

Verso la svolta computazionale della critica dantesca

Fara Autiero¹, Vittorio Celotto², Gennaro Ferrante³, Chiara Fusco⁴, Sandra Gorla⁵
Giuseppe Andrea Liberti⁶, Mariangela Palomba⁷, Serena Picarelli⁸, Stefano Angelo Rizzo⁹, Silvia Tripodi¹⁰

¹ Scuola Superiore Meridionale, f.autiero@ssmeridionale.it

² Università di Napoli Federico II, vittorio.celotto@unina.it

³ Università di Napoli Federico II, gennaro.ferrante@unina.it

⁴ Scuola Superiore Meridionale, chiara.fusco@unina.it

⁵ Università di Napoli Federico II, sandra.gorla@unina.it

⁶ Università di Napoli Federico II, giuseppeandrea.liberti@unina.it

⁷ Sapienza Università di Roma, mariangela.palomba@uniroma1.it

⁸ Università di Napoli Federico II, serena.picarelli@unina.it

⁹ Università di Catania, stefano.angelo.rizzo@gmail.com

¹⁰ Università di Napoli Federico II, silvia.tripodi@unina.it

ABSTRACT (ITALIANO)

Gli studi sulla Commedia di Dante affrontano da sempre sfide significative: l'assenza del testo originale, il successo precoce e ampio del poema, la varietà della trasmissione materiale e paleografica e la ricchezza di illustrazioni e commenti richiedono un costante dialogo tra discipline storiche e filologiche. Integrare metodi computazionali è oggi cruciale per risolvere questioni irrisolte e approfondire la comprensione della ricezione e interpretazione del poema. Basandosi sui risultati dell'Edizione Nazionale dei Commenti danteschi e dell'Illuminated Dante Project, il Naples Dante Project adotta un approccio multidisciplinare, concentrandosi su tre ambiti: medium (manoscritti, libri, frammenti), testo (Commedia, commenti) e immagine (miniature, illustrazioni, disegni). Il progetto funge da aggregatore digitale, integrando dataset sui vettori materiali, versioni testuali e caratteristiche iconografiche ed esegetiche in un archivio IIIF potenziato da tecnologie avanzate come reti neurali, imaging multimodale, modellazione semantica dei dati e computer vision. Recenti finanziamenti e collaborazioni rendono oggi possibile una vasta campagna di digitalizzazione, necessaria in virtù dell'ampliamento dei corpora, e l'innovazione dell'architettura dati. Le prime applicazioni di riconoscimento automatico della scrittura su manoscritti danteschi hanno inoltre segnato una svolta nella prassi ecdotica di tipo stemmatico. Queste acquisizioni recenti orientano il progetto verso tre obiettivi principali: 1. ILLUDANT: un catalogo delle illustrazioni della Commedia (manoscritti miniati, libri a stampa, disegni dal XIV al XVI secolo). 2. ECOMMEDIA: recensio integrale della tradizione manoscritta assistita dall'Intelligenza Artificiale per una nuova edizione critica che combina metodo stemmatico e approccio computazionale. 3. COMMENTI DANTESCHI ANTICHI: database interoperabile di commenti in volgare e latino sulla Commedia fino al XV secolo.

Parole chiave: Semantic Web; Multi-Modal Imaging; HTR; Handwriting Recognition; NLP

ABSTRACT (ENGLISH)

Toward the computational turn of Dante studies for AIUCD2025

Research on Dante's Commedia has always faced significant challenges. The absence of the original text, the poem's early and widespread success, the variety of its material and palaeographic transmission, and the wealth of illustrations and commentaries has constantly required a dialogue between historical and philological disciplines. Today, integrating computational methods into this dialogue is crucial for resolving long-standing questions and deepening understanding of the poem's reception and interpretation. Building on the achievements of the Edizione Nazionale dei Commenti danteschi and the Illuminated Dante Project, Naples Dante Project adopts a multidisciplinary approach, focusing on three key areas: medium (manuscripts, books, fragments), text (Commedia, commentaries), and image (illuminations, book illustrations, drawings). It serves as a digital aggregator, integrating datasets on material vectors, text versions, and iconographic and exegetical features into a IIIF-compliant image archive enhanced by advanced technologies such as Deep Neural Networks, Multi-modal Imaging, Semantic Data Modeling, and Computer Vision. Recent funding and collaborations have expanded digitisation efforts, enriched corpora, and restructured data architecture. Preliminary Handwritten Text Recognition on Dante manuscripts has led to a breakthrough in editorial practices based on the stemmatic method. These advancements support

