

# L'edizione digitale dei primi 16 Taccuini di Paolo Orsi

Giuseppina Monterosso<sup>1</sup>, Andrea Bolioli<sup>2</sup>, Elisa Bonacini<sup>3</sup>, Gianmario Cattaneo<sup>4</sup>, Dario Gonella, Anna Maria Marras<sup>5</sup>, Salvatore Spina<sup>6</sup>, Paola Venuti<sup>7</sup>

<sup>1</sup>Parco Archeologico e Paesaggistico di Siracusa, Eloro, Villa del Tellaro e Akrai - giuseppina.monterosso@regione.sicilia.it

<sup>2</sup>Ricercatore e consulente - andrea.bolioli@gmail.com

<sup>3</sup>Università di Bari «Aldo Moro», Dipartimento di Ricerca e Innovazione Umanistica - elisa.bonacini@uniba.it

<sup>4</sup>Università degli Studi del Piemonte Orientale - gianmario.cattaneo@uniupo.it

<sup>5</sup>Università di Torino, Dipartimento di Studi Storici - annamaria.marras@unito.it

<sup>6</sup>Università degli Studi di Catania, Dipartimento di Scienze Umanistiche - salvatore.spina@unict.it

<sup>7</sup>Ce.Di.Pa., Università degli Studi di Perugia - paola.venuti@unipg.it

## ABSTRACT (ITALIANO)

In questo articolo presentiamo l'edizione digitale dei primi 16 taccuini dell'archeologo Paolo Orsi, conservati presso il Museo Archeologico Regionale "Paolo Orsi" di Siracusa. Il progetto è stato realizzato da un team interdisciplinare utilizzando la piattaforma Transkribus. Nell'articolo descriviamo brevemente la storia dei taccuini dell'archeologo, le attività di trascrizione e codifica delle pagine, segmentazione delle immagini, annotazione dei testi, pubblicazione dei risultati, dopo una sintesi dello stato dell'arte delle edizioni digitali di diari e manoscritti.

**Parole chiave:** edizioni digitali; Transkribus; manoscritti; archeologia

## ABSTRACT (ENGLISH)

*The Digital Edition of the First 16 Notebooks of Paolo Orsi.*

In this article we present the digital edition of the first 16 notebooks of the archaeologist Paolo Orsi, preserved at the Regional Archaeological Museum "Paolo Orsi" in Syracuse. The project was carried out by an interdisciplinary team using the Transkribus platform. We briefly present the history of the archaeologist's notebooks, the activities of transcription and encoding of the pages, image segmentation, text annotation, and the publication of the results, following a summary of the current state of digital editions of diaries and manuscripts.

**Keywords:** digital scholarly editing; Transkribus; manuscripts; archaeology

## 1. INTRODUZIONE

In questo articolo presentiamo il progetto di creazione dell'edizione digitale dei primi 16 taccuini di scavo dell'archeologo Paolo Orsi. Gli originali dei taccuini sono conservati presso il Museo Archeologico Regionale "Paolo Orsi" di Siracusa, che fa parte del Parco Archeologico e Paesaggistico di Siracusa, Eloro, Villa del Tellaro e Akrai.

Il progetto dell'edizione digitale dei primi 16 taccuini è stato realizzato nell'autunno del 2023 e ha fatto parte di un progetto più ampio della Regione Siciliana denominato "Realizzazione spazio virtuale del Parco Archeologico di Siracusa" (PON Cultura e Sviluppo FESR 2014-2020).

Dopo l'analisi dei requisiti e la valutazione degli esempi e degli strumenti esistenti, abbiamo deciso di utilizzare la piattaforma Transkribus (<https://www.transkribus.org/>) sia per la correzione, l'allineamento e l'annotazione delle trascrizioni, sia per la pubblicazione online dell'edizione digitale, che si trova all'indirizzo <https://app.transkribus.org/sites/taccuini-paolo-orsi>. L'edizione digitale contiene il facsimile e la trascrizione allineata di circa 2000 pagine singole, accessibili tramite un sistema di navigazione, un motore di ricerca full text, e le annotazioni delle citazioni dei luoghi e delle date. Il lavoro è stato effettuato da un team interdisciplinare di 6 persone.

