

# Construction d'un graphe de connaissances à partir des annotations manuelles de textes de zoologie antique

Arnaud Barbe, Molka Dhouib, Catherine Faron,  
Marco Corneli, Arnaud Zucker

# Introduction

BREPOLIS All Products EN FR DE IT

## Library of Latin Texts

Quick Search Advanced Search Table of Contents Distribution of Word-f

**Query**

Modify

Include 'similar'

**Centuries filter**

s. 4 4

**Export**

Export as PDF

Select All  
Select Page  
Deselect All

Include 'Background on the Text'

Include surrounding sentences

**Filter**

All results  
Lactantius

**Top results 1-2** (Results appearing in 4 sentences of 2 titles)

Divinae Institutiones - Lactantius - s. 4 p.C.  
De opificio Dei - Lactantius - s. 4 p.C.

**Results: Sentences (sententiae) 1-4 of 4**

1 **Lactantius - Divinae Institutiones (CPL 0085)**  
 lib. : 1, cap. : 22, par. : 11, pag. : 89, linea : 18 [\*]  
Idcirco illi mulieres **in aperto** sacrificant et Bonam Deam nomini

2 **Lactantius - Divinae Institutiones (CPL 0085)**  
 lib. : 6, cap. : 7, par. : 3, pag. : 505, linea : 22 [\*]  
deus enim, ut immortale illud arcanum eius **in aperto** esset, pro sapientia et veritate, quam sine ullo duce requirebant, **in** id ipsu

3 **Lactantius - De opificio Dei (CPL 0087)**  
 cap. : 12, par. : 1, pag. : 43, linea : 4 [\*]  
12. De utero quoque et conceptione, quoniam de internis loquimur, **aperto** latent, sensum tamen atque intellegentiam latere non pot

4 **Lactantius - De opificio Dei (CPL 0087)**  
 cap. : 14, par. : 6, pag. : 49, linea : 10 [\*]  
primum ista perspicere acumen humani sensus non potest, quia

**Tip:** You can often increase the number of results by includ

Google books

## Ancient Greek and Latin Texts

As part of its mission to make the world's books searchable and discoverable, Google has digitized over five hundred ancient Greek and Latin books. We present them here downloadable as zip files of images and plain text, and as links to Google Books web pages where you can read them online in full or download PDFs. This collection was selected by Prof. Greg Crane and Alison Babeu of Tufts University, and compiled by Will Brockman and Jon Orwant of Google. Enjoy!

[A Hebrew and English lexicon of the Old Testament](#)

by Wilhelm Gesenius, Edward Robinson - 1844 - 1144 pages.

(read in your browser)

**digilibLT** Biblioteca digitale di testi latini tardoantichi

**UPO** UNIVERSITÀ DEL PIEMONTE ORIENTALE

Fondazione Università del Piemonte Orientale

REGIONE PIEMONTE

Home | The project | News | Late antiquity on the web | Help | Contacts and feedback | Reserved area | DH Day 2021

User: ptingpoo [Logout](#)

**Search**

Searchable works  
Advanced search  
Search the bibliography

**Browse and download**

Works  
 Authors  
 Bibliography

**By date**

II III IV V VI VII VIII  
Uncertain date  
All the authors

**By name**

A B C D E F G H I L M N O  
P Q R S T U V  
All the authors

**Other resources**

Modern studies on late antiquity  
Canon of late-antique authors  
Fonts and software to

**Digital library of late-antique latin texts**

The Digital Library of Late-Antique Latin Texts—DigilibLT—publishes secular prose texts written in Latin in late antiquity (from the second to the seventh century AD). The texts are annotated according to the XML-TEI standards and are offered free of charge to the public for reading and research.

Since its creation in 2010, DigilibLT has achieved significant goals: near 400 texts have already been uploaded, a number that is constantly increasing. In addition, in recent years DigilibLT has added two significant extensions to the canon of authors and works initially designed for inclusion. The first concerns the grammatical works, uploaded to the site based on the best and most recent critical editions. Since 2019, DigilibLT has started another expansion line, this time regarding late Roman law. The first step has been the completion of a Canon of Roman Law, now available on the website alongside the other canon. Uploading of texts has recently started and actually it is in full swing.

