

L’edizione scientifica digitale estesa de «Il nome della rosa»: modellizzazione, workflow e il paradigma IDEA

Christian D’Agata

Università di Catania, Catania, Italia
christian.dagata@gmail.com

Abstract

Il contributo presenta l’edizione scientifica digitale de *Il nome della rosa* di Umberto Eco caratterizzata sin dalla sua ideazione dal dialogo tra filologia digitale e critica computazionale della letteratura in un’ottica profondamente interdisciplinare che prende il nome di paradigma ‘IDEA’ (Interpretazione, Didattica, Edizione, Annotazione). Mettendo insieme variantistica d’autore, annotazione digitale e didattica della letteratura il paradigma IDEA – elaborato appositamente per la presente edizione ma estendibile a tutti i testi tutelati dal diritto d’autore – si pone l’obiettivo di superare i limiti oggettivi di pubblicazione sul web in formato aperto di contenuti di letteratura contemporanea. Il contributo descrive quindi il portale *Le varianti della rosa*, un luogo virtuale dove presentare l’apparato critico delle varianti codificato in XML-TEI (e visualizzato tramite EVT2), alcuni contenuti didattici elaborati con TRAViz, Storymap JS e Timeline JS, l’annotazione in un *Domain Specific Language* sviluppata su Euporia e ulteriori approfondimenti critici sviluppati a partire dal *distant reading* di Franco Moretti e dalla lessicografia letteraria di Giuseppe Savoca. Si espone inoltre una riflessione sulla modellizzazione dell’edizione, sul concetto di edizione estesa, sul *workflow* che ha caratterizzato il progetto e si propone una *road map* dei contenuti futuri in un’ottica di sempre maggiore integrazione tra ricerca scientifica e *public humanities*.

The paper presents the digital scholarly edition of Umberto Eco's *The Name of the Rose* characterized since its conception by the dialogue between digital philology and computational literary criticism in an interdisciplinary perspective called 'IDEA' paradigm (Interpretation, Didactics, Edition, Annotation). By bringing together authorial philology, digital annotation and literature didactics, the IDEA paradigm - elaborated specifically for this edition but extendable to all copyrighted texts - aims to overcome the objective limitations of publishing contemporary literature content on the Web in open format. The contribution then describes the portal «The Variants of the Rose», a virtual place where to present the critical apparatus of variants encoded in XML-TEI (and visualized through EVT2), some didactic content elaborated

with TRAViz, Storymap JS and Timeline JS, the annotation in a *Domain Specific Language* developed on Euporia and further critical insights elaborated from Franco Moretti's *distant reading* and Giuseppe Savoca's literary lexicography. Subsequently, a reflection on the modelling of the edition, the concept of *extended edition* and the project workflow is presented. Finally, a road map of future content is proposed with a view on the increasing integration of scientific research and Public Humanities.

1. Introduzione

Negli ultimi anni la rivoluzione digitale ha avuto un notevole impatto sulla cultura umanistica comportando un considerevole sforzo teorico volto a definire caratteristiche essenziali e limiti di quella disciplina che è stata chiamata nel tempo *Humanities Computing* e *Digital Humanities* ([64]) o, in ambito italiano, Informatica umanistica e Umanistica digitale ([10]), privilegiando nella sua definizione, di volta in volta, l'aspetto strumentale o quello teorico ed epistemologico. In particolare, nell'ambito delle applicazioni informatiche alla letteratura e alla filologia, la comunità italiana tra gli anni Ottanta e Duemila si è profondamente interrogata sugli aspetti semiotici, formali e rappresentazionali [26]; [44]; [7]), mentre dagli anni Duemila è diventata sempre più centrale una teoria fondazionale delle edizioni digitali ([45]; [39]), rinominate edizioni scientifiche digitali ([36]) in linea con il termine inglese *Digital Scholarly Editions* (DSE). Dalla centralità del paradigma digitale ([57]), all'importanza dell'aspetto collaborativo ([4]), passando per il ruolo della codifica del testo ([52]), del codice ([65]), dei diversi livelli di rappresentazione del testo ([38]), delle interfacce ([3]), dell'apparato digitale ([8]), delle ontologie e del web semantico ([11]; [17]; [14]), si è riusciti a illuminare le diverse sfaccettature di un oggetto culturale che era – e lo è tutt'ora – in divenire. La riflessione teorica ha fatto dei significativi passi in avanti, ma come ogni disciplina emergente deve scontrarsi con una continua ridefinizione dei suoi confini e lo statuto delle edizioni scientifiche digitali risulta ancor oggi problematico e aperto: anche per questo nel marzo 2022, durante l'ottava conferenza annuale dell'associazione “Digital Humanities im deutschsprachigen Raum” è stato proposto il *Manifest für digitale Editionen*, che di recente è stato tradotto in italiano ([12]). Il *Manifesto per le edizioni scientifiche digitali* rappresenta quindi l'occasione ideale per guardare indietro e fare un bilancio provvisorio, individuandone i presupposti metodologici, le dimensioni organizzative, fattuali e sociali, in un quadro teorico che vede come oggetto delle DSE ogni espressione culturale considerata come tale e come destinatari non solo la comunità scientifica ma un pubblico più vasto di lettori e appassionati. Entrando nel merito del *Manifesto*, uno degli aspetti di maggiore interesse sta nell'ampliamento del concetto di edizione rispetto alle tradizionali edizioni a stampa. Infatti si legge:

Le nuove forme di edizione vanno quindi ben oltre le consuete edizioni a stampa. Non solo le superano per ampiezza e profondità di modellizzazione, ma rendono le decisioni editoriali più trasparenti e permettono usi che vanno al di là della semplice ricezione per fini di lettura. Esse rendono possibile l'aggiunta di ulteriori materiali di ricerca e la

realizzazione di forme multiple di riproduzione a differenza dei metodi tradizionali, che sono orientati verso una sola modalità di presentazione ([12]: 104-105).

È in questo contesto – in linea con le proposte del *Manifesto* – che si sviluppa *Le varianti della rosa*, un prototipo di edizione scientifica digitale 'integrata ed estesa' de *Il nome della rosa* di Umberto Eco che vada al di là «della semplice ricezione per fini di lettura» facendo dell'«aggiunta di ulteriori materiali di ricerca» la sua stessa ragion d'essere.

Una breve storia del progetto¹ vede *Le varianti della rosa* nascere nel 2019 come luogo di sintesi tra studi filologici tradizionali e *Digital Humanities* (DH) per studiare le varianti d'autore del primo romanzo di Umberto Eco. Infatti, il semiologo, critico della cultura e scrittore alessandrino dopo più di trent'anni dalla prima edizione de *Il nome della rosa* ([20]), nel 2012, interviene sul proprio testo per eliminare alcune ripetizioni, alcune sviste, modificando la descrizione del bibliotecario e traducendo inoltre alcuni passi latini ([21]: 617-619; [54]: 90), parlando di «piccoli aggiustamenti di poco conto» ([22]). In realtà, dall'analisi sviluppata nella prima fase del progetto, risulta chiaro come gli interventi correttori riguardino circa 6500 parole e rappresentino una spinta alla semplificazione, sebbene limitata e apparentemente asistemica ([16]). *Le varianti della rosa* si proponeva – e si propone – di rispondere a tali quesiti (*Il nome della rosa è stato semplificato? Perché? Qual è stato il processo correttorio di Umberto Eco?*) sfruttando le innovazioni teoriche e metodologiche delle *Digital Humanities*.

Ciò che era necessario fare, sin dalle prime fasi del progetto, consisteva quindi nell'ipotizzare un modello di edizione scientifica digitale per un'opera del tardo Novecento che non cercasse di nascondere il problema del *copyright*² – vero e proprio limite di ogni studio critico e filologico,

-
- 1 Il progetto è iniziato nel marzo 2019 all'interno del Master di *Digital humanities* offerto dall'Università Ca' Foscari in collaborazione con l'Istituto di Linguistica Computazionale (CNR-ILC) e il CoPhiLab. Nel gennaio 2020 è stata pubblicata la prima versione del portale, nel quale si presentava il progetto e si offriva una selezione di varianti codificate in XML-TEI e visualizzate in EVT2. Con la collaborazione del Centro di Informatica Umanistica di Catania (CINUM), nel quadro di un progetto di ricerca all'interno del PhD in Scienze dell'Interpretazione, è stato sviluppato il paradigma IDEA e sono stati implementati contenuti di didattica digitale, di *public humanities* e di lessicografia letteraria, nel corso del triennio 2020-2022. È stata quindi pubblicata una seconda versione del portale nel gennaio del 2022 in linea con i risultati acquisiti e si prevede, entro la fine del 2023, la conclusione del progetto con la pubblicazione dell'edizione critica integrale delle varianti, di un database XML interrogabile, dell'annotazione realizzata su Euporia e di nuovi contenuti didattici e multimediali.
 - 2 La Legge sul diritto d'autore (L. 633/1941) chiarisce come il diritto esclusivo alla pubblicazione e all'utilizzazione economica di un'opera di ingegno creativo spetti all'autore e ai suoi eredi (per la durata di settant'anni dalla morte dell'autore). L'art.70 precisa però che: «Il riassunto, la citazione o la riproduzione di brani o di parti di opera e la loro comunicazione al pubblico sono liberi se effettuati per uso di critica o di discussione, nei limiti giustificati da tali fini e purché non costituiscano

specialmente digitale, applicato alla contemporaneità – ma che si ponesse, per così dire, *in absentia* del testo, come luogo virtuale di espansione dei confini dell’opera letteraria, attraverso analisi e approfondimenti, uno spazio che facesse dell’integrazione tra annotazione, *distant reading*, concordanza, codifica [XML-TEI](#) ([63]), filologia digitale, e *data visualization* il proprio paradigma, proponendo un modo integrato di studiare e fruire l’opera letteraria, in particolare quelle ancora oggetto di diritto d’autore.

2. Modellizzazione e il paradigma ‘IDEA’

La prima fase di ogni progetto di edizione (e di ogni lavoro scientifico) consiste nella modellizzazione, che Mancinelli e Pierazzo sintetizzano nel seguente modo:

Per modellizzazione si intende la fase preliminare di un lavoro scientifico, editoriale e non, durante la quale il ricercatore considera l’oggetto del proprio lavoro (il testo da pubblicare, per esempio), sceglie le caratteristiche essenziali e quelle opzionali per il lavoro di ricerca (i testimoni, i fenomeni testuali, gli errori) e scarta quelle che invece non sono considerate utili (i codici descritti o prodotti dopo una certa data, le varianti formali) ([36]: 48).

Si tratta dunque dello studio approfondito del proprio oggetto di ricerca al fine di individuarne gli aspetti fondamentali – come possono essere le varianti sostanziali o, più in generale, i testimoni rilevanti – in base al proprio orizzonte di ricerca, quale ad esempio la tipologia di edizione (critica, diplomatica, semidiplomatica, con commento) che si vuole allestire. Accanto a ciò – segnalano Ciula, Eide, Marras e Sahle in un volume dedicato alla modellizzazione come strategia di «*Thinking in practice*» ([13]) – bisogna sottolinearne la funzione euristica: il *modelling* è il momento privilegiato dell’indagine, dell’esplorazione, della scoperta attraverso formalismi matematici, diagrammi, *brainstorming*, metafore, in un processo non lineare, ma iterativo con ripetuti *loop* di test, feedback e revisioni ([13]: 8). Infatti, come ha mostrato Sahle, il modello può essere un diagramma, un’immagine, una narrazione, una formula, un’annotazione formalizzata, può essere fatto di parole, metallo o perfino di «carne e sangue». Si tratta allora di definire quale sia il punto di convergenza di tutti questi modelli: «we use models to refer to something, to talk about something else, to show, to simulate something else. A model stands for something [...] for understanding, communication, and exchange. For testing, analysing, producing something» ([13]: 187). In questa prospettiva un modello *sta* per qualcos’altro e la sua caratteristica principale è quella di non essere in scala 1:1, ma *più piccolo, meno dettagliato, più astratto*, rispetto all’oggetto o al dominio che sta modellando. Non solo. McCarty chiarisce come ogni modello viva in una continua tensione tra l’essere *modello di* qualcosa (*model of*) ed essere *modello per* qualcosa (*model for*), dove nel primo caso si tratta di descrivere il proprio oggetto da un determinato punto di vista, mentre nel secondo caso serve a sviluppare qualcosa di nuovo

concorrenza all’utilizzazione economica dell’opera; se effettuati a fini di insegnamento o di ricerca scientifica l’utilizzo deve inoltre avvenire per finalità illustrative e per fini non commerciali».

