

Leggere le opere del Verismo attraverso le riviste e i critici del tempo. Edizioni di testi, mappe lessicali e dizionari tematici in un portale *user oriented*

Denise Bruno

Università di Catania
denise.bruno@phd.unict.it

Giuseppe Canzoneri

Università di Catania
giuseppe.canzoneri@unict.it

Salvatore Cristofaro

CNR - Istituto di Scienze e Tecnologie della Cognizione
salvatore.cristofaro@cnr.it

Antonio Di Silvestro

Università di Catania
antonio.disilvestro@unict.it

Laura Mazzaguo

CNR - Istituto di Scienze e Tecnologie della Cognizione
laura.mazzaguo@istc.cnr.it

Daria Spampinato

CNR - Istituto di Scienze e Tecnologie della Cognizione
daria.spampinato@cnr.it

Abstract

La ricerca svolta all'interno del progetto COVerLeSS (Corpus Online del Verismo tra Letteratura, Storia e Società), collocandosi nell'ambito degli studi ascrivibili ai *Periodical Studies*, e tra i progetti di digitalizzazione e conservazione della letteratura periodica attivi da diversi anni in numerose biblioteche e istituzioni culturali italiane ed europee, mira a costruire un ambiente web integrato e *open access* nel quale trovano spazio riviste, recensioni, saggi relativi alla produzione letteraria del Verismo italiano. Il progetto mette in dialogo tecnologie e strumenti delle DH applicati agli oggetti testuali: dall'edizione digitale basata sulla marcatura XML/TEI di tipo strutturale e descrittivo con gestione e visualizzazione in TEI Publisher, all'analisi quantitativa del lessico, fino alla proposta, a partire da una serie di campi semantici, di ontologie basate su specifici elementi della codifica. La ricerca mira a costruire una vera e propria edizione-archivio, o *knowledge site*, nella quale alla lettura del testo si accompagna un'analisi dei dati sia sincronica sia diacronica, offrendo così una cartografia lessicale (il dizionario *Verbum*) basata sulla letteratura secondaria del Verismo e collegata a sua volta a quella primaria attraverso una *timeline* interattiva (*Ver-in-time*). L'ambiente web, concepito in un'ottica *user oriented* e con caratteri di inclusività, consente una pluralità di accessi a diversi livelli di utenza – dallo studente al ricercatore –, offrendo molteplici possibilità di navigazione, di fruizione dei testi, di connessione tra testualità eterogenee. Attraverso mappe lessicali e parole chiave, le varie unità testuali del progetto rivelano la coesione profonda di questa produzione letteraria e disegnano una nuova prospettiva di indagine e studio rispetto a un periodo cruciale della nostra storia culturale e letteraria come quello post-unitario.

Parole chiave: Letteratura secondaria; Verismo; Lessicografia; Filologia digitale; Edizione scientifica digitale; *Periodical Studies*; AIUCD2024

The COVerLeSS project (Corpus Online del Verismo tra Letteratura, Storia e Società) explores Italian Verismo within the field of Periodical Studies, drawing on ongoing digitization and preservation initiatives in libraries and cultural institutions across Italy and Europe. The project aims to develop an integrated, open-access web platform that brings together journals, reviews, and essays on Verismo's literary production. This platform integrates Digital Humanities tools and technologies for textual analysis, combining a digital edition with XML/TEI structural and descriptive markup, managed through TEI Publisher, quantitative lexical analysis, and the development of ontologies based on specific encoding elements. The focus is on the creation of a hybrid edition-archive, or knowledge site, where textual reading is complemented by both synchronic and diachronic data analysis. A central feature is the development of a lexical map, the Verbum dictionary, which links secondary literature on Verismo to primary sources through an interactive timeline (Ver-in-time). Designed to be user-friendly and inclusive, the platform accommodates different levels of expertise, from students to researchers, offering flexible ways to navigate, explore, and engage with diverse texts. By highlighting lexical patterns and key concepts, the project reveals the deep cohesion within this literary movement, providing a fresh perspective on a pivotal period in Italian cultural and literary history: the post-unification era.

Keywords: Secondary literature; Verism; Lexicography; Digital philology; Digital scholarly edition; *Periodical Studies*; AIUCD2024

1. Basi storico-critiche della ricerca e stato dell'arte

L'orientamento invalso nel campo di ricerca dei *Periodical Studies*, che annovera ormai un numero significativo di riviste, anche per il settore della modernità (tra queste il *Journal of Modern Periodical*

Studies),¹ è quello di riconoscere ai periodici lo statuto di oggetti culturali autonomi. Tuttavia è anche vero che l'approccio prevalente è stato ed è tuttora spesso di tipo monografico, volto al *Close Reading* di singole testate e, all'interno di queste, indirizzato ai contenuti quasi esclusivamente letterari [26:2]. Un approccio *Distant*, basato sulle 'intersezioni' tra diversi periodici cronologicamente omogenei e sull'analisi di dati legati ai testi interni, si rivela certamente più fecondo di prospettive (cfr. [26:3]), in una direzione che non può prescindere da un approccio multidisciplinare.

In questo ambito di studi, specificamente in quello dei periodici della modernità analizzati in chiave *Close e Distant* con metodi e strumenti delle Digital Humanities applicati agli oggetti testuali, si colloca la ricerca presentata in questo articolo, ricerca che si inquadra all'interno del progetto COVERLeSS, "Corpus Online del Verismo tra Letteratura, Storia e Società", finanziato su fondi PRIN 2022 PNRR. Essa, ponendo al centro dell'indagine la testualità contenuta all'interno dei periodici (recensioni, articoli, saggi), mira a conservare, valorizzare e analizzare, in un ambiente web integrato e open access, un corpus della letteratura secondaria relativa alla produzione letteraria del Verismo italiano.

L'originalità e la rilevanza culturale e scientifica della ricerca sono legate al fatto che: 1) il movimento verista accompagna il processo di stabilizzazione dell'identità nazionale italiana, e il dibattito sorto intorno a esso analizza i riflessi del processo di unificazione sulla situazione economica, sociale e culturale del Sud; 2) i testi di questo dibattito soffrono di una forte dispersione tra archivi non informatizzati, riviste non digitalizzate, biblioteche settoriali difficilmente raggiungibili; inoltre gli stessi testi brevi – soprattutto le recensioni – non sono mai presenti a testo pieno, ma per *excerpta*; 3) al di fuori di alcuni motori di ricerca bibliografici, non esistono archivi digitali strutturati di letteratura italiana secondaria, men che meno per la letteratura italiana postunitaria.

Un approccio di questo tipo richiede un'indagine capillare sui dibattiti sorti nelle riviste e nei periodici più 'monografici', impegnati in una definizione dell'idea di realtà connessa al nuovo movimento (per i repertori di riviste ottocentesche si veda [22], [5:199]). Grazie all'iniziativa di biblioteche e archivi, i progetti di digitalizzazione delle riviste si sono notevolmente incrementati (come illustrato in [21], [18]). Si pensi a testate importanti per il periodo qui indagato come «La Rassegna Settimanale» di Franchetti e Sonnino; la «Rivista Minima», presente nella collezione digitale dell'Archivio Ricordi;² il progetto "Digiteca"; o ancora, le testate digitalizzate dall'Emeroteca della Biblioteca Braidense. In tutti questi casi si tratta di progetti di digitalizzazione, utili per il reperimento della collezione completa di un periodico, ma in cui sono assenti motori di ricerca interni.

Dal punto di vista dei contenuti, la ricerca contribuisce a valorizzare, attraverso i testi, la conoscenza del dibattito culturale intorno alla situazione sociale ed economica coeva alla produzione del Verismo maggiore. L'idea da cui muove la costruzione del corpus proposto (cfr. par. 3) mira a invertire la direzione dell'indagine tradizionale (prima l'analisi delle opere e poi quella della ricezione critica e dei dibattiti sulle riviste), soffermandosi sulla capacità della letteratura secondaria di 'modellizzare' la produzione delle opere.

Il titolo scelto, COVERLeSS, letto come sostantivo, rimanda all'assenza di 'veli' nell'interpretazione dei testi del Verismo affidata ai critici e lettori contemporanei, che nei loro interventi 'scoprono', smascherano senza mediazioni i meccanismi della finzione letteraria e linguistica e li leggono nel contesto culturale e sociale di produzione dei testi stessi.

¹https://www.psupress.org/journals/jnls_jmps.html?srsltid=AfmBOoogjRVw4GZMqoB6mMuO0QkAgpr6jzZth9zuYyUQpqTFNBvKUGS

²<https://www.digitalarchivioricordi.com/it/periodicals/RivistaMinima>

L'interazione in presenza tra autori, critici e lettori fa sì che dibattiti, recensioni, saggi, rivelino in modo più esplicito le modalità in cui i vari aspetti di un'opera, non solo letterari, ma anche antropologici, storici, socio-economici, sono percepiti da chi vive nel medesimo contesto. Di conseguenza il lessico isolabile in questi testi rappresenta una vera e propria cartina di tornasole, che misura la capacità di rappresentazione e rielaborazione della realtà compiuta dagli scrittori veristi.

Da qui è nata l'esigenza di ricollocare queste testualità nel corretto e integrale contesto di produzione. Si pensi ad esempio al modo in cui sono spesso antologizzate le recensioni a un'opera, riprodotte in modo parziale, private dei riassunti della trama, delle numerose tirate moralistiche, delle digressioni aneddotiche, che non sono semplicemente segni della natura 'dialogica' di questi testi, ma identificano soprattutto l'estrazione sociale dei lettori, il loro livello culturale, la loro differente 'reattività' alle opere letterarie. Le citazioni e le allusioni peraltro rendono necessaria la contestualizzazione di questi testi entro il loro contesto di appartenenza, ossia i giornali e i periodici del tempo, all'interno dei quali essi costituiscono spesso il frammento di una discussione più ampia (si pensi al genere della polemica letteraria, o ai dibattiti che si prolungano in più numeri).

Si tratta pertanto di mettere a disposizione degli studiosi strumenti di visualizzazione e di interrogazione capaci di fare luce sui rapporti tra il dibattito politico-sociale, sulle problematiche preminenti della prima Italia unita (soprattutto con riferimento al Meridione), e sulla risonanza data a questi temi dalla nuova letteratura che in Italia prenderà concretezza con gli scritti dei veristi. Tale obiettivo richiede anzitutto che siano adeguatamente rappresentati i seguenti generi testuali: 1) recensioni contemporanee a romanzi, novelle, opere teatrali; 2) serie integrali di periodici di taglio 'monografico'; 3) saggi teorici sull'argomento. Queste tipologie differenti, dotate di registri e di un repertorio lessicale fortemente eterogenei, offrono un'immagine meno monolitica e più sfaccettata di questo periodo di storia letteraria e culturale.

Naturalmente, il *Close Reading* ai singoli testi è utile per costituire un repertorio della ricezione della letteratura verista, ma non può essere dirimente ed esclusivo. Infatti, gli strumenti di analisi quantitativa approntati, nello spirito del *Distant Reading*, rilevano la dominanza di temi e parole-chiave, nonché la loro differenziazione nei diversi generi censiti e indicizzati.

2. Obiettivi della ricerca e metodologie

La tipologia della ricerca sviluppata all'interno del progetto, mirante a sviluppare una piattaforma per la lettura, l'interrogazione per parole-chiave e/o per temi e motivi e l'analisi lessico-semantica dei testi appartenenti al corpus sopra accennato (e che verrà discusso dettagliatamente al par. 3), richiede per il perseguimento dei suoi obiettivi una serie di metodologie strettamente integrate.

Nell'ambito della filologia digitale, uno stato dell'arte completo delle edizioni scientifiche digitali si trova in Franzini³ e in Sahle.⁴ In ambito letterario sono disponibili diversi portali d'autore; tra i più recenti si segnalano: Manzoni on-line,⁵ Dante on-line,⁶ Storia d'Italia di Guicciardini⁷ e

³ <https://dig-ed-cat.acdh.oeaw.ac.at/>

⁴ <https://v3.digitale-edition.de/>

⁵ <https://www.alessandromanconi.org/>

⁶ <https://www.danteonline.it/>

⁷ <https://guicciardini-storia-italia.huma-num.fr/storia-it.html>

l'Edizione Digitale dell'Opera Omnia di Pirandello,⁸ progettata e realizzata da studiosi dell'Unità di ricerca UniCT, che rappresenta un modello di riferimento per il presente progetto.