three main objectives: 1. Illustrated Dante: A catalogue of Commedia illustrations, encompassing illuminated manuscripts, printed books, and in-series drawings from the 14th to 16th century; 2. eCommedia: An AI-assisted recensio of the Commedia's manuscript tradition, enabling a new critical edition that merges computational and traditional methods; 3. Commenti Danteschi Antichi: A database of Commedia vernacular and Latin commentaries produced by the end of 15th century, and grounded in reliable philological editions.

Keywords: Semantic Web; Multi-Modal Imaging; HTR; Handwriting Recognition; NLP

1. INTRODUZIONE

La Commedia di Dante godette di grande popolarità subito dopo la morte del poeta nel 1321. Copiata inizialmente a Ravenna e probabilmente trascritta da amanuensi bolognesi, si diffuse rapidamente in Toscana e nel resto dell'Italia centrale e settentrionale. La tradizione manoscritta diretta del poema, risalente al XIV e XV secolo, comprende oggi 580 manoscritti con almeno una cantica e 131 frammenti. Stimando un tasso di decimazione dell'85% per i manoscritti medievali, si calcola una produzione iniziale di circa 5.000 esemplari nei primi due secoli, rendendolo il testo volgare più diffuso del Medioevo occidentale, accanto ai classici latini. Questa tradizione sfruttò tutti i formati, materiali, impaginazioni e stili di scrittura disponibili, rispondendo alle esigenze di lettori laici e religiosi. Amanuensi con preparazione variabile lavorarono alla Commedia in contesti professionali e semiprofessionali, facendone un riferimento per la trasmissione materiale e grafica dei testi volgari. Tuttavia, le informazioni su questa varietà di evidenze materiali restano largamente scollegate. Gli apparati illustrativi ed esegetici presenti nei manoscritti testimoniano, inoltre, l'esigenza di spiegazioni verbali e visive per un testo semanticamente complesso. Tra XIV e XV secolo, numerosi commentatori professionisti interpretarono il poema, generando manoscritti ricchi di chiose: circa il 44% dei codici include apparati illustrativi, e l'8% presenta cicli iconografici estesi che supportano visivamente la lettura. Questo vasto metatesto richiede ancora definizione, classificazione e studio approfondito.

Ad oggi, manca un'edizione critica capace di gestire l'intera varia lectio della tradizione manoscritta. La ricostruzione del testo senza l'originale rappresenta una sfida per la critica testuale, in particolare con il metodo stemmatico. Nonostante i contributi di studiosi, le edizioni critiche si sono concentrate su tradizioni parziali e un numero limitato di manoscritti. Recentemente, Paolo Trovato ha ampliato la genealogia della tradizione (Tonello & Trovato, 2022), ma i loci critici analizzati rappresentano ancora una parte ridotta del testo. Le tecnologie digitali possono accelerare la classificazione e l'analisi di questa tradizione complessa, permettendo di studiarla per la prima volta nella sua interezza. In particolare, l'allineamento automatico delle trascrizioni integrali da HTR allo scopo di individuare lacune e vistosi errori (vd. sotto) può fornire argomenti più efficaci sulle relazioni genealogiche tra i testimoni. Un database esteso a tutta la tradizione manoscritta, collegato a un archivio digitale di facsimili e aggiornato tramite esperti e crowd-sourcing, semplificherebbe l'inserimento dei dati e fornirebbe informazioni verificabili. Tecniche avanzate di imaging diagnostico permetterebbero l'analisi dei manoscritti per dati invisibili all'occhio umano. L'intelligenza artificiale e il deep learning potrebbero migliorare la filologia materiale, testuale e attributiva, con progressi nel riconoscimento iconografico, nell'identificazione delle mani dei copisti, nella selezione delle varianti testuali e nell'analisi linguistico-stilistica di testi dall'attribuzione incerta, grazie a tecnologie come l'NLP. Il Naples Dante Project, nato da un progetto PRIN 2022 e sintesi dei progetti danteschi sviluppati a Napoli nell'ultimo decennio, mira a integrare questa complessità in un'unica piattaforma online in cui le componenti del poema, dai supporti materiali alle varianti testuali e linguistiche, saranno dinamicamente collegate a riproduzioni digitali ad alta definizione, utilizzando Deep Learning, imaging multimodale, modellazione semantica dei dati e Computer Vision.