([13]: 38). In tal senso la modellizzazione (e la sua revisione ciclica in ogni fase del progetto) dell'edizione scientifica digitale de *Il nome della rosa* è partita dalla tipica modellizzazione del filologo per superarla verso una pratica di *modelling* più ampia (*model for*), con l'elaborazione di un paradigma di edizione che potesse rispondere alle esigenze del testo preso in esame. Infatti, come ben evidenziato da Mancinelli e Pierazzo:

Ogni edizione rappresenta essa stessa un modello di un testo in quanto: 1. è basata su una selezione dei testimoni esistenti; 2. incarna il punto di vista scientifico di un editore; 3. presenta il testo secondo modelli editoriali standardizzati: edizione critica, diplomatica, semidiplomatica ecc.; 4. presenta una selezione del "fatti del testo": alcune varianti in apparato, scelta di un testo base, di una facie linguistica ecc. ([36]: 49)

Se questa analisi è ineccepibile per una edizione critica tradizionale, nel caso di un testo novecentesco ancora tutelato dal diritto d'autore la questione si fa più problematica per evidenti limiti intrinseci legati al *copyright*, dal momento che il testo non può essere pubblicato integralmente. A questo punto le strade sono due: arrendersi e non produrre edizioni di testi novecenteschi senza l'accordo delle case editrici e degli eredi, oppure cercare una soluzione differente utilizzando il digitale e il web come spazio per un discorso alternativo attraverso il quale arricchire e approfondire il testo.³ Se si sceglie questa seconda strada, bisogna modellare una realtà nuova aprendosi – nel caso delle *Varianti della rosa* – a un approccio interdisciplinare che nasca dalla sintesi tra critica computazionale della letteratura ([9]), filologia digitale ([32]), didattica della letteratura digitale ([29]) ed ermeneutica ([30]), seguendo l'intuizione che intende trasportare nell'ambito delle DH un concetto caro alla narratologia contemporanea: la narrazione estesa. Essa, infatti, caratterizza quelle narrazioni che «eccedono la misura che convenzionalmente viene attribuita al singolo racconto, costruendo mondi e personaggi segnati da coerenza e continuità, che si estendono su più piattaforme medialì, e rispondono agli interventi dei propri fruitori in modi e forme variabili» ([50]: 19). Tali narrazioni si estendono da un punto di vista intratestuale, ma anche da uno extratestuale «attraverso la produzione di informazioni collaterali eppure importanti» (*ibid.*) inglobando dentro lo stesso universo narrativo anche le fonti secondarie prodotte dai fruitori. Fuor di metafora, per superare l'ostacolo dell'indisponibilità del testo integrale, si è scelto di progettare un'edizione che presentasse solo l'apparato critico delle varianti (con un minimo di contesto di riferimento), offrendo ai lettori una panoramica complessiva delle correzioni d'autore (collegando le varianti al numero di pagina dell'edizione a stampa di riferimento) aggiungendo una serie di approfondimenti critici come l'annotazione

3 Ciò non significa che non si possa (e in parte si debba) cercare una soluzione di sintesi tra edizioni estese ed edizioni tradizionali immaginando delle espansioni *online* proposte dagli editori che detengono i diritti di pubblicazione dell'opera. Un esempio di questa possibilità è offerto dall'Edizione Nazionale dell'Opera Omnia di Luigi Pirandello, dove il portale *web* diventa il luogo dell'espansione dei contenuti cartacei ([27]). Un'ulteriore forma di integrazione è testimoniata dalla manualistica per le scuole superiori che propone nella maggior parte dei casi delle espansioni online.

delle varianti, un database XML, unità didattiche sulle varianti, analisi tipiche del *distant reading* in un'ottica concordanziale, unendo approcci diversi per approdare al modello di edizione scientifica digitale estesa (Figura 1).

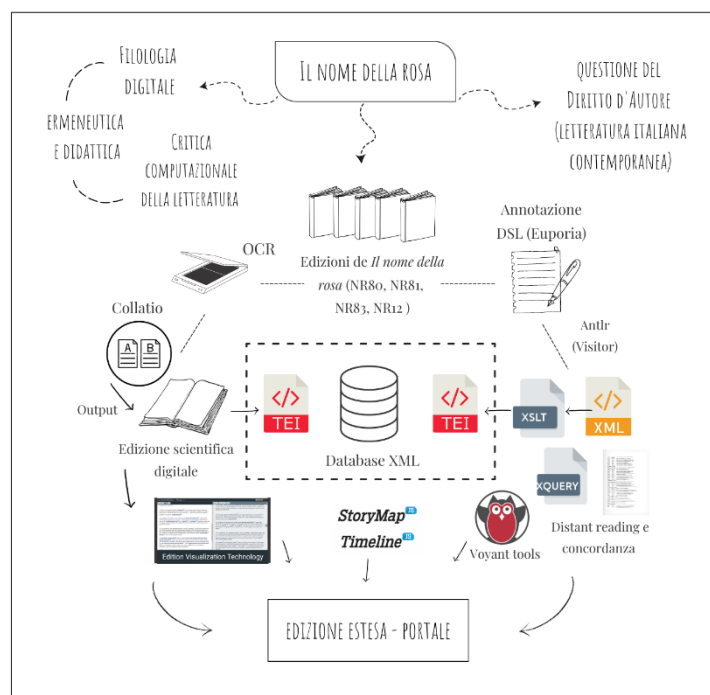


Figura 1: Modellizzazione dell'edizione scientifica digitale estesa (Fase III)

In questo senso l'edizione digitale non può e non deve sostituire l'edizione cartacea, ma le si deve accostare, ampliandola ed estendendola, in una forma ibrida tra l'approfondimento critico per specialisti e il commento per studenti e lettori.

L'utilizzo di approcci e metodologie diverse, lungi dall'essere un *vulnus* dell'edizione, diventa la stessa essenza di un nuovo paradigma di edizione, sviluppato *ad hoc* per il presente progetto: il paradigma IDEA (dalle iniziali di «Interpretazione», «Didattica», «Edizione», «Annotazione») che intende mettere al centro l'Interpretazione come atto fondante di ogni operazione critica (sia tradizionale che digitale) attraverso domande di ricerca che facciano da guida al processo di elaborazione dell'edizione. Infatti, secondo la famosa definizione di Sahle, «Scholarly digital editions are scholarly editions that are guided by a digital paradigm in their theory, method and practice» ([57]: 28) e il paradigma IDEA nasce *nel* digitale e *per* il digitale: da un lato nell'elaborazione e fruizione di contenuti digitali; dall'altro, nella progettazione che tiene conto delle specificità del digitale, ovvero la possibilità di presentare grandi quantità di dati (dall'apparato critico all'annotazione passando per la concordanza) in modo multimediale,

favorendo la *user interaction*; specificità che investono le coordinate del tempo (testo diacronico), dello spazio (tutti i testi), della forma (il testo iconico) e del *social* (il testo wiki) ([4]).

Le quattro diverse fasi del paradigma IDEA (Figura 2) non devono essere però viste come consequenziali, ma in un'ottica rizomatica e, sostanzialmente, ciclica, dove ogni fase presuppone le altre e le completa. Ad esempio, l'interpretazione fonda l'annotazione e poi i risultati di quest'ultima fanno emergere una nuova interpretazione, riprendendo in chiave digitale l'idea del circolo ermeneutico (cfr. par. 4.2) come atto fondamentale della comprensione. L'Edizione inoltre non deve essere vista soltanto come frutto di una fase specifica, prodotta a partire da una determinata codifica in XML-TEI, con una specifica visualizzazione del testo attraverso l'interfaccia scelta, ma è anche l'esito finale di tutto il paradigma, secondo un'ottica di edizione integrata ed estesa, dove per integrata ci si riferisce alla profonda connessione tra didattica, filologia, ermeneutica e DH, e con estesa si riprende, come già detto, il concetto di narrazione estesa. Così il paradigma IDEA mira all'allestimento di un portale-edizione dove far dialogare strumenti, contenuti e discipline diverse.

PARADIGMA 'IDEA' (EDIZIONE ESTESA)



Figura 2: Modellizzazione del paradigma 'IDEA' (Interpretazione-Didattica-Edizione-Annotazione)

3. Workflow dell'edizione estesa

L'edizione estesa, di cui il paradigma IDEA è l'orizzonte teorico di riferimento, è stata sviluppata seguendo diversi cicli di revisione riconducibili a tre fasi distinte: nella *Fase I* sono stati sviluppati il prototipo dell'edizione delle varianti tramite [EVT](#) ([56]) e una prima annotazione con un *Domain Specific Language* (cfr. par. 4.5) così da valutare la semplificazione linguistica dell'edizione del 2012; nella *Fase II* sono stati progettati alcuni contenuti *public* e didattici rivolti

a un pubblico di studenti, insegnanti e appassionati tramite proposte di *digital storytelling* ([51]) ed è stata inoltre intrapresa una sostanziale revisione del *front-end*; nella *Fase III* si è invece ripensata l'intera edizione, ampliando il corpus dei testimoni presi in esame, vocabolarizzando l'intera opera per analizzarne il lessico in maniera complessiva, revisionando infine l'annotazione per definire in maniera più chiara il processo correttorio di Eco. Le tre fasi sono state a loro volta suddivise in un numero variabile di attività come mostrato dal seguente schema riassuntivo (Tabella 1):

Fase	Attività	Descrizione	Programmi e Linguaggi
I	Modellizzazione	Elaborazione della domanda di ricerca (« <i>Il nome della rosa</i> è stato semplificato nel 2012?») e di un'ipotesi ermeneutica a partire da un'analisi quantitativa dei dati. Costruzione di un modello di edizione in base alle specificità del testo e agli obiettivi scientifici e didattici.	Voyant Tools , Python (NLTK), Canva , XMind
	Acquisizione e <i>Collatio</i> Edizioni NR80-NR12	Acquisizione attraverso algoritmi di OCR delle due edizioni principali (NR80, nella ristampa del 2010, e NR12), collazione automatica tramite Juxta Commons e generazione di un output codificato in TEI.	Tesseract , FineReader , Juxta Commons ⁴
	Codifica del testo	Revisione della codifica del testo secondo lo schema di codifica TEI per renderla <i>EVT compliant</i> . Correzione di eventuali errori (soprattutto trasposizioni).	XML-TEI, XSLT
	Prototipo Edizione delle Varianti	Configurazione dell'interfaccia di visualizzazione e realizzazione del prototipo attraverso EVT2 beta.	EVT2
	Annotazione ⁵	Progettazione dello schema di annotazione con un <i>Domain Specific Language</i> per validare	<i>Domain Specific Language</i> ,

