



UNIVERSITÀ  
CATTOLICA  
del Sacro Cuore



# Towards the Modeling of Polarity in a Latin Knowledge Base

R. Sprugnoli, F. Mambrini, G. Moretti, M. Passarotti

rachele.sprugnoli@unicatt.it

francesco.mambrini@unicatt.it

giovanni.moretti@unicatt.it

marco.passarotti@unicatt.it

WHiSe 2020 | June 2, 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.

**Inclusion** of a prior polarity lexicon of Latin lemmas in a Knowledge Base (KB) of interoperable linguistic resources using Semantic Web and Linked Data standards and practices:

- ▶ *LiLa: Linking Latin* project
- ▶ *Lemma Bank*
- ▶ *LatinAffectus*

**Inclusion** of a prior polarity lexicon of Latin lemmas in a Knowledge Base (KB) of interoperable linguistic resources using Semantic Web and Linked Data standards and practices:

- ▶ *LiLa: Linking Latin* project
- ▶ *Lemma Bank*
- ▶ *LatinAffectus*

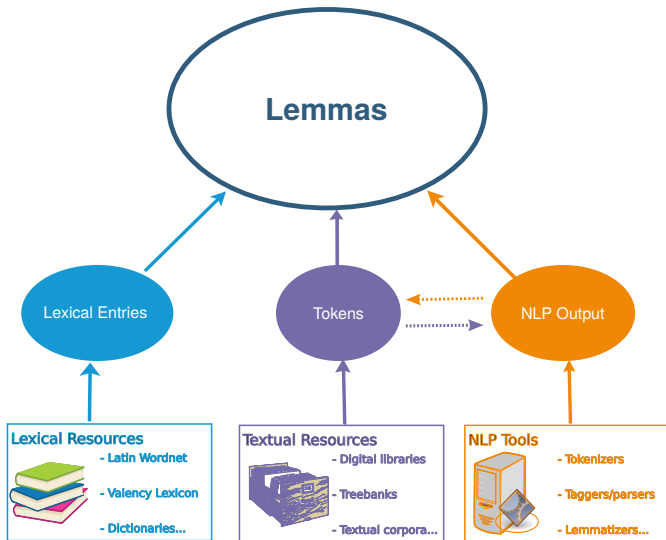
Resources are **available online**:

- ▶ Query Interface
- ▶ SPARQL Endpoint
- ▶ GitHub

- ▶ To **enrich** the information on (a subset of) lemmas of the LiLa KB
  
- ▶ To **fill a void** in the literature:
  1. NRC lexicons (Mohammad, 2018)
  2. MultilingualSentiment lexicons (Chen & Skiena, 2014)
  3. API of Latin WordNet (<https://latinwordnet.exeter.ac.uk/api>)
  
- ▶ To **enhance the interoperability** taking advantage of past work on formal representations of linguistic resources and services for sentiment analysis

- ▶ Using **empirical data** from Latin corpora and dictionaries
- ▶ Involving Latin **experts**
- ▶ Relying on a **Knowledge Base** of linguistic resources for Latin
  - ▶ corpora
  - ▶ lexicons
  - ▶ NLP tools
- ▶ Using **common vocabularies** to describe them

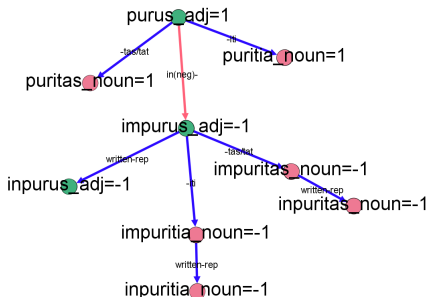




List of nouns and adjectives associated to their **out-of-context sentiment score**: from -1 (very negative) to +1 (very positive)

- ▶ **Gold Standard**: manually created by 2 Latin language and culture experts + 1 supervisor
- ▶ **Silver Standard**: automatically created by deriving new entries from the Gold Standard

- ▶ **New lemmas** added exploiting derivational, synonym and antonym relations + graphical variants of lemmas present in the Gold Standard
- ▶ Original polarity scores are **propagated** or **reversed** onto the newly derived lemmas





	PoS		Polarity			TOT
	ADJ	NOUN	POS	NEUT	NEG	
Gold Standard	454	690	231	612	301	<b>1,144</b>
Silver Standard	512	781	271	689	333	<b>1,293</b>
<b>LatinAffectus</b>	966	1,471	502	1301	634	<b>2,437</b>

Sprugnoli, R., Passarotti, M., Corbetta, D., Peverelli, A. (2020). *Odi et Amo. Creating, Evaluating and Extending Sentiment Lexicons for Latin*. In Proceedings of LREC 2020.