The implementation of the library continues in compliance with the principles stated from the beginning: clarity in the indication of the editions adopted and any editorial interventions that may have been made; enrichment of the works with short entries on late-antique authors and works; bibliographical references. The query templates and the search options remain unchanged.

Libraries, digital or physical, couldn't exist without readers. For more than ten years, we have enjoyed the encouragement of a wide

# Objectifs

- Collaboration entre Wimmics et CEPAM
- Transmission des savoirs zoologiques
- Définition d'un modèle pour représenter un corpus d'annotations manuelles
  - Extraction des entités à partir des annotations
  - Liage avec des concepts de vocabulaire de domaine
  - Transformation automatique en RDF
  - Implémentation des questions de compétences

# Questions de recherche

1. Quels types de connaissances devons-nous représenter afin d'aider les chercheurs dans leur travail d'analyse de la transmission de savoirs zoologiques à travers les textes ?
2. Quels vocabulaires existants pouvons-nous réutiliser pour représenter ces documents ?
3. Quelle approche pouvons-nous définir pour réutiliser les annotations manuelles faites par les linguistes et les rendre exploitables automatiquement?

# Description du corpus de textes annotés

## Histoire naturelle chapitres 8-11

obedientia, officiorum quae didicere memoria, amoris et gloriae voluptas, immo vero, quae etiam in homine rara, probitas, prudentia, aequitas, religio quoque siderum solis que ac lunae veneratio.

[[2]] Auctores sunt in Mauretaniae saltibus ad quendam amnem, cui nomen est Amilo, nitescente luna nova greges

Paragraphe

### Annotations = commentaire word

<b>Auteur</b>	memory
<b>Auteur</b>	pleasure
<b>Auteur</b>	comparison with human
<b>Auteur</b>	sense of justice
<b>A</b> <b>Auteur</b>	religion : astral cult : moon cult , sun cult
	<a href="#">Répondre</a> <a href="#">Résoudre</a>
<b>Auteur</b>	places : Libya : region :
<b>Auteur</b>	habitat : terrestrial habitat :
<b>Auteur</b>	places : Libya : Mauretania :
<b>Auteur</b>	intellectual authority : undefined
<b>Auteur</b>	time : moon phase : new moon
<b>Auteur</b>	gregarious

# Motifs d'annotation manuelle

<b>Motifs d'annotation</b>	<b>Exemple</b>
"concept"	elefant
"concept1; concept2; ..."	old age ; social hierarchy
"concept1 : concept2 : ..."	habitat: terrestrial:forest
"concept1 : concept2, concept3, ..."	historical character :grammarian , Aristophanes of Byzance; anecdote
"collection : concept"	Places : Italia
Combinaisons de tous ces motifs	historical character ; intellectual authority : philosopher : Antipater

# Questions de compétences

- Quels sont les animaux qui construisent un habitat ?
- Quels sont les passages du texte qui permettent d'identifier les interactions possibles entre l'homme et l'animal (la prédation, l'amitié et l'animosité)?
- Quels sont les remèdes (thérapeutiques) dont un ingrédient est une partie d'animal ?

# Vocabulaires existants

schema.org

[[TITRE]] *Historia naturalis*, LIBER VIII

[[AUTEUR]] Pline

[[EDITION]] Pline : Plinius, Naturalis Historia, ed. C.

Mayhoff, 1892-1909

[[1]] Ad reliqua transeamus animalia et primum terrestria.

Maximum est elephans proximum que humanis sensibus, quippe intellectus illis sermonis patrii et imperiorum obedientia, officiorum quae didicere memoria, amoris et gloriae voluptas, immo vero, quae etiam in homine rara, probitas, prudentia, aequitas, religio quoque siderum solis que



## Web Annotation Vocabulary



TheZoo

 **Auteur**  
10:50 Aujourd'hui

size : relative

À partir d'un document importé.