La ricerca pone le sue fondamenta nella tradizione delle Digital Scholarly Editions (DSE) [24], nella sua estensione come edizione semantica (modello ne è l'edizione dei Quaderni di Bufalini, descritta in [14]) e nei linguaggi di rappresentazione digitale XML/TEI e ontologica [38]. Essa inoltre si appoggia alla pluriennale esperienza del team di ricerca (cfr. [9], [11], [20]). Il modello di sviluppo seguito – modello CHROMA – è stato sperimentato all'interno del gruppo di ricerca CHROMA (*Computational Humanities: Representation, Organization, Management, and Analysis*)⁹ già nell'ambito del progetto *Bellini Digital Correspondence*,¹⁰ con l'obiettivo di garantire l'interoperabilità all'interno della comunità scientifica, in conformità ai principi FAIR¹¹ [42] e TRUST¹² [15], e di assicurare la sostenibilità dello sviluppo di modelli computazionali e di strumenti per le DSE.

Il modello CHROMA offre un approccio integrato e tratta il testo come un oggetto multidimensionale complesso, attraverso un processo editoriale che prevede diverse fasi dalla creazione di surrogati digitali fino alla pubblicazione e analisi in ambiente web. Questo modello collaborativo, condiviso e collaborativo assicura la consistenza dei dati, utilizzando in tutte le fasi un repository dedicato GitHub di progetto.

L'ambiente web integrato sviluppato comprende l'organizzazione e la gestione delle digitalizzazioni, la libera fruizione di esse in un portale, la possibilità di accedere a percorsi *public* pensati *ad hoc* per diverse tipologie di utenti (lettori comuni, specialisti del settore, studenti).

In particolare il lavoro si può articolare in tre macro fasi successive, ciascuna con la propria metodologia, ma tra loro correlate in un'ottica ciclica:

- a) conservazione: acquisizione di scansioni delle fonti primarie (soprattutto riviste) secondo i protocolli stabiliti dal Piano Nazionale di Digitalizzazione, metadatazione, trascrizione semi-automatica con OCR e codifica in formato XML/TEI. L'attività di digitalizzazione e di OCR segue protocolli standard e condivisi dalla comunità DH, utilizzando strumenti e metodi open source (cfr. [8], [16]);
- b) valorizzazione e fruizione: creazione di un ambiente web con sistemi di visualizzazione delle immagini e delle relative trascrizioni, con interfacce user-friendly. In questa fase si prevedono la progettazione e lo sviluppo di un portale web che offra la visualizzazione delle fonti primarie, del motore di ricerca e dei percorsi di approfondimento, a partire dall'individuazione del pubblico di riferimento (studenti, studiosi e utenti generalisti); in un'ottica di *long term preservation*, in linea con i principi FAIR e TRUST, si depositeranno i dati in infrastrutture di ricerca quali CLARIN e/o OPERAS;
- c) analisi: estrazione di dati utili per l'analisi del corpus a livello micro e macrotestuale, mediante algoritmi di *Natural Language Processing* e *Text Mining*. Si approderà pertanto a uno studio dell'impatto del Verismo sulla cultura coeva attraverso l'estrazione di termini e parole-chiave, temi e motivi, mettendoli in relazione con le opere più rappresentative (letteratura primaria). Uno spazio significativo verrà dato anche alle analisi di tipo quantitativo, con modalità di visualizzazione dei dati e i tools di *Data Visualization*.

⁸ <https://www.pirandellonazionale.it/>

⁹ <https://chroma.cnr.it>. Il modello è illustrato più nel dettaglio in [37].

¹⁰ <https://bellinicornespondence.cnr.it/>

¹¹ *Findability, Accessibility, Interoperability, Reusability*.

¹² *Transparency, Responsibility, User focus, Sustainability, Technology*.

In ogni sottofase è possibile accedere alle informazioni con granularità sempre maggiore (da *search by metadata* fino a *philological, linguistic and semantic search*).

Lo sviluppo del sistema di gestione, indicizzazione e soprattutto il Search Engine sono progettati e implementati secondo le buone pratiche di ingegneria del software e dei modelli formali per la gestione di documenti testuali, partendo da recenti esperienze di manipolazione di documenti e testi di interesse storico-letterario ([17], [12], [23]).

3. Il corpus *RiVer*

Il ventennio di cultura letteraria italiana che, ai fini della presente ricerca, si è scelto di comprendere tra gli estremi cronologici del 1872 e del 1890 si caratterizza per un dibattito straordinariamente partecipato da parte di letterati, scrittori e giornalisti, che si confrontano e riflettono, per la prima volta su scala nazionale, intorno a questioni cruciali di tipo politico-economico e antropologico di una società ancora disomogenea e in via di formazione. È un momento di straordinaria vitalità culturale, che Benedetto Croce ha definito con l'illuminante espressione di «grande conversazione» ([13: 66]), e che si svolge prioritariamente attraverso le pagine dei periodici, le cui redazioni, più o meno longeve, popolano lo scenario culturale di tutte le ex capitali dei vecchi Stati preunitari.

Il tema centrale, non solo letterario, è la fedeltà al principio del vero in arte, la rappresentazione della realtà nella sua forma più cruda e naturale, che si esprime soprattutto nel 'coraggio' dell'artista di scandagliare i temi scabrosi della sofferenza umana nei suoi aspetti più degradanti, quelli della povertà materiale, psicologica e morale.

Nel 1871, in prossimità dell'inizio della *timeline* contemplata dal progetto (vd. par. 9), iniziava la storia del romanzo naturalista con il ciclo dei Rougon-Macquart di Émile Zola [30]: un'analisi della società francese ritratta nei suoi ambienti reali – dai mercati ai sobborghi operai, dai cenacoli degli artisti alle miniere e alle campagne – e nelle sue dinamiche quotidiane. Subito dopo in Italia Luigi Capuana pubblicava la raccolta di saggi intitolata *Il teatro italiano contemporaneo* (1872), aprendo la stagione del Verismo. Non a caso nel 1879 sarà ancora Capuana, con *Giacinta*, romanzo dedicato a Zola, ad inaugurare la stagione delle grandi opere veriste, anticipando, insieme a Navarro della Miraglia che nello stesso anno pubblicava *La Nana*, il Verga maggiore; quest'ultimo, partendo dalla pubblicazione di *Nedda* nel 1874, chiudeva nel 1889, con il *Mastro Don Gesualdo*, la stagione più produttiva e vitale del Verismo.

Le opere letterarie dei veristi "maggiori" (Verga, Capuana e De Roberto), ma anche quelle che negli stessi anni andavano pubblicando i cosiddetti "minori" (Onufrio, Navarro della Miraglia, Ragusa Moleti, Chelli, Scarfoglio, Zena, Pratesi), coprono un arco cronologico lungo il quale si snoda una produzione parallela di saggi, recensioni e contributi critici, che a volte entra direttamente in dialogo con le opere letterarie, mentre in altre converge insieme a queste in un'analisi della realtà postunitaria focalizzata soprattutto sulle pesanti tare economiche e culturali del Mezzogiorno, che creavano un enorme divario tra la qualità della vita delle masse contadine meridionali e quella degli abitanti del Nord Italia.

Il corpus di testi al centro dell'analisi del progetto COVerLeSS è strutturato intorno a una nuova prospettiva di indagine, che mira a rilevare, verificare e descrivere l'esistenza di un lessico peculiare della critica letteraria verista. Su questa base la selezione dei testi si è soffermata innanzitutto sulle recensioni contemporanee alle opere più rappresentative del Verismo dei due maggiori rappresentanti di questa corrente letteraria, Giovanni Verga e Luigi Capuana, con la prospettiva di estendere l'ambito della ricerca alle recensioni coeve delle opere dei cosiddetti veristi minori, com'è il caso del citato romanzo di Emanuele Navarro della Miraglia, autore

siciliano reduce da una lunga permanenza parigina, che lo aveva reso uno dei divulgatori della cultura letteraria francese in Italia (cfr. [34:57-87]).

Allo stato attuale sono state acquisite al copus le recensioni alle opere di Verga *Nedda* (pubblicata sul numero del 15 giugno 1874 della «Rivista italiana di scienze, lettere e arti» e in volume presso l'editore milanese Brigola con il sottotitolo *Bozzetto siciliano*), *Vita dei campi* (Milano, Treves, 1880), *I Malavoglia* (Milano, Treves, 1881), *Novelle Rusticane* (Torino, Casanova, 1883), *Per le vie* (Milano, Treves, 1883), *Cavalleria rusticana* (versione teatrale; Torino, Casanova, 1884) e *Mastro-don Gesualdo* (pubblicato da Treves nel 1889 – ma datato 1890 –, e che aveva già visto la luce a puntate sulla rivista «Nuova Antologia» da luglio a dicembre 1888), attingendo alle pagine delle maggiori testate dell'epoca, quali il «Fanfulla», la «Nuova Illustrazione Universale» di Treves, la «Rivista Minima» di Ghislanzoni, pubblicata da Ricordi, il «Corriere della Sera» e il «Fanfulla della domenica», sui quali lo stesso Capuana recensisce prima *Vita dei campi* e poi *I Malavoglia* (cfr. [33]). Le firme sono quelle di Ferdinando Martini, sotto lo pseudonimo di Fantasio, Emilio Treves, con quello di Bibliofilo, e poi Salvatore Farina, Felice Camerani e altri, che avevano o avrebbero condiviso con Verga, oltre che in alcuni casi l'amicizia, riflessioni ed esperienze letterarie ed editoriali ampiamente testimoniate dall'epistolario dell'autore.

Tra le recensioni alle opere di Capuana, di prossima acquisizione, sono state selezionate quelle al romanzo *Giacinta* (sia nella prima edizione del 1879, pubblicata da Brigola, che nella nuova redazione stampata a Catania da Giannotta nel 1885-1886) e alla raccolta di racconti *Homoi!* (nelle due edizioni, Brigola del 1883 e Treves del 1888); le recensioni sono a cura di Federico Verdinois («Corriere del mattino»), Enrico Panzacchi («Nuova antologia»), Salvatore Farina («Rivista minima»), Edoardo Scarfoglio («Domenica letteraria») e dei conterranei Giorgio Arcoleo («Fanfulla della domenica»), Giovanni Alfredo Cesareo («Cronaca bizantina») e Giuseppe Pipitone Federico («Scena illustrata») (cfr. [31]).

Opere letterarie e recensioni ad esse collegate sono precedute e accompagnate, sul piano cronologico, da alcuni scritti di tipo teorico di cui sono autori, tra i primi, Francesco De Sanctis e Luigi Capuana, poi emulati da molti. Questa produzione saggistica ha contribuito, insieme alle influenze culturali di provenienza francese, alla formazione di una coscienza letteraria nuova, improntata alla ricerca e alla rappresentazione del 'vero' oltre ogni moralismo, come espressione di una compiuta maturazione del percorso dell'arte nella contemporaneità. L'ipotesi al vaglio del gruppo di ricerca è che i saggi di De Sanctis e di Capuana, insieme a una nuova esigenza-consapevolezza artistica, abbiano contribuito a fondare una sorta di lessico specialistico, che ha influenzato, consolidandosi e arricchendosi via via di termini e di immagini nuovi, il linguaggio della critica del (e in qualche caso 'al') realismo-verismo italiano. I testi selezionati per costituire questa seconda sezione del corpus sono i saggi di Luigi Capuana raccolti nei volumi *Il teatro italiano contemporaneo* (Palermo, Pedone Lauriel, 1872), *Studi sulla letteratura contemporanea (Prima serie, Milano, Brigola, 1880; Seconda serie, Catania, Giannotta, 1882) e Per l'arte* (Giannotta, 1885). Di Francesco De Sanctis i saggi *Il principio del realismo* (pubblicato sulla «Nuova Antologia» a gennaio 1876) e *Studio sopra Emilio Zola* (1877, pubblicato a puntate sul quotidiano napoletano «Roma» da giugno a dicembre 1877), inclusi entrambi nella seconda edizione dei *Nuovi saggi critici* (Napoli, Morano, 1879), la conferenza *Zola e l'«Assommoir»*, tenuta nel 1879 presso il Circolo filologico di Napoli e pubblicata nello stesso anno da Treves, e *Il darwinismo nell'arte*, conferenza tenuta presso il medesimo circolo nel 1883 e stampata nello stesso anno (Napoli, Stab. Tip. dei Classici Italiani). A questa prima selezione si aggiungeranno i saggi di Federico De Roberto raccolti nel volume *Arabeschi* (Catania, Giannotta, 1893) e gli interventi critici sul realismo di Enrico Onufrio e Girolamo Ragusa Moleti, oltre ai lavori di un critico di scuola desantisiana come Giorgio Arcoleo.

