TEI and Scholarly Digital Editions

How to make philological data easier to retrieve and visualise

Chiara Martignano
chiara.martignano@unifi.it
University of Florence, University of Siena

THE IDEA

1. In the past few decades the number of TEI-encoded scholarly digital editions (SDEs) has risen significantly, as a consequence a big amount of philologically edited data is now available in a machine-readable form. One could try to apply computational approaches, in order to further study the linguistic data, the information about textual transmission, etc. contained in multiple TEI-encoded digital editions. The problem is that **retrieving philological data through different TEI-encoded SDEs is not that simple**.

2. Every TEI-encoded edition has its **own markup model**, designed to respond to the philological requirements of that particular edition. For this reason, it is difficult to **identify** the same types of **data** through different digital editions **unambiguously**.
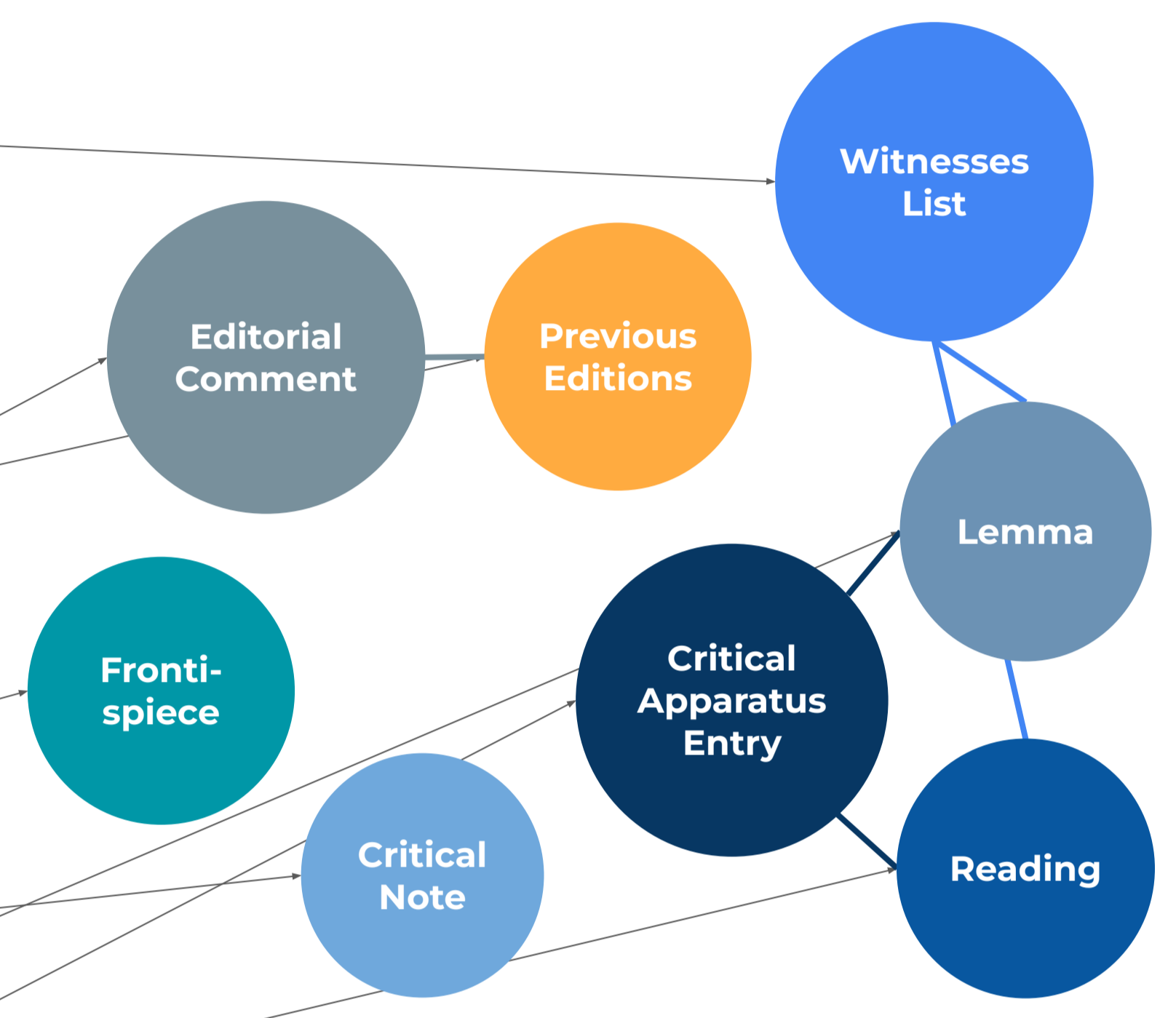
A MARKUP LAYER

Example taken from the digital critical edition of Petri de Ebulo De rebus siculis carmen, edited by Fulvio Delle Donne.

