

The digital representation of DiTMAO as a termino-ontological resource
(*Dictionnaire des Termes Médico-botaniques de L'Ancien Occitan* : communication n° 3)

Andrea Bellandi¹ – Emiliano Giovannetti¹ – Anja Weingart²
¹Istituto di Linguistica Computazionale “A. Zampolli”, CNR, Pisa
²Georg-August-Universität, Göttingen

In the talk we will present the technical aspects of the *Dictionnaire des Termes Médico-botaniques de L'Ancien Occitan* (DiTMAO)¹, in particular the conception of DiTMAO as a termino-ontological resource. In addition, we will present LexO, a collaborative editor for multilingual lexica and termino-ontological resources that has been developed in course the project (cf. Bellandi et.al 2018).

In the first part of the talk, the conceptual background of DiTMAO as ontology-lexicon is given. An ontology in computer science is a means to formally model a domain of knowledge or discourse (cf. Guarino et.al 2009). The standard languages to describe knowledge, concepts or the things in the world are RDF/OWL². The formal description, some details of which are given below and will be explained in the talk, aims at inferences and natural-language based queries, such as (in the context of DiTMAO) “*Which plants have been considered dry and hot in the humoral system?*” or “*Which variants are only documented in Hebrew script?*”

From the lexicon standpoint, all the terms in Old Occitan and other languages have been represented by means of the lemon/ontolex model, that is the primary mechanism for the representation of lexical data in the Semantic Web, developed in the context of the W3C OntoLex community group. According to the linked data principles, the descriptions of terms have been linked to different vocabularies, such as lexinfo³ (providing linguistic descriptions), and *dcterms*⁴ (containing bibliographic information). The lemon model has been also extended with respect to the needs of a multilingual and multialphabetical historical dictionary, in particular a rich domain for the documentation of lemmas and variants, cf. Weingart et.al (2016).

From the ontology standpoint, the domains of botany, human anatomy, medication, diseases, and medical instruments are being formalized by means of the Ontology Web Language (OWL), that is a formal language designed to represent rich and complex knowledge about things, groups of things, and relations between things.

The ontolex/lemon model allows to enrich the terms of the lexicon with their semantic descriptions, formally defined in the above domain ontologies. The second part focuses on the presentation of LexO, which is the editor that allows the DiTMAO team to enter and organize all information concerning the complex Old Occitan medico-botanical vocabulary. Starting with a brief introduction and the general features of LexO (such as collaboration, sorting, query and export features), the main focus will be on:

- exemplifying the DiTMAO-lemon formalization of some Old Occitan entries and their representation in the graphical user interface of LexO;
- documenting terms in manuscripts, books, maps, and so on;
- showing some query example.

The talk will end with some perspectives for the near future.

¹ The DiTMAO is being developed within the project ‘An XML-based Information System for Old Occitan Medical Terminology’, a joint project of the PIs Gerrit Bos (Universität zu Köln), Emiliano Giovannetti (Istituto di linguistica computazionale “Antonio Zampolli”, CNR; Pisa), Maria Sofia Corradini (Università di Pisa) and Guido Mensching (Georg-August-Universität Göttingen). The project is funded by the Deutsche Forschungsgemeinschaft (DFG).

² <https://www.w3.org/TR/owl-feature/>

³ <http://www.lexinfo.net/ontology/2.0/>

⁴ <http://dublincore.org/documents/dcmi-terms/>

References

Bellandi, A., Giovannetti, E., and Weingart, A. 2018. Multilingual and Multiword Phenomena in a lemon Old Occitan Medico-Botanical Lexicon. *Information*, 9(3), 52; doi:10.3390/info9030052.

Guarino, Oberle & Staab 2009: What is an ontology? In: S. Staab and Studer, R (eds): *Handbook of ontologies*, 2nd ed. Springer: Berlin, Heidelberg. 1-21

Weingart, A., Giovannetti, E. 2016. Extending the lemon Model for a Dictionary of Old Occitan Medico-Botanical Terminology. In: Sack H., Rizzo G., Steinmetz N., Mladenić D., Auer S., Lange C. (eds) *The Semantic Web. ESWC 2016. Lecture Notes in Computer Science*, vol 9989. Springer, Cham. pp. 408-421