Per completare l'analisi sulla letteratura secondaria di ambito realista-verista abbiamo scelto di accogliere nel corpus le serie integrali di quattro delle riviste dell'epoca impegnate nella diffusione dei principi del naturalismo francese e soprattutto nella teorizzazione di una versione italiana di

esso. I criteri di selezione di queste riviste mirano a restituire all'attenzione di studiosi e lettori le esperienze di alcune testate di difficile reperibilità e del tutto assenti dal web; è il caso di tre di esse: «La Farfalla» di Angelo Sommaruga,¹³ «La Fronda»¹⁴ di Emmanuele Navarro della Miraglia e «Il Momento Letterario-Artistico-Sociale»¹⁵ di Giuseppe Pipitone Federico. A queste testate è stata integrata la «Rassegna Settimanale di politica, scienze, lettere ed arti»,¹⁶ fondata da Leopoldo Franchetti e Sidney Sonnino, per via della forte vicinanza tra la linea editoriale della rivista e l'interesse dei veristi verso l'oggetto prioritario della loro rappresentazione artistica, ossia la condizione esistenziale delle masse popolari del meridione. Molti dei temi che si ritrovano nella letteratura verista si possono mettere in diretta relazione con gli articoli pubblicati nei fascicoli della «Rassegna Settimanale» (l'istruzione pubblica, il costo degli alimenti, le tasse, la malaria, per nominarne alcuni) (cfr. [4]).

In fase di realizzazione, sono emerse difficoltà legate al reperimento di alcuni periodici, in particolare «La Farfalla», la cui consistenza, dopo ripetute verifiche con la Biblioteca Braidense (che possiede la collezione integrale della rivista), è risultata significativamente superiore rispetto a quanto indicato nelle bibliografie di riferimento. Tale scostamento ha determinato un rallentamento sia nelle tempistiche di digitalizzazione a cura della biblioteca, sia nel processo di elaborazione OCR da parte del gruppo di ricerca.

4. Elementi di codifica dei testi

Il progetto adotta un modello di codifica basato sul vocabolario e le linee guida definite dal consorzio TEI¹⁷ in linguaggio XML, con un'architettura che prevede la creazione di file XML per ciascuna unità testuale e diversificati per ciascuna tipologia di testo (saggi, recensioni e articoli di rivista). La struttura dei documenti segue la consueta suddivisione gerarchica dello standard XML/TEI, con una sezione <teiHeader> dedicata ai metadati e una sezione <text> per il contenuto testuale. Nei metadati vengono registrati – oltre a informazioni di carattere generale come il titolo del contributo, l'attribuzione di responsabilità del progetto e della codifica, la data di creazione e revisione – gli elementi che caratterizzano il corpus di COVerLeSS (ad esempio, l'identificazione dell'opera recensita o citata nel caso di articoli di recensione). Particolare attenzione è dedicata alla normalizzazione dei dati bibliografici, collegati a una lista esterna di opere e autori mediante identificatori univoci. Anche le entità nominate (persone, luoghi, opere, organizzazioni) vengono codificate attraverso riferimenti incrociati a liste esterne, garantendo così un'integrazione strutturata delle informazioni senza sovrabbondanza o ripetizioni nella codifica.

La codifica del testo propriamente detto prevede – insieme all'annotazione dettagliata delle entità nominate, marcate utilizzando elementi specifici come <persName> per i nomi di persona,

¹³ Dopo una prima parentesi cagliaritano (1876-1877), la redazione venne trasferita dal fondatore, Angelo Sommaruga, a Milano, dove la rivista venne pubblicata fino al 1883 (cfr. [7]).

¹⁴ Venne pubblicata a Firenze per soli 7 numeri, tra gennaio e febbraio del 1880 (cfr. [34]: 89, 143); ospitò articoli, recensioni e opere dei maggiori letterati italiani, tra i quali Capuana e Verga.

¹⁵ Venne pubblicato a Palermo dal 1883 al 1885 in 72 numeri (cfr. [35]).

¹⁶ Pubblicata per 213 numeri dal 1878 al 1882; riuniva in un progetto finalmente unitario il dibattito sui temi politici, economici, sociali e letterari, dedicando particolare attenzione alle problematiche delle popolazioni del Mezzogiorno d'Italia (<https://rassenasettimanale.animi.it/>).

¹⁷ Text Encoding Initiative, <https://tei-c.org/>

<placeName> per i luoghi e <title> per le opere – la rappresentazione degli elementi lessicali del Verismo rilevanti da un punto di vista stilistico e linguistico, evidenziati mediante l'elemento <term>, per i quali si rimanda alla descrizione del paragrafo successivo.

Per garantire una rappresentazione accurata della struttura testuale originale, il modello prevede anche la marcatura della struttura della pagina, con l'uso di <pb> e <cb> per le interruzioni rispettivamente di pagina e di colonna, oltre alla segmentazione dei paragrafi tramite <p>. Il sistema di codifica è concepito per essere compatibile con TEI Publisher, offrendo così una visualizzazione interattiva dei testi annotati.

5. Un dizionario-repertorio tematico: *Verbum*

La codifica strutturale e quella relativa alle entità nominate costituiscono la base di un'annotazione specifica, comprendente lessemi, combinazioni sintagmatiche, qualche volta strutture fraseologiche, volte a costituire una sorta di dizionario tematico nel quale trovano spazio gli elementi caratterizzanti della dimensione sociale, della poetica, della costruzione narrativa, dello stile della letteratura verista. Non si tratta di approdare, in questo caso specifico, ad un vocabolario lemmatizzato, fondato sulla singola unità lessicale classificata grammaticalmente, quanto di isolare un campionario rappresentativo della terminologia caratterizzante il discorso critico, la recensione, la polemica letteraria, verificando il tasso di 'tecnicitá' con cui quella terminologia ricorre nelle diverse testualità esposte nel corpus. Piuttosto che di un dizionario 'reale', abbiamo preferito caratterizzare *Verbum* come un dizionario 'ideale', nel quale il lemma base, la singola parola, si trovano spesso all'interno di combinazioni sintagmatiche (ad es. sostantivo + aggettivo, verbo + complementi, locuzioni modali) che consentono di circostanziare nelle sue articolazioni il linguaggio specifico della letteratura secondaria. Ad esempio, l'utente potrà ricercare la parola *bozzetto*, trovando sia le attestazioni del singolo lessema sia alcune *iuncturae* fortemente codificate (e presenti sia nei sottotitoli di opere che in alcune dichiarazioni epistolari di Verga), come *bozzetto siciliano* e *bozzetto marinairesco*. E ancora, accanto a *colore*, *colorito*, si potrà rinvenire un sintagma chiave come *colore locale*, con la variante *colorito locale*.

La scelta di creare queste 'entrate' del dizionario ha mirato dunque a contemperare la rappresentatività dei lemmi base con la salvaguardia di una serie di sfumature semantiche che emergono dalla restituzione unitaria di sintagmi, strutture fraseologiche, locuzioni attinenti al repertorio della narrativa e critica del Verismo. Un esempio può essere dato da questa 'mappa' di termini e strutture sintagmatiche in Tabella 1 – che fanno capo al cluster @arti_visive.

Bozzetto	bozzetto dal vero
	bozzetto siciliano
Fedele	quadro fedele
	fedele pittura
Colorito	efficacia di colorito
	colorito del paesaggio
	colorito locale

	intonazione del colorito
	vivacità di colorito
Ritrarre	ritrarre la vita
	ritrarre i dialoghi
	ritrarre sentimenti
Quadro/quadretto	quadro di costumi
	quadri veri
	quadretti locali
	quadretti della vita siciliana
	quadrettino di genere campestre
	quadri variati
Dipingere	dipinge dal vero
Colore	delineare a colori vivi i contorni
	tratteggiare coi colori più smaglianti della sua tavolozza
	colore del vero
	colori vivi
	vivezza di colore
	ricchezza del colore

Tabella 1. Esempi di termini e strutture sintagmatiche afferenti al cluster “arti_visive” raggruppati per parole dominati.

Si è trattato inoltre, in virtù della natura trasversale delle tematiche affrontate attraverso i testi spogliati, di etichettare in modo univoco e riconoscibile ambiti lessicali abbastanza ampi. Per questo è opportuno descrivere brevemente la scelta fatta in ordine alla individuazione dei campi concettuali ai quali sono stati assegnati i vari insiemi di lessico riconducibili a *Verbum*.

L’ambito più ampio è quello ascrivibile alla dimensione letteraria e sociale del Verismo; in questo caso, partendo dal tag <term> – che definisce una parola singola, un’espressione composta o

una designazione metaforica considerata come termine tecnico –, abbiamo assegnato, dopo l'attributo-base @type="verbum", una serie di @subtype corrispondenti ad altrettanti ambiti concettuali che puntano a categorizzare con maggior dovizia di particolari il termine/cluster. Ognuno di essi ha dietro almeno un testo 'generativo', ossia una lettera, una prefazione, una nota d'autore.

a) Un primo ambito – caratterizzato dal @subtype="rappresentazione" – comprende tutti quei termini che si riferiscono alla descrizione degli ambienti, alla raffigurazione del personaggio anche attraverso parole e gesti, in una dimensione che è quella della resa oggettiva la quale permette di evitare, per quanto attiene ai personaggi, ciò che Verga chiamava «la descrizione, lo studio, il profilo» (lettera a Cameroni del 19 marzo 1881; vd. [41:14]). Si ritroveranno in questo campo termini come *descrizione, efficacia, evidenza, naturalezza, semplicità*.

b) Un secondo ambito – identificato dal @subtype="arti_visive" (vd. Tabella 1) –, di ampia ricorsività nei giudizi critici e nelle formulazioni di poetica, riguarda la terminologia pittorica e fotografica in gran parte mutuata dalla prefazione ai *Malavoglia*: «Basta lasciare al quadro le sue tinte schiette e tranquille, e il suo disegno semplice»; «i colori cominceranno ad essere più vivaci, e il disegno a farsi più ampio e variato»; «Persino il linguaggio tende ad individualizzarsi, ad arricchirsi di tutte le mezze tinte dei mezzi sentimenti»; «Perché la riproduzione artistica di cotesti quadri sia esatta» ([39:11,12]). Spesso le recensioni tendono nel linguaggio critico-descrittivo a una forma di mimetismo con i testi analizzati, come nel caso di quella pubblicata da Capuana su *I Malavoglia* nel «Fanfulla della domenica» del 29 maggio 1881, in cui si rintracciano lessemi ed espressioni come *dipinte con colori caldissimi, pennellate, gran quadro, stupendi bozzetti, crudezze di toni, di mezze tinte, di sfumature*.

c) Un ulteriore campo che si è voluto caratterizzare autonomamente – con @subtype="oggettività_impersonalità" – riguarda le soluzioni narrative mediante le quali l'autore dissimula la sua presenza nel racconto, attraverso la delega della voce o la 'regressione' (è il caso di Verga nei *Malavoglia*) nello status di un narratore anonimo popolare. Anche qui c'è una matrice del discorso critico, ed è la lettera-prefazione alla novella verghiana *L'amante di Gramigna*, dedicata a Salvatore Farina:

[...] la mano dell'artista rimarrà assolutamente invisibile, e avrà l'impronta dell'avvenimento reale, e l'opera d'arte sembrerà essersi fatta da sé, aver maturato ed esser sorta spontanea come un fatto naturale, senza serbare alcun punto di contatto col suo autore; che essa non serbi nelle sue forme viventi alcuna impronta della mente in cui germogliò, alcuna ombra dell'occhio che la intravvide, alcuna traccia delle labbra che ne mormorarono le prime parole come il fiat creatore; ch'essa stia per ragion propria, pel solo fatto che è come dev'essere, ed è necessario che sia, palpitante di vita ed immutabile al pari di una statua di bronzo, di cui l'autore abbia il coraggio divino di eclissarsi e sparire nella sua opera immortale [40:92]

Molti sarebbero i documenti epistolari da aggiungere a questa 'soglia' testuale così importante; basti ricordare la lettera a Francesco Torraca del 12 maggio 1881, scritta in ringraziamento della sua recensione al romanzo, apparsa sulla rivista «Il Diritto» il 9 maggio. Qui Verga afferma che il suo ideale artistico «è che l'autore s'immedesima talmente nell'opera d'arte da scomparire in essa». E aggiunge: «Vorrei quasi che un romanzo arrivasse a non portare il nome del suo autore; si affermasse da sé, come vivente per un organismo proprio e necessario» (corsivo questo dell'autore) [28:249]. Ma soprattutto egli è fermo nell'idea che «non possa sussistere un momento l'illusione della completa immedesimazione col soggetto senza dare un'uniforme intonazione a tutta l'opera, senza eclissare completamente lo scrittore» [28:250].