## 2. STATO DELL'ARTE DELLE EDIZIONI DIGITALI DI DIARI E LETTERE MANOSCRITTE

L'analisi dello stato dell'arte è stata effettuata nei mesi di settembre e ottobre 2023 e si riferisce quindi alla situazione delle edizioni digitali di manoscritti antecedenti a quella data. I 3 esempi principali che abbiamo analizzato sono stati i seguenti:

- "Bellini Digital Correspondence (BDC)", <https://bellinicorrespondence.cnr.it/> : progetto di sviluppo e fruizione dell'edizione scientifica digitale delle lettere manoscritte del compositore Vincenzo Bellini conservate presso il Museo Civico Belliniano di Catania realizzato da ricercatori di CNR-ILC, CNR-ISTC, Università di Catania, in collaborazione con studenti dell'Università di Pisa, tra il 2017 e il 2022 (Del Grosso & Spampinato, 2023; Del Grosso et al., 2019). "Il carteggio comprende 40 unità testuali distribuite su un alto numero di carte, riprodotte in oltre 100 immagini digitali" (Del Grosso & Spampinato, 2023). La piattaforma software open source utilizzata per la pubblicazione dell'edizione digitale è EVT - Edition Visualization Technology (<https://github.com/evt-project/evt-viewer/>), sviluppata a partire dal 2014 (Rosselli Del Turco, 2014; Cacioli et al., 2022).
- "Il quaderno di Paolo Bufalini", <https://projects.dharc.unibo.it/bufalini-notebook/> : un'edizione scientifica digitale semantica del quaderno di Paolo Bufalini intitolato "Appunti 1981-1991", che si pone l'obiettivo di valorizzare la complessa rete di relazioni intratestuali ed intertestuali che caratterizza il quaderno. Realizzata dal Dipartimento di Filologia Classica e Italianistica (FICLIT) e del Digital Humanities Advanced Research Center (/DH.ARC) dell'Università di Bologna (D'Aquino et al., 2020). Il quaderno si compone di 145 pagine rilegate e due carte sciolte. L'applicazione web è stata sviluppata ad hoc e si basa su dati RDF archiviati in una database a grafo (<https://projects.dharc.unibo.it/bufalini-notebook/specifications> ).
- il "Codice Pelavicino Edizione Digitale" (<https://pelavicino.labcd.unipi.it/evt/index.html> ): edizione digitale creata dal 2014 al 2020 da un team interdisciplinare guidato da Enrica Salvatori come "un lavoro i corso", nel senso che "documenti, dati e note sono stati aggiunti mano a mano che procedeva nella trascrizione, annotazione, codifica e pubblicazione. Una parte rilevante del progetto ha previsto infatti – e ancora prevede in parte – la partecipazione attiva di studiosi e appassionati nel correggere interpretazioni e aggiungere dati alle note" (Salvatori et al. 2020). Piattaforma software open source utilizzata per la pubblicazione dell'edizione digitale: EVT - Edition Visualization Technology.

La letteratura scientifica rilevante per il progetto dei Taccuini di Paolo Orsi ha incluso i lavori di trascrizione e annotazione di manoscritti epigrafici (tra i quali Calvelli et al., 2023), oltre ovviamente alla letteratura recente relativa a metodi, strumenti ed esempi di edizioni scientifiche digitali (a partire da Pierazzo, 2015; Franzini et al., 2016). Un'analisi recente della situazione delle edizioni digitali di manoscritti in Italia si trova in (Senatore et al., 2024).