4 In data 30 settembre 2022 *Juxta* non risulta più accessibile.

5 In questo contributo si usano i termini codifica e annotazione per indicare lo stesso processo di rappresentazione digitale del testo secondo due sfumature diverse: si usa codifica in riferimento alla rappresentazione degli aspetti essenziali del testo (divisioni strutturali, tag per l'apparato filologico, marcatura delle entità nominate) mentre con annotazione ci si riferisce alla formulazione di annotazioni linguistiche e di postille critiche di vario genere. La scelta lessicale implica un preciso processo di differenziazione: laddove bisogna rappresentare fenomeni strutturali del testo già altamente formalizzati in TEI con l'obiettivo di condividerli con la comunità (e visualizzarli attraverso

		l'ipotesi ermeneutica. Utilizzo del <i>tool</i> d'annotazione Euporia.	CollateX , ANTLR , Euporia
	Analisi e Interpretazione dei risultati	Analisi e interpretazione dei dati emersi dall'annotazione. Sviluppo di grafici nell'ottica della <i>data visualization</i> e del <i>distant reading</i> . Risposta alla domanda di ricerca iniziale.	XPath , XQuery , Gephi , Tableau
	Prototipo portale web e Pubblicazione	Allestimento del prototipo del portale con l'inserimento dell'edizione digitale delle varianti (EVT) e di alcuni approfondimenti critici. Raccolta di <i>feedback</i> in funzione dello <i>User-centered design</i> .	HTML, CSS
II	Ri-Modellizzazione	Rielaborazione del modello di edizione secondo il paradigma IDEA con la progettazione di un'edizione estesa e integrata.	Canva, XMind
	<i>Digital Storytelling</i>	Ideazione e sviluppo di contenuti didattici e <i>public</i> di stampa filologico (TRAViz e infografiche) e critico-storico (con le applicazioni di Knightlab)	TRAViz , StoryMap JS , TimeLine JS , Gephi
	Revisione, progettazione e pubblicazione <i>Varianti della rosa 2.0</i>	Progettazione, sviluppo e pubblicazione di un portale con una <i>User Interface</i> aggiornata, in seguito alla valutazione dei <i>feedback</i> del prototipo. Trasmigrazione in un sistema basato su un <i>Content Management System</i> (CMS) svincolando lo sviluppatore dalla gestione dell'intero sistema, con l'obiettivo di focalizzarsi esclusivamente sui contenuti, aprendo anche alla collaborazione di eventuali <i>content creator</i> esterni.	Wordpress

EVT) si è scelto di marcare tali aspetti in TEI sin dall'inizio; per rappresentare invece fenomeni meno formalizzati, fuzzy, talvolta persino tangenti al testo, ma utili al ricercatore nello sviluppo di un modello fedele alla propria domanda di ricerca e all'oggetto di studio, si è scelto di annotare costruendo un DSL per giungere allo standard TEI solo nel momento finale del paradigma. Per una riflessione sistematica sull'annotazione si rimanda a [43].

	<i>Crowdsourcing e Questionario User Interface</i>	Elaborazione di questionari per ottenere <i>feedback</i> immediati e continuativi nel tempo; implementazione di semplici moduli per coinvolgere il pubblico con la possibilità di inviare interpretazioni e contenuti.	Google Forms, Wordpress
III	Ri-Modellizzazione	Nuovi cicli di revisione: riformulazione delle domande di ricerca («Qual è il processo correttorio di Eco?», «Vi sono varianti sostanziali che precedono l'edizione del 2012?», «Qual è la specificità del lessico echiano e come cambia nel tempo?») a partire dall'ampliamento del corpus dei testimoni presi in esame e da ulteriori analisi quantitative.	Voyant Tools, Python (NLTK), Canva
	<i>Recensio, Acquisizione, Collatio</i> di tutte le edizioni notevoli	Acquisizione attraverso algoritmi di OCR di tutte le edizioni che presentano varianti (in particolare da NR80, prima edizione, a NR83, oltre la quale non sono state individuate varianti ad eccezione della revisione integrale effettuata in NR12), collazione automatica tramite VarianceViewer e generazione di un output codificato in XML-TEI.	Tesseract, FineReader, VarianceViewer
	Revisione e implementazione codifica del testo	Revisione della codifica del testo con l'aggiunta delle varianti dei testimoni presi in esame (NR80-NR83).	XML-TEI, XSLT
	Revisione e Implementazione Annotazione DSL	Progettazione di una tassonomia variantistica e revisione del DSL di annotazione. Trasformazione della codifica XML proprietaria (generata a partire dall'annotazione su Euporia) in un formato TEI <i>compliant</i> attraverso fogli di stile.	Euporia, XSLT
	Edizione Scientifica Digitale delle Varianti	Integrazione della DSE con la tassonomia implementata nell'annotazione DSL. Configurazione dell'interfaccia di visualizzazione attraverso EVT2.	EVT2

Database delle Varianti	Sviluppo del database XML comprendente la codifica di tutte le varianti (DSE) e l'output XML-TEI dell'annotazione DSL.	BaseX
Lemmatizzazione integrale e Vocabolario de <i>Il nome della rosa</i>	Progettazione ed elaborazione del vocabolario de <i>Il nome della rosa</i> con la lemmatizzazione integrale di NR83 e, in aggiunta, la lemmatizzazione delle forme uniche presenti nelle altre redazioni.	4D DBMS (LiotroConcord)
Analisi e Interpretazione dei Risultati	Analisi e interpretazione dei dati emersi dal vocabolario, dall'annotazione revisionata, dalla DSE. Sviluppo di approfondimenti critici e di grafici (<i>data visualization</i>). Risposta alle domande di ricerca iniziali.	XPath, XQuery, Python (NLTK), Voyant Tools
Pubblicazione <i>Open access</i> dell'Edizione Scientifica digitale e dell'annotazione	Rilascio dell'edizione definitiva e di tutti i dati in <i>open access</i> . In accordo coi principi FAIR ([66]) tutti i contenuti saranno rintracciabili, accessibili, interoperabili e riusabili e, in particolare, la codifica della DSE in TEI e l'annotazione DSL (esportata in TEI) saranno presenti in repository come TAPAS, GitHub e Zenodo con licenza CC BY-NC-SA.	Portale web, Tapas , Zenodo , GitHub

Tabella 1: Workflow del Progetto

4. Dal paradigma 'IDEA' all'applicazione

4.2 *Recensio, collatio e Database delle varianti*

La prima parte del lavoro è stata caratterizzata, in linea col metodo di Lachmann ([35]) dalla *recensio*, ovvero dall'individuazione di tutti i testimoni considerati rilevanti ai fini dell'allestimento dell'edizione. Se nella *Fase I*, in base alla domanda di ricerca sull'eventuale semplificazione dell'edizione Dodici, i testimoni presi in esame sono stati i testi delle due redazioni maggiori (NR80 e NR12) scegliendo, in particolare, per la prima edizione il testo della ristampa del 2010, nella *Fase III* si è scelto invece di estendere lo studio a tutte le altre redazioni con varianti (anche minime) prendendo così in considerazione dodici edizioni,⁶ di cui soltanto

⁶ NR80a (I edizione, settembre 1980), NR80b (II edizione, ottobre 1980), NR81a (III edizione, marzo 1981), NR81b (IV edizione, maggio 1981), NR81c (V edizione, luglio 1981), NR82 (IX edizione, agosto 1982), NR83a (X edizione, gennaio 1983), NR83b (XII edizione, ottobre 1983), NR84 (I

quattro significative: NR80 (I edizione, settembre 1980), NR81 (IV edizione, maggio 1981), NR83 (XII edizione, ottobre 1983) e NR12 (I edizione riveduta e corretta, gennaio 2012).

Una volta individuati i testimoni, essi sono stati acquisiti attraverso sistemi di *Optical Character Recognition* (OCR): in un primo momento ci si è affidati al software *open source* Tesseract, usato da riga di comando, per poi passare ad *Abby Fine Reader* – per la necessità di un ambiente che permettesse di revisionare il testo contestualmente all’acquisizione, affiancando alla trascrizione l’immagine ad alta risoluzione (Figura 3). La qualità delle immagini ha reso il lavoro di riconoscimento piuttosto agevole, sebbene il plurilinguismo echiano (con interi passi in latino o nel *pastiche* linguistico di Salvatore) abbia comportato una delicata fase di revisione.

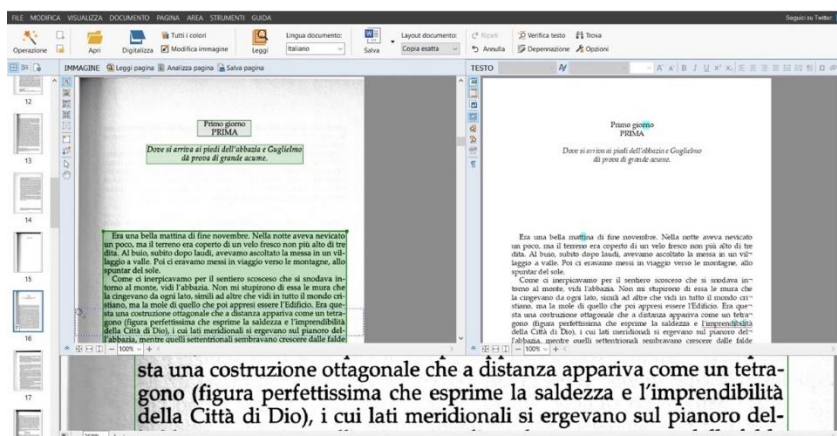


Figura 3: Schermata del software di OCR (Abby Fine Reader) con NR80a

Successivamente si è proceduto alla collazione automatica attraverso *Juxta commons* e *CollateX*: il primo è stato utilizzato nella *Fase I* per produrre l’edizione TEI *compliant* da adattare allo schema di codifica previsto da EVT, mentre il secondo era parte integrante del *tool* Euporia. A causa dello spegnimento dei server di *Juxta commons* per la fine del supporto da parte di Performant Software, nella *Fase III* si è deciso di passare a un software di collazione sviluppato dall’università di Würzburg, *Variance Viewer*,⁷ che usa lo stesso algoritmo di collazione di *Juxta* (*java-diff-utils*), ed è dotato di un’interfaccia grafica intuitiva (Figura 4). *Variance Viewer* permette di esportare i dati in formato TEI (similmente a *Juxta*, con la necessità di adattare le poche differenze, come gli attributi di `<rdg>`). Inoltre, *Variance Viewer* classifica le varianti in base al loro aspetto linguistico (dalla separazione di due parole, alle varianti con una singola differenza, passando per

edizione grandi tascabili, aprile 1984), NR90 (I edizione Bompiani per la scuola con note e introduzione di Costantino Marmo, 1990), NR80\10 (LVII edizione, marzo 2010) e NR12 (I edizione riveduta e corretta, gennaio 2012).

⁷ Per un’analisi di alcuni sistemi di collazione si veda [55].

la punteggiatura) e produce delle statistiche utili ad avere un quadro generale delle differenze tra le varie edizioni.

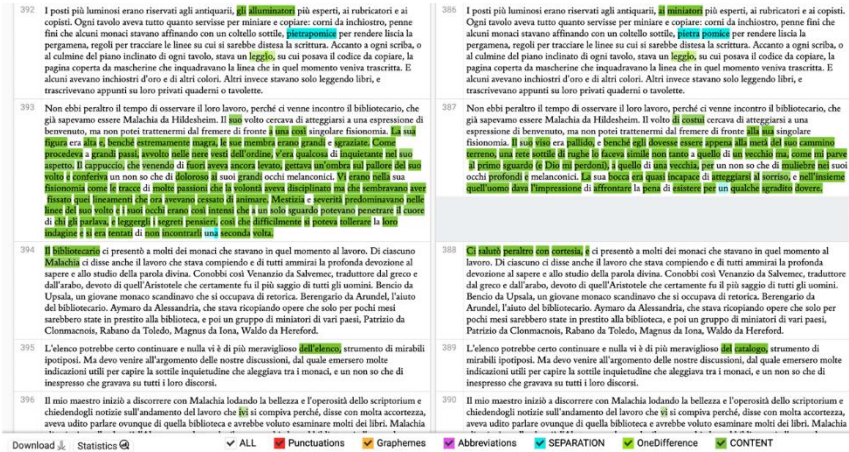


Figura 4: Variance Viewer

In questo modo è stato allestito il corpus delle varianti, scorporandole dal testo e mantenendo un minimo di contesto (di solito il paragrafo, o per i paragrafi più lunghi la frase). Il corpus è poi confluito nell'edizione scientifica digitale attraverso la codifica dell'apparato critico (esplicitando con l'attributo @wit del tag <rgd> soltanto i testimoni che presentano una lezione differente) e infine nel Database XML, sviluppato con BaseX (Figura 5) comprendente anche l'annotazione prodotta su Euporia (esportata anch'essa in TEI).

