione manuale approfondita: in questa fase si è proceduto a perfezionare la codifica, correggere imprecisioni residue e inserire *markup* specialistico ove necessario, ad esempio per evidenziare elementi testuali di rilievo, aggiungere annotazioni o segnalare varianti. Grazie a questa combinazione di automazione e intervento umano, il testo ha raggiunto un livello di accuratezza e conformità elevato.


```

<title>Orlando Furioso</title>
  <note>Trascritto e codificato da Ambra Torregrossa a partire dai frammenti autografi conservati presso la
Biblioteca Ariostea di Ferrara.</note>
</bibl>
</sourceDesc>
</fileDesc>
</teiHeader>

```

La struttura generale del documento TEI si apre con l'elemento <TEI>. Questo rappresenta la radice dell'intera codifica e racchiude sia le informazioni descrittive, sia la trascrizione del testo. All'interno di <TEI> si trova il <teiHeader>, una sezione essenziale dedicata ai metadati, cioè a tutte le informazioni bibliografiche e descrittive che permettono di identificare e contestualizzare il frammento. Il <teiHeader> contiene, tra le altre parti, il <fileDesc>, che fornisce la descrizione bibliografica di base del documento. Questa sezione è suddivisa in tre sottocomponenti:

1. <titleStmt> dove è indicato il titolo dell'opera (*Orlando Furioso*, Canto IX, Stanze 22-23) e le responsabilità, cioè chi ha svolto il lavoro di trascrizione e codifica;
2. <publicationStmt> che specifica lo stato editoriale (in questo caso "Non pubblicato")
3. <sourceDesc> che descrive in dettaglio la fonte da cui è tratto il testo.

All'interno di <sourceDesc> compare <msDesc>, che raccoglie i dati specifici sul manoscritto. Qui troviamo:

- <msIdentifier> che riporta la segnatura (<idno>Cl. I. A., It. 1.</idno>) e la collocazione fisica del documento (<repository>Biblioteca Comunale Ariostea di Ferrara</repository>);
- <msPart> che identifica la parte del manoscritto presa in esame; all'interno di <msPart> troviamo <msContents> con un <msItem> che illustra il contenuto testuale del frammento, specificando titolo, posizione nel manoscritto (<locus>carta 8r, fascicolo II, foglio 3</locus>) ed eventuali note (in questo caso, il contenuto delle due stanze in oggetto, ovvero l'episodio di Olimpia). Segue <physDesc> che fornisce la descrizione fisica del supporto, articolata in <objectDesc>, <supportDesc> e <support>. Infine, l'elemento <bibl> offre un riferimento bibliografico sintetico all'opera nella sua interezza.

```

<text>
<body>
<div type="canto" n="9">
  <head>canto nono</head>

  <!-- Numerazioni presenti sulla carta -->
  <pb n="8" type="autograph"/> <!-- numero di pagina autografo in alto a sinistra -->
  <fw type="pageNumber" place="top right" n="8">8</fw> <!-- numero di pagina non autografo in alto a destra -->
  <fw type="folioNumber" place="top right" n="3">3</fw> <!-- numero autografo del foglio (terzo) -->

```

La sezione <text> del documento TEI contiene la trascrizione vera e propria del testo. Al suo interno si trova l'elemento <body>, che racchiude i contenuti principali, e al suo interno ancora compare <div type="canto" n="9">. Questo <div> rappresenta l'unità testuale del Canto IX dell'opera: l'attributo n="9" specifica il numero del canto, mentre la struttura interna ospita la trascrizione diplomatica dei versi manoscritti. Subito all'interno di <div> troviamo <head>, che riporta il titolo della sezione, in questo caso “canto nono”, così come appare nel manoscritto. L'elemento <pb> segnala un cambio di pagina nel supporto originale. Qui l'attributo n="8" indica il numero di pagina, mentre type="autograph" specifica che si tratta di una numerazione apposta dall'autore stesso. Infine, l'elemento <fw> è usato per codificare elementi paratestuali presenti sul foglio, come numeri di pagina o di foglio. Gli attributi type e place precisano rispettivamente la tipologia di numero (ad esempio “pageNumber” o “folioNumber”) e la sua posizione fisica sulla pagina (in questo caso “top right” per indicare la parte alta a destra del foglio). In questo modo, la codifica TEI non si limita a trascrivere il testo poetico, ma documenta anche le caratteristiche materiali del manoscritto che lo conserva.

```

<!-- Stanza 22 -->
<lg type="stanza" n="22">
<l xml:id="I22-1">Io voglio<choice><abbr>ch</abbr><expn>che</expn></choice>sappiate:
<choice><abbr>ch</abbr><expn>che</expn></choice> figliuola</l>
<l xml:id="I22-2">fui<choice><abbr>dl</abbr><expn>del</expn></choice>conte d'<placeName
type="regione">olanda</placeName>: a lui si grata:</l>
<l xml:id="I22-3"><choice><abbr>Quantuq</abbr><expn>Quantunque</expn></choice> prole io
<choice><abbr>no</abbr><expn>non</expn></choice> gli fossi sola:</l>
<l xml:id="I22-4"><choice><abbr>Ch</abbr><expn>Che</expn></choice>'era da lui fratelli accompagnata:</l>
<l xml:id="I22-5"><choice><abbr>Ch</abbr><expn>Che</expn></choice>'a
<choice><abbr>quato</abbr><expn>quanto</expn></choice> io gli chiedea: da lui parola</l>
<l xml:id="I22-6">Contraria<choice><abbr>no</abbr><expn>non</expn></choice> mi fu mai replicata,</l>
<l xml:id="I22-7">Standomi lieta <choice><abbr>i</abbr><expn>in</expn></choice> questo stato:
<choice><abbr>avene</abbr><expn>avvenne</expn></choice></l>
<l xml:id="I22-8">Che ne la nostra terra un duca venne,</l>

```

```

<!-- Enjambement della stanza 22 -->
<linkGrp type="enjambment">
  <link target="#l22-1 #l22-2"/>
  <link target="#l22-3 #l22-4"/>
  <link target="#l22-5 #l22-6"/>
  <link target="#l22-7 #l22-8"/>
</linkGrp>
</lg>

<!-- Stanza 23 -->
<lg type="stanza" n="23">
  <l xml:id="l23-1">Duc<add place="above">a</add>era di<placeName type="regione">Selandia</placeName>: e
se ne giva</l>
  <l xml:id="l23-2">verso<placeName type="regione">Biscaglia</placeName> a
guerreggiar<choice><abbr>co</abbr><expn>con</expn></choice> i mori.</l>
  <l xml:id="l23-3">La bellezza e l'eta
  <choice><abbr>ch</abbr><expn>che</expn></choice>'<choice><abbr>i</abbr><expn>in</expn>
</choice> lui fioriva</l>
  <l xml:id="l23-4">E li<choice><abbr>no</abbr><expn>non</expn></choice> piu da me sentiti amori:</l>
  <l xml:id="l23-5"><choice><abbr>Co</abbr><expn>Con</expn></choice> poca guerra me gli fer captiva:</l>
  <l xml:id="l23-6">tanto piu<choice><abbr>ch</abbr><expn>che</expn></choice>
  <choice><abbr>p</abbr><expn>per</expn></choice> quel
  <choice><abbr>ch</abbr><expn>che</expn></choice>'appaia fuori</l>
  <l xml:id="l23-7"><add place="margin">Io</add>Credea: e credo: e creder credo il vero:</l>
  <l xml:id="l23-8"><choice><abbr>ch</abbr><expn>che</expn></choice>'amassi et ami me con cor
sincero.</l>

<!-- Enjambement della stanza 23 -->
<linkGrp type="enjambment">
  <link target="#l23-1 #l23-2"/>
  <link target="#l23-3 #l23-4"/>
  <link target="#l23-5 #l23-6"/>
  <link target="#l23-7 #l23-8"/>
</linkGrp>
</lg>
</div>
</body>
</text>
</TEI>

```

La codifica TEI riproduce non solo il testo poetico, ma anche la sua struttura metrica e le caratteristiche paleografiche, integrando informazioni utili per l'analisi linguistica e filologica. La struttura poetica è resa attraverso l'elemento `<lg type="stanza" n="...">`, che identifica ciascuna stanza del poema come un'unità metrica autonoma, numerata grazie all'attributo `n`. All'interno di ogni stanza troviamo una sequenza di `<l>`, ciascuno corrispondente a un singolo verso. In questo documento, ogni `<l>` possiede un attributo `xml:id` che gli assegna un

identificatore univoco: questo accorgimento è fondamentale per poter creare riferimenti incrociati, ad esempio nella marcatura degli enjambement.

La marcatura paleografica e linguistica si basa sull'uso di <choice>, un elemento che serve a rappresentare varianti testuali. Nel nostro caso, viene utilizzato per mostrare insieme la forma abbreviata presente nel manoscritto (<abbr>) e la sua espansione (<expan>), come avviene per abbreviazioni quali ch → che o dl → del. L'elemento <add> segnala invece un'aggiunta rispetto al testo di base. L'attributo place ne indica la collocazione materiale: place="margin" viene usato, ad esempio, per la parola Io inserita nel margine sinistro (st. 23, v. 7), mentre place="above" segnala un'aggiunta in interlinea, come la lettera a sovrascritta nella parola *Duca* (st. 23, v. 1).

Per quanto riguarda la strategia di codifica degli enjambement, si utilizza <linkGrp type="enjambment">, un gruppo di collegamenti che raccoglie tutti i casi di enjambement presenti in una stessa stanza. Ogni singolo enjambement è rappresentato da un <link> che collega i due versi coinvolti; l'attributo target del <link> contiene i riferimenti agli xml:id dei versi interessati (per esempio #l22-1 #l22-2). In pratica, ogni verso riceve un identificatore univoco (xml:id="l22-1", xml:id="l22-2", ecc.), e gli enjambement sono resi come relazioni esplicite tra due <l> mediante <link> all'interno del <linkGrp>. Questo metodo permette di mappare con precisione i rapporti metrici senza modificare il testo originale e di sfruttare tali informazioni in fase di analisi o visualizzazione, garantendo così una rappresentazione digitale che preserva sia la fedeltà diplomatica al manoscritto sia la possibilità di elaborazioni avanzate.

In conclusione, l'esempio fornito mostra come la TEI sia molto più di una semplice tecnologia di *markup*; essa rappresenta un linguaggio formale per la conoscenza umanistica, in grado di trasformare il testo da oggetto statico a risorsa dinamica, annotabile, analizzabile e interoperabile. La sua architettura basata su XML, unita alla modularità, alla centralità semantica e all'apertura della comunità, la rende uno strumento imprescindibile per chi intende non solo conservare un testo, ma anche analizzarlo e condividerlo in un ambiente digitale. In questo senso, la TEI si afferma come un ponte tra tradizione filologica e innovazione computazionale.

2.2.5 Progettazione dell'interfaccia utente e pubblicazione dell'edizione digitale

Nei paragrafi precedenti abbiamo visto come il modello logico più adottato per la produzione delle edizioni digitali (ivi compresa l'edizione dei frammenti ariosteschi) sia il cosiddetto “sorgente–output”, concentrandoci sulla componente del *sorgente*, ovvero sulle modalità di preparazione del testo (trascritto e codificato generalmente secondo lo standard XML–TEI) e sui problemi connessi a tali procedure. In questo paragrafo, invece, ci soffermeremo sull'output: sulle modalità della sua produzione, sulle forme della sua presentazione agli utenti e sulle conseguenze, dirette e indirette, che tale modello comporta per lettori, filologi ed editori. Un documento XML può essere facilmente convertito in diversi formati (HTML, PDF, Word, ePub, ecc.), caratteristica che ne ha favorito l'adozione all'interno della filiera editoriale per produrre versioni destinate sia alla stampa sia al web e ai dispositivi mobili. Nelle edizioni scientifiche digitali, tuttavia, la molteplicità dei formati coesiste con la possibilità di proporre il testo secondo diversi approcci scientifici: quello diplomatico, volto a riprodurre l'originale nel modo più fedele possibile, oppure quello interpretativo, che prevede una normalizzazione ortografica moderna. In questo senso i vantaggi offerti da XML appaiono evidenti: mantenere manualmente versioni distinte dello stesso testo non è realistico, poiché ogni correzione dovrebbe essere replicata in tutte le copie, con il rischio di incoerenze e perdita di controllo. L'obiettivo è quindi disporre di un unico testo matrice da cui generare, secondo le necessità del pubblico, differenti rappresentazioni. Il file sorgente funge da punto di controllo unico, mentre gli output vengono prodotti dinamicamente. Questo sistema semplifica la gestione editoriale e garantisce la flessibilità necessaria a soddisfare esigenze diverse di lettori e studiosi. In tale prospettiva, il file sorgente assume la funzione di archivio o base di conoscenza, capace di ospitare testi “potenziali” aggiornabili in forme differenti. L'edizione, intesa come output, si configura dunque come un modello del testo (*model-of*), che seleziona e organizza le caratteristiche del sorgente secondo una prospettiva ben definita. A questo concetto si collega quello di interfaccia utente; infatti, i testi digitali non si presentano soltanto come contenuti, ma anche come sistemi di funzionalità interattive. L'interfaccia costituisce il filtro tra i dati codificati e gestiti dai software e gli utenti che vi accedono: spesso assume la forma di un sito web che, occultando il codice sorgente, offre al lettore una versione leggibile e arricchita da finestre, pulsanti e opzioni che ne modulano la fruizione. Questi strumenti di interazione corrispondono a interrogazioni (*query*) inviate al software di gestione del testo, le quali determinano in tempo reale la configurazione visibile sullo schermo. La dialettica interrogazione–risposta definisce la dinamicità tipica delle edizioni digitali contemporanee,

offrendo all'utente la possibilità di svolgere letture multiple (trascrizione diplomatica, critico-interpretativa, facsimile) e usufruire di strumenti di consultazione (ad esempio, glossari, indici, ecc.). Questa dinamicità trascende le categorie editoriali tradizionali, poiché una stessa edizione può integrare elementi di più tipologie, entro i limiti fissati dall'editore, che agisce da filtro intellettuale.

Ne consegue che l'attività editoriale digitale non si limita alla codifica del testo (*constitutio textus*), ma comprende anche la progettazione delle modalità di fruizione, con importanti implicazioni teoriche e pratiche per la definizione stessa di "edizione scientifica digitale". Poiché l'output rappresenta l'espressione di una precisa tesi editoriale, l'editore non può limitarsi a costituire il testo, ma deve anche sovrintendere alla sua pubblicazione e distribuzione. Ne consegue che tale processo richiede un insieme articolato di competenze, sia filologiche sia informatiche, che difficilmente possono essere riunite in un'unica figura professionale. La progettazione di interfacce dinamiche implica infatti conoscenze tecniche avanzate, generalmente estranee alla formazione tradizionale del filologo. Per questa ragione la maggior parte delle edizioni digitali nasce dal lavoro di équipe interdisciplinari, che integrano competenze umanistiche e informatiche. Diversamente dall'editoria a stampa, in cui l'intervento dell'editore si arresta di norma alla consegna del testo, in ambito digitale il suo ruolo prosegue fino alla pubblicazione – e spesso oltre – sia per la mancanza di modelli standardizzati, sia perché l'interfaccia costituisce parte integrante della rappresentazione editoriale e veicolo privilegiato della visione scientifica sottesa all'edizione. Questa dimensione collaborativa comporta inevitabilmente la necessità di reperire risorse economiche, non sempre disponibili, in particolare per i giovani ricercatori. Un'alternativa consiste nello sviluppo autonomo di interfacce semplici, basate su linguaggi accessibili come l'HTML: soluzioni meno sofisticate sotto il profilo estetico e funzionale, ma sostenibili, longeve ed eticamente apprezzabili, poiché richiedono un investimento economico ridotto e presentano minore vulnerabilità all'obsolescenza tecnologica. Tuttavia, esse rischiano di essere percepite come prodotti poco professionali e privi di validazione accademica, soprattutto in assenza di un *paratesto digitale* in grado di garantire l'autorevolezza, a differenza di quanto accade per il cartaceo, dove la credibilità è assicurata dalla casa editrice e dalla collana. La produzione di edizioni digitali scientifiche è comunque caratterizzata da costi elevati, dovuti alla necessità di personale specializzato, all'aggiornamento dei software, all'archiviazione dei dati e alla manutenzione dei server. Come osservano Elena Pierazzo e Tiziana Mancinelli, tali oneri finiscono per limitare l'accesso soprattutto a dottorandi e giovani studiosi, con conseguenze etiche e formative per il futuro della filologia digitale: senza una massa critica di praticanti, infatti, la disciplina rischia di

non svilupparsi appieno, nonostante le sue enormi potenzialità scientifiche.¹⁴¹ Per ovviare a tali criticità, si rende necessario distinguere la ricerca testuale dall'innovazione tecnologica e promuovere strumenti di pubblicazione a bassa soglia tecnica. In questa direzione si collocano piattaforme basate sulla codifica TEI, come *EVT*, *TAPAS*, *Textual Communities* o *TEI Publisher*, che hanno contribuito a semplificare i processi e a ridurre i costi, pur presentando limiti legati a un minor grado di personalizzazione o alla necessità di competenze informatiche specifiche.¹⁴² Se da un lato tali strumenti generalisti favoriscono la democratizzazione dell'accesso, dall'altro rischiano di uniformare le pratiche, replicando modelli standardizzati simili a quelli delle collane editoriali a stampa, e lasciando così meno spazio alla sperimentazione e all'innovazione. La scelta di una piattaforma di pubblicazione si configura pertanto come una vera e propria decisione editoriale di natura scientifica, poiché ogni strumento incorpora un preciso modello intellettuale: *EVT* privilegia il rapporto tra testo e documento o la dialettica fra testo critico e testimoni; *Textual Communities* pone l'accento sul lavoro collaborativo; *TAPAS* è orientato alla conservazione a lungo termine. Non a caso, molti di questi strumenti sono stati sviluppati da filologi, a conferma del fatto che la rappresentazione testuale costituisce parte integrante dell'attività filologica stessa. Le edizioni digitali più complesse e onerose, spesso basate su software proprietari, assumono una funzione sperimentale, configurandosi come laboratori di innovazione in cui vengono sviluppate e testate soluzioni nuove, talvolta rischiose, ma potenzialmente generalizzabili e utili all'intera comunità scientifica. La dialettica tra queste edizioni "sperimentali" e gli strumenti "generalisti" consente di mantenere un equilibrio tra innovazione e accessibilità, trasformando anche le esperienze più esclusive in un patrimonio condiviso funzionale al progresso collettivo della disciplina. In tale prospettiva si colloca la nota distinzione proposta da Elena Pierazzo, che, ricorrendo alla metafora della moda, individua due modelli di edizione scientifica digitale. Le edizioni *haute couture* sono progetti altamente

¹⁴¹ T. Mancinelli e E. Pierazzo, *Che cos'è un'edizione scientifica digitale*, Roma, Carocci Editore, 2020, pp. 82-87.

¹⁴² *TAPAS* è un progetto finalizzato a offrire servizi di archiviazione, pubblicazione e accesso per dataset codificati in TEI, pensato soprattutto per singoli progetti o realtà con risorse limitate. L'obiettivo è permettere agli utenti di caricare, gestire e pubblicare in modo semplice i propri testi TEI, offrendo funzionalità di ricerca e navigazione. Per approfondire, cfr. la scheda ufficiale del progetto al seguente link: <<https://tapasproject.org>> e il report del progetto *TAPAS: Building a TEI Publishing and Repository Service*, «Journal of the Text Encoding Initiative», 5, 2013.

Textual Communities (accessibile al seguente link: <<https://textualcommunities.com/app/>>) è una piattaforma web progettata per supportare l'editing collaborativo di testi digitali. Consente la gestione condivisa di trascrizioni, versioning e strumenti per l'interazione online tra studiosi. Ad esempio, è stata adottata nel *Canterbury Tales Project* come sistema per gestire trascrizioni e revisione dei testi.

TEI Publisher (accessibile al seguente link: <<https://teipublisher.com/exist/apps/tei-publisher-home/index.html>>) è un'applicazione eXist-db che facilita la pubblicazione di testi codificati in TEI senza richiedere competenze avanzate di programmazione. Permette di caricare file TEI, personalizzare lo schema di visualizzazione tramite ODD e generare automaticamente output multipli (HTML, ePub, PDF, ecc.).

specializzati, concepiti su misura per esigenze specifiche, caratterizzati da originalità, precisione e complessità tecnica, ma anche da costi elevati e ridotta replicabilità. Le edizioni *prêt-à-porter*, al contrario, si fondano su strumenti e framework standardizzati, risultano economicamente più sostenibili e, pur offrendo un grado minore di personalizzazione, garantiscono ampia adattabilità e accessibilità, soprattutto per studiosi e progetti con risorse limitate. Pierazzo sottolinea come entrambe le tipologie siano indispensabili per un ecosistema equilibrato delle edizioni digitali: le edizioni *haute couture* promuovono sperimentazione e avanzamento metodologico, mentre quelle *prêt-à-porter* favoriscono diffusione, sostenibilità e inclusione. La metafora della moda evidenzia così il duplice volto dell'editoria digitale, in cui la tensione tra esclusività e democratizzazione trova un punto di equilibrio grazie alla complementarità dei due modelli.¹⁴³

A margine di queste osservazioni, è opportuno sottolineare che la compatibilità dell'interfaccia utente con lo standard *IIIF* (*International Image Interoperability Framework*) rappresenta un elemento fondamentale per la sostenibilità delle edizioni scientifiche digitali. Ricordiamo, infatti, che si tratta di uno standard aperto che consente di pubblicare, condividere e riutilizzare immagini ad alta risoluzione in modo interoperabile tra piattaforme diverse; grazie a IIIF, le immagini digitali non rimangono confinate in un singolo progetto o repository, ma possono essere integrate in ambienti eterogenei – come biblioteche digitali, piattaforme di annotazione, viewer avanzati o siti di ricerca collaborativa – senza duplicare file o dipendere da formati proprietari. Questo approccio non solo favorisce il riuso scientifico, ma garantisce anche maggiore longevità dei contenuti, poiché le immagini restano accessibili anche se l'infrastruttura originaria viene aggiornata o dismessa. Dal punto di vista della sostenibilità, IIIF riduce i costi di manutenzione e di migrazione tecnologica: un'immagine servita secondo lo standard può essere visualizzata e manipolata da qualsiasi viewer compatibile, come *Mirador* o *Universal Viewer*, evitando lo sviluppo di soluzioni personalizzate e non riutilizzabili. Inoltre, la compatibilità con IIIF favorisce la collaborazione istituzionale, permettendo di combinare risorse provenienti da archivi e biblioteche diverse in un unico spazio di lavoro digitale e ampliando le possibilità di analisi e comparazione. L'integrazione con lo standard IIIF rappresenta uno degli sviluppi più significativi nell'evoluzione di *EVT* (*Edition Visualization Technology*). Già con *EVT 2* è stata introdotta la possibilità di connettere i testi codificati in TEI con i facsimili digitali resi disponibili da server IIIF, consentendo così di affiancare all'edizione diplomatica o critica la riproduzione ad alta risoluzione delle fonti senza dover gestire localmente i file. Con *EVT 3*

¹⁴³ Per approfondire, cfr. E. Pierazzo, *What future for digital scholarly editions? From Haute Couture to Prêt-à-Porter*, «International Journal of Digital Humanities», 2019.

questa funzionalità è stata ulteriormente potenziata: l'interfaccia utente permette una navigazione più flessibile tra trascrizione e immagine, sfruttando appieno le API IIIF per zoom, pannellizzazione e annotazioni interoperabili. L'adozione dello standard garantisce inoltre maggiore sostenibilità e interoperabilità, poiché le edizioni prodotte con *EVT* possono integrare risorse provenienti da archivi e biblioteche diverse, offrendo agli studiosi un ambiente di lavoro dinamico, aperto e conforme alle buone pratiche internazionali.¹⁴⁴

Anche Transkribus offre una compatibilità con lo standard IIIF, anche se con alcune limitazioni. In particolare, è possibile importare direttamente i *manifest IIIF* – attualmente nella versione 2.1 della *Presentation API* – semplicemente incollando l'URL corrispondente durante la fase di caricamento. Questo permette di lavorare su immagini conservate in repository esterni senza doverle scaricare e ricaricare manualmente, semplificando notevolmente il flusso di lavoro. Al momento, però, il supporto si ferma a questa versione dello standard: se le immagini provengono da sistemi che utilizzano API più recenti, l'importazione può risultare problematica. Nonostante questo limite, l'integrazione con IIIF rappresenta già un passo importante verso una gestione più interoperabile delle immagini. In ambito accademico si discute inoltre di sviluppi futuri che potrebbero rendere questa integrazione ancora più efficace: ad esempio, la possibilità di avviare processi di riconoscimento partendo da un semplice ID IIIF, di restituire *manifest* arricchiti dopo la trascrizione, o di collegare in modo diretto le interfacce di trascrizione con sistemi di intelligenza artificiale. Queste prospettive non sono ancora realtà, ma mostrano chiaramente la direzione verso cui *Transkribus* potrebbe evolversi, rafforzando ulteriormente il dialogo con l'ecosistema IIIF. L'adozione di IIIF si configura, quindi, non solo come una scelta tecnica, ma

¹⁴⁴ Per approfondire, cfr. R. Rosselli Del Turco e P. Monella, *Extending the DSE: LOD Support and TEI/IIIF Integration in EVT*, in *Atti del IX Convegno Annuale AIUCD. La svolta inevitabile: sfide e prospettive per l'informatica umanistica* (Milano, 15-17 gennaio 2020), a cura di Cristina Marras et al., pp. 148–55, consultabile al seguente link: <<https://doi.org/10.6092/UNIBO/AMSACTA/6316>> (ultimo accesso 17 ottobre 2025). Gli attuali motori di HTR consentono lo sviluppo di modelli generalizzati (i cosiddetti “super model”). Tali modelli sono in grado di processare non una singola grafia specifica, bensì una molteplicità di scritture storiche all'interno di un medesimo sistema grafico, estendendosi su un intervallo temporale considerevole (superiore a un secolo). Per approfondire l'argomento, cfr. T. Hodel et al., *General Models for Handwritten Text Recognition: Feasibility and State-of-the Art. German Kurrent as an Example*, «Journal of Open Humanities Data», 7, 2021, consultabile al seguente link: <<https://doi.org/10.5334/johd.46>> (ultimo accesso 17 ottobre 2025); REED-COOP, *Should I use a base model when training a custom model?*, «Transkribus blog», 7 febbraio 2025, consultabile al seguente link: <<https://blog.transkribus.org/en/using-base-model-when-training-a-custom-model>> (ultimo accesso 17 ottobre 2025); REED-COOP, *Introducing Transkribus Super Models - Get Access to 'The Text Titan I'*, «Transkribus blog», 9 luglio 2023, consultabile al seguente link: <<https://blog.transkribus.org/en/introducing-transkribus-super-models-get-access-to-the-text-titan-i>> (ultimo accesso 17 ottobre 2025).

come una vera e propria scelta editoriale strategica, capace di coniugare elevata qualità scientifica con una prospettiva di conservazione e riuso a lungo termine.¹⁴⁵

Dunque, in questo capitolo si è inteso delineare una panoramica generale della metodologia che presiede alla creazione di un'edizione scientifica digitale. Tale metodologia si articola in una serie di fasi, che vanno dalla digitalizzazione, archiviazione e metadattazione dei documenti sotto forma di immagini ad alta risoluzione, alla trascrizione automatica assistita da sofisticati sistemi di HTR, fino alla codifica secondo lo standard XML-TEI e alla pubblicazione online tramite la configurazione di un'interfaccia di visualizzazione semplice e intuitiva. L'analisi dettagliata di ciascuna di queste fasi, con particolare riferimento ai frammenti autografi dell'*Orlando furioso*, sarà oggetto dei capitoli successivi. A conclusione di questa sezione introduttiva, si è ritenuto opportuno richiamare l'attenzione su un progetto digitale che – a nostro avviso – rappresenta un modello di eccellenza nell'ambito della conservazione, valorizzazione e diffusione del patrimonio culturale testuale; si tratta del *Progetto Saba*, da cui sono stati tratti preziosi spunti metodologici che hanno informato il nostro lavoro tanto sul versante dell'impostazione filologico-editoriale quanto su quello della gestione e promozione dell'iniziativa.¹⁴⁶ Il *Progetto Saba* ha avuto origine da un *project work* realizzato da Eleonora Morante nell'ambito del *Master in Digital Humanities* dell'Università Ca' Foscari di Venezia (2019), sotto la supervisione scientifica di Marina Buzzoni e con il supporto di Cristina Fenu, bibliotecaria presso la Biblioteca Civica Attilio Hortis di Trieste. Il fulcro dell'iniziativa è rappresentato dalla creazione di un'edizione genetica digitale del manoscritto menabò del *Canzoniere* di Umberto Saba, conservato presso la suddetta biblioteca. L'obiettivo perseguito dal progetto è duplice: valorizzare questo documento unico in ambito museale e renderlo accessibile anche da remoto a un pubblico ampio, non esclusivamente specialistico.

Il manoscritto in oggetto (ms. R.P. MS 1-18) è un quaderno di 214 pagine contenente 186 liriche, organizzate in sezioni e sottosezioni, che attesta la prima redazione assoluta del *Canzoniere* (nota come C19). Si tratta di un documento di straordinario valore storico e culturale, in quanto

¹⁴⁵ Per approfondire, cfr. Transkribus Help Center, *Uploading files to Transkribus*, consultabile al seguente link: <<https://help.transkribus.org/uploading-files-to-transkribus>> (ultimo accesso 17 ottobre 2025); F. Krull, G. Muehlberger e M. Terras, *Transkribus and IIF: beneficial possibilities between image sharing and Handwritten Text Recognition frameworks* (Atti della 2019 IIF Conference, Göttingen, Germany), abstract consultabile al seguente link: <<https://iif.io/event/2019/goettingen/program/26/>> (ultimo accesso 17 ottobre 2025).

¹⁴⁶ Per approfondire l'argomento, cfr. M. Buzzoni et al., *Progetto di edizione genetica digitale del Canzoniere manoscritto di U. Saba (1919-20)*, in *ME.TE. Digitali - Mediterraneo in rete tra testi e contesti*, Atti del XIII convegno annuale AIUCD (Università degli Studi di Catania, 28-30 maggio 2024), a cura di A. Di Silvestro et al., AIUCD, 2024, pp. 215-220.

testimonia una versione inedita dell'opera sabiana, precedente a quella pubblicata nel 1921 (Umberto Saba, *Il Canzoniere. 1900-1921*, Trieste, La Libreria Antica e Moderna, 1921, redazione nota come C21), che costituisce tuttora la base dell'edizione critica di riferimento, a cura di Giordano Castellani.¹⁴⁷ Il menabò, approntato dallo stesso autore, raccoglie liriche composte tra il 1900 e il 1919, alcune delle quali erano già apparse nelle raccolte *Poesie* (Firenze, Casa editrice fiorentina, 1911) e *Coi miei occhi. Il mio secondo libro di versi* (Firenze, Edizioni della «Voce», 1912). Saba non si ritenne mai pienamente soddisfatto di queste edizioni, lamentando sia l'adozione di un ordine cronologico arbitrario, sia le alterazioni apportate alla forma originaria dei testi in sede tipografica, in netto contrasto con la sua ultima volontà espressiva. L'intento, poi disatteso, di stampare questa redazione dell'opera è comprovato da due foglietti allegati al manoscritto: uno intitolato *Avvertenze per lo stampatore* e l'altro recante l'indirizzo per l'invio delle bozze. Il quaderno presenta una complessa stratificazione di interventi autoriali – quali notazioni editoriali, correzioni e riscritture – distribuiti lungo cinque fasi correttorie, ciascuna delle quali è identificata dall'uso di uno strumento grafico differente: inchiostro blu piombo per il testo base, penna blu chiaro, matita grigia, matita copiativa, penna blu oltremare, matita viola e penna rossa (quest'ultima riservata alle indicazioni tipografiche). Accanto al testo poetico, inoltre, si trovano numerosi cartigli e frammenti cartacei adesi alle pagine del quaderno, provenienti da altri taccuini o dallo stesso supporto. La varietà e la complessità degli interventi autoriali fanno di questo manoscritto un caso esemplare di documento genetico, offrendo al lettore uno spaccato prezioso del processo creativo e correttorio di Saba. Il gruppo di ricerca del *Progetto Saba* ha stimato che il 67,20% delle liriche e il 31,27% dei cartigli presenta varianti autoriali rilevanti, cosicché il documento si profila come banco di prova ideale per la sperimentazione di modelli innovativi di codifica orientati all'edizione genetica digitale.¹⁴⁸

Il progetto è il risultato della collaborazione di un team interdisciplinare con competenze che spaziano dalla filologia d'autore alle diverse branche delle *digital humanities*, e si avvale del sostegno istituzionale di numerosi enti e fondazioni. Oltre all'Università Ca' Foscari di Venezia

¹⁴⁷ Il manoscritto R.P. MS 1-18 è stato descritto e studiato da Giordano Castellani per l'edizione critica del *Canzoniere 1921* (Umberto Saba, *Il Canzoniere 1921*, edizione critica a cura di Giordano Castellani, Milano, Fondazione Arnoldo e Alberto Mondadori, 1981) e più recentemente da Silvia Vodopivec nel saggio *Le penne e le matite di Saba. Tracce di volontà autoriale perduta nel 'Canzoniere' (R.P. Ms. 1-18, Biblioteca Civica 'A. Hortis' di Trieste)* (Trieste, Università degli Studi di Trieste, 2016).

¹⁴⁸ Per approfondire, cfr. Giulia Tancredi e Cristina Fenu, *XML-TEI: Un modello per la filologia d'autore*, in *AIUCD 2022 - Culture digitali. Intersezioni: filosofia, arti, media*, a cura di F. Ciraci, G. Miglietta e C. Gatto, Lecce, Associazione per l'Informatica Umanistica e la Cultura Digitale, 2022, pp. 218-222, p. 219.

e alla Biblioteca Civica Attilio Hortis di Trieste sopra menzionate, partecipano anche il Museo della Letteratura di Trieste (LETS) e diversi atenei nazionali e internazionali, tra cui le Università di Trieste, Bologna, Pisa, Torino e il Boston College (Massachusetts). L'iniziativa si articola in due macro-fasi distinte: la prima riguarda la codifica del testo, a partire dalla digitalizzazione e dalla trascrizione diplomatica del manoscritto; la seconda è dedicata allo sviluppo delle componenti museali e promozionali del progetto. Per quanto concerne la prima fase, la codifica del manoscritto sabiano è stata realizzata mediante il linguaggio di marcatura XML secondo le linee guida della *Text Encoding Initiative* (TEI P5).¹⁴⁹ Lo schema di codifica è stato elaborato da specialisti del settore per rispondere in modo efficace alle ipotesi filologiche alla base della costituzione del testo genetico criticamente stabilito e della costruzione degli apparati critici.¹⁵⁰ In considerazione della natura stratificata del manoscritto sabiano, il sistema di marcatura contempla quattro livelli fondamentali del testo critico:

1. il livello testuale, che comprende le lezioni selezionate dal filologo per la costituzione del testo;
2. il livello intratestuale, che riguarda le varianti alternative, ovvero le forme tra cui l'autore non ha operato una scelta definitiva;
3. il livello metatestuale, che comprende le postille autoriali;
4. il livello genetico, che documenta le correzioni apportate dall'autore dopo la stesura del testo base, organizzate secondo la loro progressione cronologica.

Nel livello testuale, ogni lezione scelta dal filologo è codificata all'interno dell'apparato <app> mediante l'elemento <lem>. Questo elemento, nella prassi, identifica solitamente l'ultima delle varianti, considerata quella che l'autore intendeva destinare alla stampa. Tuttavia, qualora il filologo riconosca come espressione dell'ultima volontà autoriale una variante intermedia o il testo di base, la codifica si adatterà a tale interpretazione filologica. In fase di visualizzazione,

¹⁴⁹ Per approfondire, cfr. TEI Consortium, *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), consultabile al seguente link: <<https://www.tei-c.org/guidelines/p5/>> (ultimo accesso 17 ottobre 2025); TEI Manuscript Special Interest Group (TEI SIG MS), *An Encoding Model for Genetic Editions*, consultabile al seguente link: <<https://tei-c.org/Vault/TC/tcw19.html>> (ultimo accesso 17 ottobre 2025).

¹⁵⁰ La codifica XML-TEI è stata elaborata da un team multidisciplinare, coordinato da Cristina Fenu (Biblioteca Civica Hortis) e Giulia Fabbris (Università Ca' Foscari), con il contributo di esperti come Paola Italia, Roberto Rosselli Del Turco, Giulia Tancredi e Marina Buzzoni. La collaborazione ha permesso di rivedere e ottimizzare il modello di marcatura per adeguarlo meglio agli obiettivi scientifici dell'edizione genetica e alle potenzialità offerte dal software di visualizzazione prescelto, *EVT*.

l'estrazione automatica delle lezioni marcate con <lem> permette di presentare al lettore un testo critico "pulito", ovvero privo di interventi autoriali, e quindi di più agevole lettura.

Nel livello genetico, due sono gli aspetti fondamentali da considerare nella definizione dello schema di codifica:

1. la codifica delle correzioni, volta a mettere in evidenza i diversi strati scrittori;
2. la codifica delle varianti genetiche.

Una prima classificazione sazio-temporale riconoscibile nel manoscritto riguarda proprio gli strati scrittori, intesi come insiemi di interventi riconducibili a una specifica fase o campagna revisionale. Per rappresentare tali strati, si ricorre, all'interno del <teiHeader>, all'elemento <creation>, che accoglie una lista di <change> racchiusa in <listChange>: ciascun <change> corrisponde a uno strato scrittoriale ed è ordinato in senso cronologico.¹⁵¹ Questo modello è flessibile e può essere replicato: ad esempio, si possono creare liste distinte per isolare particolari gruppi di interventi, come lo strato relativo alle indicazioni tipografiche in penna rossa, intenzionalmente separato dalle varianti testuali propriamente dette. Nel corpo del testo (<body>), il riferimento agli strati scrittori è gestito tramite l'elemento <mod>, introdotto dal TEI SIG MS (il TEI - Special Interest Group on Manuscripts è il gruppo di lavoro che all'interno del TEI Consortium si occupa specificamente della codifica dei manoscritti) che segnala ogni modifica autoriale, includendo le micro-correzioni associate. Tra gli elementi utilizzabili si ricordano: per le cancellature, <add> per le aggiunte e <transpose> per le trasposizioni.

¹⁵¹ Il tag <teiHeader> è una componente obbligatoria di ogni documento TEI e rappresenta l'intestazione del file, ovvero la sezione che raccoglie tutte le informazioni descrittive, bibliografiche e tecniche necessarie per comprendere e contestualizzare il testo codificato. In particolare, <teiHeader> specifica: chi ha scritto, curato e codificato il testo; qual è la fonte (manoscritto, edizione a stampa, ecc.); com'è stato codificato (criteri, standard, scelte editoriali); quando e da chi è stato modificato; quali sono le caratteristiche linguistiche e testuali del contenuto. In pratica, costituisce l'equivalente digitale di una scheda catalografica, ma con maggior dettaglio e flessibilità.

Il tag <listChange> è un elemento della codifica TEI impiegato per documentare e organizzare in ordine cronologico le modifiche apportate a un documento, sia durante la fase redazionale sia nelle successive revisioni editoriali o digitali. Al suo interno vengono registrate diverse tipologie di intervento, tra cui correzioni e revisioni autoriali, modifiche editoriali (relative a edizioni a stampa o digitali) e operazioni di codifica, come la normalizzazione del testo, la marcatura linguistica o l'aggiornamento della struttura. Il contenuto di <listChange> è costituito da una serie di elementi <change>, ognuno dei quali descrive una singola modifica. Questo elemento può essere utilizzato in vari contesti, ma compare più frequentemente:

- all'interno di <teiHeader>, in particolare nella sezione <revisionDesc>, dove ha la funzione di tracciare la cronologia delle revisioni del file;
- all'interno di <creation>, quando si desidera registrare le diverse fasi di elaborazione o gli strati redazionali del testo.

Fig. 1 Codifica degli strati correttori mediante i tag <creation>, <listChange> e <change>

```

<creation corresp="#MS_C19" resp="#SV">
  <listChange ordered="true">
    <change xml:id="strato-0">Stesura base del testo, in inchiostro blu piombo; presenza di
varianti immediate. </change>
    <change xml:id="strato-1">Primo strato di varianti apportate con penna a inchiostro blu
chiaro. </change>
    <change xml:id="strato-2">Secondo strato di varianti e correzioni apportate in matita
grigia. </change>
    <change xml:id="strato-3">Terzo strato di varianti e correzioni annotate a matita copiativa.
</change>
    <change xml:id="strato-4">Quarto strato di varianti e correzioni annotate in penna blu
oltremare. </change>
    <change xml:id="strato-5">Quinto strato scrittoria con correzioni e varianti in matita
viola. </change>
  </listChange>
  <listChange ordered="false">
    <change xml:id="strato-6">Strato scrittoria con indicazioni tipografiche in penna
rossa. </change>
  </listChange>
</creation>

<creation corresp="#peccette">
  <listChange ordered="false">
    <change xml:id="strato-a">Testo base in penna nera. </change>
    <change xml:id="strato-b">Penna ocre? O penna del testo base scarica? </change>
    <change xml:id="strato-c">Penna grigio-azzurra. </change>
  </listChange>
</creation>

```

Se osserviamo il codice riportato sopra, possiamo notare la presenza di due elementi <creation>, ciascuno dei quali descrive un insieme distinto di interventi autoriali. Il primo <creation> è associato all'attributo `corresp="#MS_C19"`, che fa riferimento al manoscritto menabò del Canzoniere, identificato con il valore `#MS_C19`. Questo elemento contiene:

- cinque strati correttivi ordinati cronologicamente, distinguibili per il diverso strumento o colore utilizzato, racchiusi all'interno dell'elemento `<listChange ordered="true">`;
- uno strato aggiuntivo, relativo a indicazioni tipografiche vergate in penna rossa, la cui collocazione temporale non è certa, e che pertanto è incluso nell'elemento `<listChange ordered="false">`.

Il secondo <creation> è associato all'attributo corresp="#peccette", che fa riferimento a un altro insieme di interventi autoriali, ovvero alle pezzette correttive (strisce di carta applicate sul testo originale contenenti nuove stesure). Questo elemento include un <listChange ordered="false">, che descrive tre strati correttori non ordinati, a causa della mancanza di dati certi sulla loro successione temporale.

Una seconda classificazione spazio-temporale, invece, riguarda la sequenza delle varianti genetiche che sono codificate all'interno dell'apparato <app>; l'apparato accoglie sia la lezione messa a testo (identificata dall'elemento <lem>), sia le varianti escluse dal testo contrassegnate con l'elemento <rdg> e numerate tramite l'attributo @varSeq. A scopo esemplificativo, si può richiamare lo stesso esempio riportato da Marina Buzzoni.¹⁵²

```
<l n="5">
  <app>
    <lem varSeq="2">
      <mod change="strato-5">
        <add place="margin-left">fanno</add>
      </mod>
    </lem>
    <rdg varSeq="1">
      <mod change="strato-5">
        <del rend="strikethrough">tornan</del>
      </mod>
    </rdg>
  </app>
  fra noi
</l>
```

L'esempio illustra come la lezione selezionata dal filologo per essere messa a testo coincida con l'ultima correzione d'autore: la parola fanno, codificata all'interno dell'elemento <app> tramite <lem> e l'attributo @varSeq="2". L'elemento <lem> contiene a sua volta <mod>, corredato dell'attributo change="strato-5", che indica che la modifica è avvenuta nel quinto strato scrittoria (quello in matita viola), identificato in precedenza all'interno del <teiHeader>. All'interno di <mod> è specificata la tipologia della micro-correzione, in questo caso <add>, poiché si tratta di

¹⁵² Id., *Il progetto Saba: dare voce a un manoscritto inedito del 'Canzoniere'*, in *Informatica umanistica, digital humanities: verso quale modernità?*, a cura di M. Gatto et al., Bari, Cacucci Editore, 2024, pp. 153-165, p. 161.

un'aggiunta. È inoltre presente l'attributo *@place*, che serve a indicare il luogo fisico della correzione: in questo caso, il margine sinistro del manoscritto. Il termine *tornan*, invece, cancellato dall'autore mediante una serie di righe sovrascritte, è stato anch'esso incluso all'interno di *<app>*, ma in qualità di variante. La cancellatura testimonia chiaramente l'intenzione dell'autore di escludere questa forma dal testo finale; pertanto, essa è codificata mediante l'elemento *<rdg>*. Anche in questo caso, *<rdg>* include un elemento *<mod>*, che specifica la micro-correzione mediante il tag **, trattandosi di una soppressione.

Lo schema si ripete per tutte le varianti alternative che costituiscono il livello intratestuale: anch'esse sono racchiuse all'interno di elementi *<rdg>*, ciascuno dei quali è dotato di una codifica specifica per distinguerle tra loro. È previsto che anche queste varianti contengano una parte trascrizionale, strutturata secondo il medesimo modello: l'uso di *<mod>* con l'attributo *@change* per indicare lo strato scrittoria, *<add>* con l'attributo *@place* per localizzare l'intervento nel manoscritto e l'attributo *@seq* per indicare la sequenza. A livello metatestuale, le note autoriali sono marcate con *<note>* e dall'attributo di responsabilità *@resp* per distinguerle chiaramente dalle note inserite dai curatori dell'edizione.

Il modello di codifica adottato è stato concepito in modo da essere pienamente compatibile con il software di visualizzazione selezionato dal team del progetto: *Edition Visualization Technology (EVT)*, sviluppato da Roberto Rosselli Del Turco e dal suo gruppo di lavoro. Questo strumento consente una visualizzazione sinottica del manoscritto e del testo codificato, presentando nella colonna di sinistra la scansione digitale della pagina manoscritta, e in quella di destra la trascrizione codificata, arricchita da box colorati e collegamenti ipertestuali che facilitano l'esplorazione delle corrispondenze tra le due versioni. Tra le funzionalità più significative vi è la possibilità di filtrare le lezioni testuali mostrate al lettore. Ad esempio, è possibile visualizzare unicamente le lezioni selezionate dal filologo (ovvero quelle marcate con il tag *<lem>*), offrendo al lettore un testo "in pulito". In alternativa, si può attivare una visualizzazione stratificata, che include le varianti e ne evidenzia la successione cronologica. In questa modalità, ciascuna variante è accompagnata da un esponente numerico indicante lo strato correttorio di appartenenza, ed è collegata ad altre varianti tramite frecce direzionali, consentendo al lettore di ricostruire il processo di revisione del testo lirico. Un ulteriore livello di fruizione è offerto dalla georeferenziazione delle entità presenti nei testi, quali toponimi, antroponomi o luoghi simbolici. Nell'interfaccia del software è disponibile un elenco di entità (ad esempio *Caffè dei negozianti* o *Cimitero Ebraico*), ciascuna accompagnata da una scheda descrittiva di taglio storico-culturale e

collegata, quando disponibili i dati, a contenuti esterni o mappe interattive. Questa funzionalità contribuisce ad ampliare la dimensione divulgativa del progetto, rendendolo accessibile e interessante anche per un pubblico non strettamente accademico. Attraverso il legame tra le liriche e gli spazi urbani evocati nei versi sabiani, la piattaforma costruisce una sorta di mappa poetica della città reale, in cui testo e territorio si riflettono e si richiamano a vicenda.

In conclusione, la piattaforma digitale, fruibile sia da remoto che all'interno del Museo della Letteratura di Trieste, consentirà agli utenti di sfogliare virtualmente il manoscritto e di seguire, in chiave genetica, il processo redazionale delle singole liriche. Il lettore sarà così posto nelle condizioni di accedere al laboratorio creativo dell'autore, potendo osservare da vicino le fasi di riscrittura e le riflessioni che hanno accompagnato l'evoluzione testuale fino alla forma destinata alla stampa. In questa prospettiva, il progetto intende superare la concezione tradizionale del testo come entità statica e conclusa, per proporlo invece quale testimonianza dinamica di un processo creativo aperto, stratificato e ancora in divenire. Inoltre, offrirà al pubblico un'esperienza immersiva, mettendo in relazione la Trieste di Saba con i versi del poeta. Attraverso una biblioteca diffusa, sarà possibile esplorare luoghi, caffè, botteghe e strade cantati nei testi, navigando tra materiali testuali e audiovisivi. Il *Progetto Saba* rappresenta dunque un modello avanzato di edizione digitale, in cui la codifica XML-TEI e il software *EVT* consentono di coniugare rigore filologico e accessibilità pubblica, dando voce a un patrimonio culturale altrimenti destinato all'oblio. In quest'ottica, l'edizione genetica non si configura solo come un prodotto accademico, ma anche come uno strumento di valorizzazione partecipata, capace di coinvolgere studiosi, studenti e lettori comuni, trasformando l'atto editoriale in un'esperienza condivisa.

III

CAPITOLO

3.1 Digitalizzazione e trascrizione automatica dei frammenti autografi

Nel presente paragrafo si ripercorre l'iter che ha condotto alla digitalizzazione dei frammenti autografi, intesa sia come creazione di immagini digitali delle fonti documentarie, sia come trasformazione del contenuto testuale in formato elettronico leggibile dalle macchine (*machine-readable form*). Prima di trattare nel dettaglio questo processo, è necessario soffermarsi brevemente sul quadro giuridico di riferimento relativo a tali documenti, in quanto esso comporta una serie di implicazioni che incidono direttamente sulla realizzabilità e modalità di pubblicazione dell'edizione scientifica digitale (ESD).

I frammenti autografi rientrano nella fattispecie dei “beni culturali in pubblico dominio” che, per definizione, non sono più soggetti al diritto d'autore per scadenza dei termini temporali previsti dalla L. 633/1941. Inoltre, secondo quanto stabilisce l'art. 32-quater della stessa legge, anche le copie digitali di tali documenti, intese come “riproduzioni fedeli o mere riproduzioni” (cioè prive di carattere creativo), non sono sottoposte al diritto d'autore,¹⁵³ ma soggiacciono esclusivamente alle disposizioni del *Codice dei beni culturali e del paesaggio* (d.lgs. 42/2004). A tal proposito, si ricorda che l'art. 108, comma 3-bis del *Codice*, riformato dalla L. 124/2017, liberalizza (cioè rende gratuita ed esente da autorizzazione) sia l'esecuzione di riproduzioni di beni archivistici e librari, sia la divulgazione, con qualsiasi mezzo, delle medesime riproduzioni da parte di soggetti pubblici o privati, purché esse siano attuate senza scopo di lucro e nel rispetto delle disposizioni vigenti a tutela della riservatezza, del diritto d'autore e dell'integrità del bene culturale.¹⁵⁴

¹⁵³ La legge sul diritto d'autore (LdA 633/1941), art. 32-quater recita: «Alla scadenza della durata di protezione di un'opera delle arti visive, anche come individuate all'articolo 2, il materiale derivante da un atto di riproduzione di tale opera non è soggetto al diritto d'autore o ai diritti connessi, salvo che costituisca un'opera originale. Restano ferme le disposizioni in materia di riproduzione dei beni culturali di cui al decreto d.lgs. 42/2004».

¹⁵⁴ L'art. 108, comma 3-bis oggi in vigore rappresenta l'esito delle modifiche introdotte dalla L. 124/2017 nel testo precedente (d.lgs. 83/2014) e recita: «Sono in ogni caso libere le seguenti attività, svolte senza scopo di lucro, per finalità di studio, ricerca, libera manifestazione del pensiero o espressione creativa, promozione della conoscenza del patrimonio culturale: 1) la riproduzione di beni culturali diversi dai beni archivistici sottoposti a restrizioni di consultabilità ai sensi del capo III del presente titolo, attuata nel rispetto delle disposizioni che tutelano il diritto di

Tuttavia, il suddetto articolo specifica non solo il divieto di divulgare le riproduzioni a fini commerciali, ma anche l'obbligo di adottare modalità di pubblicazione che impediscano un loro eventuale riutilizzo a scopo di lucro. Di seguito sono illustrate, a titolo esemplificativo, le finalità d'uso delle riproduzioni di beni culturali in pubblico dominio richiamate dall'art. 108, comma 3-bis, alle quali può essere assimilata l'ESD dei frammenti autografi. Tali finalità comprendono:

- uso personale a fini di studio: ad esempio, riproduzioni effettuate da studenti, ricercatori o docenti per esigenze di studio individuale; oppure riproduzioni eseguite da professionisti che ne facciano uso nell'ambito di attività di studio connesse all'esercizio della propria professione, svolta in favore di terzi;¹⁵⁵
- attività di ricerca: ad esempio, le riproduzioni incluse all'interno di prodotti della ricerca scientifica diffusi in modalità *open access*, come tesi di laurea, master o dottorato, volumi e articoli pubblicati in riviste, accessibili gratuitamente da chiunque;
- libera manifestazione del pensiero ed espressione creativa: ad esempio, riproduzioni pubblicate su blog, social network o siti web a libero accesso.

Per quanto riguarda quest'ultima categoria, è importante sottolineare che gli obiettivi commerciali perseguiti dai principali *content provider* non rendono di per sé lucrativa la semplice condivisione di riproduzioni sulle piattaforme online che essi gestiscono. Ciò è valido a condizione che i termini d'uso di tali piattaforme riconoscano agli utenti la titolarità delle riproduzioni da essi pubblicate.

La pubblicazione online delle riproduzioni digitali comporta il potenziale rischio di un riutilizzo improprio rispetto alle finalità sopra esposte. Per prevenire tale evenienza, le *Linee guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale* (Piano Nazionale di Digitalizzazione del Patrimonio Culturale - Allegato 3) raccomandano agli istituti di conservazione titolari dei beni culturali di informare gli utenti sui termini di utilizzo delle digitalizzazioni, mediante l'adozione di una licenza conforme alla normativa prevista dal

autore e con modalità che non comportino alcun contatto fisico con il bene, né l'esposizione dello stesso a sorgenti luminose, né, all'interno degli istituti della cultura, l'uso di stativi o treppiedi; 2) la divulgazione con qualsiasi mezzo delle immagini di beni culturali, legittimamente acquisite, in modo da non poter essere ulteriormente riprodotte a scopo di lucro».

¹⁵⁵ Nel caso di riproduzioni per attività di studio finalizzate all'espletamento di pratiche professionali, l'oggetto del lucro non è da considerarsi la riproduzione del bene culturale in sé, quanto il lavoro d'opera intellettuale oggetto della prestazione professionale.

Codice.¹⁵⁶ Le licenze d'uso oggi più comunemente adottate a livello internazionale sono, com'è noto, le *Creative Commons* (CC).¹⁵⁷ Queste si rivelano particolarmente adatte per la condivisione di contenuti protetti dal diritto d'autore, ma risultano poco appropriate per la diffusione di riproduzioni di beni culturali pubblici in pubblico dominio, per le ragioni già illustrate. Tuttavia, l'organizzazione *Creative Commons* mette a disposizione anche due strumenti specifici pensati per le opere di pubblico dominio:

- il PDM (*Public Domain Mark*): un'etichetta pensata per segnalare agli utenti che un determinato contenuto non è soggetto a diritti d'autore e può quindi essere riutilizzato liberamente per qualsiasi finalità;
- il CC0: a differenza del PDM, ha valore legale e consente all'autore di rinunciare volontariamente a tutti i propri diritti, compreso il diritto di attribuzione (attributo BY, *Creative Commons*), facendo così entrare l'opera nel pubblico dominio prima della scadenza del periodo di tutela.

Nonostante ciò, anche questi strumenti si rivelano poco adatti al contesto normativo italiano. Il CC0, infatti, presuppone che vi sia un titolare originario dei diritti d'autore disposto a rinunciarvi – condizione che non si verifica nel caso delle semplici riproduzioni di beni già in pubblico dominio. Il PDM, invece, per quanto utile come indicazione generale, non risulta

¹⁵⁶ Le *Linee guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale* (consultabili al seguente link: <<https://digitallibrary.cultura.gov.it/wp-content/uploads/2023/06/PND-allegato3.pdf>>) forniscono agli istituti culturali metodologie e procedure utili per conseguire gli obiettivi strategici del PND 2022/2023 redatto dall'Istituto Centrale per la Digitalizzazione del Patrimonio Culturale - Digital Library. I principali riferimenti normativi delle *Linee guida* sono rappresentati dal *Codice dei beni culturali*, dalla direttiva europea 2019/1024 (PSI) con il relativo d.lgs. 200/2021 di recepimento e dalla direttiva europea 2019/790 (Copyright) con il relativo d.lgs. 177/2021 di recepimento.

¹⁵⁷ Le licenze CC, proposte nel 2002 da Lawrence Lessig, sono complessivamente sei e derivano dalla combinazione dei seguenti quattro attributi:

1. *Attribuzione/Attribution* (BY): l'utente è tenuto ad attribuire la paternità dell'opera nel modo indicato dall'autore stesso;
2. *Non opere derivate/No Derivatives* (ND): l'opera non può essere alterata o modificata dall'utente in nessun modo, né utilizzata per crearne una simile. È alternativa alla SA;
3. *Non commerciale/Non Commercial* (NC): l'opera non può essere sfruttata dall'utente per fini commerciali;
4. *Condividi allo stesso modo/Share Alike* (SA): l'opera può essere modificata e può circolare solo per il tramite di una licenza equivalente a quella originaria. È alternativa alla ND.

Per maggiori approfondimenti, è possibile consultare il seguente link: <<https://creativecommons.it/chapterIT/>> (ultimo accesso 17 ottobre 2025).

sufficientemente efficace, in quanto non considera le restrizioni previste dal *Codice* in merito al riutilizzo delle riproduzioni, il che potrebbe generare confusione, soprattutto a livello internazionale. Dunque, per gli istituti di conservazione si impone la necessità di fare riferimento a protocolli internazionali condivisi, come quello promosso dal consorzio RightsStatements.org, per regolamentare il riuso delle digitalizzazioni dei beni culturali.¹⁵⁸ Il consorzio RightsStatements.org mette a disposizione degli istituti di conservazione un insieme di dodici dichiarazioni standardizzate, utili per comunicare in modo chiaro al pubblico l'eventuale sussistenza di diritti di riproduzione sui beni digitalizzati. Va precisato, tuttavia, che queste dichiarazioni non costituiscono delle licenze, ma sono piuttosto delle etichette informative che possono essere associate a licenze o ad altri strumenti di condivisione che spesso non sono armonizzati a livello internazionale. In particolare, tra le dichiarazioni disponibili, *NoC-OKLR: No Copyright – Other Known Legal Restrictions* è l'unica che risulta pienamente compatibile con la normativa italiana in materia di tutela del patrimonio culturale.¹⁵⁹ Questa dichiarazione dovrà essere associata all'etichetta *MiC Standard* (attributi BY e NC, *Creative Commons*) che sintetizza le prescrizioni relative alla pubblicazione online delle riproduzioni di beni culturali in pubblico dominio appartenenti allo Stato italiano. In realtà, anche l'etichetta *MiC Standard* non costituisce una vera e propria licenza, in quanto non implica la cessione di diritti né interviene sul diritto d'autore.¹⁶⁰ Piuttosto, si limita a riassumere il contenuto delle norme attualmente in vigore in materia di riproduzione dei beni culturali pubblici, così come definite dal Ministero della Cultura (MiC). L'associazione dell'etichetta *MiC Standard* alle immagini digitali dei beni culturali – sia nei metadati esterni che interni alle immagini – serve a segnalare che la divulgazione è consentita per usi non commerciali (NC, *Creative Commons*), come previsto dall'art. 108, comma 3-bis del *Codice*. Qualsiasi uso a fini commerciali delle immagini digitali

¹⁵⁸ La condivisione di protocolli internazionali per il rilascio delle digitalizzazioni del patrimonio culturale è stata espressamente incoraggiata dalla *Raccomandazione (UE) 2021/1970* della Commissione Europea. Le dichiarazioni elaborate dal consorzio RightsStatements.org sono disponibili al seguente link: <<https://rightsstatements.org/en/>> (ultimo accesso 17 ottobre 2025). Le dichiarazioni sono state sviluppate per essere facilmente comprensibili sia dagli utenti umani sia dai sistemi automatizzati (come i motori di ricerca), grazie all'utilizzo delle tecnologie del web semantico.

¹⁵⁹ Tale dichiarazione recita: «L'uso di questo oggetto non è limitato dal diritto d'autore e/o dai diritti connessi. In una o più giurisdizioni è noto che leggi diverse dal diritto d'autore impongono restrizioni all'uso di questo oggetto. Si prega di fare riferimento all'organizzazione che ha messo l'oggetto a disposizione per maggiori informazioni». Essa è consultabile al seguente link: <<https://rightsstatements.org/page/NoC-OKLR/1.0/?language=it>> (ultimo accesso 17 ottobre 2025)

¹⁶⁰ Una licenza d'uso è un contratto attraverso il quale il titolare dei diritti (licenziante) stabilisce in anticipo quali diritti intende mantenere e quali eventualmente concedere a terzi (licenziatari) per l'utilizzo del bene.

contrassegnate da questa etichetta dovrà essere preventivamente autorizzato dagli istituti di conservazione titolari dei beni culturali e potrà essere soggetto al pagamento di un corrispettivo. Inoltre, l'etichetta *MiC Standard* riporta i dati identificativi dell'opera riprodotta, quali la provenienza (cioè l'istituto di conservazione), l'autore, il titolo e la segnatura archivistica (attributo BY, *Creative Commons*).

