

PAVES-e: Per una Hyperedizione dell'opera di Cesare Pavese

Christian D'Agata¹, Angelo Mario Del Grosso², Laura Nay³,
Giuseppe Palazzolo⁴, Antonio Sichera⁵, Daria Spampinato⁶

¹Università di Catania, Italia - christian.dagata@unict.it

²CNR Istituto di Linguistica Computazionale "A. Zampolli", Italia - angelomario.delgrosso@cnr.it

³Università di Torino, Italia - laura.nay@unito.it

⁴Università di Catania, Italia - giuseppe.palazzolo@unict.it

⁵Università di Catania, Italia - asichera@unict.it

⁶CNR Istituto di Scienze e Tecnologie della Cognizione, Italia - daria.spampinato@cnr.it

ABSTRACT

L'articolo presenta il progetto PAVES-e, finanziato con i fondi PRIN 2022, che intende creare un'edizione-archivio semantica open access, definita Hyperedizione, dell'opera di Cesare Pavese. Essa permetterà a un vasto pubblico di fruire dell'opera paveseiana attraverso un portale nel quale poter consultare le edizioni scientifiche digitali delle poesie e dei romanzi mettendole in relazione con collegamenti semantici all'epistolario e al diario (*Mestiere di vivere*), fruendo inoltre dei manoscritti, dei vocabolari d'autore e del commento multimediale. Il contributo, dopo aver presentato il progetto (distinguendo analiticamente tutte le sue fasi: *DigitalPavese*, *OntoPavese*, *PaveseInImmagini*, *PaveseInTesto*, *PaveseInParole*, *AnnotaPavese*, *BiblioPavese*), si sofferma in particolare su *PaveseInTesto*, discutendo alcune scelte di codifica e di modellizzazione delle edizioni, con alcune proposte sul workflow, sugli editor scelti e sulla interfaccia di visualizzazione.

PAROLE CHIAVE

Edizioni scientifiche digitali semantiche; codifica XML/TEI; archivi digitali d'autore; lessicografia; filologia computazionale.

1. INTRODUZIONE

Il progetto PAVES-e mira a creare una Hyperedizione dell'opera di Cesare Pavese secondo il paradigma dell'Hyperedizione sperimentato in PirandelloNazionale [5], ovvero un'edizione-archivio open-access, flessibile e profondamente integrata, che possa offrire una modalità innovativa di accedere e collegare tra loro i dati sfruttando le potenzialità delle ontologie formali. Pavese è infatti uno dei pochi autori italiani al quale da anni è dedicato un portale: si tratta di Hyperpavese (www.hyperpavese.com), che conserva le carte dell'autore concesse in comodato e suddivise in due sottofondi: Fondo Einaudi (documenti provenienti dall'omonima casa editrice) e Fondo Sini (carte originariamente conservate presso gli eredi). Hyperpavese ospita le carte dell'autore in formato digitale (circa 13.000 documenti) ma attende ancora una digitalizzazione ad alta risoluzione, uno spoglio rigoroso, una sistemazione accurata, un'indicizzazione puntuale, oltre alla possibilità di una libera fruizione. All'archivio cartaceo gli studiosi hanno attinto per la trascrizione e la pubblicazione di alcuni epistolari [12], dell'edizione critica de *Il mestiere di vivere* [10], di alcune traduzioni classiche e moderne (Dughera, Barberi, Cavallini, Pietralunga), mentre all'archivio digitale hanno recentemente guardato Barbarino per *Lavorare Stanca* [2], Grasso per *La luna e i falò* [14], Nay-Tavella per *Prima che il gallo canti* [13] e Sichera-Di Silvestro per *l'Opera poetica* [12], in cui sono state trascritte e annotate migliaia di carte inedite dell'archivio paveseiano. Hyperpavese può fornire materiali utilissimi per un 'Pavese elettronico', un PAVES-e, che dedichi all'opera del grande scrittore un'Hyperedizione, intesa come uno spazio digitale integrato, dove far dialogare testi e prospettive in una modellizzazione [4] che metta al centro il testo come dimensione virtuale di un'esperienza molteplice (filologica, critica, lessicografica), rivolta a utenti e profili diversi (lettori, studenti e studiosi).

