Supporting Ancient Historical Linguistics with EpiLexO

Valeria Quochi, CNR-ILC
co-authors: A. Bellandi, M. Mallia, C. Zavattari (CNR-ILC), A. Tommasi (UniPi)
Context

• Italian National project ItAnt: «Languages and cultures of ancient Italy. Historical Linguistics and Digital Methods» (2020-24) in collaboration with UniVe (A. Marinetti, L. Rigobianco, P. Solinas), UniFi (F. Murano, M. Ballerini, M. Zinzi, C. Toscano)

• Resources, services and software will be deposited in, hosted and run by CLARIN-IT
Project ItAnt:

- Languages of ancient Italy before linguistic Romanisation: e.g. oscan, venetic, neo-faliscan, celtic.
  - documentation time span: 8th century BC – 1st century AD
- Scarce and scarce attestations ➔
  - small number of texts (perhaps 2k all together), all in epigraphic form
  - damaged supports


https://www.britishmuseum.org/collection/object/G_1867-0508-76
Interlinked resource ecosystem
The challenge

• Complement Digital Epigraphy with “explicit linguistic” knowledge for:
  - immaterial Cultural Heritage preservation,
  - wider open fruition of such knowledge,
  - support linguistic studies of languages of fragmentary attestation
  - preserve and make widely and openly accessible such linguistic knowledge
  - foster cross-disciplinary approaches

• First step:
  - a collaborative editing platform for creating lexicons and interlinking them with (textual) testimonies, bibliography, archeological metadata, other existing resources.
The EpiLexO system

LexO-server: services for lexica & ontologies + bibliographic references

CASH-server: services for file, texts, & annotations management
The architecture

LexO-server general configuration

ItAnt configuration

- GUI Client 1
- GUI Client n

HTTP

{json}

task oriented services groups

SERVICES

INDEXES

Lucene Connector

Lucene

Apache Lucene

GraphDB

MySQL

Semantic repository

Relational DB

MANAGER

SPARQL converter

Helper

External APIs

- Corpus Annotation and Search
- Bibliography REST services
- User management REST services

CASH-server

zotero

KEYCLOAK

REST services

User management

Zotero

Corpus Annotation and Search

Bibliography REST services

External APIs

LexO-server general configuration

ItAnt configuration
The LexO-server

LexO-server API documentation V0.1

[ Base URL: /LexO-backend-beta/service ]
https://licodemo.ilc.cnr.it/LexO-backend-beta/service/swagger.json

OntoLex-Lemon lexica manager
LexO-backend

Schemes

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Path</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>POST</td>
<td>/configuration/parameter</td>
<td>Configuration parameter</td>
</tr>
<tr>
<td>GET</td>
<td>/configuration/update</td>
<td>Configuration database schema update</td>
</tr>
</tbody>
</table>

Lexicon creation

<table>
<thead>
<tr>
<th>Path</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GET</td>
<td>/lexicon/creation/lexicalEntry</td>
</tr>
<tr>
<td>GET</td>
<td>/lexicon/creation/language</td>
</tr>
<tr>
<td>POST</td>
<td>/lexicon/creation/bibliography</td>
</tr>
<tr>
<td>GET</td>
<td>/lexicon/creation/etymology</td>
</tr>
<tr>
<td>GET</td>
<td>/lexicon/creation/lexicalSense</td>
</tr>
<tr>
<td>GET</td>
<td>/lexicon/creation/lexicalConcept</td>
</tr>
</tbody>
</table>
The CASH-server

OpenAPI definition

annotation-controller
- GET /api/v1/annotation
- POST /api/v1/uploadFile
- GET /api/v1/token
- GET /api/v1/gettext
- GET /api/v1/getcontent
- POST /api/v1/searchFiles
- POST /api/crud/uploadFile
- POST /api/crud/uploadMetadata
- POST /api/crud/updateMetadata
- POST /api/crud/revert
- GET /api/crud/query
- POST /api/crud/refresh
- POST /api/crud/moveFileTo

search-controller
- POST /api/testSearch
- POST /api/searchFiles

Example Value: Schema

```json
{
  "requestUUID": "string",
  "annotations": {
    "layer": "string",
    "values": "string",
    "imported": true,
    "attributes": {
      "additionalAnnotation": {},
      "additionalProperty": {}
    }
  },
  "span": {
    "start": 0,
    "end": 1
  }
}
```
Attestations: linking to text and bib
Linking to external resources

SEE ALSO
http://ltl-enc.eu/data/leicafeleou
Linking to external resources

http://via-sec.eu/data/
Other applications using LexO services

(Giovannetti et al., 2021)
tool to access a text on a linguistic and conceptual basis
demo available at https://klab.ilc.cnr.it/talmudSearch/

(Colombo et al., 2022)
experimental tool for allowing the navigation of the lexical senses of a lexicon
demo available at https://klab.ilc.cnr.it/lexiconGraph/

Slide by Andrea Bellandi (2022)
Future work

- cross-resource, multi-level search
- develop the exploration platform
- download and export, esp. in LLOD compatible formats
- extension to different use-cases
- Assessments and enhancements to make a full-fledged CLARIN-IT service
Thank you for your attention!

Further details:
or mailto:valeria.quochi@ilc.cnr.it
Short video demo: https://youtu.be/Isdy9piqy5E

To test the platform: https://lari2.ilc.cnr.it/LexO-angular/lexicon
(test, test22)