Sulla scorta della metodologia per le ESD descritta nel capitolo precedente, abbiamo contattato gli istituti di conservazione detentori dei frammenti autografi per richiedere i facsimili digitali da pubblicare all'interno dell'edizione digitale. Malgrado vi fossero i presupposti per procedere speditamente, questa operazione preliminare ha richiesto un notevole dispendio di tempo oltre che di risorse economiche e intellettuali. Le cause che hanno complicato il processo di acquisizione dei facsimili non sono imputabili né alla carenza di infrastrutture tecnologiche o competenze tecniche adeguate, né all'assenza di politiche culturali e norme giuridiche favorevoli. Infatti, negli ultimi anni, l'ampia disponibilità di tecnologie di *imaging* digitale altamente performanti, semplici e a basso costo, l'abbondanza di operatori specializzati del settore e l'esistenza di standard condivisi come il IIIF (*International Image Interoperability Framework*), hanno considerevolmente semplificato la creazione, la pubblicazione e l'interscambio di immagini digitali ad alta definizione. Inoltre, le campagne di digitalizzazione avviate dal MiC (da ultimo il PND 2022/2023 sopra menzionato) e le modifiche apportate dalla L. 124/2017 al testo dell'art. 108, comma 3-bis del *Codice dei beni culturali* hanno contribuito ad incentivare la transizione digitale. Tali complicazioni sono riconducibili invece a ragioni di altra natura, prima fra tutte la dispersione geografica dei documenti fra tre diverse biblioteche (la Biblioteca Comunale Ariostea di Ferrara, la Biblioteca Ambrosiana di Milano e la Biblioteca Nazionale di Napoli) con le quali è stato necessario intrattenere rapporti separati. A ciò si somma la mancata riproduzione dei documenti o – quand'anche sia stata effettuata – la mancata pubblicazione online delle immagini digitali da parte degli istituti titolari dei documenti, nonostante il prestigio letterario di cui essi godono presso la comunità accademica. Al riguardo, Le *Linee guida* del PND evidenziano che se una riproduzione è già stata realizzata dall'ente culturale nell'ambito di una campagna di digitalizzazione ed è pubblicamente accessibile online, l'utente può scaricarla autonomamente, senza la mediazione dell'ente né oneri economici.¹⁶¹

¹⁶¹ Tuttavia, nel caso in cui l'accesso all'immagine sia integrato da servizi digitali avanzati (visite virtuali online, video-presentazioni di opere, download massivo, accesso a contenuti extra, interrogazioni di banche dati correlate, altissima definizione), l'ente può richiedere un contributo economico a copertura dei costi sostenuti, in linea con la direttiva europea *Public Sector Information* 2019/1024.

Dunque, è importante favorire la disponibilità online delle immagini e la possibilità di scaricarle, purché siano rispettate le seguenti indicazioni:

- le immagini devono avere una risoluzione compresa tra 2500 e 3500 pixel sul lato lungo, per garantirne l'uso a fini di studio, ricerca e riutilizzo conforme alla normativa;
- la sovrimpressione di filigrane invasive all'interno delle immagini deve essere evitata (se presenti, le filigrane possono comparire solo nella cornice esterna);¹⁶²
- le immagini devono essere accompagnate da metadati completi, sia interni che esterni, contenenti dati identificativi e informazioni sulle possibilità di riutilizzo;
- per il download delle immagini, può essere prevista la registrazione degli utenti con una procedura semplice e veloce, utile anche a fini statistici. In tal caso, è necessario prestare attenzione alla tutela della privacy, alla gestione dei dati raccolti e ai costi connessi. Questa pratica permette una migliore conoscenza del pubblico online e contribuisce a orientare in modo più efficace le politiche di valorizzazione del patrimonio culturale.

In questa prospettiva, è particolarmente positiva l'iniziativa di alcuni musei, come la Pinacoteca di Brera, che hanno deciso di rendere disponibili online immagini in media-alta definizione senza *watermark*, limitandone il riuso libero ai soli scopi non commerciali. Questa scelta rappresenta un passo concreto verso una valorizzazione autentica del patrimonio culturale, conciliando accessibilità e tutela.

Per quanto riguarda gli autografi conservati presso la Biblioteca Ariostea (ms. Cl. I. A., It. 1.), che rappresentano il nucleo più consistente del manoscritto ferrarese – 54 carte, di cui una bianca e le altre scritte su entrambi i lati – non è stato necessario richiedere la digitalizzazione ex novo, in quanto le immagini erano già disponibili nella banca dati dell'istituto e ci sono state concesse gratuitamente per finalità di ricerca scientifica.¹⁶³ Infatti, la Biblioteca aveva provveduto a

¹⁶² Si veda la *Raccomandazione della Commissione Europea sulla digitalizzazione e l'accessibilità in rete dei materiali culturali e sulla conservazione digitale* (2011/711/UE).

¹⁶³ Secondo le *Linee guida* del PND, qualora un utente privato necessiti di una riproduzione non presente nella banca dati dell'istituto che conserva il bene, può farne richiesta; la riproduzione sarà realizzata dal servizio interno o da ditte incaricate e comporterà un rimborso spese a carico dell'utente, previsto dall'art. 108, comma 3 del *Codice*, indipendentemente dalla destinazione d'uso finale dell'immagine. Il rimborso dovrà essere parametrato sulla complessità tecnica della riproduzione e della fornitura e non sulla risoluzione dell'immagine. A questo proposito, Le *Linee guida* suggeriscono di evitare la consegna di immagini a bassa risoluzione, poiché non garantiscono una

riprodurre i documenti in precedenza, intraprendendo contestualmente un'importante operazione di restauro digitale, con l'obiettivo di ovviare alle frequenti richieste di consultazione da parte degli studiosi interessati, che rischiavano di accelerare il deterioramento materiale delle carte. Tuttavia, a nostro avviso, l'obiettivo della Biblioteca è stato solo parzialmente raggiunto; se da un lato, la digitalizzazione degli originali ed il conseguente restauro sono senz'altro iniziative lodevoli e assolutamente necessarie che assicurano la tutela di questo inestimabile patrimonio culturale nel lungo periodo, dall'altro lato, l'assenza di un sito web istituzionale dedicato che ne consenta la fruizione pubblica, rappresenta un serio limite in termini di promozione e condivisione dello stesso patrimonio, con evidenti ricadute negative sul piano della ricerca scientifica.¹⁶⁴ Per di più, la mancanza di una metadatazione completa delle immagini, realizzata secondo il protocollo *MiC Standard* indicato nelle *Linee guida* del PND, comporta il rischio che – una volta pubblicate online – le immagini possano essere riutilizzate dagli utenti in modi non conformi alle finalità previste dal *Codice*. Dall'analisi condotta sui file delle immagini digitali, è

reale tutela del bene e rischiano di alimentare la diffusione online di contenuti di scarsa qualità, che ostacolano la corretta valorizzazione culturale del patrimonio.

Se invece la riproduzione di un bene è già presente nella banca dati dell'ente che lo conserva, ma non è ancora disponibile online, essa potrà essere fornita al richiedente previo pagamento di un rimborso dei costi amministrativi e gestionali sostenuti per evadere la richiesta. L'importo del rimborso può essere calcolato sulla base dei parametri della *Tabella 2 - Rimborso per riproduzioni*, contenuta nel D.M. 108/2024 recante le *Linee guida per la determinazione degli importi minimi dei canoni e dei corrispettivi per la concessione d'uso dei beni in consegna agli istituti e luoghi della cultura del Ministero della cultura* (consultabile al seguente link: <https://media.cultura.gov.it/mibac/files/boards/be78e33bc8ca0c99bffa70aa174035096/DECRETI/ANNO%202024/DM%2021%20marzo%202024%20rep.%20108-signed.pdf>).

Tuttavia, nel caso in cui la richiesta provenga da un soggetto privato, l'istituto può valutare anche la concessione gratuita delle immagini nell'ambito di un eventuale accordo di valorizzazione.

¹⁶⁴ Un modello esemplare di digitalizzazione nella medesima area geografica è rappresentato dall'*Estense Digital Library* (EDL). Questa piattaforma digitale all'avanguardia – sviluppata dalla Biblioteca Estense Universitaria di Modena con il sostegno della Fondazione di Modena e il contributo del Centro di ricerca sulle *Digital Humanities* dell'Università di Modena e Reggio Emilia – si propone di digitalizzare e rendere disponibile online l'immenso e prezioso patrimonio culturale della biblioteca, che comprende manoscritti, incunaboli, documenti musicali e carte geografiche. Un tratto distintivo dell'EDL è la sua capacità di offrire, tra le prime in Italia, non solo la consultazione di cataloghi o descrizioni, ma anche l'accesso diretto ai testi in una definizione eccezionalmente alta. Tra le sue caratteristiche principali figurano:

- la possibilità per gli utenti di accedere a oltre 700.000 pagine di manoscritti e altri materiali digitalizzati;
- l'applicazione di tecnologie avanzate, quali scanner ad alta risoluzione (come *Hasselblad*) per la digitalizzazione e protocolli di conservazione digitale a lungo termine (come i file FITS) per garantire la durabilità dei dati;
- l'uso di standard come l'*International Image Interoperability Framework* (IIIF), che facilita la condivisione e l'utilizzo dei documenti digitalizzati con altre piattaforme e istituzioni culturali nel mondo;
- l'integrazione di funzionalità avanzate per gli utenti, quali la possibilità di comparare, condividere e annotare le opere visualizzate, creando anche percorsi personalizzati di consultazione.

infatti emerso che gli unici metadati presenti sono quelli EXIF, incorporati nei file stessi.¹⁶⁵ Risultano invece del tutto assenti i metadati descrittivi, amministrativi e strutturali, fondamentali per identificare correttamente i documenti riprodotti e per specificare le condizioni di riutilizzo delle immagini.¹⁶⁶

Queste considerazioni avvalorano la necessità di pubblicare online le riproduzioni facsimilari secondo il *framework* si utilizzo IIIF (cfr. § 2.2.2); infatti, il IIIF si integra perfettamente sia con la normativa vigente sul diritto d'autore sia con il *Codice dei beni culturali*, agendo come un ponte essenziale per la gestione e la fruizione responsabile delle risorse digitali. Il suo punto di forza risiede nella capacità di incorporare metadati precisi su copyright e licenze d'uso direttamente nei *manifest* che descrivono gli oggetti digitali, affinché chiunque utilizzi tali risorse sia immediatamente e chiaramente informato sui propri diritti e obblighi.¹⁶⁷ Questo scenario costituirà la preconditione necessaria per la concreta realizzazione dell'ESD dei frammenti autografi, concepita come una piattaforma *open access*, capace di coniugare in modo efficace gli obiettivi di tutela, valorizzazione e divulgazione del patrimonio. In quest'ottica, ci auguriamo di proseguire la nostra collaborazione con la Biblioteca Ariostea, con l'obiettivo comune di rendere il progetto editoriale sempre più accessibile e perfezionabile nel tempo.

Un'altra considerazione, di genere diverso, sulla digitalizzazione dei frammenti concerne la difficoltà, sperimentata dal fruitore delle immagini digitali, di avere una cognizione esaustiva

¹⁶⁵ I metadati EXIF (*Exchangeable Image File Format*) descrivono i parametri tecnici usati al momento dell'acquisizione dell'immagine, ovvero al momento della scansione o dello scatto fotografico.

¹⁶⁶ Le *Linee guida* del PND, raccomandano agli enti culturali nazionali di codificare i metadati descrittivi, amministrativi e strutturali associati agli oggetti culturali digitali, mediante lo standard internazionale METS (*Metadata Encoding and Transmission Standard*). Il METS è alla base del protocollo di metadatazione condiviso *MiC Standard* sopra menzionato.

¹⁶⁷ Il IIIF va ben oltre la semplice visualizzazione di immagini, in quanto offre un *framework* solido e trasparente per comunicare lo status legale delle riproduzioni digitali. Ciò avviene mediante i metadati integrati direttamente nel *manifest IIIF*, un file in formato JSON che funge da vera e propria "scheda identificativa" dell'oggetto digitale e include sezioni specifiche dedicate ai diritti di utilizzo. Nel *manifest*, è possibile specificare in modo dettagliato:

- la licenza d'uso (ad esempio, *Creative Commons* o altre licenze specifiche);
- le informazioni sull'attribuzione richiesta;
- eventuali restrizioni relative all'uso commerciale o ad altre finalità;
- ogni altro dettaglio rilevante per il *copyright* e i riferimenti al *Codice dei beni culturali* per le riproduzioni di beni culturali.

Inoltre, i *viewer IIIF* (i software che permettono di visualizzare e interagire con i contenuti digitali conformi a IIIF) sono progettati per mostrare in modo chiaro ed immediato le informazioni sul copyright e le licenze, affinché l'utente venga subito a conoscenza delle condizioni d'uso dell'immagine o del contenuto digitale prima ancora di scaricarlo o riutilizzarlo. Questa visibilità immediata è fondamentale per promuovere un utilizzo lecito e rispettoso del patrimonio culturale digitalizzato.

delle caratteristiche materiali del manoscritto ferrarese (dimensione delle carte, distinzione fra recto o verso delle carte, modalità di fascicolazione e legatura, qualità dei supporti e degli altri strumenti di scrittura utilizzati), non potendo egli effettuare l'esame autoptico dell'originale cartaceo; la Biblioteca, infatti, ha imposto limitazioni stringenti all'accesso dei frammenti autografi.¹⁶⁸ Questa criticità della riproduzione digitale, a cui però – è bene precisare – fanno da contrappeso numerosi altri vantaggi, può essere parzialmente risolta attingendo alle descrizioni codicologiche presenti nelle pubblicazioni e nei repertori bibliografici, cartacei o digitali, già menzionati nel primo capitolo. D'altro canto, la ricerca di tali informazioni richiede al lettore un ulteriore sforzo intellettuale che si è cercato di minimizzare pubblicando la descrizione codicologica del manoscritto ferrarese a corredo dell'EDS. Tuttavia, per quanto utile, questa soluzione non risolve completamente il problema; sarebbe auspicabile, invece, che in futuro l'interfaccia grafica dell'edizione fosse ottimizzata, così da valorizzare pienamente la struttura del manoscritto – adottando, ad esempio, la doppia pagina come unità di rappresentazione, anziché la singola. Ad esempio, in questa direzione si colloca l'integrazione, in *EVT 3*, di *VisCol*, un software concepito per la visualizzazione e l'analisi delle relazioni tra varianti testuali e testimoni, che consente di rappresentare in modo più dinamico e interattivo la complessità dei dati codicologici e testuali. Tale sinergia potrebbe offrire nuove possibilità per una fruizione più consapevole e materialmente orientata dell'edizione digitale (cfr. § 4.2.3).

Il file fornito dalla Ariostea consta di 107 immagini di cui le prime due devono essere escluse dal computo degli autografi poiché contengono rispettivamente: una copia in miniatura del ritratto di Ludovico Ariosto attribuito a Dosso Dossi, sulla quale è stato impresso il timbro della Biblioteca; una scrittura autografa del direttore della Biblioteca, Giuseppe Agnelli, autore di una edizione fototipica dei frammenti autografi nonché di due importantissimi repertori bibliografici.¹⁶⁹ Le immagini digitali sono in Ultra HD (2160 dpi) e consentono, non solo, di apprezzare con maggiore nitidezza alcune caratteristiche materiali dei documenti come la grana della carta o le filigrane, ma anche, di leggere la scrittura di Ariosto – specialmente nelle carte

¹⁶⁸ Per approfondire, cfr. A. Campagnolo, *Codificare il codice. La materialità del libro e la sua digitalizzazione verso una codicologia digitale*, «Umanistica digitale», 17, 2024.

¹⁶⁹ La scrittura autografa è la seguente: «Carte cinquantatre autografe di Ludovico Ariosto, G. Agnelli, libro». La voce *Agnelli, Giuseppe* del *Dizionario biografico degli italiani - (Volume 1, 1960)* può essere consultata al seguente link: <[https://www.treccani.it/enciclopedia/giuseppe-agnelli_\(Dizionario-Biografico\)/](https://www.treccani.it/enciclopedia/giuseppe-agnelli_(Dizionario-Biografico)/)> (ultimo accesso 17 ottobre 2025). G. Agnelli ha curato la pubblicazione dei seguenti volumi: *I frammenti autografi dell'Orlando furioso di Ludovico Ariosto* (Roma, Fototopia Danesi, 1904); *Inventari dei manoscritti delle biblioteche d'Italia* (Firenze, Leo S. Olschki, 1933) e *Annali delle edizioni ariostee* (Bologna, Zanichelli, 1935), in collaborazione con Giuseppe Ravagnani.

più tormentate, cariche di correzioni e aggiunte – con un grado di dettaglio inimmaginabile in passato e tale da agevolare il lavoro del filologo, anche in funzione di nuove ipotesi ricostruttive sull'iter elaborativo del testo. Inoltre, poiché le immagini digitali sono state salvate solamente con il numero progressivo della presa fotografica, si è reso necessario, in prima battuta, abbinare il numero della presa fotografica al numero della carta con il rispettivo recto o verso, per poi procedere alla suddivisione delle stesse in fascicoli secondo l'ordinamento stabilito da S. Debenedetti nell'edizione del 1937.

Per quanto riguarda, invece, gli autografi (ms. S.P.33; 2; 3) e l'apografo (ms. H 55 inf.; 52) custoditi dalla Ambrosiana di Milano, abbiamo potuto visionare direttamente solo i primi, dopo aver ottenuto l'autorizzazione della direzione. Poiché le carte non sono ancora state riprodotte e pubblicate nella biblioteca digitale della Ambrosiana, è stato necessario richiedere la digitalizzazione ex novo, avvalendosi del servizio che la Biblioteca stessa mette a disposizione degli utenti sul suo sito web.¹⁷⁰ Questo servizio, gestito dalla società concessionaria Mondadori Portfolio, presenta però una grossa limitazione dal punto di vista della risoluzione delle immagini, poiché l'utente può scegliere solo fra due parametri preimpostati, ovvero fra 300 dpi o 600 dpi. Di conseguenza, queste immagini hanno una risoluzione nettamente inferiore rispetto a quella raccomandata dalle *Linee guida* del PND, ma fortunatamente riproducono copie in pulito, pronte per andare in tipografia, e pertanto facilmente leggibili. In ultima analisi, quindi, abbiamo fatto richiesta per 16 immagini a colori a 600 dpi (di cui 4 immagini riproducono le due carte autografe, mentre le altre 12 riproducono le sei carte dell'apografo) per finalità di studio che non precludono la possibilità di pubblicare online le stesse immagini, purché ciò avvenga senza scopo di lucro ai sensi dell'art. 108, comma 3-bis del *Codice*.¹⁷¹ Purtroppo, fino ad oggi non è stato possibile acquisire le immagini dell'apografo (ms. H 55 inf.; 52), a causa

¹⁷⁰ Le immagini dei fondi dell'Ambrosiana, possono essere richieste accedendo al seguente link: <<https://ambrosiana.it/scopri/riciesta-immagini/>>. Il costo delle riproduzioni è parametrato sulla risoluzione delle immagini (300 dpi o 600 dpi) e sulle finalità d'uso (studio o uso editoriale-commerciale). Il costo complessivo ammonta a 231, 80 euro.

¹⁷¹ La Biblioteca Ambrosiana è una istituzione ecclesiastica gestita dal Collegio dei Dottori composto da ecclesiastici scelti dall'Arcivescovo di Milano. Per evitare qualsiasi dubbio di carattere giuridico, è utile ricordare che l'art. 10, comma 1 del *Codice dei beni culturali* specifica: «Sono beni culturali le cose immobili e mobili appartenenti allo Stato, alle regioni, agli altri enti pubblici territoriali, nonché ad ogni altro ente ed istituto pubblico e a persone giuridiche private senza fine di lucro, ivi compresi gli enti ecclesiastici civilmente riconosciuti, che presentano interesse artistico, storico, archeologico o etnoantropologico». Di conseguenza, gli enti ecclesiastici sono tenuti a rispettare le disposizioni del *Codice* in relazione ai beni di loro proprietà. Ciò implica l'obbligo di richiedere le autorizzazioni necessarie per interventi di restauro o trasformazione, di osservare le norme sulla conservazione e sulla sicurezza, e di collaborare con le autorità competenti per garantire la tutela del patrimonio culturale.

dell'esaurimento dei fondi per la ricerca. Anche nel caso delle tre carte (ms. San Martino 353) conservate presso la Biblioteca Nazionale di Napoli, è stato necessario richiedere la riproduzione ex novo per un totale di sei immagini (4 immagini relative ai frammenti autografi dell'*Orlando furioso* e 2 immagini relative al frammento noto come *Scudo della Regina Elisa*). Tuttavia, la nostra richiesta non ha ottenuto alcun riscontro da parte della Biblioteca.

Le riflessioni sulla digitalizzazione degli autografi della Ariostea si estendono ovviamente anche alle digitalizzazioni degli autografi della Ambrosiana e della Nazionale di Napoli. Dopo l'acquisizione delle immagini digitali, il passaggio successivo è stato quello di convertire il testo contenuto nei documenti in dati leggibili dal computer (*machine-readable data*) utilizzando un formato standard come l'XML, allo scopo di consentirne l'estrazione e l'elaborazione. Sebbene questa operazione potesse essere eseguita in modo relativamente semplice e veloce attraverso la digitazione manuale del testo, si è deciso comunque di sperimentare la trascrizione automatica del testo. In effetti, le dimensioni del corpus testuale (56 carte, di cui una bianca, scritte su entrambe le facciate) non sono tali da rendere necessario il ricorso a strumenti di trascrizione automatica, che risultano invece indispensabili nel caso in cui si debbano gestire corpus testuali estremamente vasti.¹⁷² La scelta di ricorrere comunque alla trascrizione automatica è legata a motivazioni di più ampio respiro. Da un lato, essa ha rappresentato un'occasione per sperimentare una tecnologia oggi in forte espansione nel campo delle *digital humanities*; dall'altro, ha offerto l'opportunità di sviluppare un modello di *Handwritten Text Recognition* (HTR) specifico per la grafia di Ariosto, che potrebbe rivelarsi utile in due direzioni:

- per future ricerche in cui dovesse essere necessaria la trascrizione automatica di altri autografi dello stesso autore;
- per migliorare le prestazioni degli attuali modelli di HTR dedicati alla scrittura corsiva dell'italiano rinascimentale.

Per la trascrizione automatica di testi da immagini digitali, oggi i ricercatori possono contare principalmente su due piattaforme *open access*, *Transkribus* (READ-COOP) ed *eScriptorium*.¹⁷³

¹⁷² Per approfondire, cfr. S. Spina, *Artificial Intelligence in archival and historical scholarship workflow: HTR and ChatGPT*, «Umanistica Digitale», 16, 2023.

¹⁷³ Il software *eScriptorium* (disponibile al seguente link: <<https://escriptorium.inria.fr/>>) è stato sviluppato presso l'École Pratique des Hautes Études, Université Paris Sciences et Lettres (EPHE – PSL), con l'obiettivo di fornire agli studiosi delle discipline umanistiche, che non hanno familiarità con la programmazione informatica, un'interfaccia semplice per la trascrizione automatica di fonti scritte, stampate o manoscritte, in qualsiasi sistema di scrittura,

Entrambe sono molto diffuse in ambito accademico e archivistico poiché offrono GUI (*Graphic User Interface*) intuitive e non richiedono agli utenti competenze informatiche avanzate. Sia *Transkribus* che *eScriptorium* supportano l'intero flusso di lavoro che abbraccia l'importazione delle immagini, la trascrizione del testo (manuale o automatica) e la creazione di modelli di HTR personalizzati sulla base dei dati forniti dagli utenti.¹⁷⁴

La differenza principale tra le due piattaforme riguarda le possibilità d'utilizzo e le politiche di distribuzione. *eScriptorium* è un software completamente *open source* che si appoggia all'infrastruttura di *Kraken* e consente agli utenti di esportare in locale i modelli di HTR prodotti.¹⁷⁵ Tuttavia, trattandosi appunto di un software e non di un servizio, *eScriptorium* non è accessibile da un unico sito web, ma dev'essere installato localmente sul proprio hardware personale o sui server delle istituzioni che decidono di metterlo a disposizione degli utenti esterni. Se da un lato questo approccio garantisce un uso libero e sostenibile, dall'altro implica alcune limitazioni tecniche; infatti, mentre nel primo caso l'utente deve possedere le competenze informatiche necessarie per effettuare l'installazione e disporre di un dispositivo capace di sostenere un'elevata potenza di calcolo, nel secondo caso egli potrebbe trovarsi a dover fronteggiare problemi specifici del server istituzionale su cui si appoggia (come aggiornamenti o

antico o moderno. *eScriptorium* è il fulcro di un progetto più ampio di VRE (*Virtual Research Environment*) che punta ad abbracciare tutte le fasi dell'edizione digitale: dalla trascrizione, alla codifica (*markup*), all'analisi linguistica (ad esempio, la *Named Entity Recognition* o il *Natural Language Processing*), fino alla pubblicazione online di testi e immagini. Tuttavia, finora il progetto si è focalizzato soprattutto sul primo passaggio, la trascrizione (manuale o automatica), offrendo all'utente la possibilità di esportare i dati prodotti per continuare il lavoro editoriale con altri strumenti. Lo sviluppo del software è iniziato nell'ambito del progetto *Scripta*, dedicato allo studio della storia della scrittura in quasi tutte le sue forme, ed è proseguito nell'ambito del progetto *Resilience*, un'iniziativa europea tesa a costruire un'infrastruttura di ricerca a lungo termine nel campo degli studi religiosi. Per approfondire, cfr. P. Stokes et al., *The eScriptorium VRE for Manuscript Cultures*, «Classics@ Journal», 18 (1), 2021.

¹⁷⁴ Per approfondire, cfr. D. Dassonneville et al., *Compte-rendu de la journée d'étude 'Point HTR 2022' Transkribus/ eScriptorium: transcrire, annoter et éditer numériquement des documents d'archives* (Rapport de recherche, 9 juin 2022); P. Stokes e B. Kiessling, *Sharing Data for Handwritten Text Recognition (HTR)*, «Digital Humanities in Practice», 2024, consultabile al seguente link: <<https://hal.science/hal-04444641v1/document>> (ultimo accesso 17 ottobre 2025).

¹⁷⁵ Il software *eScriptorium* è stato messo a punto per interfacciarsi con il motore di HTR *Kraken*, sviluppato da Benjamin Kiessling (anch'esso membro dell'EPHE – PSL e parte del team di *eScriptorium*). Scritto in *Python*, *Kraken* è modulare e personalizzabile: ogni modulo dispone di numerosi parametri configurabili dall'utente, per adattarsi alle esigenze specifiche di ciascun progetto; inoltre, se i parametri disponibili non sono sufficienti, l'utente può sostituire qualsiasi modulo del motore con uno personalizzato. Grazie alla sua flessibilità, *Kraken* può essere utilizzato con un'ampia gamma di sistemi di scrittura. Tuttavia, richiede competenze informatiche avanzate, come la capacità di installare moduli e utilizzare la riga di comando. Per questo motivo, il motore non risulta particolarmente adatto alla maggior parte degli studiosi umanisti, ai quali è destinata l'interfaccia *user-friendly eScriptorium*. Per approfondire, cfr. P. Stokes et al., *The eScriptorium VRE for Manuscript Cultures*, «Classics@ Journal», 18 (1), 2021.

comandi bloccati), spesso senza poter contare su un’assistenza tecnica costante, essendo essa legata alla disponibilità del gestore del server. *Transkribus*, al contrario, è una piattaforma *cloud* centralizzata, ospitata presso i server dell’Università di Innsbruck. Questo approccio non richiede l’installazione locale e di conseguenza neanche particolari competenze informatiche o dotazioni tecnologiche, di modo che la piattaforma risulta più semplice da usare e immediatamente accessibile. Gli unici limiti riscontrabili riguardano l’acquisto dei cosiddetti “crediti” per l’uso delle funzioni di riconoscimento automatico e il fatto che i modelli di HTR generati dagli utenti non possono essere esportati, ma restano utilizzabili solo all’interno dello spazio *cloud* della piattaforma (cfr. Tab. 3.1).

Tabella 3.1 - Caratteristiche delle piattaforme *eScriptorium* e *Transkribus* a confronto

	<i>eScriptorium</i>	<i>Transkribus</i>
Politica di distribuzione	Completamente <i>open source</i>	Parzialmente <i>open source</i>
Motore di HTR	Kraken	PyLaia
Esportazione dei modelli di HTR	Consentita in locale	Non consentita (modelli disponibili solo nel <i>cloud</i> della piattaforma)
Accesso e installazione del software	Richiede installazione locale	Accesso immediato sul web, senza installazione (servizio <i>cloud</i> centralizzato)
Competenze informatiche richieste	Basilari (installazione del software)	Minime (registrazione sulla piattaforma)
Problemi tecnici	Possibili su server esterni, aggiornamenti non centralizzati	Gestione centralizzata, meno problemi tecnici
Diffusione	Più limitata	Ampia diffusione, numerosi utenti e istituzioni partner
Prestazioni del motore di HTR	Buone, ma inferiori alla media	Superiori grazie al continuo uso collettivo (<i>machine learning</i>)
Costi	Gratuito e sostenibile	<i>Freemium</i> (crediti gratuiti mensili limitati,

		poi a pagamento)
Punti di forza	Libertà d'uso, personalizzazione, esportazione dei modelli	Accessibilità, semplicità d'uso, alte prestazioni di trascrizione

Negli ultimi anni, questa differenza ha alimentato un acceso dibattito all'interno della comunità accademica internazionale, con numerose critiche mosse dai sostenitori dell'*open access* all'approccio parzialmente commerciale di *Transkribus*. Ciò nondimeno, bisogna riconoscere che la semplicità di utilizzo della GUI, unitamente alla disponibilità di un'infrastruttura *cloud* centralizzata pronta all'uso, ha favorito la rapida diffusione di *Transkribus*, soprattutto a partire dal 2018, con un aumento crescente degli utenti registrati e delle istituzioni partner.¹⁷⁶ Secondo Melissa Terras, «dall'ottobre del 2024, READ-COOP [la società cooperativa europea che gestisce la piattaforma] conta 227 istituti membri provenienti da 30 paesi e ha consolidato una solida base di utenti che supera i 235.000 registrati. La piattaforma *Transkribus* è stata usata per trascrivere circa 90 milioni di immagini digitali di testi storici, il che evidenzia che l'IA può essere impiegata efficacemente nell'ambito del patrimonio culturale, come attesta il conferimento dell'*Horizon Impact Award 2020* dell'Unione Europea».¹⁷⁷ Di conseguenza, dato che i sistemi di *machine learning* alla base di queste piattaforme progrediscono con la quantità di dati processati, la crescita della comunità di *Transkribus* ha condotto a prestazioni superiori (specialmente in termini di segmentazione e accuratezza della trascrizione) rispetto a quelle attualmente offerte da *eScriptorium*. In conclusione, le riflessioni fin qui esposte ci hanno portato ad adottare la

¹⁷⁶ Melissa Terras, membro del consiglio di amministrazione di READ-COOP, afferma che: «Il numero degli utenti registrati su *Transkribus* è aumentato vertiginosamente negli ultimi anni: da 2.200 nel 2015, 4.800 nel 2016, 8.500 nel 2017, 17.000 nel 2018, 30.000 nel 2019, 45.000 nel 2020 (i report di utilizzo indicano tra gli 800 e i 1400 utenti attivi a settimana). Nel febbraio 2021 la piattaforma ha raggiunto i 50.000 utenti, tra cui i principali istituti di conservazione internazionali: il Rijksmuseum, il Rahvusarhiiv (Archivio nazionale dell'Estonia), l'Arkivverket (Archivio nazionale della Norvegia), il Kansallisarkisto (Archivio nazionale della Finlandia) e la British Library. Tra le università che utilizzano *Transkribus* figurano l'Università di Cambridge, l'Università del Québec a Rimouski e l'Università Politecnica di Valencia. Il report di fine anno ha mostrato che nel 2020 c'erano 18.6264 utenti unici attivi che hanno caricato 19,9 milioni di immagini in 660.077 *jobs*, con 8443 nuovi modelli generati e 490 milioni di parole trascritte».

Per approfondire, cfr. M. Terras, *Chapter 7: Inviting AI into the Archives: The Reception of Handwritten Recognition Technology into Historical Manuscript Transcription*, in *Archives, Access and Artificial Intelligence. Working with Born-Digital and Digitized Archival Collections*, a cura di L. Jaillant, Bielefeld University Press, 2022, pp. 185-186.

¹⁷⁷ Per approfondire, cfr. M. Terras et al., *The artificial intelligence cooperative: READ-COOP, Transkribus, and the benefits of shared community infrastructure for automated text recognition*, «Open Research Europe», 2025.

piattaforma *Transkribus*, le cui caratteristiche tecniche e gestionali saranno analizzate in dettaglio nel paragrafo successivo.

3.2 La piattaforma *Transkribus* come caso esemplare di *Virtual Research Environment*

*3.2.1 Genesi, modelli di business e di governance della piattaforma*¹⁷⁸

Lo sviluppo della piattaforma *Transkribus* ha avuto inizio nel 2013, grazie ai finanziamenti del *Settimo Programma Quadro per la ricerca e lo sviluppo tecnologico* (7PQ) dell'Unione Europea, nell'ambito del progetto *tranScriptorium* (2013-2015).¹⁷⁹ L'obiettivo del progetto *tranScriptorium* era quello di mettere a punto una tecnologia di HTR innovativa, basata sull'applicazione delle reti neurali profonde (DNN), che è stata rilasciata nel 2014 mediante una *client application* gratuita (cfr. § 3.2.2) e successivamente rinominata *Transkribus*. In seguito, grazie ai finanziamenti del *Programma di ricerca e innovazione Horizon 2020* dell'Unione Europea, la piattaforma è stata sottoposta ad ulteriori aggiornamenti tecnici, nell'ambito del progetto interdisciplinare *Recognition and Enrichment of Archival Documents* (READ). Il progetto READ (2016-2019) è nato con lo scopo precipuo di implementare un'infrastruttura di ricerca online (*Virtual Research Environment*), concepita per supportare gli utenti durante l'intero flusso di lavoro che conduce all'addestramento di modelli di HTR a partire dalla creazione di dati *ground truth*.¹⁸⁰ Sotto la guida di Günter Mühlberger dell'Università di

¹⁷⁸ Per approfondire, cfr. J. Nockels et al., *Understanding the application of handwritten text recognition technology in heritage contexts: a systematic review of Transkribus in published research*, «Archival Science», 22 (3), 2022, pp. 367–392.

¹⁷⁹ Per approfondire, cfr. J. A. Sánchez et al., *Handwritten Text Recognition for Historical Documents in the tranScriptorium Project*, in *DATECH'14 Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage* (Madrid, 19-29 maggio 2014), New York, Association for Computing Machinery, 2014.

¹⁸⁰ Il termine “ground truth” è comunemente usato nell'ambito dell'apprendimento automatico per indicare dati oggettivi e accurati che sono raccolti mediante processi empirici, in opposizione ai dati che sono dedotti dalle fonti attraverso il calcolo statistico delle probabilità. Nel caso specifico di *Transkribus*, i dati *ground truth* sono raccolti trascrivendo un numero sufficiente di caratteri (di solito circa 15.000 parole o 75 pagine) relativi a un singolo *corpus* testuale; in seguito, questi dati possono essere usati per addestrare un modello di HTR applicabile a grandi volumi di dati dello stesso *corpus*. Per approfondire, cfr. REED-COOP, *What is Ground Truth?*, «Transkribus blog», 9 maggio 2023, consultabile al seguente link: <<https://blog.transkribus.org/en/what-is-ground-truth>> (ultimo accesso 17 ottobre 2025).

Innsbruck, il progetto READ ha riunito un ampio gruppo di esperti informatici e studiosi di discipline umanistiche provenienti da alcune tra le più prestigiose università a livello mondiale, avvalendosi anche della collaborazione dei principali istituti di conservazione internazionali.¹⁸¹

I servizi offerti da *Transkribus* sono destinati principalmente a quattro categorie di utenti: archivisti e bibliotecari, studiosi delle discipline umanistiche, esperti informatici e utenti generici, accomunati dall'interesse per lo studio, la conservazione e la valorizzazione dei documenti storico-letterari. Pur perseguendo obiettivi diversi, ciascuna categoria contribuisce in modo essenziale allo sviluppo dell'infrastruttura di ricerca e, al contempo, beneficia concretamente delle opportunità offerte dal suo utilizzo. La tabella seguente illustra in dettaglio le risorse intellettuali, bibliografiche, tecnologiche ed economiche che questi *stakeholder* possono condividere tra loro per soddisfare le rispettive esigenze (cfr. Tab. 3.2).¹⁸²

Tabella 3.2 - Gli *stakeholder* della piattaforma *Transkribus*

Stakeholder	Risorse	Esigenze
-------------	---------	----------

¹⁸¹ Tra le istituzioni che hanno partecipato al progetto READ, figurano: l'Università di Innsbruck (coordinatore/Austria), l'Università Politecnica di Valencia (Spagna), lo University College London (Regno Unito), il Centro Nazionale per la Ricerca Scientifica "Demokritos" (Grecia), l'Università Democrito di Tracia (Grecia), il CoSector - University of London (Regno Unito), l'Università tecnica di Vienna (Austria), l'Università di Rostock (Germania), l'Università di Lipsia (Germania), il NAFER LABS Europe (Francia), la scuola Politecnica Federale di Losanna (Svizzera), l'Archivio nazionale della Finlandia (Finlandia), l'Archivio di Stato di Zurigo (Svizzera), l'Archivio diocesano di Passau (Germania) e, più recentemente, Università di Edimburgo (Regno Unito). Per approfondire, cfr. Günter Mühlberger et al., *Transforming scholarship in the archives through Handwritten Text Recognition. Transkribus as a case study*, «Journal of Documentation», 75(5), 2019, pp. 954-976, consultabile al seguente link: <<https://doi.org/10.1108/JD-07-2018-0114>> (ultimo accesso 17 ottobre 2025); REED-COOP, *A Short History of Transkribus with Günter Mühlberger*, «Transkribus blog», 21 febbraio 2023 consultabile al seguente link: <<https://blog.transkribus.org/en/a-short-history-of-transkribus-with-gunter-muhlberger>> (ultimo accesso 17 ottobre 2025).

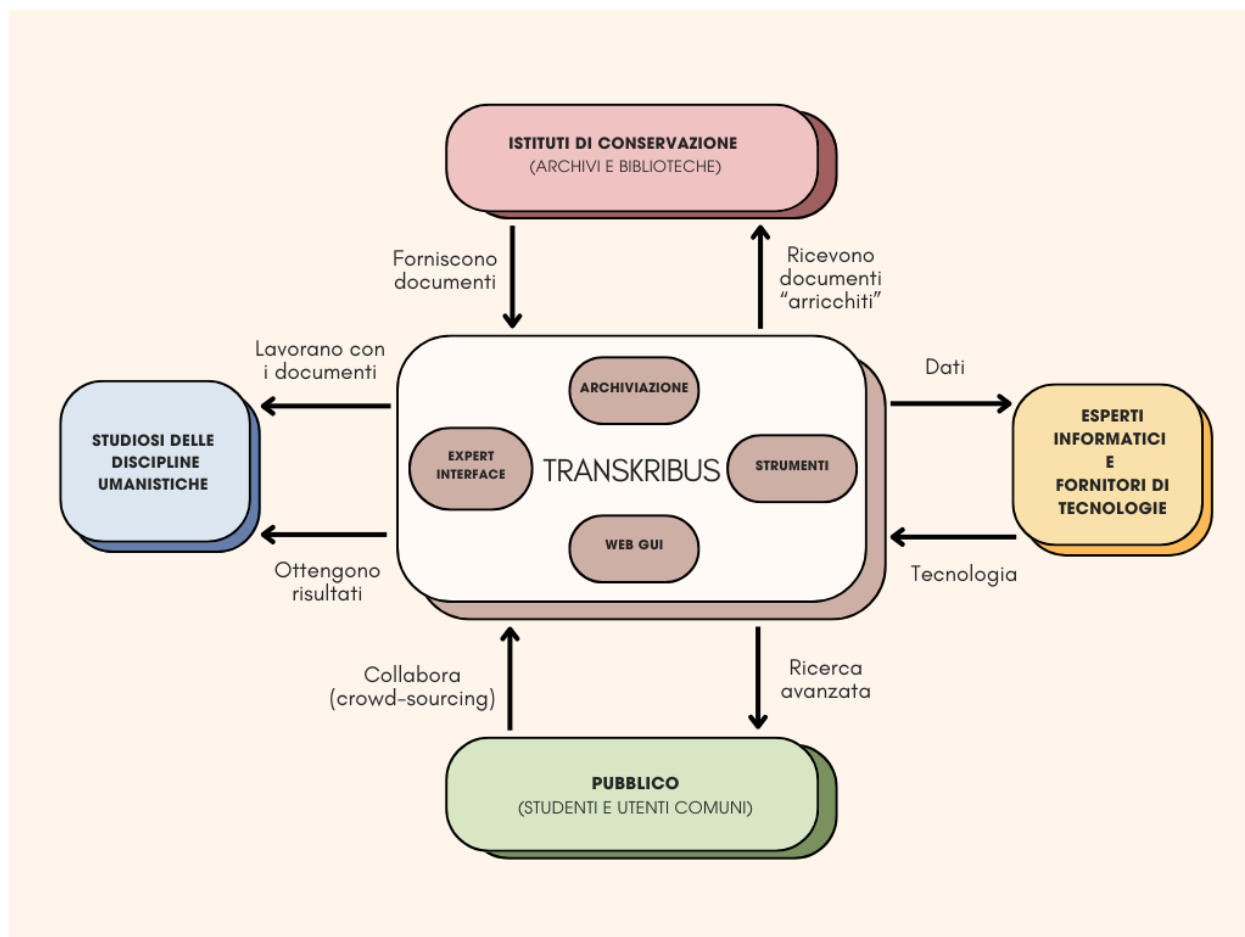
¹⁸² La Tabella 2 è un adattamento di quella riportata nel seguente articolo: P. Kahle et al., *Transkribus. A Platform for Automated Text Recognition and Searching of Historical Documents*, in *2019 15th IAPR International Conference on eScience* (San Diego, 24-27 settembre 2019), IEEE Publisher, pp. 463-466, consultabile al seguente link: <<https://ieeexplore.ieee.org/document/904176>> (ultimo accesso 17 ottobre 2025).

Archivi e Biblioteche	• Collezioni documentarie uniche; • Conoscenze settoriali; • Budget e finanziamenti regolari.	• Tecnologie per offrire servizi attrattivi agli utenti finali e agli studiosi.
Studiosi delle discipline umanistiche	• Conoscenze specialistiche; • Finanziamenti per progetti; • Rapporti con gli studenti da coinvolgere in iniziative di crowdsourcing.	• Accesso a documenti digitalizzati; • Documenti trascritti automaticamente e interrogabili.
Informatici	• Ricerca e tecnologia; • Innovazione; • Accesso a finanziamenti; • Rapporti con gli studenti da coinvolgere in iniziative di crowdsourcing.	• Dati ground truth; • Sfide di ricerca stimolanti.
Utenti comuni	• Disponibilità a contribuire con risorse proprie; • Flessibilità.	• Maggiore accesso ai documenti storico-letterari.

Dunque, mentre gli istituti di conservazione, gli umanisti e gli utenti comuni possono fornire immagini digitali e trascrizioni da usare come dati *ground truth* per l'addestramento di modelli di HTR, gli informatici svolgono le attività di ricerca necessarie per il mantenimento e l'innovazione tecnologica del sistema. Contestualmente, gli archivi e le biblioteche possono disporre della tecnologia necessaria per agevolare la consultazione delle proprie collezioni; gli umanisti possono condurre le proprie ricerche in modo più rapido ed efficace; gli utenti comuni possono avere maggiore accesso ai documenti (ad esempio, per approfondire la propria genealogia) e gli informatici possono usufruire di vaste banche dati, sotto forma di immagini e trascrizioni, per condurre i propri studi sulla HTR. Il rapporto di collaborazione sinergica che viene ad instaurarsi fra gli *stakeholder* costituisce il fondamento del modello di business su cui poggia la piattaforma *Transkribus* (cfr. Fig. 3.1).¹⁸³

¹⁸³ La *Figura 1* è un adattamento di quella riportata nel seguente articolo: Günter Mühlberger et al., *Transforming scholarship in the archives through Handwritten Text Recognition. Transkribus as a case study*, «Journal of Documentation», 75(5), 2019, pp. 954-976, consultabile al seguente link: <<https://doi.org/10.1108/JD-07-2018-0114>> (ultimo accesso 17 ottobre 2025);

Figura 3.1 – La sinergia fra gli *stakeholder* della piattaforma *Transkribus*



I presupposti tecnici per l'effettiva realizzazione di questo modello di business possono essere riassunti nei seguenti punti:

- tutti i dati e i documenti caricati su *Transkribus* rimangono privati, il che consente di limitare le restrizioni imposte dal diritto d'autore e dalle normative sulla tutela della privacy;
- tutti i dati e i documenti sono archiviati centralmente nei server dell'Università di Innsbruck, andando a formare un unico grande database che può essere usato per tutti i processi di apprendimento automatico;

- gli utenti possono gestire in completa autonomia l'intero processo di digitalizzazione, anche senza disporre di particolari competenze tecniche o di una formazione specialistica;
- gli utenti possono addestrare modelli personalizzati per il riconoscimento del testo e per l'analisi del layout (LA), utilizzando tecnologie avanzate basate su reti neurali profonde (DNN);
- le collezioni di documenti e i modelli di HTR o di LA possono essere condivisi tra gli utenti, e successivamente riutilizzati per l'addestramento di modelli di maggiore complessità (c.d. "super model");
- la piattaforma mette a disposizione modelli pubblici pre-addestrati, che consentono agli utenti di avviare immediatamente il riconoscimento del testo, senza dover creare *ex novo* un proprio modello.

In virtù di questi aspetti, che saranno trattati diffusamente nel prosieguo del capitolo, gli utenti di *Transkribus* rivestono un duplice ruolo: essi non sono soltanto meri consumatori di servizi, ma anche produttori di dati e modelli, contribuendo attivamente allo sviluppo e alla condivisione della conoscenza all'interno della comunità.

Allo scadere del periodo di finanziamento dell'UE avvenuto nel 2020, *Transkribus* è passato da un modello di business gratuito ad un modello "freemium" basato sulla gestione dei cosiddetti "crediti", al fine di garantire la sostenibilità economica della piattaforma nel lungo termine. Mentre le funzionalità di base (caricamento dei documenti, analisi del layout e trascrizione manuale) sono rimaste gratuite, la funzionalità di l'addestramento di modelli ad hoc per il riconoscimento automatico del testo e del layout è diventata a pagamento, poiché richiede una capacità di calcolo maggiore a carico dei server della piattaforma (e quindi anche costi di gestione più elevati). Gli utenti che creano un nuovo account ricevono 50 crediti gratuiti al mese (sufficienti per elaborare circa 50 pagine), esauriti i quali possono acquistare crediti aggiuntivi per continuare ad usare le funzionalità di riconoscimento automatico. L'acquisto di crediti aggiuntivi dipende sia dall'entità delle collezioni e dal tipo di materiale (manoscritto o a stampa), sia dalle specifiche esigenze di ricerca degli utenti. I crediti gratuiti offrono agli utenti la possibilità di testare la tecnologia e verificarne l'affidabilità dei risultati prima di sottoscrivere un abbonamento. D'altra parte, per i singoli ricercatori che lavorano su piccole collezioni di documenti, i crediti gratuiti potrebbero essere sufficienti per addestrare un modello di HTR abbastanza accurato, ad esempio per la *Smart Search* (cfr. § 3.2.3). Inoltre, REED-COOP (la società cooperativa europea che gestisce la piattaforma) ha istituito un programma di borse di

studio, il *Transkribus Scholarship Programme* (TSP), per supportare gli studenti nello svolgimento della propria di tesi di laurea o dottorato. L'ottenimento di una borsa di studio *Transkribus* comporta l'attribuzione gratuita di un abbonamento *Scholar* e l'assegnazione di un pacchetto di crediti fino ad un massimo di 300 al mese; la durata della borsa è stabilita in un anno. Per concorrere all'assegnazione di una borsa, gli aspiranti ricercatori sono tenuti a presentare un progetto che intendono sviluppare avvalendosi dell'abbonamento e dei crediti conferiti.¹⁸⁴ Nel 2024, anche la sottoscritta ha presentato la propria candidatura per l'ottenimento di una borsa di studio *Transkribus*, sottoponendo all'attenzione di READ-COOP il progetto dottorale sull'edizione scientifica digitale dei frammenti autografi di Ludovico Ariosto. READ-COOP, riconoscendo il notevole interesse scientifico della proposta progettuale, ha accordato un abbonamento *Scholar* unitamente a un congruo quantitativo di crediti, corrispondenti a 1.200 pagine, da destinarsi alla trascrizione automatica o ad altre funzionalità della piattaforma, entro l'arco di un anno.¹⁸⁵

Concretamente, l'attuazione del modello di business *freemium* è subordinata alla costituzione della READ-COOP SCE, sopra menzionata. Infatti, il modello di governance alla base della piattaforma consiste in una Società Cooperativa Europea (SCE) a responsabilità limitata, che è stata istituita nel 2019 con l'intento di garantire la sopravvivenza e la crescita costante dell'infrastruttura di ricerca anche oltre il termine delle sovvenzioni pubbliche europee. Questa

¹⁸⁴ Per maggiori informazioni su crediti, abbonamenti e borse di studio è possibile collegarsi al seguente link: <<https://help.transkribus.org/subscriptions-credits>> (ultimo accesso 17 ottobre 2025).

¹⁸⁵ J. Nockels et al. evidenziano che «il successo dei modelli *freemium* può variare. La versione gratuita può consentire ai consumatori di provare il prodotto prima di prendere una decisione d'acquisto e successivamente aumentare la domanda della versione a pagamento, ma può anche cannibalizzare la domanda della versione a pagamento. Detto ciò, le strategie *freemium* sono viste come un modo per andare oltre le interazioni *una tantum* con i clienti, verso un coinvolgimento a lungo termine. [...] L'indagine condotta da Terras presso la comunità degli utenti (Terras, 2022) ha evidenziato come una parte di essi temesse che l'adozione di un modello *freemium* potesse compromettere i principi fondanti della piattaforma. Un terzo degli intervistati ha espresso preoccupazioni circa il potenziale rischio per la trasparenza, la responsabilità e la natura semi-aperta della licenza dello strumento. Analogamente, la nostra precedente analisi, basata sulla consultazione degli utenti istituzionali di *Transkribus* nel 2021, aveva già posto in luce come il passaggio a un modello a pagamento fosse percepito come elemento di divisione: il 46,87% (n = 15) degli utenti aveva convenuto che il sistema a crediti avesse inciso sul loro accesso. È opportuno sottolineare che il progetto READ è stato istituito grazie a finanziamenti pubblici europei (tuttavia, nonostante il suo riconosciuto valore, non ha potuto beneficiare di ulteriori fondi pubblici per la sostenibilità). In tale contesto, la costituzione della cooperativa non rappresentava un mero atto di altruismo, bensì una necessità impellente per garantire la continuità operativa dell'infrastruttura. Intendiamo dimostrare che il TSP (*Transkribus Scholarship Programme*) incarna la strategia di READ-COOP volta a conseguire un certo grado di sostenibilità finanziaria. Attraverso il sostegno agli utenti in condizioni precarie, il TSP permette a READ-COOP di attenuare parzialmente le preoccupazioni degli utenti riguardo all'equità di accesso agli strumenti, perseguendo al contempo una sostenibilità e una redditività durature mediante un maggiore coinvolgimento a lungo termine con *Transkribus*». Per approfondire, cfr. J. Nockels et al., *Are Digital Humanities platforms facilitating sufficient diversity in research? A study of the Transkribus Scholarship Programme*, «Digital Scholarship in the Humanities», 40, 2025.

forma giuridica consente a istituzioni pubbliche e a soggetti privati di collaborare in modo strutturato, mantenendo comunque la propria autonomia, partecipando al processo decisionale tramite meccanismi democratici (come l'assemblea generale), e svolgendo attività economiche comuni. La cooperativa, che nel 2024 contava 175 azionisti in oltre 30 paesi, è aperta a nuovi membri, purché essi accettino lo statuto e acquistino un numero minimo di quote. A differenza delle società a scopo di lucro, il valore delle quote rimane fisso (non generando quindi utili per gli azionisti), il che rende la cooperativa resistente a tentativi di acquisizione, ma meno interessante per chi cerca un ritorno economico. Tuttavia, proprio questo modello di governance mira a rafforzare il senso di appartenenza e il coinvolgimento dei membri, poiché chi entra a far parte della cooperativa è al tempo stesso utente e proprietario della piattaforma. I vantaggi per i membri derivano principalmente dalla possibilità di ottenere servizi migliori e più efficienti, grazie alla condivisione di risorse e dati. Ad esempio, un'istituzione potrebbe mettere a disposizione i propri server o risorse di calcolo (GPU), ottenendo in cambio una maggiore capacità di elaborazione per i propri progetti. In sintesi, il valore generato per i membri della cooperativa non è di tipo finanziario, ma risiede nella qualità e nell'accessibilità dei servizi offerti dalla piattaforma.¹⁸⁶

3.2.2 Evoluzione tecnologica della piattaforma: *TranskribusX*, *TranskribusLite* e *Transkribus web app* ¹⁸⁷

Dal punto di vista tecnico, *Transkribus* adotta un'architettura *client-server* classica, in cui diverse applicazioni client interagiscono con un server centrale che gestisce l'archiviazione dei file e l'esecuzione degli strumenti. La comunicazione tra client e server avviene tramite un'API RESTful, cioè un'interfaccia web che espone in un formato *machine-readable* tutte le funzionalità del server: autenticazione, gestione delle collezioni, analisi del layout,

¹⁸⁶ Per approfondire, cfr. P. Kahle et al., *Transkribus. A Platform for Automated Text Recognition and Searching of Historical Documents*, in *2019 15th IAPR International Conference on eScience* (San Diego, 24-27 settembre 2019), IEEE Publisher, pp. 463-466.

¹⁸⁷ Per approfondire, cfr. P. Kahle et al., *Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents*, in *2017 14th IAPR International Conference on Document Analysis and Recognition – ICDAR* (Kyoto, 09-15 novembre 2017), IEEE Publisher, pp. 19-24.

riconoscimento del testo ed esportazione dei documenti in vari formati.¹⁸⁸ Attualmente, la maggior parte del software *Transkribus* è disponibile in modalità *open source* su *GitHub*.

Inizialmente, la piattaforma offriva due diverse interfacce client:

- *TranskribusX*: un'applicazione desktop che poteva essere scaricata gratuitamente dalla homepage della piattaforma, registrandosi con un account personale; sviluppata dal gruppo di *Digitalizzazione e Archiviazione Elettronica* (DEA) dell'Università di Innsbruck in *Java 8*, mediante lo *Standard Widget Toolkit* (SWT) di *Eclipse*, l'applicazione richiedeva l'installazione locale ed era compatibile con tutti e tre i principali sistemi operativi, *Windows*, *macOS* e *Linux*;¹⁸⁹
- *TranskribusLite*: una versione web semplificata di *TranskribusX*, progettata per offrire un'interfaccia di trascrizione intuitiva agli utenti con poca o nessuna dimestichezza nella digitalizzazione di documenti; sviluppata in *JavaScript* con minime dipendenze, *TranskribusLite* poteva essere integrata in piattaforme web di terze parti, il che la rendeva particolarmente adatta a progetti di *crowdsourcing*.

Nel 2023, REED-COOP ha rilasciato la nuova web app di *Transkribus*, che combinando la semplicità d'utilizzo di *TranskribusLite* con le funzionalità avanzate di *TranskribusX*, punta ad offrire un'esperienza utente sempre più intuitiva, completa e soddisfacente. Il lancio della web app ha rappresentato la prima tappa di un percorso evolutivo che, nel segno di un approccio

¹⁸⁸ Un'interfaccia REST (*Representational State Transfer*) è un tipo di API (*Application Programming Interface*) che consente a diverse applicazioni software di comunicare tra loro attraverso il protocollo HTTP. Basata su un'architettura definita da Roy Fielding nel 2000, la REST utilizza metodi standard come GET, POST, PUT e DELETE per operare su risorse identificate da URI (Uniform Resource Identifier). Questo approccio favorisce una comunicazione semplice, scalabile e indipendente dallo stato tra client e server, rendendo la REST una scelta popolare per lo sviluppo di servizi web e applicazioni mobili.

¹⁸⁹ Prima del lancio della nuova web app, *TranskribusX* ha rappresentato il principale punto di accesso alla piattaforma, consentendo l'esecuzione di tutte le funzionalità più importanti, quali il caricamento di documenti, la gestione delle collezioni, la trascrizione e la marcatura dei testi, l'analisi automatica del layout, l'addestramento di modelli di HTR e l'esportazione dei documenti in diversi formati. Oggi *TranskribusX* è ancora accessibile, ma non sarà sottoposto ad ulteriori aggiornamenti. Per scaricare e installare *TranskribusX*, l'utente deve registrarsi sul sito web dedicato. Dopo l'attivazione dell'account, è possibile accedere e scaricare uno dei due archivi ZIP disponibili: uno contiene i file necessari per l'esecuzione su *Microsoft Windows* e *GNU/Linux*, mentre l'altro è specifico per *macOS*; prerequisito essenziale è l'installazione di *Java 8*. Successivamente, l'archivio scaricato va estratto in una cartella in cui l'utente abbia i permessi di scrittura. La cartella risultante conterrà i file eseguibili per avviare il programma su ciascuna piattaforma (*.sh* per *Linux*, *.exe* o *.bat* per *Windows*, e *.command* per *macOS*). Per utilizzare *TranskribusX* in modalità online, è necessario che *Java* possa stabilire una connessione all'indirizzo <<https://transkribus.eu>> e che eventuali *firewall* presenti nella rete dell'utente siano configurati di conseguenza. L'applicazione supporta l'utilizzo di un server *proxy*, la cui configurazione è possibile direttamente all'interno di *TranskribusX*.

olistico, è scaturito nella creazione di un ecosistema integrato di ricerca, completamente *web-based*. La web app è accessibile da qualsiasi dispositivo, fisso o mobile, e recentemente è stata integrata con la funzionalità *Sites* e con nuovi modelli addestrabili (*Field Model* e *Table Model*), concepiti per ottimizzare il riconoscimento di layout complessi come quelli di giornali, registri, moduli, ecc. La nuova GUI dà maggior rilievo alle funzionalità d'uso frequente mediante tre ambienti di lavoro dedicati; gli ambienti di lavoro sono integrati in una barra di navigazione fissa e risultano sempre visibili da ogni altra sezione della piattaforma, di modo che l'utente può muoversi tra di essi con facilità.

Il primo dei nuovi ambienti di lavoro è *Desk*. Discostandosi dalla frammentarietà e dalla complessità che caratterizzavano le GUI di *TranskribusLite* e di *TranskribusX*, la homepage della nuova web app è stata riorganizzata in modo tale da offrire un'esperienza di navigazione molto più semplice e lineare. La sezione *Home* fornisce un accesso immediato alle collezioni e ai documenti recenti, consentendo il caricamento di nuovi materiali e l'applicazione della funzione di riconoscimento rapido del testo. Nella sezione *Collections*, invece, è possibile esplorare le raccolte di documenti preesistenti e crearne di nuove; inoltre, si possono selezionare documenti interi o singole pagine a cui applicare la funzione di analisi automatica del layout (cfr. Fig. 3.2).

Il secondo ambiente di lavoro è *Models*. Infatti, poichè l'addestramento di modelli di HTR e di LA costituisce il fulcro dell'intera infrastruttura di ricerca *Transkribus*, a questa funzionalità è stato dedicato un ambiente di lavoro specifico all'interno della nuova web app. In *Models* è possibile gestire i propri modelli e addestrarne di nuovi personalizzati. La personalizzazione è fondamentale per calibrare i modelli sui diversi stili grafici e di impaginazione che caratterizzano i documenti storico-letterari. Dunque, per addestrare un modello personalizzato, l'utente deve accedere allo spazio di lavoro *Models*, cliccare sulla sezione *Training Lab* e selezionare la funzione *Training New Model*; si aprirà un menù a tendina in cui sono elencate le diverse tipologie di modelli attualmente disponibili per l'addestramento (*Text Recognition Model*, *Baselines Model*, *Field Model* e *Table Model*); dopo aver selezionato la tipologia desiderata, si aprirà una nuova finestra in cui sono illustrate le varie fasi del processo di addestramento (cfr. Fig. 3.3).

Infine, per rendere il patrimonio storico-letterario sempre più accessibile, la nuova web app si propone non solo di fornire strumenti per la transizione *user-friendly*, ma anche di ampliare la portata della diffusione dei documenti grazie all'introduzione di una specifica funzionalità. *Sites* semplifica la creazione di siti web, ponendosi al servizio sia degli studiosi desiderosi di

condividere il frutto delle proprie ricerche, sia delle istituzioni quali archivi o biblioteche, che aspirano a rendere le proprie collezioni fruibili dal grande pubblico. Per avviare la procedura, bisogna accedere alla sezione *Sites* e configurare il sito web selezionando la funzione *Create New Site* (cfr. Fig. 3.4).¹⁹⁰

Figura 3.2 – Ambiente di lavoro *Desk*

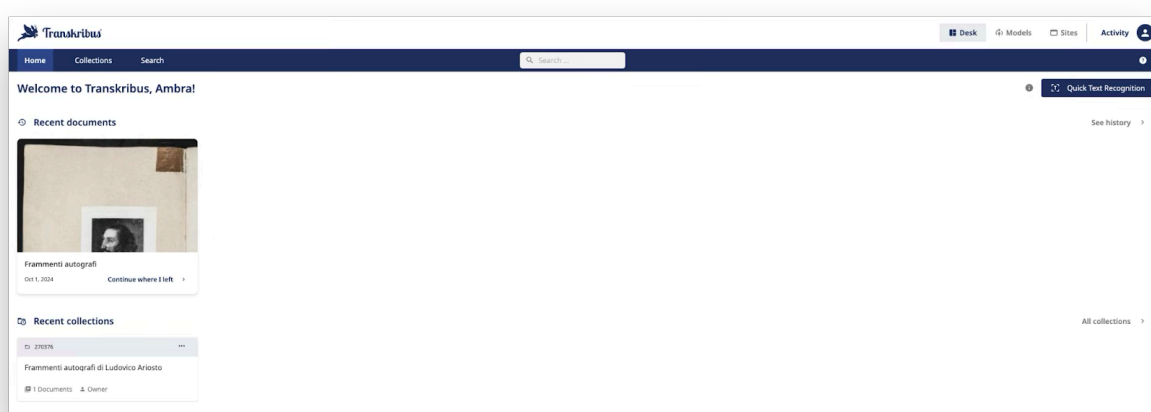


Figura 3.3 – Ambiente di lavoro *Models*

¹⁹⁰ Per maggiori informazioni sull'ambiente *Sites* e le relative modalità di configurazione, cfr. REED-COOP, *Transkribus Sites - Publishing documents made easy*, «Transkribus blog», 10 gennaio 2024, consultabile al seguente link: <<https://blog.transkribus.org/en/transkribus-sites-publishing-documents-made-easy>> (ultimo accesso 17 ottobre 2025); <<https://help.transkribus.org/sites>> (ultimo accesso 17 ottobre 2025).