d) Col @subtype="studio_osservazione" si è voluto evidenziare un tratto caratterizzante, ossia quello del romanzo come «studio sociale» (così Torraca sui *Malavoglia*) e della funzione del narratore che, con sguardo distaccato ma al contempo partecipe, si interessa «dei deboli che restano per via», dei vinti di oggi e di domani. Il campo semantico dell'osservazione e dell'osservatore è strettamente connesso con quello dell'oggettività e impersonalità. Anche in questo caso abbiamo valorizzato, ad esempio per *studio*, una serie di concrezioni sintagmatiche fortemente connotate: *studio sociale*, *studio biologico-sociale*, *studio coscienzioso*, *studio obiettivo*, *studio psicologico*. Ipotesto ideale è ancora una volta la prefazione al romanzo dell'81, dove l'artista «osservatore» è colui che meglio interpreta il senso della «fiumana del progresso»:

Chi *osserva* questo spettacolo non ha il diritto di giudicarlo; è già molto se riesce a *trarsi un istante fuori* del campo della lotta per *studiarla senza passione*, e rendere la scena nettamente, coi colori adatti, tale da dare la *rappresentazione della realtà* com'è stata, o come avrebbe dovuto essere [39:12]

e) Col @subtype="vero_verità" ci si riferisce sia al modo di rappresentare la realtà (in questo senso viene adoperata la locuzione *dal vero*, ad es. in *studi dal vero*, ma non mancano sintagmi come *riproduzione del vero* e *riproduzione verista*), sia a una caratteristica di fatti e personaggi (ad es. *fatti veri*, *tipi veri*), sia infine all'obiettivo della rappresentazione stessa, ossia il *vero* e la *verità*. A questo si connette anche la dimensione della *verosimiglianza*.

f) Infine, si è individuato un ambito concettuale riferito alla dimensione individuale e collettiva delle dinamiche socio-economiche rappresentate dagli scrittori (@subtype="evoluzione_sociale"). Un aspetto fortemente connesso alla tradizione del romanzo ciclico, che implica anche tutto quell'universo di razionalità e istinto, calcolo e appetito, ambizione e passioni che rappresentano il grande affresco della lotta per la vita. Oltre alla prefazione ai *Malavoglia*, che ci pare superfluo richiamare ancora una volta, ricordiamo la lettera di Verga all'amico avvocato Salvatore Paola Verdura del 21 aprile 1878, in cui viene rappresentato in forma di «fantasmagoria» il grande progetto dei *Vinti* (qui ancora la *Marea*) [abbiamo distinto i corsivi nostri dalle parole sottolineate da Verga]:

Ho in mente un lavoro che mi sembra bello, e grande, una specie di fantasmagoria della lotta per la vita che si stende dal cenciaiuolo al ministro ed all'artista, e assume tutte le forme, dall'*ambizione* all'*avidità del guadagno* e si presta a mille rappresentazioni del grottesco umano, *lotta providenziale* che guida l'umanità per mezzo e attraverso a tutti gli *appetiti*, alti e bassi, alla conquista della verità. Insomma cogliere il lato drammatico e ridicolo, o comico di tutte le *fisionomie sociali*, ognuna colla sua caratteristica, negli sforzi che fanno per andare avanti, in mezzo a quest'onda immensa che è spinta dai *bisogni più volgari*, o dall'*avidità della scienza* ad andare avanti, incessantemente, pena la caduta e la vita, pei deboli, e i maldestri. [...]

Ciascun romanzo avrà una *fisionomia speciale*, resa con mezzi adatti – il realismo io l'intendo così, come la schietta ed evidente manifestazione dell'*osservazione coscienziosa*, la sincerità dell'arte in una parola – potrà prendere un lato della *fisionomia della vita italiana moderna*, a partire dalle *classi infime*, dove la *lotta* è *limitata al pane quotidiano*, come nel Padron Ntoni, e a finire nelle varie ispirazioni, nelle *ideali avidità* dell'uomo di lusso (un segreto) passando per le *avidità basse* sin alle *vanità* del Mastro don

Gesualdo, rappresentante della vita di provincia, all'*ambizione* di un deputato [...] [1:260-261]

Si è stabilito di mantenere, insieme a questo lessico specifico, anche una serie di termini collaterali, che formano la costellazione storico-letteraria, di poetica e stile che ruota attorno ai termini chiave di *Verbum*. Così, è stato assegnato al @type="forma_letteraria" una serie di @subtype che identificano le varie correnti e movimenti letterari ("ismi"), gli elementi del genere letterario, comprendente ad es. novella, romanzo, dramma ("genere_letterario"), le componenti legate all'analisi del racconto, dall'intreccio al dialogo alla scena ("narrazione"), nonché le scelte di poetica e le soluzioni autoriali specificamente connesse alla dimensione del narrare ("autorialità_narratività"). E infine, attraverso gli attributi associati a "stile" ("lingua", "lessico", "sintassi", "registro") si è tentato di circostanziare i riferimenti alla lingua e allo stile che soprattutto nelle recensioni non appaiono in modo sistematico, ma abbastanza connotato nei casi in cui assumono un rilievo degno di annotazione.

Il lettore avrà comunque la possibilità, grazie all'uso del tag <seg> che permette l'annidamento del termine marcato al contesto, di poter 'navigare' tra le sfumature semantiche di ogni voce e costruirsi un proprio percorso attraverso autori, testi, poetiche.

Naturalmente è stato necessario operare una serie di distinguo, non sempre motivati onomasiologicamente, ma correlati al significato che il termine assume nel contesto di appartenenza. Così troveremo *racconto* sia sotto @subtype="genere_letterario", dove esso appare quale sinonimo di *novella*, sia sotto @subtype="autorialità_narratività", laddove il termine è inteso semplicemente come atto del narrare. E ancora, abbiamo posto sotto 'contenitori' semantici differenti i termini afferenti a *vero/verità* e a *reale/realtà*, sulla base di indicazioni provenienti dalla stessa pubblicistica del tempo (anche quella non compresa nelle riviste indicizzate nel progetto), particolarmente attenta alla distinzione tra *verismo* e *realismo* [3]. In Figura 1 si riportano le percentuali dei campi semantici riconducibili a realtà/verità.

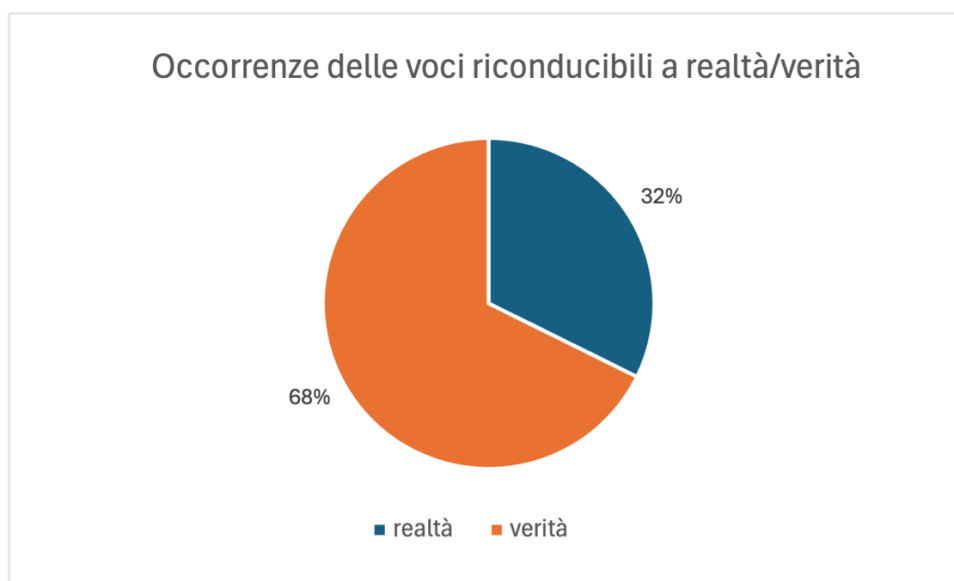


Figura 1. Il grafico mostra le percentuali dei campi semantici (singoli lemmi e strutture sintagmatiche) riconducibili a realtà/verità.

6. Gestione e trattamento dei files di codifica

Una parte peculiare del progetto COVERLeSS è quella relativa al processo di creazione dei file XML contenenti le codifiche TEI dei documenti testuali di interesse, attività che vede (e richiede) il coinvolgimento e una forte collaborazione di e tra diverse figure implicate nel progetto, anche in termini di supervisione delle varie fasi di sviluppo. Al fine di armonizzare questo processo, e di agevolare, quindi, la creazione dei file XML, e di perfezionarli anche in termini di uniformità e di aderenza agli standard informatici di codifica testuale (come quelli, in particolare, specificatamente richiesti nel web e accettati quindi come validi in pagine HTML), sono stati sviluppati degli appositi tool di *string processing* (implementati nel linguaggio di programmazione *JAVA*) per effettuare l'analisi dei file XML via via creati. Questi tool, oltre a riportare indicazioni circa eventuali *anomalie* (sintattiche) riscontrate, permettono, laddove necessario e quando possibile, di ricreare in maniera automatica (ossia, senza interventi manuali) copie aggiornate *rettificate* dei file XML che vengono così "normalizzati". In termini molto generali, questo processo di normalizzazione dei file XML consiste nell'estrarre preventivamente i contenuti dei file in forma di stringhe di caratteri (*documenti XML*), e quindi nel rimpiazzare (possibilmente attraverso più revisioni consecutive, qualora i casi lo richiedano) all'interno dei documenti XML tutte le occorrenze di ogni (sotto)stringa *S*, la cui presenza è ritenuta errata (cioè rappresenta un'anomalia), con un'occorrenza di un'opportuna *stringa normalizzante* *R*, corrispondente a *S*, rettificando così l'anomalia dovuta, appunto, alla presenza di *S*. Un caso significativo, abbastanza frequente, coinvolge ad esempio il rimpiazzamento di ogni stringa *S* (di lunghezza massimale) costituita (esclusivamente) da una sequenza di uno o più *caratteri di whitespace* qualsiasi (tra cui i comuni caratteri `\t`—*horizontal tab*—, `\r`—*carriage return*—e `\n`—*line feed*—¹⁸) con la semplice stringa " ", ossia la stringa che consiste di una singola occorrenza del carattere di *blank space*. Similmente, la presenza di certi *caratteri di controllo*, quali, in particolare, i caratteri *SOS*—*Start Of String*—e *ST*—*String Terminator*—,¹⁹ riscontrata in alcuni file XML, costituisce un'anomalia, che viene rettificata semplicemente eliminando questi caratteri.²⁰ Si osservi che in ambienti di sviluppo altamente collaborativi, come quello del progetto COVERLeSS, la presenza di anomalie sintattiche nei file XML dovute, come nei casi precedenti, alla presenza di più caratteri consecutivi di whitespace e/o dei caratteri *SOS* e *ST*, non è per niente rara. Spesso, infatti, tali anomalie vengono introdotte nei file XML senza averne cognizione, involontariamente, sia (ad esempio) a causa dell'utilizzo di editor XML che producono automaticamente indentazioni dei tag via via inseriti al fine di facilitare la visualizzazione su schermo dei documenti correntemente editati, sia anche come conseguenza di operazioni di copia e incolla a partire da sorgenti testuali quali documenti WORD o PDF, operazioni queste comunemente effettuate. La rettifica di anomalie come queste è particolarmente utile, tra l'altro, cosicché, ad esempio, quando vengono estratti i contenuti dei tag TEI da visualizzare (anche dinamicamente) online nelle pagine web (come i tag che descrivono le entità nominate, ovvero il tag `<term>` relativo al lessico), risulta più agevole formattare questi contenuti secondo i modelli di visualizzazione prestabiliti, scongiurando anche eventuali errori indesiderati di *line break*, senza necessità di effettuare ulteriori controlli sui caratteri dei file XML.²¹

¹⁸ https://en.wikipedia.org/wiki/Whitespace_character

¹⁹ https://it.wikipedia.org/wiki/Carattere_di_controllo

²⁰ Si noti che ciò corrisponde formalmente a rimpiazzare questi caratteri con la *stringa vuota*.

