

PATRIMONIVM >> PETRAE EDITOR: EPIDOC ENCODING MADE EASIER

Epigraphy.info Workshop III – Vienna – 30.05.2019

Vincent RAZANAJAO

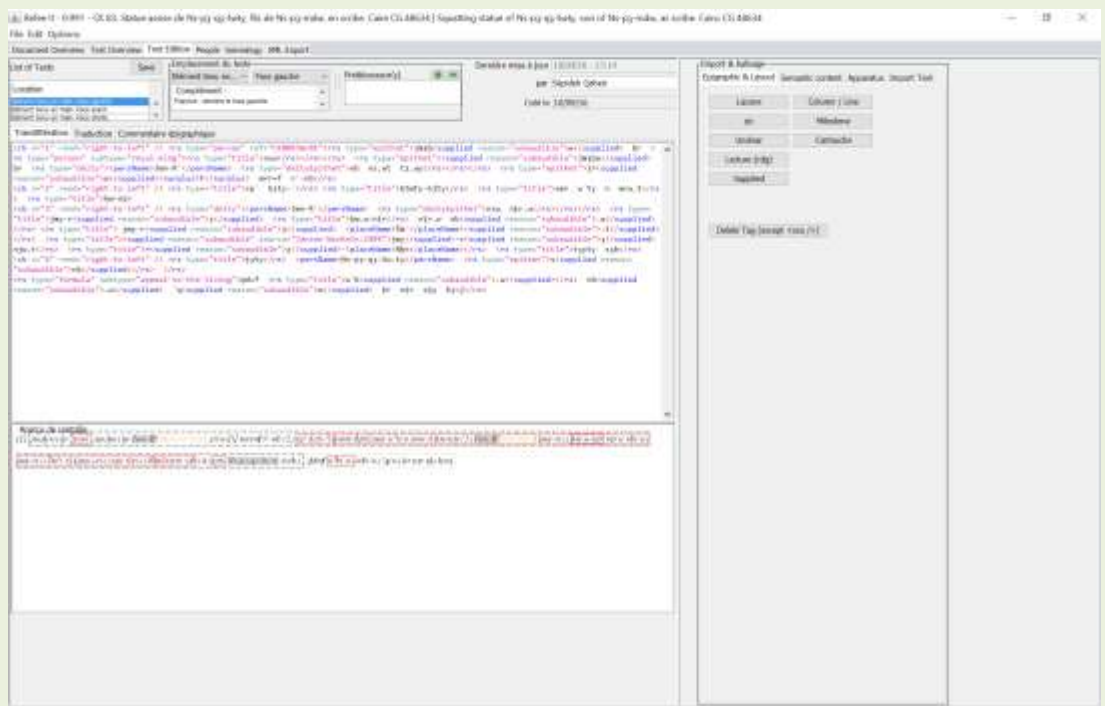
Université Bordeaux Montaigne, Institut Ausonius
PATRIMONIUM ERC-StG 716375

Some background

- Ancient Egyptian epigraphy
 - Digital corpus of the so-called ‘Ptolemaic Geographical texts
 - Karnak Cachette (<http://www.ifao.egnet.net/bases/cachette/>)

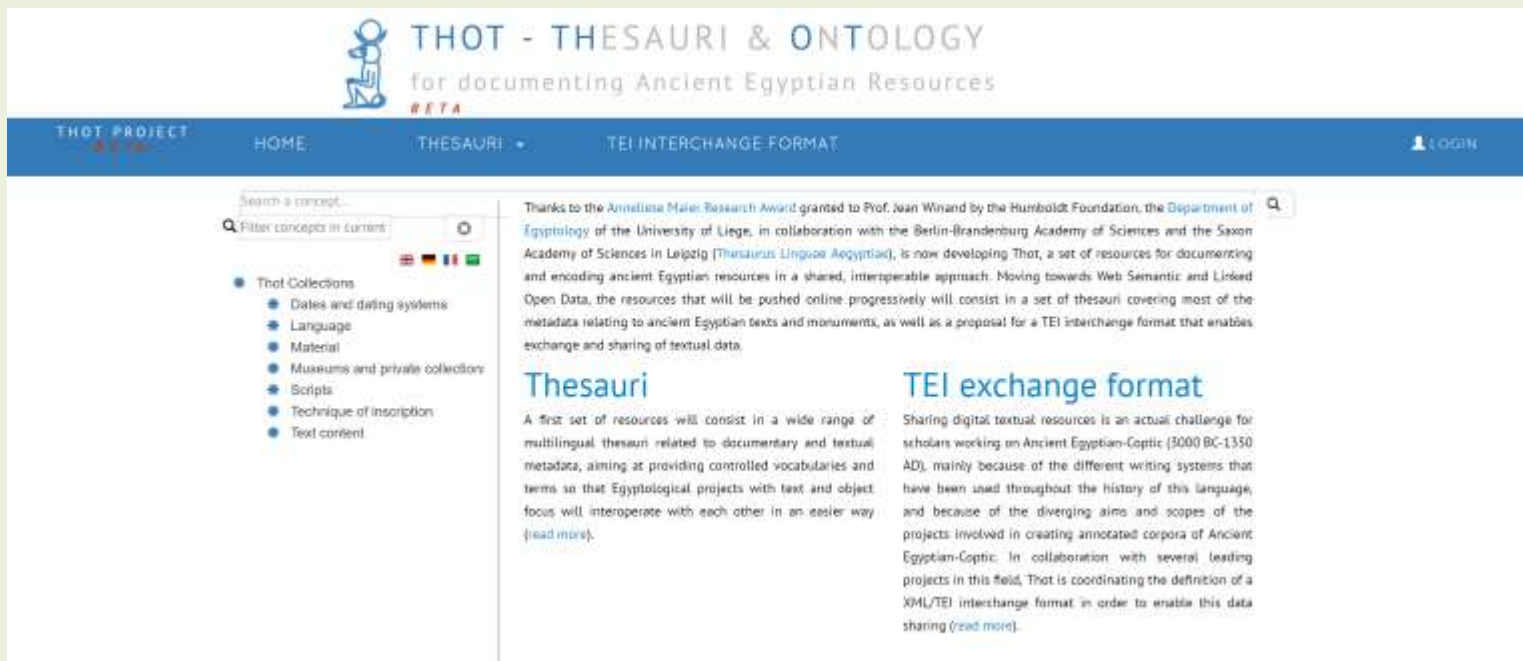


XEFEE – XML Editor for ancient Egyptian Texts
Java Desktop application



Some background

- Ancient Egyptian epigraphy
 - Digital corpus of the so-called 'Ptolemaic Geographical texts
 - Karnak Cachette (<http://www.ifao.egnet.net/bases/cachette/>)
 - XEFEE – XML Editor for ancient Egyptian Texts (Java application)
 - Thot Project: metadata for ancient Egyptian resources



