



The Digital Scholarly Edition of Angelo Poliziano's Marginal Notes to the *Epistle of Sappho to Phaon*. Encoding of Multiple Edition Levels and Implementation of Wikidata Items and CTS URNs

DATA PAPER

ANDREA LA VEGLIA 

]u[ubiquity press

ABSTRACT

The paper presents the dataset related to the digital scholarly edition (DSE) of the first part of marginal notes written by Poliziano (1454–1494) to the *Epistle of Sappho to Phaon* of Ovid's *Heroides*. The text is encoded according to the *Text Encoding Initiative* (TEI) XML schemas and the visualization of the DSE uses the *Edition Visualization Technology* (EVT) software (version 2). The encoding implements references to Linked Open Data (LOD), including *Wikidata*, and to open *corpora* via the Canonical Text Service (CTS) protocol. The semantic annotation was carried out by using *Transkribus* software and a pipeline was developed to generate a consistent TEI encoding, through XSLT stylesheets, which are available in the dataset.

CORRESPONDING AUTHOR:

Andrea La Veglia

Department of Humanities,
University of Salerno,
Fisciano, Italy

alaveglia@unisa.it

KEYWORDS:

digital scholarly edition;
edition visualization
technology (EVT); Wikidata;
semantic annotation; text
encoding initiative (TEI);
named entities

TO CITE THIS ARTICLE:

La Veglia, A. (2026). The Digital Scholarly Edition of Angelo Poliziano's Marginal Notes to the *Epistle of Sappho to Phaon*. Encoding of Multiple Edition Levels and Implementation of Wikidata Items and CTS URNs. *Journal of Open Humanities Data*, 12: 44, pp. 1–8. DOI: <https://doi.org/10.5334/johd.493>

<https://doi.org/10.5281/zenodo.17315766>

https://github.com/AndreaLaVeglia/Postille_Poliziano_Sappho_EVT

The dataset presented here will be discussed more extensively in a forthcoming paper in the second issue of the journal TeCuM – Testi e Culture del Medioevo (<https://serena.sharepress.it/TECUM/>).

CONTEXT

This dataset is related to the digital scholarly edition (DSE) of Poliziano's annotations on the *Epistle of Sappho* that I developed for my Master of Art thesis at the University of Naples "Federico II". It includes the encoding of the text, the settings of the software viewer and the XSLTs used.

Poliziano's marginal notes on the *Epistle of Sappho* can be read in the Bodleian incunabulum Auct. P 2.2 (ff. 238v-241r), which is a copy of the 1477 Parma edition of Ovid's *Opera Omnia*.¹ Among all the annotations in the incunabulum, the marginal notes to the *Epistle of Sappho* are the most conspicuous (see Brenkman, 1722) and were first studied by M. Kubo (1985) who made a draft transcription of the text that cannot be considered as a proper edition.

These annotations represent an intermediate stage of work between the study of primary sources and the academic lesson (*enarratio*) on this epistle held by Poliziano in Florence in 1481.² The comparison between the marginal notes and the *enarratio* demonstrates that the second one represents an advanced stage of research, since the marginal notes "contained little or no material substance that were not to be later utilized for the *Enarratio*" (Kubo, 1985, p. 35).³

For my dissertation I completed, revised and corrected the transcription of Kubo and then I prepared a digital encoding useful to export both a printed and a digital edition.

The DSE was created using the second version of *Edition Visualization Technology* (EVT 2), an open source software viewer for texts encoded according to the Text Encoding Initiative (TEI) guidelines, publishing the text along with related annotations and metadata (Rosselli Del Turco et al., 2014, 2019).⁴

In my edition, three view modes of EVT 2 are available: the "image-text" (*mode-imgTxt*, see Figure 1), the "text-text" (*mode-txtTxt*, see Figure 2) and the "source-text" (*mode-srcTxt*, see Figure 3) (Rosselli Del Turco et al., 2014, p. 11, 2019, p. 13, 2020). In the first mode you can select one of three parallel edition levels (facsimile, diplomatic-interpretative and critical) and see the text-image alignment; in the second mode you can compare different transcriptions and in the third mode you can explore the sources of quotes in the text.

In all view modes the named entities are highlighted and included in an interactive button that can be clicked to open a pop-up card displaying the corresponding Wikidata item's Uniform Resource Locator (URL) and a list of the entity's occurrences throughout the edition (Figure 4).

