Flying Digital Editions: A Tool for Enhancing Digital Text Exploration through Wikidata

Supervisor: Prof. Fabio Vitali Co-supervisor: PhD s. Andrea Schimmenti Presented by: Sara Vellone

March 11, 2024

Project Goal

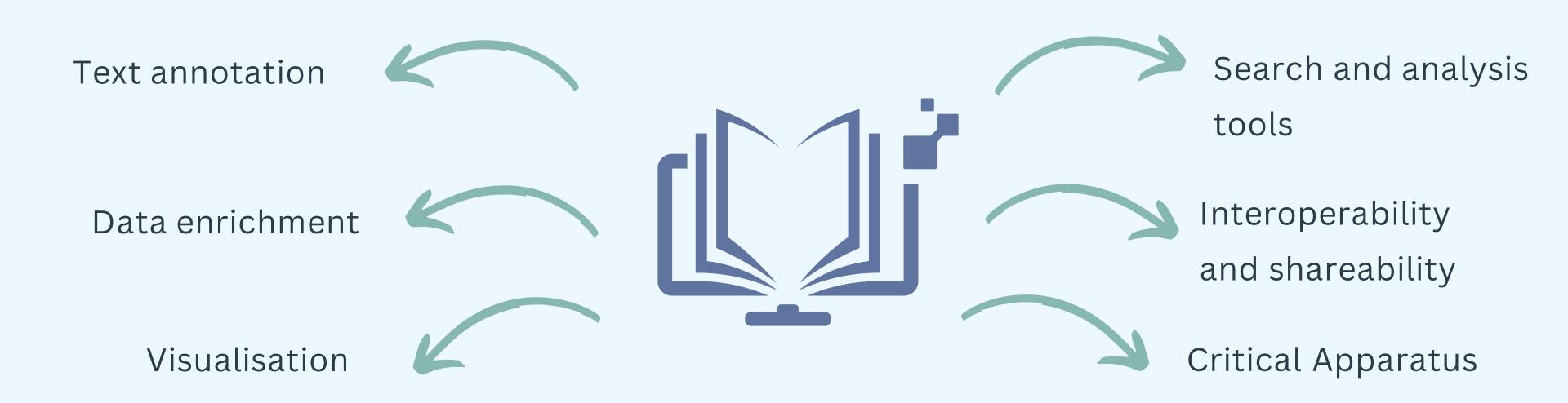
Flying Digital Editions (FDE) is a domain-agnostic tool that aims to turn any document into a *lightweight digital edition*. By reusing existing tools and leveraging large *Knowledge Bases*, FDE can annotate the text in the document with available and authoritative information.

Research Questions

- ♦ What are the main features of a Digital Edition?
- → Which of the main features of a Digital Edition can be turned into automatic processes?
- → How and at what cost is it feasible to do so?



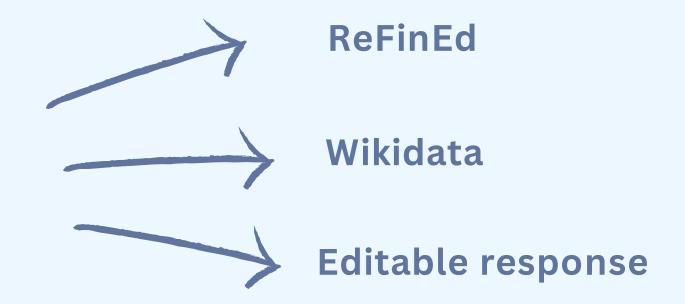
The features of a Digital Edition

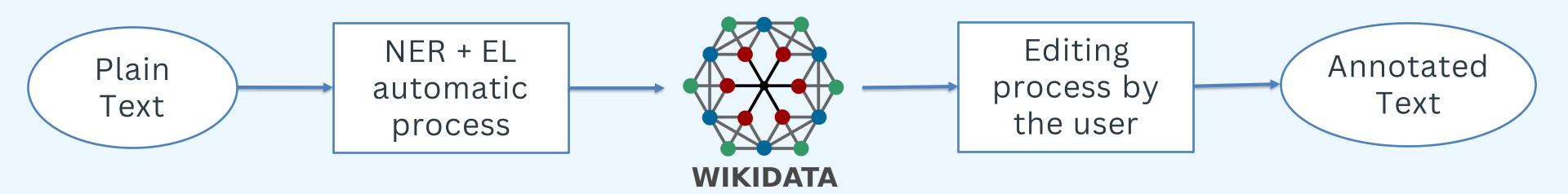


Which of these features can benefit from automation?