The screenshot shows the THOT - THESAURI & ONTOLOGY website. The header includes the logo (a stick figure with a globe) and the text "THOT - THESAURI & ONTOLOGY for documenting Ancient Egyptian Resources BETA". The navigation bar has links for "THOT PROJECT", "HOME", "THESAURI", and "TEI INTERCHANGE FORMAT", along with a "LOGIN" button. The main content area features a search bar, a filter for concepts in current, and a list of "Thot Collections" including Dates and dating systems, Language, Material, Museums and private collections, Scripts, Technique of inscription, and Text content. The main text area contains a paragraph about the project's funding and goals, followed by two sections: "Thesauri" and "TEI exchange format", each with a brief description of their purpose.

THOT - THESAURI & ONTOLOGY
for documenting Ancient Egyptian Resources
BETA

THOT PROJECT
HOME THESAURI TEI INTERCHANGE FORMAT LOGIN

Search a concept...
Filter concepts in current

Thot Collections

- Dates and dating systems
- Language
- Material
- Museums and private collections
- Scripts
- Technique of inscription
- Text content

Thanks to the Anneliese Maier Research Award granted to Prof. Joan Winand by the Humboldt Foundation, the Department of Egyptology of the University of Liege, in collaboration with the Berlin-Brandenburg Academy of Sciences and the Saxon Academy of Sciences in Leipzig (Thesaurus Linguae Aegyptiacae), is now developing Thot, a set of resources for documenting and encoding ancient Egyptian resources in a shared, interoperable approach. Moving towards Web Semantic and Linked Open Data, the resources that will be pushed online progressively will consist in a set of thesauri covering most of the metadata relating to ancient Egyptian texts and monuments, as well as a proposal for a TEI interchange format that enables exchange and sharing of textual data.

Thesauri

A first set of resources will consist in a wide range of multilingual thesauri related to documentary and textual metadata, aiming at providing controlled vocabularies and terms so that Egyptological projects with text and object focus will interoperate with each other in an easier way (read more).

TEI exchange format

Sharing digital textual resources is an actual challenge for scholars working on Ancient Egyptian-Coptic (3000 BC-1350 AD), mainly because of the different writing systems that have been used throughout the history of this language, and because of the diverging aims and scopes of the projects involved in creating annotated corpora of Ancient Egyptian-Coptic. In collaboration with several leading projects in this field, Thot is coordinating the definition of a XML/TEI interchange format in order to enable this data sharing (read more).

- Patrimonium ERC-StG 716375
 - Is NOT an epigraphical project *per se*
 - But a multidisciplinary socio-economic study of the properties of the Roman emperors from Octavian/Augustus to Diocletian
 - Database will be made of both textual AND archaeological sources
 - Ancient texts will be **encoded and enriched** with metadata (places, people, institutions, economic notions...)
 - Use of **TEI / EPIDOC**



- Patrimonium ERC-StG 716375

As it is NOT an epigraphical project *per se*...

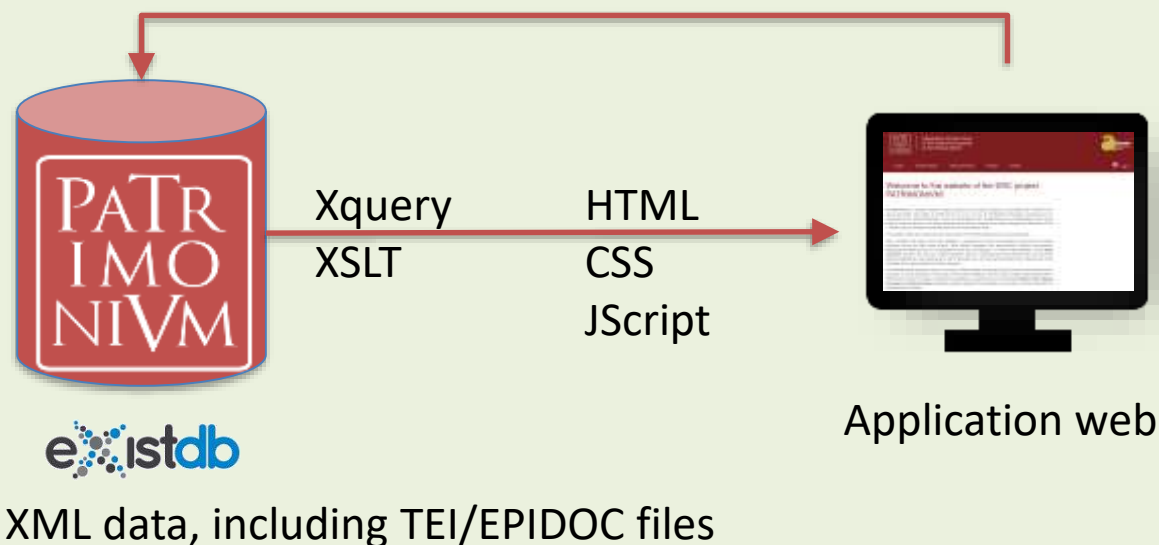
... we needed a tool to make the encoding of texts easier