2. VERSO LA SVOLTA COMPUTAZIONALE DELLA CRITICA DANTESCA

Illuminated Dante Project (2015-in corso)

Lanciato nel 2015 con il finanziamento STAR, Illuminated Dante Project (IDP) ha costituito la base pionieristica del Naples Dante Project. Intende offrire un catalogo digitale dei manoscritti della Commedia con illustrazioni del poema e combina un archivio digitale ad alta risoluzione con un database codicologico

e iconografico. Online dal 2019, IDP ospita attualmente 101 dei 274 pezzi del corpus, diventando un punto di riferimento globale per la codicologia digitale dantesca e l'iconografia antica, ampiamente utilizzato nella ricerca e nell'educazione. Ospitato dall'Università di Napoli Federico II (www.dante.unina.it/idp), IDP aderisce ai principi F.A.I.R., con metadati conformi a TEI condivisi con Manus online (Istituto Centrale del Catalogo Unico-ICCU) e manifesti JSON per IIIF. In collaborazione con Biblissima, IDP è pioniere nell'adozione dell'API IIIF in Italia, portando l'Università Federico II a ospitare la Conferenza Annuale IIIF nel 2023. Il database MySQL, sviluppato da Luigi Tessarolo, ridefinisce le descrizioni iconografiche utilizzando categorie indicizzabili come "Attribuzione stilistica", "Tipologia decorativa" e "Iconografia", analizzando le relazioni immagine-testo e i soggetti. L'interfaccia di IDP supporta l'esplorazione per biblioteche di conservazione, descrizioni dettagliate o una mappa geografica delle posizioni dei manoscritti. Nel corso di un decennio, ha fornito accesso aperto e di alta qualità a risorse digitalizzate, promuovendo avanzamenti negli ambiti codicologico, paleografico, filologico e iconografico. I risultati della ricerca (2015-20) sono stati presentati a conferenze e pubblicati su riviste peer-reviewed (Ferrante, 2019; Autiero & Ferrante & Gorla & Picarelli, 2022; Ferrante, 2024a; Ferrante, 2024b).

Nuovi corpora I. Frammenti di Commedia (2021-in corso)

Nel 2021, grazie ad un finanziamento FRA, l'area di ricerca di IDP si è aperta a Fragments of Commedia (FraC), con l'obiettivo di digitalizzare e catalogare frammenti della Commedia databili tra XIV e XV secolo e di esaminarne filologicamente la lezione. Il progetto vuole rivalutare il ruolo dei frammenti, generalmente trascurati dagli studi precedenti, nella comprensione della trasmissione materiale del poema e nelle relazioni genealogiche tra i suoi testimoni. Il team di paleografi della scuola napoletana ha stabilito il protocollo di ricerca, progettato il corpus e avviato la digitalizzazione, coprendo tutti i frammenti riutilizzati (la cosiddetta tradizione diretta) della Commedia. Questo sforzo ricostruisce testimoni perduti e indaga le circostanze del loro riutilizzo, con descrizioni che affrontano sia lo stato attuale dei frammenti che la loro provenienza originale. Dei 131 frammenti riutilizzati identificati, 40 sono già stati digitalizzati. Lo studio preliminare ha messo in discussione le idee precostituite sull'origine dei frammenti riutilizzati e ha rivelato il significato di materiali trascurati, in particolare i palinsesti. Tra le scoperte vi è un commento latino sconosciuto e un testimone monastico, che gettano luce sulla circolazione di Dante in contesti religiosi. I risultati sono stati presentati a eventi chiave, tra cui *Entre fragments i prestatges / Entre fragmentos y anaqueles* (Barcelona, 4-5/07/2024) e *Naples Dante Project Per una svolta computazionale della critica dantesca* (Napoli, 18-19/12/2024).