 **Auteur**  
10:50 Aujourd'hui

elefant

# Thesaurus TheZoo

- Représenter et structurer hiérarchiquement tous les concepts de la zoologie antique et médiévale à partir de :
  - ressources textuelles
  - ressources iconographiques
  - ressources archéologiques
- 6019 concepts
- 11 niveaux hiérarchiques
- 14 collections

The screenshot displays the Thesaurus TheZoo interface. On the left, a 'Concepts hierarchy' panel shows a tree structure with 'black eagle' selected. On the right, a detailed view for 'black eagle (en)' is shown, including its label, collection (Zoonymy Ancient class), and translations in Spanish, French, and Greek. A 'Concepts Collection' panel is also visible, showing the concept's generic and specific relationships.

Concepts hierarchy

Concepts label

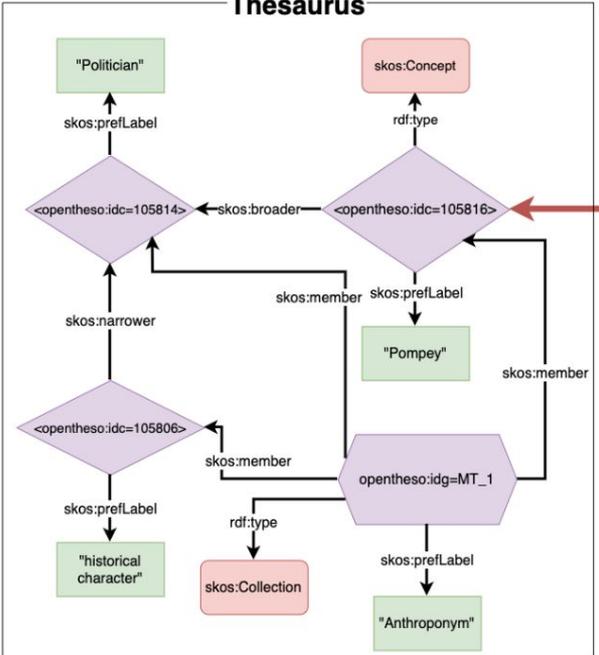
Concepts Collection

Label translations

# Modèle proposé

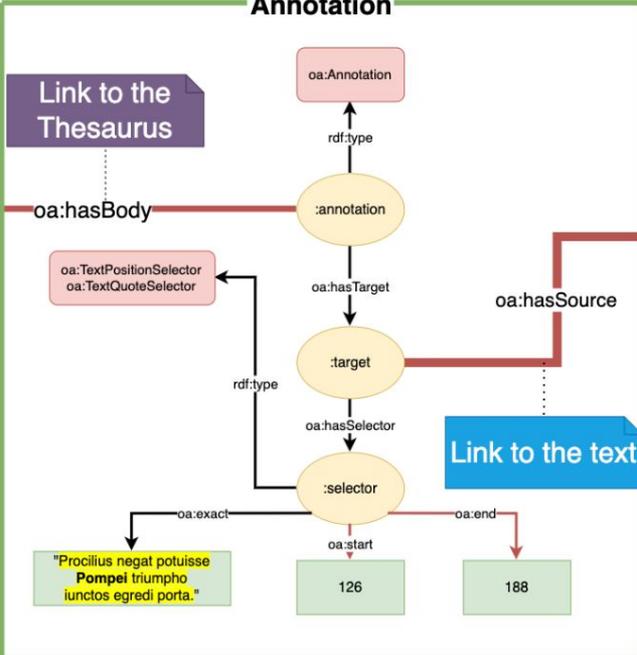
## Opentheso : TheZoo

### Thesaurus



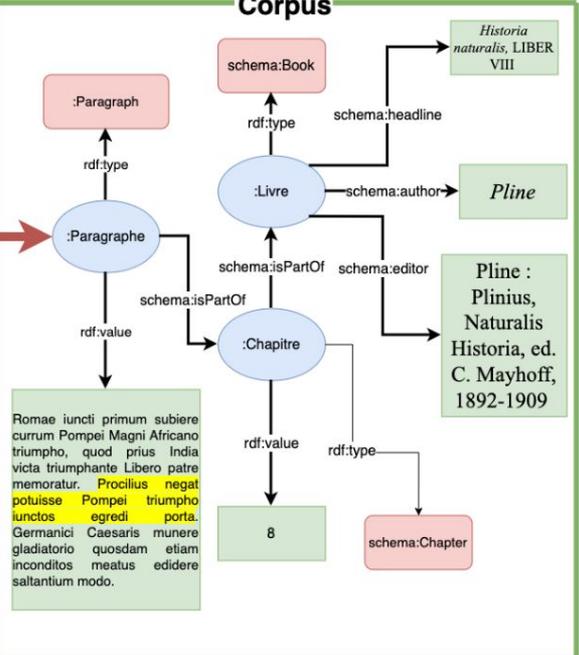
## OA: Annotation Ontology

### Annotation

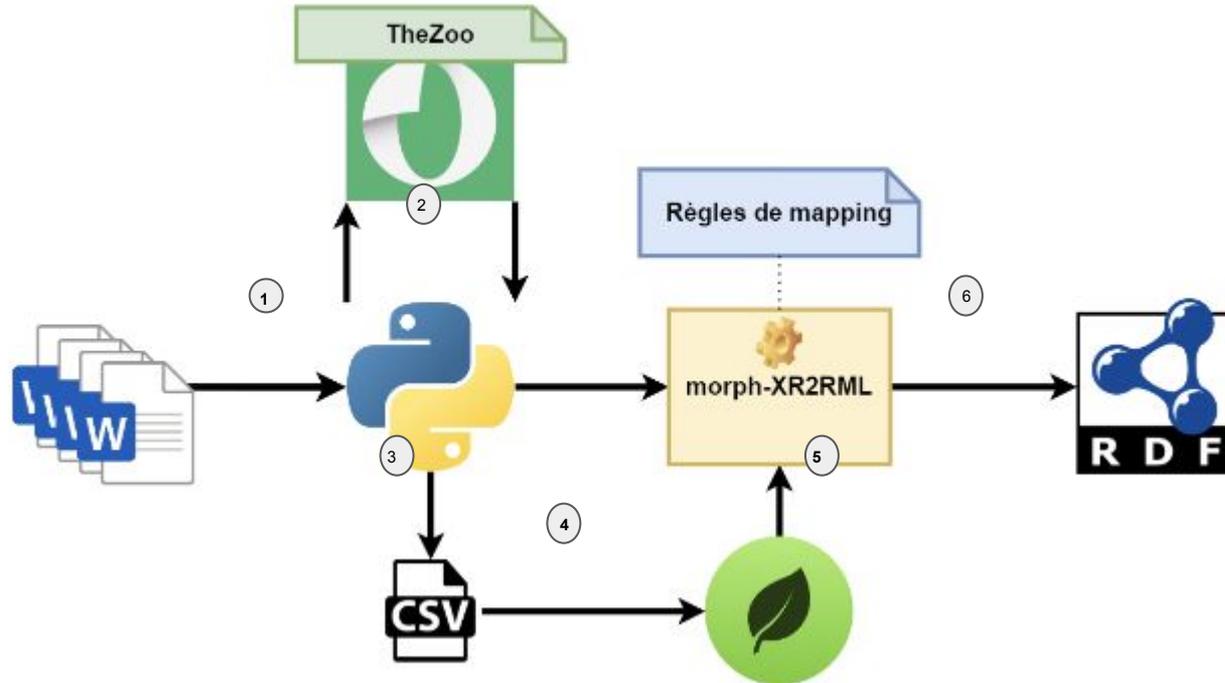


## book description

### Corpus



# Processus de construction du graphe de connaissance



# Description du graphe de connaissances produit

<b>Nbre de paragraphes</b>	911
<b>Nbre de commentaires</b>	7283
<b>Nbre d'entités reconnues</b>	13241
<b>Nbre d'entités liées</b>	11590
<b>Nbre d'entités non liées</b>	2632
<b>Nbre de triplets RDF générés</b>	88184