MODEL TITLE	LANGUAGES	WORDS	CER
The Text Titan 1 bis	ENG, GER, DUT, LAT	0	0
The Text Titan 1	GER, DUT, FRE, FIN	0	2.95%
The German Giant 1	GER	15 420 976	8.30%
The Dutchess 1	DUT	11 693 499	4.30%
Transkribus Print M1	GER, ENG, DUT, FRE	5 068 310	2.20%
Transkribus French Model 1	FRE	1 933 011	7.80%
AIP Handwritten Portuguese 15th-20th Centuries	POR	186 342	2.53%
Dabbas 1706-1711	ARA	54 519	6.73%
Scottish Custom Books V0.8	ENG	116 603	10.42%
Dutch Dean	DUT	24 721 830	3.50%
Dansk Dokumentalist	DAN	12 949 906	3.70%
Middlebeurs Amsterdam	DUT	1 542 527	5.11%
Notaire de la Nouvelle France : Adhémar de St-Martin, Antoine (1668-1714)	FRE	54 513	6.15%
1700-kallens administrativt genisk håndskrift	DAN	893 399	3.27%
Notaire de la Nouvelle France : Ragueot de Beauvillage, François (1709-1753)	FRE	52 153	5.72%

Public Model
The Text Titan 1 bis

by Transkribus Mar 19

% Languages ENG, GER, DUT, LAT

Trained on handwritten

Model ID 309593

Show Description

Figura 3.4 – Ambiente di lavoro Sites

Site Title	Published
Rettssaka mot Ole Schare i 1820-åra	Published: 3/5/2025
Iurisprudencia	Published: 13/12/2024
Regierungsratsprotokolle des Kantons St.Gallen 1803-1811	Published: 14/9/2023
Grieg (test)	Published: 10/1/2025
Tryckfrihetsförordningen	Published: 11/11/2023
Genoese notarial contracts 1300-1550	Published: 14/9/2023
Kloster Aldersbach Digital	Published: 16/12/2024
Pommersches Volksliedarchiv	Published: 16/12/2024
Marjory Fleming (1803-1811)	Published: 20/11/2024
Grenzregelungsausschuss	Published: 20/11/2024
Riksarkivet	Published: 21/11/2024
Kucak Mecmua	Published: 21/10/2023

Nella tabella che segue, sono elencate in dettaglio le caratteristiche principali dei tre ambienti di lavoro sopra menzionati (cfr. Tab. 3.3).

Tabella 3.3 – Tabella sinottica delle caratteristiche dei nuovi ambienti di lavoro *Desk*, *Models* e *Sites*

Ambiente di lavoro	Novità	Descrizione
Desk	Barra di navigazione superiore	Migliore organizzazione del lavoro grazie ad una visione d'insieme più chiara.
	Modifica facilitata di nomi e metadati di documenti e collezioni	Possibilità di modificarli direttamente dalla collezione o dal documento.
	Codifica degli stati di elaborazione ottimizzata	Stati delle pagine (in progress, done, final, ground truth, ecc.) più evidenti mediante marcatori cromatici.
	Filtraggio per stato di elaborazione	Ricerca semplificata delle pagine in base al loro stato.
	Personalizzazione della visualizzazione delle anteprime	Regolazione delle dimensioni delle anteprime nella visualizzazione a griglia o visualizzazione a tabella.
	Riduzione delle attività superflue	Struttura più efficiente (ad es. modifica diretta dei nomi delle collezioni, ecc.).
Models	Dashboard modelli privati	Panoramica generale dei modelli dell'utente.
	Nuovi modelli addestrabili	Introduzione di <i>Field Model</i> e <i>Table model</i> per un riconoscimento del layout sempre più ottimizzato.
	Semplificazione del processo di addestramento	Nuovo design e riorganizzazione delle fasi di addestramento.
	Personalizzazione dei modelli	Creazione di modelli <i>ad hoc</i> con aggiunta di descrizioni e miniature.
	Galleria modelli ottimizzata	Introduzione di filtri di ricerca tra modelli pubblici e privati.
Sites	Creazione autonoma di siti web	Configurazione del sito senza bisogno di uno sviluppatore.
	Condivisione delle collezioni di documenti	Pubblicazione online delle proprie collezioni di documenti.

Personalizzazione grafica	Possibilità di scegliere tra diversi <i>template</i> e applicare i colori del proprio brand.
Modifiche flessibili	Possibilità di aggiornare le impostazioni in qualsiasi momento.

Questa rapida panoramica della nuova GUI di *Transkribus* consentirà al lettore di orientarsi meglio nel *workflow* sotteso all’addestramento dei modelli di HTR e di LA, che sarà descritto con dovizia di particolari nel capitolo successivo.¹⁹¹

3.2.3 Il workflow del riconoscimento del testo manoscritto (HTR) sulla piattaforma Transkribus

La piattaforma *Transkribus* si avvale delle innovazioni più recenti nel campo del riconoscimento del testo manoscritto (HTR), basate su architetture di reti neurali ricorrenti (RNN).¹⁹² Come si è

¹⁹¹ Per approfondire, cfr. REED-COOP, *Introducing the new Transkribus web app*, «Transkribus blog», 29 agosto 2023, consultabile al seguente link: <<https://blog.transkribus.org/en/introducing-the-new-transkribus-web-app>> (ultimo accesso 17 ottobre 2025).

¹⁹² Nel 2017, P. Kahle et al. affermavano: «Attualmente sono integrate nella piattaforma *Transkribus* tre tecnologie di riconoscimento del testo: i caratteri latini e gotici vengono riconosciuti utilizzando il riconoscimento ottico dei caratteri (*Optical Character Recognition* – OCR) basato su *ABBYY FineReader 11 SDK*.

Per quanto riguarda i manoscritti, la piattaforma include due motori di HTR *off-line*. Il primo è stato sviluppato presso il centro di ricerca *Pattern Recognition and Human Language Technology* (PRHLT) dell'Università di Valencia (UPVLC). I modelli ottici (*Optical Model* – OM) applicati in questo approccio sono essenzialmente *Hidden Markov Models* (HMM) e contengono informazioni sull'aspetto e le caratteristiche dei caratteri in una data scrittura. Un modello linguistico (*Language Model* – LM) aggiuntivo è rappresentato da un modello *n-grammi*, basato sulla probabilità di *n* parole consecutive. La fase di addestramento nella piattaforma *Transkribus* produce sempre un OM accoppiato ad un LM appropriato, il che implica un vocabolario fisso per questa coppia. L'output del riconoscimento sono reticoli che rappresentano le trascrizioni più probabili rispetto all'OM/LM e le loro probabilità. Ogni riga di testo è rappresentata da un tale grafo di parole nell'output, un grafo aciclico diretto pesato i cui archi sono etichettati con parole. Oltre a calcolare le *n* migliori trascrizioni, i grafi di parole consentono varie applicazioni avanzate, come il *Keyword Spotting* (KWS) e la trascrizione assistita da computer di immagini di testo (*Computer Assisted Transcription of Text Images* – CATTI).

Il secondo approccio incluso nella piattaforma è stato sviluppato dal CITlab e dalla PLANET intelligent systems GmbH. Si basa su reti neurali ricorrenti (*Recurrent Neural Network* – RNN) e attualmente non si affida a un modello linguistico. Pertanto, i risultati del riconoscimento non sono vincolati a un vocabolario specifico (ma rimane la restrizione all'alfabeto addestrato). L'output di questo sistema è una *matrice di confidenza* per ogni immagine di riga di testo, contenente un vettore per ogni carattere dell'alfabeto addestrato con probabilità di occorrenza per le posizioni ispezionate nella riga. Per decodificare questa matrice, cioè calcolare un risultato di riconoscimento per una riga, si possono applicare diversi schemi. Sebbene una descrizione dettagliata di questi esulerebbe dallo scopo di questo articolo, vale la pena notare che un dizionario *1-grammi* con frequenze di parole può facoltativamente essere coinvolto qui. Anche le applicazioni KWS trarranno vantaggio dalle matrici di confidenza memorizzate».

visto nel capitolo precedente, le reti neurali possono essere istruite a riconoscere specifici stili di scrittura mediante l'analisi di immagini digitalizzate, corredate dalle rispettive trascrizioni. L'esito di tale processo di apprendimento si sostanzia in un modello di HTR, ovvero un sistema computazionale avanzato, concepito per l'automazione della trascrizione di corpora di documenti storico-letterari. Fino al 2002, *Transkribus* ha supportato il motore di riconoscimento denominato *HTR+*, sviluppato dal CITlab dell'Università di Rostock e dal suo partner, la società tecnologica tedesca PLANET. Attualmente, il motore di riconoscimento in uso è *PyLaia*. Come il motore *Laia*, che si può considerare il suo diretto antecedente, *PyLaia* è stato sviluppato dall'Università Politecnica di Valencia ed è *open source*.¹⁹³ La tecnologia di HTR si configura come linguisticamente indipendente, in quanto il processo di addestramento della rete neurale mantiene la sua intrinseca validità a prescindere dalla natura dell'alfabeto o dal periodo storico di riferimento. Questa peculiarità consente lo sviluppo di modelli specializzati per qualsivoglia tipologia di scrittura, appartenente a qualsiasi epoca. La metodologia adottata si basa su un approccio orientato alle linee di testo, in cui l'immagine di una linea di base (*baseline*) – definita come la linea orizzontale sottostante una riga di testo in un'immagine digitalizzata – e il testo corrispondente, trascritto correttamente, costituiscono l'input fondamentale per gli algoritmi di apprendimento delle reti neurali.

Dunque, appurato che le immagini digitali e le relative trascrizioni rappresentano prerequisiti indispensabili per lavorare con la tecnologia di HTR, è necessario sottoporli a una fase preliminare di elaborazione all'interno della GUI di *Transkribus*, al fine di convertirli in dati *ground truth*. I dati *ground truth* possono essere usati per l'addestramento di un modello di HTR specificamente calibrato su una determinata collezione di documenti storico-letterari, sia essa

Per approfondire, cfr. P. Kahle et al., *Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents*, in *2017 14th IAPR International Conference on Document Analysis and Recognition – ICDAR* (Kyoto, 09-15 novembre 2017), IEEE Publisher, pp. 19-24.

¹⁹³ Per maggiori informazioni sui motori di riconoscimento, cfr. REED-COOP, *The Road to the Handwritten Text Recognition*, «Transkribus blog», 14 giugno 2022, consultabile al seguente link: <<https://blog.transkribus.org/en/insights/road-to-htr-1>> (ultimo accesso 17 ottobre 2025); REED-COOP, *Try out Transkribus new recognition software PyLaia!*, «Transkribus blog», 17 settembre 2020, consultabile al seguente link: <<https://blog.transkribus.org/en/try-out-transkribus-new-recognition-software-pylaia>> (ultimo accesso 17 ottobre 2025); Felix Dietrich, *Why it took 60+ years for AI to take off*, «Transkribus blog», 2021, consultabile al seguente link: <https://readcoop.eu/insights/why-took-60-years-for-ai-to-take-off/?__hstc=48127928.03087d00981f619c0c9c7235c214c5c7.1727456760735.1746192990896.1746200166908.26&__hssc=48127928.14.1746200166908&__hsfp=1778039452> (ultimo accesso 17 ottobre 2025).

riconducibile alla produzione di un singolo autore o di una pluralità di autori accomunati da stili grafici affini (cfr. Fig. 3.5).

La creazione della *ground truth* si articola in tre fasi principali:

1. caricamento delle immagini digitali sulla piattaforma;
2. uso degli strumenti di analisi del layout (LA) per la segmentazione delle immagini in regioni di testo, linee di testo e linee di base;
3. trascrizione accurata del contenuto di ciascuna riga di testo presente nelle immagini digitali.

Le immagini digitali possono essere acquisite attraverso una vasta gamma di strumenti, dai sofisticati scanner per libri ai più comuni smartphone. Per garantire agli utenti piena autonomia e superare la dipendenza da servizi di digitalizzazione esterni, REED-COOP ha creato l'app mobile *DocScan*, che abbinata al dispositivo portatile *ScanTent*, permette di scattare foto ad alta definizione dei documenti, editarle ed importarle direttamente sulla piattaforma *Transkribus*, il tutto sfruttando la comodità di un semplice smartphone.¹⁹⁴ A questo riguardo, vale la pena menzionare l'iniziativa intrapresa dalla Bibliothèque nationale de France, che alcuni anni fa ha installato più di quaranta *ScanTent* all'interno dei propri spazi, in modo da semplificare e velocizzare il servizio di scansione dei documenti.¹⁹⁵ In un'ottica analoga di innovazione e potenziamento dei servizi di digitalizzazione, anche in Italia la Direzione generale Archivi ha recentemente avviato un progetto nazionale volto all'adozione della piattaforma *Transkribus* presso l'intera rete degli Archivi di Stato.¹⁹⁶

¹⁹⁴ L'applicazione *DocScan* può essere scaricata gratuitamente dal *Google Play Store*. Inizialmente disponibile solo per il sistema operativo *Android*, grazie all'integrazione di una modalità automatica nell'app *Doc Scan* di ifunplay, l'applicazione è ora supportata anche dal sistema operativo *iOS*.

Per maggiori informazioni sull'app *DocScan* e il dispositivo *ScanTent*, cfr. i seguenti link: <<https://help.transkribus.org/docscan>> (ultimo accesso 17 ottobre 2025); <<https://help.transkribus.org/scantent>> (ultimo accesso 17 ottobre 2025); <<https://blog.transkribus.org/en/how-to-digitise-archival-materials-with-transkribus>> (ultimo accesso 17 ottobre 2025).

¹⁹⁵ Per approfondire l'iniziativa della BnF, cfr. il seguente link: <https://readcoop.eu/success-stories/scantents-at-the-bnf/?_hstc=48127928.03087d00981f619c0c9c7235c214c5c7.1727456760735.1747675390612.1747838254223.46&_hssc=48127928.33.1747838254223&_hsfp=225572891> (ultimo accesso 17 ottobre 2025).

¹⁹⁶ Per approfondire l'iniziativa della DGA, cfr. il seguente link: <<https://icar.cultura.gov.it/archivio-notizie/notizia/la-dga-adotta-transkribus-ia-al-servizio-della-ricerca-e-dellaccessibilita-documentaria>> (ultimo accesso 17 ottobre 2025).

La piattaforma *Transkribus* supporta un'ampia gamma di formati di immagine e dispone di una capacità di archiviazione lato server atta a gestire anche vaste collezioni di documenti. Le immagini caricate su *Transkribus* rimangono private per impostazione predefinita e non sono accessibili al pubblico, a meno che il proprietario della collezione non ritenga opportuno concedere ad altri utenti l'autorizzazione a visualizzare ed elaborare i propri documenti.¹⁹⁷ I dati per l'addestramento di un modello di HTR dovrebbero essere una rappresentazione fedele di tutti gli elementi caratterizzanti della collezione, riflettendo accuratamente tutte le tipologie di layout, le varietà lessicali e gli stili di scrittura presenti. Ragion per cui, l'utente deve valutare con molta attenzione quali e quante pagine includere nella *ground truth*; ad esempio, un metodo efficace potrebbe essere quello di selezionare le pagine ad intervalli regolari all'interno della collezione. In genere, un set di dati *ground truth* pari a 15.000 parole trascritte (equivalenti a circa 75 pagine) è sufficiente per l'addestramento di un modello di HTR destinato a testi vergati da una singola mano, mentre un set di dati inferiore, pari a 5.000 parole trascritte (all'incirca 25 pagine), può andare bene per il riconoscimento di testi a stampa. A questo proposito, bisogna ricordare ancora una volta che, in linea con i principi dell'apprendimento automatico, maggiore è la quantità di dati *ground truth* forniti dall'utente, maggiore sarà il livello di accuratezza (*consistency*) raggiunto dai risultati di trascrizione. Pertanto, qualora una collezione comprenda documenti redatti da differenti mani e/o in diverse lingue, gli utenti dovranno predisporre dati *ground truth* costituiti da un numero più elevato di parole trascritte.¹⁹⁸

¹⁹⁷ Nel manuale utente, si legge: «In *Transkribus*, il termine "documento" si riferisce a un insieme di immagini che costituiscono un elemento più grande (come un fascicolo, un manoscritto o un libro). I formati accettati per le immagini sono JPEG, PNG e TIFF, con una dimensione massima di 20MB per ciascuna. Quando si effettua l'*upload*, tutte le immagini selezionate contemporaneamente sono accorpate in un unico *documento*, dove ogni immagine corrisponde a una pagina. Se, invece, si carica un file PDF, ogni pagina del PDF viene estratta e caricata come una pagina del documento. I file PDF possono avere una dimensione massima di 512 MB». Per approfondire la procedura di *upload*, cfr. il seguente link: <<https://help.transkribus.org/uploading-files-to-transkribus>> (ultimo accesso 17 ottobre 2025).

¹⁹⁸ Nel manuale utente, si legge: «Nel caso di documenti manoscritti, si consiglia di addestrare il modello con almeno 10.000 parole per ogni grafia. Modelli addestrati su un'ampia mole di dati (oltre 100.000 parole), comprendente diverse grafie dello stesso periodo storico e della stessa area geografica, dovrebbero essere in grado di riconoscere grafie mai incontrate durante l'addestramento. Tuttavia, i risultati potrebbero essere leggermente meno precisi in termini di *Character Error Rate* (misurato sui dati di validazione). La *ground truth* dovrebbe includere esempi di tutte le scritture che si desidera che il modello sia in grado di trascrivere. È possibile addestrare modelli capaci di riconoscere contemporaneamente due o più grafie, lingue, tipi di scrittura o alfabeti; tuttavia, queste varianti devono essere rappresentate in modo adeguato nella *ground truth*. Le pagine da includere nella *ground truth* sono quindi importanti, poiché influenzeranno l'efficacia del modello. Ad esempio, se si desidera addestrare un modello che riconosca le grafie di tre diversi autori, sarà necessario trascrivere circa 10.000 parole per ciascun autore. Nel caso di un autore la cui grafia è cambiata nel tempo, la *ground truth* dovrebbe comprendere pagine scritte in diversi anni che siano rappresentative di tali cambiamenti».

Per approfondire la procedura di preparazione dei dati di addestramento, cfr. il seguente link: <<https://help.transkribus.org/data-preparation>> (ultimo accesso 17 ottobre 2025).

Una volta caricate sul server di *Transkribus*, le immagini possono essere sottoposte all'analisi del layout (LA), un processo di fondamentale importanza per l'automazione della trascrizione. Recenti progressi tecnologici hanno significativamente migliorato la precisione di questa fase, agevolando l'identificazione del testo anche in documenti caratterizzati da architetture complesse (cfr. § 3.2.4).¹⁹⁹

La GUI di *Transkribus* offre sia strumenti automatici che manuali per effettuare la segmentazione delle immagini in:

- regioni di testo (*text regions*) ovvero aree, generalmente di forma rettangolare, che racchiudono blocchi di testo;
- linee di testo ovvero aree situate all'interno di una regione di testo; possono essere descritte come poligoni che racchiudono tutto il testo scritto in una riga;
- linee di base (*baselines*) ovvero polilinee che corrono lungo la parte inferiore della riga di testo; sono il punto di riferimento più importante per il riconoscimento del testo.

Il riconoscimento automatico del layout viene eseguito contestualmente all'avvio di ogni nuovo processo di riconoscimento del testo, ma può anche essere effettuato come operazione separata.²⁰⁰ Quest'ultima operazione è consigliabile nelle seguenti circostanze: quando si intende utilizzare *Transkribus* per la trascrizione manuale dei documenti; nella fase di preparazione delle trascrizioni per l'addestramento di un nuovo modello di HTR; in presenza di un layout di pagina complesso, come nel caso di documenti che contengono testo in colonna, tabelle, immagini e numerose annotazioni marginali. Per avviare il riconoscimento del layout come fase distinta, è necessario selezionare le pagine o i documenti desiderati e successivamente

¹⁹⁹ Nel caso di documenti caratterizzati da una struttura semplice e lineare (ad esempio, un diario manoscritto con righe di testo uniformi), *Transkribus* dimostra un'elevata efficacia nell'identificare e categorizzare i diversi elementi del layout. Tuttavia, l'operazione di segmentazione si complica notevolmente in presenza di documenti con layout più articolati, quali quotidiani con numerose colonne di testo, blocchi di contenuto eterogenei, titoli, immagini e didascalie. Per ovviare a questo problema, recentemente *Transkribus* ha introdotto nuove funzionalità che consentono di addestrare modelli *ad hoc* per il riconoscimento automatico del layout. L'addestramento di modelli personalizzati garantisce che la struttura intrinseca del documento sia non solo preservata ma anche compresa in modo semanticamente accurato. Per approfondire queste nuove funzionalità, cfr. i seguenti link: <https://help.transkribus.org/field-models>; <https://help.transkribus.org/tables>; <https://help.transkribus.org/baselines-models> (ultimo accesso 17 ottobre 2025).

²⁰⁰ Per approfondire il processo di riconoscimento automatico del layout, cfr. il seguente link: <https://help.transkribus.org/automatic-layout-recognition> (ultimo accesso 17 ottobre 2015).

cliccare sul pulsante *Process with AI* nella barra degli strumenti situata sopra le miniature delle pagine; sulla destra della pagina, si aprirà una finestra dove occorre selezionare l'opzione *Layout* e uno dei modelli disponibili. Il modello impostato di default è denominato *Universal Lines*. Si tratta del modello che attualmente vanta la maggiore versatilità all'interno della piattaforma, qualificandosi, di conseguenza, come una scelta raccomandabile in situazioni in cui vi sia incertezza circa il modello più appropriato per le specifiche caratteristiche del materiale analizzato. Quindi, è sufficiente cliccare sul pulsante *Start Recognition* per dare il via al processo. L'avanzamento dell'operazione può essere monitorato tramite la funzione *Jobs*. Al termine del processo, aprendo la pagina sarà possibile verificare il risultato del riconoscimento automatico del layout direttamente sull'immagine, che apparirà ora segmentata in aree di testo e linee distinte. Qualora il riconoscimento automatico del layout non avesse prodotto risultati soddisfacenti (ad esempio, omettendo alcune righe o raggruppando in modo errato le righe nelle regioni di testo), è possibile intervenire sulle impostazioni di configurazione avanzate.²⁰¹

La segmentazione automatica richiede pochi minuti per individuare regioni di testo, linee di testo e linee di base, a condizione però che le parole siano disposte in maniera relativamente ordinata sui documenti. Se, invece, i documenti presentano strutture più sofisticate, i risultati della segmentazione automatica potrebbero essere meno precisi e richiedere interventi di correzione manuali; infatti, poiché i modelli di HTR si basano principalmente sulle informazioni derivate dalle linee di base – come si è già ricordato all'inizio di questo paragrafo – è fondamentale che gli utenti verifichino il tracciato e il posizionamento delle linee di base generate automaticamente e procedano eventualmente alla loro correzione, al fine di assicurare una trascrizione accurata.²⁰² Tuttavia, bisogna constatare che quando si lavora con collezioni di documenti molto ampie, correggere manualmente le linee di base, disegnandole una ad una per adattarle alle specificità del layout in esame, può risultare un'attività molto dispendiosa in termini di tempo ed energie intellettuali. Ration per cui, nell'eventualità che il modello di riconoscimento *Universal Lines* si riveli inefficace, è possibile procedere all'addestramento di un *baseline model* specificamente dedicato alla tipologia di documenti considerati. In seguito, l'applicazione di questo modello personalizzato agli altri documenti della collezione ne

²⁰¹ Per una spiegazione dettagliata delle impostazioni di configurazione avanzate, cfr. il seguente link: <https://help.transkribus.org/advanced-layout-configuration-settings?__hstc=48127928.03087d00981f619c0c9c7235c214c5c7.1727456760735.1746641411683.1747231676041.36&__hssc=48127928.6869706968697070.1747231676041&__hsfp=1778039452> (ultimo accesso 17 ottobre 2015).

²⁰² Per approfondire l'operazione di modifica manuale del layout, cfr. il seguente link: <<https://help.transkribus.org/manual-layout-editing>> (ultimo accesso 17 ottobre 2015).

consentirà la segmentazione in modo coerente con gli esempi forniti durante la fase di addestramento. A tal proposito, è opportuno sottolineare la distinzione funzionale esistente fra i *baseline model* appena menzionati e i *field model* di cui si parlerà diffusamente nel paragrafo successivo. Mentre i *field model* sono concepiti per identificare automaticamente la struttura dei documenti mediante la segmentazione e la marcatura delle diverse aree testuali, i *baseline model* si focalizzano unicamente sul rilevamento delle linee di testo, con il vantaggio però di essere stati addestrati sullo specifico layout dei documenti in esame. Questa specializzazione conferisce loro una precisione potenzialmente superiore a quella offerta dal modello predefinito *Universal Lines*.²⁰³

La trascrizione del testo rappresenta la terza ed ultima fase del processo di creazione della *ground truth*.²⁰⁴ La trascrizione può essere effettuata manualmente o in modo semi-automatico, qualora sia disponibile un modello pubblico di riconoscimento del testo che si adatta abbastanza bene ai documenti oggetto di studio. In tal caso, l'utente può eseguire prima il modello pubblico,

²⁰³ Per addestrare un *baseline model*, bisogna innanzitutto preparare un set di dati *ground truth*. Sebbene un campione di 50 pagine rappresenti un buon punto di partenza, è importante notare che l'efficacia del modello è strettamente correlata alla complessità del layout dei documenti. Pertanto, dopo un primo ciclo di addestramento condotto su tale campione, si potrà valutare l'adeguatezza del *baseline model* ottenuto e considerare l'eventuale necessità di ulteriore materiale di addestramento. La preparazione della *ground truth* implica la segmentazione delle pagine in regioni di testo e linee di base, un'operazione che può essere eseguita sia in modalità automatica che manuale. Bisogna sottolineare che non è necessario corredare le pagine con trascrizioni testuali, in quanto l'addestramento del *baseline model* si concentra esclusivamente sulle linee di base, rendendo irrilevante la presenza di testo trascritto. Una volta completata questa fase, si può procedere all'addestramento vero e proprio. Per avviare tale processo, bisogna accedere all'ambiente di lavoro *Models*, dedicato all'addestramento di tutti i modelli di riconoscimento (*Text*, *Baseline*, *Table* e *Field*); nel caso specifico, si dovrà selezionare ovviamente l'opzione *Baselines Model*. Successivamente, è necessario indicare le pagine con il layout corretto. In sintesi, l'addestramento si articola in quattro fasi distinte: preparazione dei dati di addestramento, selezione dei dati di convalida, configurazione del modello, riepilogo e avvio.

Per approfondire questo processo nel dettaglio, cfr. il seguente link: <<https://help.transkribus.org/baselines-models>> (ultimo accesso 6 maggio 2025).

²⁰⁴ Per approfondire il processo di trascrizione manuale, cfr. il seguente link: <<https://help.transkribus.org/transcribing-manually>> (ultimo accesso 17 ottobre 2015).

Gli utenti, che sono già in possesso delle trascrizioni dei propri documenti in formato DOCX o TXT, possono usufruire dello strumento di abbinamento *Text-Image Matching* (TIM). In seguito al caricamento delle immagini e dei file di testo trascritto sulla piattaforma, l'algoritmo TIM è in grado di stabilire una corrispondenza biunivoca tra le righe presenti nelle immagini e le righe del testo trascritto; solamente le righe che raggiungono un livello di affidabilità predefinito sono incluse nei dati di addestramento. Questo strumento rappresenta, quindi, un valido ausilio per semplificare e velocizzare il processo di creazione della *ground truth* e il successivo addestramento di modelli di HTR.

Per approfondire, cfr. il seguente link <<https://help.transkribus.org/import-existing-transcriptions-with-text-image-matching>> (ultimo accesso 17 ottobre 2025).

poi correggere manualmente le trascrizioni generate automaticamente e infine salvarle come *ground truth*.²⁰⁵ In entrambi i casi, è fondamentale che le trascrizioni della *ground truth* siano il più possibile accurate, corrette e coerenti con le convenzioni editoriali adottate.²⁰⁶ La GUI di *Transkribus* presenta un *editor* di testo suddiviso in righe numerate, ognuna delle quali è associata in modo univoco alle linee di testo corrispondenti sull'immagine del documento. Per ciascuna riga, l'utente può produrre una trascrizione che replica fedelmente il testo, inclusi eventuali errori ortografici. In tal caso, si avrà una trascrizione diplomatica in cui le combinazioni di parole, i segni di interpunzione, le lettere maiuscole e minuscole, gli apici e i pedici, i simboli inconsueti e le abbreviazioni, sono trascritti esattamente così come si leggono nel documento originale. Il vantaggio di questo approccio è quello di ottenere un modello di HTR robusto, capace di trascrivere esattamente ciò che appare nell'immagine. In alternativa, l'utente può produrre una trascrizione normalizzata secondo le convenzioni ortografiche e interpuntive moderne. Infatti, se tali convenzioni sono adottate in modo coerente in tutte le pagine trascritte e la *ground truth* è abbastanza ampia, le reti neurali sono in grado di imparare a separare le parole che appaiono unite nei documenti, normalizzare l'ortografia storica, trascrivere gli apici e i pedici in linea con il resto del testo e risolvere le abbreviazioni.²⁰⁷

²⁰⁵ Quando si sceglie un modello di HTR, è necessario considerare alcuni aspetti fondamentali:

- la natura del materiale, distinguendo tra manoscritto e stampato;
- la lingua in cui il testo è redatto;
- il periodo storico di riferimento;
- la tipologia di scrittura utilizzata.

Un indicatore cruciale da valutare è il *Character Error Rate* (CER). Questa metrica esprime l'accuratezza del modello, misurando la "distanza" tra una trascrizione ideale e il risultato del riconoscimento automatico. In altre parole, il CER indica la percentuale di caratteri che il modello ha trascritto in modo errato. Un valore basso di CER denota una maggiore precisione del modello. L'insieme dei modelli disponibili per l'utilizzo è consultabile nell'ambiente di lavoro *Models*. Selezionando l'opzione *Show Description*, è possibile esaminare le caratteristiche specifiche di ciascun modello e visualizzarne le statistiche di addestramento (come il numero di parole, righe e pagine impiegate). Inoltre, è disponibile una funzione di test rapido del testo per valutare direttamente le prestazioni del modello. Si noti che la maggior parte dei modelli è stata sviluppata con l'ausilio del motore di riconoscimento *PyLaia*. Per approfondire, cfr. il seguente link <<https://help.transkribus.org/text-recognition>> (ultimo accesso 17 ottobre 2025).

²⁰⁶ Si ricorda che le convenzioni editoriali adottate nella trascrizione devono essere esplicitate nel campo *Details* del modello di HTR.

²⁰⁷ Per approfondire la procedura di preparazione dei dati di addestramento, cfr. il seguente link: <<https://help.transkribus.org/data-preparation>> (ultimo accesso 17 ottobre 2025).

In particolare, bisogna prestare attenzione ai seguenti elementi:

- segni diacritici (ad esempio, accenti, circonflessi, cediglie, trattini, tilde, ecc.);
- *i/j* e *I/J*: nei documenti storico-letterari le lettere *i* e *j* sono usate spesso in modo intercambiabile; possono essere trascritte come appaiono nei documenti o adattate all'ortografia moderna;
- *u/v* e *U/V*: nei documenti storico-letterari spesso la *v* ricorre all'inizio delle parole, mentre la *u* al centro e alla fine; possono essere trascritte come appaiono nei documenti o adattate all'ortografia moderna;
- legature: sono combinazioni comuni di lettere che formano un nuovo carattere; possono essere trascritte per esteso, usando i singoli caratteri che compongono la legatura (ad esempio, *præ* diventa *praes*);
- la lettera *s* può apparire in diverse forme: la *s* normale e la *s* lunga (con discendente) possono essere trascritte entrambe come *s* normale oppure, secondo la loro forma, come *s* o *f* (U-017F); invece, la doppia *s* o *ß* (*s* acuta o *Eszett*) devono essere trascritte secondo il testo originale;
- parole sillabate: quando le parole sillabate compaiono alla fine della riga, devono essere trascritte e separate secondo il testo originale; il segno – dev'essere aggiunto alla fine della riga solo se presente;
- stili di testo: cliccando sul pulsante *Tags*, si possono etichettare parole o porzioni di parole come grassetto, corsivo, barrato, sottolineato, apice o pedice; aggiungendo questi tag nella fase di l'addestramento del modello, essi saranno aggiunti automaticamente durante il riconoscimento di nuove pagine (per il momento, questa funzione è disponibile solo in *TranskribusX*);
- font: font diversi come *Kurrent* o *Antiqua* non sono contrassegnati in modo speciale.

Un altro elemento a cui prestare attenzione è rappresentato dalle abbreviazioni che nei documenti storico-letterari sono molto frequenti. A seconda delle esigenze, si può addestrare il modello a:

- mantenere la forma abbreviata, trascrivendo nella *ground truth* le abbreviazioni così come appaiono nei documenti, mediante i caratteri base o i caratteri speciali più simili a quelli usati dall'autore;
- trascrivere la forma estesa, risolvendo le abbreviazioni nella *ground truth* in modo uniforme;
- taggare le abbreviazioni e aggiungere l'espansione corrispondente come attributo; in tal caso, durante la creazione della *ground truth*, bisognerà trascrivere le abbreviazioni così come appaiono nei documenti, marcarle e aggiungere la forma estesa come attributo nel campo *Expansion* del tag *Abbreviation*.²⁰⁸

Recentemente, l'ottimizzazione della *web app* di *Transkribus* ha contribuito a semplificare e velocizzare il processo di trascrizione, consentendo a più utenti di lavorare in *team* sugli stessi documenti in modalità sincrona. Inoltre, gli studiosi interessati alla creazione di trascrizioni “arricchite”, destinate alla pubblicazione all'interno di edizioni scientifiche digitali, possono disporre di un'ampia gamma di strumenti di *markup*. Per semplificare, possiamo suddividere tali strumenti in due macro-categorie:

1. tag strutturali che servono, appunto, per contrassegnare la struttura dei documenti (ad esempio, titoli, paragrafi, numeri di pagina, note a margine o a piè di pagina, ecc.) e sono assegnati alle forme di layout (aree e linee di testo) nell'immagine;²⁰⁹
2. tag testuali che servono per contrassegnare parole e frasi nelle trascrizioni (ad esempio, nomi di persone, date, luoghi, eventi, abbreviazioni, ecc.) e aggiungere attributi.²¹⁰

²⁰⁸ In tal caso, l'utente dovrà ricordarsi di selezionare l'opzione relativa al tag *Abbreviation*, con la proprietà *Expansion*, durante l'addestramento del modello. Per approfondire, cfr. il seguente link: <<https://help.transkribus.org/model-setup-and-training>> (ultimo accesso 17 ottobre 2025).

²⁰⁹ È inoltre possibile stabilire tag di relazione tra le diverse forme attraverso un semplice clic con il tasto destro del mouse. Questa funzionalità permette di connettere elementi del layout che presentano un legame logico, aprendo la strada a molteplici applicazioni. I dati relativi a queste connessioni saranno inclusi nel documento XML generato. La relazione di tipo *Article* rappresenta un esempio comune e particolarmente utile. Essa si rivela efficace quando il testo di un articolo si sviluppa su più colonne, consentendo di collegare le diverse aree di testo che lo compongono e di definirne l'ordine di lettura, garantendo così la corretta ricostruzione del flusso narrativo. Per maggiori informazioni sui tag strutturali, cfr. il seguente link: <<https://help.transkribus.org/structure-tags>> (ultimo accesso 17 ottobre 2025).

²¹⁰ Gli attributi forniscono informazioni aggiuntive sul contenuto del tag e possono essere usati per estrarre dati dalla trascrizione per poi elaborarli (ad esempio, il tag *date* consente di marcare una data scritta nel documento e di aggiungere come attributo il giorno, il mese e l'anno in un formato standardizzato). È importante sottolineare che

La GUI di *Transkribus* mette a disposizione degli utenti alcuni tag predefiniti, ma consente anche di crearne di nuovi personalizzati sulla base delle caratteristiche dei documenti in esame. Per quanto riguarda i tag strutturali sopra menzionati, questi possono rivelarsi utili in diverse occasioni: per addestrare i *field model* più molte menzionati in precedenza (cfr. § 3.2.4); quando si desidera circoscrivere il riconoscimento del testo a porzioni specifiche del documento, anziché riconoscere l'intera pagina; infine, quando si desidera delimitare con precisione gli elementi costitutivi del layout, quali aree e linee di testo, ed esportare tali informazioni in formato XML, unitamente alle coordinate spaziali di ciascun elemento.

Per quanto riguarda, invece, i tag testuali – oltre che per contrassegnare le suddette entità – questi possono essere usati anche per applicare formattazioni quali grassetto, barrato, sottolineato, pedice e apice al testo. Inoltre, quando si ha a che fare con documenti antichi, segnati dal trascorrere del tempo, dall'azione degli elementi naturali e da altri eventi storici, possono rivelarsi molto utili i seguenti tag:

- *non chiaro (not clear)*: questo tag è impiegato per marcare porzioni di testo impossibili da trascrivere a causa della loro illeggibilità; evidenziando la porzione incerta nell'editor e contrassegnandola come “non chiaro”, è possibile inoltre aggiungere varianti alternative o ipotesi interpretative sulla parola illeggibile come attributo del tag;
- *lacuna (gap)*: se una porzione di testo è irrecuperabile o totalmente assente e non può essere trascritta, si inserisce il tag *gap* nel punto esatto in cui essa dovrebbe apparire all'interno dell'editor;²¹¹
- *annerimento (blackening)*: questo tag è pensato per oscurare informazioni riservate o sensibili nei file esportati; quindi, è particolarmente utile per anonimizzare dati personali in documenti destinati alla pubblicazione. L'applicazione di questo tag è complementare

non tutti i tag richiedono attributi; dipende dalle esigenze specifiche. Inoltre, si ricorda che è possibile aggiungere facilmente l'*ID Wikidata* come attributo di un tag. Dopo aver selezionato un tag, è sufficiente fare clic sul pulsante *ID Wikidata* e verrà visualizzata un'entità suggerita relativa alla parola selezionata. È possibile selezionarla se è corretta o utilizzare la barra di ricerca per trovare il termine *Wikidata* appropriato.

Per maggiori informazioni sui tag testuali, cfr. il seguente link: https://help.transkribus.org/textual-tags?_hstc=48127928.03087d00981f619c0c9c7235c214c5c7.1727456760735.1747231676041.1747235536469.37&_hssc=48127928.25.1747235536469&_hsfp=1778039452 (ultimo accesso 17 ottobre 2025).

²¹¹ Nella fase di addestramento del modello di HTR, se la trascrizione della *ground truth* contiene i tag *not clear* o *gap* per escludere specifiche parole o frasi dall'apprendimento, è necessario selezionare l'opzione corrispondente. È bene sottolineare che questa azione comporterà l'omissione non solo degli elementi taggati, ma dell'intera riga in cui essi compaiono, preservando così l'integrità dei dati di addestramento ed evitando l'apprendimento di trascrizioni imprecise.

all'uso della cosiddetta “area di annerimento”, che permette di nascondere visivamente una parola o un'intera sezione direttamente sull'immagine. È importante notare che l'area di annerimento e le funzionalità avanzate di esportazione che gestiscono l'annerimento sono attualmente disponibili esclusivamente nella versione *TranskribusX*.

La gestione dei tag testuali e strutturali avviene in modo centralizzato a livello di raccolta, il che ne agevola la modifica e ne promuove un utilizzo omogeneo da parte di tutti i collaboratori coinvolti nel progetto, garantendo coerenza e fluidità nel processo di trascrizione. Una volta completata la marcatura – qualora l'utente desideri elaborare immagini e trascrizioni al di fuori dell'ambiente *Transkribus* – la piattaforma offre la possibilità di scaricare i documenti; *Transkribus* mette a disposizione una gamma diversificata di opzioni di esportazione, consentendo all'utente di scegliere tra molteplici formati e strutture di file a seconda delle specifiche esigenze del progetto. In particolare, se le trascrizioni sono state arricchite con tag per identificare elementi testuali e strutturali, la modalità di esportazione Excel/CSV permette di raccogliere in modo sistematico tutti gli elementi marcati in un foglio di calcolo. Tale organizzazione facilita notevolmente l'accesso e l'analisi mirata di dati puntuali.²¹²

²¹² Fra le opzioni di esportazione disponibili, vi sono: file immagine (JPEG), documento DOCX, Documento PDF, file di testo TXT, file XML di pagina, foglio di calcolo EXCEL/CSV, file ALTO XML e TEI XML. Per approfondire la procedura di *download*, cfr. il seguente link: <<https://help.transkribus.org/downloading>> (ultimo accesso 17 ottobre 2025).

```
graph TD; A[CARICAMENTO DEI DOCUMENTI] --> B[SEGMENTAZIONE]; B --> C[CREAZIONE DEI DATI DI ADDESTRAMENTO]; C --> D[ADDESTRAMENTO DEL MODELLO]; D --> E[APPLICAZIONE DI MODELLI PRESISTENTI]; E --> F[RICERCA: <br/>• RICERCA FULL-TEXT<br/>• RICERCA FUZZY<br/>• RICERCA INTELLIGENTE]; E --> G[RICONOSCIMENTO DEL TESTO];
```

Il diagramma illustra il ciclo di vita di un sistema di ricerca semantica, suddiviso in sette fasi principali:

- CARICAMENTO DEI DOCUMENTI**: Fase iniziale di ingestione dei dati.
- SEGMENTAZIONE**: Processo di suddivisione dei documenti in unità significative.
- CREAZIONE DEI DATI DI ADDESTRAMENTO**: Preparazione dei dati per l'addestramento del modello.
- ADDESTRAMENTO DEL MODELLO**: Fase di apprendimento del modello di ricerca semantica.
- APPLICAZIONE DI MODELLI PRESISTENTI**: Utilizzo del modello addestrato per la ricerca e il riconoscimento.
- RICERCA:** Attività di ricerca avanzata, che include:
 - RICERCA FULL-TEXT
 - RICERCA FUZZY
 - RICERCA INTELLIGENTE
- RICONOSCIMENTO DEL TESTO**: Fase finale di estrazione e identificazione del contenuto testuale.

196

A tal proposito, ricordiamo che la “matrice di confusione” (o matrice di errore) è un metodo essenziale per valutare le prestazioni di un algoritmo di classificazione nel *machine learning*. In pratica, si tratta di una tabella che mette a confronto il numero di occorrenze reali (*ground truth*) di una specifica classe (positiva o negativa) con il numero di occorrenze predette dal modello per quella stessa classe. Nello specifico, le colonne rappresentano i valori predetti dal modello per una data classe, mentre le righe mostrano i valori effettivi (*ground truth*) di quella stessa classe. La casella in alto a sinistra fornisce il numero dei veri positivi (TP), ovvero il numero di previsioni corrette per la classe positiva. La casella sottostante riporta il numero dei falsi positivi (FP), cioè i casi di classe negativa erroneamente identificati come casi positivi; in statistica, questi sono anche definiti “errori di tipo I”. La casella in alto a destra indica il numero dei falsi negativi (FN), ovvero i casi effettivamente positivi erroneamente previsti come negativi; in statistica, questi sono anche definiti “errori di tipo II”. Infine, la casella in basso a destra riporta il numero dei veri negativi (TN), ovvero le occorrenze effettive della classe negativa previste con precisione. Sommando questi valori si ottiene il numero totale delle previsioni (cfr. Tab. 3.4).²¹³

Ad esempio, supponiamo di avere un modello che deve distinguere tra la classe positiva “lettera m” e la classe negativa “lettera n” in un set di immagini. La matrice di confusione si configura come una tabella del seguente tipo:

Tabella 3.4 – Matrice di confusione

	Predetto: lettera m	Predetto: lettera n
Reale: lettera m	Veri Positivi (VP): il modello ha correttamente previsto la lettera m quando l'immagine era effettivamente della lettera m (corretto).	Falsi Negativi (FN): il modello ha previsto la lettera n quando l'immagine era in realtà della lettera m (errore di tipo II).
Reale: lettera n	Falsi Positivi (FP): il modello ha previsto la lettera m quando l'immagine era in realtà della lettera n (errore di tipo I).	Veri Negativi (VN): il modello ha correttamente previsto la lettera n quando l'immagine era effettivamente della lettera n (corretto).

La matrice di confusione offre una visione dettagliata degli errori commessi dal modello, mostrando esattamente quali classi il modello tende a scambiare fra loro (ad esempio, se i valori

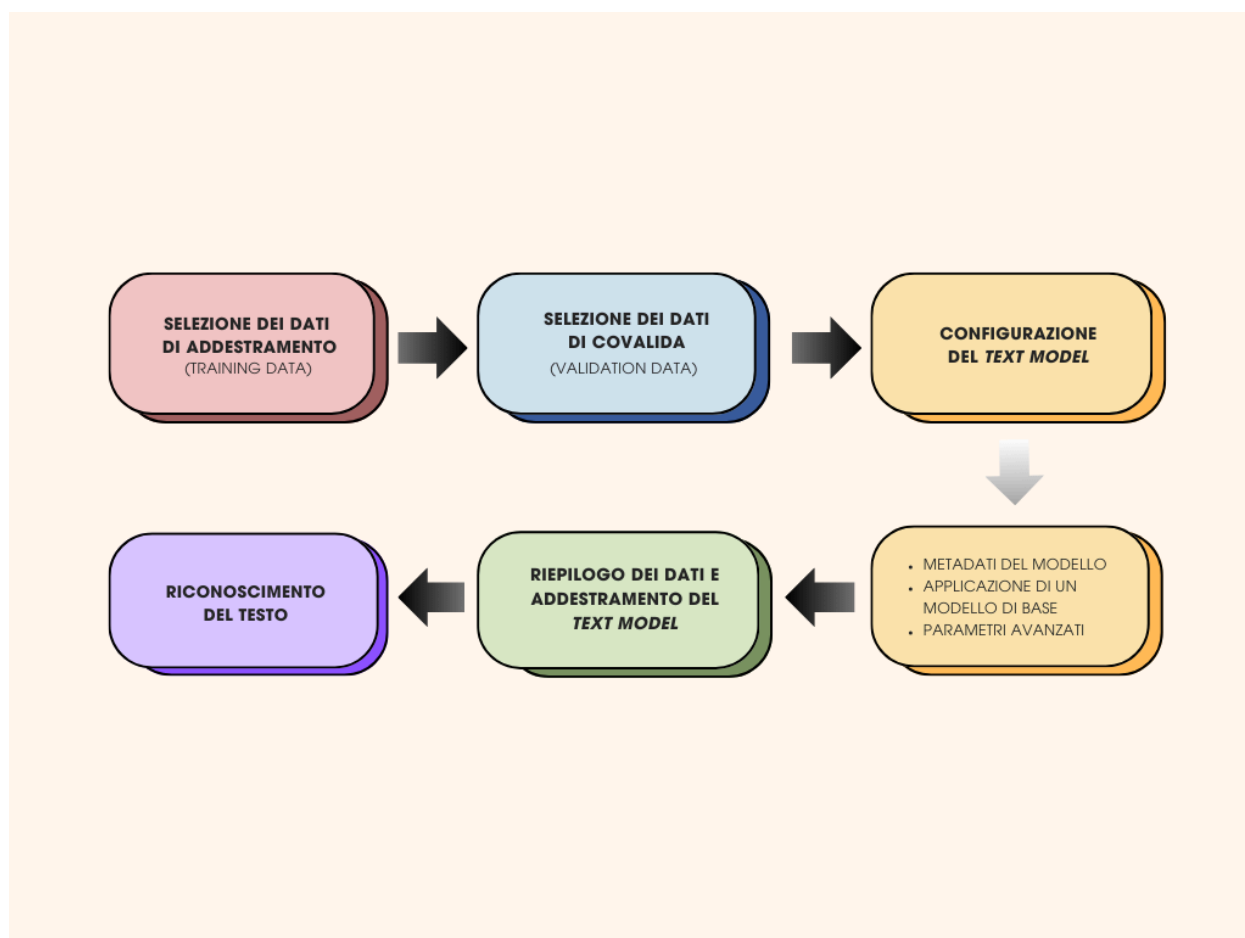
²¹³ Per approfondire l'argomento, cfr. il seguente link: <<https://www.ibm.com/it-it/think/topics/confusion-matrix>> (ultimo accesso 17 ottobre 2025).

dei FP per la classe “lettera m” sono elevati, significa che il modello tende a scambiare i casi della classe “lettera n” per quelli della classe “lettera m”) e le diverse tipologie di errori (ad esempio, distinguere tra FP e FN è molto importante, dato che le implicazioni di un determinato tipo di errore possono variare a seconda del contesto). Inoltre, essa consente di calcolare metriche più avanzate quali, l’accuratezza: $(VP + VN) / \text{Totale}$; la precisione (cioè quanto il modello è preciso quando predice una classe): $VP / (VP + FP)$; il richiamo o sensibilità (cioè quanti elementi di una classe reale il modello è riuscito a identificare): $VP / (VP + FN)$; e l’F1-Score, cioè la media armonica di precisione e richiamo. In definitiva, la matrice di confusione è un potente strumento diagnostico, in quanto non si limita a fornirci una singola misurazione di accuratezza complessiva – che a volte può essere fuorviante – ma offre un’analisi approfondita delle prestazioni di un modello di classificazione.

L’addestramento del modello di HTR può essere avviato direttamente nella sezione *Desk*. In tal caso, bisogna spuntare le caselle delle pagine trascritte correttamente; quindi cliccare sul pulsante *Process with AI* nella barra degli strumenti sopra le miniature di pagina e scegliere l’opzione *Text Recognition Model*. In alternativa, l’addestramento può essere avviato nella sezione *Models*, posizionata sul lato destro della barra di navigazione. Come si è detto, questa sezione costituisce l’ambiente di lavoro specificamente dedicato all’addestramento di tutti i modelli di IA. Accedendo alla sezione, i modelli di riconoscimento del testo disponibili, pubblici e privati, vengono visualizzati per impostazione predefinita. L’utente deve passare alla sezione *Training Lab* e cliccare sul pulsante *Train New Models* sulla sinistra; selezionando l’opzione *Text Recognition Model* si aprirà la finestra di configurazione dell’addestramento.²¹⁴

²¹⁴ Per approfondire la procedura di addestramento di un modello di HTR, cfr. il seguente link: <https://help.transkribus.org/model-setup-and-training> (ultimo accesso 17 ottobre 2025).

Figura 3.6 – Configurazione di un modello di HTR personalizzato



A questo punto, è necessario selezionare la collezione contenente i documenti *ground truth*. Durante il processo di addestramento, le pagine della *ground truth* sono suddivise in due gruppi:

- i dati di addestramento (*training data*) servono per regolare i parametri del modello, andando a formare la base su cui si sviluppa il suo apprendimento; detto altrimenti, il modello viene effettivamente addestrato utilizzando queste pagine;
- i dati di convalida (*validation data*) offrono una valutazione oggettiva del modello e servono per ottimizzarne i parametri durante l'addestramento; in pratica, queste pagine vengono messe da parte e usate per verificare la precisione del modello.

Per selezionare le pagine da includere nei dati di addestramento (*training data*), l'utente può spuntare la casella accanto al titolo del documento; in questo modo, tutte le trascrizioni disponibili all'interno del documento saranno automaticamente selezionate. In alternativa, l'utente può selezionare singole pagine specifiche cliccando prima sul pulsante *Select pages* e

poi su *Save and go back*. Inoltre, è possibile filtrare le pagine selezionate in base al loro stato. Per impostazione predefinita, la piattaforma opta per la “trascrizione più recente” (indipendentemente dallo stato di elaborazione con cui le pagine sono state salvate); ma è possibile considerare solamente le pagine che hanno raggiunto lo stato *ground truth*; le pagine prive di trascrizioni non possono essere selezionate.

Per selezionare le pagine da includere nei dati di convalida (*validation data*), l’utente può procedere manualmente o in modo automatico. La selezione manuale funziona in modo analogo a quella dei dati di addestramento. È possibile scegliere solo le pagine che contengono testo e che ovviamente non sono già state incluse nei dati di addestramento. Optando per la selezione automatica, invece, l’utente può decidere se assegnare ai dati di convalida una percentuale di pagine pari al 2%, al 5% o al 10% dei dati *ground truth*. Tuttavia, è consigliabile optare per il 10%. In linea di massima, la selezione automatica è preferibile per ottenere dati di convalida più diversificati. Ad ogni modo, è fondamentale che i dati di convalida siano rappresentativi della *ground truth* e includano tutti gli stili di scrittura, le grafie e le lingue presenti nei dati di addestramento; infatti, una scarsa varietà dei dati di convalida può distorcere la misurazione delle prestazioni del modello.

Il passaggio successivo consiste nella configurazione del modello. Innanzitutto, l’utente deve fornire alcuni metadati, quali: nome, descrizione (materiale manoscritto o stampato, periodo storico di riferimento, numero di mani coinvolte, gestione delle abbreviazioni e altre convenzioni editoriali adottate nella trascrizione), URL dell’immagine da usare come anteprima (facoltativo), lingua/e dei documenti e arco temporale coperto dai documenti. Questi metadati hanno la funzione di facilitare l’individuazione del modello in future ricerche. Successivamente, l’utente avrà la possibilità di selezionare un *base model* tra quelli personali o tra quelli pubblicamente disponibili sulla piattaforma. Il *base model* servirà come punto di partenza per l’addestramento del nuovo modello. In questo modo, l’addestramento non inizierà da zero, ma beneficerà delle conoscenze pregresse del *base model* prescelto. È fondamentale, però, che il *base model* sia stato addestrato con uno stile di scrittura simile a quello del modello che si intende creare. Questo approccio può ridurre il numero di pagine *ground truth* necessarie, accelerando il processo di addestramento e, al contempo, migliorando la qualità dei risultati di riconoscimento. Tuttavia, bisogna precisare che i benefici connessi all’uso di un *base model* non sono sempre garantiti e devono essere valutati caso per caso.²¹⁵ Inoltre, è importante sottolineare che i cosiddetti “super

²¹⁵ Per approfondire, cfr. REED-COOP, *Should I use a base model when training a custom model?*, «Transkribus blog», 7 febbraio 2025, consultabile al seguente link:

model” come *Text Titan* non possono essere usati come *base model* per l'addestramento di modelli personalizzati, a causa dell'enorme mole di dati su cui sono basati.²¹⁶

Infine, per ottimizzare le prestazioni del modello, l'utente può configurare alcuni parametri avanzati che saranno descritti in dettaglio nel prosieguo del paragrafo.

- Il parametro *cicli di addestramento* (*training cycles*) si riferisce al numero di epoche, ovvero al numero di volte in cui l'algoritmo di apprendimento esamina i dati di addestramento e si autovaluta sia sui dati di addestramento che su quelli di convalida. Il valore che viene impostato indica il numero massimo di epoche; l'addestramento si interromperà automaticamente quando il modello non migliorerà più, cioè avrà raggiunto il CER (*Character Error Rate*) più basso possibile. Inizialmente, è consigliabile mantenere l'impostazione predefinita corrispondente a 100 epoche.
- Il parametro *arresto anticipato* (*early stopping*) indica, invece, il numero minimo di epoche che il modello deve completare. Una volta raggiunto questo numero, se il CER dei dati di convalida continua a diminuire, l'addestramento proseguirà e terminerà autonomamente quando non ci saranno ulteriori miglioramenti; viceversa, se il CER non diminuisce più, l'addestramento si interromperà. In pratica, l'arresto anticipato forza il modello a eseguire almeno il numero di epoche specificato. Per la maggior parte dei

<<https://blog.transkribus.org/en/using-base-model-when-training-a-custom-model>> (ultimo accesso 17 ottobre 2025).

²¹⁶ I cosiddetti “super modelli”, come *Text Titan I*, si basano sull'architettura dei trasformatori. Questi ultimi sono una particolare tipologia di rete neurale che ha rivoluzionato il *deep learning*, specialmente nell'elaborazione del linguaggio naturale (NLP) e nella visione artificiale. A differenza delle reti neurali ricorrenti (RNN) o delle reti neurali convoluzionali (CNN), che elaborano le sequenze di dati in modo sequenziale o tramite finestre locali, i trasformatori si fondano su un meccanismo chiamato “attenzione” (*attention mechanism*). Questo meccanismo permette al modello di valutare l'importanza di ogni elemento (ad esempio, una parola in una frase) in relazione a tutti gli altri elementi della stessa sequenza di input. In pratica, quando il modello elabora una parola, non la considera isolatamente, ma ne valuta il contesto attribuendo un “peso” o un'importanza diversa a tutte le altre parole nella sequenza. Ciò consente di catturare dipendenze a lungo raggio che le RNN tradizionali faticavano a gestire.

I super modelli rappresentano un'evoluzione significativa nel riconoscimento del testo, superando le performance dei tradizionali modelli *PyLaia*. La loro efficacia deriva da una duplice architettura: una componente ottica che elabora le immagini e un modello linguistico esteso capace di gestire più lingue, incluse le loro varianti storiche e moderne. Questa struttura consente ai super modelli di eccellere con materiali misti, riconoscendo sia il testo manoscritto che quello stampato senza la necessità di cambiare modello o impostazioni. Sebbene non possano essere personalizzati o addestrati dagli utenti, offrono prestazioni immediate e di alta qualità, facilitando la creazione di dati *ground truth* per l'addestramento di modelli *PyLaia* personalizzati.

Per approfondire l'argomento, cfr. REED-COOP, *Introducing Transkribus Super Models - Get Access to 'The Text Titan I'*, «Transkribus blog», 9 luglio 2023, consultabile al seguente link: <<https://blog.transkribus.org/en/introducing-transkribus-super-models-get-access-to-the-text-titan-i>> (ultimo accesso 17 ottobre 2025).

modelli, l'impostazione predefinita di 20 epoche funziona abbastanza bene e pertanto si consiglia di non modificarla per il primo addestramento. Tuttavia, se i dati di convalida presentano una variabilità interna scarsa o assente, il modello potrebbe fermarsi troppo presto. Per ovviare a questo problema, è necessario che i dati di convalida siano sufficientemente diversificati e contengano tutti gli stili di scrittura e le tipologie di materiali presenti nei dati di addestramento. Solamente se i dati di convalida sono una quantità esigua, si può aumentare il valore di questo parametro, al fine di garantire che l'addestramento non si interrompa prima di aver elaborato tutti i dati di addestramento; bisogna tenere presente però che ciò prolungherà il tempo di addestramento.

- Il parametro *testo invertito (Reverse text - RTL)* serve per invertire il testo durante l'addestramento, qualora l'orientamento della scrittura nell'immagine sia opposto a quella della trascrizione (ad esempio, testo scritto da destra a sinistra ma trascritto da sinistra a destra).
- Il parametro *utilizza poligoni lineari esistenti per l'addestramento (use existing line polygons for training)*, se selezionato, fa sì che i poligoni lineari non vengano ricalcolati durante l'addestramento, ma vengano utilizzati quelli già presenti. Al fine di ottenere prestazioni ottimali in fase di riconoscimento, è consigliabile usare poligoni lineari simili.
- Il parametro *addestra abbreviazioni con espansione (train abbreviations with expansion)* deve essere selezionato, qualora le abbreviazioni siano state trascritte nella *ground truth* così come si leggono nelle immagini e la loro espansione sia stata indicata come attributo del tag *Abbreviation*.
- Il parametro *ometti le righe con il tag 'non chiaro' o 'lacuna' (omit lines by tag 'unclear' or 'gap')* deve essere selezionato se questi tag sono stati usati nelle trascrizioni della *ground truth* per indicare parole o frasi da escludere dall'addestramento. È importante sottolineare che l'intera riga contenente questi tag verrà omessa dai dati di addestramento.

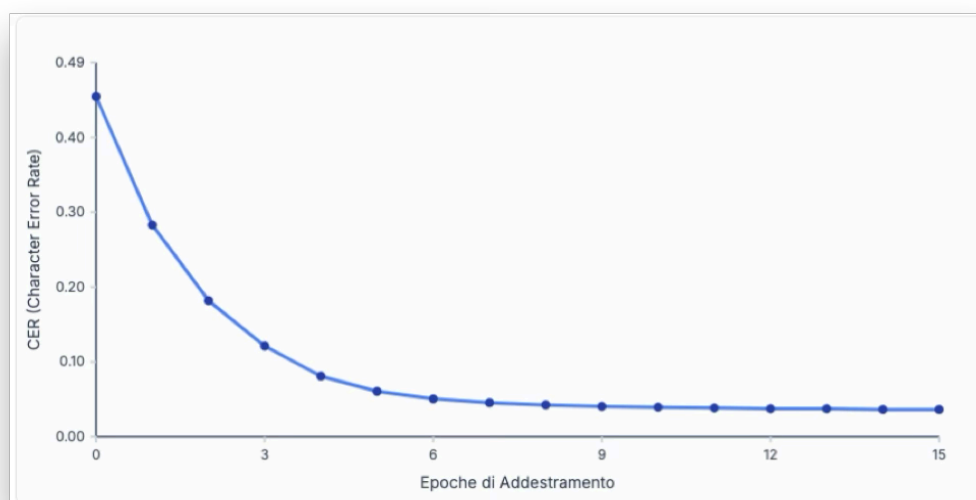
Dopo aver controllato la correttezza dei dati *ground truth* e dei parametri configurati, l'utente può avviare il processo di addestramento e monitorarne lo sviluppo tramite la funzione *Jobs* situata sulla destra della barra di navigazione. La durata del processo può variare considerevolmente, estendendosi da alcune ore a diversi giorni, in funzione della mole dei dati di addestramento e del carico gravante sull'infrastruttura informatica. Al termine dell'addestramento, l'utente riceverà una notifica via e-mail; quindi, accedendo alla sezione *My*

Models, l'utente può visualizzare il proprio modello di HTR e usarlo per trascrivere automaticamente le pagine della medesima collezione che non sono state incluse nella *ground truth*; preliminarmente, queste pagine dovranno essere caricate sulla piattaforma e sottoposte alla analisi del layout (cfr. Fig. 3.2).