Notably when enriching text with metadata



PATRIMONIUM » PETRAE EDITOR

- Web application for:
 - encoding
 - querying and browsing data
 - Curation of metadata is done via an integrated thesaurus manager
- Textual AND archaeological data are XML files
- → ALL in 1 SINGLE database





<http://exist-db.org/>

- A NoSQL (XML native) database & Application platform
- EPIDOC files can be directly created, stored and curated ‘as such’
- ‘library’ approach:
 - 1 app for you project
 - Using a library that is installed beside





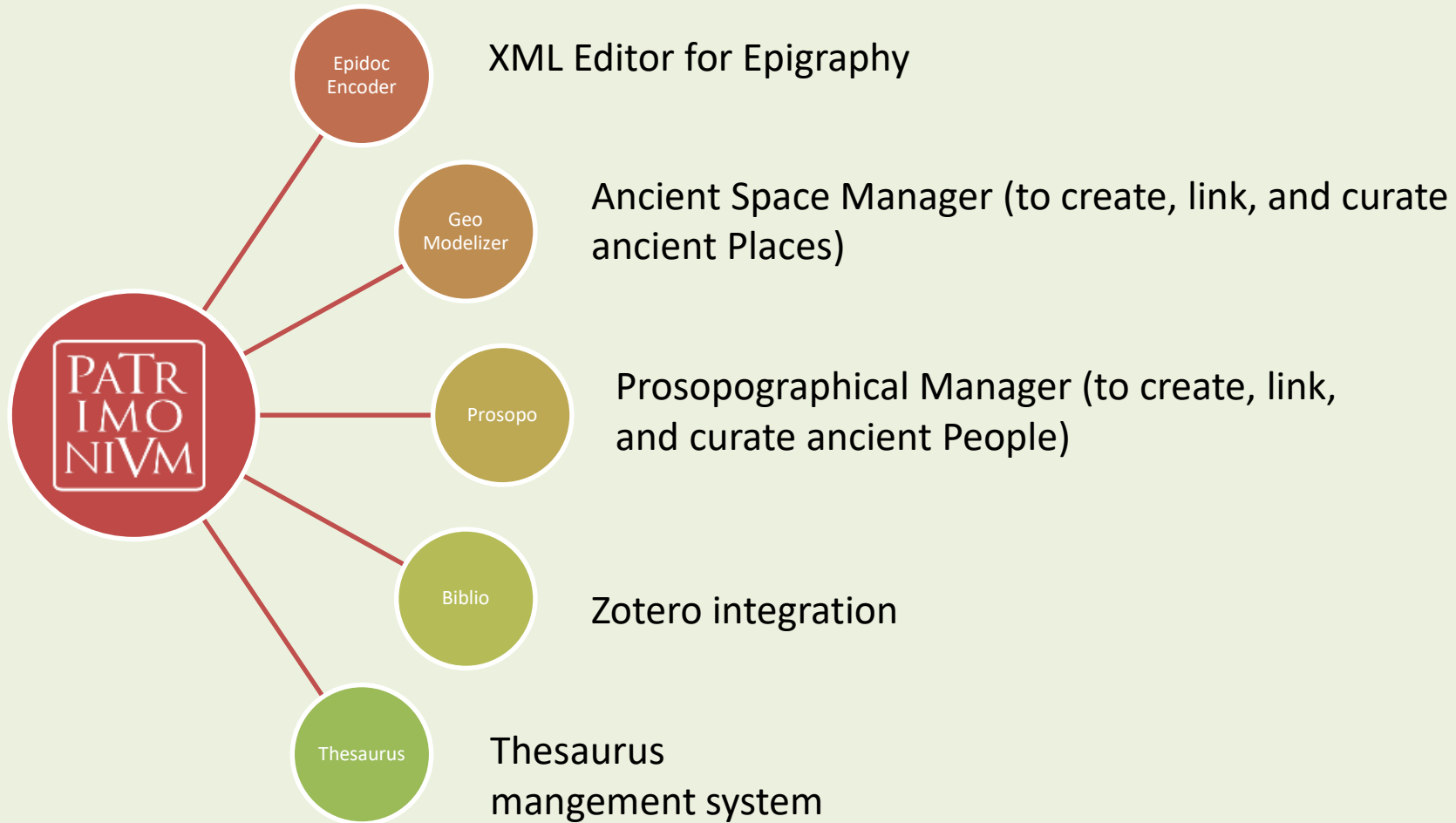
<http://exist-db.org/>

- library contains:
 - Xquery functions
 - JScript
 - Css
 - Other resources (images, fonts).
- App contains functions to build the HTML page and do things specific to the project
 - General layout
 - CSS
- Data is stored outside the app

