

From Embeddings to LLMs

Advanced Text Analysis with Python

Fall Seminar 2026

version 2026-09

Instructor:	Ass.-Prof. Hauke Licht , Ph.D.	Dates:	September 21-25, 2026
Affiliation:	University of Innsbruck	Times:	9:30-12:30 and 13:30-16:30
Email:	hauke.licht@uibk.ac.at	Place:	B6 4, 5, D-68159 Mannheim
Materials:	Code and data (Github), readings and slides (<i>tba</i>)		

Course short description This course is designed for social scientists who want to go beyond basic bag-of-words text analysis, which treat words or n -grams as distinct symbols and texts as unordered collections of such symbols, and learn how to use the latest deep learning methods, which overcome the limitations of bag-of-words text analysis, in their research. Participants will be introduced to neural text embedding techniques and learn how they enable modern applications of large language models (LLMs) and related methods in social science analyses.

About the instructor Hauke Licht is assistant professor of Computational Political Science at the Department of Political Science and the Digital Science Center at the University of Innsbruck. He has received his PhD from the University of Zurich and previously worked as post-doctoral researcher at the University of Cologne. In his research, Hauke develops and applies computational text analysis methods to study political communication, electoral competition, and democratic representation with a focus on deep learning methods, annotation, and multilingual analyses.

1 Course Description	2
2 Day-to-day Schedule and Literature	5

Course Description

This course is designed for social scientists who want to go beyond basic bag-of-words text analysis, which treat words or n -grams as distinct symbols and texts as unordered collections of such symbols, and learn how to use the latest deep learning methods, which overcome the limitations of bag-of-words text analysis, in their research. Participants will be introduced to neural text embedding techniques and learn how they enable modern applications of large language models (LLMs) and related methods in social science analyses.

Participants will learn about the conceptual motivation and methodological foundations of text embedding methods and large language models. Participants will also get hands-on experience with applying **text embedding methods** for discovering latent clusters and representative documents in large text corpora. Participants will learn how **large language models** like GPT, Claude, or Qwen work and how to use them for classifying, extracting, scoring, or ranking texts – even when participants have little or no labeled data. Further, participants will fine-tune smaller Transformer-based language models, like BERT, to replicate the annotations of human coders or LLMs and learn how to use advanced techniques like few-shot learning when labeled data is scarce.

Throughout the course, participants will work with real datasets, try out different tools and approaches, and get feedback on their own research ideas. We will focus on practical skills and real research problems, using Python throughout.

Along the way, we will discuss important issues like how to check if their models are working well, how to avoid common pitfalls, and how to think critically about the results participants get from automated text analysis. Ethical considerations and practical tips for using these methods in social science research will be a key part of our discussions.

Importantly, this is an **advanced-level course**. Participants should have prior knowledge of basic text analysis techniques. Specifically, they should have experience with standard bag-of-words pre-processing techniques and text representation approaches, such as word count-based document-feature matrices. Moreover, participants should have experience with programming in Python (see *Course Prerequisites* section below for more details). The lecturer cannot introduce or repeat basics in Python programming in the course due to limited time.

Course and Learning Objectives

By the end of the course, participants will:

- know the methodological foundations of text embedding methods, transfer learning,

Transformers, and large language models (LLMs)

- be able to apply these methods to analyze social scientific text data
- be able to apply these methods to measure social science constructs in text data
- be able to reflect critically about the application of the techniques in social science research, including relevant ethical considerations.

Prerequisites

- Basic knowledge of Python
 - creating and manipulating strings, lists, and dictionaries
 - creating and interacting with objects, classes, and methods
 - reading and manipulating data frames with pandas
 - using loops
 - defining new functions
- Prior knowledge of basic quantitative text analysis methods
 - (conceptual) knowledge of quantitative text analysis methods (e.g., dictionary analysis, topic modeling, or supervised text classification) is strongly recommended
 - bag-of-words text pre-processing (“tokenization”) and representation (i.e., how to represent document with word count vectors) can be an asset
- Basic knowledge of quantitative research methods
 - understanding of linear and logistic regression analysis
 - a basic understanding of matrix algebra might be helpful but is not required

For those who would like a primer or refresher in Python, we recommend taking the online workshop “[Introduction to Python](#)” and/or the online blended learning course “[Introduction to Computational Social Science with Python](#)”.

Software and Hardware Requirements

- **Hardware requirements**

Participants must bring their own laptop to this course. Because we will work with computationally demanding models and methods, participants must determine whether their laptop will be capable of supporting these. Please check the specification of their machine against the following minimum hardware requirements. If their laptop does not meet the requirements, participants will need a *Google Colab* license (see next point).