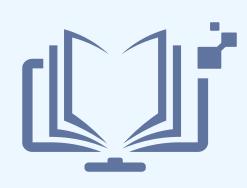
Text Annotation

- Named entity recognition and linking systems
- → Knowledge Base for retrieving information
- Control over the annotation





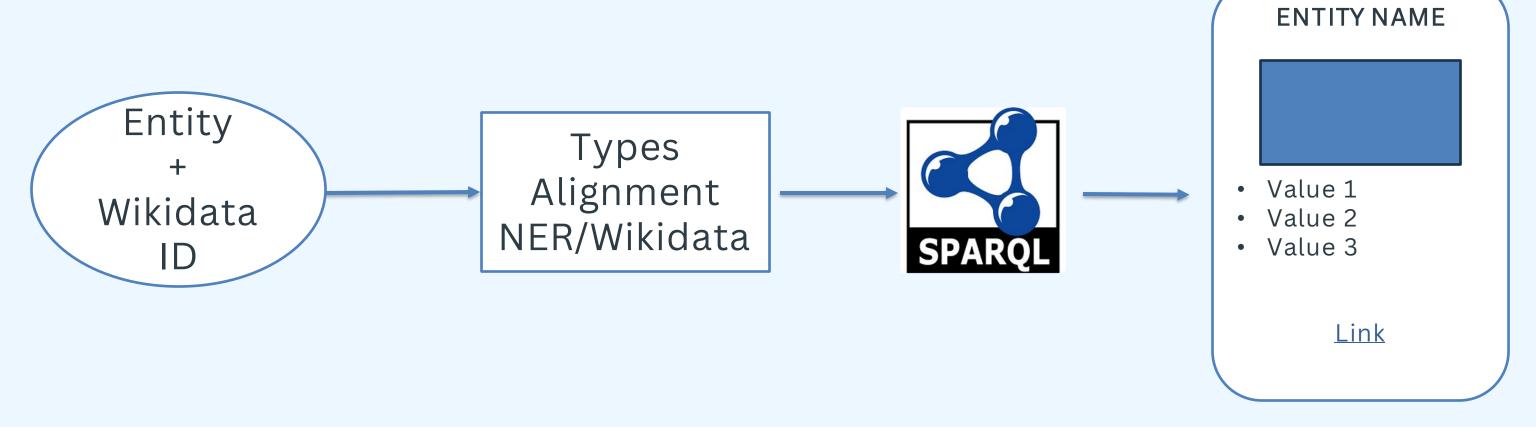
Semi-automatic process



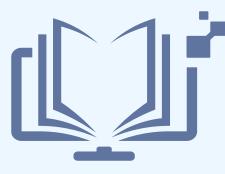
Data enrichment & Visualisation

- → Execution of SPARQL queries
- Visualisation of the data
- Standardisation of the items information