¹ The incunabulum has been studied by Owen (1889, pp. XII–XVI), Cotton (1937, p. 397), Maier (1965, pp. 350–351), Daneloni (2013, p. 311) and Villani (2018, pp. 1037–1042).

² Text published by E. Lazzeri (1971). With regard to the professorship of Poliziano at the University of Florence see: Branca (1983, pp. 73–90), Cesarini Martinelli (1996), Mandosio (2008), Orvieto (2009, pp. 324–326).

³ About this point, see also Cesarini Martinelli (1978, p. 143).

⁴ Official website of the EVT project: <http://evt.labcd.unipi.it/>.

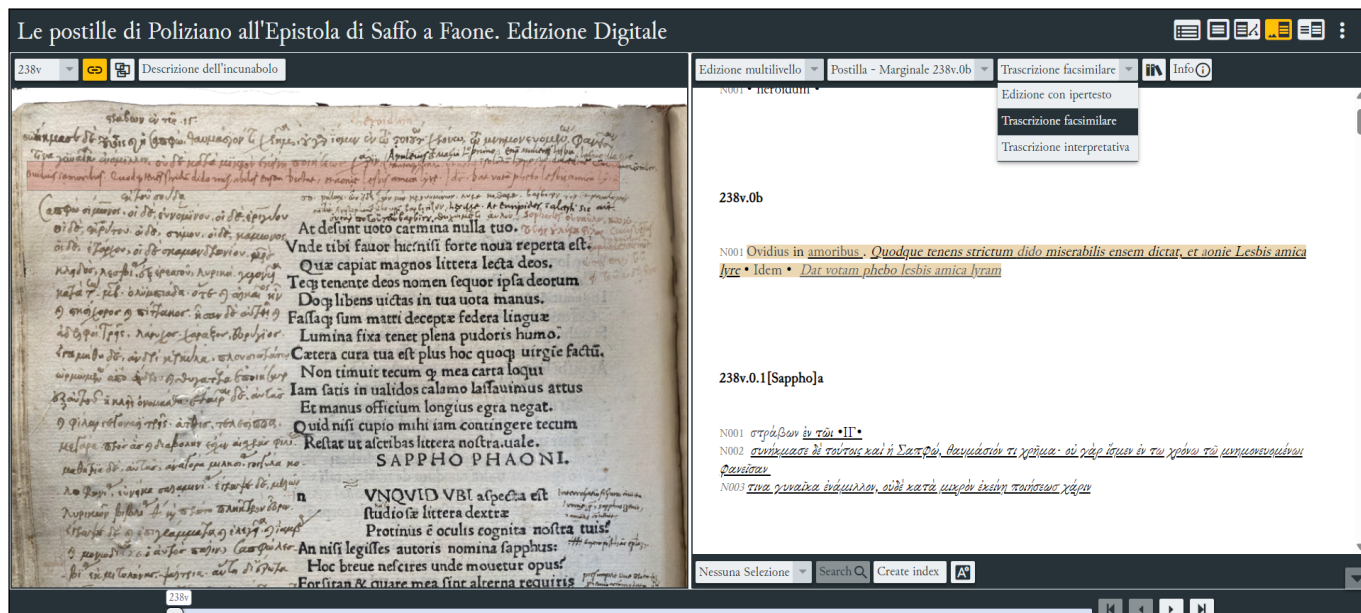


Figure 1 The image-text view mode and the selector of transcription level.