Come detto in precedenza, l'accuratezza complessiva del modello di HTR è misurata attraverso il CER (*Character Error Rate*) che esprime la “distanza” ovvero la divergenza esistente tra una trascrizione di riferimento ideale (la *ground truth*) e il testo generato dal riconoscimento automatico. In altre parole, questa metrica si riferisce alla percentuale media di caratteri trascritti erroneamente dal sistema. Il valore del CER, che può essere visualizzato nella descrizione del modello, è calcolato sul *validation set*, ovvero su quelle pagine della *ground truth* che sono state accantonate e non sono state processate durante la fase di addestramento. In genere, è necessario addestrare più modelli prima di raggiungere un risultato soddisfacente in termini di CER. Il graduale miglioramento del modello può essere rappresentato graficamente tramite una curva di apprendimento (*learning curve*) su un sistema di assi cartesiani, dove:

- l'asse delle ascisse (asse x) denota le epoche di addestramento (*Training Epochs*); tale valore incrementa in relazione al tempo di addestramento del modello;
- l'asse delle ordinate (asse y) rappresenta il CER; un valore di CER inferiore è indicativo di una migliore performance del modello.

Figura 3.7 - Curva di apprendimento del modello di HTR



La curva di apprendimento inizia da un punto relativamente elevato sull'asse y (un CER elevato) all'inizio del processo di addestramento (epoca 0 o 1). Con l'avanzare delle epoche di addestramento, il CER tende a diminuire progressivamente. In un grafico ideale come quello della Fig. 3.7, la curva mostrerebbe una rapida diminuzione nelle prime epoche, segnalando un miglioramento significativo e veloce. Successivamente, la pendenza della discesa si attenuerebbe, fino a raggiungere una fase di stabilizzazione o una diminuzione molto lenta nelle epoche successive. Ciò suggerisce che il modello ha raggiunto un plateau nelle sue prestazioni o si sta avvicinando alla sua capacità massima di apprendimento dai dati forniti. Questo andamento riflette la capacità del modello di apprendere dagli errori e di adattarsi per ottimizzare il riconoscimento dei caratteri. La valutazione del modello al termine di ogni epoca consente di monitorare questo progresso e di determinare l'efficacia della sua generalizzazione sui dati di validazione.²¹⁷

Secondo le rilevazioni di REED-COOP, l'efficienza della tecnologia di HTR varia in base alla complessità del materiale sorgente. In condizioni ottimali, ovvero con documenti manoscritti caratterizzati da stili grafici e layout semplici, l'HTR è in grado di generare trascrizioni automatiche con un CER inferiore al 5%. Per contro, i modelli addestrati su materiale a stampa dimostrano performance superiori, raggiungendo un CER compreso tra l'1% e il 2%. L'esperienza consolidata della *user community* di *Transkribus* dimostra che le trascrizioni con un CER basso consentono una revisione e correzione notevolmente più rapida, ottimizzando l'impiego delle risorse rispetto alla trascrizione manuale integrale.²¹⁸ Tuttavia, quando l'accuratezza di un modello di HTR è inferiore, con un CER superiore al 10%, la sua utilità come risorsa di ricerca autonoma diminuisce drasticamente. Diversi studi indicano che la rettifica di un elevato numero di errori diviene più onerosa in termini di tempo rispetto alla trascrizione manuale *ex novo*. Nonostante ciò, un output meno accurato non è del tutto privo di valore; infatti, l'output dell'HTR, seppur imperfetto, può comunque costituire una base robusta per l'interrogazione e l'indicizzazione di grandi *corpora* di documenti digitalizzati, fungendo da strumento complementare per la ricerca e l'analisi su vasta scala. A questo proposito, ricordiamo che la GUI di *Transkribus* integra una funzionalità avanzata di *information retrieval* denominata

²¹⁷ Per approfondire l'argomento, cfr. il seguente link: <https://help.transkribus.org/character-error-rate-and-learning-curve> (ultimo accesso 17 ottobre 2025).

²¹⁸ Per approfondire, cfr. M. Terras, *Chapter 7: Inviting AI into the Archives: The Reception of Handwritten Recognition Technology into Historical Manuscript Transcription*, in *Archives, Access and Artificial Intelligence. Working with Born-Digital and Digitized Archival Collections*, a cura di L. Jaillant, Bielefeld University Press, 2022, pp. 185-186.

“ricerca intelligente” (*Smart Search*). Questo strumento sfrutta i valori di confidenza associati ai singoli caratteri durante il processo di HTR, per identificare e recuperare tutte le occorrenze potenziali di una *query string* specificata (questo approccio è noto, appunto, come *query by string* ovvero “ricerca per stringa”). I risultati di una *Smart Search* includono sia le corrispondenze considerate ottimali dal sistema, sia ulteriori occorrenze basate su interpretazioni alternative di ciascun carattere nella pagina; ciò implica che la *Smart Search* è in grado di localizzare termini, all'interno di un corpus di documenti, anche in presenza di errori di trascrizione generati dall'HTR. Inoltre, la tecnologia permette il riconoscimento e il recupero di risultati per lemmi che presentano variazioni ortografiche, dovute sia all'evoluzione diacronica della lingua sia alle peculiarità grafiche del singolo autore. La piattaforma visualizza i risultati di una *Smart Search* come un elenco di parole trascritte. Ogni occorrenza è corredata da miniature contestualizzate delle porzioni di pagina digitalizzate in cui la parola appare e da una valutazione del livello di confidenza per ciascuna corrispondenza.²¹⁹

Per completezza ricordiamo che, oltre alla *Smart Search*, la piattaforma *Transkribus*, offre altre due funzionalità di ricerca. La prima e più semplice opzione di ricerca è quella *full-text* che consente di individuare determinati termini all'interno delle trascrizioni e delle pagine corrispondenti.²²⁰ La seconda opzione è la ricerca *fuzzy*. Si tratta di una metodologia di *information retrieval* che consente il recupero di stringhe caratterizzate da minime deviazioni rispetto al *query term*, tipicamente limitate a una o due discordanze di caratteri. Tale funzionalità risulta particolarmente vantaggiosa in contesti di errori ortografici o variazioni grafematiche. L'accesso alle opzioni di ricerca fuzzy avviene tramite la *Global Text Search*; dopo aver inserito la parola o la frase desiderata nella barra di ricerca e avviato il processo, è necessario selezionare l'opzione *Show advanced filters*; all'interno di questa sezione, è disponibile il parametro *Fuzziness*, che consente di calibrare il livello di precisione delle corrispondenze su tre

²¹⁹ Per maggiori informazioni sulla tecnologia di *Smart Search*, cfr. il seguente link: <<https://help.transkribus.org/smart-search>> (ultimo accesso 17 ottobre 2025).

²²⁰ Per impostazione predefinita, la ricerca *full-text* viene eseguita sull'intero testo delle collezioni a cui si ha accesso. È possibile filtrare i risultati per: autore: (se il nome dell'autore è presente nei metadati del documento); uploader, titolo e ID della collezione: (per limitare la ricerca a una specifica collezione). Sotto i filtri viene visualizzato il numero di documenti e dei risultati della ricerca. I risultati della ricerca vengono visualizzati in grassetto nel loro contesto, con un'anteprima delle righe in cui si trovano i termini. Selezionando un risultato, si accede alla pagina corrispondente. Nell'immagine, i risultati della ricerca sono evidenziati in grigio, mentre nella trascrizione sono sottolineati in rosso. Per maggiori informazioni sulla funzionalità di ricerca *full-text*, cfr. il seguente link: <<https://help.transkribus.org/fulltext-search>> (ultimo accesso 17 ottobre 2025).

impostazioni: l'impostazione *Low* garantisce la massima precisione, restituendo esclusivamente corrispondenze esatte al *query term* (ad esempio, una ricerca per “casa” produrrà unicamente il risultato “casa”); l'impostazione *Medium* permette lievi variazioni del *query term*, recuperando termini strettamente correlati alla stringa di ricerca, come “casa” o “cosa”; infine, l'impostazione *High* è la più permissiva e tollera variazioni significative rispetto al *query term*, inclusi addizioni, eliminazioni o sostituzioni fino a un massimo di due caratteri; ad esempio, una ricerca per “casa” potrebbe restituire risultati come “cassa” (un carattere aggiunto), “cosa” (un carattere rimosso), “case” (un carattere sostituito), “cose” (due caratteri rimossi) e “casata” (due caratteri aggiunti). I risultati della ricerca includeranno sia le corrispondenze esatte che le varianti identificate.²²¹

Una volta completato il processo di riconoscimento del testo, l'utente può scaricare le trascrizioni generate, scegliendo fra una pluralità di opzioni di esportazione.²²² Le trascrizioni possono essere usate per una molteplicità di obiettivi scientifici: possono essere integrate all'interno di edizioni scientifiche digitali; analizzate con tecniche computazionali di analisi semantica o linguistica; o nel caso di grandi fondi librari e archivistici, incorporate in sistemi di gestione dei contenuti come strumenti di interrogazione.

A conclusione di questo capitolo, proponiamo un confronto tra modelli pubblici pre-addestrati e modelli addestrati *ad hoc*. La tabella seguente illustra sinteticamente i pro e i contro connessi alla loro applicazione.

²²¹ La ricerca *fuzzy* implementata in *Transkribus* si basa sull'algoritmo di distanza di Levenshtein. Questo algoritmo è una metrica per rilevare la similarità tra due stringhe, calcolando il numero minimo di operazioni (inserzioni, cancellazioni o sostituzioni di singoli caratteri) necessarie per trasformare una stringa nell'altra. Le impostazioni del parametro *fuzziness* corrispondono direttamente al valore della distanza di Levenshtein consentita:

- *Low* corrisponde a una distanza di Levenshtein pari a 0 (questo significa che vengono recuperate solo le corrispondenze esatte);
- *Medium* corrisponde a una distanza di Levenshtein pari a 1 (in questo caso, sono ammesse variazioni che richiedono una singola operazione di modifica);
- *High* corrisponde a una distanza di Levenshtein pari a 2 (questa impostazione permette il recupero di stringhe che differiscono dal termine di ricerca per un massimo di due operazioni).

Per maggiori informazioni sulla funzionalità di ricerca *fuzzy*, cfr. il seguente link: <<https://help.transkribus.org/fuzzy-search>> (ultimo accesso 17 ottobre 2025).

²²² Per approfondire l'operazione di download, cfr. il seguente link: <<https://help.transkribus.org/downloading>> (ultimo accesso 17 ottobre 2025).

Tabella 3.5 – Modelli pubblici e modelli ad hoc: pro e contro

Modelli Pubblici	Modelli Personalizzati
Pronti all'uso: sono pre-addestrati e quindi possono essere applicati direttamente ai documenti.	Maggiore accuratezza: essendo addestrati su <i>corpora</i> specifici, riescono ad apprendere e riprodurre con maggiore fedeltà gli stili di scrittura individuali e i layout unici, il che porta a trascrizioni molto più precise rispetto a quelle ottenibili con modelli generici.
Ampi e robusti: essendo addestrati con un'ampia mole di dati, risultano efficaci con una grande varietà di testi.	Adattabilità a materiale specifico: sono in grado di gestire in modo più efficace repertori lessicali o layout particolari che potrebbero rappresentare un ostacolo per i modelli pubblici pre-addestrati; questa flessibilità li rende ideali per testi complessi o non standardizzati.
Facili da usare: sono concepiti per essere accessibili anche ai nuovi utenti, senza richiedere alcuna configurazione.	Controllo sui dati di addestramento: è possibile decidere quali dati includere nel training set; Questo controllo permette di calibrare il sistema in base alle precise tipologie di documenti su cui si sta lavorando, ottimizzando così le prestazioni del riconoscimento.
Adatti a diversi tipi di materiali: sono concepiti per riconoscere sia il testo manoscritto che quello stampato, adattandosi a scopi scientifici generali.	Ottimizzazione per progetti specifici: possono essere costantemente migliorati aggiungendo nuovi dati; questa caratteristica li rende perfetti per progetti che riguardano documenti specialistici, storici o scritti da più autori, dove la specificità del materiale richiede un'elevata personalizzazione.

In sintesi, i modelli pubblici pre-addestrati sono stati concepiti per essere usati in modo versatile con diverse lingue, tipologie di documenti e grafie. Questi modelli rappresentano la soluzione ottimale per scenari che implicano grafie comuni, materiale a stampa, o quando si ha l'esigenza di una rapida operatività iniziale che non richieda un grado di precisione specialistica elevato. I modelli personalizzati, al contrario, si rivelano indispensabili per progetti di natura specialistica, la gestione di stili di scrittura idiosincratici, o laddove i modelli pubblici non riescano a

raggiungere il livello di accuratezza desiderato. La loro utilità è particolarmente evidente in tutte quelle attività che richiedono una trascrizione di elevata qualità e coerenza.

3.2.4 Dalla marcatura semantica del documento all'addestramento di *field model* personalizzati ²²³

Riepilogando quanto detto nel paragrafo precedente, ricordiamo brevemente che la creazione di una trascrizione automatica in *Transkribus* si articola in due fasi distinte e consequenziali. Inizialmente, la piattaforma esegue l'analisi del layout del documento, individuando con precisione la disposizione spaziale del testo sulla pagina. Successivamente, *Transkribus* procede alla decifrazione dei singoli caratteri e delle parole, generando la trascrizione desiderata; questa seconda fase è propriamente definita “riconoscimento del testo”. Se i documenti presentano una *mise en page* semplice – ad esempio, un unico blocco di testo con righe diritte, aventi tutte il medesimo orientamento verticale o orizzontale – la piattaforma è in grado di riconoscere il testo facilmente; ma l'operazione di riconoscimento si complica se la *mise en page* dei documenti è complessa. In tal caso, diventa cruciale il ricorso ai cosiddetti “*field model*”, strumenti avanzati di intelligenza artificiale, concepiti per rilevare automaticamente configurazioni strutturali più articolate, come quelle presenti nei giornali, nei registri e, in generale, in tutti quei documenti dove appaiono titoli, intestazioni, testo in colonna, tabelle, *marginalia*, numeri di pagina, ecc. Dunque, così come un modello di riconoscimento del testo può essere istruito ad interpretare un determinato sistema di scrittura o stile grafico, un *field model* può essere addestrato ad identificare i vari elementi costitutivi di uno specifico layout.

In particolare, l'addestramento di *field model* personalizzati può rivelarsi utile in due circostanze. In primo luogo, quando l'utente desidera rilevare determinati elementi all'interno dei documenti (ad esempio, numeri di pagina o intestazioni) per aggiungerli ad un layout preesistente; ciò potrebbe accadere qualora i documenti siano già stati processati e in seguito emerga la necessità di marcare altre componenti strutturali. Durante la fase di riconoscimento del testo, l'utente potrà decidere di selezionare le componenti strutturali identificate dal *field model* creato ad hoc e circoscrivere il riconoscimento del testo solo a quelle aree. In secondo luogo, quando l'utente

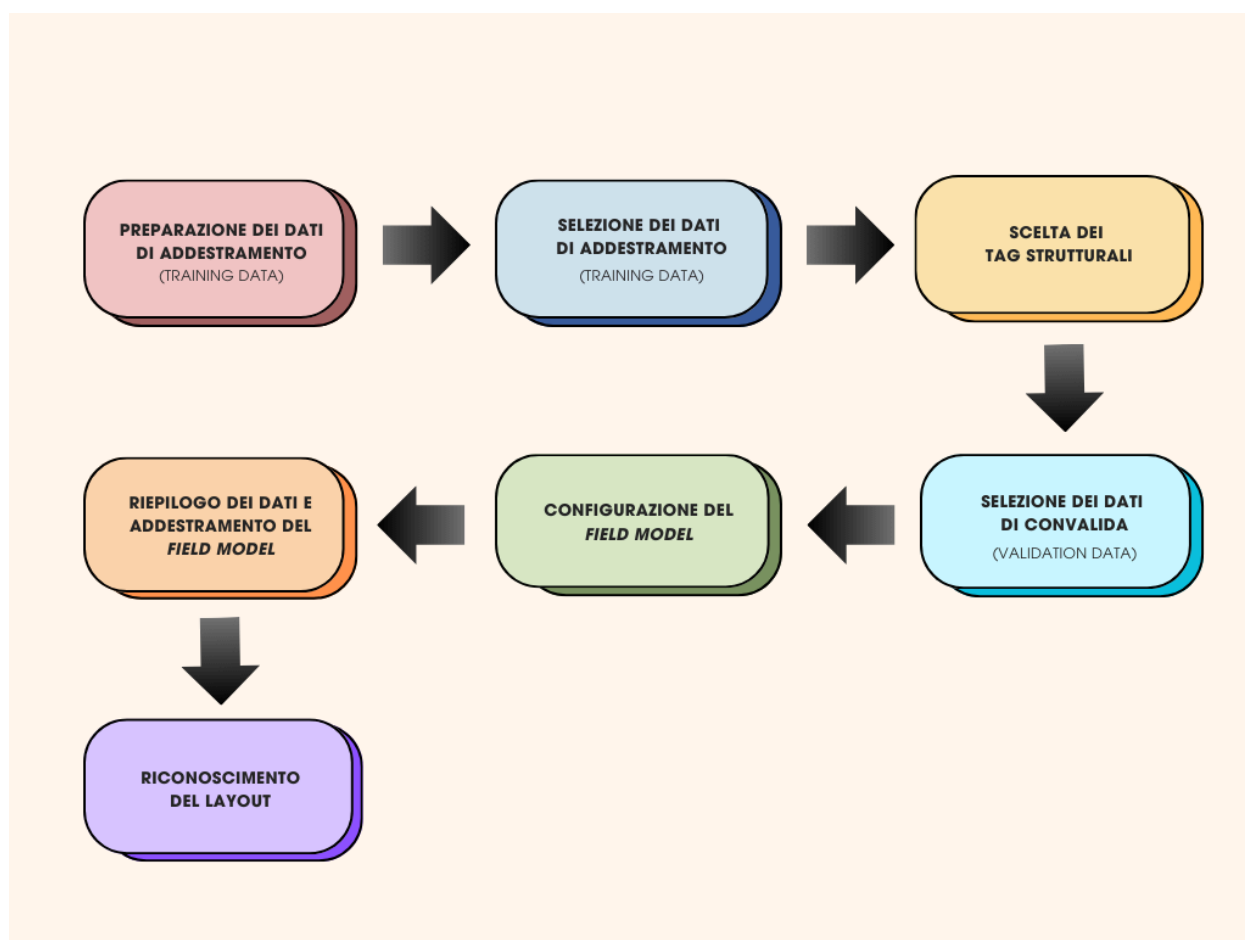
²²³ Per approfondire l'argomento, cfr. REED-COOP, *What are Field Models and how do they work?*, «Transkribus blog», 28 agosto 2024, consultabile al seguente link: <<https://blog.transkribus.org/en/what-are-field-models-and-how-do-they-work>> (ultimo accesso 17 ottobre 2025); REED-COOP, *Introducing Field Models - Trainable Layout AI in Transkribus*, «Transkribus blog», 25 luglio 2023, consultabile al seguente link: <<https://blog.transkribus.org/en/introducing-field-models-trainable-layout-ai-in-transkribus>> (ultimo accesso 17 ottobre 2025).

desidera focalizzare la propria attenzione su porzioni specifiche dei documenti (ad esempio, l'intestazione e la data) per estrapolare le informazioni che vi sono contenute. In tal caso, durante la fase di riconoscimento del testo, la piattaforma trascriverà soltanto il contenuto delle aree incluse nel modello; in questo modo, non sarà necessario trascrivere integralmente le pagine per poi isolare manualmente le aree di testo interessate. Questo approccio mirato ottimizza il flusso di lavoro, traducendosi concretamente in un notevole risparmio di tempo e risorse.

Per creare un *field model* personalizzato, è necessario predisporre, innanzitutto, i dati di addestramento (*training set*), ovvero le pagine che serviranno per istruire il modello a riconoscere specifici elementi del layout. A questo scopo, bisogna importare i documenti in una collezione dedicata all'interno d'ambiente *Desk*; quindi, aprire i documenti nell'editor per tracciare le aree di testo e le linee di base; infine procedere alla marcatura degli elementi costitutivi del layout (cfr. § 3.2.3). Dopo aver taggato un numero sufficiente di pagine (si raccomanda un minimo di 50 pagine, a seconda della complessità del layout), si può avviare la fase di addestramento vera e propria. Il processo si articola in cinque passaggi fondamentali che saranno illustrati in dettaglio nel prosieguo di questo paragrafo (cfr. Fig. 3.8).²²⁴

²²⁴ Per approfondire gli aspetti tecnici inerenti al processo di addestramento dei *field model*, cfr. il seguente link: <<https://help.transkribus.org/field-models>> (ultimo accesso 17 ottobre 2025).

Figura 3.8 – Addestramento di un *field model* personalizzato



Il primo passaggio consiste nella selezione del *training set*: accedendo all'ambiente di lavoro *Models*, l'utente deve cliccare sul pulsante *Train New Model* e scegliere ovviamente l'opzione *Field Model* per visualizzare le impostazioni di addestramento specifiche per questo tipo di modello; a questo punto, l'utente deve indicare la collezione che desidera usare e selezionare i singoli documenti all'interno di essa. Nel *training set*, è possibile includere le pagine che hanno raggiunto lo stato di elaborazione *final* (purché le aree di testo siano già ben definite), o usare solo le pagine che hanno raggiunto lo stato *ground truth*.

Il secondo passaggio consiste nella selezione dei tag strutturali definiti precedentemente, durante la fase di preparazione dei dati di addestramento. In pratica, questi tag rappresentano gli elementi che il modello imparerà a identificare. È possibile includere nel processo di addestramento anche le regioni di testo non taggate, selezionando l'apposita opzione; viceversa, qualora questa opzione non venga selezionata, le regioni prive di tag verranno escluse e il modello risultante non sarà in grado di riconoscerle. Inoltre, è importante evidenziare che i *field model* possono essere addestrati a riconoscere i cosiddetti “poligoni lineari” (ovvero le forme che delimitano

interamente il testo manoscritto su una riga), poiché l'algoritmo sottostante rimane invariato; infatti, ricordiamo che, mentre le linee di base si posizionano nella parte inferiore delle righe di testo, i poligoni lineari racchiudono l'intera estensione dei caratteri, inclusi ascendenti e discendenti. Durante la fase di riconoscimento testuale (o di addestramento), i poligoni lineari vengono elaborati automaticamente a partire dalle linee di base. In alcuni casi, tuttavia, tale processo potrebbe condurre a risultati imprecisi – ad esempio, in presenza di caratteri di grandi dimensioni, notazioni musicali o espressioni matematiche – compromettendo il riconoscimento del testo. Nell'eventualità in cui ciò accadesse, l'utente deve correggere i poligoni lineari nella *ground truth* e selezionare la funzione *Train on line polygons* per avviare l'addestramento di un *field model* specificamente progettato per i poligoni lineari.

Il terzo passaggio consiste nella selezione dei dati di convalida (*validation set*). Sono disponibili due modalità di selezione: automatica e manuale. Optando per la selezione automatica del set di convalida, è sufficiente indicare una percentuale di pagine del *dataset* da destinare a tale scopo (si raccomanda il 10%). In alternativa, è possibile definire manualmente quali pagine usare come dati di convalida.²²⁵

A questo punto, è possibile procedere alla configurazione del *field model*, fornendo innanzitutto alcune informazioni identificative, quali:

- nome del modello: si consiglia di assegnare un nome abbastanza esplicativo, usando termini che indichino con chiarezza la tipologia di layout a cui il modello è diretto;
- descrizione del modello: è opportuno fornire una breve descrizione della fisionomia del materiale usato come *ground truth* e delle finalità perseguite dal modello (ad esempio, edizioni in ottavo con note marginali manoscritte; documenti amministrativi contenenti grafici, ecc.);
- anteprima del modello: è possibile aggiungere un'immagine che funge da anteprima del modello, incollando nell'apposito campo l'URL corrispondente.

²²⁵ Bisogna precisare che i dati di addestramento e i dati di convalida rappresentano due parti distinte del medesimo insieme di dati. I primi costituiscono l'insieme di esempi usati per modulare i parametri del modello (il modello è istruito su queste pagine). I secondi, invece, offrono una valutazione imparziale delle capacità del modello (queste pagine sono usate per verificarne l'accuratezza).

Inoltre, se lo ritiene necessario, in questa fase l'utente può impostare i seguenti parametri avanzati, relativi al processo di addestramento:

- il parametro *training cycles* (cicli di addestramento) indica il numero di iterazioni (compreso tra 1.000 e 30.000) durante le quali il modello analizza i dati di addestramento per apprendere e adattarsi; se da una parte un numero maggiore di cicli può incrementare l'accuratezza del modello, dall'altra potrebbe anche aumentare il rischio di *overfitting*, ovvero di specializzazione del modello;
- il parametro *learning rate* (tasso di apprendimento) è un valore compreso tra 0,001 e 0,05, che definisce la velocità di progressione dell'addestramento da un ciclo all'altro; valori più elevati (e quindi una maggiore velocità) aumentano il rischio di ottenere un riconoscimento meno preciso;
- il parametro *depth of the backbone network* (profondità della rete backbone) consente all'utente di scegliere la profondità (misurata in livelli) della rete *backbone* sottesa al *field model*; l'opzione *Standard* (50 livelli) offre risultati più rapidi ed è indicata per documenti con layout semplici e uniformi; l'opzione *Enhanced* (101 livelli), invece, consente di rilevare elementi di layout più complessi, ma richiede tempi di elaborazione più lunghi.²²⁶

Nelle prime fasi di addestramento, è consigliabile non modificare i parametri avanzati e mantenere le impostazioni predefinite che si sono dimostrate efficaci nella maggior parte degli scenari. Dunque, dopo aver verificato che tutte le impostazioni e i dati inseriti siano corretti, si può avviare il processo di addestramento e monitorarne l'avanzamento cliccando sul pulsante *Jobs*. Al termine del processo, l'utente riceverà una notifica via e-mail.

²²⁶ Per comprendere cos'è una rete *backbone*, possiamo paragonarla alla spina dorsale umana; la spina dorsale è la struttura portante del corpo umano e ne collega le diverse parti; nell'ambito delle reti neurali, una rete *backbone* svolge un ruolo analogo. Una rete *backbone* è la componente fondamentale di una rete neurale più ampia, specialmente quando si tratta di eseguire *task* complessi come la visione artificiale (ad esempio, il riconoscimento di immagini o oggetti) o l'elaborazione del linguaggio naturale. Il suo obiettivo principale è quello di estrarre le caratteristiche più rilevanti dai dati di input e impararle. Supponiamo che un modello di IA debba riconoscere un gatto in una foto. La rete *backbone* non si preoccupa di dire "è un gatto", ma si concentra sull'individuazione e comprensione delle caratteristiche chiave del gatto: la forma delle orecchie, la consistenza del pelo, la forma degli occhi, ecc. In seguito, queste caratteristiche "apprese" sono trasferite alle parti successive della rete neurale, che le utilizzano per prendere la decisione finale (cioè, classificare l'immagine come "gatto"). In estrema sintesi, quindi, una rete *backbone* è il "cervello percettivo" di un modello di IA, la parte che si concentra sulla comprensione profonda dei dati di input, prima che qualsiasi decisione finale venga presa. Ne consegue che la sua robustezza è fondamentale per l'intero sistema di *machine learning*.

Una volta ottenuto il *field model*, è possibile testarne l'efficacia sui propri documenti: l'utente deve accedere all'area di lavoro *Desk* e selezionare il documento o le singole pagine che desidera sottoporre al riconoscimento. Ricordiamo che le prestazioni del modello migliorano su documenti che presentano caratteristiche analoghe a quelle del *training set*. Pertanto, qualora l'addestramento sia stato condotto, per esempio, su libri in formato folio contenenti unicamente corpo del testo, numero di pagina e intestazioni, è altamente probabile che il modello produca risultati insoddisfacenti se applicato a libri con intestazioni e piè di pagina. Quindi, l'utente deve aprire il documento su cui intende testare il modello e cliccare sul pulsante *Recognition* nella barra in alto a sinistra dell'editor per accedere al menu dedicato; poi deve selezionare l'opzione *Field* e, all'interno della sezione *Private Model*, individuare il *field model* da lui addestrato.

Se lo ritiene opportuno, l'utente può modificare le seguenti impostazioni avanzate relative al riconoscimento:

- il *detection confidence level* (livello di confidenza di rilevamento) è un parametro, compreso tra i valori 0,5 e 1, che definisce la soglia di sicurezza che il modello deve raggiungere prima di marcare una determinata area di testo;
- il *shape detail level* (livello di dettaglio della forma) è un parametro che determina il grado di precisione nella definizione delle forme delle aree di testo; l'opzione *High* preserva la complessità delle forme, *Low* semplifica le forme in figure basilari e *Medium* offre un equilibrio tra dettaglio e semplicità;
- la funzione *add to existing layout* (aggiungi al layout esistente), consente di aggiungere, ad un layout preesistente, nuovi tag o nuove aree di testo in base alla configurazione del *field model*.

Mentre l'accuratezza dei *text model* è misurata mediante il CER (*Character Error Rate*), quella dei *field model* può essere valutata osservando la percentuale della precisione media. La "precisione media media" (mAP) è una metrica complessa che valuta l'accuratezza con cui il sistema rileva ed etichetta le aree di testo, considerando sia se esse sono state effettivamente

rilevate sia quanto le loro dimensioni e forma corrispondono ai dati di convalida.²²⁷ Modelli con una mAP superiore al 60% possono già fornire risultati molto soddisfacenti. Tuttavia, è sempre consigliabile eseguire un test di riconoscimento con il modello e controllare *de visu* i risultati ottenuti. Inoltre, quando si istruisce un *field model* ad assegnare automaticamente i tag, è consigliabile controllare il numero di occorrenze per ciascun tag, al fine di comprendere la frequenza con cui il modello ha incontrato un dato tag durante l'addestramento; ciò contribuirà a valutare l'efficacia con cui il modello imparerà a riconoscere e applicare quel tag in situazioni nuove. Per concludere, segnaliamo che la piattaforma mette a disposizione degli utenti alcuni *field model* pronti all'uso, tra cui si segnalano *The Baroness of Block* e *The Marginalia Monarch*, per la loro versatilità e ampia possibilità di applicazione. Il primo modello è stato istruito dal team di *Transkribus* a riconoscere i blocchi di testo stampato come un'unica area piuttosto che come paragrafi separati; di conseguenza, *The Baroness of Block* risulta molto utile quando si ha a che fare con documenti che presentano una impaginazione a colonne come i giornali (cfr. Fig. 3.9).²²⁸ Il secondo modello è stato implementato sempre dal team di *Transkribus*, per riconoscere invece le annotazioni manoscritte che frequentemente si trovano ai margini dei documenti storico-letterari. Naturalmente, *The Marginalia Monarch*, necessita di essere affinato di volta in volta per adattarsi alle diverse tipologie di annotazioni a margine; ciò nondimeno, il modello costituisce una valida base di partenza per l'estrapolazione di questo importante genere di informazioni (cfr. Fig. 3.10).²²⁹

²²⁷ La *Precisione Media Media* (mAP), dall'inglese *mean Average Precision*, è una metrica fondamentale e ampiamente usata per valutare le prestazioni dei modelli di IA, in particolare quelli che si occupano del rilevamento di oggetti (*object detection*) o della segmentazione di istanze (*instance segmentation*). Quando un modello di IA deve non solo identificare cosa c'è in un'immagine (ad esempio, "c'è un gatto"), ma anche dove si trova (disegnando un riquadro attorno al gatto), le metriche tradizionali come l'accuratezza semplice non sono sufficienti. La mAP ci permette di valutare contemporaneamente due aspetti cruciali: la correttezza della classificazione (il modello ha identificato l'oggetto giusto?) e l'accuratezza della localizzazione (il riquadro disegnato dal modello attorno all'oggetto è sufficientemente preciso e si sovrappone bene all'oggetto reale?). Per approfondire l'argomento, cfr. il seguente link: <<https://www.ibm.com/docs/it/masv-and-l/maximo-vi/cd?topic=configuring-understanding-metrics>> (ultimo accesso 17 ottobre 2025).

²²⁸ Il modello *The Baroness of Block* è disponibile al seguente link: <<https://app.transkribus.org/models/layout/60963>> (ultimo accesso 17 ottobre 2025).

²²⁹ Il modello *The Marginalia Monarch* è disponibile al seguente link: <<https://app.transkribus.org/models/layout/159405>> (ultimo accesso 17 ottobre 2025).

Figura 3.9 – *The Baroness of Block*

Baroness of Blocks



by Transkribus	Mar 18, 2024
📖 Trained on	Text regions
# Model ID	60963

[See training chart](#)

Model description

The Baroness of Blocks is an AI model specifically trained on 12,000 pages of printed material, designed for the recognition of printed blocks in books and historical prints. As an initial prototype, this model excels in identifying and categorizing printed text blocks within historical documents, making it a valuable tool for researchers and archivists. However, given that it is still in its first iteration, it may not be universally applicable across all types of printed materials. Users are encouraged to test the model on their own datasets to gauge its effectiveness and suitability for specific projects.

Training data

☆
↔

Options ▾

Figura 3.10 – *The Marginalia Monarch*

Marginalia Monarch



by Transkribus	Aug 23, 2024
📖 Trained on	Text regions
📊 Training Set Size	
Pages	5 291
Tags overall	8425
Show details ▾	
📈 Mean Average Precision (mAP)	68.47%
# Model ID	159405

[See training chart](#)

Model description

The Marginalia Monarch is an AI model trained on 5,200 pages with over 8,000 instances of marginalia. This model is designed to detect and analyze handwritten notes and annotations in the margins of historical documents. It is a useful tool for researchers and archivists looking to study these often overlooked elements in historical texts. While effective, users are encouraged to test the model on their own datasets to evaluate its performance and suitability for specific projects.

Training data

☆
↔

Options ▾

3.3 Addestramento di modelli di IA personalizzati per i frammenti autografi

3.3.1 Creazione del campione *ground truth*

Dopo aver illustrato in dettaglio le caratteristiche tecniche del software *Transkribus* e i modelli di governance e business che ne regolano il funzionamento, il presente paragrafo si concentra sull'applicazione pratica dell'interfaccia web a un caso specifico di filologia d'autore: i frammenti autografi di Ludovico Ariosto. L'esperimento aveva come obiettivo lo sviluppo di modelli di intelligenza artificiale per il riconoscimento automatico del layout e della scrittura dei documenti ariosteschi, riutilizzabili non solo per il corpus in esame, ma anche per altri autografi dello stesso autore e, più in generale, per documenti coevi prodotti in un contesto storico-culturale analogo.

Il primo passo del flusso di lavoro ha riguardato la preparazione della *ground truth*. A tale scopo, sono stati caricati sulla piattaforma i facsimili digitali dei frammenti autografi, concessi per finalità di ricerca dalla Biblioteca Ariostea e dalla Biblioteca Ambrosiana, riuniti in una collezione appositamente dedicata. Sulla base delle ipotesi formulate da Santorre Debenedetti (cfr. § 1.3.1), si è proceduto innanzitutto alla ricomposizione virtuale del manoscritto ferrarese (F), correggendo l'ordinamento dei facsimili – laddove necessario – e integrando le carte mancanti. I facsimili forniti dall'Ariostea, infatti, erano stati salvati e numerati in un file secondo una sequenza corrispondente all'attuale disposizione delle carte nel mazzo, dove le cc. 18-19 (num. moderna) risultavano erroneamente collocate al centro del III fascicolo, tra le cc. 13 e 14 (num. autografe). Tali carte sono state pertanto ricollocate nella posizione corretta, ossia tra il I e il II fascicolo, precisamente fra la c. 6 (num. moderna) e la c. M 1 (num. autografa). Inoltre, il testimone F è stato integrato con i facsimili dell'Ambrosiana (M), corrispondenti alla prima e all'ultima carta del III fascicolo, opportunamente inserite nelle rispettive posizioni (cfr. Figg. 1.16 e 1.21).

È stato quindi selezionato un campione rappresentativo di 40 pagine (su un totale di 112), in grado di restituire la varietà del corpus sia dal punto di vista del layout, sia sotto il profilo linguistico e grafico. I frammenti ariosteschi presentano infatti una notevole eterogeneità elaborativa: si passa dai primi abbozzi – carichi di cassature, correzioni e *marginalia*, vergati con una grafia rapida e talora di difficile leggibilità – alle brutte copie, fino alle redazioni più accurate e alle copie in pulito destinate alla tipografia, caratterizzate da una scrittura chiara e regolare. Analogamente a quanto avviene nella costruzione di campioni di *ground truth* per

corpora prodotti da più mani, anche nella definizione del nostro *training set* si è tenuto conto di tale disomogeneità grafica – pur riconducibile a un unico autore – al fine di massimizzare l’accuratezza della trascrizione. La selezione ha seguito un criterio di significatività, volto a garantire una rappresentazione bilanciata dei sette fascicoli che compongono il manoscritto ferrarese (F).

Tabella 3.6 - Composizione del campione *ground truth*

Fascicolo	Carte (num. moderna)
I	cc. 1r. e 1 v. ; cc. 2 r. e 2 v. ; c. 3 r.
II	cc. 7 r. e 7 v. ; cc. 8 r. e 8 v. ; 9 r.
III	cc. 15 r. e 15 v.; cc. 16 r. e 16 v. ; 17 r.
IV	cc. 23 r. e 23 v. ; cc. 24 r. e 24 v. ; 25 r.
V	c. 26 v. ; cc. 27 r. e 27 v. ; cc. 28 r. e 28 v. ; cc. 29 r. e 29 v. ; c. 30 r.
VI	cc. 36 r. e 36 v. ; cc. 37 r. e 37 v. ; 38 r.
VII	cc. 44 r. e 44 v. ; cc. 45 r. e 45 v. ; cc. 46 r. e 46 v. ; c. 47 r.

Nella tabella 3.6 sono riportate le carte selezionate, indicate con la numerazione moderna (quella che negli originali è vergata in blu, in alto a sinistra della pagina), per ciascuno dei sette fascicoli (cfr. Tab. 3.6). Si noterà che da ognuno di essi sono state prelevate cinque pagine, corrispondenti alle carte iniziali dei quinterni (cfr. § 1.3.1), ad eccezione dei fascicoli V e VII, giunti integralmente, dai quali è stato prelevato un numero maggiore di pagine per preservare la coerenza del campione.

In seguito, sulle pagine campione è stata eseguita l’analisi automatica del layout come fase separata dal riconoscimento del testo vero e proprio. Si è già osservato sopra che questa

operazione risulta particolarmente utile in presenza di documenti con strutture complesse; nel nostro caso si trattava di testi disposti su colonne e corredati da numerose annotazioni marginali. Per l'analisi del layout è stato usato il modello predefinito *Universal Lines*, attualmente il più versatile disponibile sulla piattaforma e quindi consigliabile soprattutto nei casi in cui vi sia incertezza sulla soluzione più adeguata alle caratteristiche del materiale in esame. Al termine del processo, i risultati sono stati verificati direttamente sulle immagini: le pagine apparivano suddivise in aree di testo, righe e linee di base distinte. A titolo esemplificativo, si riporta l'immagine relativa alla stanza 8, c. F 1 r. (cfr. Figg. 3.11 e 3.12). In essa, la regione di testo corrisponde al rettangolo verde che racchiude la stanza 8, sia nel facsimile sia nell'editor di testo; la linea di base è evidenziata in blu sotto il verso 2 dell'immagine, e trova corrispondenza nella riga arancione presente nell'editor di testo; infine, il poligono è rappresentato dal rettangolo verde che racchiude la linea di base.

Figura 3.11 – c. F 1 r., esempio di regione di testo

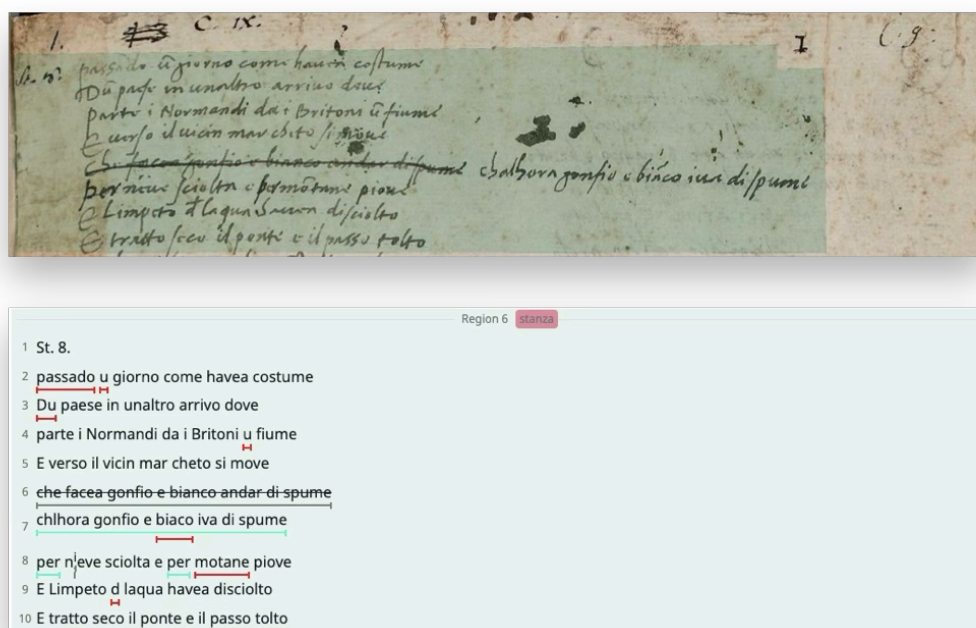
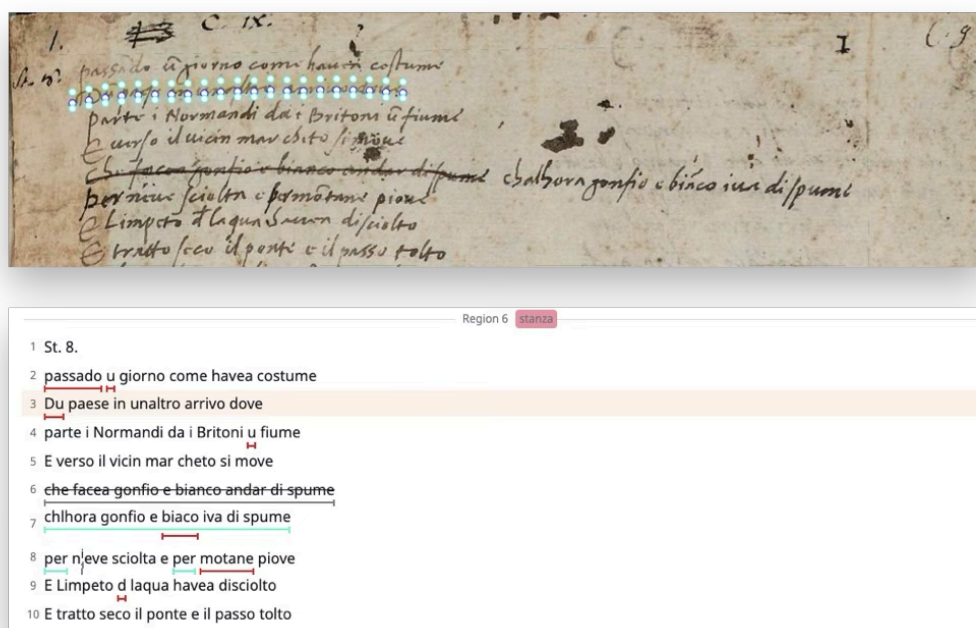


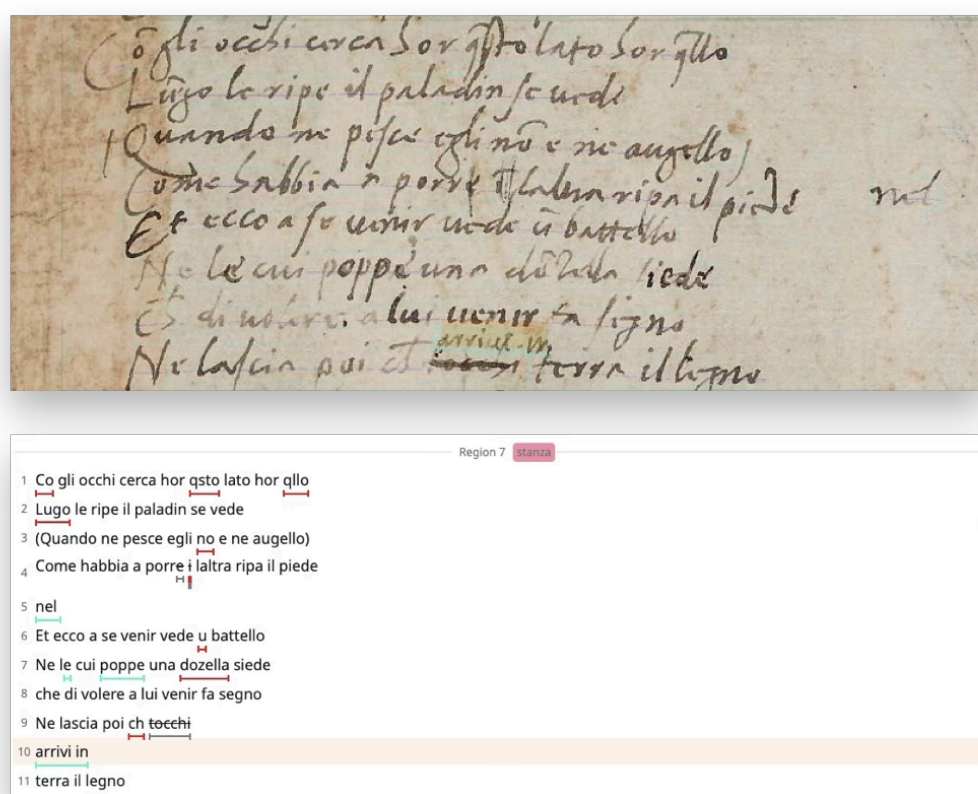
Figura 3.12 – c. F 1 r., esempio di linea di base

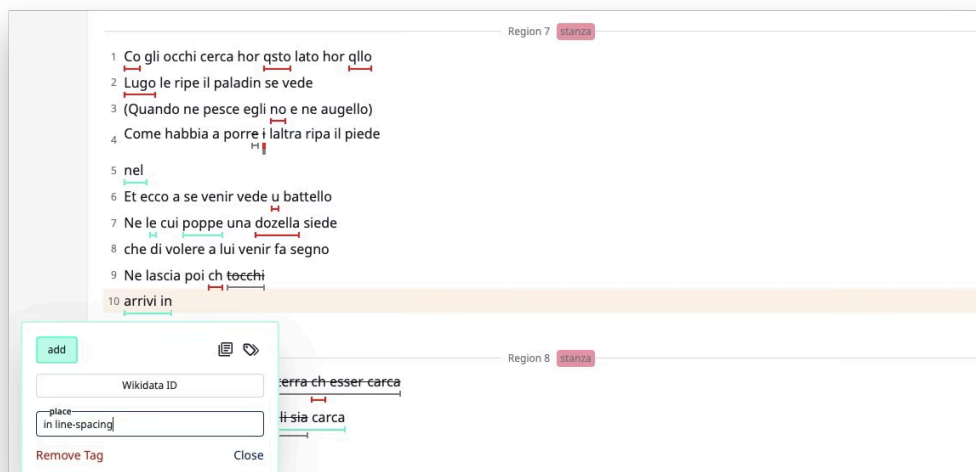


Tuttavia, non sempre l'analisi del layout ha prodotto risultati accurati e soddisfacenti, specialmente negli abbozzi e nelle brutte copie, dove la scrittura si presenta spesso disordinata e di difficile interpretazione a causa delle stratificazioni correttorie. In questi casi, si sono riscontrate linee di base imprecise, righe mancanti o erroneamente raggruppate nelle aree di testo, circostanze che hanno reso necessario un intervento di correzione manuale. Come noto, gli autografi ariosteschi presentano una suddivisione del testo in due colonne: nella colonna di sinistra Ariosto trascriveva le ottave, mentre quella di destra era riservata ad aggiunte e correzioni più lunghe e complesse, che non trovavano posto né in interlinea né nel margine sinistro del foglio. Nella colonna di sinistra le ottave si susseguono senza soluzione di continuità, e l'inizio di ciascuna è segnalato da una lettera maiuscola leggermente sporgente rispetto ai versi. Poiché il software, nella maggior parte dei casi, non ha rilevato la maiuscola iniziale o l'ha segmentata come riga autonoma, è stato necessario correggere manualmente le linee di base (e i poligoni associati, ossia le aree che ne definiscono l'estensione spaziale e i margini) dei versi incipitali delle ottave e procedere alla loro corretta rinumerazione.

In molti casi il software non ha riconosciuto le aggiunte interlineari; quindi, è stato necessario interrompere la linea di base del verso in corrispondenza dell’aggiunta, tracciare una nuova linea al di sotto dell’aggiunta e assegnarle la numerazione corretta. In questo modo, la sequenza numerica delle linee di base poteva proseguire regolarmente dal segmento successivo del verso. Un esempio significativo si trova nella c. F 1 r., st. 9, v. 8, dove Ariosto ha sostituito “tocchi” con “arrivi in”, aggiunto in interlinea (cfr. Fig. 3.13). In questo caso, la linea di base è stata suddivisa in due segmenti, inserendo “arrivi in” tra «Ne lasci poi che tocchi» e «terra il legno»; l’aggiunta è stata poi marcata con il tag *add* e l’attributo *place* con valore “in line-spacing” (editor, righe 9-11).

Figura 3.13 – c. F 1 r., esempio di segmentazione di correzione interlineare





Le correzioni effettuate in rigo sono state ricondotte al verso di riferimento sulla medesima linea di base. Diversamente, per le correzioni o aggiunte marginali è stato adottato un duplice criterio: le parole isolate o i segmenti più brevi sono stati trascritti in una linea autonoma, conservando la numerazione della regione di testo di appartenenza. Un esempio significativo si trova nella stanza 12 (c. F 1 r.). Dopo aver completato la stesura della stanza, Ariosto vi torna per modificare i versi 2 e 3: sostituisce “un scoglio iace” con “l’isola giace” al verso 2, cancella con un taglio il verso 3 e corregge «la qual per legge o ria costuma manda» (riferito a “l’isola giace”); non soddisfatto, elimina nuovamente il verso e lo riscrive come «Nomata Hebuta che per legge manda» nel margine destro del foglio (cfr. Fig. 3.14). Come mostrano le immagini, abbiamo marcato “un scoglio iace” e il verso 3 con il tag *strikethrough*, trascrivendo accanto al verso 2 la correzione “l’isola giace” con il tag *add* e l’attributo *place* con valore “on the right margin” (editor, righe 3-4). Successivamente, abbiamo riportato in una nuova riga la seconda redazione del verso 2, anch’essa contrassegnata con il tag *strikethrough* (editor, riga 5). Infine, abbiamo trascritto separatamente la terza redazione del verso, marcandola con il tag *add* e l’attributo *place* con valore “on the right margin” (editor, righe 6-7). Le aggiunte o le correzioni di maggiore estensione sono state invece raccolte in una sezione di testo separata, numerata in sequenza. Ad esempio, nella c. F 1 r., Ariosto inserisce la stanza 11 nella colonna destra e ne segnala la collocazione, mediante un apposito segno, tra la decima e la dodicesima stanza (cfr. Fig. 3.15).

Figura 3.14 – c. F 1 r., esempio di segmentazione di correzione marginale

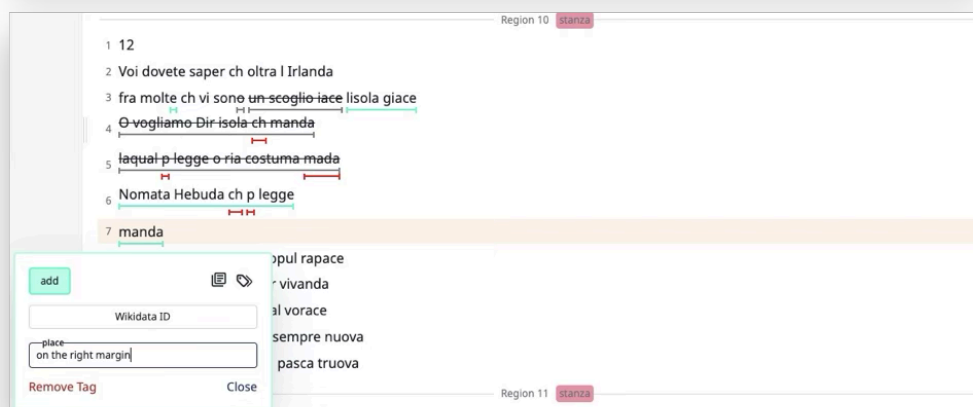
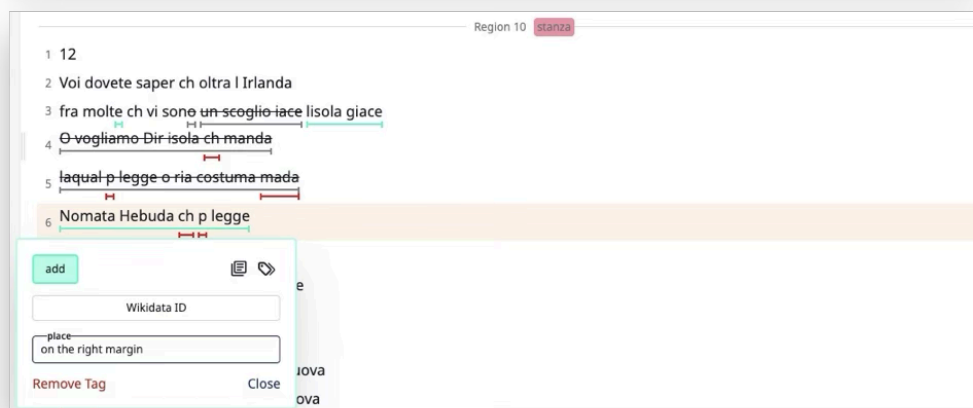
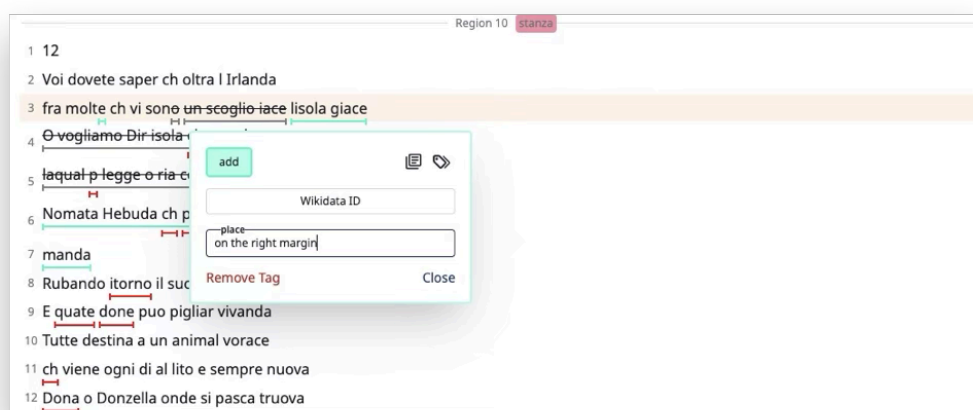
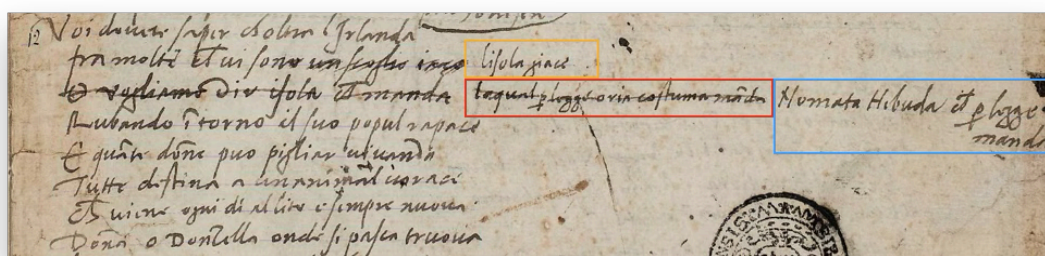
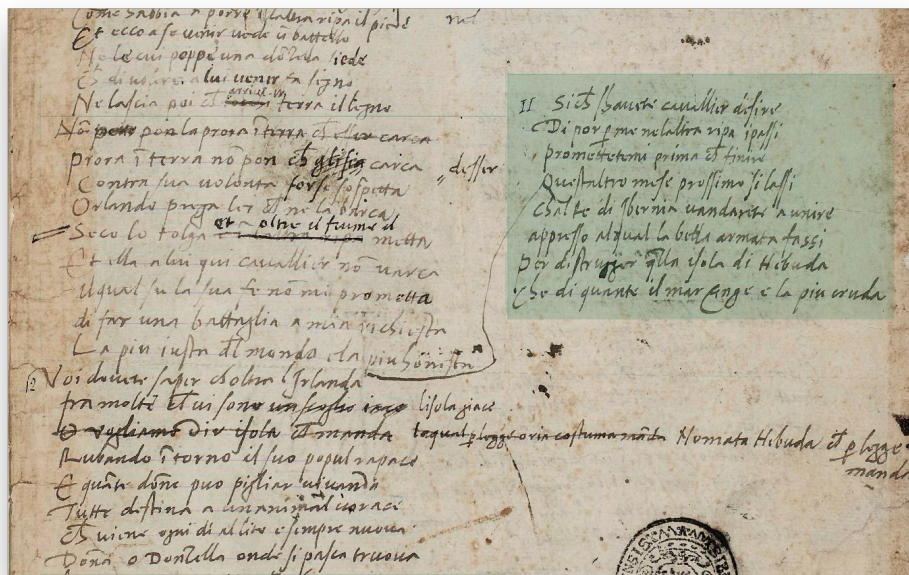


Figura 3.15 – c. F 1 r., esempio di segmentazione di aggiunta marginale come area di testo autonoma



Region 8 stanza

- 1 No mette pon-la prora-i terra-ch-esser-carca
- 2 prora i terra no pon ch gli-sia carca
- 3 desser
- 4 Contra sua volunta forse sospetta
- 5 Orlando prega lei ch ne la barca
- 6 seco lo tolga e+altra-ripa
- 7 et oltre il fiume il
- 8 metta
- 9 Et ella a lui qui cavallier no varca
- 10 ilqual su la sua fe no mi prometta
- 11 di far una battaglia a mia richiesta
- 12 La piu iusta di mondo ela piu honesta

Region 9 added stanza

- 1 II
- 2 Si ch shavete cavallier desire
- 3 Di por p me ne l'altra ripa i passi
- 4 promettetemi prima ch finire
- 5 Questaltro mese prossimo si lassi
- 6 ch al Re di Ibernica vandarete a unire
- 7 appresso al qual la bella armata fassi
- 8 per distrugger qlla isola di Hebuda
- 9 che di quante il mar Cinge e la piu cruda

Region 10 stanza

- 1 12
- 2 Voi dovete saper ch oltra l Irlanda
- 3 fra molte ch vi sono un-scoglio-iace lisola giace
- 4 O-vogliamo-Dir-isola-ch-manda
- 5 laqual-p-legge-o-ria-costuma-mada
- 6 Nomata Hebuda ch p legge
- 7 manda
- 8 Rubando itorno il suo popul rapace
- 9 E quate done puo pigliar vivanda
- 10 Tutte destina a un animal vorace
- 11 ch viene ogni di al lito e sempre nuova
- 12 Dona o Donzella onde si pasca truova

L'analisi del layout dei frammenti autografi sarà approfondita nel paragrafo dedicato all'addestramento del *field model* (cfr. § 3.3.2). Qui basti ricordare che i modelli di HTR si basano principalmente sulle linee di base e sui relativi poligoni: la loro tracciatura accurata è dunque fondamentale per ottenere una trascrizione affidabile. La correzione manuale – che comporta la ridefinizione di ciascuna linea in funzione delle peculiarità del layout – si è tuttavia rivelata un'operazione onerosa, sia in termini di tempo sia di impegno cognitivo. Il processo di creazione della *ground truth* si è concluso con la trascrizione manuale dei testi selezionati. Tenendo conto delle abitudini scritte di Ariosto (cfr. § 1.3.4), si è optato per una trascrizione diplomatica, volta a riprodurre fedelmente la lezione dell'originale, compresi eventuali errori ortografici o lessicali. Tale approccio consente di addestrare un modello di HTR più robusto, in grado di restituire con esattezza ciò che appare nell'immagine. I criteri editoriali adottati – riportati anche nel campo descrittivo del modello all'interno della piattaforma – possono essere così sintetizzati:

- le parole sono state unite o separate secondo la lezione originale;
- le forme sillabate a fine riga sono state trascritte e segmentate come nel testo di partenza; il trattino è stato mantenuto solo se effettivamente presente;
- l'uso di maiuscole e minuscole è stato rispettato quando chiaramente distinguibile; nei casi dubbi si è fatto ricorso alle norme ortografiche correnti;
- la punteggiatura è stata conservata secondo l'uso ariostesco (ad esempio, il punto impiegato in luogo della virgola, o i due punti al posto della virgola), senza introdurre segni mancanti come apostrofi o accenti;
- la distinzione tra *i* e *j* è stata mantenuta, come nel testo originale;
- le lettere *u* e *v* sono state trascritte secondo le norme grafiche contemporanee;
- la *s* normale e la *s* lunga discendente sono state entrambe uniformate alla forma *s* moderna.

Inoltre, poiché le trascrizioni sono destinate a confluire in un'edizione digitale, in questa fase si è proceduto alla marcatura del testo mediante l'impiego di *tag testuali*, sia forniti dalla

piattaforma sia appositamente creati per evidenziare alcune peculiarità dei documenti analizzati.²³⁰ In particolare:

- le cancellature sono state marcate mediante il tag *strikethrough* (cfr. Figg. 3.14 e 3.16);
- le lettere aggiunte in interlinea sono state evidenziate con il tag *superscript* (cfr. Figg. 3.17);
- le parole aggiunte in interlinea sono state etichettate con il tag *add* e l'attributo *place* con i relativi valori *in line*, *in line-spacing*, *on the right margin* e *on the left margin* (cfr. Fig. 3.14);
- le correzioni soprascritte sono state marcate con il tag *add* e l'attributo *place* con il valore "overwritten on word below" (cfr. Fig. 3.18);
- i passaggi illeggibili sono stati contrassegnati con il tag *unclear*, al quale, in alcuni casi, sono stati aggiunti attributi contenenti ipotesi interpretative e l'indicazione della causa della difficoltà di lettura (ad esempio: cancellature, inchiostro sbiadito, macchie o lacerazioni della carta). Nell'esempio riportato, la parola posta sotto la cancellatura risulta di difficile decifrazione (cfr. Fig. 3.19).