- **Windows/Linux:** i5 or Ryzen 5 CPU, > 16GB RAM, with a dedicated NVIDIA GPU with > 4GB VRAM (better: 8GB)
- **Mac:** any M-generation laptop (M1–M4) with at least 8GB of unified memory (better: 16GB+)
- **Cloud computing**
 - It is recommended that participants create a *Google Colab* account and are required to do so if their laptop does not meet the minimum hardware requirements specified above. *Google Colab* is a cloud-based alternative to using their own laptop to perform computations. GPU time on *Google Colab* is charged, but relatively inexpensive. Participants who want to use this option should budget a maximum of €20 for GPU credits (e.g. in the form of a monthly *Pro* plan subscription; the current price of the *Pro* plan (March 2026) is €11.01 per month), in case the free quota is not sufficient for the course.
 - It is recommended that participants create a *Hugging Face* (HF) account to be able to use the HF *Inference Providers* service of running open-weights LLMs. Participants should budget about 15€ for using this option (including a *Pro* subscription for the duration of the course).
- **Python**

Participants must have Python (≥ 3.11), pip, and Jupyter Notebook installed on their laptop.

Required Python libraries:

 - text analysis/NLP: transformers, sentence-transformers, BERTopic, turftopic, setfit
 - LLM inference: openai, huggingface_hub, outlines, pydantic
 - data analysis: scikit-learn, umap, hdbscan
 - others: numpy, scipy, pandas

The instructor will distribute concrete instructions for the Python setup and a comprehensive list of required libraries before the course and assist with any remaining setup problems on the first day of the course.

Organizational Structure of the Course

The course will be organized as a mixture of lectures and exercise sessions. We will switch between lectures and exercises throughout the morning and afternoon sessions of the course. In the lecture sessions, we will focus on explaining core concepts and methods. In the exercise sessions, participants will apply their newly acquired knowledge. Both will be available to answer questions and provide guidance during the entire course.

Day-to-day Schedule and Literature

Note: PDF copies of the recommended readings references below will be made available via Moodle.

Day 1 – 21. September 2026

Neural text embedding

We will begin by reviewing the limitations of count-based bag-of-words document representations and the methodological intuition that motivates embedding-based alternatives. You will learn about classic *static* word embedding methods, like Word2Vec, as well as the *contextualized* embedding techniques that power state-of-the-art transformer-based language models like BERT, GPT, and related models.

In the practical exercises, you will learn how to use transformer models for text embedding with the `transformers` and `sentence-transformers` Python libraries and how to use these embeddings for document retrieval and clustering with libraries like `scikit-learn`, `umap`, `hdbscan`, and `turftopic`. During the exercise time, the lecturer will be available for troubleshooting technical issues participants encountered when trying to set up their Python environment.

Recommended literature

goldberg_primer_2015: Yoav Goldberg (2015). *A Primer on Neural Network Models for Natural Language Processing*. arXiv: [1510.00726\[cs\]](https://arxiv.org/abs/1510.00726)

rodriguez_word_2021: Pedro L. Rodriguez and Arthur Spirling (2021). “Word Embeddings: What works, what doesn’t, and how to tell the difference for applied research”. In: *The Journal of Politics* 84.1, pp. 101–115. DOI: [10.1086/715162](https://doi.org/10.1086/715162)

jurafsky_transformers_2025: Daniel Jurafsky and James H Martin (2025). “Transformers”. In: *Speech and Language Processing: An introduction to speech recognition, computational linguistics and natural language processing*. 3rd edition. URL: <https://web.stanford.edu/~jurafsky/slp3/8.pdf>

alammar_illustrated_2018: Jay Alammar (2018). *The Illustrated Transformer*. Jay Alammar: Visualizing machine learning one concept at a time. URL: <https://jalammar.github.io/illustrated-transformer/>

huggingface_transformers_2026: Hugging Face (2026). *How do Transformers work? · Hugging Face*. URL: <https://huggingface.co/learn/llm-course/en/chapter1/4>

reimers_sentence-bert_2019: Nils Reimers and Iryna Gurevych (2019). “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks”. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint*

Conference on Natural Language Processing (EMNLP-IJCNLP). Association for Computational Linguistics. DOI: [10.18653/v1/D19-1410](https://doi.org/10.18653/v1/D19-1410)

grootendorst_bertopic_2022: Maarten Grootendorst (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. DOI: [10.48550/arXiv.2203.05794](https://doi.org/10.48550/arXiv.2203.05794). arXiv: [2203.05794\[cs\]](https://arxiv.org/abs/2203.05794)

kardos_turftopic_2025: Márton Kardos et al. (2025). “Turftopic: Topic Modelling with Contextual Representations from Sentence Transformers”. In: *Journal of Open Source Software* 10.111, p. 8183. DOI: [10.21105/joss.08183](https://doi.org/10.21105/joss.08183)

