

WL4102 | Languages and information technologies

This module will allow you to explore ways in which information technologies are used today to build dedicated tools and resources for languages, and for handling linguistic data (e.g. in dictionaries). These tools can also have implications for language learning. The module will provide you with the opportunity to use these tools directly, e.g. in developing your own projects on the basis of major corpora and databases in the languages that you are studying; the final sessions of the module will be based on your own research into available tools and resources. You will also engage with web-based information technologies by creating and maintaining your own independent website, which will be the basis of the assessment of the module.

Schedule

The module will be taught over the twelve weeks of the first semester. Discussion and lab sessions will alternate in each week.

1. Can a machine read? — Creating a Domain of One's Own
2. How does a database work? — Using Markdown
3. What is a markup language? — Using HTML and CSS
4. What are linguistic data? — Using XML
5. How do we encode linguistic data? — Using the TEI guidelines
6. What is a corpus? — Using the semantic web
7. Presentations: tools and technologies
8. What is a WordNet? — Using BabelNet
9. World languages and their dedicated linguistic resources
10. Case studies
11. Applications of language resources
12. Case studies

Technologies and tools

The main medium in which you will work will be your own domain, which you will set up through [Reclaim Hosting](#) (to which you will need to take out a subscription). You will also become familiar with the major standards using in web domains and will explore a number of the major markup languages used for this purpose. Your work will also focus on dedicated tools used in the construction of major linguistic databases and corpora. Your work will involve using relevant dedicated software (much of it open source and freely available) and also web-based tools for making content available online.

Assessment

All of your written assessments will take the form of material written for the web and published by you in the independent domain that you will create and maintain for this purpose. In addition, you will make a presentation in class on the tools and technologies which will be explored in the course of our sessions. One of the required assessments will be an extended blog essay to be completed by the end of the semester. You will also have the opportunity to write shorter pieces of content on material covered in class throughout the course of the semester.