AN ABSTRACT MODEL FORMALISED AS AN ONTOLOGY

3. A possible way to simplify the retrieval of philological information from multiple digital editions is to **link** them t**o a same model** that is able to **represent SDEs on a more abstract level**. This abstract model could be formalised as an **ontology**. Then inside different TEI-encoded editions it would be possible to add a further **markup layer** that binds each abstract component to the corresponding class of the ontology.
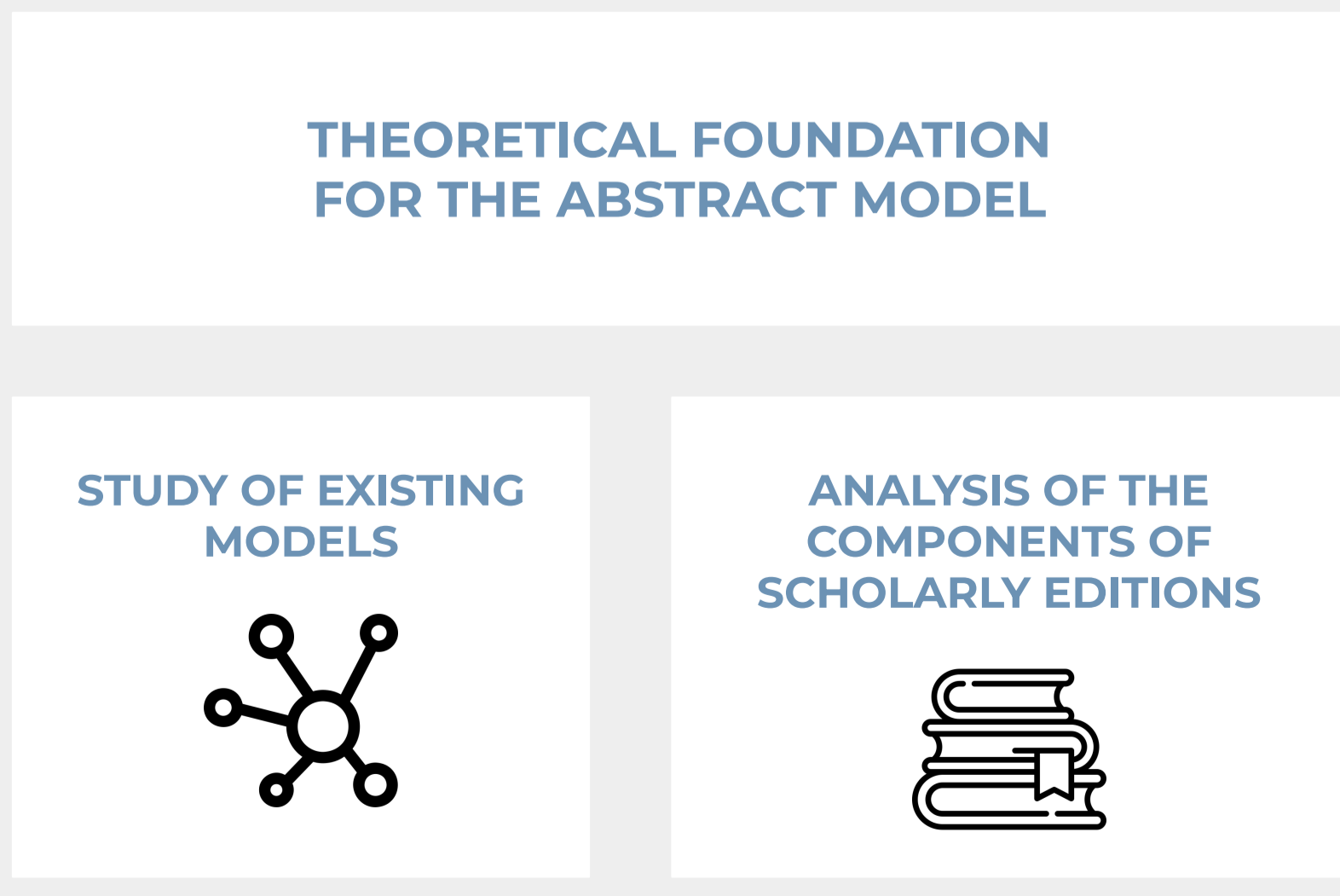
POSSIBLE APPLICATIONS

Testing the **completeness** and **expressiveness** of existing ontologies for the description of scholarly digital editions. See [2].

Queries across different TEI-encoded editions, with RDF and SPARQL. See [1], [4].

Easier **development of general purpose tools** for visualising SDEs thanks to a common abstract model. See [5].

CURRENT STATE OF THE RESEARCH

4. There are already some ontologies that were created for representing scholarly edited texts and, more specifically, the critical apparatus. My goal is to **use** these **existing models**, as well as the **TEI guidelines**, to **analyse different scholarly editions** (both digital and printed) and **identify all the abstract components** of a scholarly edition. The final goal of this preliminary research is to lay a **theoretical foundation** for an abstract model that can help make philological data more visible and easier to retrieve.

THEORETICAL FOUNDATION FOR THE ABSTRACT MODEL

STUDY OF EXISTING MODELS

ANALYSIS OF THE COMPONENTS OF SCHOLARLY EDITIONS

REFERENCES

1. Ciotti, Fabio, and Francesca Tomasi. 2016. 'Formal Ontologies, Linked Data, and TEI Semantics'. *Journal of the Text Encoding Initiative*, no. Issue 9 (September). https://doi.org/10.4000/jtei.1480.
2. Daquino, Marilena, Francesca Giovannetti, and Francesca Tomasi. 2019. 'Linked Data per le edizioni scientifiche digitali. Il workflow di pubblicazione dell'edizione semantica del quaderno di appunti di Paolo Bufalini'. Umanistica Digitale, no. 7 (December). https://doi.org/10.6092/issn.2532-8816/9091.
3. Doerr, Martin. 2003. 'The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata'. *AI Magazine* 24 (3): 75–75. https://doi.org/10.1609/aimag.v24i3.1720.
4. Eide, Øyvind. 2014. 'Ontologies, Data Modeling, and TEI'. *Journal of the Text Encoding Initiative*, no. Issue 8 (December). https://doi.org/10.4000/jtei.1191.
5. Martignano, Chiara. 2021. 'Un modello concettuale per favorire lo sviluppo e il riutilizzo di app per edizioni digitali'. *Umanistica Digitale*, no. 10 (September): 71–88. https://doi.org/10.6092/issn.2532-8816/12620.