

## **Maurizio LANA, Alice BORGNA and Nadia ROSSO**

*digilibLT – What it is and what it can bring*

DigilibLT [[www.digiliblt.uniupo.it](http://www.digiliblt.uniupo.it)] is a digital library of Late-Antique Latin texts [<http://www.digiliblt.uniupo.it/progetto.php>].

The Canon of the pagan, prose late Latin texts of the library is available at [http://www.digiliblt.uniupo.it/g\\_links\\_autori.php](http://www.digiliblt.uniupo.it/g_links_autori.php). The Grammatici Latini are also being included (they date from the classical to the late Latin period) and the inclusion of Roman Law texts is about to begin.

The works are annotated in XML-TEI, following the specifications available at [http://www.digiliblt.uniupo.it/upload/docs/note\\_di\\_trascrizione\\_e\\_di\\_codifica.pdf](http://www.digiliblt.uniupo.it/upload/docs/note_di_trascrizione_e_di_codifica.pdf).

The accuracy is very high: after OCR, the texts undergo two (and for some works, three) different complete (proof)readings. It is usual to spot typos in the printed source editions.

The editions are the most authoritative ones, chosen by the scientific committee of Latin Scholars, regardless of the time of printing because only the established text is given (no prefatory matters, apparatus(es), comments).

In the context of digilibLT a recent Latin dictionary (1980's) is available. We would like also to suggest the digitization of the dictionary, in order:

- to extract the lemmata and translations (traducenti);
- to use the structure of lemmata-translations as the basis for improving the lexical coverage of the Latin WordNet;
- to make of that same structure the skeleton over which to build a collaborative dictionary.

Global aim: a collaborative dictionary covering every period of use of Latin, deepened/enriched with NLP knowledge, conceived and built as a Linked Open Data network. The dictionary could have a back-end allowing for rich, complex, faceted construction and consultation. It would be an unicum both in itself and because it would allow for different types/levels of consultation.

The project envisages the following key features:

- date, author, title and genre of work; name of researcher, date of editing, ... are recorded and can be used as filters for the consultation;
- the dictionary is available to all libraries/projects involved in LiLa;
- all the participants in the projects, and other people specifically authorized by the different projects, can contribute to the dictionary;
- the collaborative content of the lemmata is mainly Latin passages; translations are optional; for every word/passage with 'problems' discussed in a scientific publication, the complete reference to this critical source can be (should be) given;
- in the front-end, beyond the usual textual ways of entering a dictionary, visual tools will be offered to track the distribution of (a) given linguistic phenomena over a specific period, eventually crossed with genre, author, etc.