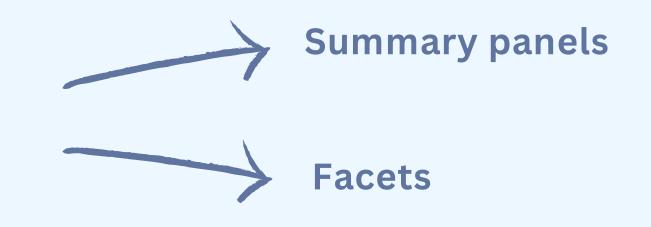


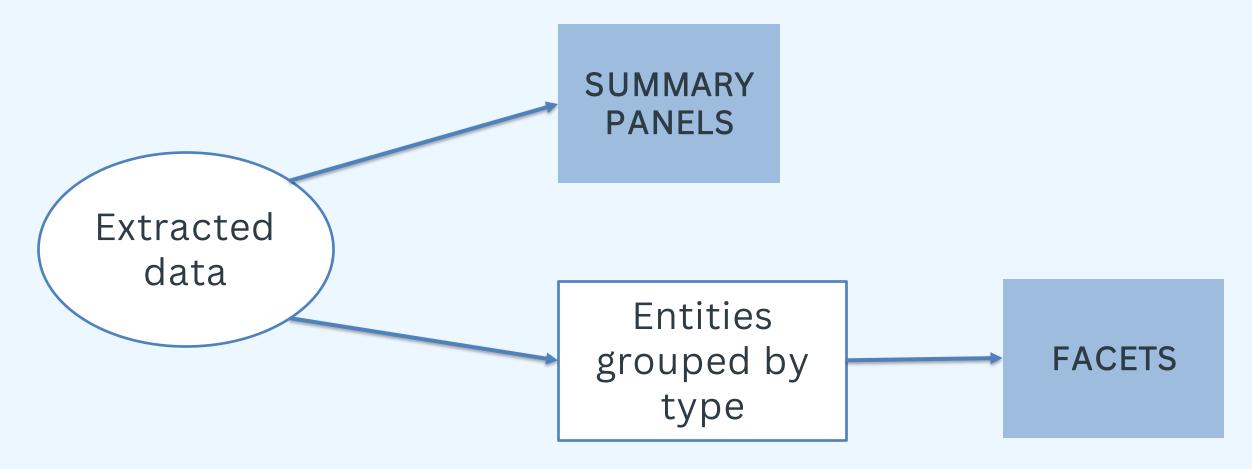
Semi-automatic process



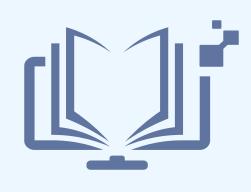
Search and analysis tools

- A way to organise the retrieved entities
- → A way to search the annotated entities