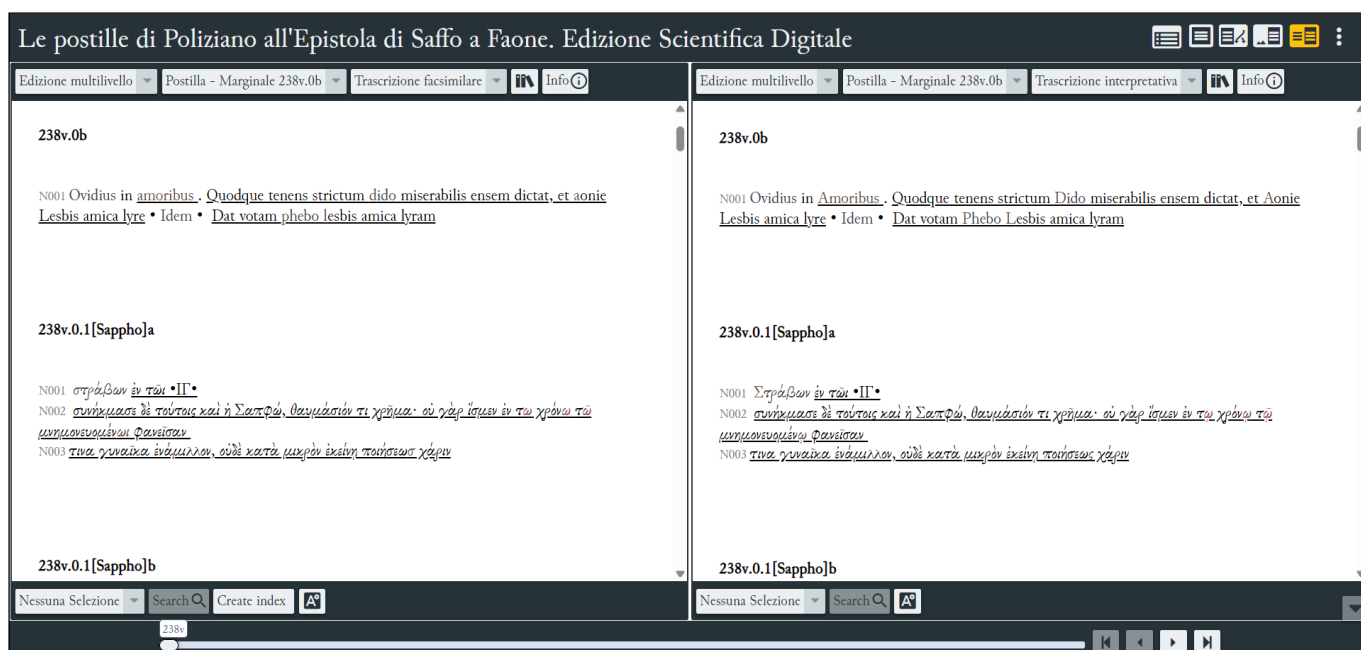


Figure 2 The text-text view mode.

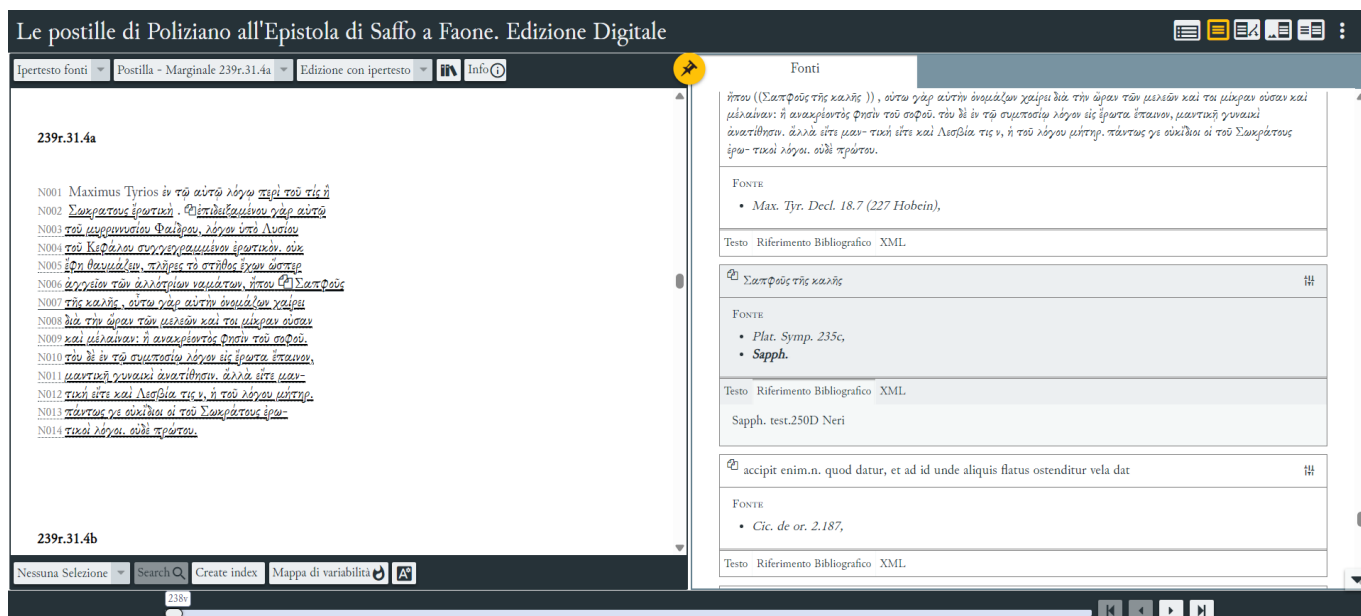


Figure 3 The source-text view mode.