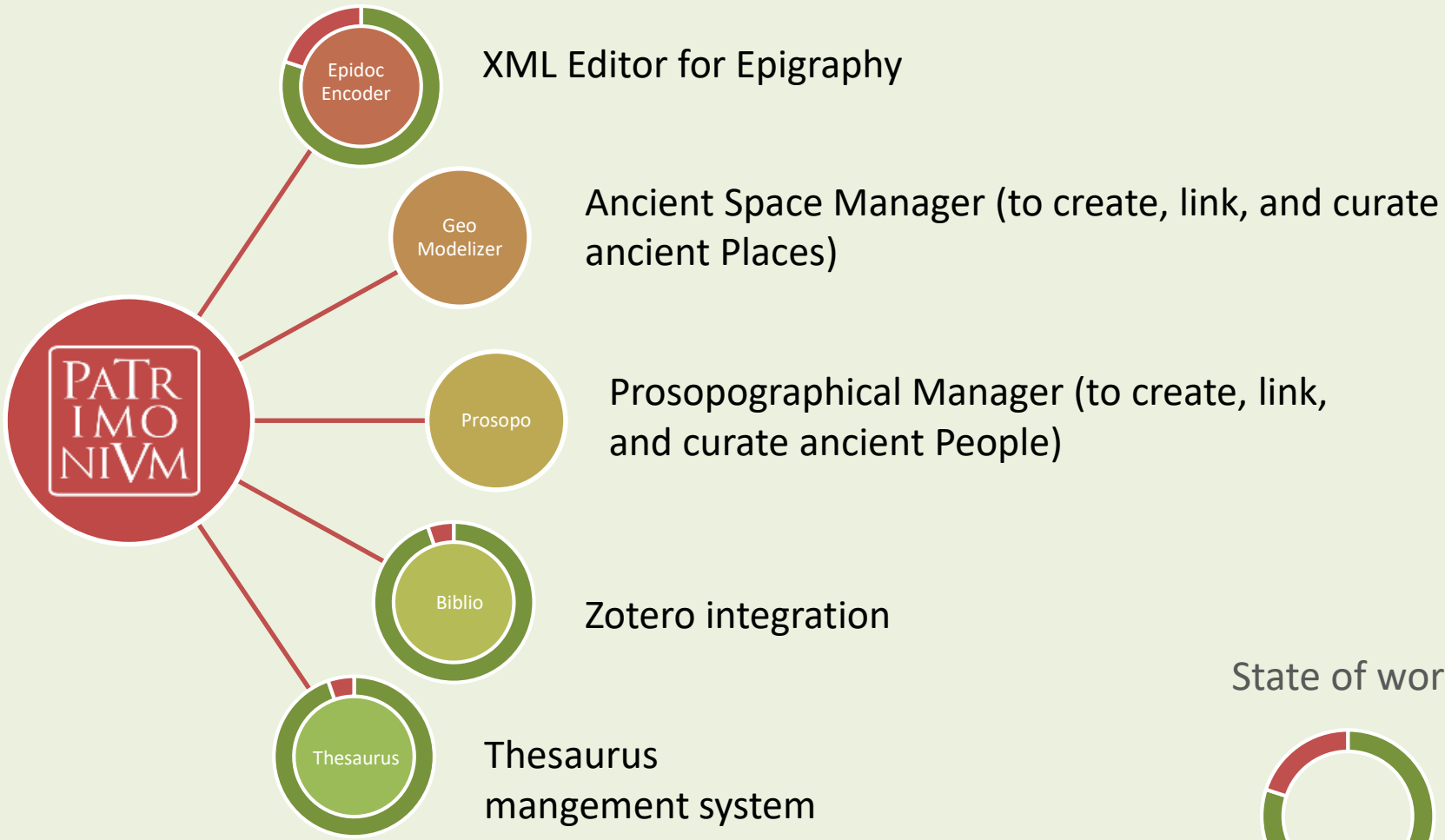
DB structure

```
db/  
|-apps/  
    |-ausohnum/  
        |-modules  
            |-teiEditor  
            |-document-editor  
            |-skosThesau  
|-ausohnum-library/  
|-ausohnumData/  
    |-concepts  
    |-documents/
```