Further reading

- *methodological foundations*: Turney and Pantel ([2010](#)), Vaswani et al. ([2023](#))
- *quantitative text analysis introductory and overview articles*: Grimmer and Stewart ([2013](#)), Boumans and Trilling ([2016](#)), Wilkerson and Casas ([2017](#)), Atteveldt and Peng ([2018](#)), Gentzkow, Kelly, and Taddy ([2019](#)), Schoonvelde, Schumacher, and Bakker ([2019](#)), Benoit ([2020](#)), Bonikowski and Nelson ([2022](#)), Grimmer, Roberts, and Stewart ([2022](#)), Ash and Hansen ([2023](#)), Nielbo et al. ([2024](#)), Bachl and Scharkow ([2024](#))
- *food for thought*: DiMaggio ([2015](#)), Baden et al. ([2022](#)),

Day 2 – 22. September 2026

From concepts to measurements with large language model (LLM) prompting

We will begin by discussing how to measure social science concepts in texts. You will learn about central ideas and principles in quantitative content analysis, such as conceptualization, text annotation and coding, reliability and validity, and how they relate to modern computational text analysis methods including text embeddings and LLMs.

You will then learn about “in-context learning,” a way to give an LLM instructions, called prompts, so it can complete text annotation tasks for you. You will apply LLMs to automate text annotation tasks. We will cover practical tips for writing clear prompts and show how different prompt styles work for different types of tasks, like classifying texts into categories, extracting information, scoring, or ranking.

In the hands-on part, you will work with real datasets from social science research. You will practice using Python tools like `openai`, `transformers`, `outlines` and `pydantic` to send prompts to LLMs, receive their responses, and format the results in a structured way. You are encouraged to bring your own ideas or problems to work on during the exercises. However, the lecturer will also provide examples for participants who might not have a concrete project of their own to work on.

Recommended literature

Content Analysis

krippendorff_conceptual_2022: Klaus Krippendorff (2022). “Conceptual Foundation”. In: *Content Analysis: An Introduction to Its Methodology*. Fourth Edition, Chapter 2. DOI:

[10.4135/9781071878781.n3](https://doi.org/10.4135/9781071878781.n3)

neuendorf_defining_2017: Kimberly A. Neuendorf (2017). “Defining Content Analysis”. In: *The Content Analysis Guidebook*. Second, Chapter 1. DOI: [10.4135/9781071802878.n1](https://doi.org/10.4135/9781071802878.n1)

LLM prompting

halterman_codebook_2026: Andrew Halterman and Katherine A. Keith (2026). “Codebook LLMs: Evaluating LLMs as Measurement Tools for Political Science Concepts”. In: *Political Analysis* 34.2, pp. 188–204. DOI: [10.1017/pan.2025.10017](https://doi.org/10.1017/pan.2025.10017)

ornstein_how_2025: Joseph T. Ornstein, Elise N. Blasingame, and Jake S. Truscott (2025). “How to train your stochastic parrot: large language models for political texts”. In: *Political Science Research and Methods* 13.2, pp. 264–281. DOI: [10.1017/psrm.2024.64](https://doi.org/10.1017/psrm.2024.64)

weber_evaluation_2023: Maximilian Weber and Merle Reichardt (2023). *Evaluation is all you need. Prompting Generative Large Language Models for Annotation Tasks in the Social Sciences. A Primer using Open Models*. DOI: [10.48550/arXiv.2401.00284](https://doi.org/10.48550/arXiv.2401.00284). arXiv: [2401.00284](https://arxiv.org/abs/2401.00284) [cs]

tornberg_best_2024: Petter Törnberg (2024). *Best Practices for Text Annotation with Large Language Models*. DOI: [10.48550/arXiv.2402.05129](https://doi.org/10.48550/arXiv.2402.05129). arXiv: [2402.05129](https://arxiv.org/abs/2402.05129) [cs]

Further reading

- *methodological foundations*: Radford et al. ([2019](#)), Brown et al. ([2020](#))
- *food for thought*: Balluff et al. ([2026](#)) McLaren et al. ([2026](#))

Day 3 – 23. September 2026

Morning session: Hands-on large language model prompting and 1-on-1s

On the third day, we will tackle real-world challenges that come up when using LLMs for text annotation in research. We will start by learning how to check if an LLM is doing a good job, using simple metrics like the F1 score to measure reliability. You will see that while newer and bigger models tend to do better, they also often consume more energy, which creates a trade-off between model performance and size arising in LLM annotation. We will discuss the implications of this trade-off to motivate the idea of using LLMs to generate high-quality annotations for fine-tuning smaller models.

