

ItAnt

LANGUAGES AND CULTURES OF ANCIENT ITALY

Historical Linguistics and Digital Models



<https://www.prim-italia-antica.unifi.it>

A. Marinetti*, F. Murano°, V. Quochi^, M. Ballerini°, F. Boschetti^, A.M. Del Grosso^, S. Piccini^, L. Rigobianco*, P. Solinas*
*Ca' Foscari University of Venice; °University of Florence; ^Institute for Computational Linguistics (GNR, Pisa)

The project investigates the cultures of ancient Italy on the basis of their **linguistic documentation**, only consisting in **epigraphic evidence**, by means of **computational tools** specifically developed for this purpose. The project focuses on **Venetic, Oscan, Faliscan** and **Cisalpine Celtic** languages.

The project plans to create a digital set of interrelated resources. The **corpus** of epigraphic **texts** will be represented according to **TEI/EpiDoc** guidelines, by an ad-hoc schema describing specific features in languages of fragmentary attestation (i.e. syllabic punctuation, multiple hypotheses of word segmentation, etc). Each text will be described by shared and standard **metadata** concerning linguistic, material and historical data, and will be provided with a **dataset of bibliographic references**.

The project additionally plans to experiment with the creation of a **semantic dataset** of interpretations by using **CRImtex** and **CRMinf** extensions of CIDOC CRM ontology. As regards the **multilingual computational lexicon**, the requirements are being investigated for the design of an efficient computational lexical model specifically dedicated to Restsprachen. We aim at adopting **Semantic Web** standards and vocabularies for providing a structured and formal representation of the lexical items and their related information as well as for allowing future linkability to other related resources. The challenges range from lemmatisation issues to sense representation, meanings being often only partially and hypothetically reconstructible.

Tools and resources will finally be made available through relevant **European RIs** (i.e. CLARIN and/or DARIAH), for long-term preservation, wider dissemination and valorisation of this heritage.

Our **challenge** is to develop a new methodology based on specific digital technologies facilitating scholars' work and collaboration, and ensuring long term preservation of the acquired knowledge.

At the end of the project it is expected that both the **methodology** as well as the **digital toolkit** and **technology** set up will be ready to be extended to the **other Restsprachen**.