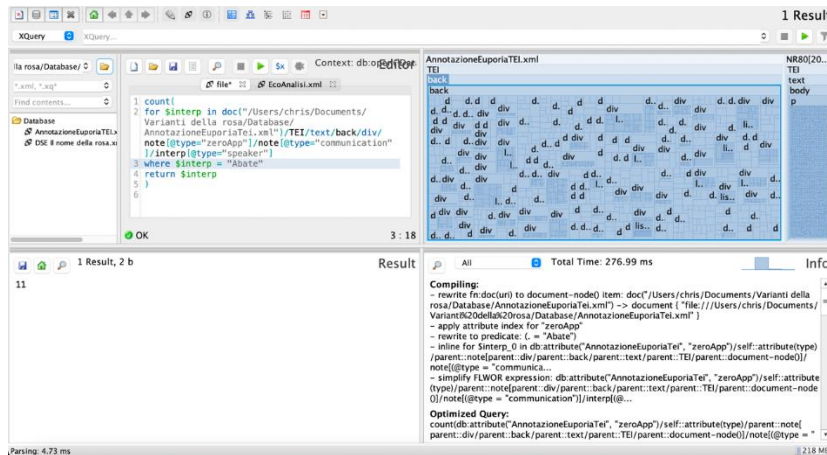


Figura 5: Database XML delle varianti e dell'annotazione (BaseX)

4.2 Interpretazione, distant reading e vocabolario

Allestire un'edizione, scegliere i testimoni, codificarli, modellarne il dominio, annotarlo, proporre approfondimenti critici e unità didattiche significa, in ultima istanza, interpretare il testo e rappresentarlo da un determinato punto di vista. Con l'avvento del digitale tale aspetto ermeneutico è stato talvolta evidenziato, principalmente quando il filologo e il critico hanno reso disponibili tutti i dati della loro ricerca, esplicitando così la propria interpretazione. Ma altrettanto spesso è stato sottaciuto, come discusso da Franco Moretti in *Falso movimento*, nel quale ha criticato fortemente la svolta quantitativa delle DH, giungendo alla conclusione che una vera e propria sintesi tra ermeneutica e quantificazione non sia in alcun modo possibile ([41]: 36). Negli ultimi anni, in realtà, diverse proposte sono state di segno opposto come, ad esempio, Andrew Piper che in *Enumerations* ([53]) pensa a un dialogo tra *close* e *distant reading* nell'elaborazione di un circolo ermeneutico digitale (Figura 6) che rinnovi quello tradizionale di Schleiermacher – secondo cui l'interpretazione deve prodursi dal continuo confronto tra la parte e il tutto ([60]) – e il circolo heideggeriano-gadameriano – che vede l'interpretazione emergere sullo sfondo della pre-comprensione, in una fusione di orizzonti dove diventa centrale l'applicazione alla situazione concreta dell'interprete ([31]; [25]). Andando oltre Piper, nel contesto delle DH il circolo ermeneutico avrebbe quindi come momenti essenziali la diade *close* (parte) e *distant* (tutto) *reading* attraverso la modellizzazione e rimodellizzazione del dominio dove la misurazione fornisce i dati (la 'cosa stessa') all'interpretazione che a sua volta dà l'orizzonte (pre-comprensione) a una nuova misurazione.

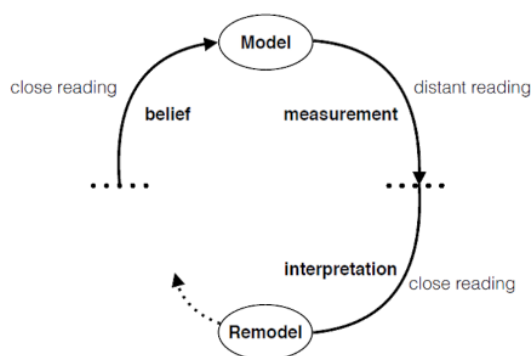


Figura 6: Circolo ermeneutico digitale ([53]: 10)

Moretti, in realtà, non ha mai negato *tout court* la possibilità che *close* e *distant reading* dialogassero tra loro, offrendosi nuovi oggetti di lavoro e innescando nuove domande, bensì, si è opposto fermamente alla loro piena collaborazione: «L'interpretazione *trasforma* tutto ciò che tocca: “questo, significa quello”. Un assoluto rispetto per i dati è il punto d'onore del lavoro quantitativo. Sono impulsi antitetici» ([41]: 36). Ciò sembra in parte contrapporsi

all'ermeneutica heideggeriano-gadameriana che parte dal presupposto che il circolo ermeneutico non è in alcun modo *vitiosus* perché l'interprete ha il compito di non lasciarsi imporre la precomprensione da pensieri estemporanei e da concetti popolari, facendo invece emergere l'interpretazione dalla 'cosa stessa' ([31]: 220-222). In questo senso si potrebbe parlare di una forma di 'ascolto ostinato del testo', o meglio, di un 'ascolto ostinato dei dati' che realizzi la sintesi tra ermeneutica e quantificazione.

Nel contesto del paradigma IDEA il circolo ermeneutico digitale si è quindi manifestato nell'interpretazione iniziale fondata su un'ipotesi (ricavata dalla lettura ravvicinata del testo e dalla letteratura secondaria) integrata, trasformata e perfino a tratti confutata dai dati emersi dal *distant reading* ([40]) e dal metodo concordanziale ([59]). Infatti, l'ipotesi ermeneutica iniziale – l'eventuale semplificazione linguistica del *Nome della rosa* – ha prodotto un'annotazione su Euporia con le marche d'uso di Tullio De Mauro ([18]) e attraverso una lettura quantitativa di questi dati è stato possibile dimostrare come ci sia stata una forma di semplificazione del lessico nelle correzioni echiane. Le due edizioni contano infatti, rispettivamente, 174047 e 169661 parole, con una riduzione di 4386 parole tra la prima e la seconda edizione (il 2.5 % della prima) e una variazione che investe 6688 parole (tra cancellazioni, sostituzioni e aggiunte). Le parole marcate ad alto uso («FO», «CO», «AU», «AD») nell'edizione del 2012 sono in un rapporto di 1:5 rispetto al corpus totale di quella marca d'uso (4 parole su 5 fanno parte dell'edizione del 1980), rapporto che diventa di 1:10 per le parole appartenenti al lessico a basso uso («TS», «LE», «BU», «OB»).⁸ Ciò conferma la presenza di una significativa riduzione del lessico nell'edizione del Dodici e contestualmente evidenzia una forma limitata di semplificazione che per essere valutata nel suo complesso deve essere messa in relazione al lessico dell'intera opera attraverso le concordanze.⁹

La concordanza è stata allestita utilizzando *LiotroConcord* – un software basato sul DBMS 4D in continuità ideale con l'applicativo di Giuseppe Savoca per IBM AS/400 – e realizzato dall'équipe di ricerca del CINUM di Catania ([19]; [27]). Il software associa ad ogni forma un lemma specifico, seguendo l'impostazione di Savoca (a sua volta modellata su quella di Esperti [23]) che comporta, tra le altre, le seguenti scelte lessicografiche:

- I verbi «essere» e «avere» vengono distinti in ausiliari o predicativi;
- gli aggettivi e i pronomi vengono lemmatizzati insieme, ponendo l'accento non sulla categoria grammaticale ma sulla loro funzione (sono presenti quindi le seguenti categorie: qualificativi, indefiniti, relativi, interrogativi-esclamativi, personali, possessivi);

8 Le categorie di De Mauro usate nell'annotazione sono: Fondamentale (FO), Comune (CO), Alto Uso (AU), Alta Disponibilità (AD), Termine specialistico (TS), Letterario (LE), Basso Uso (BU), Obsoleto (OB).

9 La lettura concordanziale delle varianti permette infatti di indagare in quali contesti viene variata una parola, se tutte le occorrenze di un determinato lemma vengono corrette, se la pratica correttoria è sistematica o rapsodica.

- i titoli di opere vengono uniti in unico lemma con categoria ‘nome proprio’, una scelta che si rivela utile per un testo come *Il nome della rosa*, perché diminuisce il rumore nella concordanza;
- i forestierismi vengono ricondotti alla lingua di appartenenza (*fr*, per i francesismi, *lt* per i latinismi, ecc.) avendo come lemma di riferimento la forma che si trova nel testo;
- Il *che* è distinto in relativo, congiunzione o interrogativo-esclamativo;
- le preposizioni semplici e articolate vengono lemmatizzate insieme e ricondotte alla preposizione semplice;
- i numerali (ordinali e cardinali) e i numeri in cifre sono ricondotti a due diverse categorie;
- le enclitiche vengono separate dai verbi (‘vederlo’ viene lemmatizzato come ‘veder’ e ‘lo’);
- i lemmi omografi vengono disambiguati e associati a due lemmi diversi.

Una volta prodotta la lemmatizzazione automatica, l’interfaccia grafica (Figura 7) permette di modificare occorrenze, forme e lemmi: unendo o dividendo le occorrenze, modificando le forme (ad esempio la lettera iniziale minuscola in maiuscola), associando una forma a un lemma diverso. La lemmatizzazione, che si configura in questo modo come semi-automatica, prevede l’intervento attivo del concordatore, il quale diviene il protagonista della lemmatizzazione, attraverso le sue correzioni, in un lungo processo di fedeltà sia all’idioletto dell’autore che al suo stesso *usus* lessicografico. Lemmatizzare diventa allora il momento aurorale dell’interpretazione in linea con l’antica pratica lessicografica per cui concentrarsi sui singoli lemmi significa entrare in rapporto diretto con la parola del testo, precedendo l’attività critica nell’uso della concordanza ([62]: 87-104).

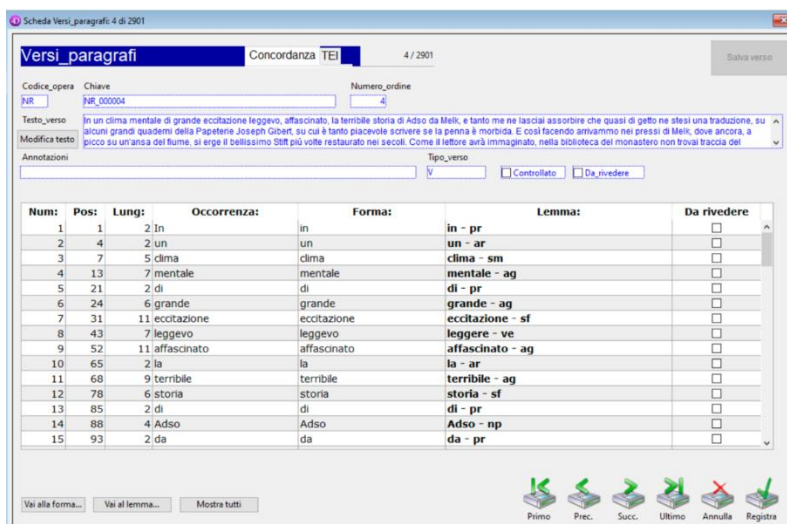


Figura 7: Schermata del software di lemmatizzazione LiotroConcord

4.3 Didattica, storytelling e ottica public

L'edizione estesa de *Il nome della rosa* ha previsto sin dalla *Fase II* l'elaborazione di unità didattiche e approfondimenti critici sulle varianti e sul lessico echiano, attenti a intercettare un pubblico più ampio. Ciò significa proporre una filologia e una critica della letteratura *public*, similmente a ciò che è accaduto per le discipline storiche con la rivoluzione imposta dalla diffusione della *public history* ([58]; [47]). La filologia, infatti, è stata spesso percepita come una disciplina rigorosa, ma fin troppo erudita e lontana dagli interessi di lettori e studenti ([28]). Mostrare in un'ottica *public* il *modus operandi* di uno scrittore, presentando le varianti in un'altra veste rispetto alle tradizionali edizioni critiche, più vicina all'orizzonte multimediale del lettore, rappresenta un tentativo di dare un impulso alle discipline filologiche alimentando l'interesse del pubblico, recuperando così il mandato sociale di una disciplina attenta al dato materiale del testo, alla sua trasmissione, che storicizza ogni interpretazione e si propone come antidoto alla post-verità, valorizzando la dimensione fattuale di ogni processo culturale.