Nuovi corpora II. Illustrated Dante-ILLUDANT (2022-in corso)

Il campo di indagine sulle immagini si è inoltre allargato con Illustrated Dante (ILLUDANT), basato su una proposta FIS 2022, al fine di indagare i primi momenti della mise-en-images del poema: a una prima fase, che coincide in linea di massima con la tradizione manoscritta ed è caratterizzata da una sperimentazione iconografica varia, influenzata da periodo, area geografica, stile o bottega, segue una seconda, in cui si evidenzia una selezione di temi dalla tradizione manoscritta e l'offerta di nuove letture, veicolate dalla tradizione a stampa dei primi due secoli (XV-XVI). ILLUDANT mira a individuare le tendenze iconografiche e a proporre una classificazione su larga scala delle illustrazioni presenti nei manoscritti, nelle edizioni a stampa e nei disegni in serie del poema dantesco.

Lo scopo del progetto consiste nella creazione di un archivio di immagini compatibile con IIIF e un database delle illustrazioni della Commedia tra XIV e XVI secolo basato sui principi dei Linked Open Data (cf. sotto). Includerà il catalogo di IDP, le illustrazioni dei libri a stampa e i disegni in serie del XVI secolo, anonimi e attribuiti a Giovanni Stradano, Alessandro Allori, Federico Zuccari, Jacopo Ligozzi, Ludovico Cigoli. Con il contributo degli storici del libro della scuola napoletana, il team ha definito il protocollo di ricerca e delineato il corpus dei libri a stampa e dei disegni (secc. XV-XVI): dei 15 incunaboli della Commedia, solo 6 sono illustrati, mentre 23 edizioni del XVI secolo contengono un apparato illustrativo. ILLUDANT ha selezionato gli esemplari di stampa da digitalizzare in base alla qualità, integrità, allineamento al progetto editoriale e caratteristiche speciali come presenza eventuale di colorazione artistica e postillati.

Il progetto si pone la sfida di rivoluzionare l'approccio all'analisi iconografica tramite l'integrazione delle più recenti acquisizioni legate all'Intelligenza Artificiale. Sulla base del corpus si prevede di addestrare un modello di riconoscimento delle immagini basato su CNN, con lo scopo di esplorare nuove strade per una classificazione iconografica su larga scala e ad alta precisione.

Nuovi corpora III. Commenti Danteschi Antichi-CoDA (2023-in corso)

Il Corpus CoDA, definito da un progetto PRIN2022, è stato integrato recentemente nell'ecosistema del Naples Dante Project. Esso mira a creare un database dei commenti antichi alla Commedia di Dante in volgare e in latino prodotti entro la fine del XV secolo. Una collaborazione con CNR-OVI e Roma La Sapienza è stata avviata per stabilire un protocollo di ricerca che includa: 1. La creazione della Bibliografia dei Commenti Danteschi (BCD) accessibile tramite PLUTO. 2. Lo sviluppo del Corpus OVI-CoDA utilizzando il software GattoWeb 3.3 con lemmatizzazione selettiva per interrogazioni lessicografiche analitiche. Il corpus include attualmente 67 testi (più della metà dell'intero corpus). Grazie ad una recente collaborazione con CNR-ILC, è stato creato un modello per codificare i testi in formato XML TEI-P5, con l'obiettivo di ampliare il sistema di query a scopi esegetici e ermeneutici, oltre che lessicografici. CoDA potenzia il preesistente Dartmouth Dante Project, offrendo una documentazione all'avanguardia sulla prima esegesi dantesca, migliorando la qualità filologica dei testi e le innovazioni tecnologiche per ottimizzare le possibilità di interrogazione.