2. STATO DELL'ARTE

Allo stato attuale, le biblioteche digitali esistenti in ambito italiano (Progetto Manuzio o Biblioteca Italiana) consentono un accesso immediato ai grandi testi della tradizione letteraria italiana, con modalità di esplorazione efficaci ma semplici, senza alcuna forma di rappresentazione delle interconnessioni semantiche tra opere letterarie e altri materiali documentari (come immagini, lettere, documenti d'archivio e risorse bibliografiche) e senza alcun supporto per l'analisi del testo e la

navigazione. I portali letterari più recenti puntano essenzialmente all'offerta di edizioni scientifiche digitali dei testi [8] o di edizioni delle carte (con trascrizione e varie forme di esplorazione accluse): Manzoni on-line¹, Dante on-line², Archivio pascoliano³ e Digital Vercelli Book⁴, dalla cui esperienza è nato il software di visualizzazione Edition Visualization Technology (EVT) [15]. Uno stato dell'arte completo per quanto riguarda le edizioni digitali si trova in Franzini⁵ [7] e in Sahle⁶. In questo contesto, l'Edizione Digitale dell'Opera Omnia di Pirandello⁷, rappresenta un modello di riferimento in quanto luogo di sintesi di esigenze scientifiche e culturali diverse: dalla pubblicazione di manoscritti e dattiloscritti (con relativa trascrizione) all'edizione digitale dei testi (con annessa sinossi); dalla vocabolarizzazione dei testi all'applicazione didattica. Filologia, lessicografia e critica si integrano in una visione plurivoca e polifunzionale dei dati testuali [16]. Per quanto riguarda le edizioni semantiche, un modello è invece l'edizione delle Lettere di Vespasiano da Bisticci⁸ [18] e l'edizione dell'Opera Omnia di Aldo Moro⁹, mentre per la codifica di epistolari e postille modello esemplare sono rispettivamente l'edizione Bellini Digital Correspondence [6] e l'edizione delle Postille di Giorgio Bassani [17]. L'organizzazione della conoscenza del progetto è in accordo a prassi ampiamente consolidate [19], mentre l'ontologia computazionale viene sviluppata sfruttando le moderne tecnologie di web semantico [1], nonché modelli concettuali scientificamente comprovati, quali CIDOC-CRM ed ontologie lessicali e reti semantiche come WordNet¹⁰ e BabelNet [9].

3. IL PROGETTO

Il progetto PAVES-e mira allo sviluppo di un'edizione-archivio che, nel rispetto dei principi FAIR (Findable, Accessible, Interoperable, Reusable) [20], costruisca un'interfaccia innovativa (a) e integri, attraverso una solida ontologia (b), le dimensioni filologiche (c, d), lessicografiche (e), critiche (f) e didattiche (g). PAVES-e prevede infatti (vd. Fig. 1):

- la progettazione e realizzazione di un'interfaccia web intuitiva e user-friendly (*DigitalPavese*);
- l'organizzazione delle informazioni e l'integrazione tra le risorse, al fine di sviluppare un'ontologia di luoghi, persone, personaggi, organizzazioni, opere comuni a tutto il patrimonio pavesiano (*OntoPavese*);
- la realizzazione di un archivio digitale, corredato da descrizioni e metadati, delle immagini di manoscritti autografi, dattiloscritti e prime edizioni (*PaveseInImmagini*);
- l'elaborazione delle edizioni digitali scientifiche dei romanzi e dell'opera poetica, dell'epistolario e del diario, codificati in XML/TEI e lemmatizzati per quanto riguarda le opere letterarie (*PaveseInTesto*);
- la creazione di un vocabolario e del relativo software di interrogazione, che comprenda i lemmi delle principali opere pavesiane (*PaveseInParole*);
- la raccolta dell'archivio bibliografico della letteratura secondaria, indicizzata e strutturata e interoperabile (*BiblioPavese*);
- la messa a punto di un tool di annotazione dei testi in XML/TEI — o la personalizzazione di un *tool* già esistente — che ne permetta un agevole riuso in prospettiva didattica (*AnnotaPavese*).