We will also talk about how to pick which texts to annotate, and how to find examples that are especially useful or different from each other, building on the topics covered on the first course day, especially

document retrieval and clustering. Moreover, you will learn practical strategies for giving LLMs a few examples, “few-shot prompting,” and for breaking down complex tasks into smaller, manageable steps, “prompt chaining.”

You are encouraged to bring your own research ideas or problems. The lecturer will provide practical materials, and you will have the chance to get individual feedback and support during 1-on-1 sessions throughout the day.

Recommended literature

on scoring/scaling

mens_scaling_2024: Gaël Le Mens and Aina Gallego (2024). *Scaling Political Texts with Large Language Models: Asking a Chatbot Might Be All You Need*. DOI: [10.48550/arXiv.2311.16639](https://doi.org/10.48550/arXiv.2311.16639). arXiv: [2311.16639\[cs\]](https://arxiv.org/abs/2311.16639)

benoit_using_2025: Kenneth Benoit et al. (2025). “Using Large Language Models to Analyze Political Texts through Natural Language Understanding”. URL: https://kenbenoit.net/pdfs/Benoit_etal_2025_AJPS.pdf

licht_scalar_2025: Hauke Licht et al. (2025b). *Measuring Scalar Constructs in Social Science with LLMs*. DOI: [10.48550/arXiv.2509.03116](https://doi.org/10.48550/arXiv.2509.03116). arXiv: [2509.03116\[cs\]](https://arxiv.org/abs/2509.03116)

on classification

ziems_can_2024: Caleb Ziems et al. (2024). “Can Large Language Models Transform Computational Social Science?” In: *Computational Linguistics* 50.1, pp. 237–291. DOI: [10.1162/coli_a_00502](https://doi.org/10.1162/coli_a_00502)

gilardi_chatgpt_2023: Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli (2023). “ChatGPT outperforms crowd workers for text-annotation tasks”. In: *Proceedings of the National Academy of Sciences* 120.30, e2305016120. DOI: [10.1073/pnas.2305016120](https://doi.org/10.1073/pnas.2305016120)

burnham_stance_2025: Michael Burnham (2025). “Stance detection: a practical guide to classifying political beliefs in text”. In: *Political Science Research and Methods* 13.3, pp. 611–628. DOI: [10.1017/psrm.2024.35](https://doi.org/10.1017/psrm.2024.35)

mellon_ais_2024: Jonathan Mellon et al. (2024). “Do AIs know what the most important issue is? Using language models to code open-text social survey responses at scale”. In: *Research & Politics* 11.1, p. 20531680241231468. DOI: [10.1177/20531680241231468](https://doi.org/10.1177/20531680241231468)

on span/information extraction

stuhler_codebooks_2025: Oscar Stuhler, Cat Dang Ton, and Etienne Ollion (2025). “From Codebooks to Promptbooks: Extracting Information from Text with Generative Large Language Models”. In: *Sociological Methods & Research* 54.3, pp. 794–848. DOI: [10.1177/00491241251336794](https://doi.org/10.1177/00491241251336794)

kasner_large_2025: Zdeněk Kasner et al. (2025). *Large Language Models as Span Annotators*. DOI: [10.48550/arXiv.2504.08697](https://doi.org/10.48550/arXiv.2504.08697). arXiv: [2504.08697\[cs\]](https://arxiv.org/abs/2504.08697)

on advanced prompting techniques

lelle_advanced_2025: Travis Lelle (2025). *Advanced Prompt Engineering: Theory, Practice, and Implementation*. URL:

<https://huggingface.co/blog/Travis-ML/advanced-prompt-engineering>

Day 4 – 24. September 2026

Morning session: Transformer encoder fine-tuning

On the fourth day, we will build on what you have learned about LLMs by focusing on how to fine-tune smaller language models using labeled data. We will start by explaining the key differences between large generative models, like GPT or Claude, and smaller encoder-only models, like BERT, DeBERTa, or ModernBERT, and why you might choose one over the other.

You will get an intuitive introduction to how the Transformer architecture works, especially how it creates contextualized word embeddings. We will use a simple example, classifying sentences into policy topics, to show how fine-tuning helps the model separate different categories in its internal map of language.

