

Introduction to Digital Text Analysis

Doctoral seminar (3 ECTS)

School of Economics and Business, University of Ljubljana

Seminar instructor

Ivan Zupic (Email: i.zupic@gold.ac.uk)

Institute of Management Studies, Goldsmiths, University of London

Seminar summary

Management and other social science researchers are increasingly using digital methods for analysing text. Methods like topic modeling and sentiment analysis have started to proliferate in recent years, yet their use is still in its infancy. The seminar will provide an introduction to these cutting edge methods and equip students with skills to start their own text analysis projects.

Seminar objectives

- Provide an overview of general principles of digital content analysis.
- Introduce students to the diverse digital methods of analysing text.
- Increase awareness of the challenges of current state-of-the-art methods of analysing text.
- Provide hands-on introduction to the process of digital text analysis.

Prerequisites

There are no prerequisites required. Part of the seminar will use R for analysis so it is desirable if students have basic knowledge of R. It is not essential to have good knowledge as the code for text analysis will be provided and will require only modest changes. If students wish to use another software, they will be able to use trial version of Leximancer for the workshop instead of R.

Students are advised to install R Studio on their laptops before the seminar.

Learning methods and schedule

The seminar will be a combination of lectures, discussions, critical analysis and hands-on training in analysing data.

Tentative outline:

Day 1:

- Introduction to digital content analysis

- Overview of methods for text analysis (topic modeling, sentiment analysis, natural language processing, ...)
- Analysis of published papers that used digital methods for text analysis
- Overview of software packages commonly used for text analysis

Day 2:

- Hands-on training for analysing text (in R or other software package)
- Overview of the challenges of using digital text analysis methods

Evaluation

Students will need to prepare a research proposal (1000-1500 words) that includes the following:

- Motivation for research
- Research questions
- Potential contributions to the literature
- Short literature review
- Descriptions of methods that will be used
- How data will be acquired

The deadline for submission will be two weeks after the seminar.

Core Literature

Antons, D., Breidbach, C. F., Joshi, A. M., & Salge, T. O. (2021). Computational literature reviews: method, algorithms, and roadmap. *Organizational Research Methods*, 1094428121991230.

Blei, D.M., 2012. Probabilistic topic models. *Communications of the ACM*, 55(4), pp.77–84.

Hannigan, T. R., Haans, R. F., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., ... & Jennings, P. D. (2019). Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, 13(2), 586-632.

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihok, G., & Den Hartog, D. N. (2018). Text classification for organizational researchers: A tutorial. *Organizational Research Methods*, 21(3), 766-799.

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2018). Text mining in organizational research. *Organizational Research Methods*, 21(3), 733-765.

Krippendorff, K., 2012. *Content Analysis: An Introduction to Its Methodology* Third edit., SAGE Publications.

Additional readings will be given before seminar start.