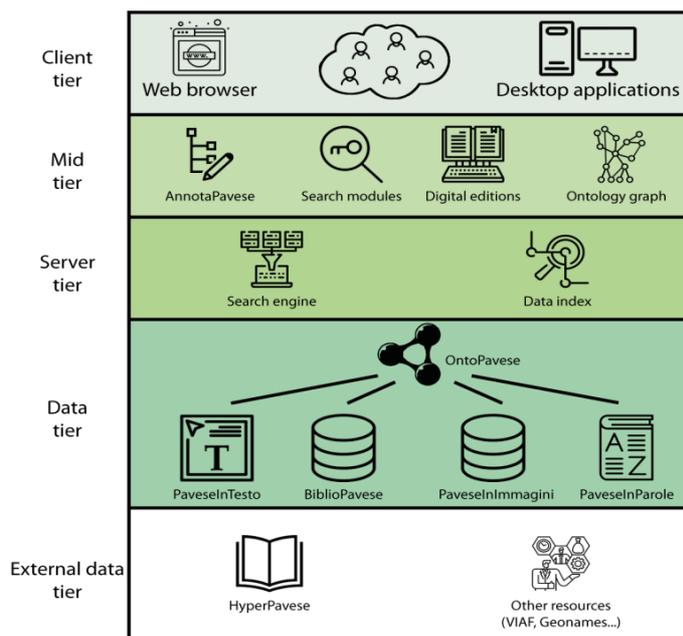


Figura 1. Architettura generale di PAVES-e

¹ <https://www.alessandromanconi.org/>

² <https://www.danteonline.it/>

³ <http://pascoli.archivi.beniculturali.it/>

⁴ <http://vbd.humnet.unipi.it/>

⁵ <https://dig-ed-cat.acdh.oeaw.ac.at/>

⁶ <https://v3.digitale-edition.de/>

⁷ <https://www.pirandellonazionale.it/>

⁸ <http://projects.dharc.unibo.it/vespasiano/>

⁹ <https://aldomorodigitale.unibo.it/>

¹⁰ <https://wordnet.princeton.edu/>

4. PAVESEINTESTO: DIGITALIZZAZIONE, ARCHIVIO E CODIFICA DEI TESTI

PaveseInTesto, in particolare, presenterà le Edizioni Scientifiche Digitali dell'opera poetica (*Lavorare stanca, Verrà la morte e avrà i tuoi occhi*), dei romanzi maggiori (*Paesi tuoi, La bella estate, Prima che il gallo canti, La luna e i falò*), dei *Dialoghi con Leucò*, del diario (*Il mestiere di vivere*) e dell'epistolario, a partire dalle edizioni di riferimento già pubblicate o in fase di pubblicazione. Le edizioni saranno allestite attraverso una codifica dell'onomastica, della toponomastica, del discorso diretto e per quanto riguarda le opere letterarie delle correzioni d'autore. Presenteranno un approfondito *TeiHeader* nel quale saranno esplicitati i metadati fondamentali (ad esempio le informazioni sulla data di stesura delle singole poesie o dei romanzi, le fasi di revisione, la responsabilità di ciascun aspetto della codifica, le informazioni bibliografiche, ecc.). La rappresentazione digitale di tutti questi aspetti sarà effettuata attraverso il sistema di codifica XML/TEI [3], che permetterà di annotare semanticamente i testi integrandoli con l'ontologia di *OntoPavese* e in modalità *Linked Open Data*. Per codificare i testi (lettere ed epistolario) il progetto ha previsto l'utilizzo di un tool pensato esplicitamente per la rappresentazione delle relazioni semantiche: *LEAF Writer*¹¹ un editor online XML & RDF per il 'Linked Editing Academic Framework'. Questo editor (vd. Fig. 2) permette attraverso una GUI il semplice inserimento di `<persName>`, `<placeName>`, `<orgName>`, `<title>`, `<rs>`, `<quote>`, `<term>`, `<date>`, a partire da un file codificato in TEI (elaborato automaticamente da un semplice script in Python), con la possibilità di aggiungere con grande immediatezza i link a VIAF, Wikidata, DBPedia, GeoNames attraverso un menù a tendina. La scelta è ricaduta su *LEAF* anche per la possibilità di editare collaborativamente in cloud appoggiandosi su GitHub, in modo da poter avere sempre un *versioning* dei documenti, e contestualmente un parser che validi il documento e permetta di visualizzare sulla destra la codifica 'raw XML'.