Un'attenzione particolare meritano le abbreviazioni, molto frequenti nei frammenti ariosteschi. Le più ricorrenti riguardano:

- le consonanti nasali *m* e *n* (sia all'interno che in fine di parola), rese con un trattino soprascritto (*non* → *no*, *con* → *co*, *bianco* → *biaco*, *donna* → *dona*, *fummo* → *fumo*, ecc.);
- la forma *che*, abbreviata in *ch* (sia come pronome relativo sia come congiunzione);
- la preposizione *per*, resa con *p* e un trattino discendente;
- la vocale *e* in fine di parola;
- gli articoli determinativi composti (*del* → *dl*, *dello* → *dlo*, *della* → *dlla*, *dei* → *d*, *degli* → *dgli*, *delle* → *dle*);

²³⁰ Si ricorda che i tag testuali possono essere utilizzati anche per l'identificazione di entità come nomi di persone, date, luoghi ed eventi, un'operazione che non è stata ancora realizzata ma che ci si auspica di intraprendere in futuro, al fine di offrire agli studiosi strumenti più articolati per l'analisi dei frammenti e per l'elaborazione di ricerche di diversa natura.

- i dimostrativi (*questo/questi* → *qsto/qsti*, *questa/queste* → *qsta/qste*, *quello/quella* → *qlo/qla*, *quelli/quelle* → *qli/qlle*);
- altre forme comuni, come *quale* → *ql*, *quanto* → *qnto*.

Le abbreviazioni sono state trascritte in forma letterale, riproducendo i caratteri più vicini a quelli impiegati dall'autore; sono state inoltre contrassegnate come tali e sciolte successivamente nel campo expansion del relativo tag (cfr. Fig. 20).

Figura 3.16 – c. F 1 r., st. 8, esempio di marcatura con il tag *strikethrough* (verso 5, «che facea gonfio e bianco andar di spume»)

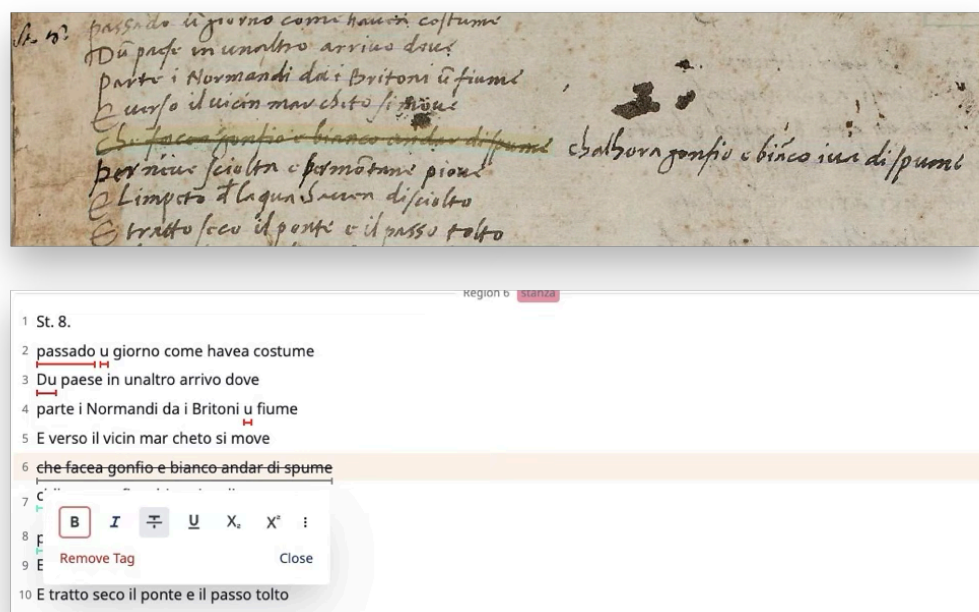


Figura 3.17 – c. F 1 r., st. 8, esempio di marcatura con il tag *superscript* (verso 6, “nieve”)

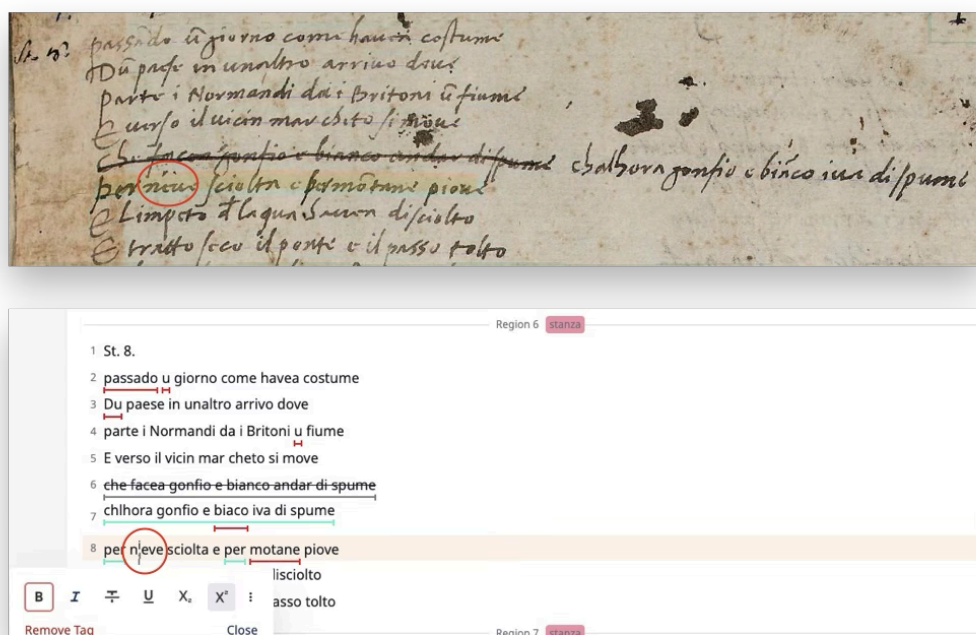


Figura 3.18 – c. F 1 r., st. 8, esempio di marcatura con il tag *add* e l'attributo *place* "overwritten on word below" (verso 6, “Per” soprascritto su “La”)

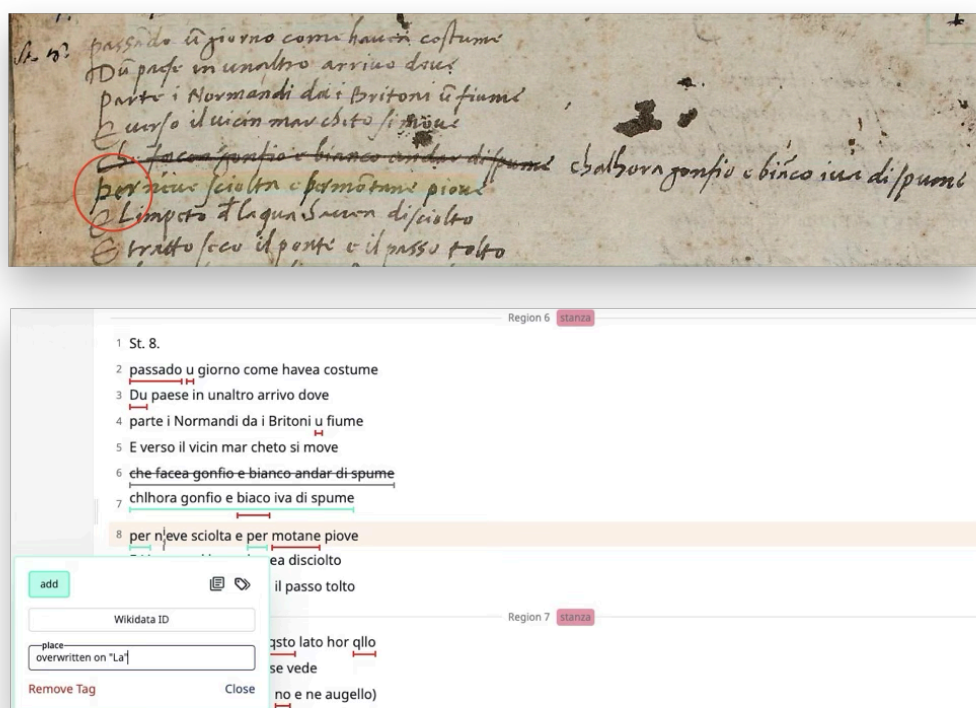


Figura 3.19 – c. F 2 r., st. 23, esempio di marcatura con il tag *unclear* e l'attributo *reason* (verso 1, forse “prusia”)

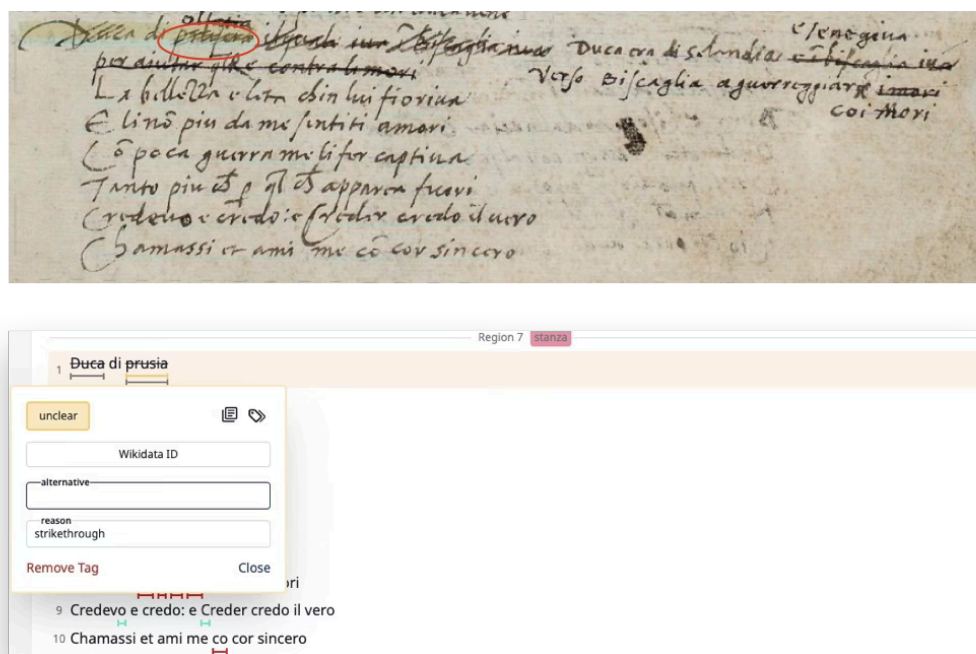
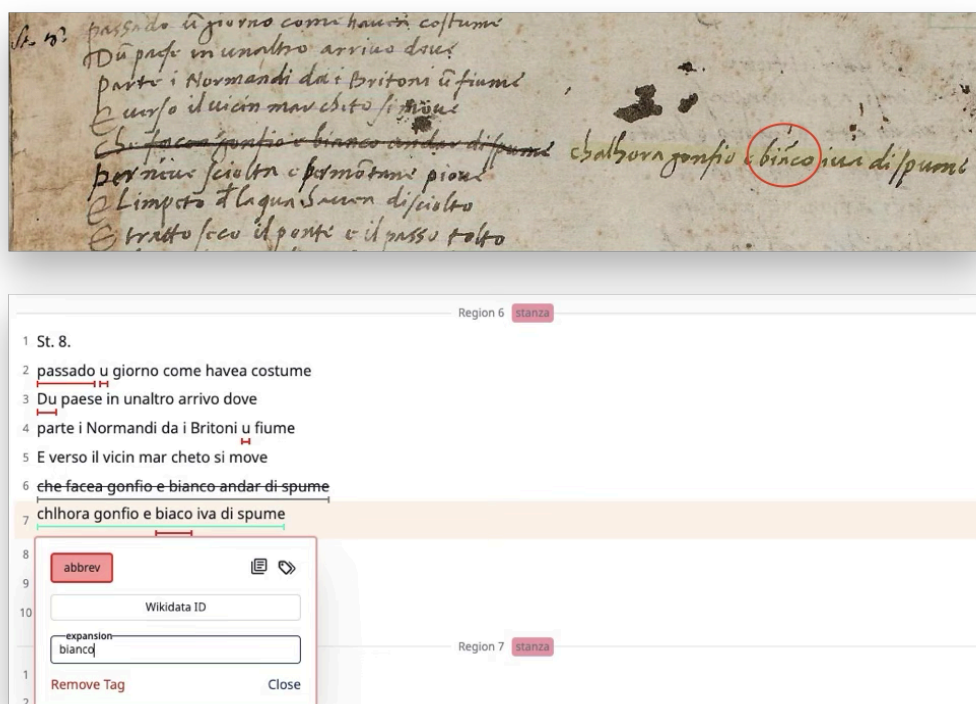


Figura 3.21 – c. F 1 r., st. 8, esempio di marcatura con il tag *abbreviation* e l'attributo *expansion* (verso 6, “biaco” per “bianco”)



3.3.2 Marcatura semantica del documento e addestramento di un *field model* per il riconoscimento automatico della *mise en page* e dei *marginalia* ariosteschi

In precedenza si è osservato che, qualora i documenti presentino una *mise en page* lineare – ad esempio costituita da un unico blocco di testo, organizzato in righe parallele e uniformemente orientate in senso verticale o orizzontale – la piattaforma riesce a svolgere l’operazione di riconoscimento testuale con relativa facilità. La situazione si complica invece quando la disposizione grafica del testo si presenta più articolata (cfr. § 3.2.4). I frammenti autografi, in particolare, offrono un caso emblematico: mentre le belle copie e le versioni definitive destinate alla tipografia tendono a caratterizzarsi per una struttura chiara e regolare, con il testo generalmente collocato in una colonna allineata a sinistra, gli abbozzi e le minute mostrano una configurazione assai più complessa. In questi ultimi, infatti, l’autore sfrutta l’intera superficie della pagina per annotare correzioni, aggiunte e varianti che si susseguono senza soluzione di continuità, sia in orizzontale sia in verticale, dispiegandosi dall’alto verso il basso e viceversa. In simili circostanze, risulta essenziale ricorrere ai cosiddetti *field model*, strumenti avanzati di intelligenza artificiale progettati per individuare automaticamente configurazioni strutturali complesse, comprendenti colonne di testo, *marginalia*, titoli, numerazioni di pagina e ulteriori componenti paratestuali. Così come un modello di riconoscimento testuale può essere istruito a decifrare un particolare sistema di scrittura o stile grafico, un *field model* può essere addestrato a distinguere e classificare i diversi elementi costitutivi di un layout. La costruzione di un *field model* personalizzato richiede innanzitutto la predisposizione del *training set*, vale a dire un corpus di pagine campione che consenta di istruire il modello al riconoscimento degli elementi strutturali specifici. A tale scopo è stato impiegato il campione di pagine descritto nel paragrafo precedente (cfr. § 3.3.1). Dopo aver corretto le linee di base (con i relativi poligoni lineari), si è proceduto a tracciare e marcare tutti gli elementi costitutivi del layout ritenuti significativi per l’edizione genetica digitale, utilizzando sia tag strutturali predefiniti sia tag appositamente elaborati. In particolare, sono stati individuate le seguenti regioni di testo:

- i numeri dei fascicoli che compongono il manoscritto ferrarese, contrassegnati con il tag strutturale *file-number* (cfr. Fig. 3.22);
- i numeri dei canti contrassegnati con il tag strutturale *canto-number* (cfr. Fig. 3.22);
- i numeri di pagina moderni contrassegnati con il tag strutturale *page-number* (cfr. Fig. 3.22);

- i numeri di pagina vergati dall'autore, contrassegnati con il tag *autograph-page-number* (cfr. Fig. 3.25);
- la numerazione autografa che ricorre ogni dieci stanze, contrassegnata con il tag strutturale *ten-stanzas-signature* (cfr. Fig. 3.26);
- la numerazione ricorrente a piè di pagina, sul recto delle cc. 3, 5, 7, 15, 19, 23 e 44, contrassegnata con il tag *signature-mark* (cfr. Fig. 3.25);
- le stanze contrassegnate con il tag strutturale *stanza* (con il relativo numero, quando presente, marcato con il tag *stanza-number*) (cfr. Fig. 3.22);
- le stanze aggiunte nella colonna di destra, contrassegnate con il tag *added-stanza* (cfr. Figg. 3.22, 3.24);
- le seconde, terze e quarte stesure, contrassegnate rispettivamente con i tag *second-draft*, *third-draft* e *fourth-draft* (cfr. Figg. 3.23, 3.24,);
- le postille ariostesche, contrassegnate con il tag *marginalia* (cfr. Fig. 3.24).

Figura 3.22 - c. F 1 r., esempio di pagina segmentata e marcata con i tag strutturali

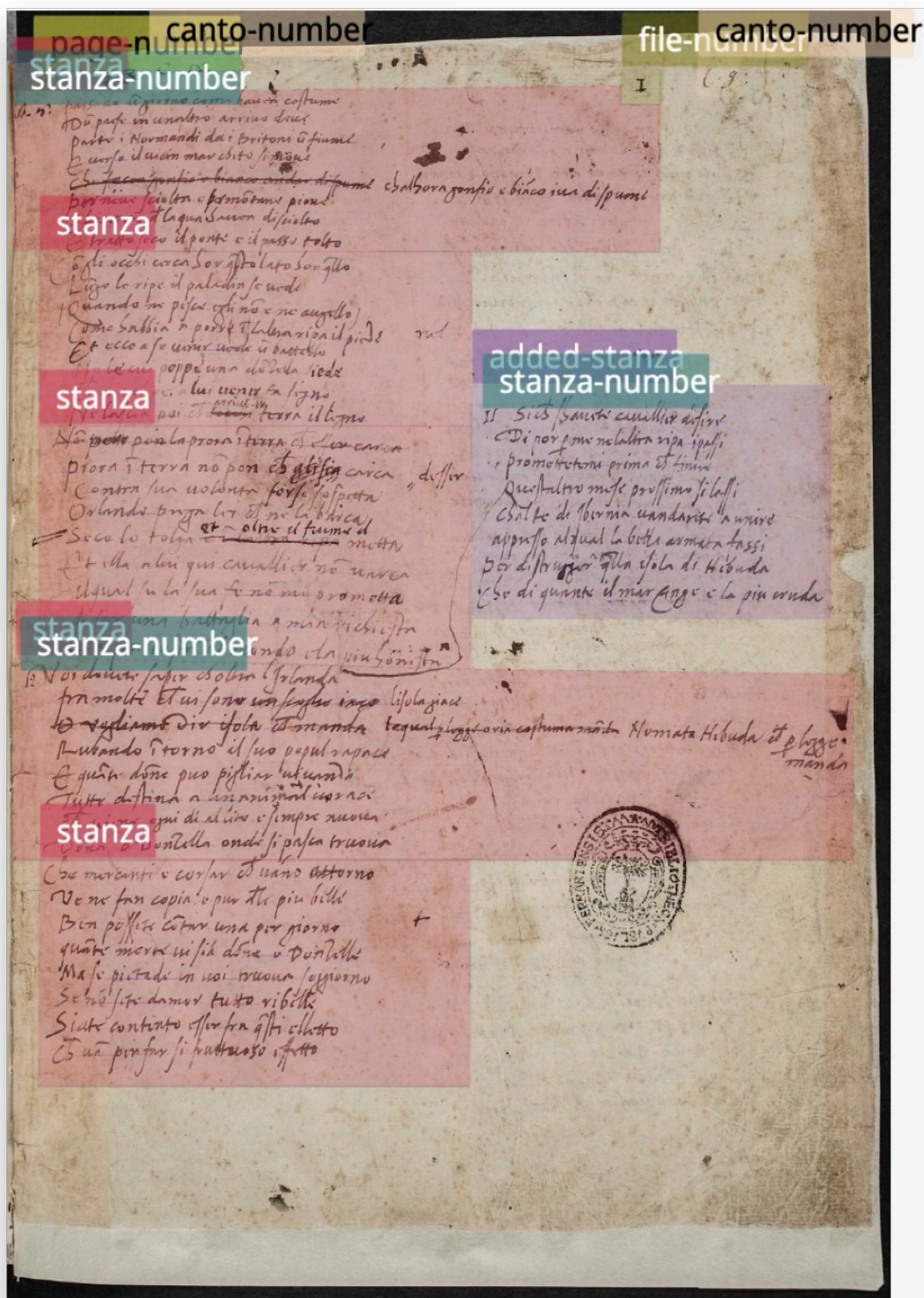
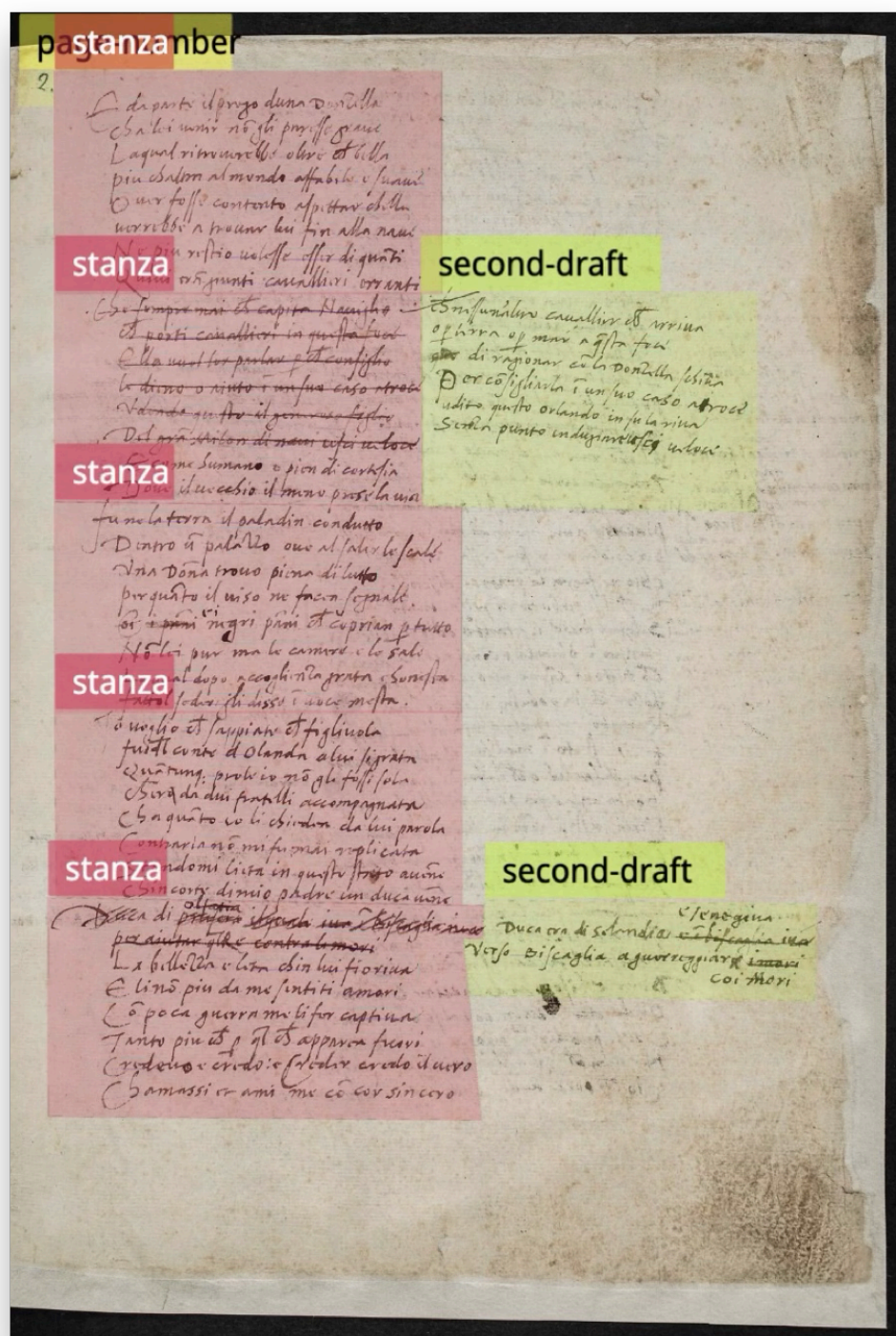
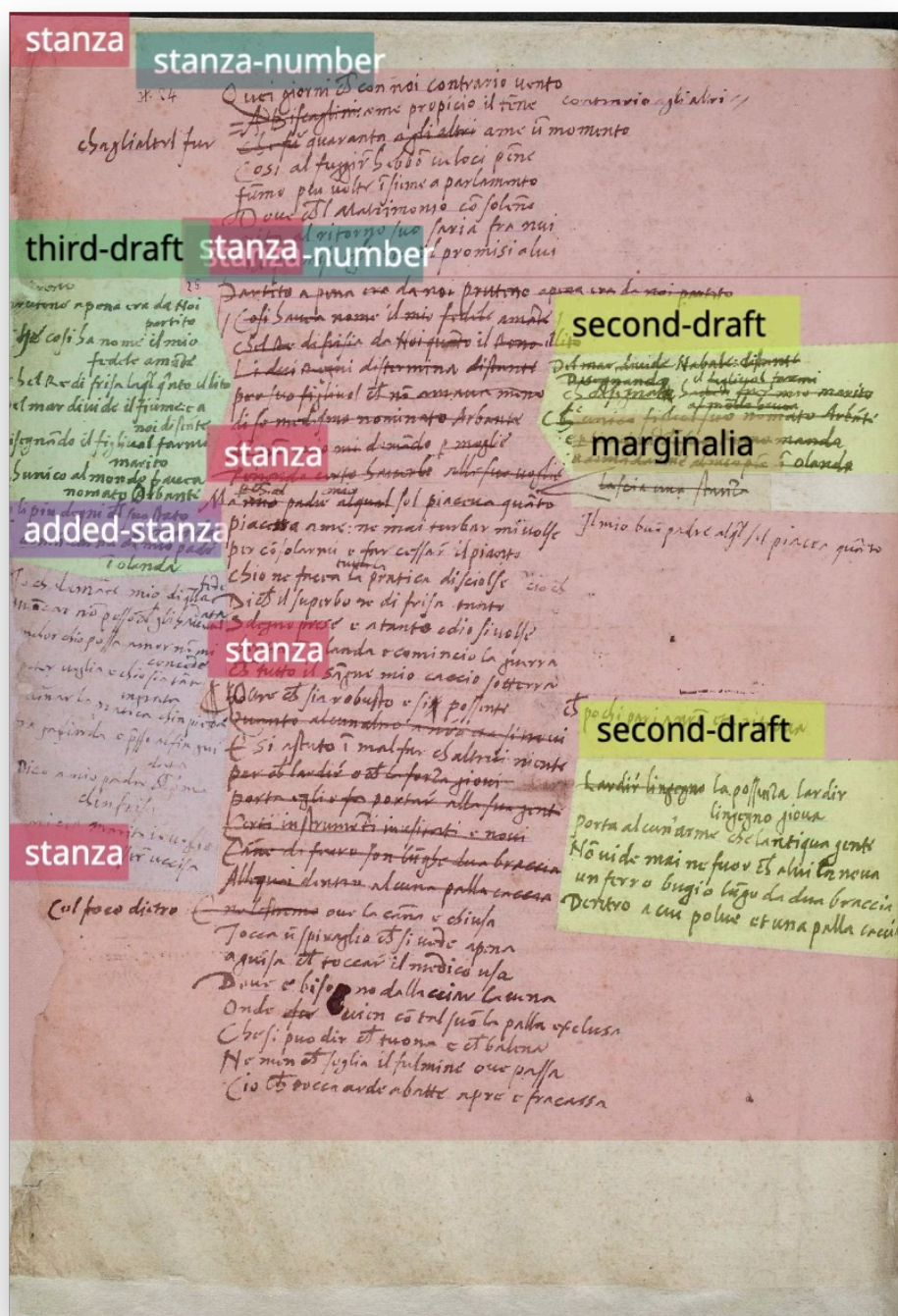


Figura 3.23 - c. F 2 r., esempio di pagina segmentata e marcata con il tag *second-draft*²³¹



²³¹ La carta costituisce la brutta copia della prima parte dell'episodio di Olimpia (canto IX) e contiene le st. 19-23. In particolare, si osserva che, dopo la prima stesura (nella colonna sinistra), Ariosto è tornato sulla st. 20, cancellando le due terzine originarie e riscrivendole nella colonna di destra, dove sono state quindi taggate come *second-draft*. Più in basso, l'autore interviene nuovamente sui versi iniziali della st. 23. Anche questa correzione è stata segmentata in una regione separata e taggata come *second-draft*. In un primo momento, Ariosto scrive «Duca era di Selandia e in Biscaglia iua»; in seguito corregge in «se ne giva verso Biscaglia a guerreggiare i Mori»; infine, Elide la -e della desinenza del verbo e sostituisce «i Mori» con «coi Mori».

Figura 3.24 - c. F 2 v., esempio di pagina segmentata e marcata i tag *second-draft*, *third-draft*, *added-stanza* e *marginalia* ²³²



²³² La carta prosegue la brutta copia della prima parte dell'episodio di Olimpia (Canto IX) e contiene le st. 24-29. In particolare, si osserva che, dopo aver redatto il testo base, Ariosto ha cancellato la st. 25, riscrivendola due volte: dapprima nel margine destro e successivamente nel margine sinistro. Più in basso, l'autore è tornato anche sulla st. 28, riscrivendola nel margine destro. Le seconde e terze stesure sono state segmentate in regioni separate e taggate, rispettivamente, come *second-draft* e *third-draft*. Si nota inoltre che, al promemoria «lascia una stanza» (taggato come *marginalia*), ha fatto seguito l'aggiunta della strofa 26 nel margine sinistro.

Figura 3.25 - c. F 7 r., esempio di pagina segmentata e marcata i tag *autograph-page-number* e *signature-mark*

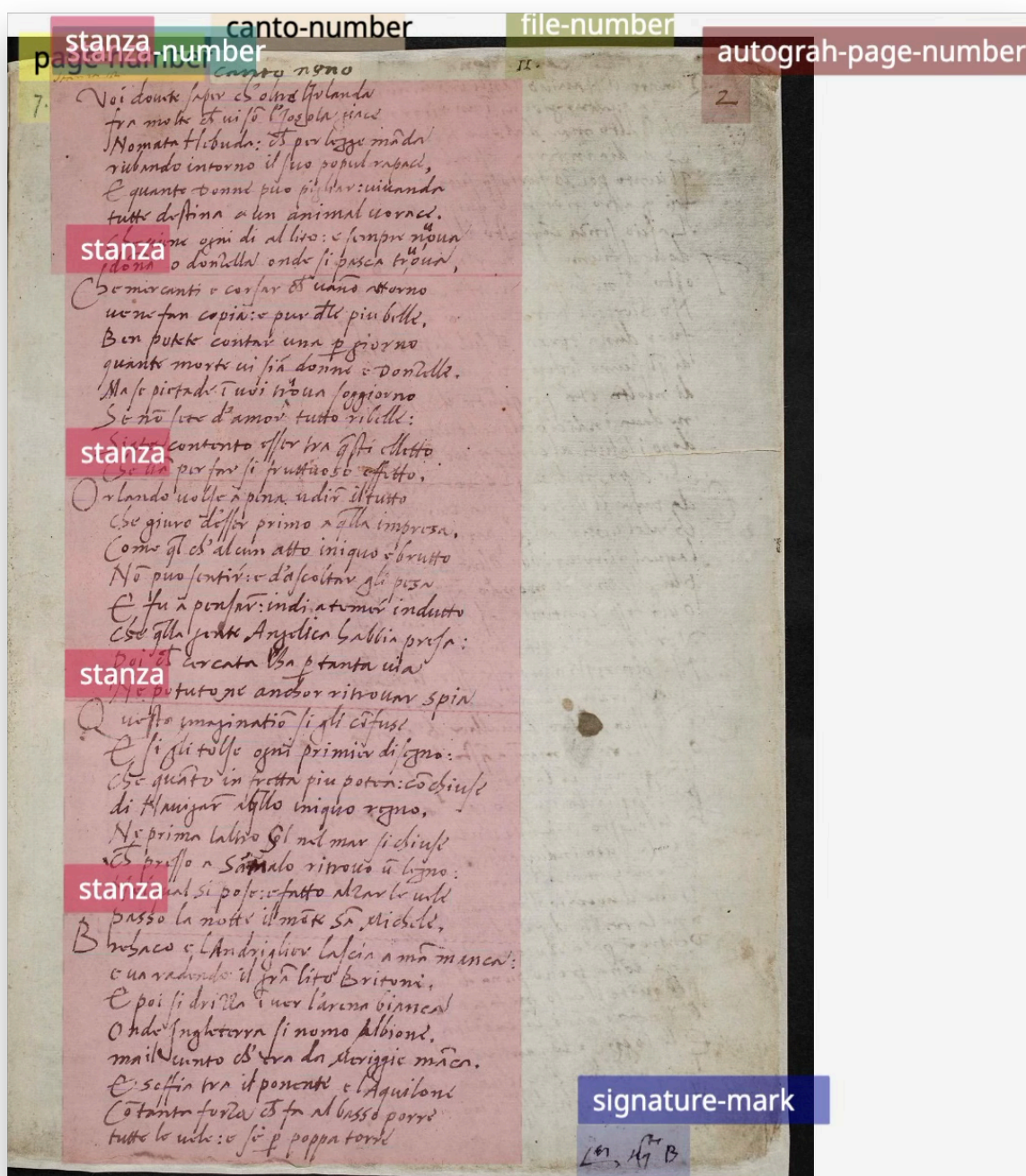
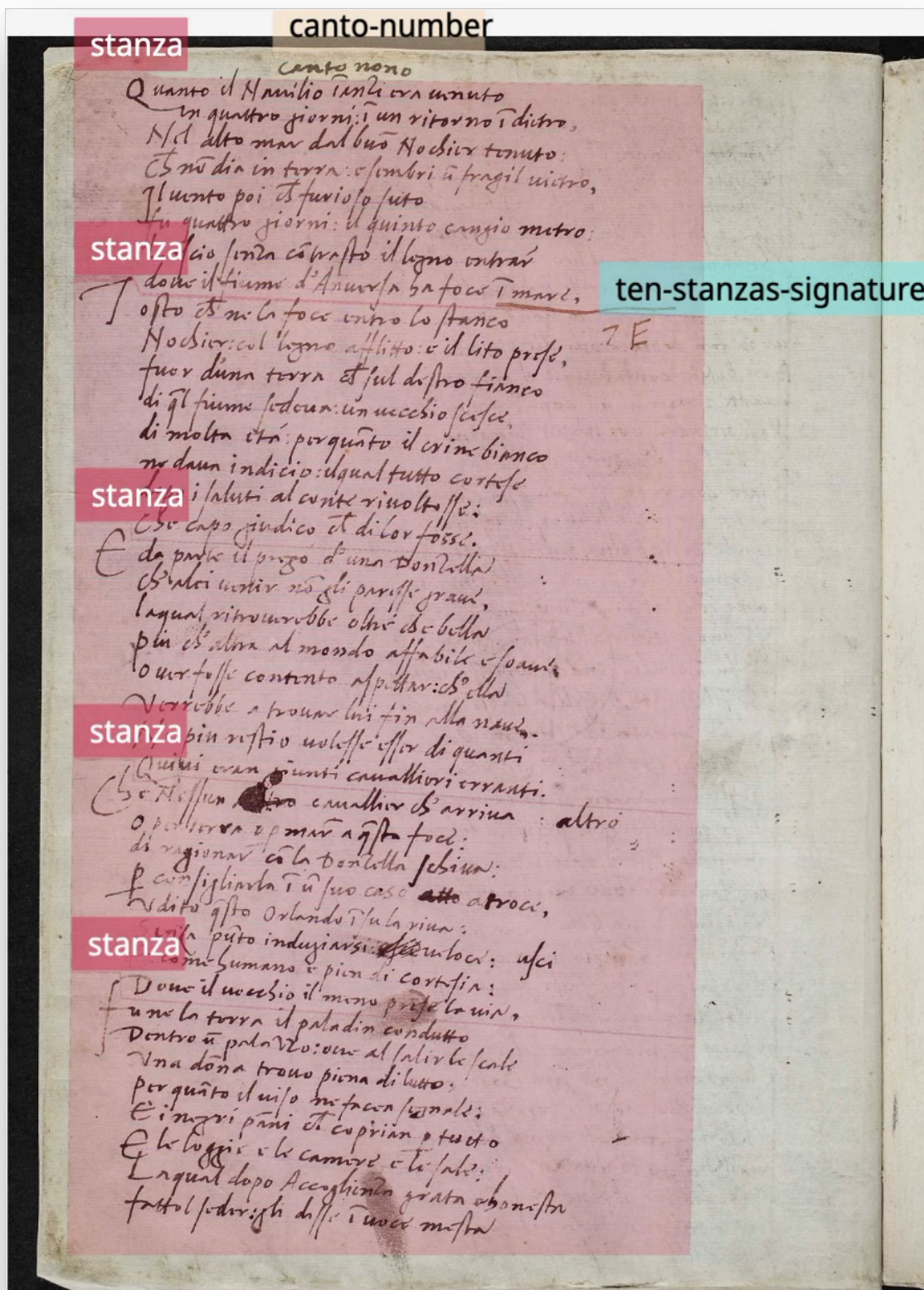


Figura 3.26 - c. F 7 r., esempio di pagina segmentata e marcata il tag *ten-stanzas-signature*



Concluse le operazioni di segmentazione e annotazione, è stata avviata la fase di addestramento del *field model* destinato al riconoscimento delle regioni di testo taggate. Il processo di addestramento si è articolato in cinque passaggi principali, descritti di seguito.

1. Selezione del *training set*, predisposto in precedenza.
2. Avvio dell'addestramento: sono stati selezionati i tag strutturali definiti nelle fasi preliminari, includendo per precauzione anche le regioni non taggate tramite l'apposita opzione.
3. Selezione automatica del *validation set*, costituito da un campione pari al 10% delle pagine del *training set*, in conformità con le raccomandazioni della documentazione ufficiale del software.
4. Configurazione del *field model*. Quest'ultima fase si articola a sua volta in due momenti fondamentali.

In primo luogo, sono state fornite le informazioni identificative del modello:

- Nome del modello (in inglese): *Ludovico Ariosto's Marginalia*
- Descrizione del modello (in inglese):

The field model named was developed with the aim of automating the recognition of the internal structure of the autograph fragments of the *Orlando Furioso*. These fragments consist of notebooks documenting the additions and modifications made by Ludovico Ariosto for the third and definitive edition of the poem (1532), exhibiting varying degrees of textual elaboration. The configuration of the manuscripts varies significantly depending on their stage of advancement. Fair copies and definitive versions, intended for printing, display a clear and regular structure, with the primary text generally arranged in a single, left-aligned column. Drafts and minutes, which represent the more complex elaboration phases, utilize the entire surface of the page. The author intervenes by annotating corrections, additions, and variants that follow one another continuously, both horizontally and vertically, with a layout that unfolds from top to bottom and vice versa. To mark and analyze the complex organization of these fragments, a detailed system of structural tags has been employed. This system allows for the identification of both macro-structural and micro-textual elements, such as quire, page, canto, and stanza numbers, additions, subsequent drafts, and authorial marginalia.

- Anteprima del modello: immagine del ritratto di Ludovico Ariosto attribuito a Tiziano, con relativo URL di riferimento.

Successivamente è stato avviato il processo di addestramento vero e proprio, articolato in diversi tentativi successivi, con l'obiettivo di raggiungere un livello di accuratezza soddisfacente,

valutato sulla base della percentuale di precisione media media (mAP). Si ricorda che la mAP (*mean Average Precision*, ovvero “precisione media media”) è una metrica di valutazione complessa e fondamentale nei modelli di riconoscimento automatico. Essa misura l’accuratezza con cui il sistema rileva ed etichetta le aree di testo, considerando non solo se tali aree sono state effettivamente individuate, ma anche quanto le loro dimensioni e forme corrispondono ai dati di convalida. In altre parole, la mAP valuta contemporaneamente la correttezza del riconoscimento e la precisione geometrica delle predizioni. Questa metrica combina in un’unica misura due indicatori chiave, *precision* e *recall*, offrendo una valutazione complessiva sia della correttezza delle predizioni sia della loro completezza. In termini pratici, una mAP elevata indica che il modello non solo identifica correttamente i tag, ma lo fa anche con un alto livello di confidenza; al contrario, una mAP bassa suggerisce la presenza di numerose predizioni errate o di tag mancanti, segnalando che il modello non ha appreso efficacemente le caratteristiche distintive delle classi.

Inoltre, durante l’addestramento, il comportamento del modello è stato costantemente monitorato attraverso le curve di *loss*, che rappresentano l’andamento della funzione di errore nel tempo. La *training loss* misura la capacità del modello di adattarsi ai dati su cui è stato addestrato, mentre la *validation loss* valuta la sua abilità di generalizzare su dati mai visti prima. Idealmente, entrambe dovrebbero diminuire progressivamente con il procedere delle epoche, indicando che il modello migliora sia nella fase di apprendimento sia nella generalizzazione. Tuttavia, se la *validation loss* si stabilizza o inizia a risalire mentre la *training loss* continua a diminuire, ciò rappresenta un chiaro indicatore di *overfitting*: il modello tende a memorizzare i dati di addestramento senza riuscire a estendere efficacemente le proprie capacità a nuovi esempi. La relazione tra mAP e *loss* non è strettamente lineare, ma può essere considerata inversamente correlata: in generale, quando la *validation loss* diminuisce, la mAP tende ad aumentare, poiché il modello migliora la propria capacità di riconoscere correttamente i tag; viceversa, quando la *validation loss* si stabilizza o cresce, la mAP tende a saturarsi o a diminuire, segnalando che il modello non sta più migliorando la propria accuratezza predittiva.

In definitiva, è possibile schematizzare le seguenti casistiche.

- Quando la *training loss* e la *validation loss* diminuiscono entrambe, il modello mostra un buon apprendimento e una corretta generalizzazione, e la mAP aumenta progressivamente.

- Quando la *training loss* continua a scendere ma la *validation loss* aumenta, si verifica un fenomeno di *overfitting*, e la mAP tende a stabilizzarsi o a ridursi.
- Quando la riduzione della *loss* è troppo lenta o piatta, il modello soffre di *underfitting*, e la mAP rimane bassa.
- Oscillazioni marcate nella *validation loss* possono derivare da un *learning rate* eccessivamente alto o da una gestione instabile dei *batch*, con conseguente instabilità anche nella mAP.

Nel complesso, l'analisi congiunta della mAP e delle curve di *loss* consente di valutare in modo approfondito sia la fase di apprendimento sia la capacità di generalizzazione del modello, fornendo un quadro chiaro dell'equilibrio tra accuratezza, stabilità e robustezza predittiva. Le tabelle che seguono riassumono, per ciascun tentativo di addestramento, i valori assegnati ai diversi parametri, la frequenza di occorrenza dei tag (costante o variabile) e la corrispondente mAP (cfr. Tab. 3.7 e 3.8).

Tabella 3.7 – Parametri di configurazione e risultati comparativi dei tentativi di addestramento del *field model* “Ludovico Ariosto’s Marginalia”

<i>Ludovico Ariosto’s marginalia</i>				
Parametro	Tentativo 1	Tentativo 2	Tentativo 3	Tentativo 4
Learning cycles	15.000	15.000	25.000	30.000
Learning rate	0,001	0,001	0,001	0,001
Depth of the backbone network	Standard (50 livelli)	Enhanced (100 livelli)	Enhanced (100 livelli)	Enhanced (100 livelli)

<i>Ludovico Ariosto's marginalia</i>				
Tag riconosciuti	Frequenza costante di ciascun tag	Frequenza costante di ciascun tag	Frequenza costante di ciascun tag	Frequenza costante di ciascun tag
mAP (mean Average Precision)	28,61%	33,32%	32,17%	34,63%
Note principali	Impostazioni predefinite, nessuna modifica ai parametri avanzati	Aumentata profondità della rete; miglioramento significativo del mAP	Aumentati i cicli di addestramento; leggero calo del mAP	Ulteriore aumento dei cicli; miglior risultato complessivo

Nel primo grafico la perdita (*loss*) diminuisce rapidamente nelle fasi iniziali dell'addestramento, stabilizzandosi poi intorno al 10–15%. Si osservano tuttavia alcuni picchi isolati nella *validation loss* circa a metà del processo, verosimilmente dovuti a fluttuazioni nei dati di validazione. Tali variazioni possono essere spiegate dalla composizione dei *batch*, ossia dei sottoinsiemi di dati utilizzati per aggiornare i pesi del modello a ogni iterazione, che possono presentare livelli diversi di eterogeneità o difficoltà in relazione alle pagine campione considerate. Nel complesso, il modello mostra un andamento convergente, ma potrebbe beneficiare di un ulteriore *fine-tuning*, volto a ottimizzare la regolarizzazione e a rendere più uniforme la convergenza. In questo modo si potrebbe migliorare la stabilità complessiva e la capacità di generalizzazione del modello (cfr. Fig. 3.27). Il secondo grafico presenta un andamento analogo ma più stabile: i picchi della *validation loss* appaiono meno pronunciati e la curva evidenzia una diminuzione più regolare, seguita da un *plateau* attorno al 10%. Tale comportamento suggerisce un *fine-tuning* più efficace, che ha consentito un processo di addestramento più pulito e coerente. Il modello mostra quindi una buona capacità di generalizzazione e i dati di validazione risultano omogenei e ben rappresentativi del dominio di addestramento (cfr. Fig. 3.28).

Figura 3.27 – Primo tentativo di addestramento del *field model* “Ludovico Ariosto’s Marginalia”

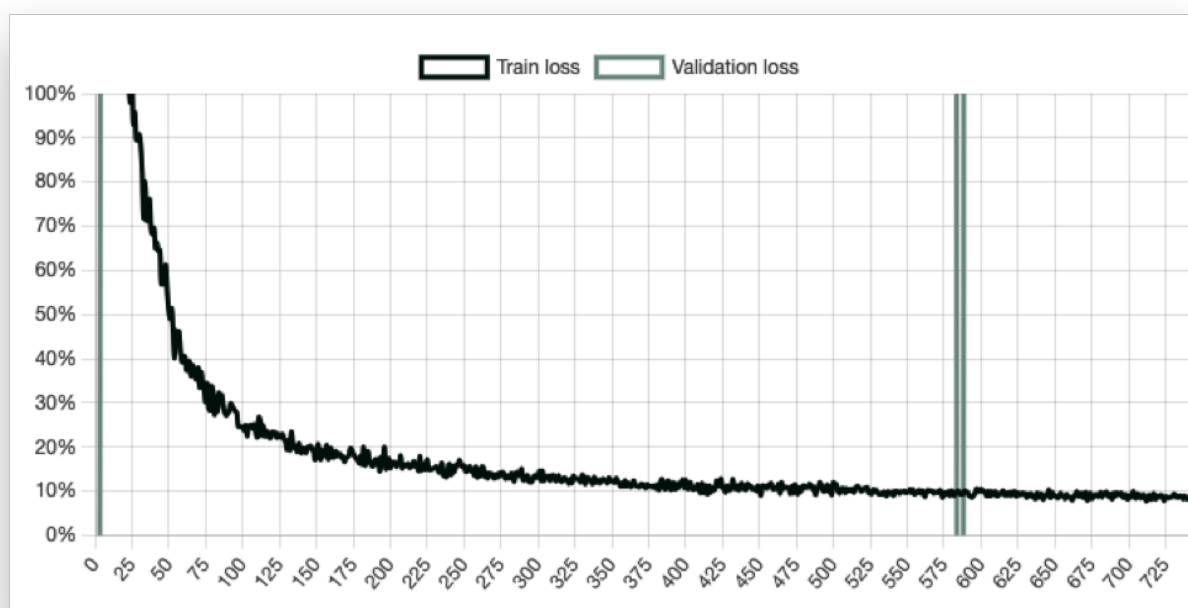


Figura 3.28 – Secondo tentativo di addestramento del *field model* “Ludovico Ariosto’s Marginalia”

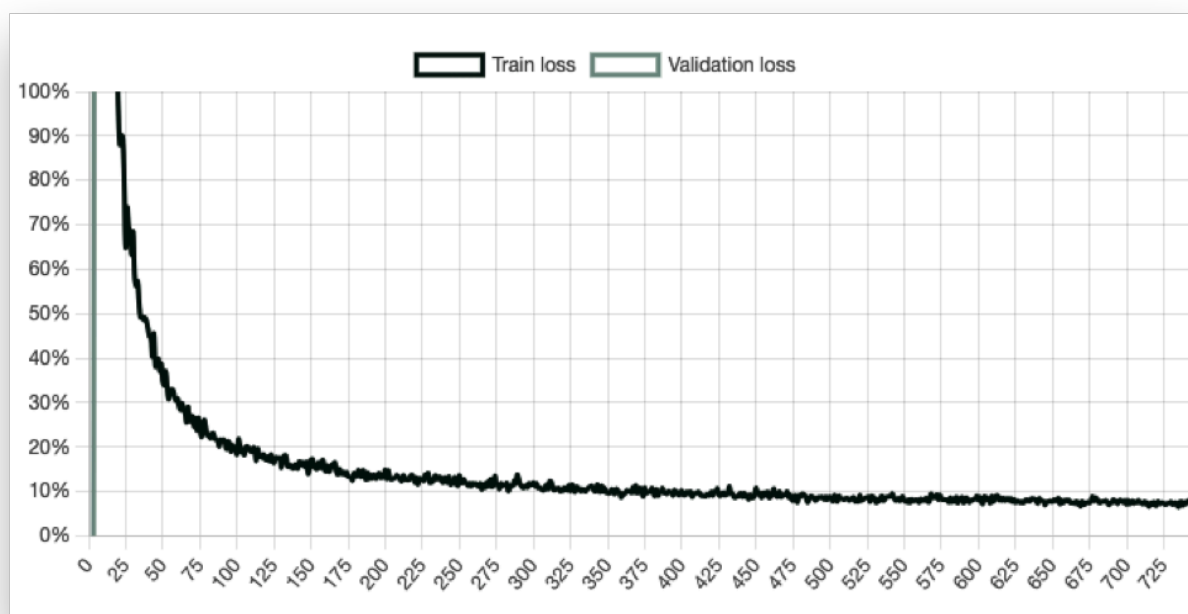


Figura 3.29 – Terzo tentativo di addestramento del *field model* “Ludovico Ariosto’s Marginalia”

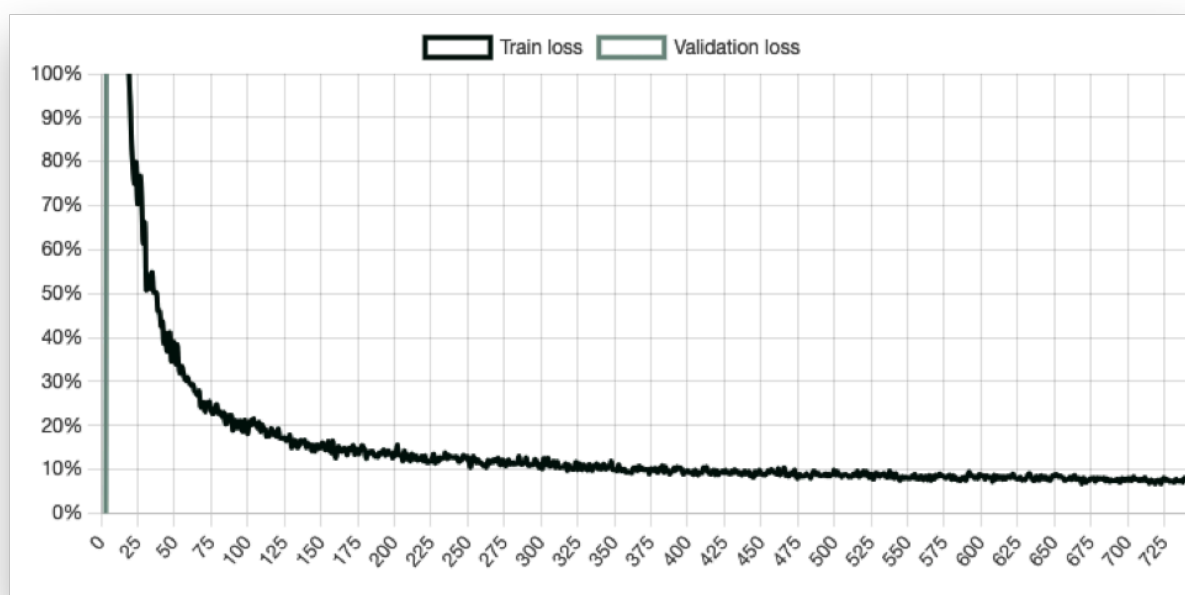
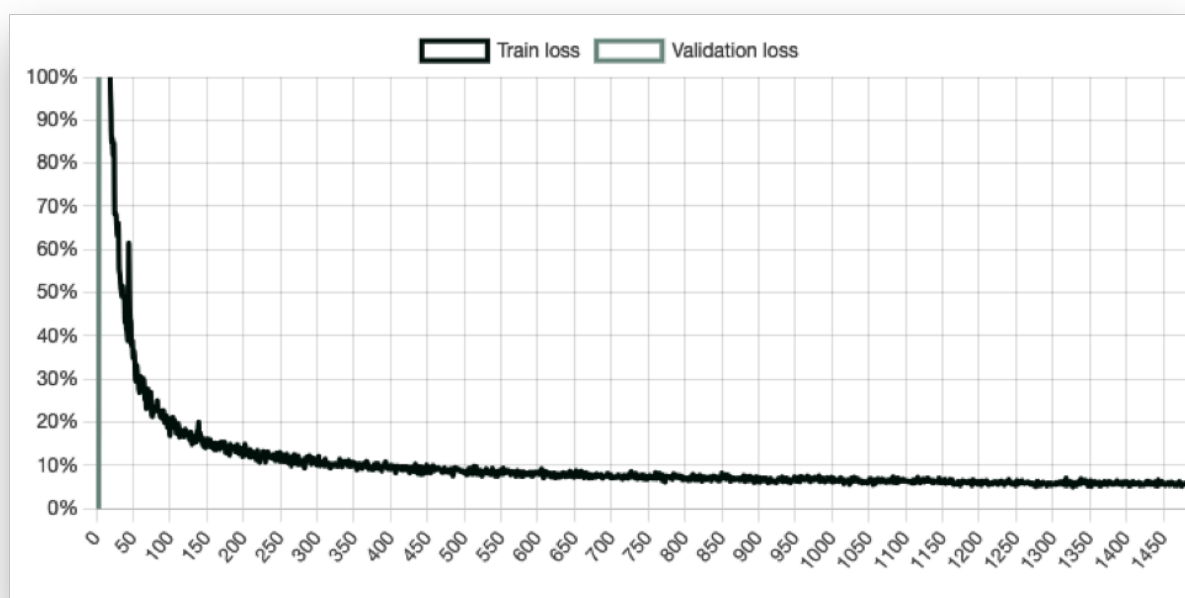


Figura 3.30 – Quarto tentativo di addestramento del *field model* “Ludovico Ariosto’s Marginalia”



Anche nel terzo grafico si osserva un andamento coerente e stabile, con una perdita finale pressoché analoga alle precedenti (circa 10%). Le curve di *training loss* e *validation loss* risultano quasi sovrapposte, un indicatore particolarmente positivo in quanto suggerisce l'assenza di fenomeni di *overfitting*. Nel complesso, questo sembra essere il modello più equilibrato tra i primi tre, presentando un buon compromesso tra capacità di apprendimento e generalizzazione, grazie a un *fine-tuning* ben calibrato (cfr. Fig. 3.29). Nel quarto grafico l'addestramento si estende su un numero maggiore di iterazioni (oltre 1400 epoche), mantenendo tuttavia un andamento coerente e regolare. La perdita diminuisce progressivamente fino a stabilizzarsi su valori molto bassi (inferiori al 10%), senza oscillazioni significative. Questo comportamento riflette un *fine-tuning* più approfondito, che ha permesso una convergenza più fine e stabile rispetto ai modelli precedenti, a fronte però di un tempo di addestramento sensibilmente maggiore. Tale configurazione risulta quindi ideale nei casi in cui il tempo di calcolo non rappresenti un vincolo rilevante (cfr. Fig. 3.30).

Alla luce dei risultati ottenuti, il modello da preferire è il quarto, poiché derivante da un *fine-tuning* più esteso. Esso presenta una convergenza progressiva e stabile, priva di oscillazioni, con una perdita finale inferiore al 10%, indice di un'ottima ottimizzazione dei parametri e di una generalizzazione efficace. Successivamente, il modello prescelto è stato testato per valutarne l'efficacia sui documenti di riferimento, configurando i parametri avanzati di riconoscimento come segue:

- *detection confidence level* (livello di confidenza di rilevamento): impostato sul valore 1, al fine di garantire la massima soglia di affidabilità richiesta al modello prima di marcare una determinata area di testo;
- *shape detail level* (livello di dettaglio della forma): impostato sull'opzione *medium*, per assicurare un equilibrio ottimale tra accuratezza del dettaglio e semplicità della rappresentazione;
- Opzione *add to existing layout* (aggiungi al layout esistente): non attivata.

Tabella 3.9 – Confronto delle prestazioni e della stabilità dei diversi addestramenti del *field model* “Ludovico Ariosto’s Marginalia”

Grafico	Durata training	Stabilità validation	Convergenza	Overfitting	Commento sintetico
1	Media	Media (qualche picco)	Buona	No	Buon risultato, ma leggermente instabile; utile un ulteriore <i>fine-tuning</i> .
2	Media	Alta	Ottima	No	<i>Fine-tuning</i> efficace, training pulito e coerente.
3	Media	Molto alta	Ottima	No	Miglior equilibrio complessivo, <i>fine-tuning</i> ben calibrato.
4	Lunga	Molto alta	Eccellente	No	Miglior risultato complessivo grazie a <i>fine-tuning</i> approfondito, ma con costo computazionale elevato.

Al termine del processo di riconoscimento è stata condotta una verifica qualitativa dei risultati, che ha confermato la buona riuscita della procedura e la corretta integrazione del modello all’interno della struttura dei documenti. Tuttavia, in alcuni casi – ad esempio quando la scrittura di Ariosto non risulta perfettamente allineata in senso orizzontale, ma tende a disporsi in modo obliquo, sconfinando nell’area della stanza successiva – si è reso necessario intervenire manualmente per correggere la segmentazione, includendo i versi nella stanza di appartenenza. Inoltre, i tag scarsamente rappresentati nel campione non sono stati riconosciuti, un limite che potrebbe essere superato, in futuro, ampliando il *training set* e incrementando la varietà delle occorrenze per ciascuna classe.

3.3.3 Addestramento di un modello di HTR personalizzato per i frammenti autografi

Dopo aver completato l'applicazione del *field model* personalizzato ai documenti in esame, è stata avviata la fase di addestramento del modello di HTR, che costituisce il fulcro dell'intero processo. Durante la fase di riconoscimento testuale, le *baselines* e i relativi poligoni lineari vengono generati automaticamente dal sistema. In alcuni casi, tuttavia, questa procedura può produrre risultati imprecisi, compromettendo la corretta segmentazione delle righe di testo. Infatti, i *layout models* si basano sul presupposto che il testo sia distribuito su righe orizzontali regolari, con oscillazioni minime. Ad esempio, le maiuscole ingrandite o ornate – che spesso si estendono verso l'alto o verso il basso oltre l'altezza media della scrittura, oppure sporgono lateralmente nel margine sinistro, come negli autografi ariosteschi – interrompono questa linearità grafica. Di conseguenza, l'algoritmo di *layout analysis* può incontrare difficoltà nel collocarle correttamente sulla stessa *baseline* del resto della riga, generando:

- linee di base frammentate o errate;
- esclusione parziale della lettera o dell'intera riga.

Tali errori di segmentazione si ripercuotono sulla successiva fase di riconoscimento testuale, riducendo la precisione complessiva dell'HTR. Per questo motivo si è deciso di condurre l'analisi delle *baselines* come operazione separata. In un primo momento è stato testato il modello pubblico *Universal Lines*, effettuando numerosi tentativi di riconoscimento nei quali sono stati variati i valori dei seguenti parametri avanzati:

- *minimal baseline length*: definisce la lunghezza minima (in pixel) che una linea di base deve avere per essere considerata valida; valori troppo bassi possono generare falsi positivi, mentre valori troppo alti rischiano di escludere linee brevi ma corrette;
- *baseline accuracy threshold*: imposta la soglia di accuratezza necessaria affinché una linea di base venga accettata dal sistema; un valore più elevato incrementa la precisione, ma può ridurre il numero complessivo di linee rilevate;
- *use trained separators*: abilita o disabilita l'utilizzo dei *separatori addestrati*, strumenti basati su modelli preesistenti che facilitano la distinzione tra righe o colonne di testo;

- *max-dist for merging baselines*: stabilisce la distanza massima (in pixel) entro la quale due linee di base vicine vengono considerate appartenenti alla stessa riga e quindi unite.

I test hanno fornito risultati insoddisfacenti e, di conseguenza, è stato addestrato un *baseline model* specificamente progettato per la tipologia di documenti in esame. A tal fine è stato utilizzato il *training set* predisposto in precedenza, mentre un *validation set* pari al 10% del campione manuale è stato selezionato automaticamente tramite l'apposita funzione del sistema. Sono stati inoltre assegnati i seguenti dati identificativi:

- Nome (in inglese): *Autograph Fragments Baselines*
- Descrizione (in inglese):

This model is designed for the recognition of baselines in the autograph fragments of the Orlando Furioso. It has been specifically trained to detect the capital initials that mark the beginning of each octave, which extend into the left margin of the page compared to the other lines.

I parametri avanzati sono stati mantenuti secondo la configurazione di *default* (cfr. Tab. 3.10).