### 3. I TACCUINI DI PAOLO ORSI

Paolo Orsi, archeologo roveretano vissuto dal 1859 al 1935, responsabile della Soprintendenza e direttore del Museo Archeologico Nazionale di Siracusa dal 1891 al 1934, è unanimemente considerato uno dei padri fondatori dell'Archeologia in Italia, e il pioniere di una nuova metodologia che iniziò ad affermarsi in Italia proprio grazie a lui. Le sue meticolose indagini dei terreni, lo studio del rapporto tra gli strati archeologici e il metodo rigoroso, con cui condusse vasti scavi di necropoli e santuari, furono presi a modello da numerosi altri studiosi, e, ancora oggi, gli esperti sottolineano l'incidenza profonda che il lavoro di Orsi ebbe nello sviluppo della Scienza archeologica. Della sua attività di archeologo, che si concentrò in particolare in territorio siciliano e magnogreco, danno soprattutto testimonianza i taccuini, che Orsi ha redatto da poco prima del suo arrivo a Siracusa, fino al suo collocamento a riposo (1888-1934). La maggior parte di questi documenti rimase di proprietà della famiglia, finché, nel 1964, furono acquistati dalla Soprintendenza Archeologica di Siracusa e destinati ad essere conservati presso l'allora Museo Archeologico Nazionale di Siracusa - oggi Museo Archeologico Regionale "Paolo Orsi". Inizia quindi il lavoro per realizzare le prime modalità di fruizione dei preziosi documenti, che vennero microfilmate, in modo da renderne consultabili le copie presso la Soprintendenza. Nello stesso tempo si avviano le prime trascrizioni, la redazione di indici e si avvia l'iter finalizzato alla edizione cartacea dei taccuini. I contatti con l'Accademia dei Lincei (determinante il ruolo della Professoressa Pelagatti, che per anni ha curato i rapporti con l'Accademia) individuata come la sede più adatta alla loro pubblicazione (Orsi appartenne all'Accademia fin dal 1896, pubblicando gran parte del suo lavoro in "Monumenti antichi dei Lincei e Notizie scavi") furono avviati da Concetta Ciurcina, direttrice del Museo archeologico di Siracusa, e proseguiti da Bice Basile. Con Gioconda Lamagna e Giuseppina Monterosso si arriva alla edizione del primo volume nel 2018 (taccuini 1-4) dopo un iter molto complesso; Giuseppina Monterosso prosegue il lavoro con la pubblicazione del secondo volume (taccuini 5-16) nel 2022 (Orsi, 2018; Orsi, 2022).