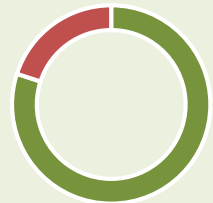
A modular platform



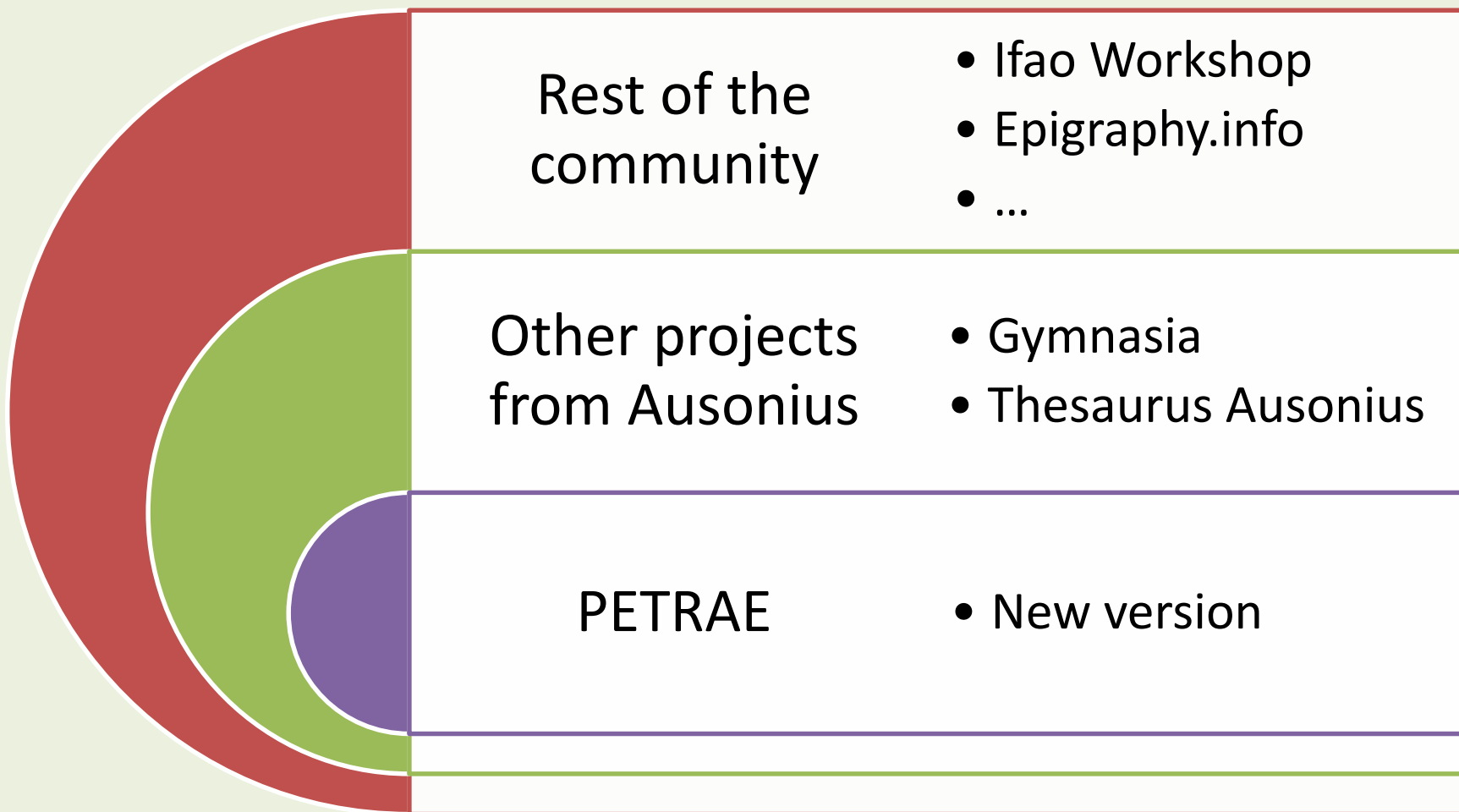
A modular platform



State of work



PATRIMONIVM » PETRAE EDITOR AND THE OTHERS



CORE PRINCIPLES

PATRIMONIVM >> PETRAE EDITOR

CORE PRINCIPLES

- App writes and reads XML/TEI files
- Front-end interface can be customized to display specific elements of a TEI document
- Editing an element is done through user-friendly input fields, dropdown lists or other widgets
- App comes with tools dedicated to text encoding and metadata enrichment

Web application interface

- Data spread over several tabs

Décret pour C. Iulius Capito

Document Metadata

Support

Text Metadata

Text Units

XML file

Editing document t1

Title



Décret pour C. Iulius Capito



Edition

Digital edition encoded by <ref target="http://gymnasia.huma-num.fr">Project Gymnasia</ref> <date when="2018">2018</date>



Document type

Main Bibliography



Cura et tutela

Trismegistos URI

123



Place of origin

Herakleia Salbakes

<https://ausonnum.huma-num.fr/place/15> <https://pleiades.stoa.org/places/638857>

Keywords



Marble trade



Manumission



Web application interface

- Data spread over several tabs

Décret pour C. Iulius Capito

Document Metadata Support Text Metadata Document t1

Title /TEI/teiHeader/fileDesc/titleStmt/title

Décret pour C. Iulius Capito /TEI/teiHeader/fileDesc/editionStmt/edition

Edition Digital edition encoded by <ref target="http://gymnasia.huma-num.fr">Project Gymnasia</ref> <date when="2018">2018</date>

Document type

Main Bibliography Cura et tutela

Trismegistos URI 123

Place of origin /TEI/teiHeader/profileDesc/textClass/keywords/term

Herakleia Salbakes
https://ausonnum.huma-num.fr/place/15 https://pleiades

Keywords <term ref="http://ausonius.huma-num.fr/concept/c19396">
Marble trade</term>

Marble trade Manumission

How elements are displayed?

- Elements are defined in an XML file and called within the editor itself
- Definitions can be customized to meet project's needs
- An element is/can be defined with:
 - Nickname
 - XPath, i.e. place of the element in the TEI file
(`/tei:teiHeader/tei:profileDesc/tei:textClass/tei:keywords/tei:term/@ref`)
 - Content type (node only; node and attribute ; attribute)
 - Cardinality
 - Label
 - link to the top-level concept from which concepts will be displayed (if thesaurus is required)

Element details

```
<teiElement>
  <nm>docKeywords</nm>
  <xpath>/tei:teiHeader/tei:profileDesc/tei:textClass/tei:keywords/tei:term/@ref</xpath>
  <contentType>textNodeAndAttribute</contentType>
  <cardinality>x</cardinality>
  <formLabel xml:lang="en">Keywords</formLabel>
  <formLabel xml:lang="fr">Mots-clés</formLabel>
  <taxold/>
  <fieldType>combobox</fieldType>
  <thesauDb>internal</thesauDb>
  <thesauTopConceptURI>http://ausonius.huma-num.fr/concept/c19365</thesauTopConceptURI>
  <attributeValueType>uri</attributeValueType>
  <xpath2ref/>
  <comments/>
</teiElement>
```

Function's call in 'document-editor.xql':

```
{ teiEditor:displayElement('docKeywords', (), (), ()) }
```

Document-editor.xql

XQuery file in project Web Application

- can be customized to meet project's needs
- produces HTML output

```
<div class="tab-pane fade in active" id="nav-metadata" role="tabpanel" aria-labelledby="nav-metadata-tab">
```

```
{teiEditor:displayElement('docTitle', (), (), ())}  
{teiEditor:displayElement('editionDesc', (), (), ())}  
{teiEditor:displayElement('docType', (), (), ())}  
{teiEditor:principalBibliography()}  
{teiEditor:displayElement("tmNumber", (), (), ())}  
  
{teiEditor:displayElement("placeOfOrigin", (), (), ())}  
{teiEditor:displayElement("dateOfOriginNotBefore", (), (), ())}  
{teiEditor:displayElement("dateOfOriginNotAfter", (), (), ())}  
  
{ teiEditor:displayElement('docKeywords', (), (), ()) }  
{ teiEditor:displayElement('docGeneralCommentary', (), (), ()) }
```

```
</div>
```

INTERFACE

Web application interface

- Data spread over several tabs

Décret pour C. Iulius Capito

Document Metadata

Support

Text Metadata

Text Units

XML file

Editing document t1

Title



Décret pour C. Iulius Capito



Edition

Digital edition encoded by <ref target="http://gymnasia.huma-num.fr">Project Gymnasia</ref> <date when="2018">2018</date>



Document type

Main Bibliography



Cura et tutela

Trismegistos URI

123



Place of origin

Herakleia Salbakes

<https://ausonnum.huma-num.fr/place/15> <https://pleiades.stoa.org/places/638657>

Keywords



Marble trade



Manumission



Décret pour C. Iulius Capito

Editing document t12

Document Metadata Support Text Metadata Text

Title Décret pour C. Iulius Capito

Edition Digital edition encoded by <ref text="http://gymnasia.huma-num.fr">Project Gymnasia</ref> <date when="2018">2018</date>