Tabella 3.10 - Configurazione dei parametri di addestramento avanzati del *baseline model* personalizzato

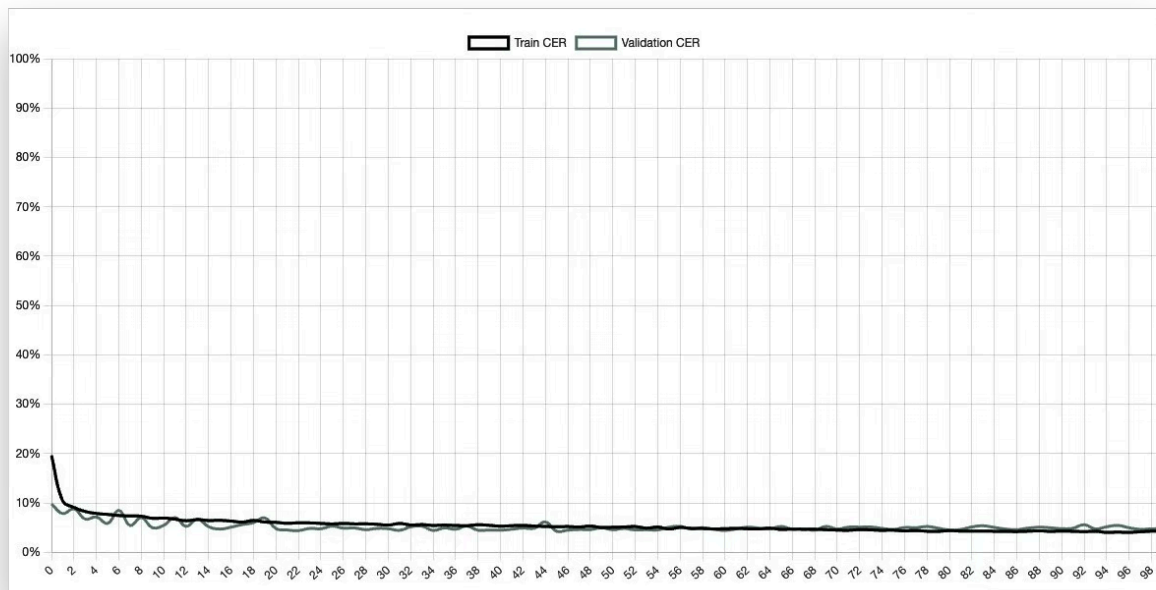
<i>Autograph Fragments Baselines</i>		
Parametro	Descrizione	Valore
Training Cycles	Numero totale di cicli di addestramento del modello.	100
Learning Rate	Velocità di apprendimento, ossia l'entità con cui il modello aggiorna i propri pesi durante l'addestramento.	0.001
Scaling Adjustment	Definisce l'intervallo di ridimensionamento casuale delle immagini, utile per aumentare la variabilità <i>training set</i> e migliorare la capacità di generalizzazione del modello.	min 0.2 / max 0.5

<i>Autograph Fragments Baselines</i>		
Validation	Percentuale del dataset utilizzata per la validazione interna durante l'addestramento, al fine di monitorare la qualità del modello e prevenire fenomeni di overfitting.	0.33
Tilt & Shift	Applica una leggera distorsione geometrica alle immagini (rotazione o traslazione minima) per simulare inclinazioni o spostamenti del testo, migliorando la robustezza del modello.	0.025
Wobble Effect	Disattiva l'effetto di oscillazione artificiale delle linee, che può essere usato in altri contesti per aumentare la variabilità dei tratti grafici.	Non attivo
Image Rotation	Disattiva la rotazione casuale delle immagini, mantenendo il testo nella sua posizione originale.	Non attivo
Thin or Thicken Lines	Parametro che regola lo spessore delle linee di testo: un valore pari a 1 mantiene lo spessore originale, senza alterazioni.	1

Al termine del processo di addestramento, il modello risultante ha mostrato prestazioni complessivamente soddisfacenti, come evidenziato dal grafico allegato (cfr. Fig. 3.31). Quest'ultimo illustra l'evoluzione del tasso di errore durante le diverse epoche di addestramento, distinguendo tra la curva del *Train CER* (in nero) e quella del *Validation CER* (in grigio chiaro). L'osservazione del grafico consente di rilevare una rapida diminuzione del CER nelle prime epoche – da valori iniziali intorno al 18–20% fino a circa il 5–6% – segno di un apprendimento efficace sin dalle fasi iniziali. Dopo circa la decima epoca, entrambe le curve si stabilizzano con oscillazioni minime, indicando che il modello ha raggiunto una fase di convergenza. La sostanziale sovrapposizione tra i valori di *training* e di *validation* suggerisce inoltre che non si

siano verificati fenomeni di *overfitting*, e che il modello sia in grado di generalizzare correttamente anche su dati non inclusi nel set di addestramento. Il mantenimento di un CER di validazione costantemente basso (intorno al 5–7%) conferma, in tal senso, una buona capacità di riconoscimento delle linee di base.

Figura 3.31 – Addestramento del *baseline model* “Autograph Fragments Baselines”



La mAP, che misura la coerenza complessiva delle predizioni rispetto al *ground truth*, costituisce un ulteriore indicatore della qualità del modello. Sebbene il valore registrato (pari al 4,78%) risulti contenuto, esso deve essere interpretato alla luce della complessità grafica dei documenti analizzati: la presenza di maiuscole sporgenti, irregolarità di impaginazione e discontinuità nelle linee dei versi rende infatti più difficoltosa una segmentazione perfettamente coerente con l’originale. Nonostante ciò, il comportamento complessivo del CER e la stabilità delle curve di apprendimento indicano un modello solido, coerente e ben adattato alle caratteristiche del corpus ariostesco.

In conclusione, il modello mostra un apprendimento stabile e una buona capacità di generalizzazione, rivelandosi in grado di gestire efficacemente le peculiarità grafiche dei manoscritti analizzati. Pur con margini di miglioramento nella precisione assoluta (mAP), esso rappresenta un risultato significativo nel processo di ottimizzazione del riconoscimento delle linee di base nei frammenti autografi. Per incrementare in futuro il valore della mAP, sarà

opportuno ampliare e diversificare il *training set*, includendo un numero maggiore di esempi rappresentativi delle varianti grafiche e delle irregolarità di impaginazione. Inoltre, sarà utile introdurre tecniche di *data augmentation* (come leggere rotazioni o variazioni di scala) e perfezionare i parametri di addestramento – in particolare il *learning rate* e il numero di *training cycles* – così da migliorare la capacità del modello di riconoscere correttamente le linee di base anche nei casi più complessi.

Dunque, il *baseline model* addestrato stato applicato sull'intera collezione di documenti, avendo cura di selezionare l'opzione *Keep existing text regions*, al fine di preservare la segmentazione del layout precedentemente ottenuta mediante il *field model* personalizzato. I parametri avanzati sono stati configurati come segue:

- *Minimal baseline length*: low (10 px)
- *Baseline accuracy threshold*: medium (55 px)
- *Use trained separators*: sometimes (5 px)
- *Max-dist for merging baselines*: medium
- Split lines on region border: non attiva
- *Line overlap fraction*: 30%

In particolare, il parametro *Line overlap fraction*, impostato al valore del 30%, determina la soglia di sovrapposizione verticale ammessa tra due linee di testo affinché queste siano considerate appartenenti alla medesima regione. In altri termini, qualora due linee si sovrappongano per una porzione pari o superiore al 30% della loro altezza, esse vengono interpretate come parte della stessa area testuale; viceversa, una sovrapposizione inferiore a tale soglia induce il sistema a trattarle come linee appartenenti a regioni distinte. Tale parametro riveste un ruolo significativo nella gestione di documenti caratterizzati da righe tra loro ravvicinate o da una struttura grafica complessa, poiché consente di modulare la sensibilità del modello nella distinzione tra blocchi di testo contigui. In questo modo si riduce il rischio di errori di segmentazione, garantendo una più accurata identificazione delle aree testuali e, di conseguenza, una migliore qualità del riconoscimento complessivo.

Infine è stato addestrato il modello di HTR. L'addestramento si è articolato in cinque passaggi fondamentali:

1. selezione del *training set* predisposto in precedenza;

2. selezione automatica del *validation set*, mediante la scelta casuale di un campione pari al 10% delle pagine del *training set*, in conformità con le raccomandazioni della documentazione ufficiale del software;
3. configurazione del modello, articolata a sua volta in alcuni sottopassaggi.

In primo luogo, sono state inserite le informazioni identificative necessarie per agevolare l'individuazione del modello in eventuali ricerche future:

- Nome (in inglese): *Ludovico Ariosto's Handwriting*
- Descrizione (in inglese):

The model was trained on a selection of autograph manuscripts by Ludovico Ariosto that preserve several episodes later incorporated into the third edition of the *Orlando Furioso*. The texts contained in these documents display various stages of elaboration, ranging from initial drafts to intermediate copies and fair copies intended for print. The manuscripts date from the period between the poem's second edition (1521) and its third (1532). For the preparation of the ground truth, a diplomatic transcription was adopted, developed according to the following editorial conventions: words were joined or separated according to the original reading; syllabified forms at the end of a line were transcribed and segmented as in the source text; hyphens were retained only when actually present; the use of uppercase and lowercase letters was respected whenever clearly distinguishable, while in doubtful cases modern orthographic conventions were applied; punctuation was preserved according to Ariosto's original usage (for example, the period used in place of the comma, or the colon instead of the comma), without introducing missing marks such as apostrophes or accents; the distinction between *i* and *j* was maintained, as in the original text; the letters *u* and *v* were transcribed according to modern graphic conventions; both the short *s* and the long descending *s* were normalized to the modern *s* form; abbreviations were transcribed using characters most similar to those employed by the author and subsequently expanded using the appropriate tag.

- Lingua dei documenti: italiano del XVI sec.
- Arco cronologico dei documenti: 1521-1532

In secondo luogo, è stato selezionato come *base model* il modello di HTR denominato *Transkribus Italian Handwriting M1*, tra quelli pubblicamente disponibili sulla piattaforma.²³³ Questo approccio ha consentito di ridurre il numero di pagine di *ground truth* necessarie, di accelerare il processo di addestramento e, al contempo, di migliorare la qualità complessiva dei

²³³ REED-COOP, *Should I use a base model when training a custom model?*, «Transkribus blog», 7 febbraio 2025, consultabile al seguente link: <<https://blog.transkribus.org/en/using-base-model-when-training-a-custom-model>> (ultimo accesso 17 ottobre 2025).

risultati di riconoscimento.²³⁴ Il *base model* prescelto è stato addestrato su un *training set* costituito da 4.652 pagine e un *validation set* di 655 pagine, per un totale di 653.630 parole e 277.202 righe, con un CER pari al 6,70%. Il modello è descritto come segue:

Italian handwriting, 16th-19th century. This is a generic model trained on a diverse dataset. Such model provide good results without the need for any extra training work. However, the best results can usually be achieved by training a special model for homogenous material, e. g. texts written by the same person or from a narrow historical period. Curated by the Transkribus team, this model is occasionally updated with community data for continuous improvement.

Successivamente è stata avviata la fase di addestramento del modello di HTR, articolata in diversi tentativi successivi con l'obiettivo di raggiungere un livello di accuratezza soddisfacente nel riconoscimento della scrittura ariostesca. Si ricorda che l'accuratezza del modello è stata valutata attraverso il CER, indice che misura la divergenza tra la trascrizione di riferimento (*ground truth*) e il testo riconosciuto automaticamente, esprimendo la percentuale media di caratteri errati. Il valore è calcolato sul *validation set*, utilizzato per verificare la capacità di generalizzazione del modello. Segue una tabella riassuntiva dei valori assegnati ai diversi parametri avanzati nel corso dei vari esperimenti di fine-tuning (cfr. Tab. 3.11).

²³⁴ REED-COOP, 4 AI models for transcribing old Italian handwriting and typewritten texts, «Transkribus blog», 27 novembre 2024, consultabile al seguente link: <<https://blog.transkribus.org/en/4-ai-models-for-transcribing-old-italian-handwriting-and-typewritten-texts>> (ultimo accesso 17 ottobre 2025).

Tabella 3.11 – Parametri di configurazione e risultati comparativi dei tentativi di addestramento del *field model* “Ludovico Ariosto’s Handwriting”

<i>Ludovico Ariosto’s Handwriting</i>				
Parametro	Tentativo 1	Tentativo 2	Tentativo 3	Tentativo 4
Learning cycles	100	200	100	200
Early stopping	20	30	20	30
Use existing line polygons for training	Sì	Sì	Sì	Sì
Train abbreviations with expansion	Sì	Sì	Sì	Sì
Omit lines by tag 'unclear' or 'gap'	Sì	Sì	Sì	Sì
CER	4,32%	6,99%	8,04%	7,31%
Note	Miglior risultato complessivo. Ottimo equilibrio tra numero di cicli e generalizzazione; assenza di overfitting.	Addestramento prolungato; inizio di overfitting dopo ~100 cicli. CER più alto rispetto al Tentativo 1.	Apprendimento stabile ma meno efficace; probabile minore coerenza dei dati.	Leggero miglioramento rispetto al Tentativo 2, ma persistono segni di overfitting; modello meno generale.

I risultati dei quattro tentativi di *fine-tuning* sono illustrati nei grafici seguenti, che rappresentano l'andamento del CER sul *training set* e sul *validation set* in funzione dei cicli di apprendimento. L'analisi congiunta dei grafici e dei valori finali di CER consente di valutare l'efficacia del processo di ottimizzazione in termini di apprendimento, convergenza e fenomeni di *overfitting*. Nel primo tentativo, impostato su 100 cicli di apprendimento e *early stopping* a 20, il grafico mostra un rapido decremento del CER nei primi cicli, seguito da una fase di stabilizzazione. Le curve di training e validazione rimangono molto vicine, segno di un apprendimento equilibrato e privo di *overfitting*. Il valore finale di CER pari al 4,32% indica un'ottima accuratezza e una solida capacità di generalizzazione. Questo tentativo può essere considerato il più efficace nell'ambito del *fine-tuning* condotto (cfr. Fig. 3.32). Nel secondo tentativo, i cicli di apprendimento sono stati aumentati a 200 e l'*early stopping* portato a 30. Il grafico evidenzia un andamento iniziale analogo al precedente, ma con una crescente divergenza tra le curve di *training* e *validation* a partire dalla metà del processo. Dopo circa 100 cicli, la curva di validazione tende a stabilizzarsi, mentre quella di training continua a diminuire, segnalando l'inizio di un fenomeno di *overfitting*. Il CER finale, pari al 6,99%, risulta superiore a quello del primo esperimento, indicando che un numero eccessivo di cicli di *fine-tuning* non migliora le prestazioni complessive del modello, ma anzi può comprometterne la capacità di generalizzazione (cfr. Fig. 3.33).

Figura 3.32 – Primo addestramento del modello di HTR “Ludovico Ariosto’s Handwriting”

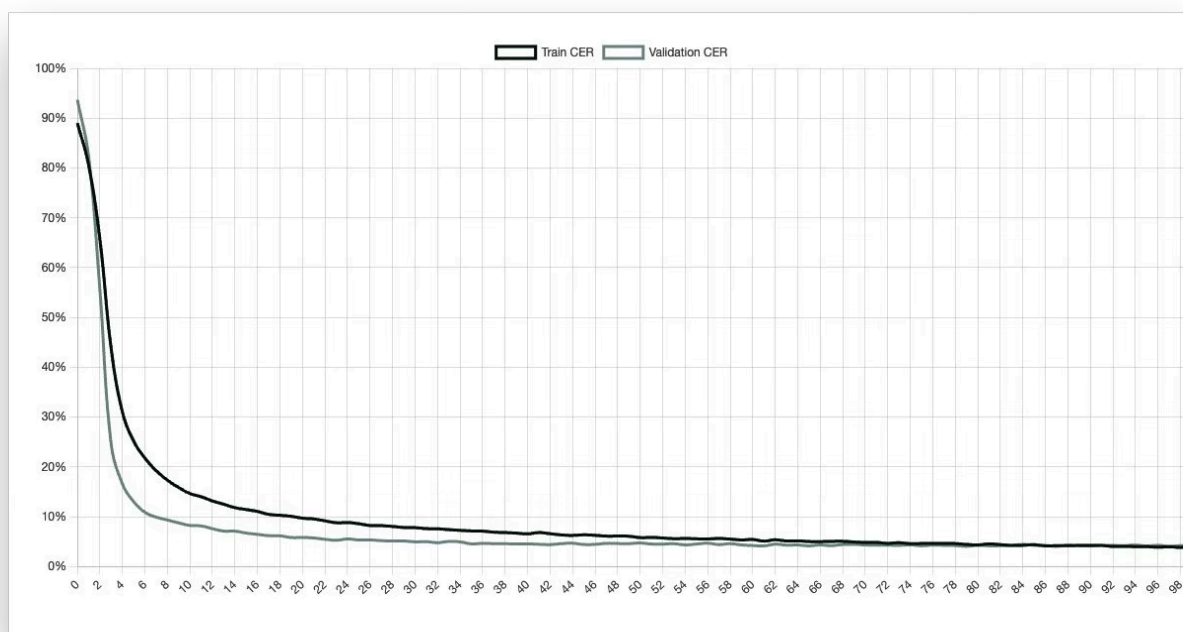


Figura 3.33 – Secondo addestramento del modello di HTR “Ludovico Ariosto’s Handwriting”

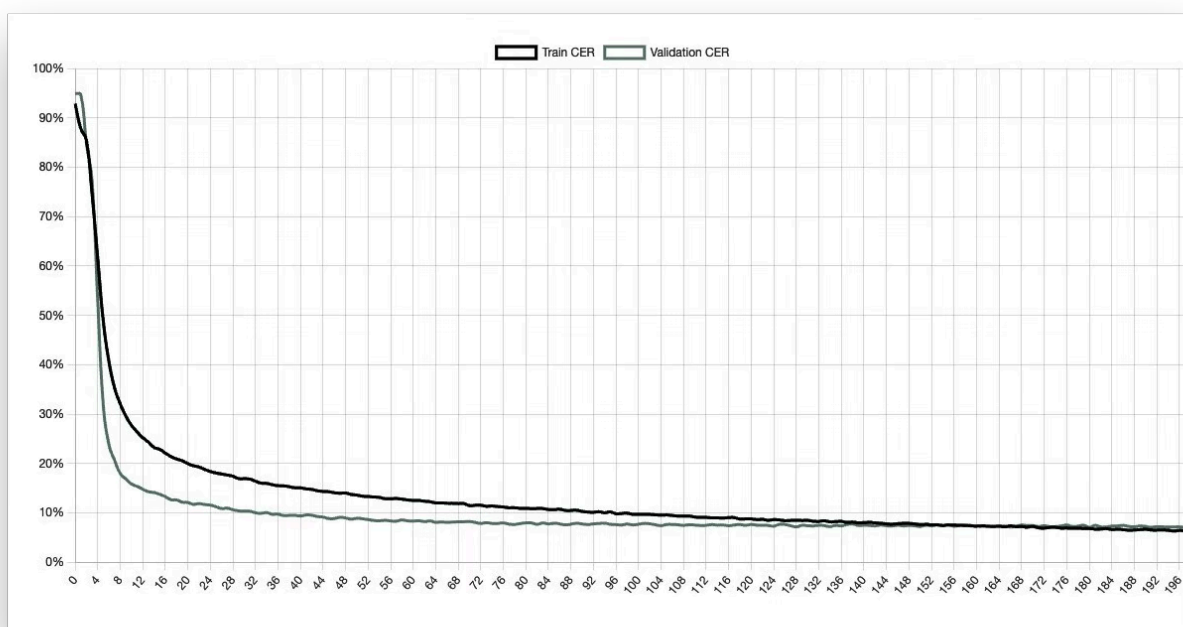


Figura 3.34 – Terzo addestramento del modello di HTR “Ludovico Ariosto’s Handwriting”

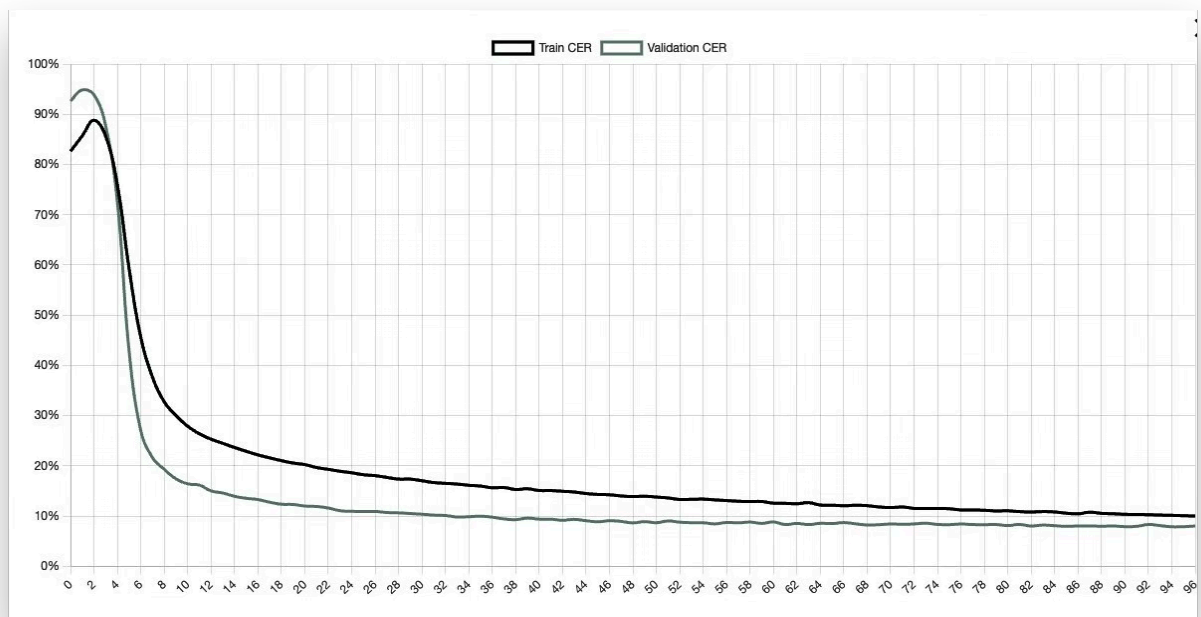
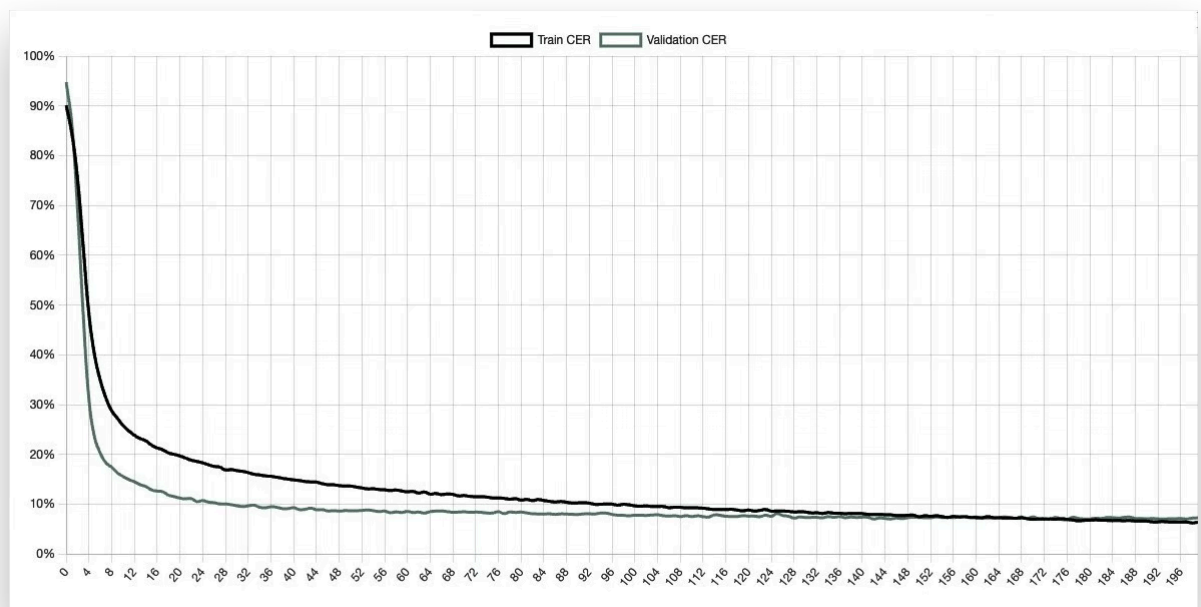


Figura 3.35 – Quarto addestramento del modello di HTR “Ludovico Ariosto’s Handwriting”



Il terzo tentativo, condotto nuovamente su 100 cicli con *early stopping* a 20, mostra un calo più graduale del CER rispetto ai casi precedenti. Le curve di *training* e *validation* seguono un andamento regolare ma restano più distanziate, suggerendo una minore efficacia del processo di *fine-tuning*. Il valore finale di CER (8,04%) indica una convergenza accettabile ma meno performante (cfr. Fig. 3.34). Infine, nel quarto tentativo, con 200 cicli e *early stopping* a 30, il grafico mostra un'evoluzione simile a quella del secondo esperimento. Le curve di *training* e *validation* rimangono relativamente vicine nella prima parte dell'apprendimento, ma la riduzione del CER procede più lentamente e tende a stabilizzarsi dopo circa 150 cicli. Il valore finale di CER pari al 7,31% rappresenta un leggero miglioramento rispetto al secondo tentativo, pur senza raggiungere le prestazioni ottimali del primo (Fig. 3.35).

In sintesi, l'analisi comparativa dei grafici e dei valori di CER evidenzia come il primo tentativo di *fine-tuning* rappresenti il miglior compromesso tra durata dell'addestramento, stabilità del modello e accuratezza complessiva. Come mostrano i tentativi successivi, l'aumento dei cicli di apprendimento non determina un miglioramento delle prestazioni, ma tende piuttosto a favorire un progressivo *overfitting*, riducendo la capacità del modello di generalizzare correttamente sui dati di validazione. Un ulteriore elemento che incide in modo significativo sulle prestazioni è rappresentato dalla gestione dei tag testuali durante l'addestramento su *Transkribus* – e, più in generale, nei modelli di HTR. Tali tag (XML o TEI) svolgono una funzione essenziale, poiché consentono di segnalare fenomeni come abbreviazioni ed espansioni, lacune o passaggi illeggibili (ad esempio <gap> e <unclear>), nonché altre annotazioni editoriali quali correzioni, aggiunte o note marginali. La scelta di includere o escludere determinati tag nelle trascrizioni influisce direttamente sulla coerenza tra le immagini di input e le trascrizioni di riferimento (*ground truth*), con ripercussioni sui valori del CER e sulla capacità di generalizzazione del modello. In particolare, un incremento del CER può derivare da incoerenze nella gestione dei tag tra il training set e il validation set: se, ad esempio, il modello è addestrato ignorando specifici tag (come <gap> o <unclear>), ma questi sono presenti nella fase di validazione, esso tenderà a produrre trascrizioni divergenti rispetto alla *ground truth*, con conseguente aumento del tasso di errore. Altri fattori che possono contribuire a tale incremento riguardano la trattazione non uniforme delle abbreviazioni e la presenza di tag inline non normalizzati (come <supplied>, <add>, , <choice>), i quali possono generare disallineamenti tra addestramento e validazione. Analogamente, la rimozione selettiva di righe contenenti tag complessi può essere utile per ridurre il rumore nei dati, ma se applicata in modo non omogeneo nei due set, compromette la capacità del modello di generalizzare in modo efficace. Pertanto, una gestione

coerente e normalizzata dei tag testuali risulta indispensabile per mantenere l'allineamento tra immagine e trascrizione, garantendo la stabilità dei risultati, contenendo l'aumento del CER e migliorando complessivamente l'affidabilità e la capacità predittiva del modello.

In conclusione, il modello, una volta ritenuto soddisfacente in termini di accuratezza e stabilità, è stato applicato all'intera collezione dei documenti. In questa fase è stata prestata particolare attenzione alla configurazione dei parametri avanzati, selezionando le opzioni *Keep original line polygons* e *No polygon simplification*. La prima consente di preservare la forma e la posizione originaria dei poligoni di linea, evitando che il sistema modifichi automaticamente i contorni delle aree di testo individuate durante la segmentazione; la seconda, invece, impedisce la semplificazione geometrica dei poligoni, mantenendo intatta la precisione dei margini. L'attivazione di entrambe le impostazioni risulta cruciale per garantire la fedeltà spaziale tra le immagini originali e le relative trascrizioni, evitando alterazioni nei limiti delle righe che potrebbero compromettere l'allineamento tra testo e supporto grafico. Questa scelta metodologica assicura una maggiore coerenza del riconoscimento su tutta la collezione, preservando la qualità del lavoro di segmentazione preliminare e la consistenza dei risultati finali.

Le trascrizioni generate dal modello risultano complessivamente accurate, salvo alcuni refusi riscontrabili nei punti in cui la grafia ariostesca si presenta particolarmente disordinata o di difficile decifrazione a causa delle stratificazioni correttive. In tali casi è stato necessario intervenire manualmente per correggere gli errori di riconoscimento. Un'ulteriore criticità riguarda la gestione dei tag multipli. Si ricorda che in *Transkribus*, i tag fungono da metadati descrittivi dei fenomeni testuali presenti nel documento, quali abbreviazioni ed espansioni, cancellature o aggiunte. L'assegnazione di un tag a una determinata porzione di testo ne comporta la registrazione nel file XML-TEI della trascrizione, dove esso è rappresentato come elemento di *markup* associato a uno specifico intervallo testuale.

Una delle principali limitazioni del sistema consiste nell'impossibilità di gestire correttamente più tag applicati alla stessa sequenza di testo; infatti, *Transkribus* non supporta la sovrapposizione o la nidificazione dei tag, ossia la coesistenza di più marcature sul medesimo intervallo. Quando l'utente assegna due tag distinti a una porzione di testo (ad esempio abbreviazione e cancellatura), il sistema tenta di rappresentarli entrambi nel file XML, ma ciò genera un conflitto di *markup*, poiché solo uno dei due può essere serializzato correttamente. Durante l'esportazione o l'addestramento dei modelli di trascrizione automatica, la struttura

testuale viene quindi semplificata: *Transkribus* seleziona un unico tag, solitamente quello applicato per ultimo o con priorità maggiore. Nel processo di riconoscimento HTR viene elaborato esclusivamente il contenuto testuale, mentre i tag sono gestiti in una fase successiva dal modulo di *layout & tagging*. Di conseguenza, in presenza di tag sovrapposti, nel *dataset* XML finale rimane un solo elemento, con perdita delle marcature multiple originarie.

Per preservare entrambi i tag e garantire una rappresentazione coerente, possono essere adottate alcune strategie operative:

- Separare leggermente le aree di applicazione dei tag, anche di un solo carattere, così da evitare la sovrapposizione completa e permettere al sistema di registrarli entrambi nel file TEI.
- Creare un tag combinato (ad esempio, “abbrev_cancellatura”) nei casi in cui i due fenomeni coesistano in modo inscindibile.
- Gestire la gerarchia dei tag direttamente nel codice XML, operando al di fuori dell’interfaccia di *Transkribus*.

Quest’ultima soluzione è quella ritenuta preferibile e sarà approfondita nel capitolo successivo, dedicato allo schema di codifica XML-TEI elaborato per i frammenti autografi.

IV

CAPITOLO

4.1 Prototipo di modello di codifica XML-TEI per l'edizione genetica digitale dei frammenti autografi

La trascrizione dei frammenti autografi ariosteschi rappresenta la fase preliminare alla loro codifica digitale secondo lo standard XML-TEI (*Text Encoding Initiative*), oggi riconosciuto come riferimento internazionale per la rappresentazione digitale dei testi nell'ambito delle *digital humanities*. Basata su un linguaggio gerarchico e leggibile sia dall'uomo sia dalla macchina, la TEI consente di descrivere con rigore la struttura logica, semantica e materiale dei documenti, garantendo al contempo flessibilità, interoperabilità e durabilità dei dati. La sua architettura modulare e la logica partecipativa che ne caratterizza lo sviluppo permettono di adattare lo schema di codifica alle esigenze specifiche di ciascun progetto, mantenendo la coerenza con le pratiche condivise dalla comunità scientifica.²³⁵

Nel Capitolo II è stata illustrata, a titolo esemplificativo, l'applicazione della TEI alle stanze 22 e 23 del Canto IX dell'*Orlando Furioso*, conservate nella c. 8r del II fascicolo del manoscritto autografo (ms. Cl. I. A., It. 1) custodito presso la Biblioteca Comunale Ariostea di Ferrara (F). Si tratta di una copia in bella, probabilmente destinata alla stampa, la cui codifica ha consentito di integrare la trascrizione diplomatica con la marcatura degli aspetti linguistici, metrici e paleografici, restituendo una rappresentazione digitale accurata e analitica del testo. Il presente capitolo intende approfondire tale prospettiva, proponendo un'analisi più articolata dei moduli TEI P5 (ovvero degli insiemi tematici di elementi e attributi XML) più pertinenti al caso di studio dei frammenti autografi e funzionali alla loro rappresentazione e modellizzazione all'interno di un'edizione genetica digitale. L'obiettivo è gettare le basi per un prototipo di

²³⁵ Per approfondire, cfr. i seguenti testi: AA.VV., *Creating a Digital Scholarly Edition with the Text Encoding Initiative*, a cura di M. Burghart, DEMM (Digital Editing of Medieval Manuscripts), 2017; L. Burnard, *What Is the Text Encoding Initiative?*, Marseille, OpenEdition Press, 2014; L. Burnard e C. M. Sperger-McQueen, *Il manuale TEI Lite. Introduzione alla codifica elettronica dei testi letterari*, a cura di F. Ciotti, Milano, Edizioni Sylvestre Bonnard, 2005.

modello di codifica, ovvero individuare un insieme di strategie in grado di restituire i documenti in modo fedele, completo e computazionalmente sostenibile, considerando i diversi livelli di analisi: bibliografico (collocazione e descrizione delle fonti), codicologico e paleografico (composizione e fisionomia materiale dei documenti), linguistico e metrico, genetico-autoriale (processo di elaborazione del testo). L'applicazione sistematica di tale modello all'intero corpus documentario sarà oggetto di una fase successiva della ricerca, poiché richiede una riflessione complessa e di lungo periodo, in ragione della natura stratificata e non lineare del materiale oggetto di studio.

La codifica in XML-TEI sarà realizzata secondo un approccio misto, combinando strumenti automatici e intervento umano. In una prima fase, il testo è stato elaborato con *Transkribus*, che effettua il riconoscimento automatico del contenuto e genera un file di output in formato TEI, fornendo così una struttura di base conforme agli standard. Successivamente, il file sarà sottoposto a un validatore automatico per individuare eventuali errori formali o discrepanze rispetto allo schema adottato. Infine, il documento sarà oggetto di una revisione manuale approfondita, volta a perfezionare la codifica, correggere imprecisioni residue e inserire markup specialistico ove necessario – ad esempio, per evidenziare elementi testuali di rilievo, aggiungere annotazioni o segnalare varianti. Grazie a questa combinazione di automazione e controllo umano, si mira a ottenere un livello elevato di accuratezza e aderenza agli standard internazionali di codifica testuale.

4.1.1 Il modulo TEI Header

Nel linguaggio di codifica TEI, il modulo *TEI Header* (o *teiHeader*) rappresenta la sezione introduttiva e descrittiva di ogni documento TEI. Esso non contiene il testo letterario vero e proprio, ma tutte le informazioni metatestuali e documentarie che descrivono il contenuto, la provenienza, la struttura e le modalità di creazione e gestione del file. In altre parole, svolge un ruolo analogo a quello della scheda catalografica o di una scheda di metadattazione archivistica: serve a garantire la tracciabilità, la leggibilità a lungo termine e la standardizzazione dei dati all'interno di un corpus digitale o di un progetto di edizione scientifica digitale. Il modulo si compone di quattro sezioni principali, ciascuna con una funzione specifica:²³⁶

²³⁶ Per approfondire, cfr. § 2 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 21-72.

1. L'elemento <fileDesc> (*file description*) è obbligatorio e contiene la descrizione bibliografica e catalografica del testo e del file digitale. Include:
 - <titleStmt>, con le informazioni sul titolo dell'opera e sui responsabili (autore, editore, curatore, trascrittore, ecc.);
 - <editionStmt>, con i dati relativi all'edizione o alla versione digitale;
 - <extent>, con l'indicazione dell'estensione o dimensione del documento;
 - <publicationStmt>, con le informazioni sulla pubblicazione o diffusione del file;
 - <sourceDesc>, che descrive la fonte o le fonti utilizzate (manoscritti, edizioni a stampa, copie digitali, ecc.).
2. L'elemento <encodingDesc> (*encoding description*) spiega i criteri di codifica adottati, le convenzioni editoriali e le norme utilizzate per abbreviazioni, correzioni o normalizzazioni testuali. Può contenere elementi come <projectDesc> (*project description*) o <editorialDecl> (*editorial declaration*).
3. L'elemento <profileDesc> (*profile description*) fornisce informazioni linguistiche, stilistiche, tematiche o relative all'autore. Include sottosezioni come <langUsage> (*language usage*) o <textClass> (*text classification*, cioè classificazione tematica o di genere).
4. L'elemento <revisionDesc> (*revision description*) registra la cronologia delle modifiche effettuate al file, con date, nomi e motivazioni, assicurando la tracciabilità dell'evoluzione del documento digitale.

Gli elementi del *teiHeader* possono contenere numerosi attributi per garantire precisione e interoperabilità, tra cui:

- @xml:id: identificatore univoco dell'elemento;
- @n: numero o etichetta interna;
- @type: specifica la tipologia (ad es. "autograph", "manuscript", "digitalEdition");
- @lang o @xml:lang: lingua del contenuto;
- @resp: indica la responsabilità di un'azione (ad es. "trascrittore", "revisore", "codificatore").

All'interno del *teiHeader*, una posizione di rilievo è occupata dall'elemento <listWit> (*list of witnesses*), che ha la funzione di elencare e descrivere i testimoni di un testo, ossia i manoscritti, le edizioni, le copie d'autore o le altre fonti che documentano le diverse lezioni del testo. Questo

elemento è particolarmente importante per le edizioni scientifiche digitali, specie quelle critico-genetiche, in quanto fornisce la base di riferimento per l'apparato critico e per la gestione delle varianti testuali. In un documento TEI, <listWit> si trova solitamente all'interno di <sourceDesc>, che a sua volta appartiene alla sezione <fileDesc> del <teiHeader>. In questa posizione, esso raccoglie in modo strutturato tutti i testimoni principali utilizzati nell'edizione. Ogni testimone è rappresentato da un elemento <witness>, identificato da un attributo univoco @xml:id, che può contenere una descrizione bibliografica o archivistica, spesso racchiusa in un elemento <bibl>. In alternativa, <listWit> può essere collocato anche nel corpo del testo (<text>), ad esempio in una sezione introduttiva o in appendice, quando si vuole elencare i testimoni pertinenti a una sola parte dell'opera o a una specifica sezione dell'edizione digitale. L'elemento <listWit> è strettamente collegato al modulo *Critical Apparatus*, che gestisce la rappresentazione delle varianti testuali. Gli elementi dell'apparato quali, <app> (*apparatus entry*), <lem> (*lemma*) e <rdg> (*reading*), utilizzano l'attributo @wit per fare riferimento ai testimoni elencati in <listWit>. In questo modo, si crea un sistema coerente di rimandi interni che collega direttamente le lezioni testuali alle loro fonti documentarie

4.1.6 Il modulo *Critical Apparatus*

Il modulo *Critical Apparatus* è uno dei moduli specializzati previsti dallo schema TEI per la codifica delle varianti testuali, siano esse di copia o d'autore, all'interno dell'apparato critico. Esso fa parte dei moduli opzionali che possono essere inclusi in un documento TEI quando l'edizione digitale prevede la registrazione di testimonianze multiple, congetture editoriali o fenomeni di variazione tra redazioni plurime di uno stesso testo.²³⁷ Dal punto di vista strutturale, il modulo *Apparato* si integra nel corpo principale del documento TEI, che è costituito dall'elemento radice <TEI>, contenente due sezioni principali:

1. la testata <teiHeader>, che descrive i metadati e la documentazione editoriale;
2. il testo vero e proprio, racchiuso in <text>.

All'interno dell'elemento <text>, il modulo *Apparato* può essere usato in qualunque punto del corpo testuale (ovvero all'interno di <body>, <div>, <p> o altri elementi di livello testuale) per segnalare un punto di variazione attraverso l'uso di <app>. In altre parole, <app> e i suoi

²³⁷ Per approfondire, cfr. § 13 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 461-484.

elementi figli (<lem>, <rdg>, <rdgGrp>, ecc.) sono inseriti direttamente nel flusso del testo, nel punto esatto in cui si verifica la variante, garantendo così un collegamento semantico e strutturale tra il testo base e l'apparato critico. L'apparato può inoltre collegarsi ai testimoni elencati nella sezione <listWit>, generalmente collocata nel <teiHeader> o in un'appendice del documento, attraverso gli attributi di riferimento (@wit, @source, @resp), stabilendo così un sistema coerente di rimandi interni tra testo, apparato e descrizione delle fonti. Gli elementi costitutivi del modulo sono:

- <app> (*apparatus entry*) che racchiude un punto di variazione testuale;
- <lem> (*lemma*) che contiene la lezione adottata nel testo base;
- <rdg> (*reading*) che rappresenta una lezione alternativa;
- <rdgGrp> (*reading group*) che consente di raggruppare più lezioni;
- <wit> (*witness*) che rinvia ai testimoni che attestano una lezione;
- <witDetail> e <note> che aggiungono informazioni contestuali o critiche su testimoni o varianti.

Ogni punto dell'apparato (<app>) può contenere una o più lezioni (<lem>, <rdg>) e riferimenti ai testimoni (@wit). La logica di annidamento e referenzialità tra gli elementi garantisce una rappresentazione coerente del rapporto tra testo e fonti. L'apparato è quindi integrato direttamente nel testo (piuttosto che collocato in un'area separata), e la sua gestione è delegata ai software di visualizzazione o ai criteri editoriali scelti dal progetto.

4.1.2 Il modulo *Manuscript Description*

Il modulo *Manuscript Description* (o *msdescription*) della TEI P5 è dedicato alla descrizione dei manoscritti e di altri testimoni materiali di un testo (autografi, apografi, miscellanee, quaderni, frammenti, ecc.). Si tratta di un modulo fondamentale per la rappresentazione digitale del patrimonio manoscritto, in quanto consente di strutturare in modo normalizzato tutte le informazioni necessarie a identificare, descrivere e contestualizzare un manufatto scrittorio, con un approccio simile a quello dei repertori e cataloghi biblioteconomici (es. ISBD, REICAT, ICCU, CERL, Manuscriptorium).²³⁸ Il suo elemento radice è <msDesc>, che può essere inserito:

²³⁸ Per approfondire, cfr. § 11 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 358-401.

- all'interno di un <sourceDesc> nel <teiHeader> (per descrivere la fonte di un testo digitale o trascritto);
- oppure come documento autonomo, se si vuole creare una scheda digitale di catalogo.

La struttura del modulo segue una logica gerarchica, che rispecchia le pratiche catalografiche tradizionali. L'elemento <msDesc> (*manuscript description*) è il contenitore principale. I suoi attributi principali sono:

- @xml:id: identificatore univoco della descrizione nel file TEI;
- @type: specifica il tipo di manoscritto (ad es. "autograph", "copy", "fragment", ecc.).

Al suo interno si trovano cinque sezioni principali.

1. L'elemento <msIdentifier> contiene i dati identificativi del manoscritto e la sua collocazione fisica, espressi mediante i seguenti sottocampi:

- <country>, paese di conservazione;
- <region>, area o regione;
- <settlement>, città;
- <repository>, istituto o biblioteca;
- <collection>, fondo o collezione;
- <idno>, segnatura o numero di inventario.

2. L'elemento <msContents> descrive il contenuto testuale del manoscritto mediante i seguenti sottocampi:

- <msItem> rappresenta una singola unità testuale (opera, frammento, sezione) di cui è possibile specificare: <locus>, la posizione nel manoscritto; <title>, il titolo dell'opera o del frammento; <author>, l'autore del testo; <incipit> e <explicit>, le prime e ultime parole del testo; <note>, le osservazioni sul contenuto.

3. L'elemento <physDesc> contiene la descrizione fisica, paleografica e codicologica del manoscritto, articolata nei seguenti sottoinsiemi tematici.

- Oggetto fisico: <objectDesc @form>, tipo di oggetto (ad es. codice, foglio, bifoglio, ecc.), che include:

- <supportDesc> con <support> (tipo di supporto come carta, pergamena, papiro) e, se pertinente, <watermark> per la descrizione di filigrane;
 - <layoutDesc> / <layout @columns @ruledLines @writtenLines>, disposizione del testo sulla pagina, numero di colonne, rigatura e righe scritte, presenza di rigatura o foratura (*pricking*).
- Scrittura e mani: <handDesc>, descrizione delle mani scriventi; <handNote>, caratteristiche grafiche; <scriptDesc>, tipologia di scrittura (ad es. gotica, umanistica, cancelleresca).
 - Decorazione: <decoDesc> con eventuali <decoNote>, descrizione delle decorazioni (iniziali, rubriche, miniature, fregi).
 - Legatura e stato materiale: <bindingDesc> / <binding>, legatura attuale e legature precedenti; <condition>, stato di conservazione; <accMat>, materiali d'accompagnamento (ad es. note, custodie, inserti).
 - Struttura dei fascicoli e segnature: <collation>, formula e descrizione dei fascicoli (*quires*); <signatures>, sistema delle segnature; <catchwords>, sistema dei richiami (*catchwords*). In trascrizione, è possibile segnalare l'inizio di un fascicolo mediante il milestone <pb/>.²³⁹
 - Foliazione e impaginazione: <foliation>, sistema di foliazione o numerazione delle carte adottato.
4. L'elemento <history> registra le informazioni storiche e la provenienza del manoscritto mediante i seguenti sottocampi:
- <origin>, luogo e data di produzione;
 - <provenance>, informazioni sulla storia del possesso o sul contesto di conservazione;

²³⁹ Nel linguaggio di codifica TEI, una *milestone* è un marcatore puntuale e privo di contenuto che serve a segnalare un confine o un punto di discontinuità all'interno del testo. A differenza degli elementi strutturali che racchiudono un contenuto (come <div> o <p>), le *milestone* si limitano a indicare un punto preciso, senza delimitare un intervallo di testo. Questo tipo di elemento risulta particolarmente utile in tre situazioni. In primo luogo, consente di registrare confini fisici del supporto materiale, come il cambio di riga, di pagina o di colonna. In secondo luogo, permette di rappresentare strutture che non possono essere espresse attraverso un annidamento gerarchico regolare, come nel caso delle cosiddette gerarchie sovrapposte, in cui una sezione o un'unità testuale comincia all'interno di un paragrafo e termina in un altro. Infine, le milestone risultano utili quando si desidera indicare limiti specifici di un'edizione o applicare convenzioni editoriali particolari senza alterare la struttura logica del testo.

L'elemento generico <milestone/> può essere impiegato per segnare confini personalizzati di vario tipo, mentre la TEI prevede anche una serie di elementi *milestone-like* già definiti, destinati a usi più specifici: <lb/> (*line break*) per il cambio di riga, <pb/> (*page break*) per il cambio di pagina e <cb/> (*column break*) per il cambio di colonna. A questi si aggiunge l'elemento <anchor/>, anch'esso privo di contenuto, che serve a marcare punti di riferimento nel testo da collegare tramite relazioni esplicite.

- <acquisition>, notizie sull'acquisizione da parte dell'istituto attuale.

5. L'elemento <additional> raccoglie dati amministrativi, bibliografici e di consultazione mediante i seguenti sottocampi:

- <adminInfo>, condizioni di accesso, stato di conservazione;
- <availability>, restrizioni d'uso o licenze;
- <listBibl>, bibliografia relativa al manoscritto.

Gli elementi del modulo *msdescription* possono contenere vari attributi standard TEI, tra cui quelli segnalati nella seguente tabella (cfr. Tabella 4.1):

Tabella 4.1 - Attributi associati agli elementi del *msdescription*

Attributo	Descrizione	Esempio
@xml:id	Identificatore univoco dell'elemento nel file	xml:id="ms_ferrara"
@n	Numero progressivo o identificativo interno	n="2"
@type	Tipologia o categoria (autograph, copy, fragment, palimpsest, ecc.)	type="fragment"
@resp	Responsabilità di un'azione (trascrizione, descrizione, ecc.)	resp="#trascrittore"
@when, @notBefore, @notAfter	Datazione o intervallo cronologico	notBefore="1521" notAfter="1532"
@lang o @xml:lang	Lingua del contenuto	xml:lang="it"
@form, @columns, @ruledLines, @writtenLines	Attributi di layout e struttura	form="codex" columns="2"

4.1.3 Il modulo *Facsimile*

Il modulo *Facsimile* della TEI è un'estensione del modello di codifica concepita per collegare la trascrizione testuale alle riproduzioni digitali del supporto materiale. Esso consente di associare in modo sistematico il testo codificato alle immagini digitali del manoscritto, della stampa o di altro artefatto documentario e, se necessario, di suddividere tali immagini in aree specifiche

(*zone*), rappresentando così la disposizione spaziale del testo e degli elementi paratestuali. Gli elementi del modulo *Facsimile* si collocano al di fuori della sezione testuale, nella parte strutturale del documento TEI. La collocazione canonica è la seguente:²⁴⁰

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">

  <teiHeader> ... </teiHeader>

  <facsimile> ... </facsimile>

  <text> ... </text>

</TEI>
```

L'elemento `<facsimile>`, posto allo stesso livello di `<text>` e `<teiHeader>`, svolge la funzione di componente descrittiva globale, accessibile da qualsiasi parte della trascrizione mediante riferimenti tramite l'attributo `@facs`. La struttura interna del modulo Facsimile rispecchia la relazione tra il documento fisico e la sua rappresentazione digitale. Essa si articola in tre livelli principali: `<facsimile>`, `<surface>` e `<zone>`, che modellano in modo progressivo la corrispondenza tra testo e immagine.

1. L'elemento `<facsimile>` costituisce la radice del modulo e funge da contenitore generale delle superfici digitali del documento. Può includere uno o più elementi `<surface>` o `<surfaceGrp>`, rappresentanti rispettivamente singole superfici o insiemi di superfici (ad es. un bifoglio o una sequenza di carte). In questo modo `<facsimile>` diventa il punto di riferimento centrale per tutte le risorse iconiche associate a un testimone manoscritto o a un documento a stampa.
2. L'elemento `<surface>` rappresenta una singola superficie fisica (una pagina, una carta o una facciata) e costituisce l'unità fondamentale all'interno di `<facsimile>`. Può contenere uno o più elementi `<graphic>`, che rimandano a file d'immagine esterni tramite l'attributo `@url`, ed eventualmente elementi `<zone>` per la suddivisione in aree specifiche. Ad esempio:

```
<surface xml:id="f1r" n="1r">

  <graphic url="images/f1r.jpg"/>

</surface>
```

²⁴⁰ Per approfondire, cfr. § 12.1 e 12.2 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 402-417.

L'attributo `@xml:id` assegna un identificatore univoco alla superficie, mentre `@n` indica la segnatura o numerazione convenzionale.

3. L'elemento `<zone>`, figlio di `<surface>`, definisce una porzione specifica della superficie come una colonna, un riquadro decorativo, un blocco di testo o una singola riga. Le zone vengono descritte tramite attributi spaziali: `@ulx`, `@uly`, `@lrx`, `@lry` (coordinate dei vertici) o `@points` (insieme di coordinate). Questi parametri consentono di delineare aree rettangolari o poligonali all'interno dell'immagine, permettendo una mappatura precisa della disposizione grafica e testuale. L'uso di `<zone>` risulta essenziale per la rappresentazione della mise en page, per l'analisi paleografica e per la visualizzazione sinottica immagine-testo.

Il legame tra trascrizione e immagini è stabilito mediante l'attributo globale `@facs`, che può essere usato su diversi elementi testuali all'interno della sezione `<text>` o, più precisamente, all'interno di `<body>`. `@facs` accetta come valore un riferimento (URI o identificatore interno) che punta a un elemento `<facsimile>`, `<surface>` o `<zone>`. Considerando l'esempio sopra, l'attributo `@facs` può essere applicato a diversi livelli della gerarchia testuale:

- Su un elemento di alto livello, collega l'intero testo al facsimile designato:

```
<text facs="#facsimile1">
```

- Su un elemento di delimitazione, indica che il cambio di pagina corrisponde alla superficie identificata come `f1r`:

```
<pb facs="#f1r" n="1r"/>
```

- Su un elemento di dettaglio, mette in relazione una singola parola con la zona corrispondente dell'immagine:

```
<w facs="#z12">in</w>
```

In questo modo, `@facs` permette di stabilire un collegamento preciso tra i diversi livelli della trascrizione e le rispettive porzioni del facsimile digitale.

In conclusione, il modulo *Facsimile* fornisce alla TEI una struttura coerente e flessibile per integrare testo e immagine, rendendo possibile una rappresentazione completa del documento sia nella sua dimensione linguistica sia in quella materiale.

4.1.4 I moduli *Default Text Structure* e *Verse*

Nel sistema di codifica TEI, la rappresentazione degli aspetti strutturali e formali di un testo poetico, come la suddivisione in canti, le strutture strofiche, la disposizione materiale sulla pagina e i fenomeni retorico-sintattici come l'enjambement, si realizza attraverso l'impiego coordinato di diversi moduli, ciascuno con una funzione specifica. Nel caso di un poema epico-narrativo come l'*Orlando furioso* di Ludovico Ariosto, è fondamentale adottare una strategia di codifica che consenta di preservare al contempo la complessità metrica, la struttura gerarchica del testo e le peculiarità stilistiche.

Il modulo *Default Text Structure* (o *textstructure*) ha come obiettivo principale l'articolazione macrostrutturale del testo, ossia la definizione delle sue principali unità gerarchiche. L'elemento di base è <text>, al cui interno si trova <body>, che costituisce il corpo effettivo del testo. Quest'ultimo può essere ulteriormente suddiviso mediante <div>, che, nel caso di testi poetici di ampio respiro come l'*Orlando furioso*, può avere l'attributo type="canto" per rappresentare la suddivisione in canti. Ogni canto può contenere un'intestazione <head> con il titolo o la denominazione del canto stesso, seguita da sequenze poetiche strutturate tramite <lg> (*line group*, ossia gruppo di versi o strofa) e <l> (verso), elementi definiti dal modulo *Verse*.²⁴¹

Il modulo *Verse* è dedicato alla rappresentazione delle strutture strofiche e delle informazioni metriche. L'elemento <lg> (*line group*) viene utilizzato per rappresentare le stanze o ottave, (ad es. <lg type="ottava">), mentre i singoli versi sono racchiusi in <l>. Nel caso dell'*Orlando furioso*, la struttura metrica di riferimento è l'ottava rima, composta da otto versi endecasillabi con schema di rime ABABABCC. Prendendo in esame la stanza 22 del Canto IX dell'*Orlando Furioso*, riportata in precedenza, la relativa struttura testuale può essere codificata come segue:

```
<lg type="stanza" subtype="ottava" n="22" met="11" rhyme="ABABABCC">
```

```
<l n="1">Io voglio ch sappiate: ch figliuola</l>
```

```
<l n="2">fui dl conte d'olanda: a lui si grata:</l>
```

```
<l n="3">Quantuq prole io no gli fossi sola:</l>
```

²⁴¹ Per approfondire, cfr. § 4 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 174-205.

<l n="4">Ch'era da lui fratelli accompagnata:</l>
 <l n="5">Ch'a quato io gli chiedea: da lui parola</l>
 <l n="6">Contraria no mi fu mai replicata,</l>
 <l n="7">Standomi lieta i questo stato: avene</l>
 <l n="8">Che ne la nostra terra un duca venne,</l>
 </lg>

Gli attributi @met e @rhyme derivano dal gruppo *att.metrical*, che permette di esplicitare, in modo sistematico, le informazioni relative al metro, alla realizzazione effettiva e allo schema rimico. Inoltre, la dichiarazione metrica generale dell'opera può essere formalizzata attraverso l'elemento <metDecl>, mentre l'uso di <rhyme> con l'attributo @label consente di specificare i diversi suoni rimici (ad esempio a, b, c), garantendo una codifica coerente e riutilizzabile anche a fini di analisi automatica o comparativa.²⁴²

Per la rappresentazione della disposizione materiale del testo sulla pagina, la TEI mette a disposizione una serie di elementi “core” che permettono di descrivere l'impaginazione e i cambi di supporto, come <pb/> (*page break*, cambio di pagina), <cb/> (*column break*, cambio di colonna) e <lb/> (*line break*, a capo tipografico). Questi elementi, corredati da attributi quali @n, @ed, @edRef o @type, permettono di documentare la struttura fisica del testo, aspetto cruciale soprattutto nelle trascrizioni diplomatiche di testimoni manoscritti o a stampa. Ulteriori informazioni sul layout possono essere fornite tramite il modulo *msdescription*, che include elementi come <layout> e <layoutDesc>, utili per descrivere la configurazione materiale della pagina (numero di colonne, linee tracciate e linee scritte), offrendo così un quadro completo della dimensione visiva del testo.²⁴³

Accanto alla rappresentazione metrica e strutturale, un aspetto di rilievo nella codifica del Furioso è la marcatura degli enjambement, ossia la prosecuzione del senso sintattico oltre la fine del verso. Tale fenomeno può essere rappresentato in TEI mediante diverse strategie, a seconda del livello di dettaglio desiderato. Una soluzione consiste nell'utilizzare l'elemento <seg>

²⁴² Per approfondire, cfr. § 6 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp.231-246.

²⁴³ Per approfondire, cfr. § 3 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 73-172.

(oppure <phr>) con l'attributo type="enjambement" per segnalare il punto di interruzione sintattica, come nell'esempio seguente:

```
<lg type="stanza" subtype="ottava" n="22" met="11" rhyme="ABABABCC">
<l n="1">Io voglio ch sappiate: <seg type="enjambement" n="1-2">ch figliuola</seg></l>
<l n="2"><seg type="enjambement" n="1-2">fui dl conte d'olanda</seg>:<seg type="enjambement"
n="2-3">a lui si grata</seg>:</l>
<l n="3"><seg type="enjambement" n="2-3">Quantuq prole io no gli fossi sola</seg>:</l>
<l n="4"><seg type="enjambement" n="3-4">Ch'era da lui fratelli accompagnata</seg>:</l>
<l n="5"><seg type="enjambement" n="4-5">Ch'a quato io gli chiedea</seg>:<seg
type="enjambement" n="5-6">da lui parola</seg></l>
<l n="6"><seg type="enjambement" n="5-6">Contraria no mi fu mai replicata</seg>,</l>
<l n="7"><seg type="enjambement" n="6-7">Standomi lieta i questo stato</seg>:<seg
type="enjambement" n="7-8">avene</seg></l>
<l n="8"><seg type="enjambement" n="7-8">Che ne la nostra terra un duca venne</seg>,</l>
</lg>
```

In questo esempio, la marcatura avviene a cavallo dei versi: la stessa coppia di elementi <seg> condivide l'attributo n, che indica il salto metrico (ad esempio n="1-2" per l'enjambement tra i vv. 1–2). In tal modo, l'elemento <seg> non solo evidenzia la discontinuità grafica del verso, ma anche la continuità semantica e sintattica del periodo, consentendo analisi automatiche e visualizzazioni che mettano in luce le relazioni tra versi contigui.

In alternativa, è possibile utilizzare l'elemento <lb> (line break), corredato da un attributo personalizzato come @enjamb="yes", per segnalare che il passaggio di riga corrisponde a un enjambement. Nell'esempio seguente, l'intero testo è racchiuso in un singolo elemento <l> o, più opportunamente, in un blocco di versi <lg>. Ciascun elemento <lb> indica non solo la presenza dell'enjambement, ma anche la coppia di versi coinvolta, tramite l'attributo aggiuntivo @enjambPair. In tal modo, ogni interruzione di verso che coincide con un enjambement è annotata in forma esplicita, ad esempio: enjamb="yes" enjambPair="1-2".²⁴⁴

²⁴⁴ Per approfondire, cfr. § 6 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp.231-246.