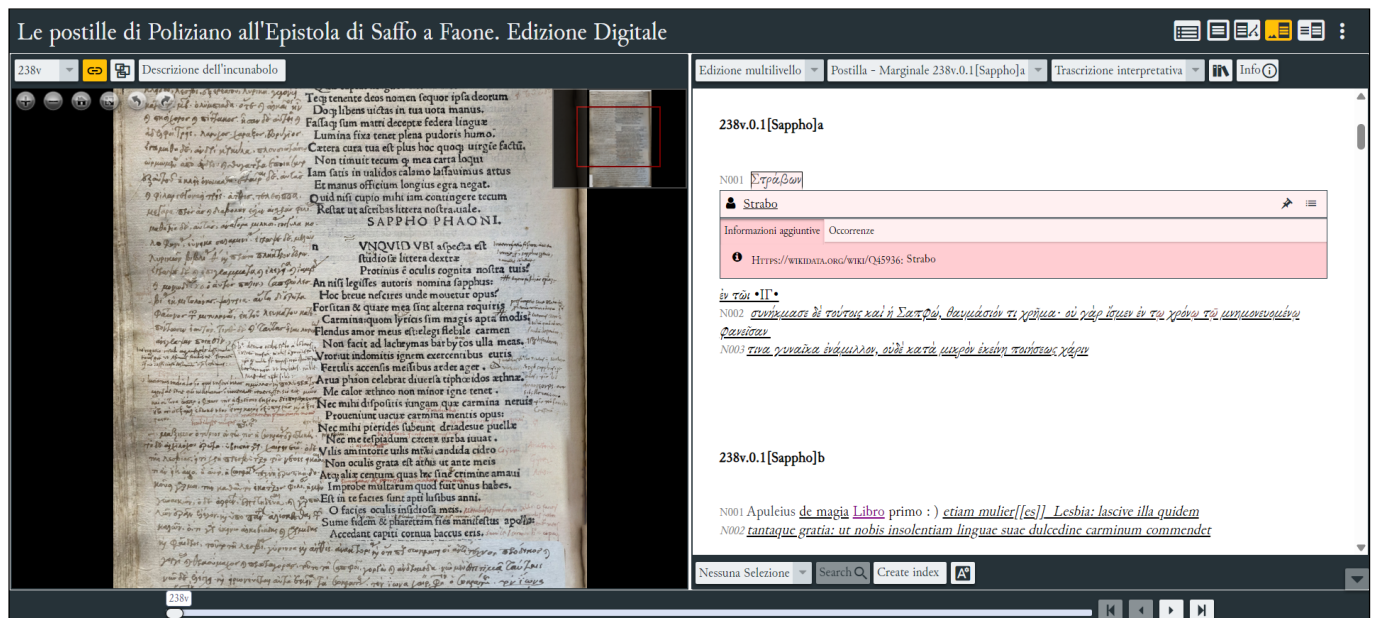


Figure 4 The Wikidata implementation for the named entity Strabo.

(2) METHOD

The digital space has enabled a multidimensional representation of texts, overcoming the limitations of the printed page, which is confined by its nature to a bidimensional space (see Reggiani, 2025). In this case study, I faced the challenging editing of an apparatus of marks in a book (for genre definition see Stoddard, 1985; Petrella, 2022), namely all the annotations made in a book by its reader: due to its complexity and asystematicity, the edition of such a text is only possible if it is conceived on multiple layers of representation.