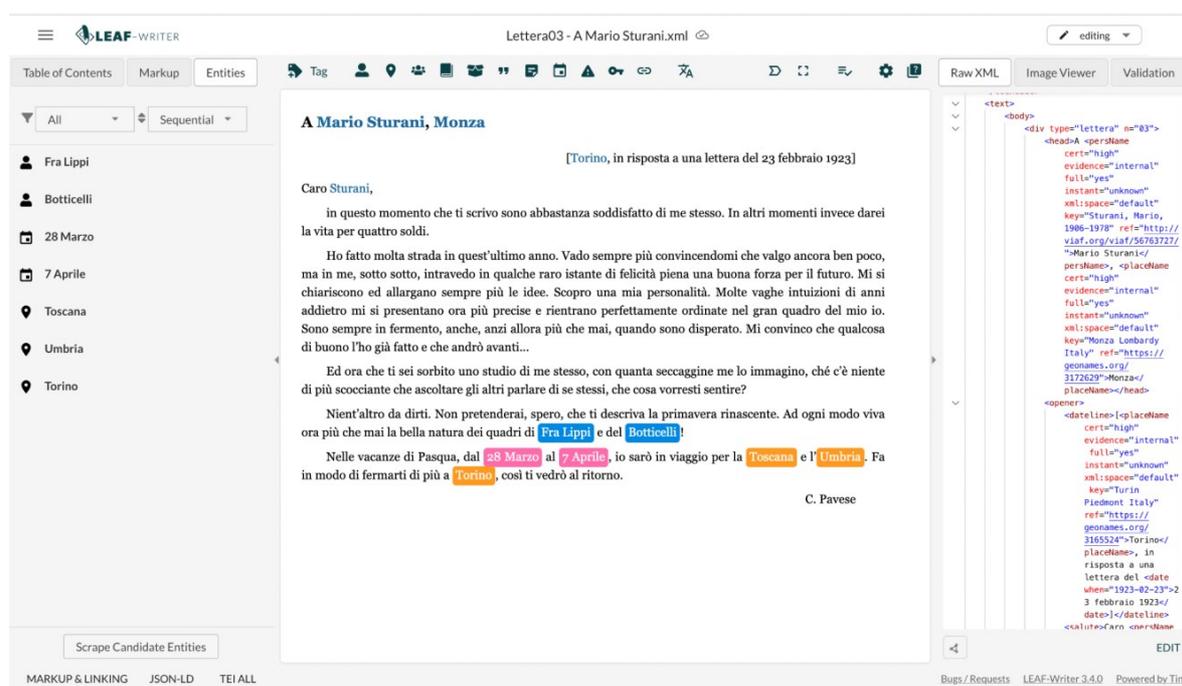


Figura 2. Schermata di visualizzazione del tool di codifica LEAF-writer

Per la codifica dei testimoni e degli apparati critici delle poesie e dei romanzi, ovvero per la rappresentazione dell'aspetto filologico, si è però scelto di utilizzare come editor Oxygen in quanto *LEAF* è pensato specialmente per la rappresentazione semantica dei *Linked Open Data* e presenta ancora dei limiti nella rappresentazione degli apparati e, in generale, di tutti quei dati che presentano diversi livelli di gerarchia. Infatti, ad esempio, una poesia come *I mari del Sud* in *Lavorare stanca* presenta tre testimoni (A1, A2 e A3; vd. Fig. 3), di cui uno particolarmente lavorato, com'è anche possibile leggere dall'edizione critica tradizionale (vd. Fig. 4)