Inoltre, le *Digital Humanities* rappresentano una sfida e un'opportunità anche per la didattica ([46]): facendosi promotrici di un apprendimento attivo e "situato" ([34]), destrutturando il contesto classe e aprendolo allo spazio virtuale dove «il docente come produttore e mediatore di contenuti» ([29]: 90) si serve degli archivi digitali, dell'eccezionale mole di contenuti e di dati reperibili *online* per progettare lezioni che abbiano al centro l'interdisciplinarietà e le esigenze educative dello studente attraverso la fruizione di video o di altri contenuti multimediali. La sfida già abbastanza ardua per le materie scientifiche non è certamente più semplice per coloro che si occupano di letteratura, dal momento che «lo studioso e l'insegnante di letteratura devono sempre attuare un dialogo, e duplice: con il testo e con gli altri lettori del testo» ([37]: 18), avendo sempre come obiettivo la significatività del processo educativo, ancor più oggi che «l'insegnamento della letteratura, a scuola, ma anche nell'odierna università, ha invece destinatari meno rigorosi [rispetto all'accademia], e non necessariamente interessati, ai quali va costantemente motivato il senso dell'oggetto dello studio, giustificata l'importanza dei testi che si chiede di leggere» (*ibid.*).

Per tali motivi si è deciso di proporre all'interno de *Le varianti della rosa* contenuti che valorizzassero la portata *public* della filologia e della letteratura, attraverso una selezione di varianti per lettori occasionali (con infografiche e visualizzazioni *ad hoc*) e tramite unità didattiche.

Nel primo caso, come è possibile vedere in Figura 8, le varianti più significative sono state rielaborate tramite l'uso della libreria web *open-source Text Re-use Alignment Visualization* (TRAViz)¹⁰ ([33]).

10 L'uso di TRAViz è un esperimento volto a rendere gli apparati di varianti maggiormente fruibili per i lettori non specialistici, sebbene la natura delle varianti (2 soli testimoni) ne limiti l'efficacia dal momento che TRAViz ha una visualizzazione più efficace quando sono presenti molti testimoni.

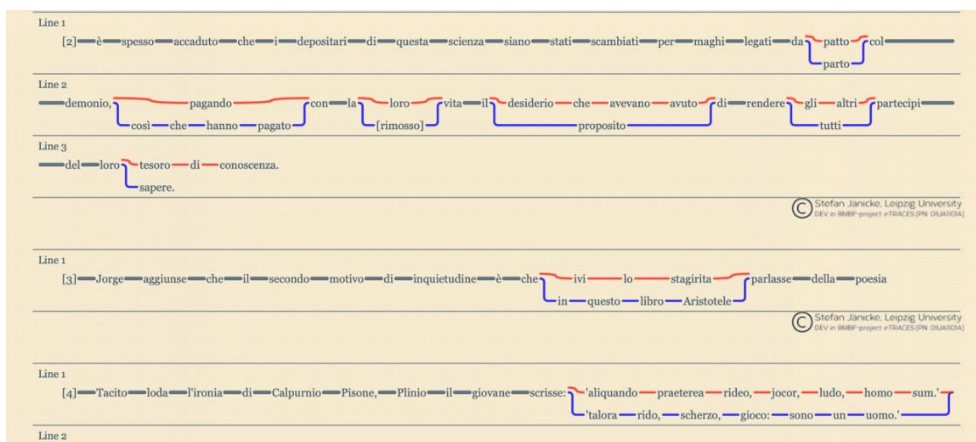


Figura 8: Dettaglio della 'Top 10' realizzata con TRAViz

TRAViz, sviluppato dal *team* di Stefan Janicke sulla base di *Variante Graphs* ([61]) utilizza file codificati in TEI per rappresentare le varianti come fossero sentieri che si biforcano: le parole vengono automaticamente allineate e ciascun testimone viene segnalato da un colore differente, così da migliorare la visualizzazione delle varianti anche per un non specialista.

Per quanto riguarda gli approfondimenti didattici si è invece progettata una unità didattica sulla storia dell'eresia, poiché l'annotazione ha mostrato come il contesto della maggior parte delle varianti riguardi proprio la trattazione dell'eresia (Figura 9).

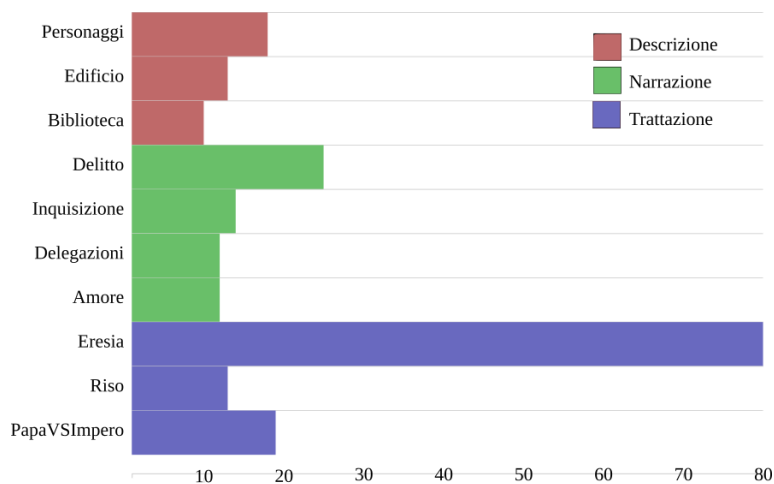


Figura 9: Analisi quantitativa dell'annotazione

Raccontare la storia dell'eresia significa riportare alla luce uno dei centri propulsivi della correzione echiana, tesa a semplificare porzioni di testo sentite come troppo pedanti. Ricostruirne le linee generali attraverso StorymapJS e TimelineJS può significativamente avvicinare il lettore e gli studenti a uno dei temi della narrazione, a tratti obliata dallo stesso autore. In particolare, StorymapJS permette di 'geolocalizzare' gli eventi (Figura 10) in una sorta di presentazione interattiva che favorisce l'esplorazione con link esterni, video, immagini. Si sono così sottratte le parole echiane – che fanno dello sfondo storico il campo privilegiato di discussione dei problemi del presente – alla pura erudizione storica per incarnarle in una mappa degli eventi, con immagini tratte dalla realtà e con documenti storici, come ad esempio la targa in ricordo di Fra Dolcino a Vercelli.



Figura 10: Schermata dell'unità didattica sulla storia dell'eresia (StorymapJS)

Se StorymapJS ha come dimensione privilegiata quella spaziale, TimelineJS, sviluppato anch'esso da Knightlab, preferisce la dimensione temporale inserendo gli eventi in una linea del tempo. Anch'essa permette di costruire relazioni a tratti inesplorate e inusuali, come è possibile vedere in Figura 11 con il passo echiano che parla di una *Historia fratris Dulcini Heresiarche* e l'incunabolo del XXVIII canto dell'*Inferno* (commentato da Cristoforo Landino e datato agosto 1481) in cui Dante cita Fra Dolcino, protagonista dell'eresia principale raccontata ne *Il nome della rosa*. La temporalità diventa così esperienza per un confronto di testi anche distanti, ma vicini al lettore, come accade nel caso del parallelismo tra le parole di Ubertino su Fra Dolcino e la rilettura datane da Dario Fo in *Mistero buffo* ([24]).

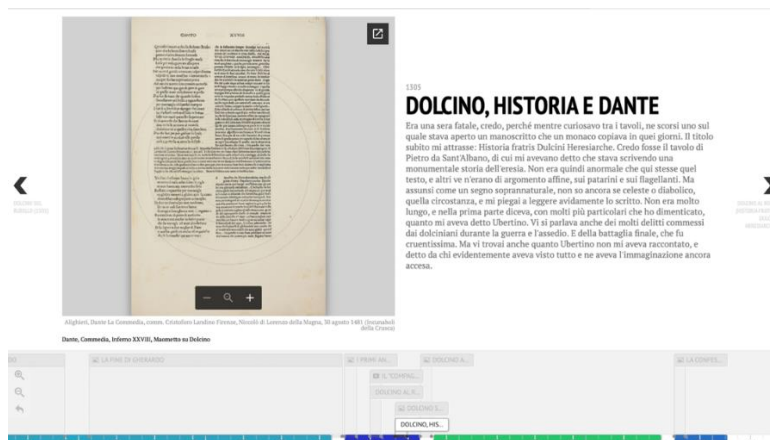


Figura 11: Schermata dell'unità didattica sulla storia dell'eresia (Timeline JS)

4.4 Edizione, interfaccia e visualizzazione

L'edizione, *stricto sensu*, è relativa alla sola edizione dell'apparato critico delle varianti, reso disponibile attraverso l'uso di EVT2 tramite una codifica XML-TEI con la seguente struttura:

```
<app xml:id="NR_0626_01">
  <lem wit="#NR1980">ivi lo stagirita</lem>
  <rdg wit="#NR2012" type="sostanziale" cause="mutamentoOrizzonti">in
  questo libro Aristotele</rdg>
</app>
```

Codifica del testo che poi in EVT2 viene visualizzata nel seguente modo, con la possibilità di cliccare sulla parola e avere una sorta di apparato a comparsa:

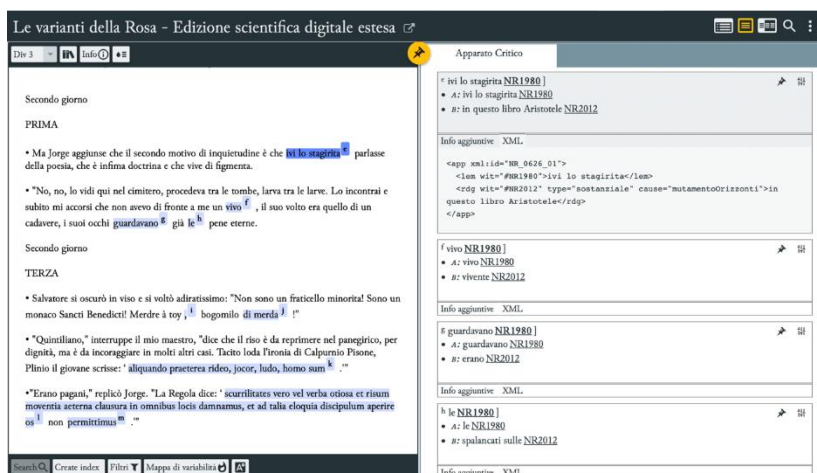


Figura 12: EVT2

EVT2 permette inoltre di raggruppare le varianti (ad esempio il gruppo NR1980, comprendente la prima edizione e le ristampe successive, e il gruppo NR2012 che presenta la revisione complessiva di Eco del 2012), aggiungere delle note editoriali e visualizzare il codice XML. In particolare, il modello della codifica rispetta le indicazioni delle *Guidelines TEI* sulla rappresentazione dell'apparato con il *parallel segmentation method*, usando il tag `<lem>` per rappresentare la lezione a testo e il tag `<rdg>` per le varianti in apparato, aggiungendo gli attributi `@type` (con valori quali «sostanziale» e «formale») e `@cause` (con valori quali «semplificazione», «mutamento orizzonte d'attesa», «evitare ripetizione», «alleggerimento», «traduzione latinismi», ecc.), e in alcuni casi il tag `<note>` per il commento critico. La scelta della prima edizione in una delle sue ultime ristampe (NR80\2010) – sostanzialmente uguale all'ultima ristampa che presenta correzioni (NR83) – come testimone base è stata determinata dalla volontà di allestire un'edizione critica che restituisse il testo frutto dell'ispirazione originaria dell'autore, recuperando la *facie* linguistica di quella prima edizione divenuta *bestseller* e letta per più di trent'anni, presentando invece con un apparato critico genetico le lezioni presenti esclusivamente in NR80 e/o NR81 e con un apparato evolutivo le varianti della redazione di NR12.