Transizione computazionale I. Acquisizione digitale (2017-in corso)

La durabilità e la referenzialità di un progetto come IDP è stata possibile anche grazie ai numerosi accordi istituzionali e scientifici che hanno accompagnato il suo sviluppo. In particolare quello con la Direzione Generale Biblioteche e diritto d'autore (2017–2023), l'ICCU e le varie biblioteche statali e non statali (ad es. Trivulziana) e importanti istituzioni internazionali (ad es. Bibliothèque nationale de France, Oxford Bodleian Libraries) per digitalizzare manoscritti e frammenti della Commedia miniati, concedendo licenze di riproduzione e riuso gratuito in cambio di metadati strutturati. L'evoluzione al Naples Dante Project e l'ampliamento dei corpora implicherà il rinnovo e l'estensione di questo tipo di accordi. Dal 2017, il team di ricerca napoletano collabora con l'infrastruttura View dell'Università di Lovanio per applicare tecniche di imaging diagnostico - Imaging Multispettrale, XRF e Reflectance Transformation Imaging (RTI) - ai manoscritti di IDP. Nel 2019, il Microdome Multispettrale (MSMD) è stato testato sul manoscritto Dante Filippino (CF 2.16), insieme ad una selezione di codici miniati conservati alla Biblioteca dei Girolamini. A novembre 2024, una parte del team si è unita al gruppo di tecnici di Lovanio a Stoccarda per acquisire immagini di un manoscritto della Commedia conservato presso la Württembergische Landesbibliothek.

Transizione computazionale II. Architettura dati (2023-in corso)

Contemporaneamente alla stipula di accordi e collaborazioni e alla definizione di nuovi corpora, si è reso necessario un riesame completo dell'architettura dei dati per allinearla ai modelli astratti e ai principi di Linked Open Usable Data. La nuova architettura del Naples Dante Project organizzerà i dati in cinque modelli descrittivi: manoscritti danteschi (Dante_ms), libri (Dante_bk), frammenti (Dante_frg), disegni (Dante_drw) e iconografia (Dante_icon), collegati con trascrizioni testuali (Dante_txt). Per gestire framework diversificati di database e archivi di immagini (database relazionali e No-SQL, IIIF) e formati di output (XML-TEI, ALTO, JSON, RDF), è stato adottato il framework CADMUS back-end, sviluppato da Daniele Fusi con l'Università di Venezia. Questo sistema consente: a. Gestione di dati eterogenei tramite modelli indipendenti dall'implementazione. b. Modelli dinamici per un'espansione senza soluzione di continuità. c. Un ambiente unificato per database centralizzati, interfacce grafiche, editing, validazione e ricerche in tempo reale. Tale architettura, concepita per essere replicabile nelle sue componenti, può aprire la strada a future iniziative umanistiche computazionali.

Transizione computazionale III. Modellazione semantica (2024-in corso)

Nel suo originario sviluppo del PRIN 2022, NDP ha avviato una collaborazione con il Dipartimento di Matematica e Informatica di Catania per sviluppare una gerarchia concettuale per modelli relativi ai media danteschi (codicologici, bibliografici e figurativi) e all'indicizzazione iconografica (immagine). Il framework semantico mira a integrarsi con i cinque modelli di dati definiti dal gruppo di lavoro sull'Architettura dati (cf. sopra), utilizzando gli standard CIDOC CRM e LRMoo. Una mappatura preliminare del modello Dante_ms ha portato ai seguenti avanzamenti: a. Modellazione gerarchica delle opere e delle espressioni: la distinzione di Work (F1), Expression (F2), Manifestation (F3) e Item (F5) secondo LRMoo chiarisce diversi aspetti della Commedia, con sottoclassi per edizioni critiche, varianti testuali e dettagli macrotestuali (es. Cantiche e Canti). b. Descrizioni avanzate di manoscritti e testi: i Linguistic Objects (E33), collegati a Strings (E62), rappresentano colofoni, dediche e rubriche, applicabili ai manoscritti. c. Unità composite: LRMoo accoglie manoscritti con progetti codicologici o strati testuali multipli. d.