¹¹ <https://leaf-writer.leaf-vre.org/>

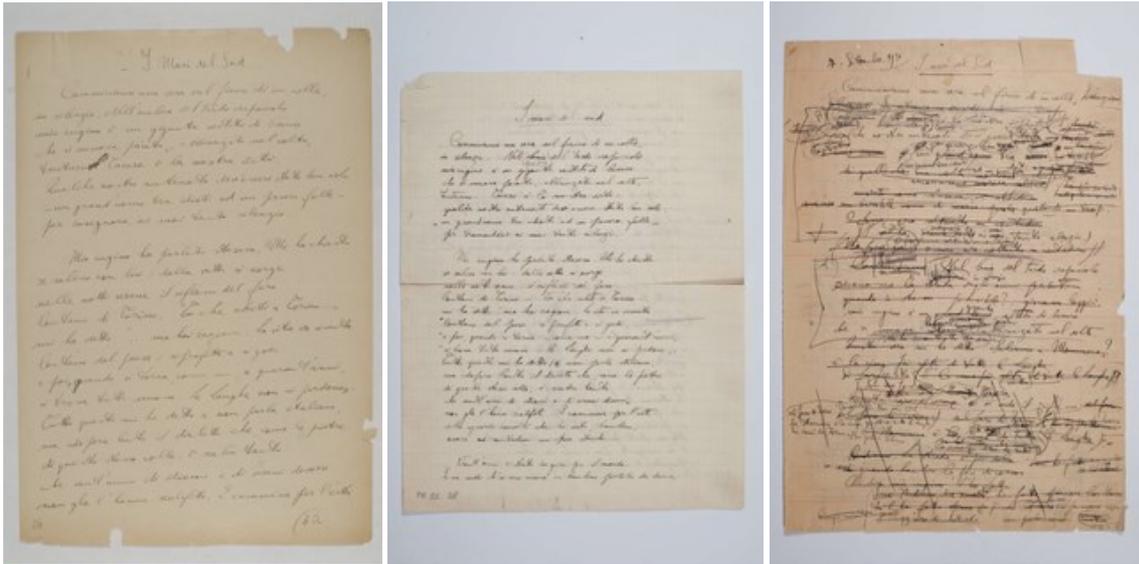


Figura 3. Immagini della digitalizzazione presente su HyperPavese della prima carta delle tre diverse redazioni di Mari del Sud

I MARI DEL SUD

Qualche nostro antenato dev'essere stato ben solo⁶
 – un grand'uomo tra idioti o un povero folle –⁷
 per insegnare ai suoi tanto silenzio.⁸

Mio cugino ha parlato stasera. Mi ha chiesto⁹
 se salivo con lui: dalla vetta si scorge¹⁰
 nelle notti serene il riflesso del faro¹¹

⁶ A1 Qualche nostro antenato dev'essere [stato un bandito / o uno scemo]
 [stato uno scemo] sup.
 [stato un bandito] -sup.
 [stato un idiota] -dx.
 [stato un reietto] inf.
 [morto.] -sx.sup.
 stato sup. ben [triste] inf.
 [solo] -inf.
 triste -dx.

A2 qualche nostro antenato dev[e] essere stato ben solo
 dev'essere

A3 Qualche nostro antenato dev'essere stato ben solo

Figura 4. Edizione critica dei Mari del Sud

Volendo distinguere il momento della rappresentazione del testo da quello della presentazione dei dati attraverso un'interfaccia, ci si è posti il problema su come rendere, innanzitutto nella fase di modellizzazione della codifica, i diversi livelli del testo: quello documentale e quello dell'apparato critico. Tali livelli, infatti, darebbero frutto a due modelli di edizione diversi, l'edizione diplomatica del manoscritto e l'edizione critica con apparato. Per distinguere questi due aspetti si è scelto innanzitutto di rappresentare ciascun testimone di una poesia in <sourceDoc>, che è un tagset figlio di <TEI> in alternativa a <facsimile>, attraverso i tag di aggiunta, cassatura, sostituzione (<add>, , <mod>), inserendo invece nel <text> il testo accolto nell'edizione critica. La visualizzazione di ciascun testimone è poi demandata a TeiPublisher¹² (vd. Fig. 5) (opportunamente adattato e implementato per le visualizzazioni degli aspetti semantici e dei Linked Open Data) che permetterà la consultazione contemporanea sia del facsimile digitale delle carte, sia del testo critico, affiancando le informazioni sui personaggi, luoghi, organizzazioni, opere citate.