L'edizione, però, *lato sensu*, è quella che emerge dall'integrazione delle varie sezioni del portale, quella che abbiamo definito 'edizione estesa' che nasce dal dialogo tra didattica e scientificità, tra interfaccia, *data visualization* e annotazione.

4.5 Annotazione, Euforia e interrogazione

Infine, l'annotazione si configura come il momento culminante del paradigma IDEA, la fase in cui discutere l'interpretazione in modo esplicito, recuperando la lezione analogica di Barthes ([2]). Ne *Le varianti della rosa* l'annotazione è stata condotta su Euforia, uno strumento *web-based*, progettato e implementato dal CoPhiLab del CNR-IILC, che permette di annotare il testo attraverso un *Domain Specific Language* ([49]), ovvero un linguaggio formale – descritto da una *Context-Free Grammar* (CFG) che ne definisce lessico e regole di formazione, consentendo di esprimere in maniera sintetica ed efficace informazioni relative a un dominio di conoscenza. L'idea da cui essa è scaturita è quella di un approccio all'annotazione che sia il più possibile vicino alle esigenze dell'annotatore, come è stato scritto da Boschetti e Del Grosso ([5]: 78):

Nell'approccio che adottiamo, la CFG viene realizzata insieme agli studiosi o agli studenti che effettuano le annotazioni per rispondere a un'ipotesi di ricerca. Il DSL è un linguaggio familiare all'annotatore. Nei nostri progetti ciò significa, da un lato, essere vicini alle convenzioni tradizionali adottate nell'ambito degli studi umanistici (per apparati critici, etc.) e, dall'altro, essere vicini alle nuove convenzioni diffuse sui social media, come l'uso di *hashtag* per marcare parole chiave e trend topics.

Dal punto di vista tecnico, per rendere le annotazioni computabili, Euforia sfrutta un compilatore di compilatori (*compiler compiler*), ANTLR, che a partire dalla *Context-Free*

Grammar «genera un *parser* per il linguaggio stesso, che viene utilizzato sia per la validazione delle annotazioni, sia per la serializzazione in XML con uno schema proprietario conforme alle regole della grammatica del DSL» ([6]: 90). Ciò significa che l'annotatore ha la possibilità di concentrarsi sul proprio dominio verificando contestualmente la correttezza formale della propria, con la possibilità infine di avere come output un file XML ulteriormente trasformabile attraverso fogli di stile.

La potenzialità dei DSL sta chiaramente nella grande versatilità – a patto di comprenderne le regole di base – con Euporia che è stato utilizzato per rappresentare, oltre alle varianti del *Nome della rosa*, gli apparati filologici del *Qohelet* ([1]), le formule dei rituali greci presenti nella tragedia classica ([42]), le tre redazioni della traduzione leopardiana, in sestine, della *Batracomiomachia* (condotta da Matteo Cazzato e descritta in [5]: 26-27) e l'attività scolastica di traduzione e annotazione di classici greci in un'esperienza di didattica inclusiva e collaborativa presso il Liceo Gargallo di Siracusa e il Liceo Galilei di Pisa ([15]).¹¹

In linea con l'idea dello *user-centered design*, i primi prototipi di Euporia si sono basati su interfacce diverse, a seconda delle esigenze dell'annotatore: nel caso de *Il nome della rosa* il layout dell'interfaccia si è articolato in due finestre (Figura 13): sulla sinistra il testo con le lezioni dell'Ottanta in rosso e quelle del Dodici in blu – attraverso una collazione implementata con *CollateX*–; e sulla destra un *editor* di testo che ne validava l'annotazione (in base alla CFG).

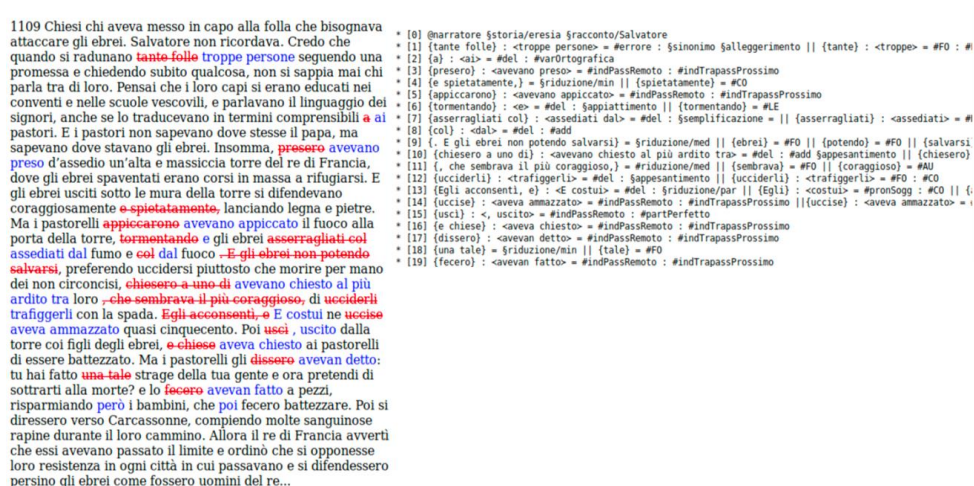


Figura 13: Schermata di Euporia durante la fase di annotazione

¹¹ Sui dettagli di Euporia e dell'annotazione DSL si rimanda quindi a [42]; [1]; [5], che presentano in maniera più esplicita le caratteristiche delle CFG, dei parser e dell'*user-centered design* che fa da sfondo alla progettazione di Euporia.

La CFG è stata invece progettata avendo come obiettivo quello di rappresentare sinteticamente informazioni di tipo sia filologico che interpretativo riassumibili nel seguente schema:

- l'uguale (=) è stato utilizzato per collegare la lezione presa in esame con la relativa annotazione linguistica e interpretativa;
- le parentesi graffe ({}) per indicare i passi rimossi dall'edizione del 1980 e le parentesi uncinate (< >) invece per quelli aggiunti del Dodici;
- i due punti (:) per indicare la sostituzione di un passo dell'Ottanta con uno del Dodici, mettendo in relazione i lemmi sostituiti con le annotazioni interpretative corrispondenti secondo un formalismo tipico delle proporzioni matematiche (lezione NR80 : lezione NR12 = annotazione NR80 : annotazione NR12);
- l'*hashtag* (#) per descrivere i fenomeni linguistici e per inserire le marche d'uso di De Mauro (espresse con due lettere maiuscole, ad es. «LE» per «Letterario»);
- il simbolo § per i descrittori interpretativi;
- la chiocciola (@) per indicare i personaggi (parlanti e relativi interlocutori) coinvolti nella variante (se più di uno, divisi da barra, con il primo individuato come *sender* e il secondo come *recipient*);
- la barra obliqua (/) per specificare il tag che lo precede (nella maggior parte dei casi per circoscriverne il tema);
- la doppia barra verticale (| |) introduce una *subannotation* dei singoli lemmi.

Con questo formalismo si è così descritta la pratica correttoria echiana avendo sempre come orizzonte la domanda di ricerca iniziale. Sebbene a una prima lettura possa sembrare complessa, la potenzialità espressiva è enorme, come è possibile vedere nel seguente esempio completo, dove alla lezione originale dell'Ottanta «ivi lo stagirita» viene sostituita nel Dodici la lezione «in questo libro Aristotele»:

```
&626
[0] @Bencio/Guglielmo §Aristotele/Poetica
[1] {ivi lo stagirita} : <in questo libro Aristotele> =
$delRiferimentoErudito : $semplificazione $mutamentoOrizzonti
$esplicitazione || {ivi} = #LE || {stagirita} = #TS/stor ||
<questo> = #FO || <libro> = #FO || <Aristotele> = #nomeProprio
```

L'annotazione si riferisce al paragrafo 626 del testo, segnalato con la &, ed è divisa in due sezioni fondamentali: la prima – con esponente «[0]» – contiene le informazioni sui personaggi coinvolti nella variante e sull'argomento principale del paragrafo; la seconda – con esponente «[1]» – consiste nell'annotazione della variante e dei singoli lemmi, usando le marche d'uso di De Mauro per le parole autosemantiche, ovvero dotate di significato, rimuovendo alcune *stop words*

(soprattutto i verbi ausiliari «essere» e «avere», i verbi servili come «potere» e «dovere», più altre parole ad altissima occorrenza presenti in ogni testo).

Una volta completata l'annotazione, grazie ad ANTLR si passa al *parsing* di tutti i nodi e viene prodotto un file XML conforme alla struttura ad albero della *Context-Free Grammar*.

Si veda la serializzazione XML del passo sopracitato generata automaticamente (Figura 14).

```

<anno>
<refNum>626</refNum>
<criticalNote>
  <ref>[<refNum>0</refNum>]</ref>
  <communication>
    <sender>@<text>Bencio</text></sender>
    <recipient>/<text>Guglielmo</text></recipient>
  </communication>
  <interpTag>§<text>Aristotele</text></interpTag>
  <themTag>/<text>Poetica</text></themTag>
</criticalNote>
<app>
  <ref>[<refNum>1</refNum>]</ref>
  <annotation>
    <subst>
      <del>{<text>ivi lo stagirita</text>}</del>
      <add><text>in questo libro Aristotele</text>></add>
    </subst>
    <substAnno>
      <delAnno>
        <interpTag>§<text>delRiferimentoErudito</text></interpTag>
      </delAnno>
      <addAnno>
        <interpTag>§<text>semplificazione</text></interpTag>
        <interpTag>§<text>mutamentoOrizzonti</text></interpTag>
        <interpTag>§<text>esplicitazione</text></interpTag>
      </addAnno>
    </substAnno>
  </annotation>
  <subannotations>
    <annotation>
      <del>{<text>ivi</text>}</del>
      <delAnno><lingTag>#<text>LE</text></lingTag></delAnno>
    </annotation>
    <annotation>
      <del>{<text>stagirita</text>}</del>
      <delAnno><lingTag>#<text>TS</text></lingTag><themTag>/<text>stor</text></themTag></delAnno>
    </annotation>
    <annotation>
      <add><text>questo</text></add>
      <addAnno><lingTag>#<text>F0</text></lingTag></addAnno>
    </annotation>
    <annotation>
      <add><text>libro</text></add>
      <addAnno><lingTag>#<text>F0</text></lingTag></addAnno>
    </annotation>
    <annotation>
      <add><text>Aristotele</text>></add>
      <addAnno><lingTag>#<text>nomeProprio</text></lingTag></addAnno>
    </annotation>
  </subannotations>
</app>
</anno>

```

Figura 14: Serializzazione XML generata a partire dall'annotazione di Euporia

La codifica così generata, che esplicita tutta la semantica e la struttura gerarchica dell'annotazione, risulta molto più verbosa, di difficile comprensione, sebbene ancora, entro certi limiti, leggibile. Il riferimento al paragrafo si serve del tag <refNum> mentre le due sezioni dell'annotazione sono rappresentate da <criticalNote> – per la codifica delle informazioni contestuali sul parlante (sender), interlocutore (recipient) e argomento (interpTag e themTag) – e da tanti <app> quanto sono le entrate di apparato. La codifica di ciascun <app> è suddivisa in <annotation> e <subAnnotation>, con i tag <subst>, e <add> per indicare rispettivamente sostituzioni, cancellazioni e aggiunte, mentre i tag <substAnno>, <delAnno> e <addAnno> per indicare annotazioni specifiche (con all'interno <lingTag>, <interpTag> e <themTag>, per le annotazioni linguistiche, interpretative e tematiche).

La codifica XML proprietaria permette già un'interrogazione attraverso *XPath* e *XQuery*, ma risulta poco funzionale sia all'integrazione nel Database (che presenta al suo interno il file dell'edizione in TEI), sia per la gestione FAIR dei dati di ricerca, non essendo un tipo di codifica pensato per l'interoperabilità e per la riusabilità da parte di altri studiosi. Per questo motivo attraverso un foglio di stile XSLT, creato *ad hoc*, si è proceduto a trasformare l'annotazione rendendola TEI *compliant*. Dopo aver applicato il foglio di stile l'esempio di Aristotele risulta così codificato in TEI (Figura 15).