We re-use **formal representation frameworks** to describe the lexical resource and the sentiment properties of each entry:

- ▶ lemon
- ▶ lime
- ▶ Ontolex
- ▶ Marl

<https://lila-erc.eu/data/lexicalResources/LatinAffectus/Lexicon>

## LatinAffectus

<http://lila-erc.eu/data/lexicalResources/LatinAffectus/Lexicon>

**dcterms:description** Gold: prior polarity lexicon of Latin lemmas created by two experts of Latin language and culture following a multi-stage process and an extensive reconciliation phase. It follows a five-way classification: 1 (fully positive), 0.5 (somewhat positive), 0 (neutral), -0.5 (somewhat negative), -1 (fully negative) Silver: prior polarity lexicon built by deriving new entries through synonym, antonym and derivational relations with the entries in the gold standard.

<b>rdfs:label</b>	LatinAffectus
<b>dcterms:title</b>	LatinAffectus - sentiment lexicon for latin.
<b>rdf:about</b>	<a href="https://github.com/CIRCSE/Latin_Sentiment_Lexicons">https://github.com/CIRCSE/Latin_Sentiment_Lexicons</a>
<b>rdf:type</b>	crm:E31 — lime:Lexicon
<b>lime:entry</b>	< <a href="http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_4858">http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_4858</a> > — < <a href="http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_6578">http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_6578</a> > — < <a href="http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_19360">http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_19360</a> >

## malus

[http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma\\_111418](http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_111418)

<b>rdfs:label</b>	malus
<b>rdf:type</b>	ontolex:LexicalEntry
<b>ontolex:canonicalForm</b>	< <a href="http://lila-erc.eu/data/id/lemma/111418">http://lila-erc.eu/data/id/lemma/111418</a> > ↳ malus
<b>ontolex:sense</b>	< <a href="http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalSense/lemma_111418">http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalSense/lemma_111418</a> > ↳ Prior sense of malus

### INVERSE RELATIONS

is lime:entry of	1 resource < <a href="http://lila-erc.eu/data/lexicalResources/LatinAffectus/Lexicon">http://lila-erc.eu/data/lexicalResources/LatinAffectus/Lexicon</a> > ↳ LatinAffectus
------------------	--

## Prior sense of malus

[http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalSense/lemma\\_111418](http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalSense/lemma_111418)

**rdfs:label**

Prior sense of malus

[<http://www.gsi.dit.upm.es/ontologies/marl/ns#polarityValue>](http://www.gsi.dit.upm.es/ontologies/marl/ns#polarityValue)

-1.0E0

**rdf:type**

ontolex:LexicalSense

[<http://www.gsi.dit.upm.es/ontologies/marl/ns#hasPolarity>](http://www.gsi.dit.upm.es/ontologies/marl/ns#hasPolarity)

[<http://www.gsi.dit.upm.es/ontologies/marl/ns#Negative>](http://www.gsi.dit.upm.es/ontologies/marl/ns#Negative)

### INVERSE RELATIONS

is ontolex:sense of

1 resource

[<http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma\\_111418>](http://lila-erc.eu/data/lexicalResources/LatinAffectus/id/LexicalEntry/lemma_111418)

↳ malus

**Semi-automatic linking** of the LatinAffectus lexical entries to their corresponding lemmas in the LiLa KB:

1. automatic matching of 2.084 lemmas (85.5%)
2. manual linking of:
  - ▶ 107 Medieval or New Latin lemmas: e.g. *praesuppositio* “assumption”
  - ▶ 246 ambiguous lemmas

## fidēs

*feminine noun V declension*

[View the declension of this word](#)

- 1 faith, loyalty
- 2 honesty
- 3 credit
- 4 confidence, trust, belief
- 5 good faith

## fidēs

*feminine noun III declension*

[View the declension of this word](#)

- 1 chord, instrument string
- 2 constellation Lyra
- 3 stringed instrument
- 4 lyre



How the interoperability between the resources connected in LiLa, and particularly LatinAffectus, can **support** research in the Humanities?



How the interoperability between the resources connected in LiLa, and particularly LatinAffectus, can **support** research in the Humanities?

### *Index Thomisticus* Trebank (ITTb):

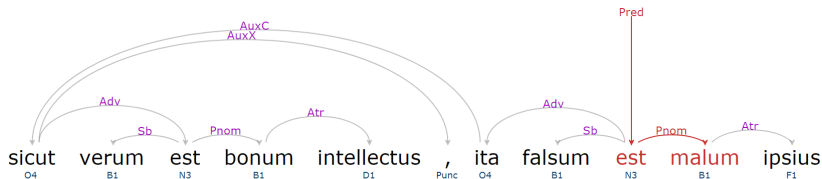
- ▶ first three books of the treatise *Summa contra Gentiles* of the philosopher Thomas Aquinas (XIII Century)
- ▶ 277,547 tokens
- ▶ morphological and syntactic annotation

## What is the nature of Evil?

What is the nature of Evil?