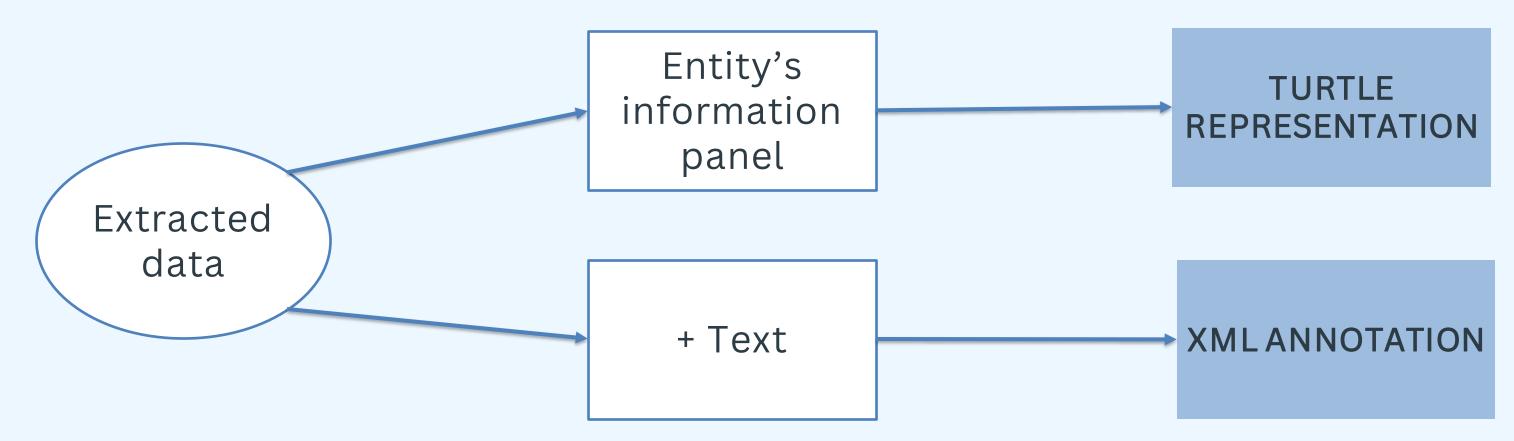


Automatic process

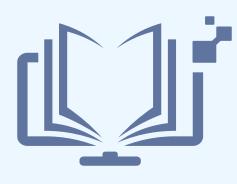


Interoperability and shareability

→ Sharing data and information
Export of the text in TEI-XML



Automatic process



Development

Server - Python APIs

- Two endpoints
- Annotation process
- → SPARQL queries templates

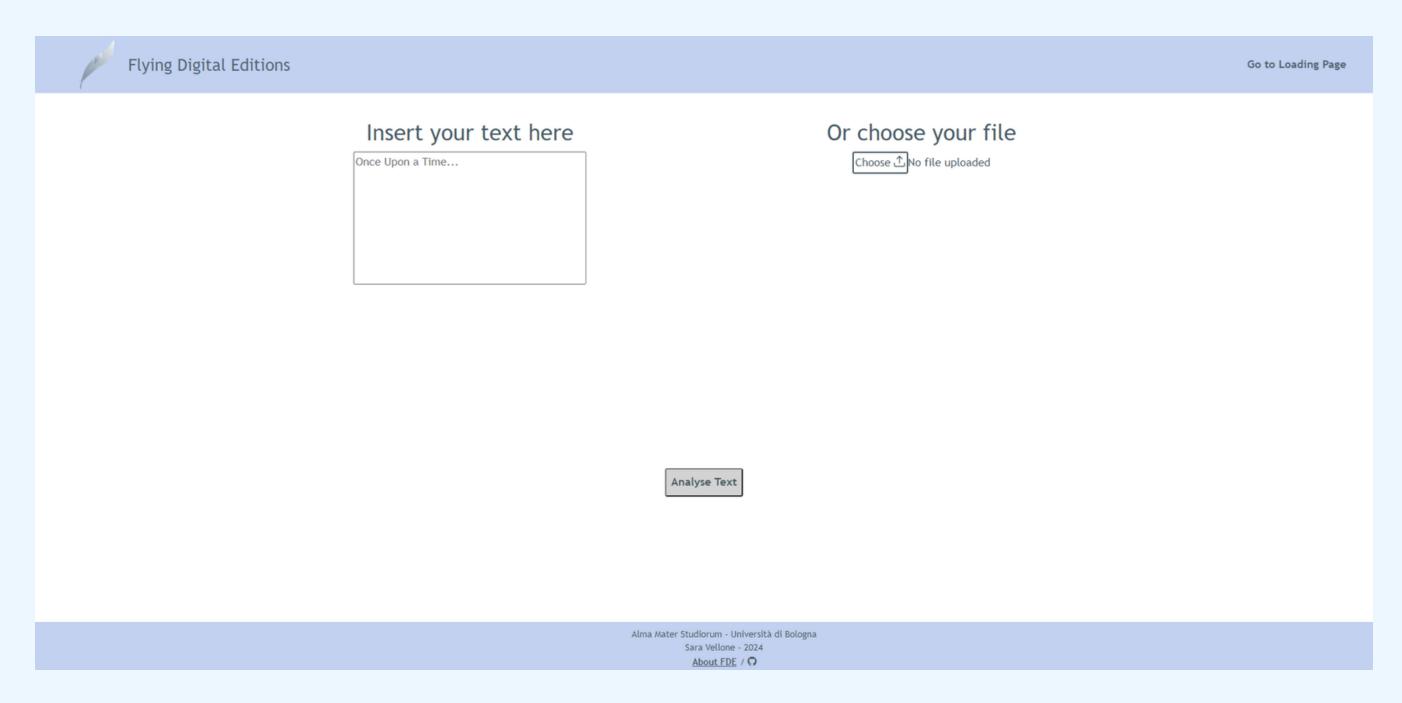
Client - Angular Framework

- Loading-mode page
- → Edit-mode page
- → View-mode page
- → Information panel



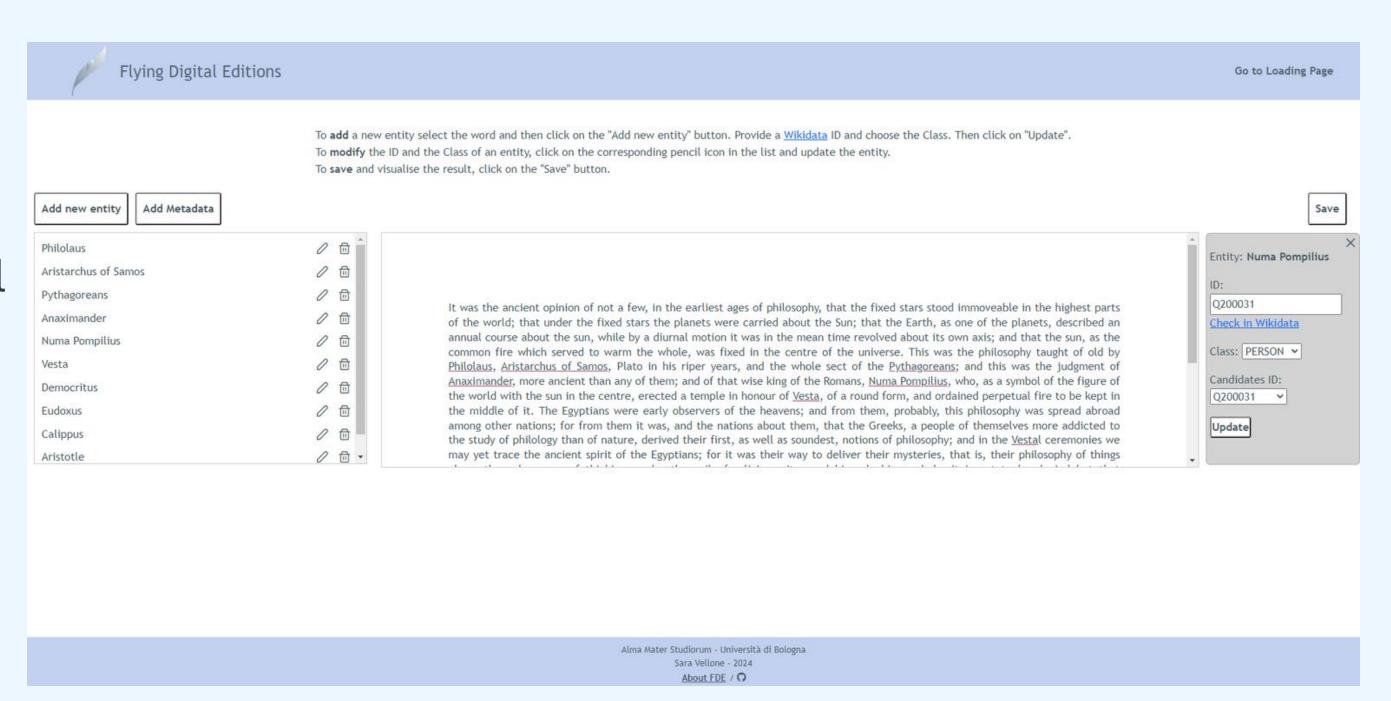
Development Loading-Mode

- Few familiarelements
- → Multi-input



Development Edit-Mode

- Guidelines
- Summary panel
- → Text-centered
- Editing card



Development View-Mode

- Searching facet
- Visual highlighting
- → Information panel
- Export available

Flying Digital Editions

Go to Loading Page

Person Other

- Philolaus
- Aristarchus of Samos
- Anaximander
- Numa Pompilius

Back to Edit-mode

- Vesta
- ⊕ Democritus
- ⊕ Eudoxus
- Calippus
- ⊕ Aristotle

Principia

Isaac Newton

It was the ancient opinion of not a few, in the earliest ages of philosophy, that the fixed stars stood immoveable in the highest parts of the world; that under the fixed stars the planets were carried about the Sun; that the Earth, as one of the planets, described an annual course about the sun, while by a diurnal motion it was in the mean time revolved about its own axis; and that the sun, as the common fire which served to warm the whole, was fixed in the centre of the universe. This was the philosophy taught of old by Philolaus, Aristarchus of Samos, Plato in his riper years, and the whole sect of the Pythagoreans; and this was the judgment of Anaximander, more ancient than any of them; and of that wise king of the Romans, Numa Pompilius, who, as a symbol of the figure of the world