In this DSE the following information is encoded in compliance with TEI guidelines:

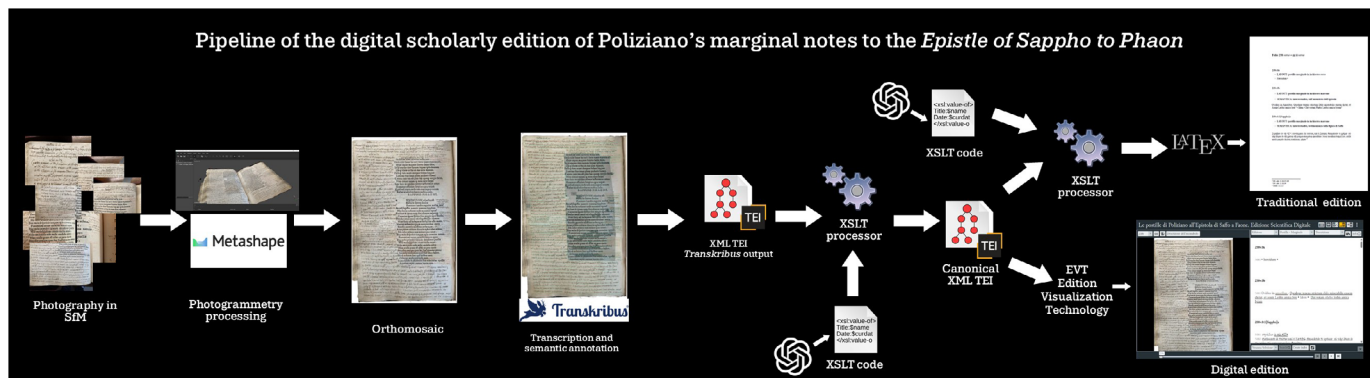
- semantic annotation of layout with attributes of <zone> in <facsimile>;
- codicological information with <msDesc>;
- categorization of each unit of text with <taxonomy> and attributes of <div> (taking inspiration from the model of Siciliano & Del Grosso, 2022);
- double level of transcription using the <choice>;
- palaeographical analysis with <add> and <unclear>;
- identification of sources with <quote> connected to a <cit> in <back>, with encoding of reference to online corpora via Uniform Resource Names (URNs) of Canonical Text Services (CTS);⁵
- annotation of named entities by using <persName> and <placeName> nested within a <ref> element which refers to the Wikidata URL through the @target attribute;⁶
- identification of signs used by Poliziano to relate each annotation to the main text of incunabulum by using <metamark> and its @target attribute and a nested <g> element related to a series of <glyph> declared in <charDecl>.

STEPS

The work was initiated from a sequence of close-up photographs taken by Prof. Giovan Battista D'Alessio at the Bodleian Library, which I processed in *Metashape* to produce HD pictures of full pages.

5 The Canonical Text Service (CTS) is a protocol to create machine-readable citations by using Uniform Resource Names (URNs). The CTS protocol is based on the CITE ("Collections, Indices, Texts, and Extensions") Architecture that is a framework for identifying citations by considering the text as an "Ordered Hierarchy of Citable Objects" (OHCO) (with regard to URN: Moats, 1997; with regard to OHCO: Smith & Weaver, 2009; with regard to CITE architecture and CTS URNs: Blackwell & Smith, 2019; Tiepmar & Heyer, 2019; Berti, 2021, pp. 105–114, 2023, pp. 318–319).

6 See the TEI P5 guidelines for implementation of Wikidata references in <listPerson> and <listPlace> at <https://www.tei-c.org/Vault/P5/4.9.0/doc/tei-p5-doc/en/html/examples-listPerson.html> and <https://www.tei-c.org/Vault/P5/4.9.0/doc/tei-p5-doc/en/html/examples-listPlace.html>.



Afterward, I made use of *Transkribus* (© ReadCoop)⁷ for a manual transcription and a semantic annotation of layout and text. In addition to native tags, I created custom tags according to TEI guidelines (Figure 5).

Figure 5 Pipeline of the project.

Regarding the named entities, I took advantage of the *Transkribus* tool “Wikidata ID”, which allows connecting each name with its related Wikidata item and store its identifier number as an attribute. Precisely, in the web app, after choosing a tag (e.g. ‘person’ or ‘place’) by clicking the Wikidata ID button, the system will automatically suggest an entity linked to the selected term. The user can select the suggested item or alternatively use the search field to look up the correct Wikidata entry.⁸ Since the version released in May 2025, it has been possible to automatically tag all occurrences of the same term in the *Transkribus* document, including the Wikidata ID attribute.⁹

After completing the transcription and annotation, I addressed the issues related to data export. While *Transkribus* enables the export of transcriptions and tags in TEI XML format, the resulting files do not always conform to EVT 2, and customized tags are not converted into canonical TEI elements. An additional encoding regularization step was therefore necessary. To carry out this transformation, I developed a set of XSLT stylesheets, with assistance from *ChatGPT*, as documented in the *GitHub* repository.

