
האוניברסיטה העברית בירושלים

סילבוס

ניתוח טקסטואלי כמותי - 50069

תאריך עדכון אחרון 09-02-2025

נקודות זכות באוניברסיטה העברית: 2

תואר: מוסמך

היחידה האקדמית שאחראית על הקורס: תקשורת ועיתונאות

השנה הראשונה בתואר בה ניתן ללמוד את הקורס: 0

סמסטר: סמסטר ב'

שפת ההוראה: אנגלית

קמפוס: הר הצופים

מורה אחראי על הקורס (רכז): כריסטיאן באדן

דוא"ל של המורה האחראי על הקורס: c.baden@mail.huji.ac.il

שעות קבלה של רכז הקורס: יום רביעי 14-16

מורי הקורס:

תאור כללי של הקורס:

In our contemporary world, virtually every part of life involves digital texts: Whether we read for our studies or just a cookbook, write work emails, use entertainment media or communicate with our friends, and even when we order stuff online or leave that scathing review on yelp, texts contain immense amounts of information that may be of interest for social scientific research. In this class, we learn how to recover that information from large bodies of text, using quantitative textual analysis. Following some foundational considerations about quantitative content analysis, we address a variety of current developments and challenges in quantitative text analysis. We practice different techniques, including both manual and computational analysis, and see how different strategies can be used to extract different information from digital texts. The class includes an introduction to text mining and semantic network analysis, and discusses different ways in which qualitative and quantitative strategies can be combined to improve our ability to glean more information from texts.

מטרות הקורס:

The class aims to build a firm understanding of the procedures, challenges and capabilities of quantitative text analysis; to discuss important developments in the analysis of digital texts of different kinds (from document archives and news discourse to interactive communication and the variety of texts generated through our daily use of digital media); to enable participants to understand and evaluate existing research done in content analysis, text mining and semantic network analysis; to practice the development of sound strategies for textual analysis and their application in both academic and professional contexts.

תוצרי למידה

בסיומו של קורס זה, סטודנטים יהיו מסוגלים:

- *Understand the principles & procedures, quality criteria and validation strategies needed in quantitative textual research*
- *Distinguish different kinds of analytic options, their respective strengths and weaknesses and key applications*
- *Contextualize insights gained from text-analytic research, including contemporary approaches to text mining, to gain a founded understanding of its implications and limitations*
- *Critically evaluate present developments in quantitative text analysis and its underlying challenges*
- *Perform high-quality content-analytic research, and be able to employ basic software tools for computational analysis to aid the investigation*
- *Design and validate suitable research strategies for quantitative textual research*

- 24.03.25 Getting to terms: Why quantify textual contents?
31.03.25 What's in a word: The classification of meaning *
07.04.25 To code or not to code? The making of intersubjective criteria *
Pessach Break
21.04.25 What could possibly go wrong? Validity, validation, & measurement issues *
28.04.25 Say it exactly: Rule-based approaches to computer-assisted text analysis *
05.05.25 Something like this: (Semi-)Supervised approaches & machine classification *
12.05.25 Probably meaningful: Unsupervised approaches & topic modeling *
19.05.25 Bring your own problem: Challenges in text analysis
26.05.25 No class: Individual consultations
Shavuot
09.06.25 Beyond the text: Finding frames and other latent meaning *
16.05.25 No class: Individual consultations
23.06.25 Trust the magic: (How) Will Generative Large Language Models change everything?
30.06.25 Presentation of measurement strategies & Wrap-up

- Krippendorff, K. (2004). Content analysis: An introduction to its methodology. Thousand Oaks, CA: Sage. Chapter 2 (pp. 18-43)
Mourao, R. R., & Robertson, C. (2019). Fake news as discursive integration: An analysis of sites that publish false, misleading, hyperpartisan and sensational information. Journalism Studies.
Baden, C. & Sharon. T. (2019). Blinded by the Lies? Toward an operational strategy for delineating conspiracy theories in interactive discourse. Paper presented at the 69th ICA Annual Conference, Washington, D.C.
Krippendorff, K. (2004). Content analysis: An introduction to its methodology. Thousand Oaks, CA: Sage. Chapter 7 (pp. 125-149)
Boxman-Shabtai, L. & Shifman, L. (2013). When ethnic humor goes digital. New Media & Society, 17(4), 520-539.
Skalski, P. D., Neuendorf, K. A. & Cajigas, J. A. (2017). Content analysis in the interactive media age. In K. A. Neuendorf (Ed.), The content analysis guidebook (pp. 201-242). Thousand Oaks, CA: Sage.
Muddiman, A., & Stroud, N. J. (2017). News values, cognitive biases, and partisan

- incivility in comment sections. *Journal of Communication*, 67, 5866-5609.
- van Cuilenburg, J. J., Kleinnijenhuis, J., & de Ridder, J. A. (1988). Artificial intelligence and content analysis: Problems of and strategies for computer text analysis. *Quality & Quantity*, 22, 65-97.
- Burscher, B., Vliegthart, R. & de Vreese, C. H. (2015). Using supervised machine learning to code policy issues: Can classifiers generalize across contexts? *The Annals of the American Academy of Political and Social Science*, 659(1), 122-131.
- Hirst, G., Riabinin, Y., Graham, J., Boizot-Roche, M., & Morris, C. (2014). Text to ideology or text to party status? In B. Kaal, I. Maks, & A. van Elfrinkhof (Eds.), *From text to political positions: Text analysis across disciplines* (pp. 93-115). Amsterdam, Netherlands: John Benjamins.
- Günther, E., & Quandt, T. (2016). Word counts and topic models. *Digital Journalism*, 4(1), 75-88.
- Jacobi, C., van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, 4(14), 89-106.
- Doerfel, M. L. (1998). What constitutes semantic network analysis? A comparison of research and methodologies. *Connections*, 21(2), 16-26.
- Kleinnijenhuis, J. & van Atteveldt, W. (2014). Positions of parties and political cleavages between parties in texts. In B. Kaal, I. Maks, & A. van Elfrinkhof (Eds.), *From text to political positions: Tekst analysis across disciplines* (pp. 1-20). Amsterdam, the Netherlands: John Benjamins.
- Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, 58, 258-279.
- Papacharissi, Z., & de Fatima Oliveira, M. (2008). News frames terrorism: A comparative analysis of frames employed in terrorism coverage in U.S. and U.K. newspapers. *The International Journal of Press/Politics*, 13, 52-74.
- Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit. *Digital Journalism*, 4(1), 8-23.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage. Chapter 14 (pp. 339-364)
- Baden, C., Kligler-Vilenchik, N. & Yarchi, M. (2019). Hybrid Content Analysis: Toward a strategy for the computer-assisted classification of large text corpora using topic modeling. Paper presented at the 69th ICA Annual Conference, Washington, D.C.
- Nicholls, T. & Bright, J. (2018). Understanding news story chains using information retrieval and network clustering techniques. *Communication Methods & Measures*.

חומר לקריאה נוספת:

Will be specified in the syllabus

מרכיבי הציון הסופי:

הגשת עבודה מסכמת / פרויקט גמר / מטלת סיכום / מבחן בית / רפרט 50 %
מצגת / הצגת פוסטר / הרצאה / סמינר / פרוסמינר / הצעת מחקר 30 %

השתתפות פעילה / עבודת צוות 20 %

מידע נוסף / הערות: