Life after ERC

What's New and What's Next in the LiLa Knowledge Base

Marco Passarotti

SwissText 2025
Winterthur, May 14-15, 2025

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.

Overview

Fundamentals of LOD and Introduction to LiLa

- Mission and Architecture

What's New

- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next

- Discussion and Conclusion

Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

Research question
State of affairs

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

Research question
State of affairs

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

- Textual Resources

Research question
State of affairs

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

- Textual Resources
- Lexical Resources

Research question
State of affairs

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

- Textual Resources
- Lexical Resources
- NLP Tools

Research question
State of affairs

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

The Linked Data Principles

…just to be FAIR

The Linked Data Principles
…just to be FAIR

- Use URIs for things (e.g. an entry in a lexicon, a token in a corpus)

The Linked Data Principles
…just to be FAIR

- Use URIs for things (e.g. an entry in a lexicon, a token in a corpus)
- Use HTTP URIs to allow people (and machines) to look up things

The Linked Data Principles

…just to be FAIR

- Use URIs for things (e.g. an entry in a lexicon, a token in a corpus)
- Use HTTP URIs to allow people (and machines) to look up things
- Use web standards to represent/query (meta)data, such as RDF and SPARQL

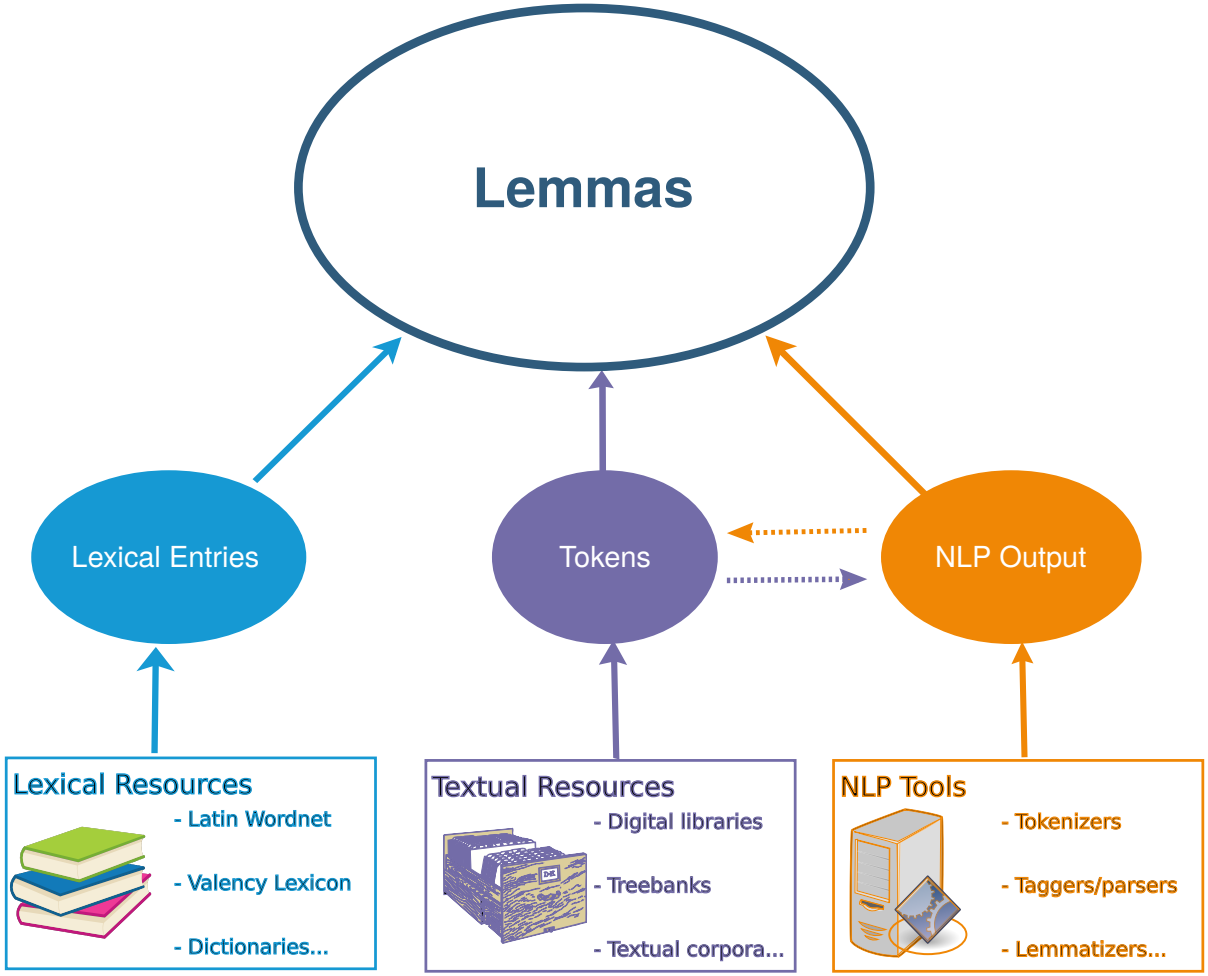
The Linked Data Principles

…just to be FAIR

- Use URIs for things (e.g. an entry in a lexicon, a token in a corpus)
- Use HTTP URIs to allow people (and machines) to look up things
- Use web standards to represent/query (meta)data, such as RDF and SPARQL
- Include links to other URIs