²¹ Sottolineiamo che le anomalie sintattiche presenti e riscontrate all'interno file XML non si limitano soltanto ai casi descritti prima che coinvolgono i caratteri di whitespace e di controllo. Un altro caso interessante, abbastanza comune, riguarda la presenza di caratteri formalmente differenti

Dopo la rettifica delle anomalie sintattiche nei file XML, viene (parzialmente) testata (in maniera semi-automatizzata, attraverso l'uso di appositi tool sviluppati ad hoc) l'aderenza delle annotazioni TEI contenute nei file al modello di codifica prestabilito. In linea di principio, questa fase funziona come segue. Ogni documento XML viene inteso come la rappresentazione formale di una (istanza di una) *entità* (nella fattispecie, il reale documento di testo da codificare) caratterizzata da diversi *attributi* (es., nome dell'autore del documento di testo, data di creazione del documento, luogo di creazione, etc.), dove ogni attributo può assumere (per ogni data istanza dell'entità) uno (ed uno solo) tra i valori di un fissato insieme di possibili (tipi di) valori, designati per quell'attributo; ciascuno di questi attributi A corrisponde, all'interno della *struttura ad albero* T del documento XML, a un determinato nodo XML N_A (cioè un tag TEI)—*nodo-attributo*—il cui contenuto rappresenta in forma di stringa il valore dell'attributo A , e individuato, questo nodo, dal percorso (unico) che dalla radice dell'albero T conduce ad esso. Durante il controllo di aderenza alle annotazioni TEI, i documenti XML vengono quindi testati in termini di fedeltà alla suddetta rappresentazione formale di entità che essi stessi sono appunto intesi costituire: si estrae ogni singolo nodo-attributo N_A , utilizzando il percorso che lo individua all'interno dell'albero T , e se ne verifica formalmente l'ammissibilità del contenuto come rappresentazione di un possibile valore che può assumere l'attributo A .

7. Analisi semantica di *Verbum*

L'analisi automatica dei dati ha fatto sì che emergesse una particolarità legata all'intersezione non vuota degli elementi di uno stesso campo semantico in cluster diversi. Questo dato, frutto non di una scelta a priori, è interessante da indagare, poiché fa emergere relazioni sia tra i singoli termini che tra le varie combinazioni sintagmatiche (in particolare quelle di sostantivo + aggettivo, verbo + complementi, locuzioni modali). Si riporta a titolo d'esempio in Figura 2 la terminologia relativa al 'vivere' con le sue sfumature caratterizzanti del linguaggio di dominio etichettate con l'elemento <term>, estratta dalle recensioni finora analizzate.

tra loro, aventi cioè un diverso valore di *code point* (https://en.wikipedia.org/wiki/Code_point#In_character_encoding), ma utilizzati in maniera interscambiabile col medesimo ruolo funzionale, errore probabilmente commesso a causa della difficile distinguibilità grafica di questi caratteri che ne favorisce appunto la confusione. Un esempio specifico è rappresentato dai caratteri con valori di code point 8211 e 8212, usati entrambi per indicare segni di inciso all'interno di frasi: in realtà il primo dovrebbe essere usato per indicare invece dei range. Poiché nel caso del progetto COVerLeSS, come specificatore di range (ad es. nel caso di date) si adotta il carattere avente code point 45 (il comune carattere del "segno meno"), la presenza del carattere di code point 8211 viene pertanto considerata un'anomalia, che viene rettificata rimpiazzando questo carattere con il carattere di code point 8212.

| al vivo
 | colori più vivi
 | colori vivi
 | convenzioni del vivere sociale
 | creature vive
 | forme viventi
 | impressione così viva e così immediata
 | lingua viva
 | lingua viva e parlata
 | persona viva
 | persone vive
 | rendere vive
 | viva
 | vivacità
 | vivacità di colorito
 | vivamente
 | vive
 | viventi
 | viventi creature
 | vivezza
 | vivezza di colore
 | vivi
 | vivido
 | vivissime
 | vivo
 | vivo e vero

Figura 2. Termini e strutture sintagmatiche riconducibili al campo semantico del ‘vivere’

Infatti, la Tabella 2 mostra come i termini riportati nella Figura 2 afferenti al campo semantico del ‘vivere’ siano presenti in cluster differenti a seconda del contesto di riferimento.

Campo semantico del vivere		
arti_visive	6	verbum
evoluzione_sociale	1	
rappresentazione	35	
vero_verità	2	
lingua	5	stile

Tabella 2. Occorrenze dei termini del campo semantico del ‘vivere’ nei vari cluster

Questi scenari suggeriscono l’idea di poter definire un’ontologia formale in cui ogni termine viene visto come un individuo (in senso ontologico, appunto) relazionato con i diversi valori di @type e, in particolar modo, @subtype, che ne delineano i contesti, questi ultimi anch’essi trattati come individui governati da specifiche relazioni gerarchiche di inclusione (ad es., ogni termine con un certo @subtype X, sarà anche un termine con un certo altro @subtype Y) e/o assiomi di disgiunzione (ad es., nessun termine con un certo @subtype V può essere un termine con un

certo altro @subtype W) che ne descrivono proprietà semantiche. Inoltre, si potrebbe introdurre una “ragionevole” relazione di similarità (ad es., sintattica) tra termini (ad es. i termini vivo e vivacità) in modo da partizionare questi ultimi in singoli cluster (di termini simili) e testare quindi in maniera automatica, tramite reasoner, le differenze semantiche di termini appartenenti ad uno stesso cluster rispetto ai valori di @(sub)type ad essi attribuiti (e quindi, ad es., il fatto che due termini, benché simili, occorrono in realtà in contesti diversi). È infatti in corso un’analisi circa la prospettiva di una simile modellazione ontologica. A titolo esemplificativo, in Fig. 3 si riporta la rappresentazione in OWL Functional Syntax (delle parti salienti) di una prima tentata bozza dell’ontologia formale *OntoTerm* che realizza questa modellizzazione.²² Le object property *hasType* e *hasSubtype* implementano, rispettivamente, le relazioni secondo cui un termine ha attribuito un certo type, ovvero un certo subtype; mentre la relazione di similarità tra due termini è descritta dalla object property *isSimilarTo*. Quest’ultima relazione viene asserita come simmetrica (Axiom_1) e (localmente) riflessiva (Axiom_2) (caratteristiche minimali di relazioni di similarità). Gli assiomi Axiom_3 e Axiom_4 stabiliscono le proprietà di inclusione e disgiunzione tra specifici valori di @subtype.

```

ObjectPropertyDomain(<OntoTerm#hasSubtype> <OntoTerm#TERM>)
ObjectPropertyRange(<OntoTerm#hasSubtype> <OntoTerm#SUBTYPE>)
ObjectPropertyDomain(<OntoTerm#hasType> <OntoTerm#TERM>)
ObjectPropertyRange(<OntoTerm#hasType> <OntoTerm#TYPE>)
ObjectPropertyDomain(<OntoTerm#isSimilarTo> <OntoTerm#TERM>)
ObjectPropertyRange(<OntoTerm#isSimilarTo> <OntoTerm#TERM>)
ClassAssertion(<OntoTerm#SUBTYPE> <OntoTerm#X>)
ClassAssertion(<OntoTerm#SUBTYPE> <OntoTerm#Y>)
ClassAssertion(<OntoTerm#SUBTYPE> <OntoTerm#V>)
ClassAssertion(<OntoTerm#SUBTYPE> <OntoTerm#W>)
# Axiom_1
  SymmetricObjectProperty(<OntoTerm#isSimilarTo>)
# Axiom_2
  SubClassOf(<OntoTerm#TERM> ObjectHasSelf(<OntoTerm#isSimilarTo>))
# Axiom_3
  SubClassOf(ObjectSomeValuesFrom(<OntoTerm#hasSubtype> ObjectOneOf(<OntoTerm#X>))
ObjectSomeValuesFrom(<OntoTerm#hasSubtype> ObjectOneOf(<OntoTerm#Y>)))
# Axiom_4
  SubClassOf(ObjectIntersectionOf(ObjectSomeValuesFrom(<OntoTerm#hasSubtype>
ObjectOneOf(<OntoTerm#V>)) ObjectSomeValuesFrom(<OntoTerm#hasSubtype>
ObjectOneOf(<OntoTerm#W>))) owl:Nothing)

```

Figura 3. Rappresentazione in OWL Functional Syntax di una bozza dell’ontologia *OntoTerm*

Al fine di organizzare semanticamente al meglio la conoscenza di dominio e quindi le proprietà tra i termini e le strutture sintagmatiche sono stati annotati, allo stato attuale, i <term> in 57 recensioni alle maggiori opere di Giovanni Verga [32]. I testi sono stati trattati secondo lo standard XML/TEI. Il segmento di corpus codificato comprende 9 recensioni riguardanti la novella *Nedda* (1874), 12 il romanzo *I Malavoglia* (1881), 4 la novella *Pane Nero* (1882), 9 la raccolta *Vita dei Campi* (1880), 10 le *Novelle Rusticane* (1882-1883) e 13 il romanzo *Mastro-don Gesualdo* (1889).

²² Una descrizione del linguaggio OWL Functional Syntax è consultabile in <https://www.w3.org/TR/owl2-syntax/>.

Per quanto attiene al <term> @type="Verbum", centrale nella nostra analisi, l'estrazione dei dati è stata eseguita con il tool NormaTEI,²³ sviluppato dal gruppo CHROMA del CNR ISTC nell'ambito del progetto *BellinInRete*.²⁴ Questo tool permette di analizzare in modo complessivo tutti i file XML/TEI di un'edizione, e in esso è possibile effettuare sia ricerche rapide (per individuare eventuali refusi o effettuare normalizzazioni delle codifiche), che complesse, attraverso delle interfacce intuitive. Completano il tool un ambiente statistico e sistemi di esportazione dati (txt, xls)[10]. In questo caso NormaTEI è stato utilizzato per esemplificare un campione descrittivo di *Verbum*.

Nella Tabella 3 è possibile avere uno sguardo d'insieme sulla distribuzione quantitativa dei cluster di <term>, con i loro specifici @type e @subtype, in correlazione al numero delle recensioni. In particolare, l'utilizzo del tool ha consentito di isolare, all'interno delle 57 recensioni, 1257 forme ed espressioni fraseologiche, di cui ben 277 appartengono all'area semantica delle arti visive, e in generale all'atto della narrazione come rappresentazione; 94 sono legate al campo semantico dell'oggettività della prosa narrativa; 263 alla rappresentazione che afferisce a un'area semantica fortemente caratterizzata del linguaggio verista; 306 alla peculiare forma di 'studio' verista, strettamente legato all'osservazione; 206 richiamano la dimensione socio-economica; 111, infine, sono riferite all'area tematica del vero e della verità.

Classificazione dei termini	Sottocategoria	Occorrenza	Numero di recensioni
verbum	arti_visive	277	49
	evoluzione_sociale	206	31
	oggettività_impersonalità	94	25
	rappresentazione	263	53
	studio_osservazione	306	43
	vero_verità	111	40
forma_letteraria	autorialità_narratività	219	43
	genere_letterario	175	43
	ismi	66	23
	narrazione	184	40
stile	lessico	13	8
	lingua	53	16
	registro	1	1
	sintassi	27	13

Tabella 3. Presenza di <term> in 57 recensioni

Un altro aspetto rilevante che le estrazioni di dati hanno permesso di isolare riguarda la distribuzione diacronica di *Verbum* nell'arco cronologico preso in esame. La Figura 4 infatti ci permette di monitorare le occorrenze del lessico dei critici e dei suoi relativi cluster negli anni

²³ <https://github.com/pierpaolosichera/NormaTEI>

²⁴ <http://bellininrete.istc.cnr.it>

che vanno dal 1874 al 1890, sottolineando una maggiore incisività in concomitanza con la pubblicazione de *I Malavoglia*, risalente al 1881.

