

ELTeC in TextGrid



ELTeC

The European Literary Text Collection (ELTeC) is a state-of-the-art, open access multilingual collection of corpora of European literary texts (novels, 1840-1920) for the development of tools and methods in Computational Literary Studies. Currently, ELTeC contains more than 2000 full-text novels in XML-TEI in 21 languages. They are distributed via multiple platforms: GitHub, Zenodo and by now already 1365 full-texts in 15 languages in TextGrid Repository.

Advantages for ELTeC in TextGrid

TextGrid enables transformation of XML-TEI to HTML (for traditional reading) or plain text (for analysis). ELTeC's findability is enhanced by persistent identifiers, authority file identifiers for authors, and subject classification. Texts in TextGrid Repository are findable through DARIAH Community Gateway in OpenAIRE (and the EOSC Marketplace) or the CLARIN Virtual Language Observatory and Collection Registry, accessible through the repository's APIs, and become more interoperable and reusable through connected tools such as the shelf function, Voyant Tools or to the CLARIN Switchboard.

> 2K

Texts: XML-TEI
Genre: Novels
Period: 1840-1920

21

European Languages



<https://textgridrep.de/project/TGPR-99d098e9-b60f-98fd-cda3-6448e07e619d>



Authors: Nanette Rißler-Pipka¹, José Calvo Tello², Stefan E. Funk², Carolin Odebrecht⁴, Christof Schöch³, Ubbo Veenijer²
¹Max Weber Foundation, Bonn; ²Göttingen State and University Library; ³University of Trier; ⁴Humboldt-Universität zu Berlin