¹² <https://teipublisher.com/>

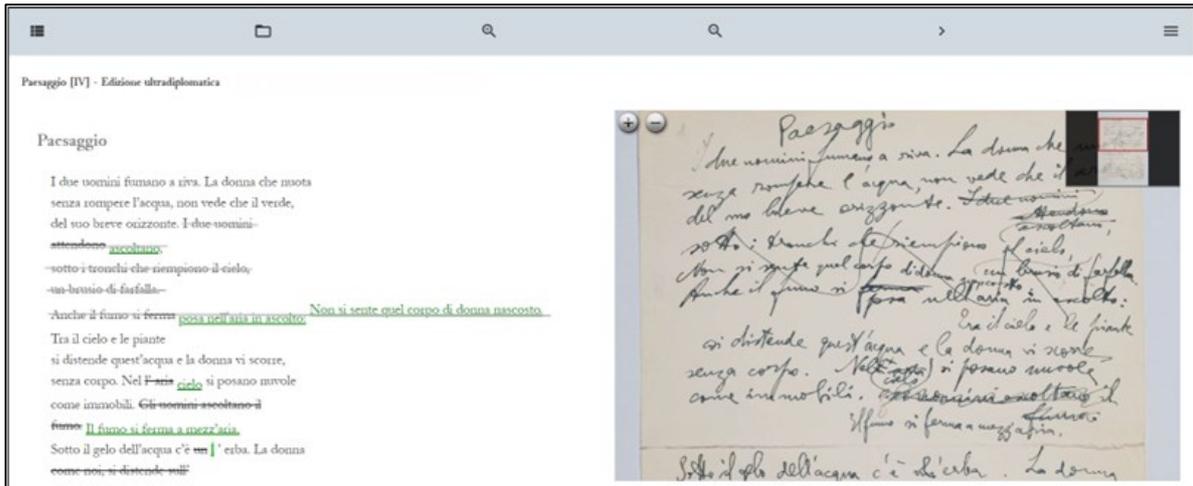


Figura 5. Prototipo di visualizzazione di una carta della poesia Paesaggio [VI] con TeiPublisher

5. ULTERIORI SVILUPPI

PavesesInTesto rappresenta certamente *in nuce* le caratteristiche principali di PAVES-e, trattandosi del momento fondamentale di allestimento delle edizioni scientifiche digitali, ma la specificità del progetto e dell’Hyperedizione è quella di prevedere un ampliamento dei tradizionali confini del testo attraverso la lemmatizzazione, il commento multimediale, la rappresentazione semantica dei vari aspetti del progetto, il riuso di tool, lo sviluppo di interfacce ad hoc. In sintesi, a partire dagli elementi già tratteggiati nell’articolo, è attualmente previsto negli sviluppi futuri:

1. Implementazione di un’istanza di LEAF-writer su un server dedicato e sviluppo di funzionalità aggiuntive come la marcatura degli apparati e una migliore gestione dei vari livelli di gerarchia (*AnnotaPaveses*).
2. Integrazione e visualizzazione efficace dell’aspetto documentale (<sourceDoc>) dei testimoni con l’apparato critico del testo (*PavesesInTesto*).
3. Gestione e collegamento dei dati della lemmatizzazione con la codifica del testo (*PavesesInParole*).
4. Sviluppo di un’interfaccia per la visualizzazione e l’interrogazione dell’ontologia (*OntoPaveses*).
5. Sviluppo dell’interfaccia dell’Hyperedizione che permetta di collegare tutti i dati, da quelli filologici, a quelli multimediali e lessicografici, in un’unica modalità di lettura progettata ad hoc per una visualizzazione modulare basata sull’user-centered design (*DigitalPaveses*).

6. RINGRAZIAMENTI

Il progetto PAVES-e è finanziato dal Ministero dell’Università e della Ricerca - bando PRIN 2022. Gli autori, inoltre, ringraziano Adriana Damico, Laura Mazzagufò e Alberto Luca Zuliani per il lavoro svolto nell’ambito del progetto.