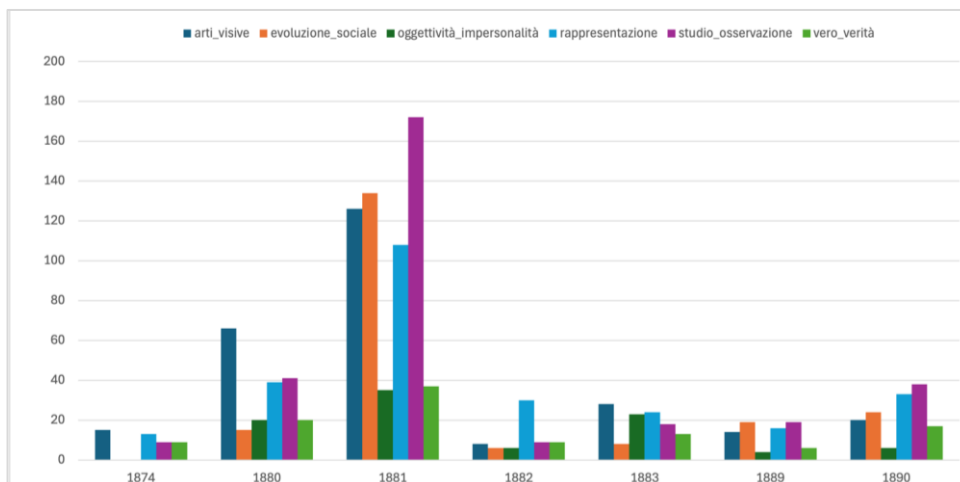


Figura 4. Occorrenze dei @subtype di Verbum in diacronia

La Figura 5 invece mostra in quante recensioni appaiono termini appartenenti ai vari cluster di *Verbum*. Anche questo dato assume una forte rilevanza, dal momento che conferma la presenza più o meno costante di questo lessico caratterizzante negli anni in cui il movimento verista si afferma.

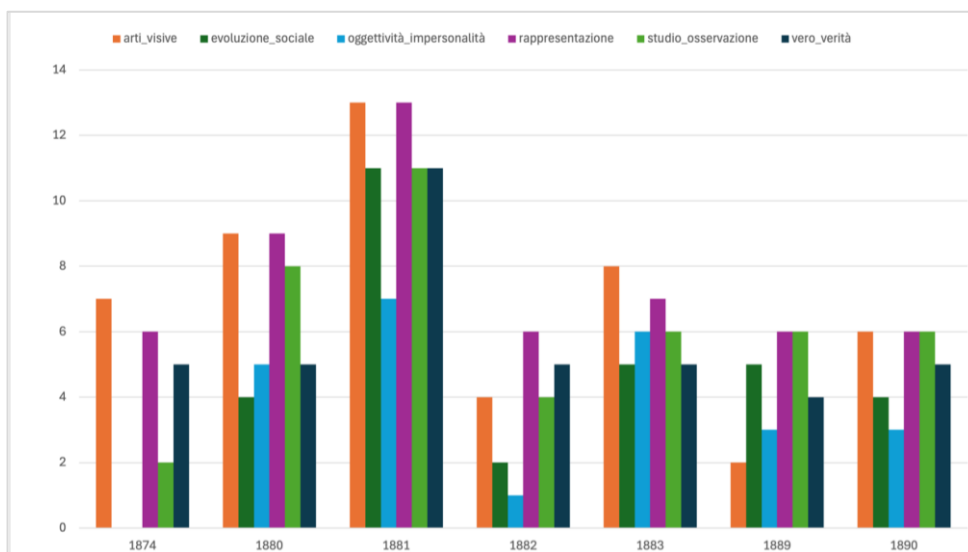


Figura 5. Distribuzione dei cluster di Verbum nelle recensioni

Lo spoglio progressivo dei testi amplia, dunque, non poco la già ricca cartografia lessicale e semantica di un eccezionale fenomeno del nostro post-risorgimento letterario quale è stato il Verismo, fornendo dati nuovi, i quali, incardinati in categorie più ampie, mostrano le sfaccettature di una lingua che via via diviene ‘paradigma’ per i recensori, ma anche per quegli

autori – che spesso la critica definisce ‘minori’ [25] – i quali corredano spesso le opere con prefazioni e postfazioni che richiamano puntualmente la tassonomia della pubblicistica. Un esempio lo ritroviamo nelle pagine di Lucini, il quale nel *Commiato a Gian Pietro da Core* dichiara che, immergendosi nella ricerca del fatto, l’autore non può più sottrarsi «allo incessante spettacolo di quanto [gli] accade intorno» [27:IX]; o ancora, nella *Prefazione ai Cento racconti popolari lucchesi*, Idelfonso Nieri avverte il lettore di aver ritratto «caratteri e parlate prese dal vero» [29:VI], usando altresì la parlata lucchese, poiché mutandole sarebbe «svanito il colore locale». Anche la Serao nel *Commiato a Dal vero* riporta «quanto era frutto di osservazione» [36:306], mentre Chelli, nella *Prefazione a L’Eredità Ferramonti*, scrive: «ho intrapreso un’opera vasta di osservazione» [6:3].

Le occorrenze emerse grazie all’estrazione effettuata mettono in evidenza come l’intento categoriale sia coerente e coeso. Prendiamo come esempio l’interpretazione dei dati estrapolati dal @subtype=“studio_osservazione”, il quale mostra una significativa occorrenza di termini i quali riflettono l’immagine di un linguaggio che acquista solidità nel tempo, con ‘picchi’ in corrispondenza delle opere maggiori di Verga. Si propongono di seguito, nella Tabella 4, alcuni esempi:

analisi	analisi dei sentimenti
	analisi psicologica
documenti	documenti umani
fatto	fatti diversi
	fatti umani
fisiologia	fisiologia della vita siciliana
	studio fisiologico
osservare	osservatore
	osservare con diligenza
	finezza d’osservazione
	osservazione attenta, acuta, minuta
	osservazione diligente, scrupolosa
	campo d’osservazione
realtà	realtà della vita
	realtà delle cose
studio	studio coscienzioso
	studio sociale
	studio biologico-sociale
	studio diretto
	studio psicologico

Tabella 4. Esempi di @subtype=“studio_osservazione”

Più nel dettaglio, interessante è il termine *fisionomia/fisionomia*, il quale, apparso per la prima volta nella lettera all'amico Salvatore Paola Verdura del 21 aprile 1878, diventa una parola chiave, o meglio 'modello', *in primis* per lo stesso Verga, che lo utilizzerà nella lettera a Rod del 18 aprile 1881, in cui afferma di aver provato a «rendere nettamente la *fisionomia* caratteristica di quei contorni siciliani nella lingua». Inoltre, il termine ricorrerà significativamente nelle recensioni di Francesco Torraca, il quale lo userà a proposito delle "eccezioni", che «non sono ancora compiuti caratteri, ma hanno una *fisionomia* spiccata»; e ancora, lo impiegherà Piucco, sulla «Gazzetta di Venezia» nell'aprile dell'81, parlando dell'abilità del Verga nel «cogliere la vera *fisionomia* di un ambiente sociale» [33:247] e, infine lo adopererà Remigio Zena, il quale, recensendo *Per le vie* nel settembre dell'83, parlerà di «*fisionomia* dei luoghi» [43:191], oltre che dei personaggi.

Il senso dell'analisi dei dati offerti dallo spoglio linguistico – effettuato grazie alla codifica di *Verbum* – alla caratterizzazione della lingua dei recensori, supportata altresì da 'paratesti', i quali si costituiscono come 'documenti di metodo', può essere espresso da una recensione di De Roberto al romanzo *Lydia* di Neera (pseudonimo di Anna Radius Zuccari): l'autore dei *Viveré* sottolinea in quest'opera «l'arte più nuda ma più forte, esaminante tutte le *manifestazioni* della realtà, procedente più *obbiettivamente* e fondata sull'*osservazione*» [19].

8. Visualizzazione e Interfaccia

La gestione e visualizzazione dei dati sin qui descritti è demandata a una piattaforma web, strutturata in moduli interattivi e integrati, progettati per fornire strumenti avanzati di consultazione, interrogazione e analisi.²⁵ Questa infrastruttura tecnologica mira a promuovere una trasformazione significativa nell'accesso e nella comprensione della letteratura secondaria sul Verismo, supportando l'interpretazione critica e stimolando nuove prospettive di ricerca, in un'interazione simbiotica con i dati relativi alla letteratura primaria. L'home page di COVERLeSS è accessibile all'URL <https://coverless.cnr.it>, di cui la Figura 6 presenta un parziale screenshot. Attraverso l'integrazione di funzionalità di esplorazione, ricerca e visualizzazione interattiva dei dati, il portale consente agli utenti di approcciarsi al corpus con un metodo semplificato ma al contempo efficace.

L'annotazione e la codifica dei testi costituiscono il prerequisito essenziale per garantire la visualizzazione e la consultazione dell'edizione sul web, progettata su un'istanza personalizzata del software open-source TEI Publisher²⁶ (versione 9.0.0). TEI Publisher si basa sul database noSQL XML eXist-db, installato su un server CNR ISTC di Catania. L'utilizzo di TEI Publisher risponde al primo modulo (consultazione dei dati) del portale web e consente l'elaborazione delle edizioni digitali dei documenti del corpus con la massima flessibilità rispetto alle peculiarità di ciascuna tipologia testuale, senza l'obbligo di modellare i contenuti entro un framework "one-size-fits-all".

²⁵ Nel rispetto delle linee guida di design per i siti internet e i servizi digitali della Pubblica Amministrazione sviluppate dall'AGID - Agenzia per l'Italia Digitale (<https://www.agid.gov.it/>).

²⁶ <https://teipublisher.com/exist/>



Figura 6. Home page di COVERLeSS

In particolare, è stato progettato un layout specifico che risponde alle esigenze dei testi del corpus identificati come recensioni, diverso dalla visualizzazione ideata per la pubblicistica, corredata dalla riproduzione fotografica della fonte primaria. In entrambi i casi, un menu laterale a scomparsa (Figura 7) permette l'accesso ai metadati del testo, codificati all'interno dell'elemento <teiHeader> di ciascuna unità testuale.

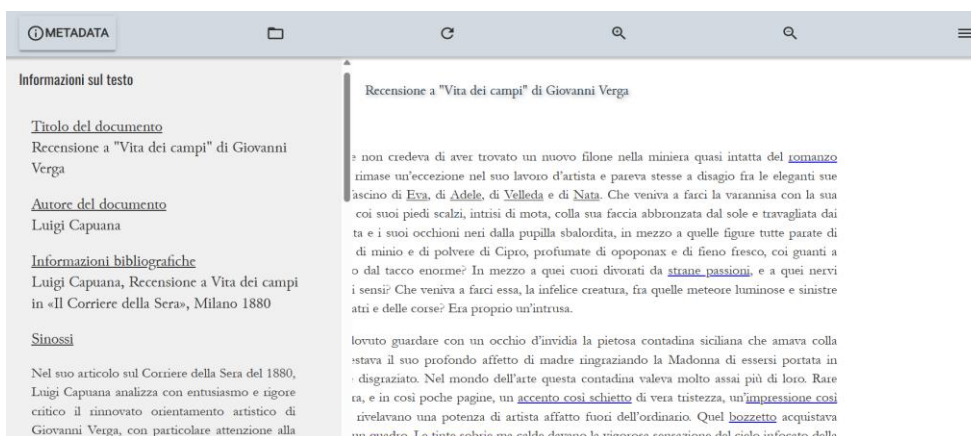


Figura 7. Prototipo della visualizzazione in TEI Publisher di un articolo di recensione apparso nel «Corriere della Sera», con attivazione del menu laterale che registra i metadati del testo

Per le recensioni è stato adottato un modello di visualizzazione (Figura 8) con una trascrizione a colonna singola, che offre la possibilità di consultare commenti e note esplicative, riferimenti bibliografici dell'opera recensita (insieme ad eventuali collegamenti ipertestuali) e un elenco delle entità nominate nel testo, organizzate in liste (nomi di persona, luoghi, organizzazioni, opere della bibliografia primaria).