Document type

Main Bibliography +
Dalla Rosa 2017, p. 78

TM Number xxx

Keywords +

`/TEI/text/body/div[@type="bibliography"][@subtype="edition"]`

Can search and add bibliographical references from project's zotero library

Insert a Bibliographical Reference

Search in Zotero Group 2094917

Dalla Rosa 2017, Les successeurs désignés du prince et leur patrimoine. À propos de la délimitation d'Er-Rahel en Maurétanie Césarienne (CIL, VIII 21663) (Zotero key: IPGTYFBA)

Camilli 2018-02-14, Contributi per un'archeologia di cantiere: i bolli laterizi dalla necropoli di Porto all'Isola Sacra (Zotero key: V94EZZB)

Dalla Rosa 2014, Cura et tutela: l'origine del potere imperiale sulle province proconsolari (Zotero key: M5JC8SVU)

Dalla Rosa 2016, From Exploitation to Integration: Imperial Quarries, Estates and Freedmen, and the Integration of Rural Phrygia (Zotero key: 2KFHUUWD)

Dalla Rosa 2017, Propriété familiale, pouvoir impérial : origine et gestion du patronat d'Augusta en Asie Mineure (Zotero key: WWH75V9)

Date: 2017

Title: Les successeurs désignés du prince et leur patrimoine. À propos de la délimitation d'Er-Rahel en Maurétanie Césarienne (CIL, VIII 21663)



URL: undefined

IPGTYFBA



Add reference Cancel


Document Metadata **Support** Text Metadata Text Units XML file

Support Type +

Stèle  












Material +

Marble  

State of preservation (desc.) Physical condition: <rs ref="fno_entry-see-text">complet</rs> 

State of preservation (keyword) +

Text Support

Findspot	Dimensions (mm)	Current & past locations
Place - 	Height - 	Country - 
Discovery date (not before) - 	Width - 	Place - 
Discovery date (not after) - 	Depth - 	Place +
Finder +		Repository - 
Geo-coordinates - 		Main identifier - 
Certainty		

Document Metadata

Support

Text Metadata

Text Units

XML file

Abstract Text Metadata

This TEI document has **1 msItem** Element

msItem 1

Modern title ⓘ

Free text



Main language ⓘ

Ancient Greek



Other language(s) ⓘ



Ref. to canonical text ⓘ

ref-to-canonical-text-uri



Writing Support(s)+

This TEI document has **1 /sourceDoc/surface** element

1 **main-face** | **t1-surface1** |

Figures +

Type ⓘ



Text Layout Metadata

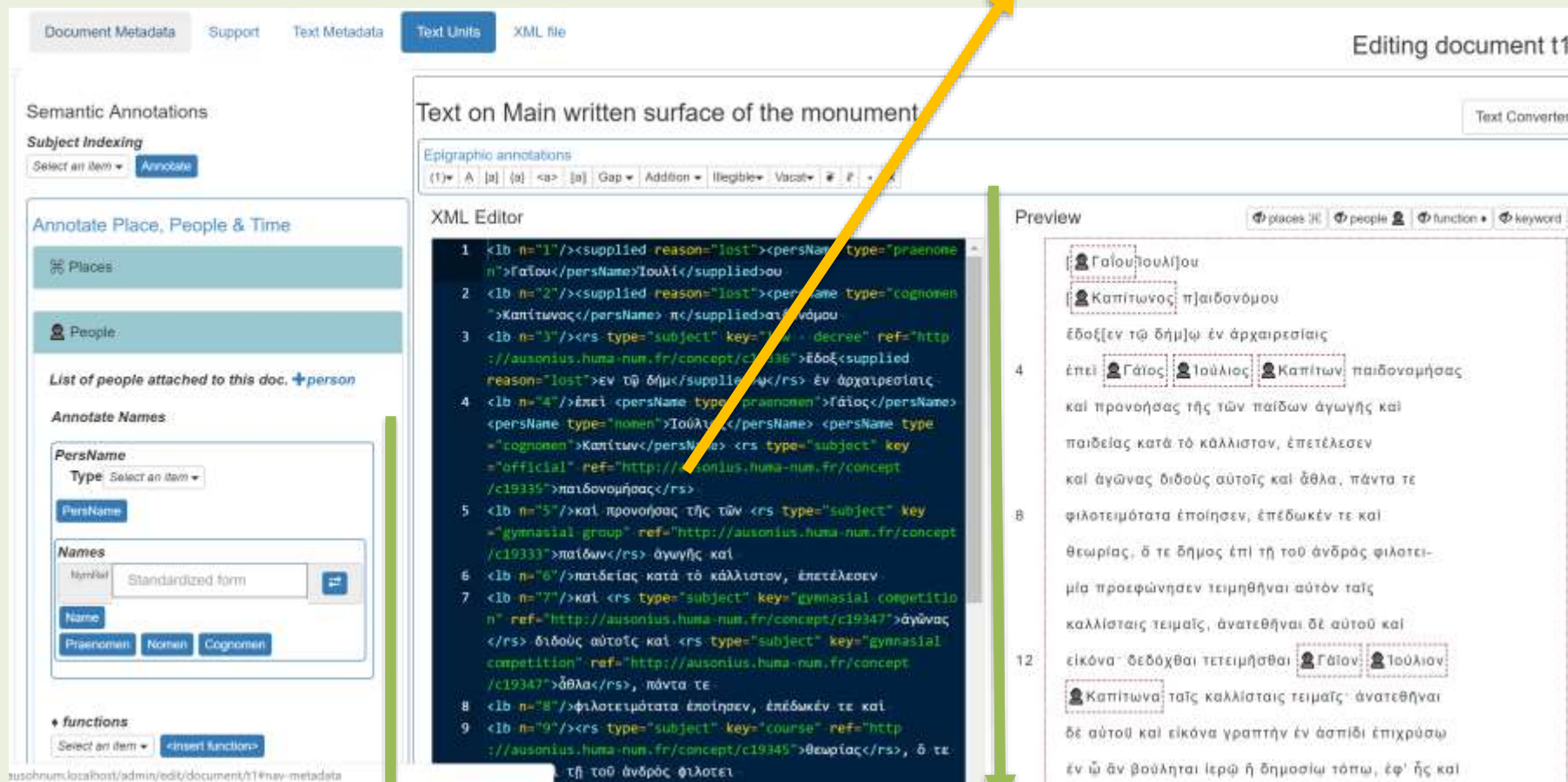
This TEI document has **1 layout** Element