In the hands-on part, you will practice fine-tuning text classifiers with the `transformers` library in Python. The lecturer will provide real labeled datasets for a variety of tasks, so you can apply your new skills to real research problems, including multilabel classification, scoring, pairwise comparisons, and phrase extraction.

Recommended literature

wankmuller_introduction_2022: Sandra Wankmüller (2022). “Introduction to Neural Transfer Learning With Transformers for Social Science Text Analysis”. In: *Sociological Methods & Research*, p. 004912412211345. doi: [10.1177/00491241221134527](https://doi.org/10.1177/00491241221134527)

do_augmented_2022: Salomé Do, Étienne Ollion, and Rubing Shen (2022). “The Augmented Social Scientist: Using Sequential Transfer Learning to Annotate Millions of Texts with Human-Level Accuracy”. In: *Sociological Methods & Research* 53, pp. 1167–1200. doi: [10.1177/00491241221134526](https://doi.org/10.1177/00491241221134526)

wang_finetuning_2023: Yu Wang (2023). “On Finetuning Large Language Models”. In: *Political Analysis*, pp. 1–5. doi: [10.1017/pan.2023.36](https://doi.org/10.1017/pan.2023.36)

griswold_stay_2026: Max Griswold, Michael W. Robbins, and Michael S. Pollard (2026). “Stay Tuned: Improving Sentiment Analysis and Stance Detection Using Large Language Models”. In: *Political Analysis* 34.3, pp. 392–411. doi: [10.1017/pan.2025.10023](https://doi.org/10.1017/pan.2025.10023)

licht_detecting_2025: Hauke Licht and Ronja Szczepanski (2025). “Detecting Group Mentions in Political Rhetoric A Supervised Learning Approach”. In: *British Journal of Political Science* 55, e119. doi: [10.1017/S0007123424000954](https://doi.org/10.1017/S0007123424000954)

lin_using_2025: Gechun Lin (2025). “Using cross-encoders to measure the similarity of short texts in political science”. In: *American Journal of Political Science*. DOI: [10.1111/ajps.12956](https://doi.org/10.1111/ajps.12956)

timoneda_bert_2025: Joan C. Timoneda and Sebastián Vallejo Vera (2025). “BERT, RoBERTa, or DeBERTa? Comparing Performance Across Transformers Models in Political Science Text”. In: *The Journal of Politics* 87.1, pp. 347–364. DOI: [10.1086/730737](https://doi.org/10.1086/730737)

Further reading

- *methodological foundations*: Vaswani et al. ([2023](#)) Devlin et al. ([2019](#))
- *prompting vs. fine-tuning evidence*: Bucher and Martini ([2024](#)), Brandt et al. ([2026](#)), Licht et al. ([2025a](#))

Day 5 – 25. September 2026

Few-shot learning, further methodological considerations, and 1-on-1s

On the last day of the course, we will dive into advanced techniques for fine-tuning encoder models when you have only a small amount of labeled data, “few-shot learning.” You will learn about two practical approaches: Natural Language Inference (NLI) and Sentence Transformer Fine-Tuning (SetFit).

In the afternoon, we will discuss important issues that come up when using LLMs or classifiers to generate measurements for research. You will learn how errors in classification, like false positives and false negatives, can affect the results of statistical analyses, and why using LLMs can sometimes introduce risks similar to “p-hacking.” We will talk about strategies to recognize and reduce these risks.

Throughout the day, you will have the opportunity to work on your own research projects and get individual feedback and support from the lecturer during 1-on-1 sessions.

Recommended literature

burnham_political_2024: Michael Burnham et al. (2024). *Political DEBATE: Efficient Zero-shot and Few-shot Classifiers for Political Text*. DOI: [10.48550/arXiv.2409.02078](https://doi.org/10.48550/arXiv.2409.02078). arXiv: [2409.02078](https://arxiv.org/abs/2409.02078) [cs]

laurer_less_2022: Moritz Laurer et al. (2022). “Less Annotating, More Classifying – Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI”. in

tunstall_efficient_2022: Lewis Tunstall et al. (2022). *Efficient Few-Shot Learning Without Prompts*. DOI: [10.48550/arXiv.2209.11055](https://doi.org/10.48550/arXiv.2209.11055). arXiv: [2209.11055](https://arxiv.org/abs/2209.11055) [cs]

egami_using_2024: Naoki Egami et al. (2024). *Using Imperfect Surrogates for Downstream*