LiLa Knowledge Base

Lexically-based architecture and (meta)data sources



LiLa and Ontolex Lemon

A *de facto* W3C standard for publishing lexical data as LLOD

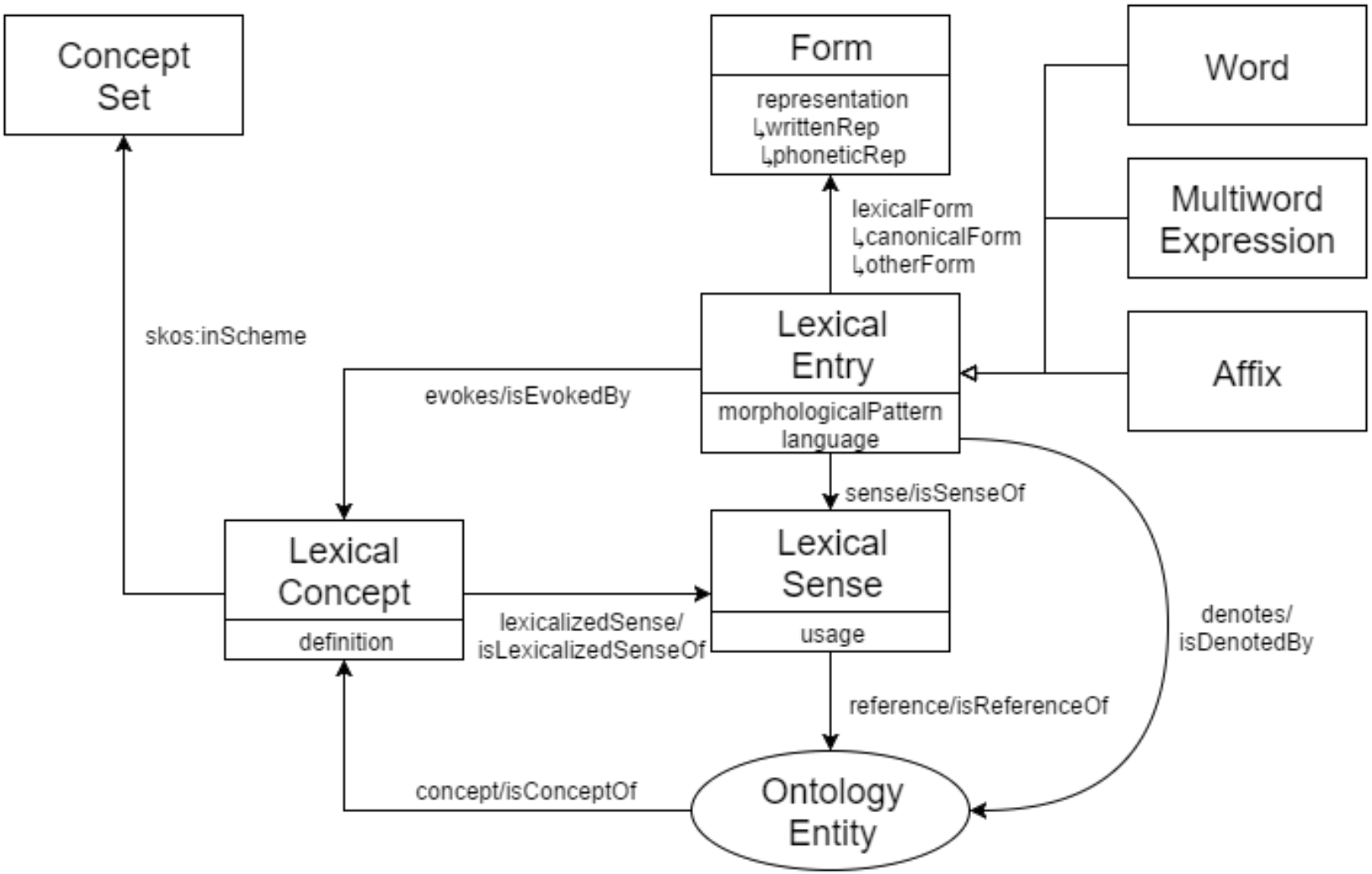


Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

Language Resources Interlinked in LiLa

...and some numbers

Resources connected so far

`https://lila-erc.eu/data-page/`

- **Lemmas** in the Lemma Bank: approx. 170K (Written Reps: approx. 270K)
- **Lexical Entries** in Lexical Resources: approx. 170K
- **Tokens** in Textual Resources: approx. 12M
- **Triples**: approx. 140M

Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

Lemma Bank Query Interface

https://lila-erc.eu/query/

Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

LiLa Search Platform

https://lila-erc.eu/LiLaLisp/

Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

SPARQL Access Point

`https://lila-erc.eu/sparql/`

Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

TextLinker

To process raw textual data and link tokens to the Lemma Bank

TextLinker

https://lila-erc.eu/LiLaTextLinker/

Table of Contents

Fundamentals of LOD and Introduction to LiLa
- Mission and Architecture

What's New
- Language Resources
- Lemma Bank Query Interface
- LiLa Search Platform
- SPARQL Access Point
- TextLinker

What's Next
- Discussion and Conclusion

What's Next
Discussion and Conclusion

- Automatic disambiguation of linking for textual resources

What's Next
Discussion and Conclusion

- Automatic disambiguation of linking for textual resources
- Meta-linking of LOD resources: images, geodata and prosopography for the ancient world

What's Next
Discussion and Conclusion

- Automatic disambiguation of linking for textual resources
- Meta-linking of LOD resources: images, geodata and prosopography for the ancient world
- Multilingual interoperability and Lemma Banks for other languages (LiITA, etymological resources)

What's Next
Discussion and Conclusion

- Automatic disambiguation of linking for textual resources
- Meta-linking of LOD resources: images, geodata and prosopography for the ancient world
- Multilingual interoperability and Lemma Banks for other languages (LiITA, etymological resources)
- Integrating NLP Tools

What's Next
Discussion and Conclusion

- Automatic disambiguation of linking for textual resources
- Meta-linking of LOD resources: images, geodata and prosopography for the ancient world
- Multilingual interoperability and Lemma Banks for other languages (LiITA, etymological resources)
- Integrating NLP Tools
- Exploiting explicit information stored in knowledge graphs to fine-tune LLMs

Thank you

Let's interlink!

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.