**Syllabus of an educational component of a degree programme**

|  |  |
| --- | --- |
| Name of unit conducting a component | ***Doctoral School of Social Sciences*** |
| Name of an educational component | **Qualitative Data Analysis - focus on Texts** |
| Language of education | English |
| Goals of education | **Understanding of:**   * Text Preprocessing techniques * Various Techniques for Text Analysis   **Practical knowledge:**   * Ability to use advanced Text Mining Tool * Drawing conclusions about the text corpus and text data samples |
| Learning outcomes of an educational component | Theoretical & Practical knowledge on Text Mining Techniques and how to use them in real life data analysis cases with Advanced Text Mining Tools (CAQDAS – Computer Assisted Qualitative Data Analysis Software). |
| Verification methods and assessment criteria of learning outcomes obtained by students | Each student needs to complete tasks assigned to them during each meeting. Course ends with comprehensive project being build from a scratch (defining goals, exploring text mining techniques, completing the analysis, building visuals, writing analytical report). |
| Type of an educational component (obligatory/optional) | Obligatory |
| Year of study | 2nd, 1st |
| Semester | Summer |
| Name and surname of the coordinator of a component and/or person/s conducting a component | Ph.D. Krzysztof Tomanek |
| Name and surname of person/s conducting an examination or granting credit in the case when this sposóis other person than conducting a component | Ph.D. Krzysztof Tomanek |
| Manner of completion | Distance learning - exercises via e-learning platform.  E-learning conversion centre.  **Windows operating system required** |
| Preliminary and additional requirements | Basic understanding on Qualitative Data Analysis vs Quantitative Data Analysis. |
| Type and number of hours of courses requiring  direct participation of academic staff and students, if in a given component such courses are included | 30 |
| Number of ECTS credits assigned to a component | 3 |
| Balance of ECTS credits | 3 |
| Applied teaching methods | Interactive workshops with CAQDAS installed on student’s computers.  Conversations, problem solving, explorations through text mining techniques, and mixed-method approached. |
| Form and conditions of passing a component, including conditions of allowing to take an examination, as well as form and conditions of passing each type of courses included in a given component | Each student needs to complete Text Mining Project that will be presented during the course final meeting. |
| Content of an educational module (with division into forms of courses completion) | 1. Qualitative vs Quantitative Methods. 2. Mixed-Methods approach in Social Sciences. 3. Defining research problems, hypothesis, data sources. 4. Text sampling and building text corpus. 5. Validating text quality: text preprocessing techniques (stemming, lemmatization) and text postprocessing. 6. Building code-book vs research problems and hypothesis. 7. Defining codes, looking for proper unit of analysis. 8. Texts coding phase. 9. Validating coding reliability, coding agreement. 10. Exploring texts: simple searches, bag of words, similarity searches, stop list. 11. Codes – variables: transformations from and into qualitative or quantitative data. 12. Similarity analysis: similar texts, similar topics. 13. Coding co-occurrence: dendrogram and similarity coefficient, link analysis, proximity analysis. 14. Sequence analysis and the role it plays in data explorations. 15. Advanced search techniques: cluster coding, fuzzy string matching, 16. Automations in analysis: logical and statistical rules in automated analytics. 17. TF\*IDF, topic modeling and automated classification methods. 18. Dictionaries as main analytical tools in text mining-oriented analytics. |
| List of basic as well as supplementary literature, knowledge of which is required in order to pass a given component | The Expression of Emotions in 20th Century Books:  <https://www.researchgate.net/publication/236078895_The_Expression_of_Emotions_in_20th_Century_Books>  Thou Art in a Deal: The Evolution of Religious Language in the Public Communications of Donald Trump  <https://www.researchgate.net/publication/344337560_Thou_Art_in_a_Deal_The_Evolution_of_Religious_Language_in_the_Public_Communications_of_Donald_Trump>  Impacts of COVID-19 on global tourism industry: A cross-regional comparison  <https://www.researchgate.net/publication/344254867_Impacts_of_COVID-19_on_global_tourism_industry_A_cross-regional_comparison>  CONTENT ANALYSIS OF MEDIA REPORTING ON PROTESTS FOLLOWING THE DEATH OF GEORGE FLOYD  <https://provalisresearch.com/blog/content-analysis-media-protests-george-floyd/>  Using Content Analysis to Characterise the Sensory Typicity and Quality Judgements of Australian Cabernet Sauvignon Wines <https://www.mdpi.com/2304-8158/8/12/691/htm>  HOW WE ANALYZED U.S. PRIMARY DEBATES WITH WORDSTAT  <https://provalisresearch.com/blog/primary-debates-text-analytics/>  Art Moves in Self Organizing Social Media: Virtual Dynamics Impact Tourism Reality  <https://www.researchgate.net/publication/332422796_Art_Moves_in_Self_Organizing_Social_Media_Virtual_Dynamics_Impact_Tourism_Reality>  QDA Miner & Wordstat manuals  <https://provalisresearch.com/resources/manuals/> |