L'annotazione viene compresa in un <div> con un proprio @xml:id che può essere collegato attraverso l'attributo @target al <div> corrispondente dell'edizione. La sezione <criticalNote> viene trasformata in un tag <note> di tipo *criticalNote* (con all'interno dei tag <interp> di tipo *speaker* per il parlante, *recipient* per il destinatario, *topic* per l'argomento) mentre tutte le altre entrate di apparato usate i tag appropriati <listApp> e <app>. Le lezioni dall'edizione dell'Ottanta (secondo la ristampa NR80\2010) – in linea con l'edizione scientifica digitale – vengono codificate con <lem>, mentre le lezioni di redazioni successive con <rdg>. Il resto delle annotazioni viene rappresentato da <note> nidificati e da <interp> di tipo *linguistic*, per le marche d'uso e le altre annotazioni linguistiche, di tipo *hermeneutic*, per le annotazioni maggiormente interpretative e soggettive, e di tipo *thematic*, per esplicitare il tema dell'annotazione.¹²

¹² L'attributo thematic è infatti utilizzato prevalentemente per specificare meglio il tema (o argomento) del tag <interp> all'interno dello stesso <note>.

```

<div type="nota" xml:id="note_NR_0626">
  <note type="loc">626</note>
  <note type="criticalNote">
    <note type="communication">
      <interp type="speaker">Bencio</interp>
      <interp type="recipient">Guglielmo</interp>
    </note>
    <interp type="topic">Aristotele</interp>
    <interp type="thematic">Poetica</interp>
  </note>
  <listApp>
    <desc><num>1</num></desc>
    <app>
      <lem wit="NR1980">ivi lo stagirita</lem>
      <rdg wit="NR2012">in questo libro Aristotele</rdg>
      <note type="substAnno">
        <note type="delAnno">
          <interp type="hermeneutic">delRiferimentoErudito</interp>
        </note>
        <note type="addAnno">
          <interp type="hermeneutic">semplificazione</interp>
          <interp type="hermeneutic">mutamentoOrizzonti</interp>
          <interp type="hermeneutic">esplicitazione</interp>
        </note>
      </note>
    </app>
    <listApp type="subannotations">
      <app>
        <lem wit="NR1980">ivi</lem>
        <note type="delAnno">
          <interp type="linguistic">LE</interp>
        </note>
      </app>
      <app>
        <lem wit="NR1980">stagirita</lem>
        <note type="delAnno">
          <interp type="linguistic">TS</interp>
          <interp type="thematic">stor</interp>
        </note>
      </app>
      <app>
        <rdg wit="NR2012">questo</rdg>
        <note type="AddAnno">
          <interp type="linguistic">FQ</interp>
        </note>
      </app>
      <app>
        <rdg wit="NR2012">libro</rdg>
        <note type="AddAnno">
          <interp type="linguistic">FQ</interp>
        </note>
      </app>
      <app>
        <rdg wit="NR2012">Aristotele</rdg>
        <note type="AddAnno">
          <interp type="linguistic">nameProprio</interp>
        </note>
      </app>
    </listApp>
  </listApp>
</div>

```

In questo modo attraverso il linguaggio *XQuery* è possibile interrogare i dati dell'annotazione facendo emergere interpretazioni utili ad ampliare l'edizione estesa e a rimodellarla, come ad esempio è stato fatto per confermare la presenza di un certo grado di semplificazione nell'edizione del Dodici ([16]). A titolo puramente esemplificativo, in linea con l'esempio qui esplicitato, si

veda la *query* per quantificare il numero di correzioni interpretate come mutamenti di orizzonte d'attesa del pubblico (similmente a 'stagirita' corretto in 'Aristotele'):

```
count (  
for $interp in doc("/Varianti della  
rosa/Database/AnnotazioneEuporiaTei.xml")//interp[@type="herm  
eneutic"]  
where $interp = "mutamentoOrizzonti"  
return $interp  
)
```

Figura 15: Output XML-TEI dell'annotazione DSL

In questo modo, interrogando l'annotazione e quantificando le proprie interpretazioni (insieme alla lemmatizzazione del testo, all'edizione scientifica digitale, alle visualizzazioni alternative e alle unità didattiche che presentano attività critiche tese a dare nuova luce ai testi letterari del passato rendendoli attuali), si può provare a gettare quel ponte tra quantificazione ed ermeneutica che è stato qui avanzato a proposito del circolo ermeneutico digitale. Ponte forse impossibile secondo Moretti ([41]), ma che vale la pena tentare.

5. Verso la nuova edizione «riveduta e corretta»

Il prototipo di edizione digitale qui esposto trova la sua ragion d'essere nell'integrazione tra le diverse modalità di approccio al testo, alcune delle quali solitamente marginali nello studio critico, ma centrali in un'ottica di diffusione degli studi specialistici anche a un pubblico non accademico. Se si pensa che la critica letteraria digitale non si debba limitare a riflettere soltanto su opere del lontano passato, ma debba interrogarsi e proporre contenuti sui testi significativi del presente, il paradigma IDEA dell'edizione estesa mira ad arricchire la conoscenza di opere di letteratura contemporanea attraverso approfondimenti in rete. Ma la potenzialità del digitale sta anche nel continuo ampliamento dei suoi confini, con nuove proposte e riflessioni. Diventa centrale quindi il *Versioning* e una *road map* dei contenuti futuri. Una versione 3.0 del portale-edizione (Figura 16) è in sviluppo con:

- una nuova annotazione con il possibile sviluppo di un'ontologia secondo una tassonomia delle varianti ben definita e un motore di ricerca delle varianti per tema, personaggio, intervento correttivo;
- integrazione completa tra la codifica TEI e l'annotazione DSL;
- migrazione a EVT3 per l'interfaccia della DSE;

- altre unità didattiche su temi fondamentali del romanzo, sulle singole varianti (verso una ‘top20’ con infografiche e approfondimenti sulle venti varianti più significative), sul rapporto tra varianti del romanzo e varianti tra film e serie tv, nell’ottica di uno *storytelling* transmediale;
- sviluppo di una sezione del portale aperta al contributo di *content creator* – soprattutto studiosi e chiani e studenti – nell’ottica di una ricerca collaborativa;
- alcuni esperimenti di *gamification* con dei quiz pensati per studenti di liceo e dell’università;
- rilascio di tutti i dati e tutti i contenuti (*open content*) nel rispetto dei principi FAIR.

Così come Umberto Eco, trentadue anni dopo la pubblicazione del suo capolavoro, decideva di intervenire e correggere il proprio testo, con «un’edizione riveduta e corretta», così *Le varianti della rosa* nella sua stessa essenza si propone di essere in continuo aggiornamento, con la consapevolezza che «stat rosa pristina nomine, nomina nuda tenemus» ([20]: 503).

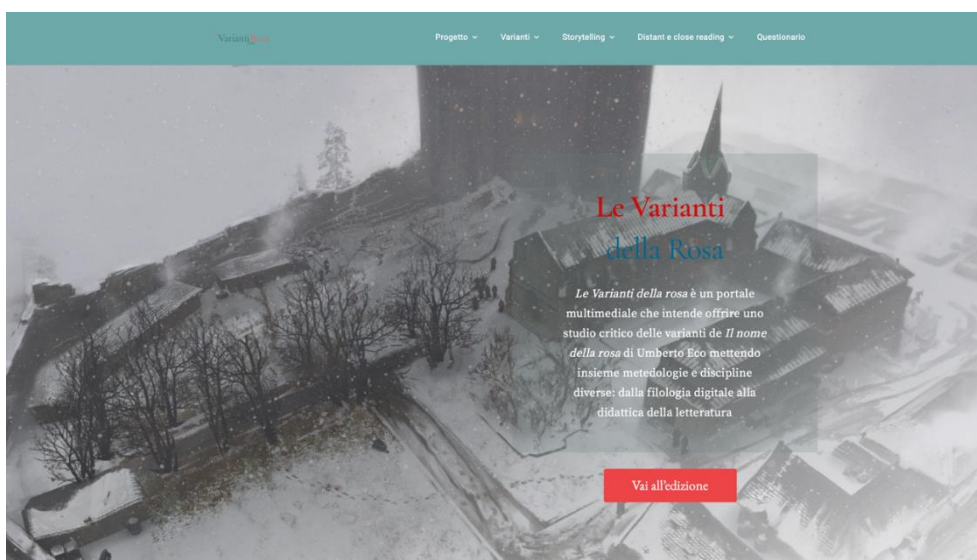


Figura 16: Il Portale-edizione “Le varianti della rosa”

Ringraziamenti

Si ringrazia il CNR-ILC (e Federico Boschetti) per l’elaborazione del *Domain Specific Language* e della CFG, per lo sviluppo dell’interfaccia personalizzata di Euporia e per la serializzazione in XML; Roberto Rosselli del Turco e gli sviluppatori di EVT; e infine il CINUM per la cornice

interpretativa dell'edizione integrata e Pietro Sichera per l'utilizzo del software di concordanza *LiotroConcord*.

References

- [1] Bambaci, Luigi, Federico Boschetti e Riccardo Del Gratta. 2019. "Qohelet Euporia: a Domain specific Language for the Encoding of the critical Apparatus." *International Journal of Information Science & Technology*, vol. 3 no. 5: 26-37. <http://dx.doi.org/10.1109/CIST.2018.8596332>.
- [2] Barthes, Roland. 1970. *S/Z*. Paris: Seuil.
- [3] Bleier, Roman, Martina Bürgermeister, Helmut W. Klug, Frederike Neuber e Gerlinde Schneider (ed.). 2018. *Digital Scholarly Editions as Interfaces*. Norderstedt: BoD.
- [4] Bonsi, Claudia, Angelo Di Iorio e Paola Italia. Fabio Vitali. 2015. "Manzoni's Electronic Interpretations." *The Mechanic Reader*, LIII (2015/2): 91-99.
- [5] Boschetti, Federico e Angelo M. Del Grosso. 2020. "L'annotazione di testi storico-letterari al tempo dei social media." *Italica Wratislaviensia*, 11 (1): 65-99. <http://dx.doi.org/10.15804/IW.2020.11.1.03>.
- [6] Boschetti, Federico e Gloria Mugelli. 2021. "Il metodo Euporia per creare nuovi archivi digitali sulla tragedia greca." *FuturoClassico*, 7: 83-113. <https://doi.org/10.15162/2465-0951/1381>.
- [7] Buzzetti, Dino. 2002. "Digital Representation and the Text Model." *New Literary History*. Winter, Vol. 33, No. 1: 61-88.
- [8] Buzzoni, Marina. 2016. "A Protocol for Scholarly Digital Editions? The Italian Point of View." In *Digital Scholarly Editing. Theories and Practices*. Edited by M. J. Driscoll and E. Pierazzo, 59-82. Cambridge, UK: Open Public Publishers.
- [9] Ciotti, Fabio. 2017. "Modelli e metodi computazionali per la critica letteraria: lo stato dell'arte." In *L'Italianistica oggi: ricerca e didattica*, Atti del XIX Congresso dell'ADI - Associazione degli Italianisti (Roma, 9-12 settembre 2015), a cura di B. Alfonzetti, T. Cancro, V. Di Iasio, E. Pietrobon, Roma, Adi editore (http://www.italianisti.it/Atti-di-Congresso?pg=cms&text=p&cms_codsec=14&cms_codcms=896).
- [10] Ciotti, Fabio. 2018. "From Informatica Umanistica to Digital Humanities and return: a conceptual history of Italian DH". *TESTO & SENSO*. 19. 2018.
- [11] Ciotti, Fabio. 2018. "A Formal Ontology for the Text Encoding Initiative". *Umanistica Digitale*, 2 (3). <https://doi.org/10.6092/issn.2532-8816/8174>.
- [12] Ciotti, Fabio, Elena Corradini, Elisa Cugliana, Giulia D'Agostino, Lorenzo Ferroni, Franz Fischer, Maurizio Lana, Paolo Monella, Torsten Roeder, Roberto Rosselli Del Turco e Patrick Sahle. 2022. "Manifesto per le edizioni scientifiche digitali". *Umanistica Digitale*, 12, 103-108. <https://doi.org/10.6092/issn.2532-8816/14814>.