Interoperabilità con CADMUS: l'integrazione basata su dati JSON consente una elaborazione automatizzata dei dati conforme ai principi F.A.I.R.

Transizione computazionale IV. Handwriting Text Recognition (2024-in corso)

L'analisi automatica di manoscritti medievali e moderni tramite deep learning sta acquisendo importanza per migliorare e velocizzare la ricerca paleografica. Tuttavia, i progressi nel riconoscimento degli scribi sono stati limitati dalla mancanza di grandi quantità di dati sufficientemente annotati, poiché il processo di annotazione è complesso e dispendioso in termini di tempo. Sfruttando la classificazione paleografica dei codici danteschi già acquisita dalla bibliografia e in corso di acquisizione per la catalogazione dei manoscritti e dei frammenti di NDP, è stata recentemente avviata una collaborazione con Paolo Merialdo e il gruppo di ricerca in Paleografia Computazionale di Roma 3, che ha ottenuto risultati promettenti con i dati della Biblioteca Vaticana (Lastilla & Ammirati & Firmani & Komodakis & Merialdo & Scardapane, 2022). La collaborazione mira ad addestrare un tool di features paleografici basato su CNN utilizzando dati di Ground Truth (GT) sulla paleografia dantesca fornita dal gruppo napoletano. Un esperimento preliminare prevede il pre-addestramento della CNN su dati GT provenienti da 17 manoscritti di tre copisti di primo Trecento attivi digraficamente nelle tipologie della textualis e della cancelleresca.

Transizione computazionale V. HTR-Dante (2024-in corso)

NDP ha recentemente adottato tecniche avanzate di Deep Learning per trascrivere, allineare e riconoscere la più antica varia lectio della Commedia. Col supporto del PRIN 2022, del PNRR-Changes 2023-25 e del Dipartimento di Eccellenza 2023-27, tra ottobre e dicembre 2024 Thibault Clérice, ricercatore all'INRIA di Parigi e esperto di tecniche avanzate di Handwritten Text Recognition e Natural Language Processing, è stato invitato presso il Laboratorio di Digital Humanities dell'Università Federico II (<https://www.fedhlab.unina.it>), avviando ufficialmente il progetto HTR_Dante, con l'obiettivo di addestrare modelli HTR sulla piattaforma eScriptorium. Si tratta del primo progetto su larga scala per un'opera letteraria medievale italiana. Quanto all'annotazione per il training di Kraken - il pacchetto OCR alla base dell'ambiente eScriptorium - HTR_Dante si conforma alle linee guida generali del gruppo di ricerca HTRomance dell'INRIA, e intende rispettare le consegne, per ora ancora ufficiose, della recentemente costituita Commissione Nazionale per la Trascrizione e l'Edizione dei Testi in Italiano. Il gruppo di lavoro ha definito protocolli specifici per l'HTR di codici danteschi che hanno dovuto affrontare situazioni speciali come la risoluzione dello scriptio continua, la gestione delle lettere maiuscole per le iniziali delle terzine, e le variazioni formali dei copisti. Per il training dei modelli HTR sono stati isolati quattro tipi di scrittura come peculiari della tradizione manoscritta della Commedia: a. Scrittura cancelleresca (Chancery script). b. Scrittura mercantesca (Merchant script). c. Gotica/Testuale (Gothic/Textualis script). d. Scrittura umanistica (Humanistic script). In questa ricerca preliminare, la scrittura cancelleresca è stata la prima ad essere analizzata utilizzando sei manoscritti del gruppo stemmatico a₀ delineato da Tonello (Tonello, 2018). Dopo aver risolto gli iniziali problemi di trascrizione automatica attraverso l'addestramento del modello, l'HTR realizzato con eScriptorium e il modello adattato CATmuS Medieval 1.5 ha raggiunto un'accuratezza superiore al 90%. Il workflow sarà esteso agli altri tre tipi di scrittura, con l'obiettivo di ottimizzare l'HTR per la tradizione dantesca e verificare le ipotesi stemmatiche più recenti.