Inference: Design-based Supervised Learning for Social Science Applications of Large Language Models. DOI: [10.48550/arXiv.2306.04746](https://doi.org/10.48550/arXiv.2306.04746). arXiv: [2306.04746](https://arxiv.org/abs/2306.04746) [stat]

baumann_large_2025: Joachim Baumann et al. (2025). *Large Language Model Hacking: Quantifying the Hidden Risks of Using LLMs for Text Annotation*. DOI: [10.48550/arXiv.2509.08825](https://doi.org/10.48550/arXiv.2509.08825). arXiv: [2509.08825](https://arxiv.org/abs/2509.08825) [cs]

mclaren_magic_2026: Lorcan McLaren et al. (2026). *Magic Words or Methodical Work? Challenging Conventional Wisdom in LLM-Based Political Text Annotation*. DOI: [10.48550/arXiv.2603.26898](https://doi.org/10.48550/arXiv.2603.26898). arXiv: [2603.26898](https://arxiv.org/abs/2603.26898) [cs.CL]

References

- Alammar, Jay (2018). *The Illustrated Transformer*. Jay Alammar: Visualizing machine learning one concept at a time. URL: <https://jalammar.github.io/illustrated-transformer/>.
- Ash, Elliott and Stephen Hansen (2023). “Text Algorithms in Economics”. In: *Annual Review of Economics* 15 (Volume 15, 2023), pp. 659–688. DOI: [10.1146/annurev-economics-082222-074352](https://doi.org/10.1146/annurev-economics-082222-074352).
- Atteveldt, Wouter van and Tai-Quan Peng (2018). “When Communication Meets Computation: Opportunities, Challenges, and Pitfalls in Computational Communication Science”. In: *Communication Methods and Measures* 12.2, pp. 81–92. DOI: [10.1080/19312458.2018.1458084](https://doi.org/10.1080/19312458.2018.1458084).
- Bachl, Marko and Michael Scharkow (2024). *Computational text analysis*.
- Baden, Christian, Christian Pipal, Martijn Schoonvelde, and Mariken A. C. G van der Velden (2022). “Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda”. In: *Communication Methods and Measures* 16.1, pp. 1–18. DOI: [10.1080/19312458.2021.2015574](https://doi.org/10.1080/19312458.2021.2015574).
- Balluff, Paul, Justin Chun-ting Ho, Johannes B. Gruber, Sean Palicki, Alexis Palmer, Luca Rossi, Irina Shklovski, and Chung-hong Chan (2026). “Newer, Larger, Better? A Critique of the Unreflective LLM Adoption in Communication Research”. In: *Political Communication* 0.0, pp. 1–10. DOI: [10.1080/10584609.2026.2618486](https://doi.org/10.1080/10584609.2026.2618486).
- Baumann, Joachim, Paul Röttger, Aleksandra Urman, Albert Wendsjö, Flor Miriam Plaza-del Arco, Johannes B. Gruber, and Dirk Hovy (2025). *Large Language Model Hacking: Quantifying the Hidden Risks of Using LLMs for Text Annotation*. arXiv: [2509.08825](https://arxiv.org/abs/2509.08825) [cs].
- Benoit, Ken (2020). “Text as Data: An Overview”. In: pages 461–497 and pages 461–497. *The SAGE Handbook of Research Methods in Political Science and International Relations*. Ed. by Luigi Curini and Robert Franzese. 2 vols. URL: <https://methods.sagepub.com/hnbk/edvol/research-methods-in-political-science-and-international-relations/chpt/26-text-as-data-overview>.
- Benoit, Kenneth, Scott De Marchi, Conor Laver, Michael Laver, and Jinshuai Ma (2025). “Using Large Language Models to Analyze Political Texts through Natural Language Understanding”. URL: https://kenbenoit.net/pdfs/Benoit_etal_2025_AJPS.pdf.
- Bonikowski, Bart and Laura K. Nelson (2022). “From Ends to Means: The Promise of Computational Text Analysis for Theoretically Driven Sociological Research”. In: *Sociological Methods & Research* 51.4, pp. 1469–1483. DOI: [10.1177/00491241221123088](https://doi.org/10.1177/00491241221123088).
- Boumans, Jelle W. and Damian Trilling (2016). “Taking Stock of the Toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars”. In: *Digital Journalism* 4.1, pp. 8–23. DOI: [10.1080/21670811.2015.1096598](https://doi.org/10.1080/21670811.2015.1096598).

