

Inicio > Computational Linguistics

## **Computational Linguistics**



We apply computational methods to study language and analyse texts, specialising in NLP techniques for text mining, data processing, and historical document transcription. Our current focus is automating medieval manuscript transcription and extracting structured data using LLMs.

## Summary

Computational Linguistics is often defined as the scientific study of language from a computational perspective. Our research line focuses on investigating new ways of modeling linguistic phenomena, addressing linguistic questions through computational methods, and analyzing large text corpora using natural language processing techniques. Our research aims to achieve a better understanding of language as a means of cultural expression and its relationship with both past and present societies.

We collaborate with other research lines in the Computational Humanities group (Historical Archives, Museums and Cultural Heritage, and Computational Archaeology), contributing new methodological approaches to analyzing large text corpora: from data collection, selection, and processing to linguistic and data analysis, text mining, and information extraction.

Currently, we are working with the Historical Archives research line to develop a tool for the automatic transcription of medieval manuscripts and creating LLM-based strategies to extract structured information from these documents. Our project aim is to make the knowledge contained in these documents accessible and understandable to both experts and the public, enabling researchers from various fields (e.g., linguistics, social sciences, economic history, medical history) to gain easier access to new knowledge about the past

and establish new connections for historical analysis.

## Objectives

- Study language using computational methods.
- Collaborate on text analysis and automatic transcription of medieval manuscripts.
- Make historical knowledge more accessible to experts and the public.
- Facilitate new insights across fields like linguistics, social sciences, and history.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 15 Mayo 2025 - 08:53): <u>https://www.bsc.es/es/research-development/research-areas/social-simulation/computational-linguistics</u>