# Evaluation du graphe produit

## Evaluation du processus d'extraction de connaissances

- Précision = 1
- Rappel = 0.814

# Evaluation du graphe produit

## Implémentation en SPARQL des questions de compétence

Quels sont les animaux  
qui construisent un habitat?

```
PREFIX oa: <http://www.w3.org/ns/oa#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT DISTINCT ?paragraph ?name_animal ?name_construction
WHERE {?annotation1 a oa:Annotation; oa:hasBody ?animal;
      oa:hasTarget [oa:hasSource ?paragraph; oa:hasSelector [oa:exact?mention_animal]].
      ?annotation2 oa:hasBody ?construction; oa:hasTarget [oa:hasSource ?paragraph;
      oa:hasSelector [oa:exact ?mention_construction]].
      ?animal a skos:Concept; skos:prefLabel ?name_animal.
      <https://opentheso.huma-num.fr/?idg=MT_10&idt=th310> skos:member ?animal.
      ?construction skos:prefLabel ?name_construction;
      skos:broader+ <https://opentheso.huma-num.fr/?idc=105466&idt=th310>.
      FILTER (lang(?name_animal) = "en"). FILTER (lang(?name_construction) = "en")}
ORDER BY ?paragraph
```

Paragraph	Text	Animal Concept	Mention Animal	Construction concept	Mention construction
<a href="http://www.zoomathia.com/10-annoter/PL/10/">http://www.zoomathia.com/10-annoter/PL/10/</a>	Plurimum volant quae apodes, quia careant usu pedum, ab aliis cypseli appellantur, hirundinum specie, nidificant in scopulis. hae sunt quae toto mari cernuntur, nec umquam tam longo nave tamque continuo cursu recedunt a terra, ut non circumvolent eas apodes. cetera genera residunt et insistent, his quies nisi in nido nulla: aut pendunt aut iacent.	swallow1	hirundinum specie	nest	in nido

Lors d'un clic d'une ligne

Chapitre 10

[114] Plurimum volant quae apodes, quia careant usu pedum, ab aliis cypseli appellantur, **hirundinum specie**, nidificant in scopulis. hae sunt quae toto mari cernuntur, nec umquam tam longo nave tamque continuo cursu recedunt a terra, ut non circumvolent eas apodes. cetera genera residunt et insistent; his quies nisi **in nido**, nulla: aut pendunt aut iacent.

**swallow1**

**nest**

# Conclusion

- Construction d'un graphe de connaissances à partir des annotations manuelles de l'Histoire Naturelle de Plin l'Ancien
- Capture du contexte d'apparition des différentes annotations et structuration avec Open Annotation
- Liage des annotations manuelles avec TheZoo
- Interrogation du graphe en exploitant les contextes d'apparition et les liens entre les concepts du TheZoo
- Des problèmes de modélisation soulevés au niveau de TheZoo

# Perspectives

- Automatisation de la tâche d'annotation des textes
- Développement d'interfaces de visualisation plus intelligibles et intuitives pour les experts

**Merci pour votre attention!**

# Extraction et désambiguïsation d'entités nommées

## Requête SPARQL de recherche de concepts dans TheZoo à partir de labels

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?x WHERE {
  {?x a ?type; skos:prefLabel ?label.
   ?y skos:prefLabel ?collection; skos:member ?x.
   FILTER(lang(?label) = "en").
   FILTER("{label}" in (ucase(str(?label)),lcase(str(?label)), str(?label))).
   FILTER("{parent}" in (ucase(str(?collection)), lcase(str(?collection)), str(?collection))).
 } UNION {?x a ?type; skos:prefLabel ?label;skos:broader+ ?y.
   ?y skos:prefLabel ?concept;
   FILTER(lang(?label) = "en").
   FILTER("{label}" in (ucase(str(?label)),lcase(str(?label)), str(?label))).
   FILTER("{parent}" in (ucase(str(?concept)),lcase(str(?concept)), str(?concept))).
}}
```