“as the true is the good of the intellect, so the false is its evil”

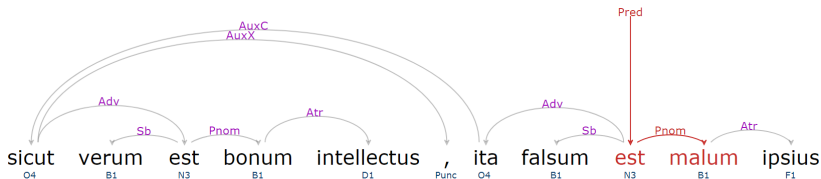


*Summa contra Gentiles, lib. 1, cap. 61, num. 8*

What is the nature of Evil?



“as the true is the good of the intellect, so the false is its evil”



*Summa contra Gentiles*, lib. 1, cap. 61, num. 8



What are the lemmas with a negative polarity that are the subject of the verb *sum* (“to be”) in a copular construction?

```
PREFIX ontollex: <http://www.w3.org/ns/lemon/ontollex#>
PREFIX lila: <http://lila-erc.eu/ontologies/lila/>
PREFIX powla: <http://acoli.cs.uni-frankfurt.de/resources/powla/owl/powla.owl#>
PREFIX lime: <http://www.w3.org/ns/lemon/lime#>
PREFIX marl: <http://www.gsi.dit.upm.es/ontologies/marl/ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX deprel: <http://lila-erc.eu/data/corpora/ITTB/id/synFunction/>

SELECT ?lablemma (count(?lemma) as ?tot) WHERE {
  <http://lila-erc.eu/data/lexicalResources/LatinAffectus/Lexicon> lime:entry ?lexentry .
  ?lexentry ontollex:canonicalForm ?lemma ;
  | ontollex:sense [ marl:hasPolarity marl:Negative ]

  SERVICE <http://localhost:3030/corpora/sparql> {
    ?t a powla:Terminal ;
    | lila:hasLemma ?lemma ;
    | powla:hasParent ?head ;
    | lila:hasDepRel deprel:Sb ;
    | powla:hasStringValue ?form .
    ?head lila:hasLemma <http://lila-erc.eu/data/id/lemma/126689> ;
    | powla:hasChild [ lila:hasDepRel deprel:Pnom ] .
  }

  SERVICE <http://localhost:3030/lemmaBank/sparql> {
    ?lemma rdfs:label ?lablemma
  }
}
group by ?lemma ?lablemma
order by desc(?tot)
```

The most frequent **negative subjects** of copular constructions in the ITTB

Lemma	English translation	Occurrences
<i>malum</i>	'evil'	37
<i>corruptio</i>	'destruction', 'decay', 'corruption'	7
<i>fornicatio</i>	'fornication'	4
<i>labor</i>	'labour', 'toil', 'exertion'	3

- ▶ *malum*: technical word for the concept of evil itself
- ▶ *corruptio*: destruction can be a good thing if it implies the destruction of evil
- ▶ *fornicatio*: always subject of a copular construction where *peccatum* ("sin") is the predicate nominal

### ► Coverage of *LatinAffectus*

- *privatio* “deprivation” (25 occurrences)
- *paupertas* “poverty” (9)
- *peccatum* “sin” (7)
- *defectus* “difect” (6)

### ► Lack of word sense disambiguation

## lăbŏr

*masculine noun III declension*

[View the declension of this word](#)

- 1 labour, toil, exertion, effort or work
- 2 task, undertaking
- 3 production
- 4 childbirth
- 5 preoccupation and concern
- 6 struggle, suffering, distress, hardship and stress

- ▶ No need of **pre-processing**
  
- ▶ **Interoperability** leads to a fruitful crossing between different resources to allow a broad range of corpus-based researches:
  - a dependency treebank
  - the *LatinAffectus* lexicon
  - the collection of lemmas of the LiLa KB



- ▶ To **extend** the coverage
- ▶ To **clean** *Latin WordNet*

lemma	synset_id	definition
capitolium	n#06188340	the federal government of the United States
voco	v#00720710	send a message or attempt to reach someone by radio, phone, etc; make a signal to in order to transmit a message [...]

- ▶ To **add** *Latin WordNet* to the LiLa KB -> to assign different polarities to different synsets

It's time to find a way to harmonize differences in data and metadata and achieve interoperability:

- ▶ to extract maximum benefit from the research investments
- ▶ to make resources and tools really impact and improve life of Classicists

**From unconnected and scattered information to knowledge**

## Università Cattolica del Sacro Cuore CIRCSE Research Centre



<https://lila-erc.eu>



[info@lila-erc.eu](mailto:info@lila-erc.eu) / [rachele.sprugnoli@unicatt.it](mailto:rachele.sprugnoli@unicatt.it)



[@ERC\\_LiLa](https://twitter.com/ERC_LiLa) / [@RSprugnoli](https://twitter.com/RSprugnoli)



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.