- [13] Ciula, Arianna, Øyvind Eide, Cristina Marras e Patrick Sahle (eds.). 2018. *Models and Modelling between Digital and Humanities: A Multidisciplinary Perspective*. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-62883-7>.
- [14] Cristofaro, Salvatore e Daria Spampinato. 2021. "OntoBelliniLetters: A Formal Ontology for a Corpus of Letters of Vincenzo Bellini." In *Metadata and Semantic Research*. MTSR 2020. Communications in Computer and Information Science, vol 1355. Edited by E. Garoufallou, M. Ovale-Perandones. Cham: Springer. https://doi.org/10.1007/978-3-030-71903-6_19.
- [15] Crucitti, Marilena, Michela Benedetti, Greta Maneschi, Roberta Mirandola, Antonella Soldani, Ludovico Amato, Filippo Lepori, Andrea Taddei e Federico Boschetti. 2021. "La collaborazione inclusiva: un'esperienza didattica di annotazione tramite Euporia." *Umanistica Digitale*. 11. <http://doi.org/10.6092/issn.2532-8816/13680>.
- [16] D'Agata, Christian. 2021. "I Nomi della Rosa. Un'analisi testuale informatica delle varianti del Nome della Rosa tra Distant e Close reading." In *Letteratura e Scienze. Atti delle sessioni parallele del XXIII Congresso dell'ADI Roma*: Adi editore. https://www.italianisti.it/publicazioni/atti-di-congresso/letteratura-e-scienze/1.ADI_19_D_AGATA.pdf.
- [17] Daquino, Marilena, Francesca Giovannetti e Francesca Tomasi. 2019. "Linked Data per le edizioni scientifiche digitali. Il workflow di pubblicazione dell'edizione semantica del quaderno di appunti di Paolo Bufalini." *Umanistica Digitale*. 3 (7). <https://doi.org/10.6092/issn.2532-8816/9091>.
- [18] De Mauro, Tullio. 1999-2000. *Grande dizionario italiano dell'uso* con CD-Rom, Torino: Utet.
- [19] Di Silvestro, Antonio, Christian D'Agata, Giuseppe Palazzolo e Pietro Sichera. 2022. "Conservazione e fruizione di banche dati letterarie: l'archivio della poesia italiana dell'Otto/Novecento di Giuseppe Savoca." In *Culture digitali. Intersezioni: filosofia, arti, media. Precedings della 11a conferenza nazionale*. A cura di F. Ciraci, G. Miglietta, & C. Gatto. Lecce: AIUCD, 98-104.
- [20] Eco, Umberto. 1980. *Il nome della rosa*. Milano: Bompiani.
- [21] Eco, Umberto. 2012. *Il nome della rosa*. I edizione riveduta e corretta. Milano: Bompiani.
- [22] Eco, Umberto e Paolo Di Stefano. 2012. *Eco: così ho rivisto «Il nome della rosa» ma salvatemi dai critici militanti*, Corriere della sera, 31 gennaio 2012. https://www.corriere.it/cultura/libri/12_gennaio_31/di-stefano-eco-nome-rosa_2f91ca48-4c25-11e1-8f5b-8c8dfe2e8330.shtml.
- [23] Esperti, Pietro. 1979. "Grammatichetta della lingua italiana ad uso del calcolatore." In *Al servizio del vocabolario della lingua italiana*, A cura di S. D'Arco Avalle. Firenze: Accademia della Crusca.
- [24] Fo, Dario. 2018. *Mistero buffo*. Milano: Guanda. Prima pubblicazione Cremona: Nuova Scena, 1969.
- [25] Gadamer, Hans Georg. 2000. *Verità e metodo*. Trad. e apparati di G. Vattimo, Milano: Bompiani. Prima pubblicazione Tübingen: J.C.B. Mohr, 1960.
- [26] Gigliozzi, Giuseppe (ed.). 1987. *Studi di codifica e trattamento automatico di testi*. Roma: Bulzoni.
- [27] Giuffrida, Milena, Christian D'Agata, Laura Giurdanella e Pietro Sichera. 2021. "Pirandello Nazionale: per un nuovo modello di edizione digitale, collaborativa e integrata." In *AIUCD 2021 - DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale. Raccolta degli abstract*

- estesi della 10a conferenza nazionale*. A cura di F. Boschetti, A. M. Del Grosso, E. Salvatori. Pisa: AIUCD, 207-214.
- [28] Giunta, Claudio. 2011. "La filologia d'autore non andrebbe incoraggiata." *Ecdotica*. 1/2011: 104-119. DOI: 10.7385/99118.
- [29] Giusti, Simone. 2015. *Didattica della letteratura 2.0*. Roma: Carocci.
- [30] Grondin, Jean. 2020. *L'ermeneutica*. Brescia: Queriniana.
- [31] Heidegger, Martin. 2015. *Essere e tempo*. Traduzione a cura di Alfredo Marini. Milano: Mondadori. Prima pubblicazione Tübingen: J.C.B. Mohr, 1960. Prima pubblicazione Tübingen: Max Niemeyer Verlag, 1927,
- [32] Italia, Paola. 2020. *Editing Duemila. Per una filologia dei testi digitali*. Roma: Salerno Editrice.
- [33] Janicke, Stefan, Annette Geßner, Greta Franzini, Melissa Terras, Simon Mahony e Gerik Scheuermann. 2015. "TRAViz: A Visualization for Variant Graphs." *Digital Scholarship in the Humanities*, Volume 30: 83-89.
- [34] Lave, Jean e Etienne Wenger. 1990. *Situated Learning: Legitimate Peripheral Participation*. Cambridge (UK): Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>.
- [35] Maas, Paul. 1952. *Critica del testo*. Traduzione a cura di Nello Martinelli. Firenze: Le Monnier.
- [36] Mancinelli, Tiziana e Elena Pierazzo. 2020. *Che cos'è un'edizione scientifica digitale*. Roma: Carocci.
- [37] Manganaro, Andrea. 2020. *Per una didattica della letteratura italiana*, Acireale-Roma: Bonanno.
- [38] Monella, Paolo. 2018. "Livelli di rappresentazione del testo nell'edizione del De nomine di Orso Beneventano." *Umanistica Digitale*, 2 (2). <https://doi.org/10.6092/issn.2532-8816/7286>.
- [39] Mordenti, Raul. 2012. "Domande teoriche sul concetto di edizione (nel nome di Giuseppe Gigliozzi)". In *Dall'Informatica umanistica alle culture digitali. Atti del Convegno di studi (Roma, 27-28 ottobre 2012) in memoria di Giuseppe Gigliozzi*. In "Quaderni Digilab", vol.2, a cura di F. Ciotti e G. Crupi, 167-182. http://digilab-epub.uniroma1.it/index.php/Quaderni_Digilab/article/view/32/29. DOI: 10.7357/DigiLab-22.
- [40] Moretti, Franco. 2020. *A una certa distanza*. Roma: Carocci. Prima pubblicazione London: Verso, 2013.
- [41] Moretti, Franco. 2022. *Falso movimento. La svolta quantitativa nello studio della letteratura*. Milano: notte tempo.
- [42] Mugelli, Gloria, Federico Boschetti, Angelo M. Del Grosso, Fahad Khan e Andrea Taddei. 2016. "A user-centred design to annotate ritual facts in ancient Greek tragedies." *Bulletin of the Institute of Classical Studies*, 59: 103-120.
- [43] Nantke, Julia. Schlupkothen Frederik (eds.). 2020. *Annotations in scholarly editions and research*. Berlin-Boston: De Gruyter.
- [44] Orlandi, Tito. 1992. "Informatica umanistica: realizzazioni e prospettive." In *Calcolatori e scienze umane, Scritti del Convegno organizzato dall'Accademia Nazionale dei Lincei e dalla Fondazione IBM Italia*. Milano: Etas Libri, 1-22.
- [45] Orlandi, Tito. 2007. "Teoria e prassi di una edizione computazionale." In *Digital philology and medieval texts*, a cura di A. Ciula e F. Stella. Pisa: Pacini, 85-92.
- [46] Pace, Rosaria. 2015. *Digital Humanities, una prospettiva didattica*, Roma: Carocci.

- [47] Paci, Deborah. 2021. "Conoscere è partecipare: digital public history, wiki e citizen humanities." *Umanistica Digitale*. 10. 235–249. <https://doi.org/10.6092/issn.2532-8816/12555>.
- [48] Parr, Terence. 2010. *Language implementation Patterns: Create Your Own Domain-Specific and General Programming Languages: Techniques for Implementing Domain-Specific Languages*, Raleigh (NC): Pragmatic Bookshelf.
- [49] Parr, Terence. 2013. *The Definitive ANTLR 4 Reference*. 2nd ed. Raleigh (NC): Pragmatic Bookshelf.
- [50] Pescatore, Guglielmo. 2018. *Ecosistemi narrativi*. Roma: Carocci.
- [51] Petrucco, Corrado e Marina De Rossi. 2009. *Narrare con il digital storytelling a scuola e nelle organizzazioni*. Roma: Carocci.
- [52] Pierazzo, Elena. 2005. *La codifica dei testi*. Roma: Carocci.
- [53] Piper, Andrew. 2018. *Enumerations. Data and Literary Studies*. Chicago: Chicago University Press.
- [54] Pischedda, Bruno. 2016. *Eco: guida al Nome della rosa*. Roma: Carocci.
- [55] Roeder, Torsten. 2020. "Review of 'Juxta Web Service, LERA, and Variance Viewer. Web based collation tools for TEI'." *RIDE*. 11. DOI: 10.18716/ride.a.11.5.
- [56] Rosselli Del Turco, Roberto, Chiara Di Pietro e Chiara Martignano. 2020. "Progettazione e implementazione di nuove funzionalità per EVT 2: lo stato attuale dello sviluppo." *Umanistica Digitale*. 7.
- [57] Sahle, Patrick. 2016. "What is a Scholarly Digital Edition?". In *Digital Scholarly Editing. Theories and Practices*. Edited by M. J. Driscoll and E. Pierazzo, 19-40. Cambridge, UK: Open Public Publishers.
- [58] Salvatori, Enrica. 2017. "Digital (Public) History: la nuova strada di una antica disciplina". *RiMe*. 1: 57-94. DOI: 10.7410/1291.
- [59] Savoca, Giuseppe. 2000. *Lessicografia letteraria e metodo concordanziale*. Firenze: Olschki.
- [60] Schleiermacher, Friedrich. 1996. *Ermeneutica*. Traduzione e apparati di Massimo Marassi. Milano: Rusconi. Prima pubblicazione Berlin: Reimer, 1838.
- [61] Schmidt, Desmond e Robert Colomb. 2009. "A data structure for representing multi-version texts online." In *International Journal of Human-Computer Studies*, 67, 6: 497-514. <https://doi.org/10.1016/j.ijhcs.2009.02.001>.
- [62] Sichera, Antonio. 2019. *Ermeneutiche. Punti di vista dal confine*. Leonforte (en): Euno edizioni.
- [63] *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 4.3.0. Last updated on 31st August 2021. TEI Consortium. <http://www.tei-c.org/Guidelines/P5/>.
- [64] Terras, Melissa, Julianne Nyhan e Edward Vanhoutte (eds.). 2013. *Defining Digital Humanities - A Reader*. Williston: Ashgate.
- [65] van Zundert, Joris J. 2019. *Why the Compact Disc Was Not a Revolution And «Cityfish» Will Change Textual Scholarship, or What is a Computational Edition?*. *Ecdotica*. 2019. Vol. 15, 129–156.
- [66] Wilkinson, Mark, Michel Dumontier, Ijsbrand Aalbersberg, et al. 2016. "The FAIR Guiding Principles for scientific data management and stewardship." *Sci Data*. 3. <https://doi.org/10.1038/sdata.2016.18>.