```

<lg type="stanza" subtype="ottava" n="22" met="11" rhyme="ABABABCC">
  Io voglio ch sappiate: ch figliuola<lb enjamb="yes" enjambPair="1-2"/>
  fui dl conte d'olanda: a lui si grata:<lb enjamb="yes" enjambPair="2-3"/>
  Quantuq prole io no gli fossi sola:<lb enjamb="yes" enjambPair="3-4"/>
  Ch'era da lui fratelli accompagnata:<lb enjamb="yes" enjambPair="4-5"/>
  Ch'a quato io gli chiedea: da lui parola<lb enjamb="yes" enjambPair="5-6"/>
  Contraria no mi fu mai replicata,<lb enjamb="yes" enjambPair="6-7"/>
  Standomi lieta i questo stato: avene<lb enjamb="yes" enjambPair="7-8"/>
  Che ne la nostra terra un duca venne,
</lg>

```

Infine, un'ulteriore soluzione, più astratta ma estremamente flessibile, consiste nell'utilizzare l'elemento `<linkGrp>` per raccogliere e formalizzare le relazioni di enjambement tra versi, senza intervenire direttamente nel testo. In questo caso, la connessione sintattica è rappresentata come un collegamento tra due o più elementi identificati.²⁴⁵ Tale modalità consente di mantenere il testo pulito e leggibile, concentrando la rappresentazione delle relazioni sintattiche e poetiche in una sezione dedicata dell'apparato strutturale. Inoltre, l'impiego di `<linkGrp>` favorisce analisi quantitative e visualizzazioni reticolari delle relazioni tra versi o strofe, senza modificare la codifica del testo di base (cfr. § 2.2.4). In sintesi, le tre strategie (tramite `<seg>`, tramite `<lb>` e tramite `<linkGrp>`) offrono diversi livelli di granularità e astrazione nella rappresentazione dell'enjambement, permettendo di distinguere chiaramente tra il confine grafico e quello sintattico, in coerenza sia con la tradizione filologica sia con le esigenze dell'analisi digitale del testo poetico.

²⁴⁵ Oltre alle milestone, la TEI offre diverse strategie complementari per rappresentare discontinuità o relazioni testuali. Gli elementi `<seg>` o `<span>` possono essere usati per evidenziare intervalli interni al testo, quando un singolo punto non basta. Il metodo *stand-off*, basato su `<link>` e `<linkGrp>`, consente invece di descrivere relazioni o corrispondenze tra parti del testo in una sezione separata, mantenendo il testo di base libero da marcature analitiche. Un'ulteriore possibilità è l'uso combinato di `<anchor>` e `<link>`, in cui due ancore identificate da attributi `xml:id` distinti vengono collegate esternamente tramite un elemento `<link target="#a1 #a2"/>`. In sintesi, le *milestone* e le strategie affini offrono strumenti flessibili per rappresentare confini, relazioni e discontinuità nel testo, preservando al tempo stesso la coerenza strutturale e la leggibilità della codifica TEI. Per approfondire, cfr. § 17 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 578-584.

Dunque, la combinazione dei moduli *textstructure*, *verse* e dei *core elements* dedicati alla materialità del testo, unita a una marcatura interpretativa degli enjambement, consente di rappresentare in TEI in modo completo e articolato la complessa architettura poetica dell'*Orlando furioso*. Tale strategia permette di mantenere una netta distinzione tra livello metrico, strutturale e sintattico, rispettando i principi di interoperabilità e riusabilità propri della TEI, e al tempo stesso offre una base solida per analisi quantitative, visualizzazioni e studi filologici nell'ambito delle *digital humanities*.

4.1.5 Il modulo *Representation of Primary Sources*

Il modulo *Representation of Primary Sources* della TEI P5 fornisce un insieme di elementi destinati alla trascrizione diplomatica e semi-diplomatica dei testi manoscritti e dei documenti originali. Il suo obiettivo è consentire una rappresentazione fedele della materialità e della stratificazione grafica di un testo, preservando le tracce delle diverse fasi di scrittura e di revisione, caratteristiche dei manoscritti autografi o dei testimoni annotati. In altri termini, tale modulo consente di rendere visibile la genesi testuale, ovvero il processo di elaborazione del testo attraverso cancellature, riscritture, aggiunte, ripensamenti e altre modifiche d'autore che costituiscono il laboratorio della scrittura.²⁴⁶

1. L'elemento `<add>` (*addition*) serve a segnalare un testo aggiunto rispetto alla stesura originale, sia esso scritto sopra la riga, in interlinea o a margine. Può essere accompagnato da attributi che specificano la posizione, la mano e il tipo di intervento, quali:

- `@place`: indica la collocazione dell'aggiunta ("inline", "above", "below", "left", "right", "margin", "overstrike");
- `@hand`: rinvia a un identificatore di mano (ad es. #ariosto);
- `@type`: distingue la natura dell'aggiunta ("interlinear", "marginal", "overwritten", "later addition").

2. L'elemento `<del>` (*deletion*) rappresenta il testo cancellato o soppresso e può essere accompagnato dai seguenti attributi:

²⁴⁶ Per approfondire, cfr. § 12 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 402-460.

- @rend: descrive la forma grafica della cancellatura (“overstrike”, “erased”, “crossedOut”, “blotted”, ecc.);
 - @hand: identifica la mano che ha effettuato la cancellatura;
 - @type: consente di distinguere la natura dell’intervento (“autograph”, “editorial”, “mechanical”).
3. L’elemento <subst> (*substitution*) permette di rappresentare una correzione complessa, in cui un testo è cancellato e un altro è aggiunto in sua sostituzione; contiene al proprio interno uno o più e <add>.
4. Gli elementi <sic> e <corr> servono a distinguere rispettivamente tra la forma originale e la forma proposta dal curatore o dall’autore stesso.
5. L’elemento <choice> è usato per rappresentare varianti o alternative tra più lezioni di un testo, specialmente quando si tratta di diverse campagne redazionali o fasi di revisione. Esso può contenere diverse coppie o insiemi di sottoelementi che descrivono relazioni tra forme testuali alternative. Tra i più comuni si trovano:
- <orig>, che indica la forma originale di una parola o espressione;
 - <reg>, che rappresenta la forma regolarizzata o modernizzata;
 - <sic>, che conserva una forma errata o anomala così come appare nel testo di base;
 - <corr>, che offre la forma corretta o normalizzata corrispondente;
 - e infine <abbr> e <expan>, usati per la gestione delle abbreviazioni e delle loro espansioni.

Un aspetto di particolare rilievo nella codifica dei testi manoscritti, in particolare nei frammenti ariosteschi, riguarda la gestione delle abbreviazioni. L’elemento <abbr> viene utilizzato per trascrivere fedelmente la forma abbreviata così come appare nel manoscritto, mentre <expan> ne contiene l’espansione completa o la lettura interpretativa. La TEI consente inoltre di arricchire la codifica attraverso metadati descrittivi: l’attributo @type può essere impiegato per specificare la tipologia dell’abbreviazione (ad es. “sospensione”, “contrazione” o “sigla”), mentre @resp permette di indicare la responsabilità dell’espansione o dell’interpretazione. In questo modo, la marcatura risulta più trasparente e rigorosamente documentata dal punto di vista filologico ed editoriale.

6. Gli elementi <gap> e <supplied> indicano rispettivamente una parte mancante del testo e l’integrazione o ricostruzione editoriale di un passaggio lacunoso.

7. Gli elementi `<handList>` e `<handNote>` consentono di identificare e descrivere le diverse mani presenti in un documento, risultando particolarmente utili nei casi in cui siano attestate più fasi redazionali o revisioni successive.
8. L'elemento `<seg>` (segment) è usato per delimitare e qualificare porzioni di testo che rivestono particolare interesse analitico o critico. Quando corredato dagli attributi `@type="revision"` o `@ana`, esso diventa uno strumento fondamentale per individuare le revisioni d'autore e per riconoscere le diverse campagne redazionali intervenute nel processo di composizione.
- L'attributo `@type="revision"` specifica che il segmento rappresenta una modifica apportata in fase di revisione (ad esempio, una riscrittura, un'aggiunta o una soppressione rispetto alla stesura precedente).
 - L'attributo `@ana`, invece, collega il segmento a un identificatore (URI o riferimento interno) che rinvia a una descrizione più dettagliata della campagna di revisione o del momento genetico in cui l'intervento si colloca.

Di conseguenza, questi attributi consentono di mappare la storia evolutiva del testo, stabilendo relazioni esplicite tra i singoli interventi nel corpo del testo e la loro descrizione storico-filologica collocata in un'apposita sezione del documento TEI.

Per documentare in modo sistematico e coerente le diverse fasi di revisione o di trasformazione testuale, la TEI prevede due sezioni specifiche e complementari: `<listChange>` e `<revisionDesc>`.

1. L'elemento `<listChange>` si colloca nel corpo del documento TEI, al di fuori del `<teiHeader>`, e di norma all'interno della sezione `<text>`, preferibilmente nel `<back>` o in uno spazio descrittivo autonomo (ad esempio, in `<standOff>`). Tale collocazione riflette la sua funzione di descrizione genetico-filologica delle revisioni autoriali o redazionali del testo, in quanto parte integrante della rappresentazione del documento d'autore e non della sua edizione digitale. L'elemento `<listChange>` raccoglie una serie ordinata di elementi `<change>`, ciascuno dei quali descrive una fase di intervento – una correzione, una riscrittura, una revisione globale, ecc. – collegata al testo tramite identificatori (quali `@ana`, `@change`, ecc.). Dal punto di vista filologico, `<listChange>` svolge pertanto la funzione di registro cronologico e strutturato delle campagne redazionali, consentendo di correlare i dati testuali (ad esempio, un elemento `<seg>` con `@ana="#ch2"`) alle informazioni metatestuali relative

alla revisione cui tali dati appartengono. Ciascun elemento <change> può inoltre descrivere, oltre alla natura dell'intervento correttivo:²⁴⁷

- il soggetto che lo ha effettuato (@who);
- la datazione o il periodo (@when, @notBefore, @notAfter);
- l'ambito o l'oggetto della revisione (ad esempio, un canto, un capitolo o un passo specifico);
- eventuali note esplicative.

2. L'elemento <revisionDesc> è collocato all'interno del <teiHeader> e ha la funzione di documentare le modifiche apportate al file TEI o all'edizione digitale, quali versioni, aggiornamenti e interventi di natura editoriale. Si tratta, dunque, di un elemento concepito per la descrizione metatestuale del documento digitale, volto a registrare la storia editoriale e tecnica della codifica. In ambito filologico-genetico, tuttavia, l'uso di <revisionDesc> può essere esteso per rappresentare le fasi di trasformazione autoriale di un testo, intese come vere e proprie revisioni storiche del processo compositivo. In questo senso, l'elemento può fungere da metadato storico-editoriale, a condizione che tale riappropriazione metodologica sia dichiarata esplicitamente nella documentazione del progetto o nella sezione <encodingDesc>, al fine di evitare ambiguità tra il livello della codifica e quello del testo d'autore. Ciascun intervento registrato all'interno di <revisionDesc> è rappresentato da un elemento <change>, analogo a quelli contenuti in <listChange>. Mentre, però, <listChange> si colloca nel corpo del documento e descrive le revisioni autoriali o redazionali in senso strettamente genetico, <revisionDesc> mantiene la propria collocazione nel header, garantendo così una separazione funzionale tra la documentazione metatestuale (relativa all'edizione digitale) e la documentazione genetica (inerente al testo d'autore).²⁴⁸

In un progetto complesso è pertanto possibile mantenere entrambi gli elementi:

- <revisionDesc> per la storia editoriale digitale;
- <listChange> per la storia redazionale autoriale.

L'adozione di tali strutture consente di costruire una rappresentazione cronologicamente coerente e scientificamente tracciabile dell'evoluzione testuale. Ogni intervento (una

²⁴⁷ Per approfondire, cfr. § 12 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 402-460.

²⁴⁸ Per approfondire, cfr. § 2 del *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), pp. 21-72.

sostituzione, una riscrittura, una cancellatura) non è più considerato come episodio isolato, ma come parte di un sistema relazionale che collega:

- il testo fenomenico (il segno sul manoscritto);
- il contesto storico (la revisione o campagna di lavoro);
- la dimensione interpretativa (il commento o la nota critica).

Nel caso dei manoscritti ariosteschi, ad esempio, un sistema fondato sull'uso degli elementi `<seg>` e `<listChange>` consente di visualizzare la progressiva evoluzione poetica degli episodi aggiunti, seguendo il percorso che va dai primi abbozzi alle brutte copie, fino alle stesure definitive. In questo modo è possibile tracciare in modo analitico le revisioni che documentano la maturazione stilistica e compositiva dell'autore.

In conclusione, il modulo *Representation of Primary Sources* permette di restituire con rigore filologico la dimensione genetica del testo, documentandone le fasi di elaborazione, le trasformazioni progressive e gli interventi d'autore che ne scandiscono la storia redazionale. Attraverso gli strumenti previsti dal modulo è possibile riconoscere e distinguere le diverse mani presenti in un manoscritto, nonché mappare le campagne di revisione e le riscritture successive, offrendo una rappresentazione analitica e stratificata del processo compositivo. Parallelamente, la codifica TEI garantisce un'elevata interoperabilità dei dati, favorendo l'integrazione tra edizioni critiche digitali, riproduzioni facsimilari e archivi specialistici. Le funzionalità di visualizzazione dinamica – come la possibilità di mostrare o nascondere cancellature, aggiunte e correzioni – offerte da software per la pubblicazione di edizioni digitali scientifiche (quali EVT) ampliano le prospettive di indagine e rendono più consapevole la fruizione critica del documento. Ne consegue che il modulo si configura come uno degli strumenti più rilevanti nell'ambito della filologia d'autore digitale. Un esempio particolarmente significativo è costituito dai manoscritti ariosteschi, nei quali cancellature, varianti e aggiunte rappresentano tracce essenziali per la ricostruzione del laboratorio creativo dell'Orlando Furioso e per la comprensione della sua complessa stratificazione redazionale (cfr. Tab. 4.2).

Tabella 4.2 - schema riassuntivo del *modulo Representation of Primary Sources*

Elemento	Descrizione
<add>	Funzione principale: segnala un'aggiunta rispetto alla stesura originaria del testo. Attributi principali: @place, @hand, @type Collocazione nel documento TEI: all'interno del corpo testuale (<body> o <text>). Relazione con altri elementi: spesso associato a (in <subst>); può essere collegato a <change> tramite @ana o @change.
	Funzione principale: indica una cancellatura o soppressione di testo. Attributi principali: @rend, @hand, @type Collocazione nel documento TEI: all'interno del corpo testuale (<body> o <text>). Relazione con altri elementi: può comparire da solo o insieme a <add> all'interno di <subst>; può rimandare a una revisione documentata in <change>.
<subst>	Funzione principale: rappresenta una sostituzione complessa, in cui un testo è cancellato e un altro aggiunto in sua vece. Attributi principali: (nessuno specifico, ma contiene e <add>). Collocazione nel documento TEI: all'interno del corpo testuale (<body> o <text>). Relazione con altri elementi: contiene e <add>; le sue parti possono essere collegate a una revisione in <change>.
<choice>	Funzione principale: raggruppa varianti o alternative di lezione (forme originali, corrette, regolarizzate, ecc.). Attributi principali: nessuno obbligatorio (spesso usato con <sic>, <corr>, <orig>, <reg>). Collocazione nel documento TEI: nel corpo testuale o in apparati di variante. Relazione con altri elementi: può coesistere con <seg> per distinguere fasi redazionali; le scelte possono riferirsi a revisioni indicate in <change>.
<seg>	Funzione principale: delimita un segmento testuale di particolare interesse (ad esempio, una revisione o una fase compositiva). Attributi principali: @type, @ana Collocazione nel documento TEI: all'interno del corpo testuale (<body> o <text>). Relazione con altri elementi: collega porzioni di testo ai metadati descrittivi in <listChange> o a singoli <change> mediante @ana.
<listChange>	Funzione principale: elenca in modo strutturato le revisioni o campagne redazionali di un testo. Attributi principali: nessuno specifico (contiene <change>). Collocazione nel documento TEI: nel corpo del documento (<text>), di norma in <back> o <standOff>. Relazione con altri elementi: contiene <change>; collegato ai segmenti testuali tramite @ana o @change presenti in <seg>, <add>, .
<change>	Funzione principale: descrive una singola revisione, correzione o campagna redazionale (autoriale o editoriale). Attributi principali: @who, @when, @notBefore, @notAfter, @n, @type Collocazione nel documento TEI: all'interno di <listChange> oppure di <revisionDesc>. Relazione con altri elementi: identificato da xml:id, può essere richiamato da <seg>, <add>, tramite @ana o @change.

Elemento	Descrizione
<revisionDesc>	Funzione principale: documenta le revisioni e modifiche del file TEI o dell'edizione digitale. Attributi principali: nessuno specifico (contiene <change>). Collocazione nel documento TEI: all'interno del <teiHeader>. Relazione con altri elementi: parallelo a <listChange>: descrive la storia editoriale digitale, mentre <listChange> descrive la storia redazionale autoriale.

4.2 Pubblicazione online dell'edizione genetica digitale dei frammenti autografi

Una volta completata l'operazione di codifica testuale, il processo di modellizzazione della nostra edizione scientifica digitale giunge al suo compimento con la selezione dello strumento di visualizzazione più idoneo agli obiettivi accademici prefissati e con la conseguente pubblicazione dell'edizione stessa. In tale fase, la scelta dell'ambiente di pubblicazione è stata orientata dal principio di equilibrio fra rigore scientifico, accessibilità e sostenibilità tecnologica. Alla luce della distinzione proposta da Elena Pierazzo tra edizioni *haute couture* e *prêt-à-porter*, abbiamo deliberatamente optato per una soluzione che privilegiasse la funzionalità e la sostenibilità rispetto alla personalizzazione estrema. Secondo Pierazzo, infatti, le edizioni *haute couture* si configurano come prodotti altamente specializzati, costruiti su misura in relazione a esigenze specifiche di ricerca, e caratterizzati da un elevato grado di libertà progettuale, ma anche da costi di sviluppo e manutenzione considerevoli. Le edizioni *prêt-à-porter*, al contrario, si fondano su modelli e infrastrutture preesistenti, meno flessibili ma più facilmente accessibili e sostenibili nel medio e lungo periodo, sia dal punto di vista tecnico sia economico.²⁴⁹

In tale prospettiva, abbiamo individuato in *Transkribus Sites* una soluzione adeguata alla realizzazione della nostra edizione digitale. Questo servizio, integrato nell'ecosistema di *Transkribus*, offre una piattaforma *user-friendly* che consente la pubblicazione online di trascrizioni e immagini sincronizzate, senza richiedere competenze avanzate di sviluppo web o di programmazione. L'interfaccia intuitiva e la gestione semplificata dei contenuti rappresentano un indubbio vantaggio per progetti accademici di dimensioni medio-piccole, facilitando la fruizione dei risultati della ricerca anche a un pubblico non specialistico. Dal punto di vista

²⁴⁹ Elena Pierazzo, *Digital Scholarly Editing: Theories and Practices* (Farnham, Ashgate, 2015), pp. 61–73.

dell'utente finale, l'accesso a un'edizione pubblicata tramite *Transkribus Sites* risulta estremamente agevole. L'edizione può essere resa pubblica, divenendo così accessibile mediante un semplice collegamento web, senza la necessità di autenticazione o dell'installazione di software specifici. Una volta aperto il sito, l'utente ha la possibilità di consultare le immagini digitalizzate e le trascrizioni allineate, di effettuare ricerche testuali, di visualizzare i metadati e di navigare tra i documenti attraverso un'interfaccia chiara, responsiva e intuitiva. Nel caso in cui il curatore opti per una modalità di accesso riservato, l'utente dovrà invece disporre di un account *Transkribus* e ottenere l'autorizzazione all'accesso da parte del responsabile del progetto. Questa configurazione è generalmente adottata per edizioni ancora in fase di revisione o soggette a vincoli di diritto d'autore. L'architettura di accesso differenziata consente di bilanciare in modo efficace le esigenze di disseminazione dei risultati scientifici con quelle di tutela e controllo dei contenuti. Tuttavia, la natura commerciale del servizio comporta alcune limitazioni: la dipendenza da un'infrastruttura esterna e l'assenza di pieno controllo sul codice e sui dati potrebbero rappresentare un vincolo per chi miri alla completa autonomia editoriale o desideri garantire la piena riproducibilità e trasparenza scientifica del processo di pubblicazione.²⁵⁰

Come possibile alternativa a soluzioni commerciali per la pubblicazione di edizioni digitali, il software EVT (*Edition Visualization Technology*) rappresenta una piattaforma *open source* altamente personalizzabile, più prossima al paradigma *haute couture*, pur mantenendo un equilibrio fra rigore metodologico e accessibilità tecnica. EVT costituisce infatti uno degli strumenti più significativi sviluppati nell'ambito della filologia digitale per la creazione e la navigazione di edizioni scientifiche basate su codifica TEI/XML, in conformità con gli standard e le *Guidelines* del *Text Encoding Initiative Consortium*. Il progetto nasce nell'ambito del *Vercelli Book Digitale* (VBD), promosso dall'Università di Pisa sotto la direzione di Roberto Rosselli Del Turco, con la collaborazione di un team interdisciplinare attivo nei settori della filologia digitale e dell'informatica umanistica. L'iniziativa è finalizzata alla creazione di un'edizione digitale del *Vercelli Book* (Archivio e Biblioteca Capitolare di Vercelli, ms. CXVII), un codice membranaceo della fine del X secolo, tra i più importanti testimoni della tradizione poetica in lingua inglese antica. Proprio per la sua rilevanza storica e linguistica, il progetto ha

²⁵⁰ La documentazione ufficiale su *Transkribus Sites*, è disponibile al seguente link: <<https://www.transkribus.org/it/sites>> (ultimo accesso 10 ottobre 2025).

richiesto uno strumento capace di garantire longevità tecnologica, trasparenza editoriale e indipendenza dalle piattaforme proprietarie.²⁵¹

Fin dalle origini, EVT è stato sviluppato secondo i principi dell'*open source* e della sostenibilità a lungo termine, ed è costruito interamente su tecnologie web standard (HTML, CSS, JavaScript), per garantire la compatibilità con i principali browser e la persistenza nel tempo sulla rete. Alcune componenti specifiche, come il visualizzatore di immagini, si basano su moduli esterni *open source* ampiamente supportati, tra cui *Openseadragon* per la gestione dello zoom e della navigazione, in modo da ridurre il rischio di incompatibilità future.²⁵² L'architettura generale del software è modulare, così che ogni componente possa essere sostituita o aggiornata facilmente senza compromettere la stabilità complessiva del sistema.²⁵³

4.2.1 EVT 1 (2010–2019)

La prima generazione del software, nota come *EVT 1*, è stata sviluppata tra il 2010 e il 2013 e rilasciata ufficialmente nel 2013, raggiungendo la sua versione finale (EVT 1.3) nel dicembre 2019. Questa versione è dedicata prevalentemente a edizioni diplomatiche o critiche su testimone unico ed è stata utilizzata con successo in diversi progetti, tra cui l'edizione digitale del *Codice Pelavicino*.²⁵⁴ EVT 1 si basa su una catena di trasformazioni XSLT 2.0 applicate ai file TEI/XML, che generano automaticamente un sito web statico pronto per la consultazione locale o online.²⁵⁵ La struttura modulare del software si articola attorno al foglio principale *evt_builder.xsl*, il quale coordina una serie di moduli XSLT contenuti nella directory *builder_pack*. Tali moduli si dividono in due categorie:

²⁵¹ La documentazione sul *Vercelli Book Digitale*, è disponibile al seguente link: <<http://vbd.humnet.unipi.it/>> (ultimo accesso 10 ottobre 2025).

²⁵² La documentazione sul visualizzatore *Openseadragon* è disponibile al seguente link: <<https://openseadragon.github.io/>> (ultimo accesso 10 ottobre 2025).

²⁵³ È possibile consultare il sito web del progetto al seguente link: <<http://evt.labcd.unipi.it>> (ultimo accesso 10 ottobre 2025). La repository GitHub generale è consultabile al seguente link: <<https://github.com/evt-project>> (ultimo accesso 10 ottobre 2025).

²⁵⁴ La documentazione sul *Codice Pelavicino* è disponibile al seguente link: <<https://pelavicino.labcd.unipi.it/il-progetto/il-software-evt/>> (ultimo accesso 10 ottobre 2025).

²⁵⁵ È opportuno ricordare che XSLT (*Extensible Stylesheet Language Transformations*) è un linguaggio destinato alla trasformazione di documenti XML in altri formati, quali XML con struttura differente, HTML o testo semplice. EVT 1 si fonda su una catena di trasformazioni XSLT 2.0, che prevede l'applicazione sequenziale di più fogli di stile: ciascuna trasformazione utilizza come input l'output della precedente, modificando progressivamente il documento.

1. moduli di costruzione, responsabili della generazione del sito HTML;
2. moduli di elaborazione, incaricati di estrarre e formattare il testo dell'edizione a partire dai marcatori di pagina (<pb/>).

Le edizioni realizzate con EVT 1 supportano due livelli testuali – diplomatico e diplomatico-interpretativo – e garantiscono un'elevata compatibilità con i file TEI, codificati secondo gli standard comunemente adottati, senza la necessità di interventi di personalizzazione complessi. Le funzionalità principali comprendono la sincronizzazione testo-immagine, lo zoom, la lente d'ingrandimento e la navigazione dei facsimili, assicurando un'interazione fluida e intuitiva tra il lettore e il documento digitale. Nel corso della sua evoluzione, EVT 1 ha progressivamente ampliato le proprie potenzialità, integrando strumenti concepiti per migliorare l'esperienza di consultazione e l'efficacia filologica delle edizioni digitali. Tra le principali innovazioni si annoverano:

- Il supporto per le traduzioni, che consente di affiancare alla trascrizione del testo originale una o più versioni tradotte, visualizzabili in modo sincronizzato o alternato. Tale funzione risponde all'esigenza di rendere l'edizione accessibile anche a un pubblico non specialista, pur preservando la struttura critica e l'impianto ecdotico del testo di partenza.
- La visualizzazione di versi come prosa, che permette di mostrare un testo poetico in forma continua, senza interruzioni di verso, pur mantenendo la codifica TEI dei singoli versi (<l>, <lg>). Questa modalità facilita una lettura linguistica o traduttiva del contenuto, indipendente dalla sua struttura metrica, offrendo una prospettiva di consultazione complementare a quella filologico-metrica.
- L'integrazione di *VisColl* (*Visual Collation Modeler*), che consente di rappresentare graficamente la struttura materiale del manoscritto, visualizzando la composizione dei fascicoli, delle carte e dei bifogli. Tale funzionalità, sviluppata originariamente presso l'Università della Pennsylvania, rende possibile documentare e comprendere la fisicità e l'organizzazione codicologica del testimone digitale, arricchendo l'edizione di una dimensione descrittiva e strutturale di tipo codicologico.²⁵⁶

²⁵⁶ Per maggiori informazioni su *VisColl*, è possibile consultare il seguente link: <<https://viscoll.org/>> (ultimo accesso 10 ottobre 2025).

Tuttavia, la crescente complessità del codice XSLT e la difficoltà di estendere ulteriormente il framework hanno portato, dopo il 2019, alla necessità di una completa riscrittura del software.²⁵⁷

4.2.2 EVT 2 (2016–2020) ²⁵⁸

La seconda generazione, EVT 2, rappresenta un vero e proprio *reboot* del progetto, basato su una nuova architettura MVC (*Model-View-Controller*) implementata tramite il framework *AngularJS*.²⁵⁹ La beta 2 è stata pubblicata nel giugno 2020, ma lo sviluppo era iniziato già nel 2016. EVT 2 introduce una gestione più flessibile dei dati TEI, grazie a *parser* dedicati in grado di convertirli in strutture JSON temporanee, rendendo possibile la manipolazione dinamica dei contenuti e la visualizzazione comparata di più testimoni (cfr. Figura 4.1).²⁶⁰

²⁵⁷ R. Rosselli Del Turco, *Designing an advanced software tool for Digital Scholarly Editions. The inception and development of EVT (Edition Visualization Technology)*, «Textual Cultures», 12 (2), 2019, pp. 91-111; R. Rosselli Del Turco e C. Di Pietro, *Between Innovation and Conservation: The Narrow Path of User Interface Design for Digital Scholarly Editions*, in *Digital Scholarly Editions as Interfaces*, vol. 12, Norderstedt, BoD, 2018, pp. 133–163; R. Rosselli Del Turco et al., *Edition Visualization Technology: A Simple Tool to Visualize TEI-based Digital Editions*, «Journal of the Text Encoding Initiative», 8, 2014.

²⁵⁸ La repository EVT 2 è disponibile al seguente link: <<https://github.com/evt-project/evt-viewer>> (ultimo accesso 10 ottobre 2025).

²⁵⁹ L'architettura *Model-View-Controller* (MVC) è un modello di progettazione software che separa un'applicazione in tre componenti principali, al fine di favorire la modularità, la manutenibilità e la riusabilità del codice.

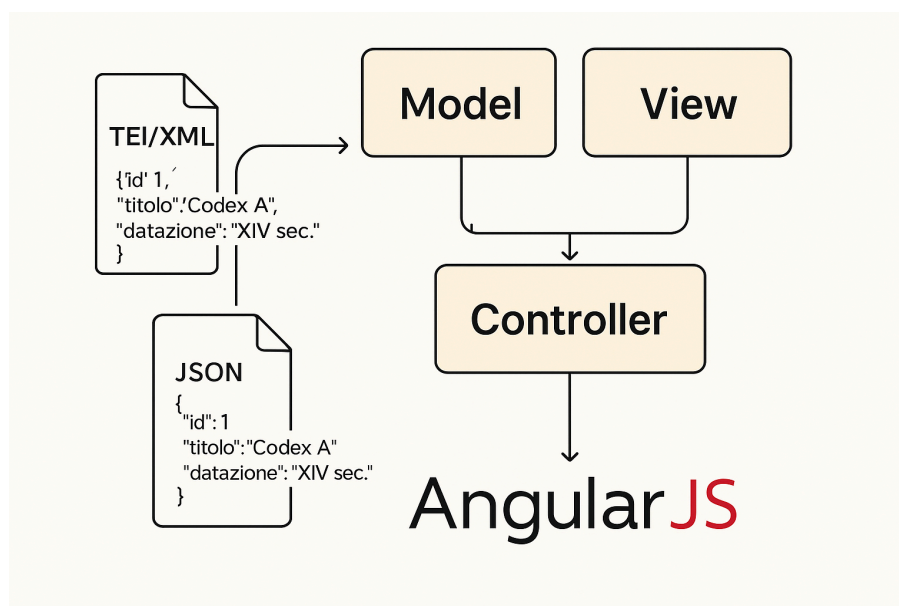
- *Model*: gestisce i dati e la logica applicativa (in EVT, i dati TEI convertiti in JSON);
- *View*: si occupa della presentazione visiva e dell'interfaccia utente;
- *Controller*: funge da intermediario tra Model e View, ricevendo gli input dell'utente, aggiornando i dati e richiedendo il rendering della vista aggiornata.

Questa architettura è alla base di *AngularJS*, un framework JavaScript open source sviluppato da Google, che consente di creare applicazioni web dinamiche in cui i dati e l'interfaccia utente si aggiornano automaticamente e in tempo reale grazie al *data binding* (collegamento bidirezionale tra *Model* e *View*). In EVT 2, *AngularJS* fornisce la struttura logica per gestire la complessità del progetto e migliorare l'interattività della visualizzazione.

²⁶⁰ È opportuno ricordare che un *parser* è un componente software che analizza un documento strutturato (come un file TEI/XML) e lo traduce in una rappresentazione interna comprensibile e manipolabile dal programma. Nel contesto di EVT 2, i parser dedicati interpretano i file TEI e ne estraggono le informazioni rilevanti, preparandole per la trasformazione in un formato più flessibile come JSON.

Il JSON (*JavaScript Object Notation*) è un formato di rappresentazione dei dati leggero e leggibile, basato su coppie chiave-valore e strutture annidate (oggetti e array). È ampiamente utilizzato per lo scambio di dati tra client e server, soprattutto nelle applicazioni web. In EVT 2, i dati TEI vengono temporaneamente convertiti in strutture JSON per poter essere gestiti e manipolati dinamicamente dal framework *AngularJS*.

Figura 4.1 – L'architettura Model-View-Controller e il framework AngularJS



La distinzione tra dati testuali e strumenti di visualizzazione consente di separare nettamente la componente filologica da quella tecnologica, favorendo la riusabilità e l'interoperabilità dei dati. Tra le principali innovazioni figurano:²⁶¹

- un modulo avanzato di collegamento testo-immagine;
- l'integrazione del protocollo IIIF (*International Image Interoperability Framework*), che consente la connessione diretta con repository di immagini distribuiti;
- un'interfaccia più configurabile e personalizzabile;
- e una maggiore stabilità nella gestione dei grandi corpora di immagini e trascrizioni.

In particolare, l'integrazione del protocollo IIIF rappresenta una delle evoluzioni più significative introdotte con EVT 2. Questa innovazione consente al software di interagire direttamente con repository di immagini digitali conformi allo standard IIIF, superando i limiti della gestione locale dei facsimili e aprendo a un'infrastruttura realmente distribuita e interoperabile. L'adozione di IIIF offre un duplice vantaggio: da un lato, la possibilità di caricare e visualizzare in tempo reale immagini ospitate su server esterni, evitando la duplicazione dei

²⁶¹ R. Rosselli Del Turco e C. Di Pietro, *La visualizzazione di edizioni digitali con EVT: una soluzione per edizioni diplomatiche e critiche*, «Ecdotica», 16, 2020, pp. 148–173; R. Rosselli Del Turco, C. Di Pietro e C. Martignano, *Progettazione e implementazione di nuove funzionalità per EVT 2: lo stato attuale dello sviluppo*, «Umanistica Digitale», 7, 2019.

dati e favorendo l'uso di risorse condivise; dall'altro, la potenzialità di integrare metadati semantici mediante tecnologie *Linked Open Data* (LOD), che arricchiscono la rappresentazione testuale con riferimenti espliciti a persone, luoghi e opere. La combinazione TEI/IIIF implementata in EVT si basa su una logica di *mapping* tra markup testuale e risorse iconografiche interoperabili: ogni porzione del manoscritto, identificata nel TEI come <zone> o <surface>, può essere collegata a un frammento dell'immagine definito nel manifest IIIF. Questo approccio trasforma il visualizzatore in un ambiente semantico integrato, nel quale la componente grafica e quella testuale dialogano dinamicamente, restituendo al lettore la complessità stratificata dell'oggetto editoriale digitale.²⁶²

4.2.3 EVT 3 (2021–oggi) ²⁶³

La terza generazione, EVT 3, è in corso di sviluppo dal 2021 e ha visto la pubblicazione di una versione alpha basata su Angular 9 all'inizio del 2021, seguita da un primo prototipo stabile nel 2023. Questa nuova versione segna il passaggio definitivo a *TypeScript* e a una struttura interamente modulare, mantenendo il paradigma MVC già introdotto con EVT 2.²⁶⁴ L'obiettivo principale di EVT 3 è fornire una piattaforma interdisciplinare e scalabile, capace di gestire non solo documenti TEI, ma anche formati e schemi alternativi. Ogni componente (viewer, editor, apparato critico, bookreader) è concepita come modulo indipendente, riutilizzabile in contesti diversi. Le funzioni già operative includono la visualizzazione sinottica multi-pannello, la gestione avanzata del collegamento testo-immagine, la visualizzazione in modalità *bookreader* e le prime applicazioni in ambito di filologia d'autore e critica genetica. EVT 3 si configura dunque come una piattaforma di riferimento per la filologia digitale di nuova generazione,

²⁶² R. Rosselli Del Turco e P. Monella, *Extending the DSE: LOD Support and TEI/IIIF Integration in EVT*, in *Atti del IX Convegno Annuale AIUCD. La svolta inevitabile: sfide e prospettive per l'informatica umanistica* (Milano, 15-17 gennaio 2020), a cura di Cristina Marras et al., pp. 148–55.

²⁶³ La repository EVT 3 è disponibile al seguente link: <<https://github.com/evt-project/evt-viewer-angular>> (ultimo accesso 10 ottobre 2025).

²⁶⁴ Il *TypeScript* è un linguaggio di programmazione *open source* sviluppato da Microsoft, che estende JavaScript introducendo un sistema di tipi statici e altre funzionalità pensate per rendere più sicuro, leggibile e strutturato lo sviluppo di applicazioni complesse. I tipi statici consentono di dichiarare e controllare il tipo di dato (ad esempio stringa, numero, booleano, oggetto, ecc.) già in fase di compilazione, prima che il programma venga eseguito. In questo modo è possibile individuare errori e incongruenze nel codice in anticipo, migliorando l'affidabilità del software e riducendo i bug dovuti a operazioni su dati non compatibili. Le versioni moderne di *Angular* (successive ad *AngularJS*) sono state riscritte interamente in *TypeScript*, poiché il linguaggio favorisce un'organizzazione modulare del codice e un controllo rigoroso dei tipi, caratteristiche fondamentali per la gestione di applicazioni web di grandi dimensioni, come EVT 2, che si basa su architetture analoghe.

orientata alla collaborazione, alla trasparenza dei processi e alla piena interoperabilità con altre infrastrutture accademiche e archivistiche.²⁶⁵

In conclusione, la preparazione di un'edizione digitale richiede l'applicazione congiunta di strumenti di marcatura e ambienti di sviluppo capaci di garantire una gestione coerente e validata dei documenti XML. Tra gli editor disponibili, *Oxygen XML Editor* rappresenta la soluzione più completa e diffusa, sebbene sia un software a pagamento, utilizzabile gratuitamente solo per un periodo di prova di trenta giorni. In alternativa, *Visual Studio Code*, integrato con l'estensione *open source Scholarly XML* sviluppata da R. Viglianti, si è rivelato l'opzione più flessibile e performante in ambito gratuito, consentendo la validazione automatica dei file tramite schemi *RelaxNG* e l'integrazione di ulteriori estensioni dedicate. Per quanto riguarda la realizzazione dell'edizione digitale, abbiamo inoltre testato EVT 2, seguendone le principali fasi operative:

1. la codifica del testo secondo gli standard TEI XML;
2. l'inserimento dei dati nelle relative *directory del progetto* (ad esempio, le immagini nella cartella *data/images/single*);
3. la configurazione del sistema mediante la modifica del file *config.json* e, facoltativamente, del foglio di stile *custom-style.css*;
4. infine, la verifica dell'edizione attraverso il caricamento del file *index.html* in un browser, secondo le istruzioni previste per la navigazione locale.

L'esperienza condotta con *Visual Studio Code* ed *EVT 2* ha confermato la solidità e l'affidabilità di questa metodologia, che si configura come una prassi efficace e pienamente conforme agli standard internazionali per la realizzazione di edizioni digitali testualmente e strutturalmente coerenti. Tuttavia, è opportuno ricordare che, qualora si scelga di utilizzare il software *EVT*, sarà necessario predisporre uno spazio di *hosting* adeguato (con o senza dominio dedicato). *EVT*, infatti, genera una versione statica dell'edizione digitale, costituita da file in formato *HTML*, *CSS*, *JavaScript* e dalle relative immagini. Tali file devono essere caricati su un server web per poter essere consultabili online; risulta dunque indispensabile disporre di un servizio di *hosting*, che può consistere:

²⁶⁵ R. Rosselli Del Turco et al., *There and back again: what to expect in the next EVT version*, in *AIUCD 2022 - Intersezioni: filosofia, arti, media*, Atti dell'XI Convegno Nazionale dell'AIUCD (Lecce, 1-3 giugno 2022), a cura di F. Ciraci et al., AIUCD, 2022, pp. 212-217.

- in un *hosting* istituzionale, come un server universitario o bibliotecario;
- in una piattaforma gratuita, ad esempio *GitHub Pages* o *GitLab Pages*;
- oppure in un dominio dedicato, nel caso in cui si desideri dotare il progetto di un indirizzo personalizzato e autonomo.

4.3 Disseminazione online dei risultati della ricerca: il sito web dedicato all'edizione digitale

Nell'ambito della ricerca dottorale dedicata all'edizione digitale dei frammenti autografi dell'*Orlando furioso* di Ludovico Ariosto, è stato realizzato un sito web concepito come strumento di disseminazione scientifica e di valorizzazione dei risultati conseguiti. Il progetto è stato ideato come prodotto digitale complementare alla tesi di dottorato, in un'ottica di integrazione transmediale tra testo accademico e piattaforma interattiva.²⁶⁶ Il sito non sostituisce dunque l'elaborato scritto, ma ne rappresenta un'estensione, progettata per potenziare la fruizione, la consultabilità e la condivisione dei contenuti di ricerca in ambiente digitale. Attualmente, il sito risulta completo e pienamente operativo all'interno dell'editor del CMS (*Content Management System*) adottato per lo sviluppo, ossia Wix.com. Tuttavia, è richiesta la predisposizione di uno spazio web per la pubblicazione del sito di presentazione dell'edizione, eventualmente associato a un dominio dedicato, se ritenuto opportuno. L'acquisto di un dominio non rappresenta comunque un requisito imprescindibile, poiché molte istituzioni quali, università, biblioteche o enti di ricerca, offrono già spazi web o sottodomini riservati ai progetti accademici. In alternativa, è possibile ricorrere a servizi di *hosting* gratuiti, che forniscono indirizzi standard, oppure collocare il sito di presentazione all'interno dello stesso dominio che ospita l'edizione digitale, qualora ciò risulti tecnicamente e istituzionalmente appropriato.

L'architettura informativa del sito è stata progettata secondo criteri di chiarezza, coerenza semantica e usabilità, con l'obiettivo di agevolare la navigazione e garantire un accesso intuitivo ai contenuti testuali e multimediali.²⁶⁷ La struttura si articola in tre sezioni principali:

²⁶⁶ Per approfondire, cfr. D. Fiormonte, *Digital Humanities: Teorie, metodi, strumenti*, Roma, Carocci, 2014.

²⁶⁷ Per approfondire, cfr. J. Nielsen, *Designing Web Usability: The Practice of Simplicity*, Indianapolis, New Riders, 2000.

1. *Header* (intestazione): costituisce la parte superiore della pagina e include il logo, il titolo dell'edizione digitale e un menu di navigazione progettato per consentire un passaggio agevole tra le diverse aree del sito. Il menu è organizzato in modo gerarchico, così da rendere immediata la transizione tra i contenuti e assicurare un'esperienza d'uso coerente (cfr. Fig. 4.1).
2. *Corpo principale* (homepage e pagine secondarie): rappresenta l'area centrale del sito e comprende due sezioni distinte:
 - una sezione dedicata alla metodologia editoriale adottata, articolata in cinque fasi corrispondenti ad altrettante pagine. Ciascuna fase è accessibile sia tramite i collegamenti presenti nel menu a tendina sia attraverso i pulsanti associati alle icone dell'homepage. Ogni pagina contiene un abstract del capitolo corrispondente della tesi di dottorato, disponibile per il download in formato PDF (cfr. Figg. 4.2, 4.3, 4.4, 4.5);
 - una sezione dedicata alla piattaforma dell'edizione digitale vera e propria, accessibile mediante un apposito pulsante collocato nella stessa area (cfr. Fig. 4.6).
3. *Footer* (piè di pagina): costituisce la parte inferiore del sito e raccoglie i contatti, le icone delle istituzioni accademiche o culturali potenzialmente coinvolgibili in partnership scientifiche, nonché le informazioni di copyright e la dichiarazione di accessibilità, redatta in conformità con le *Web Content Accessibility Guidelines* (WCAG) (cfr. Fig. 4.7).²⁶⁸

²⁶⁸ Per approfondire, cfr. World Wide Web Consortium (W3C), *Web Content Accessibility Guidelines (WCAG) 2.1*, 2028, disponibile al seguente link: < <https://www.w3.org/TR/WCAG21/> >

Figura 4.1 - header del sito web

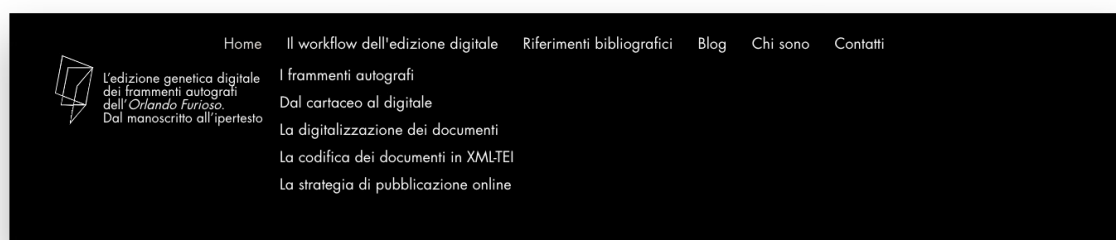


Figura 4.2 - homepage, sezione dedicata alla metodologia editoriale

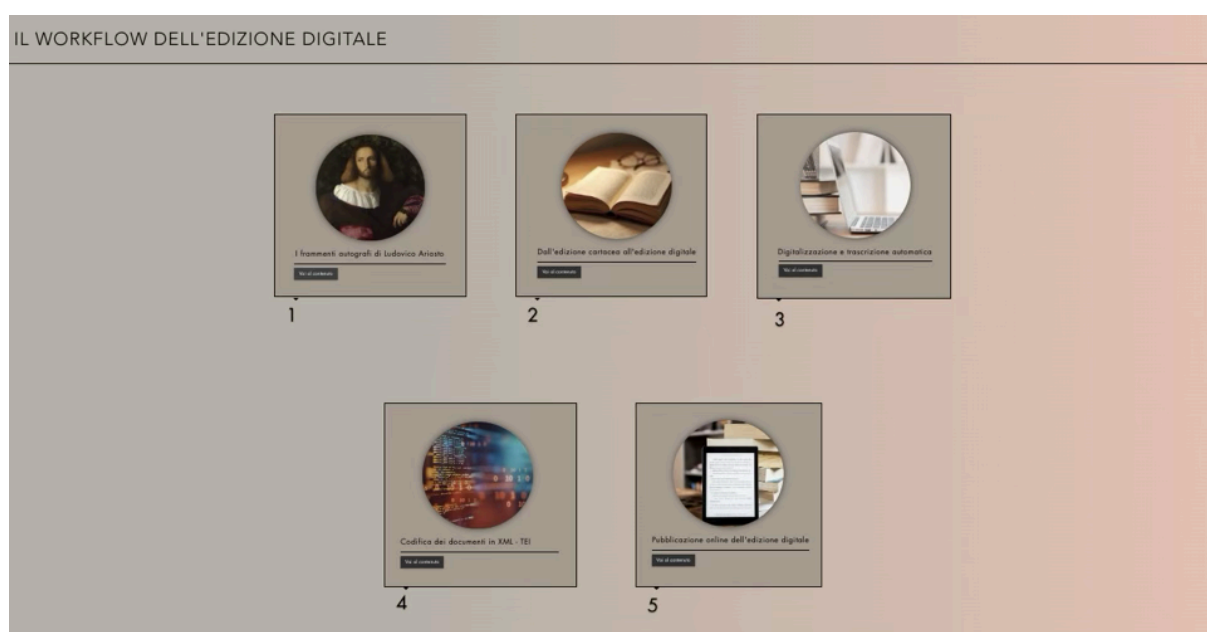


Figura 4.3 - pagina secondaria dedicata ai *frammenti autografi*

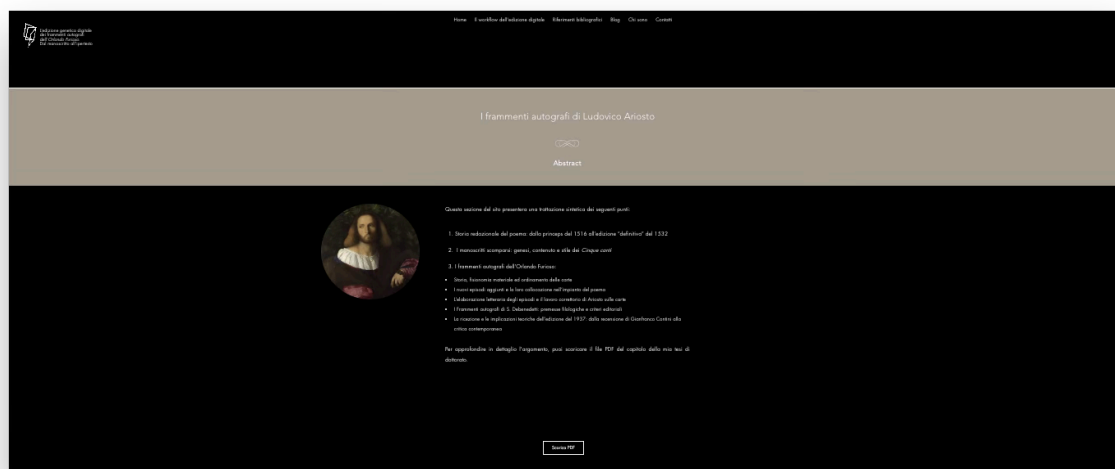


Figura 4.4 - pagina secondaria dedicata ai riferimenti bibliografici

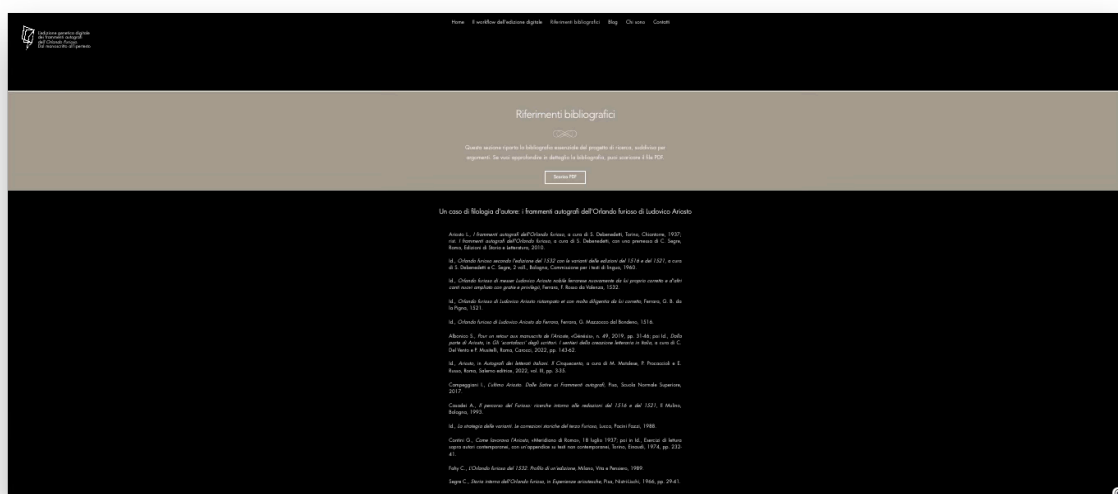


Figura 4.5 - pagina secondaria dedicata al blog

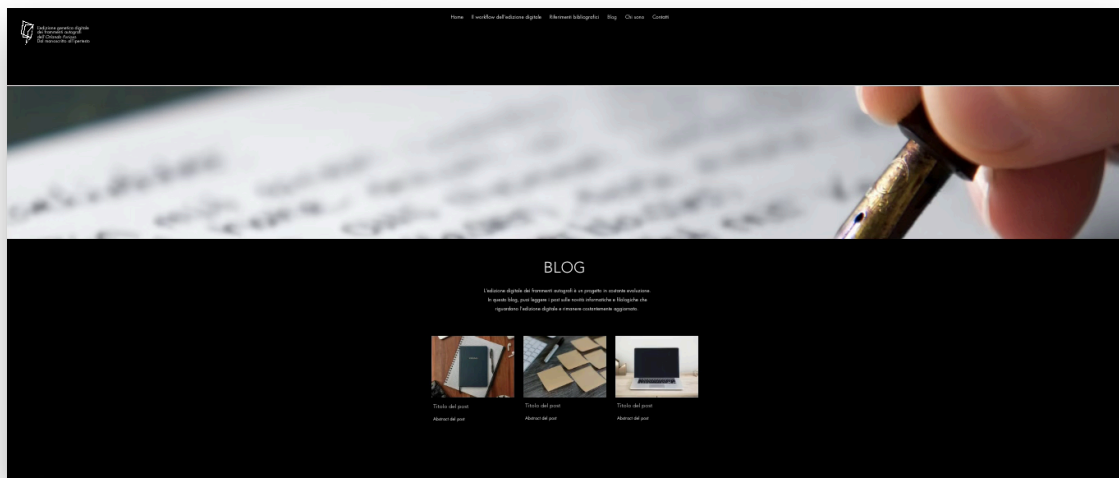


Figura 4.6 - homepage, sezione dedicata alla piattaforma dell'edizione digitale

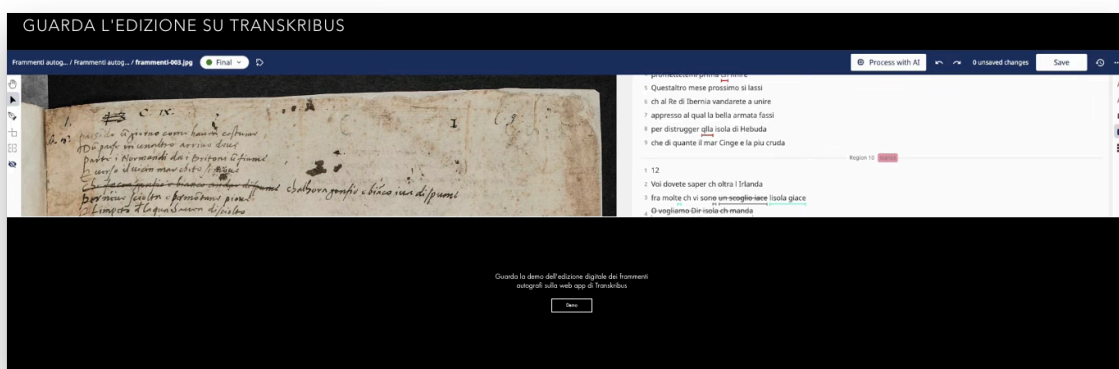
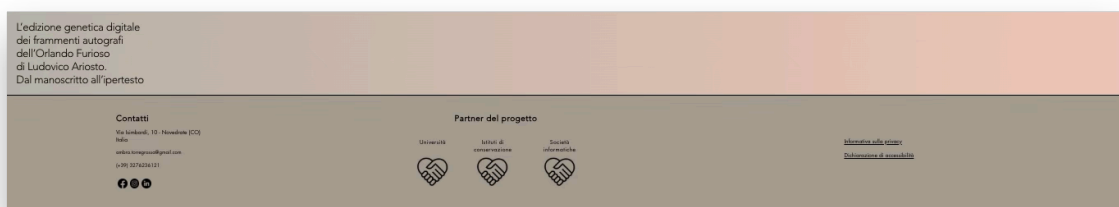


Figura 4.7 - footer del sito web



Dal punto di vista concettuale e metodologico, il sito è stato concepito come una vera e propria “rimediazione digitale” della tradizionale nota al testo tipica delle edizioni critiche a stampa.²⁶⁹ È opportuno ricordare, infatti, che l’edizione proposta si configura come una *scholarly edition*, vale a dire una pubblicazione fondata su criteri metodologici e filologici rigorosi, documentati e riproducibili.²⁷⁰ Tale impostazione risponde al principio di verificabilità e replicabilità che costituisce il fondamento della ricerca scientifica. Nelle edizioni critiche a stampa, la verificabilità delle scelte editoriali – intesa come possibilità per il lettore di risalire alle decisioni dell’editore attraverso l’analisi delle varianti d’autore o di copista – si fonda tradizionalmente sulla nota al testo e sull’apparato critico, strumenti che consentono di ricostruire e, se necessario, discutere le scelte operate. Nell’ambito delle edizioni digitali, invece, la verificabilità dei risultati assume una configurazione più articolata, distribuita su due livelli complementari: *front-end* e *back-end* del sito web.²⁷¹ Sul piano del *front-end*, l’affidabilità scientifica è garantita dalla presenza di un’interfaccia che funge da nota al testo digitale, integrata da un apparato critico costruito secondo criteri coerenti con la tradizione filologica.²⁷² Sul piano del *back-end*, la possibilità di accedere al codice sorgente dell’edizione e di analizzare il modello di codifica adottato dall’editore costituisce un ulteriore livello di trasparenza. Di conseguenza, la verifica si estende anche alla dimensione infrastrutturale, consentendo di valutare la coerenza tra la rappresentazione dei dati e i principi metodologici dichiarati, e di promuovere una piena tracciabilità scientifica dell’intero processo editoriale.²⁷³

Concludendo, è opportuno sottolineare la necessità di monitorare attentamente la diffusione digitale dei risultati della ricerca, in particolare dei capitoli della tesi dottorale resi disponibili in formato PDF attraverso il sito web, al fine di prevenire eventuali usi impropri o fenomeni di plagio. A tale scopo, possono essere adottate strategie integrate di natura tecnica, legale e analitica, volte a garantire un equilibrio tra l’accessibilità dei contenuti, la tutela dei diritti

²⁶⁹ Per approfondire, cfr. J.D. Bolter e R. Grusin, *Remediation: Understanding New Media*, Cambridge (MA), MIT Press, 1999.

²⁷⁰ Per approfondire, cfr. P. Sahle, *What is a Scholarly Digital Edition?* in *Digital Scholarly Editing: Theories and Practices*, a cura di M. Boot et al. Cambridge, Open Book Publishers, 2016, pp.

²⁷¹ Per approfondire, cfr. E. Pierazzo, *Digital Scholarly Editing: Theories, Models and Methods*, Farnham, Ashgate, 2015.

²⁷² Per approfondire, cfr. J. McGann, *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction*, Cambridge (MA), Harvard University Press, 2013.

²⁷³ Per approfondire, cfr. R. Siemens e S. Schreibman, *A Companion to Digital Literary Studies*, Oxford, Blackwell, 2013.

d'autore e il monitoraggio dell'impatto scientifico delle pubblicazioni digitali.²⁷⁴ Una prima misura di tutela potrebbe consistere nell'inserimento di un watermark digitale, visibile o invisibile, all'interno dei file PDF distribuiti.²⁷⁵ Tale marcatura consentirebbe di associare in modo permanente al documento alcune informazioni identificative, come il nome dell'autore, la data di pubblicazione, il DOI o l'URL del sito di riferimento, nonché la licenza d'uso (ad esempio, *Creative Commons CC BY-NC-ND 4.0*), senza comprometterne la leggibilità.²⁷⁶ Per la realizzazione dei watermark si possono impiegare strumenti commerciali, come *Adobe Acrobat Pro* o *PDF-XChange Editor*, oppure soluzioni open source, come *pdftk* e *Ghostscript*, che consentono di automatizzare il processo e integrarlo nei flussi di pubblicazione digitale.

Parallelamente, è opportuno rendere esplicite le condizioni d'uso dei documenti attraverso note di copyright e licenze chiaramente visibili. In ciascun file PDF si potrebbe inserire, nella prima pagina e nel piè di pagina, una dichiarazione che specifichi i diritti dell'autore e le limitazioni d'uso, come nel seguente esempio:

© 2025, Ambra Torregrossa. Tutti i diritti riservati.

Il documento è disponibile per consultazione e citazione scientifica secondo la licenza CC BY-NC-ND 4.0.

È consentita la riproduzione e citazione a fini scientifici, con obbligo di attribuzione e divieto di uso commerciale o modifica del contenuto.

Le medesime informazioni dovrebbero essere incluse anche nei metadati del file PDF e nel codice HTML del sito, così da garantirne la tracciabilità e la lettura automatica da parte dei motori di ricerca e dei repository accademici. L'impiego di licenze aperte ma regolamentate,

²⁷⁴ Per approfondire, cfr. i seguenti testi: P. Bertolini, *Monitoraggio digitale e tutela dell'autorialità: watermark, fingerprint e metriche d'impatto nella comunicazione scientifica*, Roma, Carocci Editore, 2023; G. Serra e C. Ricci, *Gestione dei diritti e monitoraggio dell'impatto delle pubblicazioni scientifiche online*, Firenze, Firenze University Press, 2022; L. De Santis e A. Gallo, *Metadati, identità digitale e tracciabilità della ricerca: strumenti e metodologie per l'editoria accademica*, Milano, Ledizioni, 2021.

²⁷⁵ Per approfondire, cfr. F. Montanari, *Watermarking e protezione dei documenti digitali: tecniche, strumenti e applicazioni*, Roma, FrancoAngeli, 2018.

²⁷⁶ La licenza *Creative Commons Attribution – Non Commercial – No Derivatives 4.0 International* (CC BY-NC-ND 4.0) consente la libera consultazione e condivisione dell'opera, a condizione che venga sempre riconosciuta la paternità dell'autore. Essa vieta tuttavia ogni uso a fini commerciali e qualsiasi forma di modifica, adattamento o elaborazione del contenuto. Si tratta di una delle licenze più restrittive tra quelle proposte da Creative Commons, in quanto tutela l'integrità dell'opera pur favorendone la diffusione per finalità di studio e ricerca. L'obiettivo è garantire un equilibrio tra l'accesso aperto ai risultati scientifici e la salvaguardia dei diritti morali e patrimoniali dell'autore, nel rispetto delle norme internazionali sul diritto d'autore.

come quelle proposte dal progetto *Creative Commons*, permetterebbe di conciliare la logica dell'*open access* con la tutela del diritto d'autore.²⁷⁷

Oltre alla protezione dei file, è utile implementare un sistema di monitoraggio dei download e del traffico web per valutare la diffusione dei materiali e il loro impatto sulla disseminazione scientifica. Su *Wix.com* questa attività può essere realizzata tramite *Google Analytics 4* (GA4), che permette di analizzare visite, tempi di permanenza, provenienza geografica degli utenti e frequenza dei download. L'integrazione di GA4 è possibile attraverso *Google Tag Manager*, inserendo il codice di tracciamento nelle impostazioni del sito, oppure utilizzando le funzionalità di *Analytics* già presenti nel pannello di controllo di *Wix*. Attraverso *Tag Manager* è inoltre possibile impostare eventi personalizzati per monitorare con precisione il numero di download e la loro distribuzione geografica. In alternativa, si può ricorrere a *Matomo Cloud*, una piattaforma con funzionalità simili a GA4 ma con una maggiore attenzione alla privacy, che include anche strumenti come le *heatmap* per analizzare il comportamento di navigazione degli utenti. Poiché *Wix* non consente l'installazione diretta di plugin esterni, il monitoraggio dei download deve essere gestito tramite eventi configurati in GA4 o attraverso le dashboard di *Wix Analytics*. Per aumentare il controllo sull'accesso ai documenti, è possibile prevedere meccanismi di limitazione del download diretto. L'utente può, ad esempio, accedere ai file solo dopo la visualizzazione di un *disclaimer* legale o tramite una procedura di autenticazione leggera, come l'inserimento del proprio indirizzo e-mail in un modulo *Wix*. Inoltre, *Wix* consente di gestire link di accesso controllato o aree riservate per i file, riducendo così il rischio di diffusione non autorizzata dei contenuti digitali.

Infine, per prevenire e contrastare eventuali fenomeni di plagio, si può prevedere l'attivazione di sistemi di monitoraggio automatizzato in grado di rilevare la presenza online di copie non autorizzate dei materiali. Servizi come *Google Alerts*, *Copyscape* o *Turnitin Author Feedback* permettono di individuare la pubblicazione di testi identici o parzialmente corrispondenti in altri siti web. L'attivazione di alert basati su frasi chiave estratte dai documenti consentirebbe di ricevere notifiche tempestive in caso di riproduzione non autorizzata. Integrata con le misure di watermark, licenza e tracciamento, tale strategia contribuirebbe a rafforzare la tracciabilità, la

²⁷⁷ Per approfondire, cfr. i seguenti testi: Creative Commons Italia, *Guida alle licenze Creative Commons. Versione 4.0 – Come scegliere e applicare una licenza CC*, Torino, Centro Nexa su Internet & Società, 2020; S. Casellato e M. Rossi, *Diritto d'autore e Open Access nella ricerca scientifica*, Bologna, Il Mulino, 2019; L. Lessig, *Cultura libera. Un equilibrio fra anarchia e controllo, contro l'estremismo della proprietà intellettuale*, Milano, Apogeo, 2005

trasparenza e la protezione autoriale dei materiali diffusi online, assicurando una gestione scientificamente responsabile dei contenuti digitali.

Le pratiche di tutela e monitoraggio delineate in questa sezione evidenziano come la pubblicazione digitale dei risultati di ricerca non possa prescindere da una riflessione integrata sui temi dell'accesso aperto, della proprietà intellettuale e della responsabilità scientifica. L'obiettivo non è limitare la circolazione del sapere, bensì costruire un ecosistema editoriale trasparente, in cui i contenuti siano al contempo accessibili, riconoscibili e tutelati. In prospettiva, l'adozione di strumenti interoperabili per la tracciabilità dei download, la standardizzazione dei metadati e l'impiego di licenze aperte ma regolamentate (come quelle *Creative Commons*) potrebbe rappresentare un modello di buone pratiche per le edizioni digitali in ambito accademico. Allo stesso modo, lo sviluppo di sistemi di monitoraggio automatizzato basati su intelligenza artificiale e *machine learning* potrebbe in futuro migliorare la capacità di individuare riusi non autorizzati dei contenuti, rendendo la tutela del patrimonio scientifico più efficiente e proattiva. In questa prospettiva, la tutela digitale non andrebbe intesa come una barriera, ma come un principio metodologico volto a garantire la correttezza, la verificabilità e la sostenibilità della ricerca scientifica nell'ecosistema digitale contemporaneo.

CONCLUSIONE