- Brandt, Patrick T., Sultan Alsarra, Vito D’Orazio, Dagmar Heintze, Latifur Khan, Shreyas Meher, Javier Osorio, and Marcus Sianan (2026). “Extractive versus Generative Language Models for Political Conflict Text Classification”. In: *Political Analysis* 34.3, pp. 344–372. DOI: [10.1017/pan.2025.10027](https://doi.org/10.1017/pan.2025.10027).
- Brown, Tom B. et al. (2020). *Language Models are Few-Shot Learners*. arXiv: [2005.14165\[cs\]](https://arxiv.org/abs/2005.14165).
- Bucher, Martin Juan José and Marco Martini (2024). *Fine-Tuned ‘Small’ LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification*. arXiv: [2406.08660\[cs\]](https://arxiv.org/abs/2406.08660).
- Burnham, Michael (2025). “Stance detection: a practical guide to classifying political beliefs in text”. In: *Political Science Research and Methods* 13.3, pp. 611–628. DOI: [10.1017/psrm.2024.35](https://doi.org/10.1017/psrm.2024.35).
- Burnham, Michael, Kayla Kahn, Ryan Yank Wang, and Rachel X. Peng (2024). *Political DEBATE: Efficient Zero-shot and Few-shot Classifiers for Political Text*. arXiv: [2409.02078\[cs\]](https://arxiv.org/abs/2409.02078).
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv: [1810.04805\[cs\]](https://arxiv.org/abs/1810.04805).
- DiMaggio, Paul (2015). “Adapting computational text analysis to social science (and vice versa)”. In: *Big Data & Society* 2.2, p. 2053951715602908. DOI: [10.1177/2053951715602908](https://doi.org/10.1177/2053951715602908).
- Do, Salomé, Étienne Ollion, and Rubing Shen (2022). “The Augmented Social Scientist: Using Sequential Transfer Learning to Annotate Millions of Texts with Human-Level Accuracy”. In: *Sociological Methods & Research* 53, pp. 1167–1200. DOI: [10.1177/00491241221134526](https://doi.org/10.1177/00491241221134526).
- Egami, Naoki, Musashi Hinck, Brandon M. Stewart, and Hanying Wei (2024). *Using Imperfect Surrogates for Downstream Inference: Design-based Supervised Learning for Social Science Applications of Large Language Models*. arXiv: [2306.04746\[stat\]](https://arxiv.org/abs/2306.04746).
- Gentzkow, Matthew, Bryan Kelly, and Matt Taddy (2019). “Text as Data”. In: *Journal of Economic Literature* 57.3, pp. 535–574. DOI: [10.1257/jel.20181020](https://doi.org/10.1257/jel.20181020).
- Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli (2023). “ChatGPT outperforms crowd workers for text-annotation tasks”. In: *Proceedings of the National Academy of Sciences* 120.30, e2305016120. DOI: [10.1073/pnas.2305016120](https://doi.org/10.1073/pnas.2305016120).
- Goldberg, Yoav (2015). *A Primer on Neural Network Models for Natural Language Processing*. arXiv: [1510.00726\[cs\]](https://arxiv.org/abs/1510.00726).
- Grimmer, Justin, Margaret E. Roberts, and Brandon M. Stewart (2022). *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press.
- Grimmer, Justin and Brandon M. Stewart (2013). “Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts”. In: *Political Analysis* 21.3, pp. 267–297. DOI: [10.1093/pan/mps028](https://doi.org/10.1093/pan/mps028).
- Griswold, Max, Michael W. Robbins, and Michael S. Pollard (2026). “Stay Tuned: Improving Sentiment Analysis and Stance Detection Using Large Language Models”. In: *Political Analysis* 34.3, pp. 392–411. DOI: [10.1017/pan.2025.10023](https://doi.org/10.1017/pan.2025.10023).
- Grootendorst, Maarten (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. arXiv: [2203.05794\[cs\]](https://arxiv.org/abs/2203.05794).
- Halterman, Andrew and Katherine A. Keith (2026). “Codebook LLMs: Evaluating LLMs as Measurement Tools for Political Science Concepts”. In: *Political Analysis* 34.2, pp. 188–204. DOI: [10.1017/pan.2025.10017](https://doi.org/10.1017/pan.2025.10017).
- Hugging Face (2026). *How do Transformers work?* · Hugging Face. URL: <https://huggingface.co/learn/llm-course/en/chapter1/4>.
- Jurafsky, Daniel and James H Martin (2025). “Transformers”. In: *Speech and Language Processing: An introduction to speech recognition, computational linguistics and natural language processing*. 3rd edition. URL: <https://web.stanford.edu/~jurafsky/slp3/8.pdf>.
- Kardos, Márton, Kenneth C. Enevoldsen, Jan Kostkan, Ross Deans Kristensen-McLachlan, and Roberta Rocca (2025). “Turftopic: Topic Modelling with Contextual Representations from Sentence Transformers”. In: *Journal of Open Source Software* 10.111, p. 8183. DOI: [10.21105/joss.08183](https://doi.org/10.21105/joss.08183).