The first XSLT (*Pipeline-XSLT/01-regularize_TEI.xsl*) regularizes the elements `<facsimile>` (according to EVT 2 expectations), `<choice>`, `<sic>` and `<del>` and creates both the source apparatus and the index of named entities. The transformation of each element is described in the comments of the XSLT sheet.

Subsequently, the division of the whole text in glosses was managed by a second XSLT (*Pipeline-XSLT/02-create_div.xsl*) and I manually attributed a semantic ID for each gloss. Then, with a third XSLT (*Pipeline-XSLT/03-order_div.xsl*) all `<div>` elements were arranged in ascending order.

At the conclusion of this XSLT sequence, a line of validation code was added to ensure compliance with `tei_all` schema. This was followed by the manual categorization of each text division and the description of metatextual signs. Finally, all codicological information was recorded into the `<msDesc>` module, along with additional metadata in the `<fileDesc>`.

At the end of this regularization of the XML TEI encoding, I configured the visualization in EVT 2 by editing the settings in the config folder of the software (*DigitalEditionEVT/config*). Subsequently, the print edition was produced by exporting a LaTeX document from the XML encoding (*Pipeline-XSLT/06-LaTeX_export.xsl*), which has been useful to generate a Portable Document Format (PDF) (Figure 5).

(3) DATASET DESCRIPTION

REPOSITORY NAME

GitHub (https://github.com/AndreaLaVeglia/Postille_Poliziano_Sappho_EVT); Zenodo (<https://doi.org/10.5281/zenodo.17315766>).

⁷ <https://www.transkribus.org/>.

⁸ See <https://help.transkribus.org/textual-tags>.

⁹ See the change log (viewable only by registered users) <https://app.transkribus.org/changelog>, release of 07 May 2025.

OBJECT NAME

The Digital Scholarly Edition of Marginal Notes of Poliziano to the Epistle of Sappho to Phaon.

La Veglia
Journal of Open
Humanities Data
DOI: 10.5334/johd.493

6

FORMAT NAMES AND VERSIONS

XML, XSL

CREATION DATES

15-02-2025–13-02-2026

DATASET CREATORS

Andrea La Veglia (University of Salerno): creation and curation of dataset; Giovan Battista D'Alessio (University of Rome "La Sapienza"): image acquisition and supervision of transcriptions and philological work; Gennaro Ferrante (University of Naples "Federico II"): supervision of workflow and technical work; Roberto Rosselli Del Turco (University of Turin): creator and developer of software "EVT 2"; Luigi Bambaci (University of Bologna): assistance in developing of XSLT transformation from XML TEI to LaTeX; Angelo Mario Del Grosso (CNR-ILC): supervision and validation of XML TEI encoding.

LANGUAGE

English, Italian, Latin, Ancient Greek.

LICENSE

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International.

PUBLICATION DATE

2025-10-10.

The repository on GitHub is linked to Zenodo, according to the best practice for publishing digital editions, just as in the case of the *PerseusProject* (Cerrato, Babeu, et al., 2025; Cerrato, Clérice, et al., 2025) and of *BelliniDigitalCorrispondence* (Del Grosso & Spampinato, 2023).

The project includes a creation of a Wikidata item in case of missing correspondent for named entities, as in the case of the work *De Elocutione* by Demetrius the Rhetorician (<https://www.wikidata.org/wiki/Q131412205>).

(4) REUSE POTENTIAL

The first potential reuse of the dataset concerns the pipeline. The sequence of XSLT stylesheets can be applied to any TEI XML file exported from a *Transkribus* document, if it uses the same customized tags adopted in this project. In this way *Transkribus* can be used as an effective environment for producing scholarly digital editions. Moreover, as I have shown, once the digital edition is encoded, generating a print edition directly from this encoding is also feasible.

In addition, the encoding introduces innovative applications of the TEI guidelines for representing the relationships between main texts and paratexts (or specifically *postillae*) and proposes a model for encoding other similar paratexts rich in information distributed across multiple layers of analysis.

Finally, the inclusion of references to Wikidata items and CTS URNs contributes to making the edition both open and interoperable, enabling meaningful dialogue and data exchange with other projects.

ACKNOWLEDGEMENTS

The data paper presents the results of my Master of Arts dissertation presented at the University of Naples "Federico II" in February 2025, under the supervision of Prof. Giovan Battista D'Alessio and Prof. Gennaro Ferrante whom I thank for their guidance. My research benefited

from the activities organised by the FeDHLab, Digital Humanities laboratory at the Department of Humanities of the University of Naples.