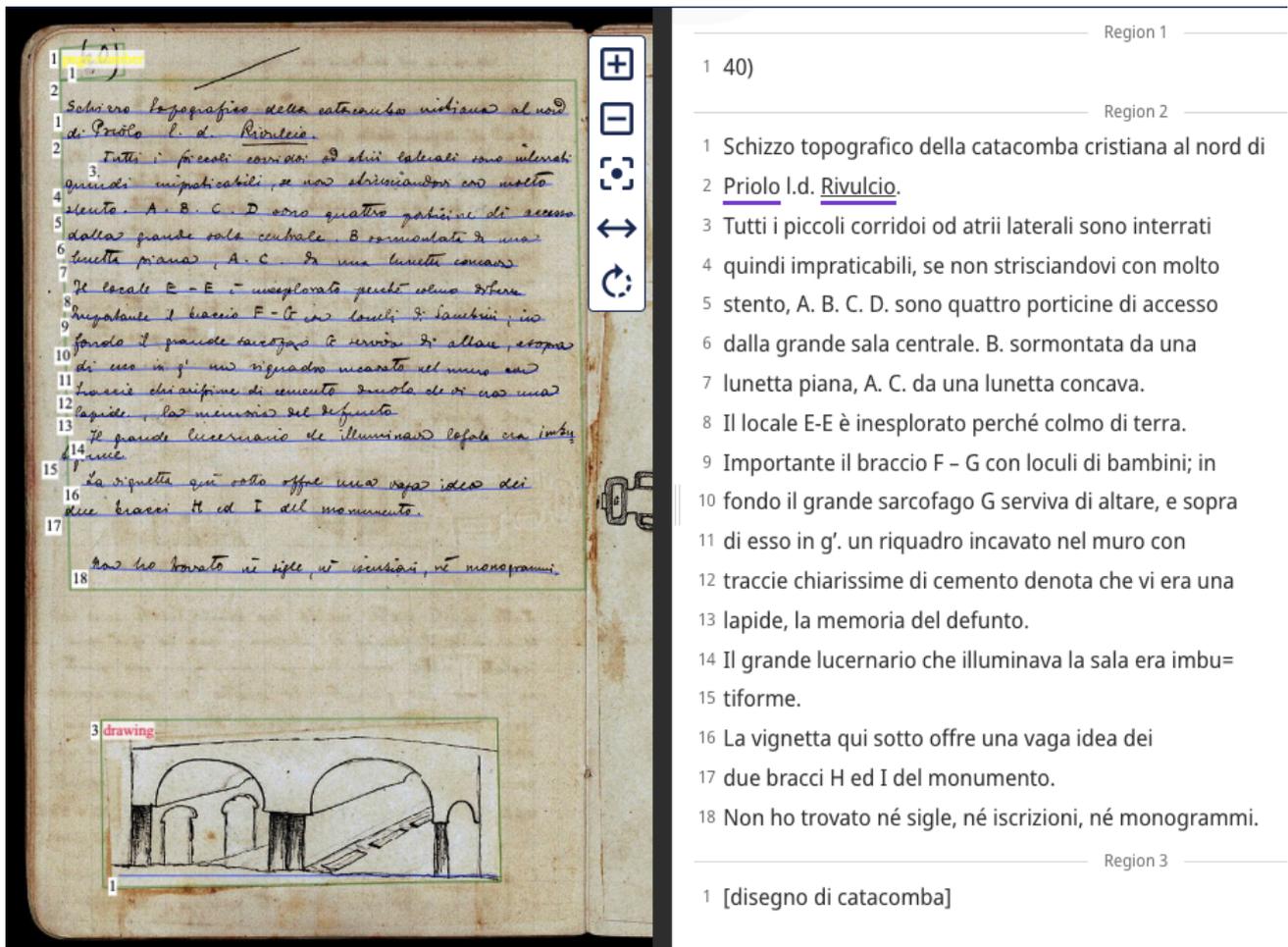
Durante la sua attività di ricerca, Orsi redige 151 taccuini. Di questi, 137 rimasero in possesso degli eredi, ed entrarono a far parte del patrimonio del Museo siracusano nel 1964. Altri 11, mai posseduti dalla famiglia e depositati presso la Soprintendenza Archeologica di Reggio Calabria, furono restituiti nel 1985. I taccuini n. 41 e 49 risultano invece dispersi. Si tratta di piccoli quaderni di formato tascabile (le dimensioni sono di circa 150/160 mm. di altezza e poco meno di 100 di larghezza), con pagine a quadretti tenute insieme da una cucitura a macchina e con coperta di carta telata nera, rivestita di carta più leggera, nei vari toni del beige o del rosa.

I primi 16 taccuini coprono un arco cronologico che va dal 1888 agli inizi del 1893. In quel periodo, Orsi si occupò soprattutto degli scavi archeologici a Siracusa e Megara Hyblea, in Sicilia, e Locri Epizefiri, in Calabria. Nei taccuini vengono registrate le sue giornate di lavoro, i frequenti spostamenti da uno scavo e all'altro, descrivendo e disegnando i manufatti ritrovati, o quelli già conservati nei musei da lui visitati. Orsi realizza schizzi topografici delle aree di scavo, mette a confronto le informazioni ricavabili dalle fonti greche e latine con i suoi ritrovamenti, dà notizia di pubblicazioni coeve di interesse archeologico; non mancano appunti più coloriti riguardanti la sua quotidianità.

#### **4. L'EDIZIONE DIGITALE IN TRANSKRIBUS**

Per il progetto dell'edizione digitale dei primi 16 Taccuini di Paolo Orsi abbiamo scelto di utilizzare la piattaforma Transkribus (<https://www.transkribus.org/>): sviluppata a partire dal 2016, Transkribus è una piattaforma che consente la trascrizione e l'annotazione manuale collaborativa dei documenti (Kahle et al, 2017), il riconoscimento automatico delle grafie manoscritte e dei caratteri a stampa. Negli ultimi anni, inoltre, è stata aggiunta la possibilità di pubblicare online l'edizione digitale, creando il sito web per la consultazione e l'accesso alle informazioni (<https://www.transkribus.org/sites>), ed aggiornala in tempo reale.