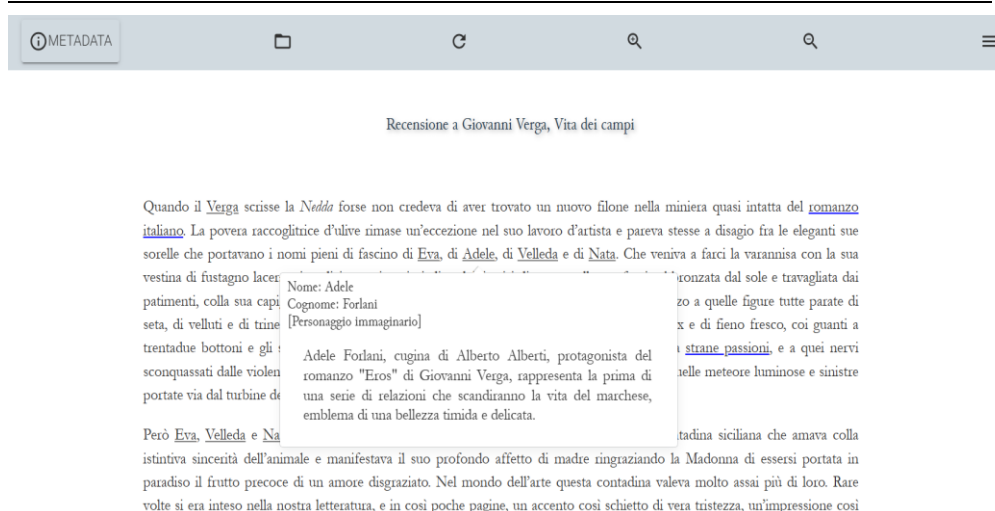


Figura 8. Prototipo della visualizzazione in TEI Publisher del medesimo articolo di recensione, con un riquadro informativo che si attiva al passaggio del mouse, finalizzato a illustrare le informazioni codificate per ciascuna entità nominata nel testo

Oltre alle funzionalità di base per la navigazione e per un'agevole fruizione dei contenuti (zoom personalizzato, scansione dei paragrafi in conformità con la fonte primaria), le entità nominate sono state sottolineate con colorazioni diversificate per facilitarne il riconoscimento. I colori adottati seguono la convenzione cromatica adoperata nel progetto *Verismo Digitale* [2], in modo da garantire una ideale continuità visiva tra la consultazione della bibliografia primaria e secondaria. Si segnala inoltre che in blu sono evidenziati termini ed espressioni relativi al lessico verista, codificati nei file XML mediante l'elemento <term>.

Per l'edizione degli articoli in rivista e dei saggi è stata prevista, in aggiunta, una modalità di visualizzazione che integra l'immagine facsimilare (Figura 9) accanto alla trascrizione del testo. Un archivio del materiale fotografico acquisito dalle biblioteche è ospitato su un server CNR ISTC di Catania, con dominio dedicato. Su tale dominio è stato installato un volume del container Cantaloupe,²⁷ destinato alla gestione e all'archiviazione delle immagini secondo il protocollo IIIF (International Image Interoperability Framework),²⁸ uno standard aperto e universale progettato per garantire l'accesso e l'interoperabilità di immagini e risorse digitali.

²⁷ <https://cantaloupe-project.github.io/>

²⁸ <https://iiif.io/>

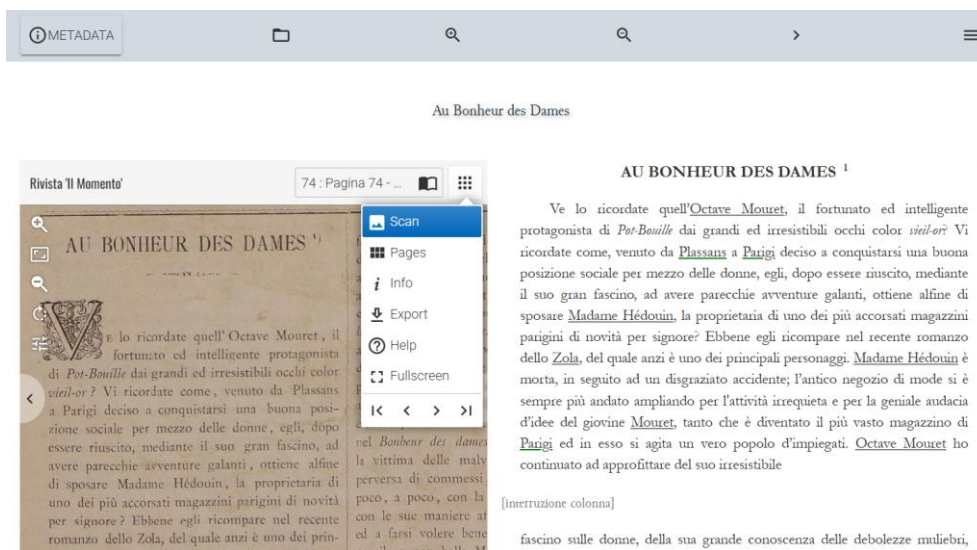


Figura 9. Prototipo della visualizzazione in TEI Publisher di un articolo in rivista, con accanto l'immagine corrispondente

9. Timeline *Ver-in-time*

La struttura della linea del tempo è stata pensata per fornire un supporto grafico di sintesi, intuitivo ed esaustivo, sui contenuti di COVERLeSS, funzionando al contempo come strumento didattico e come vero e proprio indice cliccabile dei contenuti del sito. Come lo stesso titolo, *Ver-in-Time*, vuole suggerire, gli elementi appartenenti alle tre macro aree costituenti il corpus COVERLeSS (recensioni, saggi e riviste) vengono collocati nel proprio contesto storico, mettendoli in relazione agli eventi più significativi (politici, sociali e culturali), nazionali e internazionali, come avvenimenti anch'essi, ma di una microstoria, che è quella tutta da delineare della critica e della pubblicistica verista.

Come già indicato nel par. 3, si è scelto di comprendere l'analisi del dibattito tra gli estremi cronologici del 1872 e del 1892. *Ver-in-Time* si compone di tre parti principali – *Linea dei Saggi*, *Linea delle Recensioni*, *Linea delle Riviste* –, suddivise in intervalli temporali di un anno, come mostra la Figura 10. In corrispondenza di ciascun anno lungo la linea del tempo, nella parte bassa viene inoltre mostrata la lista di alcuni tra i principali eventi storici verificatisi in quell'anno. Sono riportate tre diverse tipologie di eventi storici (*eventi politici e sociali italiani*; *eventi politici e sociali esteri*; *eventi scientifici, artistici e culturali italiani ed esteri*), rappresentate con tre diversi colori, come indicato nella legenda in basso a sinistra. Si osservi che per rendere l'aspetto grafico della linea del tempo più confortevole, la descrizione dei principali eventi storici viene talvolta presentata in forma abbreviata: posizionandosi su di essa, a seguito di un click del mouse è possibile visualizzarla in forma estesa.

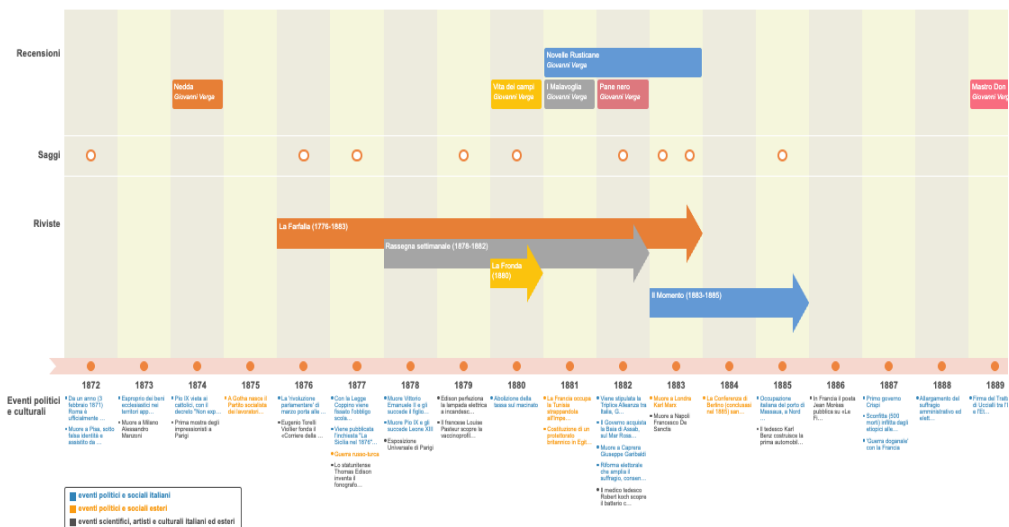


Figura 10. Prototipo della linea del tempo Ver-in-time

Lungo la *linea dei saggi* vengono riportati dei cerchietti segnaposto che individuano i saggi la cui collocazione temporale ricade all'interno di uno specifico anno, e che, quando *hoverati*, visualizzano il titolo del saggio corrispondente; cliccando su quest'ultimo, l'utente viene rimandato ad una pagina descrittiva di dettaglio. Nel caso della *linea delle recensioni*, il passaggio del mouse su uno dei titoli (*Nedda*, *Vita dei Campi*, *I Malavoglia*, *Pane nero*, *Novelle Rusticane*, *Mastro Don Gesualdo*) permette di visualizzare la lista delle recensioni di quella specifica categoria, come si vede in Figura 11. Come per i saggi, ogni item di questa lista è cliccabile con conseguente rimando ad una pagina del portale COVERLeSS contenente una descrizione dettagliata. Un click su una *freccia* nella *linea delle riviste* permette, poi, di esaminare la pagina della rivista che contiene tutti i numeri che si collocano negli anni individuati dalla freccia.

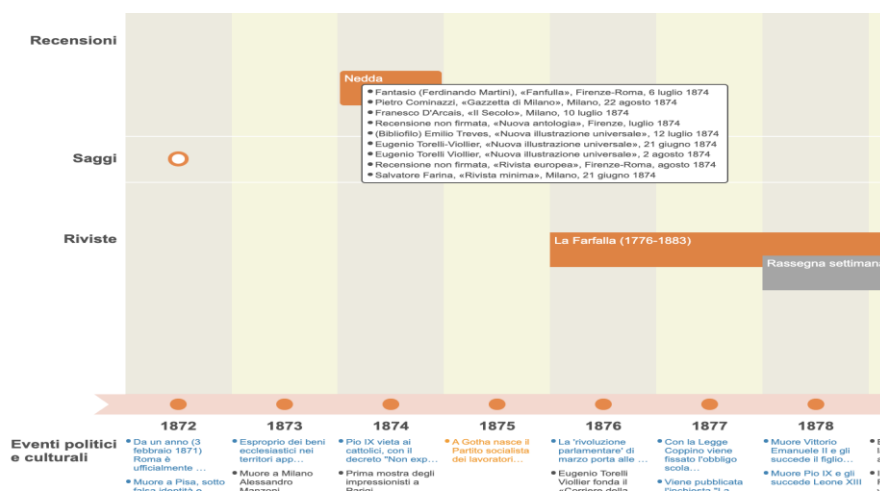


Figura 11. Particolare del prototipo Ver-in-time con il pop-up delle recensioni a Nedda

Dal punto di vista implementativo, la linea del tempo *Ver-in-time* è stata realizzata attraverso la definizione, e quindi l'inclusione all'interno degli elementi HTML, di appositi fogli di stile CSS per modellare le varie forme grafiche (shape) e applicare i diversi effetti visivi (colori, formattazione del testo, etc.). La parte interattiva viene gestita da funzioni implementate in (puro) JavaScript accoppiate all'utilizzo di selettori e combinatori CSS. Buona parte dei dati testuali visibili che popolano *Ver-in-time* sono stati estratti, e quindi incorporati in *Ver-in-time*, in maniera automatica a partire dai file XML grazie all'utilizzo di alcune procedure (in JAVA), create ad hoc, che prelevano questi dati direttamente dai file XML e li "scrivono" in forma strutturata all'interno di file sorgenti js contenenti apposite funzioni JavaScript che usano appunto questi dati per inizializzare la linea del tempo in fase di *on load*.²⁹ Questo processo automatizzato garantisce la dovuta corrispondenza tra le informazioni codificate nei file XML con quelle visibili e riprodotte in *Ver-in-time*, scongiurando, tra l'altro, eventuali errori e discrepanze dovuti ad una trascrizione manuale di queste informazioni.

10. Conclusioni

Su un piano di un avanzamento testuale, la ricerca sviluppata all'interno di COVeRLeSS intende dimostrare le potenzialità delle DH applicate all'edizione critica di testi, alla ricerca lessicale-tematica e alla navigazione strutturata all'interno di una testualità più 'liquida', sfuggente, intrinsecamente plurivoca e dialogica quale quella dei periodici. Di fronte all'affollarsi di materiali eterogenei, spesso frammentari e frammentati, l'unica possibilità di farli dialogare è quella di 'convocare' metodologie e strumenti diversi: l'edizione testo-immagine, l'interfaccia di ricerca, l'analisi del lessico e delle parole-chiave, etc. Va da sé che l'analisi del lessico specifico, di cui abbiamo tentato di realizzare uno specimen rappresentativo dal punto di vista quantitativo e qualitativo, ma con già una prima ipotesi di strutturazione in chiave di ontologia, riflette il taglio fortemente interdisciplinare delle tematiche, in quanto esso si costituisce come immagine della letteratura del secondo Ottocento in chiave letteraria, storico-sociale, antropologica. I percorsi tematici, che costituiscono la successiva implementazione del progetto, costruiti con i materiali testuali editi e quelli lessicali categorizzati e analizzati, faranno ancor di più di COVeRLeSS uno spazio di approfondimento interpellabile dagli utenti a diversi livelli, restituendo un'immagine autenticamente 'plurale' della letteratura del nascente Stato unitario, e fornendo una rappresentazione dinamica del processo socio-culturale che ha accompagnato la nascita di una 'nuova' letteratura.