Summary ⓘ



1. **t1-layout1** Surface: main-face | **Description:**

/TEI/text/body/div[@type='edition']/div[@type='textpart']



The screenshot displays the Epigraphic Edition web application. At the top, there are navigation tabs: Document Metadata, Support, Text Metadata, Text Units, and XML file. The main area is titled "Text on Main written surface of the monument" and contains an "XML Editor" with a list of lines of XML code. A yellow arrow points from the URL above to the XML Editor. To the left, there is a "Semantic Annotations" sidebar with "Subject Indexing" and "Annotate Place, People & Time" sections. To the right, there is a "Preview" section showing the rendered text with various annotations like person icons and boxes. A green arrow points from the XML Editor to the Preview section.

Toolbox for encoding "epigraphical" features (lacuna, restitutions...)

Toolbox for encoding "semantic" metadata (placenames, people, topic indexing...)

Document Metadata

Support

Text Metadata

Text Units

XML file

Editing document t1



```

1 <TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="t1" n="1">
2 <!-- @xml:id must be URI beyond the project and unique at corpus level, and should include reference to project whenever it is possible-->
3 <!-- @n should be used to specify the source project-->
4 <teiHeader xml:lang="en">
5 <!-- @xml:lang should be used so that multilingual metadata on a same document could be compiled from different projects-->
6 <fileDesc>
7 <!-- Each file allows for the encoding of 1 to n texts (as witnesses) attested on an (originally) unique document. It is important to keep in mind that the unit encoded is the original document-->
8 <titleStm>
9 <!-- Metadata about the ancient text(s), as a discrete witness(es) in its/their material context (i.e. not the abstract text, for which see /TEI/teiHeader/fileDesc/sourceDesc/msDesc/m-->
10 <title>Décret pour C. Julius Capito</title>
11 </titleStm>
12 <editionStm>
13 <!-- Metadata about the current digital edition-->
14 <edition>Digital edition encoded by <ref target="http://gymnasia.huma-num.fr">Project Gymnasia</ref> <date when="2018">2018</date>
15 </edition>
16 <!-- Plain text summarizing the purpose of the current, digital file.-->
17 <!-- A <date>, with a @when in order to be machine readable, should embrace any date related to the edition-->
18 <!-- A <ref> should point to web page's project-->
19
20 <principal>
21 <persName ref="#pfrölich">Pierre Frölich</persName>
22 <!-- @ref should be used to point to person in /TEI/teiHeader/profileDesc/particDesc/listPerson[@type="editorial"], below-->
23 </principal>
24 <!-- @ref should be used to point to person(s) listed in /TEI/teiHeader/profileDesc/particDesc/listPerson[@type='editorial'], below-->
25 <respStm>
26 <resp>Metadata</resp>
27 <!-- Name(s) of people responsible for metadata input-->
28 <persName ref="#personID">Name</persName>
29 <persName ref="#personID">Name</persName>
30 </respStm>
31 <respStm>
32 <resp>Text transliteration</resp>
33 <!-- Names of the people responsible for the transliteration of the Egyptian text-->
34 <persName ref="#personID">Name</persName>
35 <persName ref="#personID">Name</persName>
36 </respStm>

```


Editing data



Input field



Dropdown menu populated with concepts from internal SKOS thesaurus

Editing data

Document Metadata Support Text Metadata **Text Units** XML file Editing document t1

Semantic Annotations

Subject Indexing
Select an item Annotate

Annotate Place, People & Time

Text on Main written surface of the monument

Text Converter

Epigraphic annotations
(1) A [a] [a] [a] <a> [a] Gap Addition Illegible Vacat

XML Editor Preview

places people function keyword

Text converter

You can paste any text written according to main standards in use in epigraphy

Select a script Ancient Greek

Select source format PHI

Paste your text below

```

1 ἔδοξεν τῆι βουλῆι καὶ τῶι δήμῳ· ἐπ[ι]
2 στεφανηφόρο Δεϊνῶνος, μηνός Κολλυ-
3 ριώνος· Εὐσθένης ἐπρυτάνευε· Ἀναξαγό-
4 ρης εἶπε· Λυκινάδῃ Λευνίδεω κνι[δ]ίω[ι]
5 ἄναγράψαι τῆμ προξενίην καθά[ερ]
6 προγόνους ἀνεγέγραπτο ἐπ' εὐε[ρ]γεσί[ηι]
7 [τῆ]ς πόλεως τῆς Ἰασέων καὶ [εἶ]ναι αὐτ[ῶ]ι
8 [ὥς?] τοῖς προκ[υ]β[ό]νοις ἀτελεῖν κ[α]ὶ προ[ε]δρ[ί]την
9 καὶ ἔσπλον καὶ ἔκπλον ἀουλ[ε]ῖ καὶ ἀσπ[ο]ν]-
10 10 δεῖ καὶ ἐν εἰρήνῃ καὶ ἐμ πολέμῳ, καὶ Ἰα[σέ]-
11 α εἶναι ἕαμ βόληται· ταῦτα δὲ εἶναι καὶ [αὐ]-