Abbiamo utilizzato le fotografie delle pagine dei taccuini realizzate dai fotografi del Museo nel corso degli anni, evitando quindi una nuova digitalizzazione. Abbiamo effettuato un editing delle fotografie, creando una immagine per ogni pagina dei taccuini, in modo uniforme. Abbiamo utilizzato anche le trascrizioni della riproduzione anastatica, effettuandone una revisione completa. Abbiamo quindi proceduto alla segmentazione delle immagini, all'allineamento dei testi, al tagging manuale delle citazioni di luoghi e date. La codifica del testo è in formato TEI-XML ed è esportabile dalla piattaforma in vari formati. Contemporaneamente alla codifica dei testi, abbiamo costruito il sito web attraverso l'app "Transkribus Sites" (<https://app.transkribus.org/sites/taccuini-paolo-orsi>). Nella figura seguente (Fig. 1) si vede un esempio di pagina segmentata in regioni e linee a sinistra, e a destra la trascrizione e le annotazioni (pag 40 del Taccuino 4).



**Figura 1. Esempio di pagina (pag 40 del Taccuino 4) segmentata in regioni e linee a sinistra, con a destra la trascrizione e le annotazioni, nella piattaforma Transkribus.**

Il motivo principale della scelta della piattaforma Transkribus è stato la necessità di lavorare collaborativamente e velocemente alla trascrizione e codifica delle pagine, segmentazione delle immagini, annotazione dei testi e contemporaneamente alla pubblicazione dei risultati, per consentire la visione e il controllo da parte delle curatrici del museo e le responsabili del progetto. I dati annotati in modo standard potranno essere riutilizzati in altri sistemi di visualizzazione di TEI-XML ed eventualmente integrati o corretti.

## 5. SVILUPPI FUTURI

La creazione dell'edizione digitale dei primi 16 taccuini di Paolo Orsi vuole essere la prima fase di un progetto collaborativo più ampio di pubblicazione online di tutti i taccuini manoscritti dell'archeologo. L'esigenza di una edizione digitale è stata manifestata da molti ricercatori nel corso degli anni; tra questi citiamo ad es. Jonathan Prag in occasione della pubblicazione cartacea dei primi 4 taccuini: "Given the existence already of high quality digital scans, and the desirability, over time, of providing supporting information or annotation (rather than requiring any scholar who approaches the volumes to repeat the work of identification and bibliographic investigation which others may already have undertaken), a digital solution, perhaps using the increasingly widely adopted International Image Interoperability Framework, which would facilitate viewing, comparison and annotation of the pages, might ultimately be of greater value" (Prag, 2021).

Negli ultimi anni, la svolta digitale e le tecnologie di Intelligenza Artificiale hanno ampliato le frontiere della ricerca umanistica, che può avvalersi di tali tecnologie in diversi momenti della ricerca, quali l'analisi dei testi, la ricostruzione di network, la visualizzazione delle informazioni in mappe e timeline. Tra le varie tecnologie, quelle di Handwritten Text Recognition (HTR) rappresentano un'innovazione utile perché

consentono la trascrizione semi-automatica dei documenti d'archivio manoscritti, e facilitano quindi la realizzazione di edizioni digitali (Spina, 2023).

In questo progetto sui primi 16 taccuini di Paolo Orsi abbiamo provato ad addestrare un modello di HTR sulla scrittura di Paolo Orsi, che potrà essere perfezionato e quindi utilizzato nella trascrizione semi-automatica dei successivi taccuini.

## RINGRAZIAMENTI

Il progetto dell'edizione digitale dei primi 16 taccuini ha fatto parte di un progetto più ampio della Regione Siciliana denominato "Realizzazione spazio virtuale del Parco Archeologico di Siracusa" (PON Cultura e Sviluppo FESR 2014-2020). Si ringrazia la Dott.ssa Rosa Lanteri e lo staff del Parco Archeologico e Paesaggistico di Siracusa, Eloro, Villa del Tellaro e Akrai.