Per parafrasare un paio di immagini care al massimo autore del Verismo, l'intento che accompagna costantemente le fasi della ricerca e del progetto è quello di sollecitare lettori curiosi, utenti appassionati, aficionados della letteratura della modernità, ma anche studiosi esperti, a guardare la grande letteratura post-unitaria dall'altro lato del cannocchiale, immedesimandosi, anzi "mettendosi nella pelle" dei lettori contemporanei, diventando essi stessi protagonisti del dibattito intorno alle grandi opere e scrutandone messaggi inediti, prospettive nuove, immagini sorprendenti.

Una pluralità di metodi, una molteplicità di ingressi al testo è stata sempre la *conditio sine qua non* dell'armonizzazione di oggetti eterogenei, osservati con la lente della lettura, della scomposizione lessicale, della relazione tra ambiti semantici, dell'attraversamento spazio-temporale. Tutti aspetti

²⁹ Per la creazione della timeline non è stato utilizzato nessuno dei tool adibiti a questo scopo tra quelli ampiamente disponibili in circolazione; ciò al fine di poter gestire in maniera più efficace l'implementazione delle varie funzionalità e personalizzazioni, avendo maggiore controllo su di essa.

che sono gli ideali capitoli di una ricerca la quale, nel nostro intento, vuole essere aperta a nuove integrazioni, inclusiva, metodologicamente dialogica, e soprattutto aperta alle più diverse esperienze testuali che scuola e università propongono e proporranno ai loro giovani discenti. E ancora, *last but non least*, di fronte ai progetti di digitalizzazione massiva, fortunatamente gestiti in Italia con cautela e rigore, e senza gli estremismi delle grandi ‘multinazionali’ delle biblioteche digitali, abbiamo voluto gettare un ponte con le tante biblioteche e istituzioni culturali della nostra penisola che oggi, tra le sfide della resilienza digitale, assolvono il difficile compito di comunicare alle nuove generazioni un patrimonio testuale moderno di sorprendente e ancora inesplorata ricchezza.

Ringraziamenti

Progetto PRIN PNRR 2022 COVerLeSS finanziato dall’Unione europea – Next Generation EU, Missione 4 Componente 2 CUP E53D2301888 0001.

Si ringraziano Milena Giuffrida per la consulenza scientifica sui testi di Capuana e Mariagiusti Polizzi per la codifica delle recensioni.

Il contributo è frutto di una riflessione congiunta da parte degli autori in tutte le sue parti e fasi di elaborazione. Nello specifico Denise Bruno è co-responsabile di §5 e §7; Giuseppe Canzoneri è responsabile di §3 e co-responsabile di §9; Salvatore Cristofaro è responsabile di §6 e co-responsabile di §7 e §9; Antonio Di Silvestro è responsabile di §1 e co-responsabile di §5 e §10; Laura Mazzagufo è responsabile di §8 e co-responsabile di §4; Daria Spampinato è responsabile di §2 e co-responsabile di §4, §7 e §10.

Bibliografia

- [1] Alfieri, Gabriella. – il realismo io l’intendo così, ..., in una parola –. Ritocchi e postille alla lettera a Salvatore Paola su «La Marea», in «Annali della Fondazione Verga», 13, Catania, 2020.
- [2] Barbarino, Liborio Pietro; Conti, Elisa; D’Agata, Christian; Grasso, Miryam; Martines, Ninna Maria Lucia; Vitale, Eliana. Verismo digitale. Per un’edizione digitale commentata delle opere di Verga, Capuana, De Roberto. In A. Di Silvestro, D. Spampinato (a cura di), AIUCD 2024 Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings. ISBN 978-88-942535-8-0. DOI 10.6092/unibo/amsacta/7927, pp. 252-259.
- [3] Bigazzi, Roberto. I colori del vero: vent’anni di narrativa: 1860-1880. Pisa, Nistri-Lischi, 1969.
- [4] Borrelli, Marco. Nell’officina del verismo: la novellistica della Rassegna Settimanale. Napoli, Loffredo, 2023.
- [5] Briganti, Alessandra; Cattarulla, Camilla; D’Intino, Franco. I periodici letterari dell’Ottocento: indice ragionato (collaboratori e testate). Milano: F. Angeli, 1990.
- [6] Chelli, Gaetano Carlo. L’eredità Ferramonti, Roma: Sommaruga, 1884.
- [7] Chemello, Adriana. La Farfalla di Angelo Sommaruga. Storia e indici. Roma: Bulzoni, 1977.

- [8] Chiffolleau, Floriane. Keeping It Open: A TEI-based Publication Pipeline for Historical Documents. *Journal of the Text Encoding Initiative* [Online], Issue 15 | 2024, Online since 20 November 2024, connection on 04 February 2025. URL: <http://journals.openedition.org/jtei/5306>; DOI:<https://doi.org/10.4000/12s01>
- [9] Cristofaro, Salvatore; D'Agata, Christian; Di Silvestro, Antonio; Palazzolo, Giuseppe; Sichera, Pierpaolo; Spampinato, Daria. 2021. DEMOTICON. Per un'edizione semantica dei Malavoglia. In Federico Boschetti, Angelo Mario Del Grosso, Enrica Salvatori (Eds.). *AIUCD2021 Book of Extended Abstracts*. In: *Quaderni di Umanistica Digitale*, 2021, 471-473. DOI:[10.6092/unibo/amsacta/6712](https://doi.org/10.6092/unibo/amsacta/6712).
- [10] Cristofaro, Salvatore; Del Grosso, Angelo Mario; Mazzagufo, Laura; Sichera, Pietro; Spampinato, Daria. Implementing Collaborative Digital Scholarly Editions: Insights from Bellini Digital Correspondence. *International Journal of Information Science and Technology*, vol. 9, no. 1, 2025. [in corso di stampa]
- [11] Cristofaro, Salvatore; Sichera, Pietro; Spampinato, Daria. An ontology proposal for a corpus of letters of Vincenzo Bellini: formal properties of physical structure and the case of rotated texts. *IJMSSO* 15(4), 2021, 269-279.
- [12] Cristofaro, Salvatore; Spampinato, Daria. Aspetti funzionali e implementativi del Museo epigrafico digitale EpiCUM. *Umanistica Digitale* 4(9) 2020, 61-77. DOI: <https://doi.org/10.6092/issn.2532-8816/9973>.
- [13] Croce, Benedetto. *Storia d'Italia dal 1871 al 1915*, Bari, Laterza, 1939.
- [14] Daquino, Marilena; Giovannetti, Francesca; Tomasi, Francesca. 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) 2019. DOI: <https://doi.org/10.6092/issn.2532-8816/9091>.
- [15] Dawey, Lin; et al. The TRUST principles for digital repositories. *Scientific data*, 7 (2020). <https://doi.org/10.1038/s41597-020-0486-7>
- [16] Del Grosso, Angelo Mario; Bellandi, Andrea; Giovannetti, Emiliano; Marchi, Simone; Nahli, Ouafae. Scanning is Just the Beginning: Exploiting Text and Language Technologies to Enhance the Value of Historical Manuscripts. *IEEE 5th International Congress on Information Science and Technology*, 2018. DOI: [10.1109/CIST.2018.8596373](https://doi.org/10.1109/CIST.2018.8596373).
- [17] Del Grosso, Angelo Mario; Giovannetti, Emiliano; Marchi, Simone. Il Modello a Microkernel Di Omega Nello Sviluppo Di Strumenti per Lo Studio Dei Testi: Dagli ADT Alle API. In "Book of Extended Abstracts" *AIUCD*, 2017: 199-205.
- [18] De Pasquale, Andrea. The Italian national digital newspaper library. *Bibliothecae.It* 7(2) 2018: 348-370. DOI: <https://doi.org/10.6092/issn.2283-9364/8951>.
- [19] De Roberto, Federico. Neera, in «Letteratura Contemporanea», 22 aprile 1888.
- [20] Di Silvestro Antonio; D'Agata, Christian; Palazzolo, Giuseppe; Sichera, Pierpaolo. Conservazione e fruizione di banche dati letterarie: l'archivio della

- poesia italiana dell'Otto/Novecento di Giuseppe Savoca. Atti del Convegno AIUCD2022, 2022: 98-104.
- [21] D'Orsogna, Fabio; Palanga, Giulio. Riviste digitali e digitalizzate italiane (RIDI): a reconnaissance for the national newspaper library. *JLIS.It* 13(1) 2022: 374-389. DOI: <https://doi.org/10.4403/jlis.it-12734>.
- [22] Farinelli, Giuseppe. *La pubblicistica nel periodo della scapigliatura*. Milano: IPL, 1984.
- [23] Fusi, Daniele, 2020. Text Searching Beyond the Text: a Case Study. «*Rationes Rerum*» Vol. 15, 2020, pp. 199-230.
- [24] Italia, Paola; Tomasi, Francesca. Filologia digitale. Fra teoria, metodologia e tecnica. *Ecdotica* 11, 2014: 112-131. DOI: 10.7385/99218.
- [25] I verismi regionali. Atti del Congresso Internazionale di Studi, Catania, 27-29 aprile 1992 (2 Voll.). Catania: Fondazione Verga, 1996.
- [26] Jutta, Ernst; Scheiding, Oliver. Periodical Studies as a Transpistemic Field. *Studies in Periodical Studies Today. Multidisciplinary Analyses*, ed. by Jutta Ernst, Dagmar von Hoff, Oliver Scheiding, Brill, Leiden/Boston, 2022.
- [27] Lucini, Gian Pietro. *Gian Pietro da Core: storia della evoluzione della Idea*. Milano, Galli, 1895, p. IX.
- [28] Melis, Rossana. *La bella stagione del Verga*. Catania: Biblioteca della fondazione Verga, 1990.
- [29] Nieri, Idelfonso. *Cento racconti popolari lucchesi*. Livorno: Giusti, 1908.
- [30] Pellini, Pierluigi. *Naturalismo e Verismo. Zola, Verga e la poetica del romanzo*. Firenze: Le Monnier Università, 2010.
- [31] Raya, Gino. *Bibliografia di Luigi Capuana (1839-1968)*. Roma: Ciranna, 1969.
- [32] Raya, Gino. *Bibliografia Verghiana (1840-1871)*. Roma: Ciranna, 1972.
- [33] Rappazzo, Felice; Lombardo, Giovanna. *Giovanni Verga fra i suoi contemporanei*. Soveria Mannelli: Rubbettino, 2016.
- [34] Romano, Cinzia. *Emanuele Navarro della Miraglia: un percorso esemplare di secondo Ottocento*. Catania: Biblioteca della Fondazione Verga, 1998.
- [35] Saja, Giuseppe. *Il Momento. Identità di una rivista di fine Ottocento con gli indici del periodico (1883-1885)*. Caltanissetta-Roma: Sciascia, 2004.
- [36] Serao, Matilde. *Dal vero*. Milano: Perussia e Quadrio, 1879.
- [37] Sichera, Pierpaolo; Cristofaro, Salvatore; Mazzagufo, Laura; Spampinato, Daria; Del Grosso, Angelo Mario. CHROMA model for H2IOSC. Poster for Bazaar in CLARIN Annual Conference 2024 (CLARIN2024), Barcelona, Spain, <<https://doi.org/10.5281/zenodo.13913607>>.
- [38] Spadini, Elena; Tomasi, Francesca; Vogeler, Georg. (Eds.). *Graph Data-Models and Semantic Web Technologies in Scholarly Digital Editing*. Norderstedt: Herstellung und Verlag, 2021.

- [39] Verga, Giovanni, *I Malavoglia*, a cura di Ferruccio Cecco, Novara: Fondazione Verga-Interlinea, 2014.
- [40] Verga, Giovanni. *Vita dei Campi*, a cura di Riccardi Carla, Novara: Fondazione Verga-Interlinea, 2021.
- [41] Verga, Giovanni. *Carteggi con Felice Cameroni, Salvatore Farina e Ferdinando Martini*, a cura di Vitale Melania Maria, Novara: Fondazione Verga-Interlinea, 2023.
- [42] Wilkinson, Mark D.; et al. The FAIR guiding principles for scientific data management and stewardship. *Scientific data* 3, no. 1 (December 2016): 160 018. DOI: <https://doi.org/10.1038/sdata.2016.18>.
- [43] Zena, Remigio. *Verismo polemico e critico*, Roma: Silva Editore, 1971.