BIBLIOGRAFIA

- [1] Allemang, Dean, e Jim Hendler. *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL*. Amsterdam: Elsevier, 2011.
- [2] Barbarino, Liborio P. *Il primo “Lavorare stanca” di Pavese (1936). Edizione critica*. Avellino: Sinestesie, 2020.
- [3] Burnard, Lou. *What Is the Text Encoding Initiative?: How to Add Intelligent Markup to Digital Resources*. Marseille: OpenEdition Press, 2014.
- [4] Ciula, Arianna, Øyvind Eide, Cristina Marras, e Patrick Sahle. *Modelling Between Digital and Humanities Thinking in Practice*. Cambridge, UK: Open Book Publisher, 2023. <https://doi.org/10.11647/OBP.0369>.
- [5] D’Agata, Christian, Antonio Di Silvestro, e Antonio Sichera. «Edizione critica, edizione digitale, hyperedizione. “Il fu Mattia Pascal” come paradigma dell’Edizione digitale dell’Opera Omnia di Luigi Pirandello». *Bollettino. Centro di studi filologici e linguistici siciliani* 33 (2022): 263–80.
- [6] Del Grosso, Angelo Mario, Erica Capizzi, Salvatore Cristofaro, Maria R. De Luca, Emiliano Giovannetti, Simone Marchi, Graziella Seminara, e Daria Spampinato. «Bellini’s Correspondence: a Digital Scholarly Edition for a Multimedia Museum». *Umanistica Digitale* 3, fasc. 7 (2019): 23–47. <https://doi.org/10.6092/issn.2532-8816/9162>.
- [7] Franzini, Greta, Melissa Terras, e Simon Mahony. «A Catalogue of Digital Editions». In *Digital Scholarly Editing: Theories, Models and Methods*, a cura di Elena Pierazzo. Farnham, Surrey: Ashgate, 2015.
- [8] Mancinelli, Tiziana, e Elena Pierazzo. *Che cos’è un’edizione scientifica digitale*. Roma: Carocci, 2021.

- [9] Missikoff, Michele, Roberto Navigli, e Paola Velardi. «Integrated approach to web ontology learning and engineering». *Computer* 35, fasc. 11 (2002): 60–63.
- [10] Pavese, Cesare. *Il mestiere di vivere 1935-1950*. (a cura di) M. Guglielminetti e L. Nay. Torino: Einaudi, 1990.
- [11] Pavese, Cesare. *La luna e i falò. Edizione critica*. (a cura di) M. Grasso. Avellino: Sinesthesie, 2020.
- [12] Pavese, Cesare. *L'Opera poetica. Testi editi, inediti e traduzioni*. (a cura di) A. Sichera e A. Di Silvestro. Milano: Mondadori, 2021.
- [13] Pavese, Cesare. *Prima che il gallo canti*. (a cura di) L. Nay e C. Tavella. Milano: BUR, 2021.
- [14] Pavese, Cesare, e Bianca Garufi. *Una bellissima coppia discorde. Il carteggio tra Cesare Pavese e Bianca Garufi (1945-1950)*. (a cura di) M. Masoero. Firenze: Olschki, 2011.
- [15] Rosselli del Turco, Roberto, e Chiara Di Pietro. «La visualizzazione di edizioni digitali con EVT: una soluzione per edizioni diplomatiche e critiche». *Ecdotica* 1, fasc. 2019 (2019): 148–73. <https://doi.org/10.7385/99301>.
- [16] Savoca, Giuseppe. *Lessicografia letteraria e metodo concordanziale*. Firenze: Olschki, 2000.
- [17] Siciliano, Angela, e Angelo Mario Del Grosso. «From print to digital: an encoding model for the scholarly edition of Giorgio Bassani's notes». *Umanistica Digitale* 6, fasc. 13 (2022): 21–48. <https://doi.org/10.6092/issn.2532-8816/13688>.
- [18] Tomasi, Francesca, (a cura di). *Vespasiano da Bisticci, Lettere. Knowledge Site 3.0*. Bologna: DH.arc., 2020.
- [19] Tomasi, Francesca, Elena Spadini, e Georg Voegler. *Graph Data-Models and Semantic Web Technologies in Scholarly Digital Editing*. Norderstedt: Institut für Dokumentologie und Editorik, 2021.
- [20] Wilkinson, Mark, Michel Dumontier, Ijsbrand J. Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. «The FAIR Guiding Principles for scientific data management and stewardship». *Scientific Data* 3, fasc. 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.