## BIBLIOGRAFIA

- Cacioli, G., Cerretini, G., Chiara, D. P., Maenza, S., Rosselli Del Turco, R., & Zenzaro, S. (2022). There and back again: what to expect in the next evt version. In *AIUCD 2022-Culture digitali. Intersezioni: filosofia, arti, media. Preceedings della 11a conferenza nazionale*, Lecce, 2022 (pp. 212-217). Associazione per l'Informatica Umanistica e la Cultura Digitale.
- Calvelli, L., Boschetti, F., & Tommasi, T. (2023). EpiSearch. Identifying Ancient Inscriptions in Epigraphic Manuscripts. *Journal of Data Mining & Digital Humanities*, (Sciences of Antiquity and digital humanities).
- Codice Pelavicino. Edizione digitale, a cura di E. Salvatori, E. Riccardini, R. Rosselli del Turco, L. Balletto, C. Alzetta, C. Di Pietro, C. Mannari, R. Masotti, A. Miaschi, 2a ed., 2020  
<<http://pelavicino.labcd.unipi.it>> ISBN 978-88-944430-2-8 DOI:  
<https://doi.org/10.13131/978-88-944430-2-8>
- Daquino, M., Dello Buono, M., Giovannetti, F., & Tomasi, F. (2020). Paolo Bufalini, Appunti (1981-1991) [Edizione scientifica digitale semantica]. Digital Humanities Advanced Research Centre (/DH.arc), Bologna, <https://doi.org/10.6092/unibo/amsacta/6415>
- Del Grosso, A. M., & Spampinato, D. (Eds.). 2023. *Bellini Digital Correspondence*. CNR Edizioni, 2023. ISBN: 978-88-8080-562-5.
- Del Grosso, A. M., Capizzi, E., Cristofaro, S., De Luca, M. R., Giovannetti, E., Marchi, S., Seminara, G. & Spampinato, D. (2019, dicembre). Bellini's Correspondence: A Digital Scholarly Edition for a Multimedia Museum. *Umanistica Digitale*, 7, 23-47. <http://doi.org/10.6092/issn.2532-8816/9162>
- Franzini, G., Mahony, S., and Terras, M. (2016), 'A Catalogue of Digital Editions', In: Pierazzo, E. and Driscoll, M. J. (eds) *Digital Scholarly Editing: Theories and Practices*. Cambridge: Open Book Publishers. DOI: 10.11647/OBP.0095.09
- Kahle, P., Colutto, S., Hackl, G., & Mühlberger, G. (2017, November). Transkribus-a service platform for transcription, recognition and retrieval of historical documents. In *2017 14th iapr international conference on document analysis and recognition (icdar)* (Vol. 4, pp. 19-24). IEEE.
- Orsi, P. (2018). I Taccuini, I. Riproduzione anastatica e trascrizione dei Taccuini 1-4. A cura di Gioconda Lamagna, Giuseppina Monterosso. *Monumenti Antichi Accademia Nazionale dei Lincei*, 75, Giorgio Bretschneider Editore, ISBN: 9788876892998
- Orsi, P. (2022). I Taccuini, II. Riproduzione anastatica e trascrizione dei Taccuini 5-16. *Monumenti Antichi Accademia Nazionale dei Lincei*, 84, Giorgio Bretschneider Editore, ISBN: 9788876893407
- Pierazzo, E. (2015). *Digital Scholarly Editing: Theories, Models and Methods*. Farnham, Surrey: Ashgate.
- Prag, J. R. (2021). Paolo Orsi, «un modestissimo cultore della epigrafia»: Paolo Orsi's contribution to Sicilian epigraphy as illustrated by the newly published 'taccuini'. *Atti della Accademia Nazionale del Lincei. Rendiconti della Classe di Scienze Morali, Storiche, Filologiche*, 32(1-2)
- Rosselli Del Turco, R., Buomprisco, G., Pietro, C. D., Kenny, J., Masotti, R., & Pugliese, J. (2014). Edition visualization technology: A simple tool to visualize TEI-based digital editions. *Journal of the Text encoding initiative*, (8).
- Senatore, C., Del Turco, R. R., & Pisano, P. (2024). Digital Resources for Manuscripts: Between Fragmentation and Development Prospects. *Mimesis Journal*, 13(2), 104-118
- Spina, S. (2023). Handwritten Text Recognition as a digital perspective of Archival Science. *Aidainformazioni*, (1-2), 115-132.