Faculty of Social Sciences

SIMM71, Computational Content Analysis for the Social Sciences, 7.5 credits

Datavetenskaplig innehållsanalys för samhällsvetenskap, 7,5 högskolepoäng

Second Cycle / Avancerad nivå

Details of approval

The syllabus was approved by Graduate School Board on 2022-02-22 to be valid from 2022-08-29, autumn semester 2022.

General Information

The course is offered as an interdisciplinary single subject course in the Social Sciences at the second-cycle level, and as an optional course within the Master Programmes in Social Scientific Data Analysis, Global Studies, Middle Eastern Studies, Social Studies of Gender and Development Studies.

Language of instruction: English

Main field of studies	*Depth of study relative to the degree requirements*
Political Science	A1F, Second cycle, has second-cycle course/s as entry requirements
Social Anthropology	A1F, Second cycle, has second-cycle course/s as entry requirements
Gender Studies	A1F, Second cycle, has second-cycle course/s as entry requirements
Sociology of Law	A1F, Second cycle, has second-cycle course/s as entry requirements
Sociology	A1F, Second cycle, has second-cycle course/s as entry requirements
Social Work	A1F, Second cycle, has second-cycle course/s as entry requirements
Human Geography	A1F, Second cycle, has second-cycle course/s as entry requirements
Development Studies	A1F, Second cycle, has second-cycle course/s as entry requirements

Learning outcomes

Upon completion of the course, the student shall be able to:

Knowledge and understanding

- demonstrate knowledge of computational methods such as natural language processing and computer vision, and their application to communication and media content;
- demonstrate understanding of how communication and media content can be analyzed to investigate current social and organizational issues;

Competence and skills

- exemplify skills in extracting content features from texts through computational methods including but not limited to text classification;
- exemplify skills in extracting content features from images and/or video through computational methods including but not limited to object detection;
- demonstrate the ability to summarize and visualize results from computational content analysis of text and images;

Judgement and approach

- independently and critically reflect on the ethical implications of computational content analysis of communication and media material;
- independently and critically reflect on, and make informed decisions with regard to, methodological choices in the application of computational content analysis.

Course content

This course is aimed towards students who have some prior knowledge of quantitative research methods and wish to further develop their understanding of content analysis, and ability to independently apply computational methods of extracting content features from digital texts and images. The course focuses on content analysis as a method to gain knowledge about current and historical social issues by analyzing observations of communication and media messages. Some of the computational content analysis techniques most commonly used within the social sciences, such as natural language processing and computer vision, are presented and practiced, and their connection to the field of artificial intelligence and machine learning is discussed. The focus lies on applying these methods and techniques and presenting the results through data visualizations.

Course design

Teaching includes lectures and teacher assisted exercises in practical computational content analysis and data visualization (computer lab work). Unless there are valid reasons to the contrary, compulsory participation is required in two examination seminars where students have the opportunity to present results from their content analyses by using data visualizations. Students who have been unable to participate due to circumstances such as accidents or sudden illness will be offered the opportunity to compensate for or re-take compulsory components. This also applies to students who have been absent because of duties as an elected student

representative.

Assessment

Content analysis techniques for texts and images are examined separately in two "data visualization reports" presented at seminars. The research reports should also include some contextual information, e.g. reflections on ethical aspects of the analysis techniques employed. Each report is examined individually and is worth 50% of the course grade.

The course includes opportunities for assessment at a first examination, a re-sit close to the first examination and a second re-sit for courses that have ended during that school year. Two further re-examinations on the same course content are offered within a year of the end of the course. After this, further reexamination opportunities are offered but in accordance with the current course syllabus.

The examiner, in consultation with Disability Support Services, may deviate from the regular form of examination in order to provide a permanently disabled student with a form of examination equivalent to that of a student without a disability.

Subcourses that are part of this course can be found in an appendix at the end of this document.

Grades

Marking scale: Fail, E, D, C, B, A.
The grade for a non-passing result is Fail. The student's performance is assessed with reference to the learning outcomes of the course. For the grade of E the student must show acceptable results. For the grade of D the student must show satisfactory results. For the grade of C the student must show good results. For the grade of B the student must show very good results. For the grade of A the student must show excellent results. For the grade of Fail the student must have shown unacceptable results.

The grade for the entire course consists of the average grade of all assessed assignments (A = 5, B = 4, C = 3, D = 2, E = 1). Thus, if the grade C is received on the first report and the grade B is received on the second report, the final course grade would be averaged and rounded up to a B. For a grade of Pass on the entire course, the student must have been awarded at least E on all assessments for which the grading scale A–E+Fail applies.

Entry requirements

To be eligible for the course students must have a Bachelor's degree with a major (i.e. at least 90 ECTS credits) in Development Studies, Gender Studies, Human Geography, Political Science, Social Anthropology, Social Work, Sociology or Sociology of Law, or the equivalent.

To be admitted to the course, students must have completed the course SIMM61 Quantitative Data Analysis in R, 15 credits, or otherwise demonstrated basic working knowledge of relevant programming language. A further requirement is that students have taken at least 7.5 credits of quantitative methods at the advanced level, e.g. SIMM16 Introduction to Quantitative Methods, 7.5 credits, or equivalent.

Oral and written proficiency in English equivalent to English 6/B (advanced) from

Swedish upper secondary school is a requirement. International qualifications will be assessed in accordance with national guidelines.

Subcourses in SIMM71, Computational Content Analysis for the Social Sciences

Applies from H22

2201 Computational Content Analysis for the Social Sciences, 7,5 hp
Grading scale: Fail, E, D, C, B, A