



UNIVERSITÀ
CATTOLICA
del Sacro Cuore



LiLa: Linking Latin

Building a Knowledge Base
of Interlinked Linguistic Resources for Latin

The LiLa team

Conference *Linked Pasts 6*
University of London and British Library
December 2-16, 2020



This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme - Grant Agreement No. 769994.

Research question

State of affairs



We have built and collected (for Latin and other languages):

We have built and collected (for Latin and other languages):

- ▶ Textual Resources

We have built and collected (for Latin and other languages):

- ▶ Textual Resources
- ▶ Lexical Resources

We have built and collected (for Latin and other languages):

- ▶ Textual Resources
- ▶ Lexical Resources
- ▶ NLP Tools

We have built and collected (for Latin and other languages):

- ▶ Textual Resources
- ▶ Lexical Resources
- ▶ NLP Tools

Scattered and unconnected

ERC Consolidator Grant 2018-2023

A collection of multifarious, interoperable linguistic resources described with the same vocabulary for knowledge description (by using common data categories and ontologies)

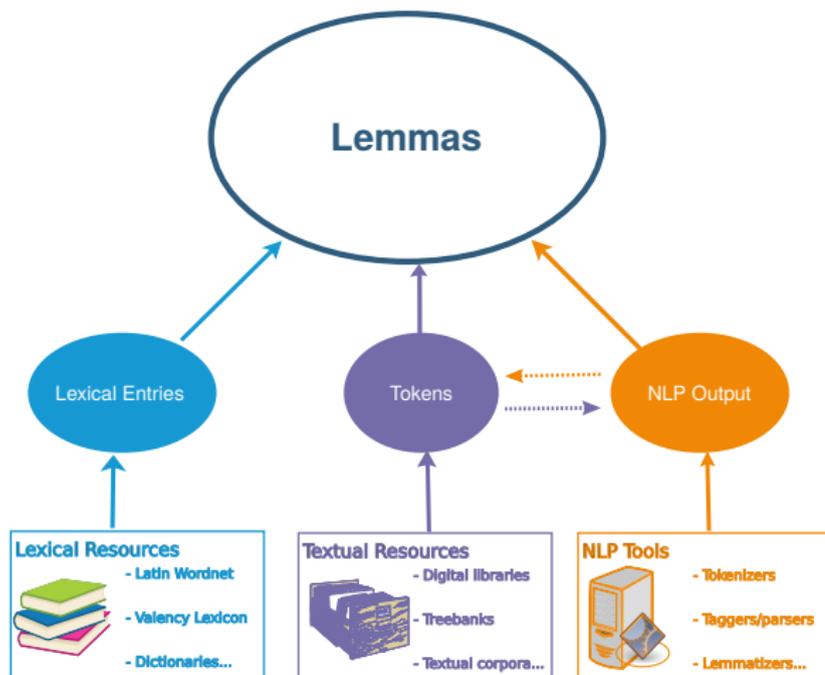
Interlinking as a Form of Interaction



Infrastructure



Interoperability



Lemma *admiror* 'to admire, to respect'

<https://lila-erc.eu/data/id/lemma/87541>

► Corpora

- Index Thomisticus Treebank (*Summa contra Gentiles*): ca. 450,000 nodes
- Dante Search (700th death anniversary coming up!): ca. 46,000 tokens
- Querolus sive Aulularia*: ca. 17,000 tokens
- PROIEL and LLCT treebanks
- Computational Historical Semantics Corpus

► Lexica

- Word Formation Latin: ca. 46,000 lemmas (Classical Latin)
- BRILL Etymological dictionary of Latin and the other Italic Languages: ca. 1,400 entries
- LatinAffectus (sentiment lexicon for Latin): ca. 2,300 entries
- Index Graecorum Vocabulorum in Linguam Latinam: ca. 1,800 entries
- Latin WordNet & Latin Vallex 2.0: ca. 1,000 manually checked entries
- Lewis & Short Dictionary

► NLP tools

- LEMLAT (lemma bank): ca. 150,000 lemmas

► TOTAL: approximately 10 million triples

The LiLa Activity at Linked Pasts 6

Friday 11th December. 9:00-16:00 (GMT) = 10:00-17:00 (CET)



The LiLa Activity at Linked Pasts 6

Friday 11th December. 9:00-16:00 (GMT) = 10:00-17:00 (CET)



- ▶ **WHAT:** to show the workflow we employ to connect a linguistic resource for Latin to the LiLa Knowledge Base, and to demonstrate the way in which LiLa can be queried, with practical examples of real research use-cases

The LiLa Activity at Linked Pasts 6

Friday 11th December. 9:00-16:00 (GMT) = 10:00-17:00 (CET)



- ▶ **WHAT:** to show the workflow we employ to connect a linguistic resource for Latin to the LiLa Knowledge Base, and to demonstrate the way in which LiLa can be queried, with practical examples of real research use-cases
- ▶ **HOW:** to provide participants with a Latin raw text and to teach them how to perform automatic lemmatization and RDF-ization in order to link it to LiLa

The LiLa Activity at Linked Pasts 6

Friday 11th December. 9:00-16:00 (GMT) = 10:00-17:00 (CET)



- ▶ **WHAT:** to show the workflow we employ to connect a linguistic resource for Latin to the LiLa Knowledge Base, and to demonstrate the way in which LiLa can be queried, with practical examples of real research use-cases
- ▶ **HOW:** to provide participants with a Latin raw text and to teach them how to perform automatic lemmatization and RDF-ization in order to link it to LiLa
- ▶ **WHO-1:** everyone who wishes to publish Latin texts on the web, using their own repositories or tools (like Recogito)

The LiLa Activity at Linked Pasts 6

Friday 11th December. 9:00-16:00 (GMT) = 10:00-17:00 (CET)



- ▶ **WHAT:** to show the workflow we employ to connect a linguistic resource for Latin to the LiLa Knowledge Base, and to demonstrate the way in which LiLa can be queried, with practical examples of real research use-cases
- ▶ **HOW:** to provide participants with a Latin raw text and to teach them how to perform automatic lemmatization and RDF-ization in order to link it to LiLa
- ▶ **WHO-1:** everyone who wishes to publish Latin texts on the web, using their own repositories or tools (like Recogito)
- ▶ **WHO-2:** everyone interested in the different aspects involved in the construction of a LLOD knowledge base

The LiLa Activity at Linked Pasts 6

Friday 11th December. 9:00-16:00 (GMT) = 10:00-17:00 (CET)



- ▶ **WHAT:** to show the workflow we employ to connect a linguistic resource for Latin to the LiLa Knowledge Base, and to demonstrate the way in which LiLa can be queried, with practical examples of real research use-cases
- ▶ **HOW:** to provide participants with a Latin raw text and to teach them how to perform automatic lemmatization and RDF-ization in order to link it to LiLa
- ▶ **WHO-1:** everyone who wishes to publish Latin texts on the web, using their own repositories or tools (like Recogito)
- ▶ **WHO-2:** everyone interested in the different aspects involved in the construction of a LLOD knowledge base
- ▶ **INFO & REGISTRATION** (deadline: 9th December):
<https://lila-erc.eu/linked-pasts-6-activity/>

Thanks!

Get in touch



LiLa: Linking Latin

Università Cattolica del Sacro Cuore
CIRCSE Research Centre



info@lila-erc.eu



<https://github.com/CIRCSE>



<https://lila-erc.eu>



@ERC_LiLa



Largo Gemelli 1, 20123 Milan, Italy



This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme - Grant Agreement No. 769994.