Il percorso di ricerca qui presentato ha mostrato come l'incontro tra filologia d'autore e tecnologie dell'intelligenza artificiale possa produrre risultati significativi, tanto sul piano metodologico quanto su quello operativo. L'applicazione di modelli di AI per il riconoscimento del layout documentario e della scrittura ariostesca ha confermato la possibilità di integrare strumenti automatizzati nei processi di analisi e trascrizione dei materiali autografi, senza rinunciare al controllo e alla sensibilità interpretativa propri del lavoro filologico. Gli esperimenti condotti hanno evidenziato buoni livelli di accuratezza nel riconoscimento, esito di un processo di addestramento complesso, articolato in più fasi di *fine-tuning*, necessario per adattare i modelli alle specificità grafiche, linguistiche e materiali dei frammenti ariosteschi. In prospettiva, tali modelli potranno essere ulteriormente perfezionati grazie all'ampliamento del *training set*, alla disponibilità di nuove immagini ad alta risoluzione e all'impiego di tecniche di *transfer learning* – metodologie del *machine learning* che consentono di riutilizzare conoscenze acquisite da modelli pre-addestrati su compiti affini per migliorare le prestazioni su un nuovo dominio specifico, riducendo al contempo la quantità di dati necessari e i tempi di addestramento.

Parallelamente, anche il prototipo di modello di codifica XML-TEI sviluppato per l'edizione genetica digitale rappresenta un risultato significativo di questa ricerca. Si è così delineata la possibilità di strutturare in modo modulare e interoperabile i dati testuali e documentari, consentendo di rappresentare in ambiente digitale le peculiarità linguistiche e metriche dei testi in esame, le diverse fasi del processo creativo ariostesco con le loro complesse stratigrafie correttive, e di restituire la fisionomia materiale degli originali. Tuttavia, questo costituisce soltanto l'inizio di un percorso più lungo e articolato, nel quale il modello dovrà essere ulteriormente affinato e testato nella fase di implementazione effettiva, al fine di garantire la piena compatibilità con gli standard internazionali di edizione digitale, la coerenza semantica dei metadati e la massima accessibilità dei contenuti. L'interazione diretta con i dati reali e con i

sistemi di pubblicazione online costituirà, infatti, un banco di prova indispensabile per valutarne la flessibilità e l'efficacia rappresentativa.

In questa prospettiva, il lavoro presentato non si configura come un punto d'arrivo, ma come un prototipo integrato di edizione genetica digitale, destinato a evolversi in dialogo con le più recenti infrastrutture di ricerca dedicate a Ludovico Ariosto. Tra queste, un ruolo centrale è svolto dal progetto *Ariosto Online*, nato presso l'Università "La Sapienza" di Roma come portale digitale monografico, la cui finalità principale è raccogliere, organizzare e rendere accessibile, in un ambiente scientificamente affidabile e digitalmente interoperabile, l'intero patrimonio letterario, manoscritto e documentario legato al poeta.²⁷⁸ Il portale si configura come una vera e propria infrastruttura di ricerca, concepita per valorizzare le opere di Ariosto – comprese le varianti, i testimoni manoscritti, le lettere e le poesie – all'interno di un contesto filologico e critico rinnovato. La piattaforma è stata pensata per un pubblico ampio e diversificato (studiosi, docenti, studenti e appassionati di letteratura rinascimentale) e offre strumenti di consultazione avanzati, repertori corredati da metadati accurati e servizi testuali che consentono un accesso integrato alle fonti e ai risultati della ricerca. Un aspetto centrale della missione di *Ariosto Online* è la produzione e l'ospitalità di edizioni digitali scientifiche, codificate secondo gli standard TEI-XML, corredate da apparati critici e strumenti di navigazione tra testimoni, varianti e redazioni successive. In prospettiva, il portale accoglierà anche i testi della nuova *Edizione Nazionale delle Opere di Ariosto*, promossa dall'Accademia dell'Arcadia e attualmente in preparazione. Tale integrazione consentirà di aggiornare e ampliare progressivamente i materiali digitali, assicurando continuità tra tradizione editoriale e innovazione tecnologica.

²⁷⁸ Il portale *Ariosto Online* (consultabile al seguente link: <<https://www.ludovicoariosto.org/>>) è stato realizzato nell'ambito dello *Spoke 3 'Digital Philology'* del progetto *PNRR PE/5 CHANGES* e ha integrato le risorse e le competenze dell'unità di ricerca della Sapienza in filologia digitale. Le attività di sviluppo del portale, avviate nel 2023, si sono articolate in diverse fasi. I lavori sono iniziati con la schedatura delle lettere ariostesche, a cura di Chiara De Cesare, e con il censimento delle rime, condotto da Nicole Volta in collaborazione con i progetti digitali *Epistulae* (consultabile al seguente link: <<https://epistulae.unil.ch/>>) e *Lyra* (diretto da Simone Albonico presso l'Università di Losanna e consultabile al seguente link: <<https://lyra.unil.ch/>>). Parallelamente, si è proceduto al riversamento dei risultati del censimento dei manoscritti autografi ariosteschi, frutto delle ricerche dello stesso Albonico, e si sono avviate le prime marcature digitali dei testi, insieme alla redazione delle schede di presentazione delle opere, pubblicate progressivamente online. Nel corso del 2025 è previsto il rilascio di una nuova fase del progetto, che include l'integrazione dei testi dell'*Edizione Nazionale delle Opere di Ariosto* e il consolidamento della struttura filologica complessiva. Inoltre, *Ariosto Online* si inserisce nel più ampio quadro del progetto *Filologia digitale: Autografi e portali d'autore*, promosso dalla Fondazione CHANGES e dalla Sapienza. L'iniziativa mira a creare una rete di interoperabilità tra diversi portali monografici, come *Tasso Online* (consultabile al seguente link: <<https://www.torquatotasso.org/>>) e il *Portale degli Autografi dei Letterati Italiani* (consultabile al seguente link: <<http://www.autografi.net/it/>>), favorendo la condivisione dei dati e la valorizzazione delle risorse digitali dedicate alla tradizione letteraria italiana.

Dal punto di vista strutturale, il portale è articolato in sezioni tematiche che riflettono la molteplicità degli interessi e delle opere ariostesche. La sezione *Biografia* ricostruisce il profilo umano e intellettuale dell'autore nel contesto storico-culturale del Rinascimento; la sezione *Opere* ospita schede critiche dedicate ai principali testi ariosteschi, corredate da informazioni sui testimoni, sulle edizioni e sulle varianti; la sezione *Manoscritti* offre un catalogo dei principali testimoni autografi e non autografi, con descrizioni codificate e, quando possibile, riproduzioni digitali ad alta definizione realizzate in collaborazione con biblioteche e archivi conservatori. Completano il portale le sezioni *Lettere e Poesie*, dedicate rispettivamente alla schedatura e alla metadattazione dell'epistolario e delle rime, e la sezione *Percorsi*, pensata per la divulgazione dei risultati della ricerca attraverso approfondimenti tematici e collegamenti con altri progetti digitali affini. Ogni scheda testuale è concepita per favorire una navigazione ipertestuale tra testo, varianti e apparati, rendendo immediato il passaggio tra le diverse testimonianze e la loro contestualizzazione critica. L'obiettivo è offrire al lettore un'esperienza di consultazione dinamica, in cui il testo letterario sia percepito nella sua dimensione processuale e stratificata. Benché alcune sezioni del portale siano ancora in costruzione, l'obiettivo è rendere gradualmente accessibili le risorse secondo un modello di sviluppo aperto e incrementale.

In tale contesto, il portale rappresenta un ambiente di riferimento ideale per la presente ricerca, poiché offre l'infrastruttura scientifica e tecnologica entro cui potrà collocarsi il prototipo di edizione genetica digitale dei frammenti autografi. L'integrazione del prototipo nel portale consentirà non solo di valorizzarne i risultati, ma anche di metterli in relazione con le altre risorse ariostesche digitalizzate, contribuendo alla costruzione di un ecosistema testuale coerente e condiviso. Tale prospettiva richiede tuttavia un coinvolgimento stabile e coordinato delle istituzioni detentrici dei materiali, in particolare delle biblioteche che conservano i testimoni autografi, la cui collaborazione è indispensabile per garantire l'accesso, la qualità delle riproduzioni e la sostenibilità a lungo termine del progetto.

L'esperienza maturata nel corso di questa ricerca conferma che la filologia digitale non è soltanto un campo di applicazione tecnologica, ma un terreno di dialogo disciplinare, capace di rinnovare il modo in cui si studia, si rappresenta e si comunica il testo letterario. L'edizione dei frammenti autografi dell'*Orlando furioso* si configura così come un vero e proprio laboratorio di sperimentazione, in cui la memoria materiale dell'opera incontra le potenzialità del digitale, e in cui il gesto autoriale di Ariosto, restituito nella sua concretezza grafica e nel suo dinamismo creativo, continua a interrogare il nostro presente tecnologico.

BIBLIOGRAFIA PER ARGOMENTI

Un caso di filologia d'autore: i frammenti autografi dell'*Orlando furioso* di Ludovico Ariosto

Edizioni dell'*Orlando furioso* di riferimento

Ariosto L., *Orlando furioso. Secondo la princeps del 1516*, a cura di M. Dorigatti, con la collaborazione di G. Stimato, Firenze, Olschki, 2016.

Id., *Orlando furioso e Cinque Canti*, a cura di R. Ceserani e S. Zatti, Torino, Utet, 1997, 2 voll.

Id., *Lettere*, a cura di A. Stella, in L. Ariosto, *Tutte le opere, vol. III: Satire Erbolato Lettere*, a cura di C. Segre, Milano, Mondadori, 1984.

Id., *I frammenti autografi dell'Orlando furioso*, a cura di S. Debenedetti, Torino, Chiantorre, 1937; rist. *I frammenti autografi dell'Orlando furioso*, a cura di S. Debenedetti, con una premessa di C. Segre, Roma, Edizioni di Storia e Letteratura, 2010.

Id., *Frammenti autografi dell'Orlando furioso*, a cura di G. Agnelli, Roma, Danesi, 1904.

Id., *Orlando furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, a cura di S. Debenedetti e C. Segre, 2 voll., Bologna, Commissione per i testi di lingua, 1960.

Id., *Orlando furioso di m. Ludovico Ariosto*, Venezia, Valgrisi, 1556, 2 voll.

Id., *Orlando furioso di m. Ludovico Ariosto*, Venezia, Gabriel Giolito de' Ferrari, 1548.

Id., *Cinque canti di un nuovo libro di m. Ludovico Ariosto*, Venezia, Aldo Manuzio, 1545.

Id., *Orlando furioso di m. Ludovico Ariosto*, F. Rosso da Valenza, 1532.

Id., *Orlando furioso di Ludovico Ariosto ristampato et con molta diligentia da lui corretto*, Ferrara, G. B. da la Pigna, 1521.

Id., *Orlando furioso di Ludovico Ariosto da Ferrara*, Ferrara, G. Mazzocco dal Bondeno, 1516.

Storia redazionale dell'*Orlando furioso*

Ascoli A., *Fede e riscrittura. Il Furioso del 1932*, «Rinascimento», 43, 2003, pp. 93-130, consultabile al seguente link: https://www.academia.edu/1821739/_Fede_e_riscrittura_il_Furioso_del32_ (ultimo accesso 17 ottobre 2025).

Casadei A., *La tradizione delle opere di Ariosto: Orlando Furioso, Frammenti, Cinque Canti*, in *Storia della letteratura italiana, vol. X: La tradizione dei testi*, a cura di E. Malato, C. Ciociola, Roma, Salerno Editrice, 2001, pp. 827-831.

Id., *Il percorso del Furioso: ricerche intorno alle redazioni del 1516 e del 1521*, Il Mulino, Bologna, 1993.

Dalla Palma G., *Dal secondo al terzo Furioso: mutamenti di struttura e movimenti ideologici*, in *Ludovico Ariosto. Lingua, stile e tradizione*, Atti del congresso organizzato dai comuni di Reggio Emilia e Ferrara (12-16 ottobre 1974), a cura di C. Segre, Milano, Feltrinelli, 1976, pp. 95-105.

Debenedetti S., Segre C., *Nota al testo*, in L. Ariosto, *Orlando furioso secondo l'edizione del 1532 con le varianti delle edizioni del 1516 e del 1521*, Bologna, Commissione per i testi di lingua, 1960, vol. II, pp. 1649-1697.

Dorigatti M., *Introduzione*, in L. Ariosto, *Orlando Furioso. Secondo la princeps del 1516*, Firenze, Olschki, 2016, pp. 21- 37.

Id., *Nota al testo*, in L. Ariosto, *Orlando Furioso. Secondo la princeps del 1516*, Firenze, Olschki, 2016, pp. 41- 180.

Fahy C., *L'Orlando furioso del 1532. Profilo di un'edizione*, Milano, Vita e Pensiero, 1989.

Harris N., *Filologia e bibliologia a confronto nell'Orlando furioso del 1532*, in AA.VV, *Libri, tipografi, biblioteche. Ricerche storiche dedicate a Luigi Balsamo*, Firenze, Olschki, 1997, pp. 105-122.

Montecchi G., Sorella A., *I nuovi modi della tradizione fra Quattro e Cinquecento*, in *Storia della letteratura italiana*, vol. X: *La tradizione dei testi*, a cura di E. Malato, C. Ciociola, Roma, Salerno Editrice, 2001, pp. 633-673.

Moretti W., *L'ultimo Ariosto*, Bologna, Patron Editore, 1977.

Pigna G. B., *La vita di messer Ludovico Ariosto*, in *I romanzi*, Venezia, Valgrisi, 1554; rist. *I romanzi*, a cura di E. Ritrovato, Bologna, 1997, consultabile al seguente link: https://www.google.it/books/edition/I_romanzi_di_M_Giouan_Battista_Pigna_al/cJviW2EmwS8C?hl=it&gbpv=1&dq=giovan+battista+pigna+i+romanzi&printsec=frontcover (ultimo accesso 17 ottobre 2025).

Rajna P., *Le fonti dell'Orlando furioso*, Firenze, Sansoni Editore, 1876.

Sangirardi G., *Ludovico Ariosto*, Firenze, Le Monnier, 2006.

Segre C., Zampolli A., *Le concordanze diacroniche dell'Orlando furioso*, in *Ludovico Ariosto. Lingua, stile e tradizione*, Atti del congresso organizzato dai comuni di Reggio Emilia e Ferrara (12-16 ottobre 1974), a cura di C. Segre, Milano, Feltrinelli, 1976, pp. 231-264.

Segre C., *Storia interna dell'Orlando furioso*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 29-41.

Zatti S., *Il lavoro di tutta una vita. L'elaborazione del Furioso e i Cinque Canti*, in L. Ariosto, *Orlando furioso e Cinque Canti*, Torino, Utet, 1997, vol. I, pp. 27-38.

Autografi dell'Orlando furioso

Albonico S., *Pour un retour aux manuscrits de l'Arioste*, «Génésis», n. 49, 2019, pp. 31-46; poi Id., *Dalla parte di Ariosto*, in *Gli 'scartafacci' degli scrittori. I sentieri della creazione letteraria in Italia*, a cura di C. Del Vento e P. Musitelli, Roma, Carocci, 2022, pp. 143-62.

Id., *Ariosto*, in *Autografi dei letterati italiani. Il Cinquecento*, a cura di M. Motolese, P. Procaccioli e E. Russo, Roma, Salerno editrice, 2022, vol. III, pp. 3-35.

Campeggiani I., *L'ultimo Ariosto. Dalle Satire ai Frammenti autografi*, Pisa, Scuola Normale Superiore, 2017.

Id., *I frammenti autografi dell'Orlando Furioso: un'ipotesi per lo 'scrittoio' di Ariosto*, in *Dreaming again on things already dreamed: 500 years of Orlando furioso (1516-2016)*, Atti del convegno tenuto a Oxford nel 2016 in occasione del 5° centenario dell'*Orlando furioso*, a cura di M. Dorigatti e M. Pavlova, Oxford, Peter Lang, 2019, pp. 133-143.

Cerlini A., *Scrittura e punteggiatura negli autografi dell'Ariosto*, «Cultura neolatina», 4(5), 1944-45, pp. 37-60.

Debenedetti S., *Introduzione*, in L. Ariosto, *I frammenti autografi dell'Orlando furioso*, Roma, Edizioni di Storia e Letteratura, 2010.

Dorigatti M., *Il manoscritto dell'Orlando furioso (1505-1515)*, in *L'uno e l'altro Ariosto. In corte e nelle delizie*, a cura di G. Venturi, Firenze, Olschki, 2011, pp. 1-44.

Gritti V., *I manoscritti ambrosiani della Rocca di Tristano e della Storia d'Italia e la rappresentazione degli originali perduti*, «Filologia italiana», 17, 2020, pp. 197-265.

Motta U., *Ariosto e dintorni*, in *Tra i fondi dell'Ambrosiana. Manoscritti italiani antichi e moderni*, Atti del convegno (Milano, 15-18 maggio 2007), a cura di M. Ballarini, G. Barbarisi, C. Berra e G. Frasso, Milano 2008, vol. I, pp. 205-27.

Cinque Canti

Casadei A., *I 'Cinque Canti' o l'ultima eredità di Boiardo*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 113-128.

Id., *Sulla situazione testuale e sulla cronologia dei 'Cinque Canti'*, in *Il percorso del 'Furioso'. Ricerche intorno alle redazioni del 1516 e del 1521*, Bologna, Il Mulino, 1993, pp. 179-192.

Ceserani R., Zatti S., *Cinque Canti*, in L. Ariosto, *Orlando furioso e Cinque Canti*, Torino, Utet, 1997, vol. II, pp. 1634-1651.

Gritti V., *Per l'edizione critica dei 'Cinque Canti' di Ariosto*, «Filologia italiana», 13, 2016, pp. 139-191, consultabile al seguente link: <https://sfera.unife.it/retrieve/e309ade3-2399-3969-e053-3a05fe0a2c94/113922360214.pdf?utm_source=chatgpt.com> (ultimo accesso 17 ottobre 2025).

Larosa S., *Poesia e cronologia nei Cinque Canti. Una nuova ipotesi*, Rende, Centro Editoriale e Librario, Università della Calabria, 2001.

Segre C., *Studi sui Cinque Canti*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 121-177.

Filologia ariostesca e critica delle varianti d'autore nell'*Orlando furioso*

Casadei A., *La strategia delle varianti. Le correzioni storiche del terzo Furioso*, Lucca, Pacini Fazzi, 1988.

Contini G., *Come lavorava l'Ariosto*, «Meridiano di Roma», 18 luglio 1937; poi in *Id.*, *Esercizi di lettura sopra autori contemporanei, con un'appendice su testi non contemporanei*, Torino, Einaudi, 1974, pp. 232-41.

Dorigatti M., *Momenti della filologia ariostesca nel Cinquecento*, in *La filologia in Italia nel Rinascimento*, a cura di C. Caruso e E. Russo, Roma, Edizioni di Storia e Letteratura, 2018, pp. 193-215.

Isella D., *Le varianti d'autore (critica e filologia)*, in *Id.*, *Le carte mescolate vecchie e nuove*, a cura di S. Isella Brusamolino, Torino, Einaudi, 2009, pp. 7-28.

Italia P., Raboni G., *Che cos'è la filologia d'autore*, Roma, Carocci, 2010.

Italia P., Raboni G. et al., *What is Authorial Philology*, Cambridge, Open Book Publishers, 2021, consultabile al seguente link: < <https://doi.org/10.11647/OBP.0224> > (ultimo accesso 17 ottobre 2025).

Morini L., *Ruscelli e le pretese varianti ariostesche al Furioso del '32*, in *In ricordo di Cesare Angelini. Studi di letteratura e filologia*, a cura di F. Alessio e A. Stella, Milano, Il Saggiatore, 1979, pp. 160-84.

Ruscelli G., *Mutationi, et miglioramenti che M. Lodovico Ariosto havea fatti per mettere nell'ultima impressione del Furioso*, in L. Ariosto, *Orlando furioso di m. Ludovico Ariosto, tutto ricorretto, et di nuove figure adornato. Al quale di nuovo sono aggiunte le annotationi, gli avvertimenti, & le dichiarazioni di Girolamo Ruscelli...*, Venezia, Valgrisi, 1556, 2 voll., cc. 3d1r-3e1v, consultabile al seguente link: https://www.google.it/books/edition/Orlando_furioso_Di_m_Lodouico_Ariosto_tu/d796a9orNHgC?hl=it&gbpv=1&dq=Orlando+furioso+di+m.+Ludovico+Ariosto,+tutto+ricorretto,+et+di+nuove+figure+adornato.+Al+quali+di+nuovo+sono+aggiunte+le+Annotationi,+gli+auuertimenti,+&+le+dichiarazioni+di+Girolamo+Ruscelli&printsec=frontcover (ultimo accesso 17 ottobre 2025).

Segre C., *Petrarca, Ariosto e la critica delle varianti nel Cinquecento*, in *Storia della letteratura italiana, vol. XI: La critica letteraria dal Due al Novecento*, a cura di E. Malato e P. Orvieto, Roma, Salerno Editrice, 2003, pp. 353-67.

Id., *Critique des variantes et critique génétique*, «Génésis», 7, 1995, pp. 29-45; poi in versione italiana, con il titolo *La genesi del testo: critica delle varianti e critica genetica*, in *La costruzione del testo in italiano. Sistemi costruttivi e testi costruiti*, Atti del convegno (Barcellona, 24-29 aprile 1995), a cura di M. de las Nieves Muñiz Muñiz e F. Amelia, Firenze 1996, pp. 11-21.

Spagnolo L., *Variantistica ed ecdotica dell'Orlando furioso*, «Tipo-filologia», 1, 2008, pp. 61-87, consultabile al seguente link: <https://www.torrossa.com/it/resources/an/2247839#>> (ultimo accesso 16 dicembre 2024).

Trovato P., *Con ogni diligenza corretto. La stampa e le revisioni editoriali dei testi letterari italiani (1470-1570)*, Ferrara, UnifePress, 2009, rist. anastatica dell'ed., Bologna, Il Mulino, 1991.

Poetica e critica dell'*Orlando furioso*

Binni W., *Il problema critico dell'Orlando Furioso*, in *Metodo e poesia di Ludovico Ariosto ed altri studi ariosteschi*, a cura di Rosanna Alhaique Pettinelli, Firenze, La Nuova Italia, 1996, parte I, cap.V.

Id., *Introduzione alla poetica ariostesca*, in *Metodo e poesia di Ludovico Ariosto ed altri studi ariosteschi*, a cura di Rosanna Alhaique Pettinelli, Firenze, La Nuova Italia, 1996, parte I, cap.VI.

Id., *Metodo e poesia nell'Orlando furioso*, in *Metodo e poesia di Ludovico Ariosto ed altri studi ariosteschi*, a cura di Rosanna Alhaique Pettinelli, Firenze, La Nuova Italia, 1996, parte I, cap.VII.

Id., *Il capolavoro*, in *Metodo e poesia di Ludovico Ariosto ed altri studi ariosteschi*, a cura di Rosanna Alhaique Pettinelli, Firenze, La Nuova Italia, 1996, parte II, sezione II.

Id., *Storia della critica ariostesca*, in *Metodo e poesia di Ludovico Ariosto ed altri studi ariosteschi*, a cura di Rosanna Alhaique Pettinelli, Firenze, La Nuova Italia, 1996, parte IV, cap.I.

Caretti L., *Ariosto e Tasso*, Torino, Einaudi, 1961, pp. 17-44.

Casadei A., *Nuove prospettive su Ariosto e sul Furioso*, «Italianistica», 3, 2008, pp. 167-192.

Ceserani R., Zatti S., *Introduzione*, in L. Ariosto, *Orlando furioso e Cinque Canti*, Torino, Utet, 1997, vol. I, pp. 61-79.

Jossa S., *Ariosto*, Bologna, Il Mulino, 2009.

Segre C., *La poesia dell'Ariosto*, in *Esperienze ariostesche*, Pisa, Nistri-Lischi, 1966, pp. 9-27.

Zatti S., *Leggere l'Orlando furioso*, Bologna, Il Mulino, 2016.

Evoluzione linguistica di Ludovico Ariosto

Gritti V., *Intorno a una serie di correzioni nell'Orlando Furioso. I metaplasmi e le forme eteroclite*, «Storie e Linguaggi», 8, 2022, pp. 115-196.

Stella A., *Note sull'evoluzione linguistica dell'Ariosto*, in *Ludovico Ariosto. Lingua, stile e tradizione*, Atti del congresso organizzato dai comuni di Reggio Emilia e Ferrara (12-16 ottobre 1974), a cura di C. Segre, Milano, Feltrinelli, 1976, pp. 49-64.

Informatica umanistica e filologia digitale

Digital humanities: metodi, strumenti, prospettive

AA.VV., *Informatica umanistica, digital humanities: verso quale umanità?*, a cura di M. Gatto, A. Squeo e S. Silvestri, Bari, Cacucci, 2024.

AA.VV., *Digital humanities. Metodi, strumenti, saperi*, a cura di F. Ciotti, Roma, Carocci, 2023.

Alshater M., *Exploring the Role of Artificial Intelligence in Enhancing Academic Performance: A Case Study of ChatGPT*, «SSRN Electronic Journal», 2022, consultabile al seguente link: <<https://doi.org/10.2139/ssrn.4312358>> (ultimo accesso 17 ottobre 2025).

Balbo A. , *L'antichista nell'era digitale: un quadro di riflessione teorico-pratica nel terzo millennio*, in AA.VV., *DH.22. Digital Humanities 2022. Per un confronto interdisciplinare tra saperi umanistici a 30 anni dalla nascita del World Wide Web*, Atti del convegno internazionale (Università degli Studi dell'Aquila, 26 - 27 ottobre 2022), a cura di M. Di Maro, V. Merola e T. Nocita, Roma - Bristol, «L'ERMA» di BRETSCHNEIDER, 2023, pp. 13-24.

Berlangieri M. G., *Archivi e Data Visualization. Esplorazione ed esposizione delle fonti mediante la multidimensionalità dei dati digitali e le intelligenze artificiali*, «DigItalia», 18(2), 2023, pp. 173-183, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00085>> (ultimo accesso 13 marzo 2025).

Buzzetti D., *Alle origini dell'informatica umanistica. Humanities computing e/o digital humanities?*, «Rendiconti Lincei Scienze Morali Storiche e Filologiche», 30(9), 2019, pp. 71-103.

Campagnolo A., *Codificare il codice. La materialità del libro e la sua digitalizzazione verso una codicologia digitale*, «Umanistica digitale», 17, 2024, consultabile al seguente link: <<https://umanisticadigitale.unibo.it/article/view/16694>> (ultimo accesso 17 ottobre 2025).

Id., *Book Conservation and Digitization: The Challenges of Dialogue and Collaboration*, ARC Humanities Press, 2020.

Cerullo L., Negri A., *L'infrastruttura software per il patrimonio culturale (ISPC) come abilitatore di un Ecosistema digitale nazionale del patrimonio culturale*, «DigItalia», 18(1), 2023, pp. 38-50, consultabile al seguente link: <<https://digitalia.cultura.gov.it/article/view/3008>> (ultimo accesso 17 ottobre 2025).

Ciotti F., *Chat GBT: un pappagallo stocastico può essere di aiuto a un vero ricercatore (umanistico)?*, in *La memoria digitale. Forme del testo e organizzazione della conoscenza*, Atti del XII Convegno Annuale AIUCD (Università di Siena, 5-7 giugno 2023) a cura di E. Carbé et al., AIUCD, 2023, pp. 245-250.

Id., *Minerva e Il Pappagallo. IA generativa e modelli linguistici nel laboratorio dell'umanista digitale*, «Testo E Senso», 26, 2023, pp. 289-315, consultabile al seguente link: <<https://doi.org/10.58015/2036-2293/671>> (ultimo accesso 17 ottobre 2025).

Gil T. L., *Sui big data e il neopositivismo digitale nella ricerca storica*, «Umanistica Digitale», 18, 2024, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/20365>> (ultimo accesso 17 ottobre 2025)).

Grillo M., *Preservazione del patrimonio culturale e cloud computing: caratteristiche e criticità*, in *ME.TE. Digitali - Mediterraneo in rete tra testi e contesti*, Atti del XIII convegno annuale AIUCD (Università degli Studi di Catania, 28-30 maggio 2024), a cura di A. Di Silvestro et al., AIUCD, 2024, pp. 117-121.

Kaplan F., Di Lenardo I., *Big Data of the Past*, «Frontiers in Digital Humanities», 4, 2017, consultabile al seguente link: <<https://doi.org/10.3389/fdigh.2017.00012>> (ultimo accesso 17 ottobre 2025).

Mulligan J., *Digital Humanities Application Development in the Cloud*, in *HARC '19: Proceedings of the Humans in the Loop: Enabling and Facilitating Research on Cloud Computing* (Chicago, 19 luglio 2019), New York, Association for Computing Machinery, 2019, consultabile al seguente link: <<https://dl.acm.org/doi/10.1145/3355738.3355753>> (ultimo accesso 17 ottobre 2025).

Roncaglia G., *L'architetto e l'oracolo. Forme digitali del sapere da Wikipedia a ChatGPT*, Bari, Laterza, 2023.

Tomasi F., *Organizzare la conoscenza: digital humanities e web semantico*, Milano, Editrice Bibliografica, 2022.

Terras M., *The Rise of Digitization: An Overview*, in *Digitization Perspectives*, a cura di Ruth Rikowski, Leiden, Paesi Bassi, 2011, pp. 1-20, consultabile al seguente link: <https://www.researchgate.net/publication/251382472_The_Rise_of_Digitization> (ultimo accesso 17 ottobre 2025).

Filologia digitale ed edizioni scientifiche digitali

AA.VV., *Manifesto per le edizioni scientifiche digitali*, «Umanistica digitale», 12, 2022, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/14814>> (ultimo accesso 17 ottobre 2025).

AA.VV., *Teoria e forme del testo digitale*, a cura di M. Zaccarello, Roma, Carocci, 2019.

AA.VV., *Digital Scholarly Editing: Theory and Practice*, a cura di M. J. Driscoll e E. Pierazzo, Cambridge, Open Book Publisher, 2016, consultabile al seguente link: <<https://www.openbookpublishers.com/product/483>> (ultimo accesso 17 ottobre 2025).

AA.VV., *Edizioni critiche digitali. Edizioni a confronto*, a cura di P. Italia e C. Bonsi, Roma, Sapienza Università Editrice, 2016, consultabile al seguente link: <<https://doi.org/10.13133/9788893770033>> (ultimo accesso 17 ottobre 2025).

Bozzi A., Corradini M. S. e Ferrari G., *Percorsi Di Linguistica e Di Filologia Computazionali*, Edizioni ETS, 2019.

Buzzetti D., *Digital Editions and Text Processing*, in *Text Editing, Print and the Digital World*, a cura di Marilyn Deegan e Kathryn Sutherland, Digital Research in the Arts and Humanities, Surrey, Ashgate Publishing, 2009, pp. 45-61.

Id., *Digital Representation and the Text Model*, «New Literary History», 33(1), Project MUSE, 2002, pp. 61-88.

Buzzoni M. et al., *Progetto di edizione genetica digitale del Canzoniere manoscritto di U. Saba (1919-20)*, in *ME.TE. Digitali - Mediterraneo in rete tra testi e contesti*, Atti del XIII convegno annuale AIUCD (Università degli Studi di Catania, 28-30 maggio 2024), a cura di A. Di Silvestro et al., AIUCD, 2024, pp. 215-220.

Id., *Il progetto Saba: dare voce a un manoscritto inedito del 'Canzoniere'*, in *Informatica umanistica, digital humanities: verso quale modernità?*, a cura di M. Gatto et al., Bari, Cacucci Editore, 2024, pp. 153-165.

Cappa A., *Il Centro nazionale per lo studio del manoscritto verso una nuova dimensione: dai microfilm al digitale grazie ai fondi del PNRR*, «DigItalia», 18(1), 2023, pp. 9-15, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00056>> (ultimo accesso 17 ottobre 2025).

Fiormonte D., *Per una critica del testo digitale: letteratura, filologia e rete*, Roma, Bulzoni, 2018.

Id., *Scrittura e filologia nell'era digitale*, Torino, Bollati Boringhieri, 2003.

Gabler H. W., *Text Genetics in Literary Modernism and Other Essays*, OpenBook Publishers, 2018.

Id., *Genetic Texts - Genetic Editions - Genetic Criticism or: Towards Discoursing the Genetics of Writing*, in *Problems of Editing*, a cura di Christa Jansohn, De Gruyter Brill, 2012, consultabile al seguente link: <<https://doi.org/10.1515/9783110939958.59>> (ultimo accesso 17 ottobre 2025).

Italia P., *Editing Duemila. Per una filologia dei testi digitali*, Roma, Salerno Editrice, 2020.

Id., *Filologia d'autore digitale*, «Ecdotica», 1, 2019, pp. 202 - 216.

Id., *Fake texts e Wiki edizioni. Per una filologia digitale sostenibile*, «Storie e Linguaggi», 5(1), 2019, pp. 277-298.

Id., *Filologia d'autore digitale e multidisciplinare. Dall'Authorship alla Fotonica*, in *Saperi umanistici nella Contemporaneità*, a cura di G. Sampino e F. Scaglione, Palermo, Palumbo, 2018, pp. 322-334.

Id., *Editing 2.0. Quali testi leggiamo e leggeremo in rete?*, «Nuovi Argomenti», 73, 2016, pp. 80-86.

Italia P., Tomasi F., *Filologia digitale. Fra teoria, metodologia e tecnica*, «Ecdotica», 11, 2014, pp. 112 - 130.

Lippolis A. S., Totaro G., *Variants Mining. Computational investigations on authorial variants: a comparison between Leopardi and Manzoni*, «Umanistica digitale», 10, 2021, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/12627>> (ultimo accesso 17 ottobre 2025).

Leonardi L., *Filologia digitale del Medioevo italiano*, «Griseldaonline», 20 (2), 2021, consultabile al seguente link: <<https://doi.org/10.6092/issn.1721-4777/12817>> (ultimo accesso 17 ottobre 2025).

Longo L., *Dai testi cartacei ai testi virtuali: potenzialità dell'ecdotica digitale*, «Umanistica digitale», 3, 2018, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/8146>> (ultimo accesso 17 ottobre 2025).

Malaspina E., *Il futuro dell'edizione critica (cioè lachmanniana), più o meno digitale. Riflessioni (in)attuali*, «Storie e Linguaggi», 5(1), 2019, pp. 35-60.

Mancinelli T., Pierazzo E., *Che cos'è un'edizione scientifica digitale*, Roma, Carocci Editore, 2020.

Martignano C., *Critical Edition Ontology: a conceptual model for digital critical editions*, «Umanistica digitale», 17, 2024, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/19469>> (ultimo accesso 17 ottobre 2025).

McGann J., Buzzetti D., *Critical Editing in a Digital Horizon*, in *Electronic Textual Editing*, a cura di L. Burnard, K. O'Brien O'Keeffe e J. Unsworth, The Modern Language Association of America, 2006, pp. 51-71, consultabile al seguente link: <https://www.researchgate.net/publication/268208603_Critical_Editing_in_a_Digital_Horizon> (ultimo accesso 8 aprile 2025).

Michelone F., *L'edizione critica tra digitale e stampa: riflessioni metodologiche*, «Umanistica digitale», 10, 2021, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/12626>> (ultimo accesso 17 ottobre 2025).

Milanese G., *Filologia, letteratura, computer. Idee e strumenti per l'informatica umanistica*, Milano 2020.

Modolo M., *Il riuso delle immagini dei beni culturali pubblici (1962-2022): un percorso a ostacoli*, «DigItalia», 18(2), 2023, pp. 124-133, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00080>> (ultimo accesso 17 ottobre 2025).

Pierazzo E., Stokes P. A., *Old Books, New Books and Digital Publishing*, in *The Bloomsbury Handbook to the Digital Humanities*, a cura di James O'Sullivan, Londra, Bloomsbury Academic, 2022, pp. 233-44.

Pierazzo E., *Quale infrastruttura per le edizioni digitali? Dalla tecnologia all'etica*, «Textual Cultures», 12(2), 2019, pp. 5-17, consultabile al seguente link: <<https://www.jstor.org/stable/26821533>> (ultimo accesso 17 ottobre 2025).

Id., *Edizione documentaria digitale: rinuncia intellettuale o opportunità scientifica?*, «Ecdotica», 1, 2019, consultabile al seguente link: <<https://hal.science/hal-03461314>> (ultimo accesso 17 ottobre 2025).

Pierazzo E. et al, *Les publications numériques de corpus d'auteurs. Guide de travail, grille d'analyse et recommandations (VI-Novembre 2018). Rapport de recherche*, Huma-Num, 2018, consultabile al seguente link: <<https://shs.hal.science/halshs-01932519>> (ultimo accesso 17 ottobre 2025).

Pierazzo E., *Il testo è morto: lunga vita ai testi. Pluralismo testuale e edizioni digitali*, «Prassi Ecdotiche della Modernità Letteraria», 3, 2018, pp. 322-344, consultabile al seguente link: <<https://doi.org/10.13130/2499-6637/10633>> (ultimo accesso 17 ottobre 2025).

Id., *Digital Scholarly Editing: Theories, Models and Methods*, Farnham, Ashgate Publishing, 2015, consultabile al seguente link: <<https://hal.univ-grenoble-alpes.fr/hal-01182162/document>> (ultimo accesso 17 ottobre 2025).

Id., *Digital Documentary Editions and the Others*, «Scholarly Editing: the Annual of the Association for Documentary Editing», 35, 2014, consultabile al seguente link: <<https://scholarlyediting.org/2014/essays/essay.pierazzo.html>> (ultimo accesso 17 ottobre 2025).

Pierazzo E., André J., *Le codage en TEI des brouillons de Proust: vers l'édition numérique*, «Genesis», 36, pp. 155-161, consultabile al seguente link: <<https://hal.univ-grenoble-alpes.fr/hal-01182150>> (ultimo accesso 17 ottobre 2025).

Pierazzo E., *A Rationale of Digital Documentary Editions*, «Literary and Linguistic Computing», 26 (4), 2011, pp. 463-477.

Id., *Digital Genetic Edition: the Encoding of Time in Manuscript Transcription*, in *Text Editing, Print and the Digital World*, a cura di Marilyn Deegan e Kathryn Sutherland, Digital Research in the Arts and Humanities, Surrey, Ashgate Publishing, 2009, pp. 169-186.

Price K. M., *Electronic Scholarly Editions*, in *A Companion to Digital Literary Studies*, John Wiley & Sons, 2013, pp. 434-450.

Price K. M., *Edition, Project, Database, Archive, Thematic Research Collection: What's in a Name?*, «DHQ: Digital Humanities Quarterly», 3(3), 2009, consultabile al seguente link: <<https://digitalcommons.unl.edu/englishfacpubs/69>> (ultimo accesso 17 ottobre 2025).

Robinson P., *Towards a Theory of Digital Editions*, «The Journal of the European Society for Textual Scholarship», 2013, pp. 105– 131.

Id., *The History of Scholarly Digital Editions, Plc*, «Papers of The Bibliographical Society of Canada», 51(1), 2013, consultabile al seguente link: <<https://jps.library.utoronto.ca/index.php/bsc/article/view/20763/16998>> (ultimo accesso 17 ottobre 2025).

Id., *Some principles for the making of collaborative scholarly editions in digital form*, «Digital Humanities Quarterly», 11(2), 2017, consultabile al seguente link: <<https://dhq-static.digitalhumanities.org/pdf/000293.pdf>> (ultimo accesso 17 ottobre 2025).

Id., *Electronic Editions for Everyone*, in *Text and Genre in Reconstruction. Effects of Digitalization on Ideas, Behaviours, Products and Institutions*, a cura di W. McCarty, Open Book Publisher, 2010, pp. 145-163, consultabile al seguente link: <<https://www.openbookpublishers.com/books/10.11647/obp.0008>> (ultimo accesso 17 ottobre 2025).

Id., *What is a critical digital edition?*, «Variants: The Journal of the European Society for Textual Scholarship», 1, 2002, pp. 43–62, consultabile al seguente link: <<https://doi.org/10.5281/zenodo.6533168>> (ultimo accesso 17 ottobre 2025).

Schmidt D., *True interoperability for digital scholarly editions*, «Umanistica Digitale», 10, 2021, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/12604>> (ultimo accesso 17 ottobre 2025).

Stella F., *Testi letterari e analisi digitale*, Roma, Carocci, 2018.

Zaccarello M., *Testo, teoria, edizione. Come cambia la filologia nel contesto digitale*, «Ecdotica», 1, 2019, pp. 186–201.

Trascrizione automatica del testo: software di OCR e HTR

Bastianello E., *Digitalizzazione, trascrizione, citazione: le fonti testuali per le pubblicazioni digitali*, «DigItalia», 18(2), 2023, pp. 162-172, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00084>> (ultimo accesso 17 ottobre 2025).

Id., Baumgartner R., *L'applicazione del riconoscimento testi neurale per la realizzazione di ristampe digitali*, in *La memoria digitale. Forme del testo e organizzazione della conoscenza*, Atti del XII Convegno Annuale AIUCD (Università di Siena, 5-7 giugno 2023) a cura di E. Carbé et al., AIUCD, 2023, pp. 15-22

Bazzaco S., *La trascrizione automatica di documenti a stampa antichi. Appunti per un modello di riconoscimento della tipografia in corsivo*, «DigItalia», 1, 2024, pp. 63-86, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00094>> (ultimo accesso 17 ottobre 2025).

Cordell R., Smith D., *Report: A Research Agenda for Historical and Multilingual Optical Character Recognition*, Northeastern University Library, 2018, consultabile al seguente link: <<http://hdl.handle.net/2047/D20297452>> (ultimo accesso 17 ottobre 2025).

Cordell R., *Q i-jtb the Raven': Taking Dirty OCR Seriously*, «Book History», 20, 2017, pp. 188-225, consultabile al seguente link: <<https://scispace.com/pdf/q-i-jtb-the-raven-taking-dirty-ocr-seriously-4joz3t45nt.pdf>> (ultimo accesso 17 ottobre 2025).

Dassonneville D. et al., *Compte-rendu de la journée d'étude 'Point HTR 2022' Transkribus/eScriptorium: transcrire, annoter et éditer numériquement des documents d'archives* (Rapport de recherche, 9 June 2022), consultabile al seguente link: <<https://hal.science/hal-03692413>> (ultimo accesso 17 ottobre 2025).

Dunley R., *Machines Reading the Archive: Handwritten Text Recognition Software*, «The National Archives Blog», (19 marzo 2018) consultabile al seguente link <<https://blog.nationalarchives.gov.uk/machines-reading-the-archive-handwritten-text-recognition-software/>> (ultimo accesso 17 ottobre 2025).

Günter Mühlberger et al., *Transforming scholarship in the archives through Handwritten Text Recognition. Transkribus as a case study*, «Journal of Documentation», 75(5), 2019, pp. 954-976, consultabile al seguente link: <<https://doi.org/10.1108/JD-07-2018-0114>> (ultimo accesso 17 ottobre 2025).

Hodel T. et al., *General Models for Handwritten Text Recognition: Feasibility and State-of-the-Art. German Kurrent as an Example*, «Journal of Open Humanities Data», 7, 2021, consultabile al seguente link: <<https://doi.org/10.5334/johd.46>> (ultimo accesso 17 ottobre 2025).

Kahle P. et al., *Transkribus. A Platform for Automated Text Recognition and Searching of Historical Documents*, in *2019 15th IAPR International Conference on eScience* (San Diego,

24-27 settembre 2019), IEEE Publisher, pp. 463-466, consultabile al seguente link: <https://ieeexplore.ieee.org/document/904176> (ultimo accesso 17 ottobre 2025).

Kahle P., *Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents*, in *2017 14th IAPR International Conference on Document Analysis and Recognition – ICDAR* (Kyoto, 09-15 novembre 2017), IEEE Publisher, consultabile al seguente link: <https://doi.org/10.1109/ICDAR.2017.307> (ultimo accesso 17 ottobre 2025).

Mancinelli T., *Early printed edition and OCR techniques: what is the state-of-art? Strategies to be developed from the working-progress Mambrino project work*, «Historias Fingidas», 4, 2016, pp. 255-260.

Mansutti S., Spina S., *Trascrivere, analizzare e pubblicare documenti d'archivio: Transkribus e l'intelligenza artificiale al servizio dello storico*, in *La memoria digitale. Forme del testo e organizzazione della conoscenza*, Atti del XII Convegno Annuale AIUCD (Università di Siena, 5-7 giugno 2023) a cura di E. Carbé et al., AIUCD, 2023, pp. 410-415.

Nockels J. et al., *Understanding the application of handwritten text recognition technology in heritage contexts: a systematic review of Transkribus in published research*, «Archival Science», 22 (3), 2022, pp. 367–392, consultabile al seguente link: <https://doi.org/10.1007/s10502-022-09397-0> (ultimo accesso 17 ottobre 2025).

Nockels J., *Are Digital Humanities platforms facilitating sufficient diversity in research? A study of the Transkribus Scholarship Programme*, «Digital Scholarship in the Humanities», 40, 2025, consultabile al seguente link: https://academic.oup.com/dsh/article/40/Supplement_1/i46/7646909 (ultimo accesso 17 ottobre 2025).

Perdiki E., *Preparing Big Manuscript Data for Hierarchical Clustering with Minimal HTR Training*, «Journal of Data Mining and Digital Humanities», 2023, consultabile al seguente link: <https://hal.science/hal-03880102v4> (ultimo accesso 17 ottobre 2025).

REED-COOP, *Should I use a base model when training a custom model?*, «Transkribus blog», 7 febbraio 2025, consultabile al seguente link: <https://blog.transkribus.org/en/using-base-model-when-training-a-custom-model> (ultimo accesso 17 ottobre 2025).

Id., *4 AI models for transcribing old Italian handwriting and typewritten texts*, «Transkribus blog», 27 novembre 2024, consultabile al seguente link: <https://blog.transkribus.org/en/4-ai-models-for-transcribing-old-italian-handwriting-and-typewritten-texts> (ultimo accesso 17 ottobre 2025).

Id., *How to retrain a model in Transkribus*, «Transkribus blog», 8 novembre 2024, consultabile al seguente link: <https://blog.transkribus.org/en/how-to-retrain-a-model-in-transkribus> (ultimo accesso 17 ottobre 2025).

Id., *What are Field Models and how do they work?*, «Transkribus blog», 28 agosto 2024, consultabile al seguente link: <https://blog.transkribus.org/en/what-are-field-models-and-how-do-they-work> (ultimo accesso 17 ottobre 2025).

Id., *Transkribus Sites - Publishing documents made easy*, «Transkribus blog», 10 gennaio 2024, consultabile al seguente link: <https://blog.transkribus.org/en/transkribus-sites-publishing-documents-made-easy> (ultimo accesso 17 ottobre 2025).

Id., *Preserving Cultural Heritage: Transkribus Integration with Wikimedia Projects*, «Transkribus blog», 4 ottobre 2023, consultabile al seguente link: <https://blog.transkribus.org/en/preserving-cultural-heritage-transkribus-integration-with-wikimedia-projects> (ultimo accesso 17 ottobre 2025).

Id., *Introducing the new Transkribus web app*, «Transkribus blog», 29 agosto 2024, consultabile al seguente link: <https://blog.transkribus.org/en/introducing-the-new-transkribus-web-app> (ultimo accesso 17 ottobre 2025).

Id., *Introducing Field Models - Trainable Layout AI in Transkribus*, «Transkribus blog», 25 luglio 2023, consultabile al seguente link: <https://blog.transkribus.org/en/introducing-field-models-trainable-layout-ai-in-transkribus> (ultimo accesso 17 ottobre 2025).

Id., *Introducing Transkribus Super Models - Get Access to 'The Text Titan I'*, «Transkribus blog», 9 luglio 2023, consultabile al seguente link:

<https://blog.transkribus.org/en/introducing-transkribus-super-models-get-access-to-the-text-titan-i>> (ultimo accesso 17 ottobre 2025).

Id., *What is Ground Truth?*, «Transkribus blog», 9 maggio 2023, consultabile al seguente link: <https://blog.transkribus.org/en/what-is-ground-truth>> (ultimo accesso 9 aprile 2025).

Id., *How to digitise archival materials with Transkribus*, «Transkribus blog», 3 aprile 2023, consultabile al seguente link: <https://blog.transkribus.org/en/how-to-digitise-archival-materials-with-transkribus>> (ultimo accesso 17 ottobre 2025).

Id., *A Short History of Transkribus with Günter Mühlberger*, «Transkribus blog», 21 febbraio 2023, consultabile al seguente link: <https://blog.transkribus.org/en/a-short-history-of-transkribus-with-gunter-muhlberger>> (ultimo accesso 17 ottobre 2025).

Id., *How to read cursive handwriting in historical documents*, «Transkribus blog», 15 dicembre 2022, consultabile al seguente link: <https://blog.transkribus.org/en/how-to-read-cursive-handwriting-in-historical-documents>> (ultimo accesso 17 ottobre 2025).

Id., *What is handwriting recognition and how does it work?*, «Transkribus blog», 11 agosto 2022, consultabile al seguente link: <https://blog.transkribus.org/en/what-is-handwriting-recognition-and-how-does-it-work>> (ultimo accesso 17 ottobre 2025).

Id., *OCR vs. HTR or 'What is AI, actually?'*, «Transkribus blog», 9 maggio 2021, consultabile al seguente link: <https://blog.transkribus.org/en/insights/ocr-vs-htr>> (ultimo accesso 17 ottobre 2025).

Sánchez J. A. et al., *Handwritten Text Recognition for Historical Documents in the tranScriptorium Project*, in *DATeCH'14 Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage* (Madrid, 19-29 May 2014), New York, Association for Computing Machinery, 2014, consultabile al seguente link: <https://doi.org/10.1145/2595188.2595193> (ultimo accesso 17 ottobre 2025).

Seaward L., Kallio M., *Transkribus: Handwritten Text Recognition Technology for Historical Documents*, in *Proceedings of Digital Humanities Conference* (Montréal, 8-11 August 2017), consultabile al seguente link: < <https://dh-abstracts.library.cmu.edu/works/4193> > (ultimo accesso 17 ottobre 2025).

Schwarz-Ricci V. I., *Handwritten Text Recognition per registri notarili (secc. XV- XVI): una sperimentazione*, «Umanistica Digitale», 13, 2022, pp. 171-181, consultabile al seguente link: <<https://umanisticadigitale.unibo.it/article/view/14926>> (ultimo accesso 17 ottobre 2025).

Singh A. et al., *A Survey of OCR Applications*, «International Journal of Machine Learning and Computing», 2(3), 2012, consultabile al seguente link: <https://www.researchgate.net/publication/269395249_A_Survey_of_OCR_Applications> (ultimo accesso 17 ottobre 2025).

Spina S., *Digitality as a longue durée historical phenomenon*, «Umanistica Digitale», 18, 2024, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/19420>> (ultimo accesso 17 ottobre 2025).

Id., *Artificial Intelligence in archival and historical scholarship workflow: HTR and ChatGPT*, «Umanistica Digitale», 16, 2023, consultabile al seguente link: <<http://doi.org/10.6092/issn.2532-8816/17205>> (ultimo accesso 17 ottobre 2025).

Stauder F., *Introducing the first major Transkribus release of 2025: What's changed and why?*, «Transkribus blog», 26 febbraio 2025, consultabile al seguente link: <<https://blog.transkribus.org/en/february-release-2025>> (ultimo accesso 17 ottobre 2025).

Stokes P. et al., *The eScriptorium VRE for Manuscript Cultures*, «Classics@ Journal», 18 (1), 2021, consultabile al seguente link: <<https://classics-at.chs.harvard.edu/classics18-stokes-kiessling-stokl-ben-ezra-tissot-gargem/>> (ultimo accesso 17 ottobre 2025).

Stokes P., Kiessling B., *Sharing Data for Handwritten Text Recognition (HTR)*, «Digital Humanities in Practice», 2024, consultabile al seguente link: <<https://hal.science/hal-04444641v1/document>> (ultimo accesso 17 ottobre 2025).

Terras M. et al., *The artificial intelligence cooperative: READ-COOP, Transkribus, and the benefits of shared community infrastructure for automated text recognition*, «Open Research

Europe», 2025, consultabile al seguente link: <<https://open-research-europe.ec.europa.eu/articles/5-16>> (ultimo accesso 17 ottobre 2025).

Terras M., *Chapter 7: Inviting AI into the Archives: The Reception of Handwritten Recognition Technology into Historical Manuscript Transcription*, in *Archives, Access and Artificial Intelligence. Working with Born-Digital and Digitized Archival Collections*, a cura di L. Jaillant, Bielefeld University Press, 2022, pp. 179-204, consultabile al seguente link: <<https://doi.org/10.14361/9783839455845-008>> (ultimo accesso 17 ottobre 2025).

Vinciarelli A., *Offline Cursive Handwriting: From Word To Text Recognition*, IDIAP Research Record, 2003, consultabile al seguente link: <<https://infoscience.epfl.ch/record/82879>> (ultimo accesso 17 ottobre 2025).

Weichselbaumer N. et al., *New Approaches to OCR for Early Printed Books*, «DigItalia», 15(2), 2020, pp.74-87, consultabile al seguente link: <<https://doi.org/10.36181/digitalia-00015>> (ultimo accesso 17 ottobre 2025).

Sitografia

Transkribus Help Center, disponibile al seguente link: <<https://help.transkribus.org/getting-started>> (ultimo accesso 17 ottobre 2025).

Codifica XML-TEI e pubblicazione di edizioni digitali con EVT

AA.VV., *Creating a Digital Scholarly Edition with the Text Encoding Initiative*, a cura di M. Burghart, DEMM (Digital Editing of Medieval Manuscripts), 2017, consultabile al seguente link: <https://ec.europa.eu/programmes/erasmus-plus/project-result-content/cfb03c8e-6765-477e-aa4d-e2b162f57068/IO2_Creating%20a%20digital%20edition%20with%20the%20TEI.pdf> (ultimo accesso 17 ottobre 2025).

Brett B., *TEI, the Walt Whitman Archive, and the Test of Time*, «Journal of the Text Encoding Initiative», 13, 2020, consultabile al seguente link: <<https://doi.org/10.4000/jtei.3249>> (ultimo accesso 17 ottobre 2025).

Burghart M., *The TEI Critical Apparatus Toolbox: Empowering Textual Scholars through Display, Control, and Comparison Features*, «Journal of the Text Encoding Initiative», 10, 2016, consultabile al seguente link: <<http://jtei.revues.org/1520>> (ultimo accesso 17 ottobre 2025).

Burnard L., Sperger-McQueen C. M., *Il manuale TEI Lite. Introduzione alla codifica elettronica dei testi letterari*, a cura di F. Ciotti, Milano, Edizioni Sylvestre Bonnard, 2005, consultabile al seguente link: <https://tei-c.org/Vault/P5/1.0.1/doc/tei-p4-lite/teiu5_it.html> (ultimo accesso 17 ottobre 2025).

Burnard L., *What Is the Text Encoding Initiative?*, Marseille, OpenEdition Press, 2014, consultabile al seguente link: <<https://books.openedition.org/oep/426>> (ultimo accesso 17 ottobre 2025).

Cummings J., *A world of difference: Myths and misconceptions about the TEI*, «Digital Scholarship in the Humanities», 34, 2019, pp. 158-179, consultabile al seguente link: <<https://doi.org/10.1093/llc/fqy071>> (ultimo accesso 17 ottobre 2025).

Id., *Opening the Book: Data Models and Distractions in Digital Scholarly Editing*, «International Journal of Digital Humanities», 1(2), 2019, pp. 179-93, consultabile al seguente link: <<https://doi.org/10.1007/s42803-019-00016-6>> (ultimo accesso 22 marzo 2025).

Dillen W., *Sequentiality in Genetic Digital Scholarly Editions. Models for Encoding the Dynamics of the Writing Process*, in *Digital Humanities 2016. Conference Abstracts* (Cracovia, 11-16 luglio 2016), a cura di Maciej Eder e Jan Rybicki, Alliance of Digital Humanities Organizations (ADHO), 2016, pp. 174-175, consultabile al seguente link: <<https://doi.org/10.17613/M6GB9B>> (ultimo accesso 17 ottobre 2025).

Estill L., *Encoding the Edge: Manuscript Marginalia and the TEI*, «Digital Literary Studies», 1(1), 2016, consultabile al seguente link: <<https://journals.psu.edu/dls/article/view/59715/59912>> (ultimo accesso 17 ottobre 2025).

Milone F., *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.

Muñoz T., Viglianti R., *Texts and Documents: New Challenges for TEI Interchange and Lessons from the Shelley-Godwin Archive*, «Journal of the Text Encoding Initiative», 8, 2014,

consultabile al seguente link: <<https://doi.org/10.4000/jtei.1270>> (ultimo accesso 17 ottobre 2025).

Nava B., *Siamo tutti bédieriani? Prospettive per le edizioni genetiche digitali*, «Umanistica Digitale», 6 (14), 2022, pp. 19-40, consultabile al seguente link: <<https://doi.org/10.6092/issn.2532-8816/14949>> (ultimo accesso 17 ottobre 2025).

Nava B., Priore B., *Codificare (al)'Infinito*, in *La memoria digitale. Forme del testo e organizzazione della conoscenza*, Atti del XII Convegno Annuale AIUCD (Università di Siena, 5-7 giugno 2023) a cura di E. Carbé et al., AIUCD, 2023, pp. 421-423.

Rosselli Del Turco R. et al., *Introduzione all'edizioni digitali: preparazione con codifica XML-TEI e visualizzazione con il software EVT*, in *La memoria digitale. Forme del testo e organizzazione della conoscenza*, Atti del XII Convegno Annuale AIUCD (Università di Siena, 5-7 giugno 2023) a cura di E. Carbé et al., AIUCD, 2023, pp. 410-415.

Rosselli Del Turco R. et al., *There and back again: what to expect in the next EVT version*, in *AIUCD 2022 - Intersezioni: filosofia, arti, media*, Atti dell'XI Convegno Nazionale dell'AIUCD (Lecce, 1-3 giugno 2022), a cura di F. Ciracì et al., AIUCD, 2022, pp. 212-217.

Rosselli Del Turco R., Monella P., *Extending the DSE: LOD Support and TEI/IIIF Integration in EVT*, in *Atti del IX Convegno Annuale AIUCD. La svolta inevitabile: sfide e prospettive per l'informatica umanistica* (Milano, 15-17 gennaio 2020), a cura di Cristina Marras et al., pp. 148–55, consultabile al seguente link: <<https://doi.org/10.6092/UNIBO/AMSACTA/6316>> (ultimo accesso 17 ottobre 2025).

Rosselli Del Turco R., Di Pietro C., *La visualizzazione di edizioni digitali con EVT: una soluzione per edizioni diplomatiche e critiche*, «Ecdotica», 1, 2019, pp. 148–73.

Rosselli Del Turco R., *Designing an advanced software tool for Digital Scholarly Editions. The inception and development of EVT (Edition Visualization Technology)*, «Textual Cultures», 12 (2), 2019, pp. 91-111, consultabile al seguente link: <<https://doi.org/10.14434/textual.v12i2.27690>> (ultimo accesso 17 ottobre 2025).

Rosselli Del Turco R., Di Pietro C., Martignano C., *Progettazione e implementazione di nuove funzionalità per EVT 2: lo stato attuale dello sviluppo*, «Umanistica Digitale», 7, 2019,

consultabile al seguente link: <<https://doi.org/10.6092/ISSN.25328816/9322>> (ultimo accesso 17 ottobre 2025).

Id., *Between Innovation and Conservation: The Narrow Path of User Interface Design for Digital Scholarly Editions*, in *Digital Scholarly Editions as Interfaces* (Pubblicazioni dell'Istituto per la Documentologia e l'Editing, vol. 12), a cura di R. Bleier, M. Bürgermeister et al., Norderstedt, BoD, 2018, pp. 129-160.

Rosselli Del Turco R. et al., *Edition Visualization Technology: A Simple Tool to Visualize TEI-based Digital Editions*, «Journal of the Text Encoding Initiative», 8, 2014, consultabile al seguente link: <<https://doi.org/10.4000/jtei.1077>> (ultimo accesso 17 ottobre 2025).

Siciliano A., Del Grosso A. M., *From Print to Digital: An Encoding Model for the Scholarly Edition of Giorgio Bassani's Notes*, «Umanistica Digitale», 6(13), 2022, pp. 21-48, consultabile al seguente link: <<https://umanisticadigitale.unibo.it/article/view/13688>> (ultimo accesso 17 ottobre 2025).

Tancredi G., Fenu C., *XML-TEI: un modello per la filologia d'autore*, in *AIUCD 2022 - Intersezioni: filosofia, arti, media*, Atti dell'XI Convegno Nazionale dell'AIUCD (Lecce, 1-3 giugno 2022), a cura di F. Ciraci et al., AIUCD, 2022, pp. 218-222.

TEI Consortium, *TEI P5: Guidelines for Electronic Text Encoding and Interchange* (P5 Version 4.9.0. Last updated on 24th January 2025), consultabile al seguente link: <<https://www.tei-c.org/guidelines/p5/>> (ultimo accesso 17 ottobre 2025).

TEI Manuscript Special Interest Group (TEI SIG MS), *An Encoding Model for Genetic Editions*, consultabile al seguente link: <<https://tei-c.org/Vault/TC/tcw19.html>> (ultimo accesso 17 ottobre 2025).

Sitografia

Corso online *Digital Scholarly Editions: Manuscripts, Texts, and TEI Encoding*, a cura di M. Burghart, E. Pierazzo et al, disponibile al seguente link: <<https://teach.dariah.eu/course/view.php?id=32>> (ultimo accesso 17 ottobre 2025).