Transizione computazionale VI. Riconoscimento automatico delle varianti (2024-in corso)

La fertile collaborazione in situ con Thibault Clérice (INRIA), Angelo Del Grosso e Federico Boschetti (CNRILC) ha aperto un ulteriore campo di indagine per automatizzare la collazione dei testi processati con HTR al fine di identificare errori significativi nella varia lectio integrale della tradizione e verificare le relazioni genealogiche tra i testimoni. Basandosi sul principio metodologico lachmanniano che le lacune testuali sistematiche possono rivelare congiunzioni significative tra i testimoni di un testo (Chiesa, 2022), Clérice ha sviluppato un algoritmo per rilevare lacune nei codici finora trascritti dal gruppo di lavoro HTR_Dante, utilizzando la versione XML dell'edizione critica di Petrocchi del 1966 come Ground Truth. Entro la fine di novembre 2024, l'algoritmo è stato perfezionato e applicato con successo a due manoscritti della Commedia, identificando tutte le loro lacune testuali (da singoli versi a più terzine) e variazioni significative mai riconosciute negli studi filologici precedenti. Questo rappresenta un progresso rivoluzionario nella metodologia della critica testuale, consentendo la ricostruzione dei testi poetici diffusi in ampie tradizioni manoscritte. I risultati preliminari, presentati alla conferenza internazionale di dicembre 2024 (Naples Dante Project. Per una svolta computazionale della critica dantesca), hanno dimostrato il potenziale di questo approccio. Il team prevede di applicare il modello a 20 manoscritti, perfezionarne le

capacità e presentare i risultati alla conferenza Computational Humanities Research 2025. Un'ulteriore fase consisterà nell'esame automatico delle varianti sul più ampio numero possibile di trascrizioni digitali dei codici della Commedia (obiettivo ideale: 400 testimoni su 580 della tradizione non frammentaria) con comunicazione dei risultati su riviste di alto profilo internazionale (es. Nature)

Transizione computazionale VII. Dante Critical Texts (2024-in corso)

Sempre nell'ambito del progetto PRIN 2022, e nell'ottica dello studio comparato del testo criticamente emendato della Commedia, NDP ha inaugurato Dante Critical Texts (DCT), uno strumento progettato per confrontare i testi della Commedia di Dante a partire da edizioni fondate su un esame di tipo stemmatico della tradizione manoscritta. Il progetto mira a rendere accessibili i principali risultati della filologia dantesca (emendatio), dall'edizione di Karl Witte del 1862 fino alle più recenti. DCT offre tre principali funzioni: a. Search: consente agli utenti di cercare occorrenze verbali. b. Texts: permette l'analisi comparativa tra le edizioni, evidenziando divergenze testuali. c. Editions: fornisce descrizioni concise e dettagliate delle varie edizioni. La funzione Texts include attualmente l'edizione di Petrocchi del 1966 (Petrocchi, 1966-67), quella di Malato del 2021 (Malato, 2021) e i testi critici dell'Inferno editi da Vandelli (Vandelli, 1921) e Tonello-Trovato (Tonello & Trovato, 2022). La pubblicazione del Purgatorio di Vandelli (Vandelli, 1921) è prevista a breve. La funzione Editions include già una descrizione dell'edizione di Petrocchi del 1966.

