

- Project
- [People](#)

Dating Medieval Texts through Machine Learning:



The LexiChron project is exploring methods for determining the chronology of otherwise undated ancient and medieval texts. In the pre-Modern period, we are often reliant on attributions to known authors to establish a text's historical timeframe, but in many cultures texts may lack any authentic attribution or other internal indication of their date. Accurate dating is essential because historical and literary texts that lack an agreed and precise chronology cannot be situated within their correct social, political, historical and intellectual context.

Scholars working in these areas are often almost wholly reliant on linguistic dating, particularly where works survive only in later manuscripts. Traditional linguistic methods for dating texts are enormously time consuming and often lead to substantially varying results. The LexiChron project is exploring the potential of computer-assisted document dating to provide a chronology for large numbers of texts. Quite apart from the increased capacity and speed of these methods, electronic dating can provide scholars with verifiable levels of confidence in the dates supplied.

The task, known variously as Text Dating, Diachronic Text Evaluation, Temporal Text Classification, and Document Dating, is both theoretically and practically very interesting. Automated text dating systems are trained from a corpus of texts annotated with time stamps. Ordinal regression and multi-class classification are the popular algorithms in current use, and these are being developed further by the LexiChron project to improve performance for the dating of ancient and medieval texts.

The project has potential to impact on dating methodologies across a wide range of ancient and medieval cultures, and more generally on historical linguistics. Emergent fields, such as computational forensics and computational journalism, and more traditional tasks, such as discourse similarity, sense shifting, readability and narrative frameworks, may also benefit from a system capable of dating texts automatically.

Contacts: x.han@qub.ac.uk g.toner@qub.ac.uk

University Road, Belfast BT7 1NN, Northern Ireland, UK