I am very grateful to all experts that provided me invaluable support for the development of this work and for the paper: Roberto Rosselli del Turco, Angelo Mario Del Grosso, Camillo Carlo Pellizzari di San Girolamo, Luigi Bambaci, Nicola Reggiani, Thibault Clérice, Sara Mansutti, Monica Berti, Michele Giovanni Silani, Riccardo Montalto, Paolo Monella and Aanandavardhan Nyaupane.

I thank Andrea Farina, the editorial team of the JHOD and the anonymous reviewers for their suggestions and assistance in the publishing process.

COMPETING INTERESTS

The author was supported by the *Transkribus* Scholarship program for one year and is currently regional coordinator for *Wikimedia Italia*.

AUTHOR CONTRIBUTIONS

Andrea La Veglia: Conceptualization, Data curation, Investigation, Methodology.

AUTHOR AFFILIATIONS

Andrea La Veglia  orcid.org/0000-0001-9590-2610
Department of Humanities, University of Salerno, Fisciano, Italy

REFERENCES

- Berti, M. (2021). Digital Editions of Historical Fragmentary Texts. In *Propylaeum*. Propylaeum. <https://doi.org/10.11588/propylaeum.898>
- Berti, M. (2023). L'antichità greco Romana e le tecnologie digitali. In F. Ciotti (Ed.), *Digital humanities: Metodi, strumenti, saperi* (pp. 312–324). Carocci.
- Blackwell, C. W., & Smith, N. (2019). The CITE Architecture: A Conceptual and Practical Overview. In M. Berti (Ed.), *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution* (pp. 73–94). De Gruyter Saur. <https://doi.org/10.1515/9783110599572-006>
- Branca, V. (1983). *Poliziano e l'umanesimo della parola*. Einaudi.
- Brenkman, H. (1722). *Historia Pandectarum seu Fatum exemplaris florentini accedit Gemina dissertatio de Amalphi*. Guilielmus van de Water.
- Cerrato, L., Babeu, A., Almas, B., TDBuck, ahanhardt, Stella, Dik, H., Clérice, T., Crane, G., Fleischman, S., McCallum, N., Munson, M., Berra, A., Pletcher, C., Wegner, J., KATEBHN, Palladino, C., Mittmann, A., Sowell, E., ... Beine, J. J. (2025). *PerseusDL/canonical-greekLit: 0.0.19833599998* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.17781753>
- Cerrato, L., Clérice, T., Babeu, A., Almas, B., Jovanović, N., annettegessner, nkprasad12, Burns, P. J., Stella, mkonieczny9805, Munson, M., Jøhndal, M., Celano, G. G. A., maryam-foradi, zachhimes, Matt, MMernitz, Seydi, M., Bolner, T., ... TDBuck. (2025). *PerseusDL/canonical-latinLit: 0.0.19906058753* [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.17807341>
- Cesarini Martinelli, L. (1978). In margine al commento di Angelo Poliziano alle Selve di Stazio. *Interpres. Rivista Di Studi Quattrocenteschi*, 1, 96–145.
- Cesarini Martinelli, L. (1996). Poliziano professore allo studio fiorentino. In R. Fubini (Ed.), *La Toscana al tempo di Lorenzo il Magnifico: Politica, Economia, Cultura, Arte* (Vol. 2, pp. 463–482), Pacini Editore.
- Cotton, J. M. S. (1937). Ex Libris Politiani: II. Incunabula Bodleiana. *The Modern Language Review*, 32(3), 394–399. <https://doi.org/10.2307/3715916>
- Daneloni, A. (2013). Angelo Poliziano (Angelo Ambrogini). In F. Bausi, M. Campanelli, S. Gentile, & J. Hankins (Eds.), *Autografi dei letterati italiani. Il Quattrocento* (Vol. 1, pp. 295–392). Salerno Editrice.
- Del Grosso, A. M., & Spampinato, D. (2023). *Bellini Digital Correspondence* [Cnr Edizioni]. <https://bellinidigitalcorrespondence.cnr.it/>, <https://doi.org/10.5281/zenodo.8167516>
- Kubo, M. (1985). Sappho-Ovidius-Renaissance. A Description of the Marginal Notes to Sappho's Letter in the Parmensis 1477 (Bodl. Auct. P. II. 2). *Mediterraneus. Annual Report of the Collegium Mediterranistarum*, 8, 3–51.
- Lazzeri, E. (Ed.). (1971). *Angelo Poliziano. Commento inedito all'epistola ovidiana di Saffo a Faone*. Sansoni.
- Maier, I. (1965). *Les manuscrits d'Ange Politien: Catalogue descriptif, avec dix-neuf documents inédits en appendice*. Librairie Droz. <http://openurl.droz.org/?id=9782600030021>