Il quadro delle ricerche appena delineato dimostra che Naples Dante Project sta puntando a integrare la complessità del suo campo di indagine in un unico aggregatore, in cui i vettori materiali del poema, le varianti testuali e gli apparati iconografici ed esegetici siano classificabili con precisione, collegabili in modo dinamico a riproduzioni ad alta risoluzione, oltre che ricercabili in contesto LOD. L'adozione di tecniche di Deep Learning, Imaging Multimodale, Modellazione Semantica dei Dati e di Computer Vision per interconnettere caratteristiche materiali, testuali, meta-testuali, linguistiche e figurative, offriranno una rappresentazione articolata e completa dell'opera dantesca. La divulgazione dei risultati, dagli archivi digitali ai modelli di intelligenza artificiale, in open access e tramite un'unica piattaforma, garantirà una piena condivisione dei risultati del progetto non soltanto con la comunità accademica e i ricercatori ma soprattutto con l'utenza non specialistica, in un'ottica di comunità di ricerca inclusiva e globale che possa confrontarsi con la tradizione testuale e iconografica della Commedia.

RINGRAZIAMENTI

Il progetto è finanziato da: PRIN 2022 - NAPLES DANTE PROJECT; Spoke 3 del PNRR PE 5 Changes - Digital Libraries, Archives and Phylogeny; Dipartimento di Eccellenza 2023-2027, Dipartimento di Studi Umanistici, Università degli Studi di Napoli Federico II.

BIBLIOGRAFIA

- Autiero, F., & Ferrante, G., & Gorla, S., & Picarelli, S. (2022). Dante e Petrarca allo (stesso) scrittoio. Per lo sviluppo di un'ontologia di IDP a partire dall'istanza manoscritti di Itinera. Culture digitali. Intersezioni: filosofia, arti, media. Precedings della 11a conferenza nazionale AIUCD. Ciraci, F., & Miglietta, G., & Gatto, C. (Eds.). 169-171.
- Chiesa, P. (2022). Elementi di critica testuale. Bologna: Patròn.
- Ferrante, G. (2019). Illuminated Dante Project. Per un archivio digitale delle più antiche illustrazioni della Commedia. I. Un case study quattrocentesco (mss. Italien 74, Riccardiano 1004 e Guarneriano 200). Dante visualizzato. Carte ridenti ii: XV secolo. Prima parte. Ciccuto, M., & Livraghi, L.M.G. (Eds.) Firenze: Cesati, 229-255.
- Ferrante, G. (2024). Premessa. La Commedia di Dante nell'antica tradizione manoscritta. Per uno studio integrato e computazionale dei codici, dei testi e delle immagini. Ferrante, G. (Ed). Roma-Padova: Editrice Antenore, 7-11.
- Ferrante, G. (2024). Verso il Naples Dante Project. Recensio digitale integrale dei codici della Commedia di Dante. "A guisa d'uom che 'n dubbio si raccerta". Vecchie e nuove prospettive per la biografia e l'opera dantesca. Bardi, R., & Canova, L. (Eds.). Firenze: Cesati, 145-160.

- Lastilla, L., & Ammirati S., & Firmani D., & Komodakis N., & Merialdo P., & Scardapane S. (2022). Selfsupervised learning for medieval handwriting identification: A case study from the Vatican Apostolic Library. *Information Processing & Management*, 59, 3, 1-20.
- Malato, E. (Ed.). 2021. Dante Alighieri. *Inferno*. Roma: Salerno Editrice.
- Petrocchi, G. (Ed). 1966-67. Dante Alighieri. *La Commedia secondo l'antica vulgata*. Milano: Mondadori.
- Tonello, E. (2018). *Sulla tradizione tosco-fiorentina della 'Commedia' di Dante (secoli XIV-XV)*. Padova: libreriauniversitaria.it.
- Tonello, E., & Trovato, P. (Eds.). 2022. Dante Alighieri. *Commedia. Inferno*. Padova: libreriauniversitaria.it.
- Vandelli, G. (Eds.). 1921. Dante Alighieri. *La Divina Commedia*. Firenze: Bemporad.
- Tonello, E., & Trovato, P. (Eds.). 2022. Dante Alighieri. *Commedia. Inferno*. Padova: libreriauniversitaria.it.
- Vandelli, G. (Eds.). 1921. Dante Alighieri. *La Divina Commedia*. Firenze: Bemporad.