with the sun in the centre, erected a temple in honour of Vesta, of a round form, and ordained perpetual fire to be kept in the middle of it. The Egyptians were early observers of the heavens; and from them, probably, this philosophy was spread abroad among other nations; for from them it was, and the nations about them, that the Greeks, a people of themselves more addicted to the study of philology than of nature, derived their first, as well as soundest, notions of philosophy; and in the vestal ceremonies we may yet trace the ancient spirit of the

Numa Pompilius Legendary second king of rome, Description succeeding romulus Birth Date 13 Apr 752 CE Birth Place Death Date 1 Jan 670 CE Death Place Gender Ancient roman politician, ancient Occupation

Export Text in XML

Alma Mater Studiorum - Università di Bologna Sara Vellone - 2024

About FDE / O

Results - General summary

Pros

- Leveraging Wikidata
- Editable annotation response
- SPARQL templates to fit types
- Coherence for the visualised data
- Turtle format available for download
- Text's XML format available for download
- An easy interface for non-expert users
- → Flexibility of the code

Cons

- → Entity detection and entity linking issues
- Generic categorisation of the entities
- → Items not mapped in Wikidata
- Weight of the models



Conclusions

Which of the main features of a Digital Edition can be turned into automatic processes?

How and at what cost is it feasible to do so?

It's not an impossible task, but...



- → Generic annotation
- → Human effort is still needed



Future Works

- → Features that are common in SDE, but have not been implemented.
- → Features that are implementable using available models.
- → Other features, such as authentication system and dynamically generated SPARQL queries.