- Mandosio, J.-M. (2008). Un enseignement novateur. Les cours d'Ange Politien à l'université de Florence (1480–1494). *Histoire de l'éducation*, 120, 33–52. <https://doi.org/10.4000/histoire-education.1830>
- Moats, R. (1997). *URN Syntax* (Request for Comments No. RFC 2141). Internet Engineering Task Force. <https://doi.org/10.17487/RFC2141>
- Orvieto, P. (2009). *Poliziano*. Salerno Editrice.
- Owen, S. G. (1889). *Ovidii Tristium libri V*. E typographeo Clarendoniano. <https://catalog.hathitrust.org/Record/011594603>
- Petrella, G. (2022). *Scrivere sui libri: Breve guida al libro a stampa postillato*. Salerno Editrice.
- Reggiani, N. (2025). The Digital Critical Edition of the Papyri: Topics, Issues and Perspectives. In N. Reggiani (Ed.), *Digital Papyrology III: The Digital Critical Edition of Greek Papyri: Issues, Projects, and Perspectives* (pp. 1–48). De Gruyter. <https://www.degruyterbrill.com/document/doi/10.1515/9783111070162-003/html>
- 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. <https://doi.org/10.4000/jtei.1077>
- Rosselli Del Turco, R., Di Pietro, C., & Masotti, R. (2020, June 30). *EVT 2.0—User Manual*. SourceForge. <https://sourceforge.net/projects/evt-project/files/EVT%20js%20-%20User%20manual.pdf/download>
- Rosselli Del Turco, R., Pietro, C. D., & Martignano, C. (2019). Progettazione e implementazione di nuove funzionalità per EVT 2: Lo stato attuale dello sviluppo. *Umanistica Digitale*, 3(7). <https://doi.org/10.6092/issn.2532-8816/9322>
- Siciliano, A., & Del Grosso, A. M. (2022). Dalla stampa al digitale: Un modello di codifica per l'edizione delle postille di Giorgio Bassani. *Umanistica Digitale*, 6(13), 21–48. <https://doi.org/10.6092/issn.2532-8816/13688>
- Smith, D., & Weaver, G. (2009). Applying Domain Knowledge from Structured Citation Formats to Text and Data Mining: Examples Using the CITE Architecture. *Computer Science Technical Reports*. https://digitalcommons.dartmouth.edu/cs_tr/326
- Stoddard, R. E. (Ed.). (1985). *Marks in Books, Illustrated and Explained*. Houghton Library.
- Tiepmar, J., & Heyer, G. (2019). The Canonical Text Services in Classics and Beyond. In M. Berti (Ed.), *Digital Classical Philology* (pp. 95–114). De Gruyter. <https://doi.org/10.1515/9783110599572-007>
- Villani, L. (2018). Poliziano lettore dello pseudo Ovidio: Le glosse alla Nux pseudo ovidiana. In C. Cocco, C. Fossati, A. Grisafi, F. Mosetti Casaretto, & G. Boiani (Eds.), *Itinerari del testo per Stefano Pittaluga: Vol. II* (pp. 1037–1050). Università di Genova, Scuola di scienze umanistiche, Dipartimento di antichità, filosofia e storia (Sezione D.AR.FI.CL.ET.).

TO CITE THIS ARTICLE:

La Veglia, A. (2026). The Digital Scholarly Edition of Angelo Poliziano's Marginal Notes to the *Epistle of Sappho to Phaon*. Encoding of Multiple Edition Levels and Implementation of Wikidata Items and CTS URNs. *Journal of Open Humanities Data*, 12: 44, pp. 1–8. DOI: <https://doi.org/10.5334/johd.493>

Submitted: 02 December 2025

Accepted: 29 January 2026

Published: 13 March 2026

COPYRIGHT:

© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Journal of Open Humanities Data is a peer-reviewed open access journal published by Ubiquity Press.