```

XML preview

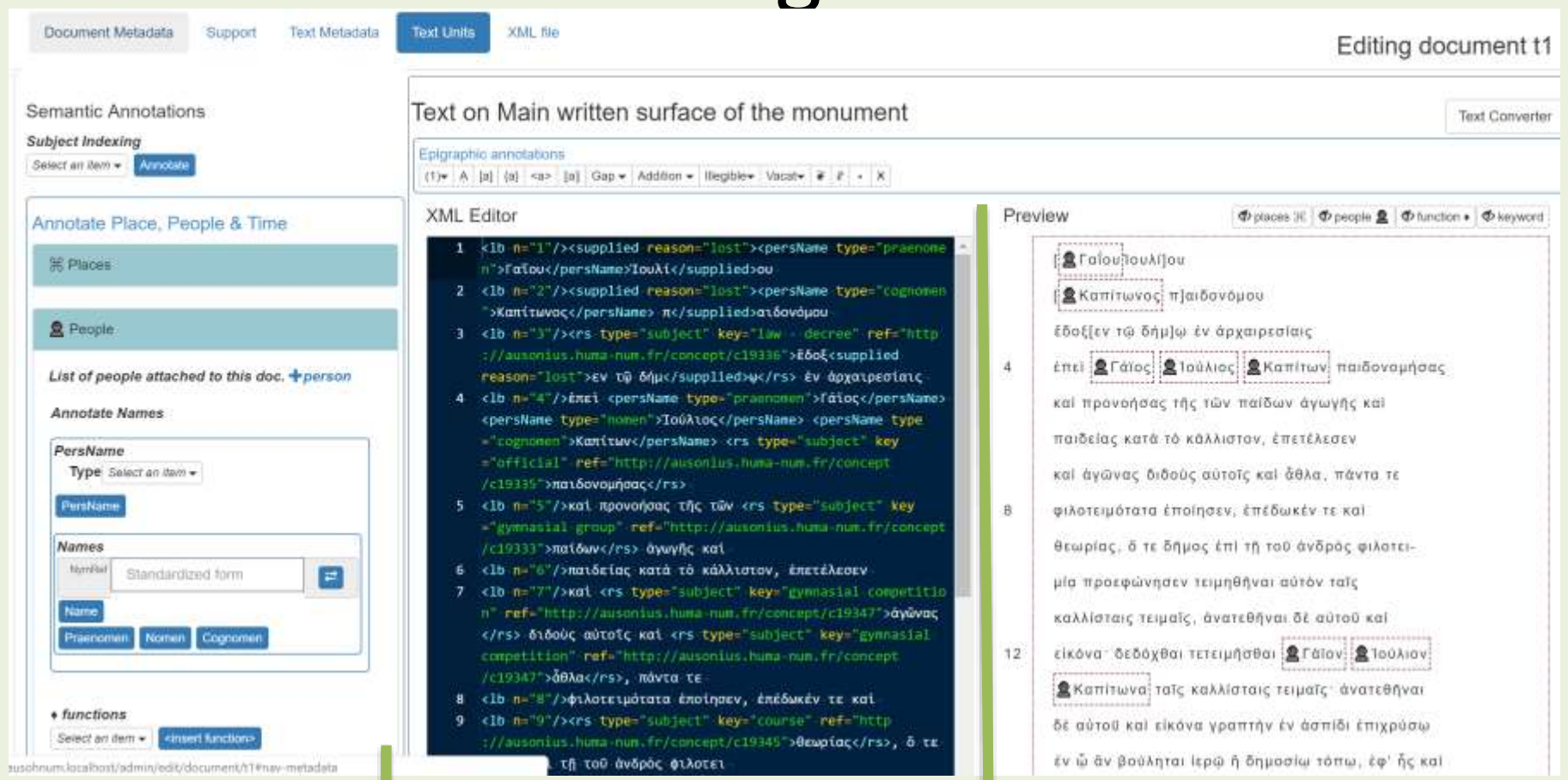
```

1 <lb n="1"/>ἔδοξεν τῆι βουλῆι καὶ τῶι δήμῳ· ἐπ<supplied reason="lost">ι</supplied>-
2 <lb n="2"/>στεφανηφόρο Δεϊνῶνος, μηνός Κολλυ-
3 <lb n="3" break="no">ριώνος· Εὐσθένης ἐπρυτάνευε· Ἀναξαγό-
4 <lb n="4" break="no">ρῆς εἶπε· Λυκινάδῃ Λευνίδεω κνι<supplied reason="lost">δ</supplied>ίω<supplied reason="lost">ι</supplied>-
5 <lb n="5"/>ἀναγράψαι τῆμ προξενίην καθά<supplied reason="lost">ερ</supplied>-
6 <lb n="6"/>προγόνους ἀνεγέγραπτο ἐπ' εὐε<supplied reason="lost">ρ</supplied>γεσί<supplied reason="lost">ηι</supplied>-
7 <lb n="7"/><supplied reason="lost">τῆ</supplied>ς πόλεως τῆς Ἰασέων καὶ <supplied reason="lost">εἶ</supplied>ναι <supplied reason="lost">αὐ</supplied>τ<supplied reason="lost">ῶ</supplied>ι-
8 <lb n="8"/><supplied reason="lost">ὥς</supplied>?<supplied reason="omitted">γό</supplied>
</supplied>νοις ἀτελεῖν κ<supplied reason="lost">αὶ προ</supplied>εδρ<supplied reason="lost">ι</supplied>την</supplied>-
9 <lb n="9"/>καὶ ἔσπλον καὶ ἔκπλον ἀουλ<supplied reason="lost">ε</supplied>ῖ καὶ ἀσπ

```

'standard' text (e.g. from a Word file, a HTML page...)
is pasted in the top panel...
...and converted on the fly
in TEI/XML/Epidoc

Editing data



JavaScript Ace Editor to edit text

HTML preview

Toolbox for encoding "epigraphical" features (lacuna, restitutions...)

Toolbox for encoding "semantic" metadata (placenames, people, topic indexing...)

CREATION OF A DOCUMENT

Creation from template

- Creation of a document is based on a template
- templates are TEI files with minimal TEI/Epidoc elements
- Templates are customizable to meet a project needs

Creation from external resource (beta)

- It is planned to make possible direct import from external resources, when XML/TEI files are available
- Implemented for EDH