- Kasner, Zdeněk, Vilém Zouhar, Patrícia Schmidtová, Ivan Kartáč, Kristýna Onderková, Ondřej Plátek, Dimitra Gkatzia, Saad Mahamood, Ondřej Dušek, and Simone Balloccu (2025). *Large Language Models as Span Annotators*. arXiv: [2504.08697\[cs\]](#).
- Krippendorff, Klaus (2022). “Conceptual Foundation”. In: *Content Analysis: An Introduction to Its Methodology*. Fourth Edition, Chapter 2. doi: [10.4135/9781071878781.n3](#).
- Laurer, Moritz, Wouter van Atteveldt, Andreu Casas, and Kasper Welbers (2022). “Less Annotating, More Classifying – Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI”. In.
- Lelle, Travis (2025). *Advanced Prompt Engineering: Theory, Practice, and Implementation*. URL: <https://huggingface.co/blog/Travis-ML/advanced-prompt-engineering>.
- Licht, Hauke, Tarik Abou-Chadi, Pablo Barberá, and Whitney Hua (2025a). “Measuring and Understanding Parties’ Anti-elite Strategies”. In: *The Journal of Politics*. doi: [10.1086/730711](#).
- Licht, Hauke, Rupak Sarkar, Patrick Y. Wu, Pranav Goel, Niklas Stoehr, Elliott Ash, and Alexander Miserlis Hoyle (2025b). *Measuring Scalar Constructs in Social Science with LLMs*. arXiv: [2509.03116\[cs\]](#).
- Licht, Hauke and Ronja Sczepanski (2025). “Detecting Group Mentions in Political Rhetoric A Supervised Learning Approach”. In: *British Journal of Political Science* 55, e119. doi: [10.1017/S0007123424000954](#).
- Lin, Gechun (2025). “Using cross-encoders to measure the similarity of short texts in political science”. In: *American Journal of Political Science*. doi: [10.1111/ajps.12956](#).
- McLaren, Lorcan, James Cross, Zuzanna Krakowska, Robin Rauner, and Martijn Schoonvelde (2026). *Magic Words or Methodical Work? Challenging Conventional Wisdom in LLM-Based Political Text Annotation*. arXiv: [2603.26898\[cs.CL\]](#).
- Mellon, Jonathan, Jack Bailey, Ralph Scott, James Breckwoldt, Marta Miori, and Phillip Schmedeman (2024). “Do AIs know what the most important issue is? Using language models to code open-text social survey responses at scale”. In: *Research & Politics* 11.1, p. 20531680241231468. doi: [10.1177/20531680241231468](#).
- Mens, Gaël Le and Aina Gallego (2024). *Scaling Political Texts with Large Language Models: Asking a Chatbot Might Be All You Need*. arXiv: [2311.16639\[cs\]](#).
- Neuendorf, Kimberly A. (2017). “Defining Content Analysis”. In: *The Content Analysis Guidebook*. Second, Chapter 1. doi: [10.4135/9781071802878.n1](#).
- Nielbo, Kristoffer L., Folgert Karsdorp, Melvin Wevers, Alie Lassche, Rebekah B. Baglini, Mike Kestemont, and Nina Tahmasebi (2024). “Quantitative text analysis”. In: *Nature Reviews Methods Primers* 4.1, pp. 1–16. doi: [10.1038/s43586-024-00302-w](#).
- Ornstein, Joseph T., Elise N. Blasingame, and Jake S. Truscott (2025). “How to train your stochastic parrot: large language models for political texts”. In: *Political Science Research and Methods* 13.2, pp. 264–281. doi: [10.1017/psrm.2024.64](#).
- Radford, Alec, Jeff Wu, R. Child, D. Luan, Dario Amodei, and I. Sutskever (2019). “Language Models are Unsupervised Multitask Learners”. In.
- Reimers, Nils and Iryna Gurevych (2019). “Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks”. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics. doi: [10.18653/v1/D19-1410](#).
- Rodriguez, Pedro L. and Arthur Spirling (2021). “Word Embeddings: What works, what doesn’t, and how to tell the difference for applied research”. In: *The Journal of Politics* 84.1, pp. 101–115. doi: [10.1086/715162](#).
- Schoonvelde, Martijn, Gijs Schumacher, and Bert N. Bakker (2019). “Friends With Text as Data Benefits: Assessing and Extending the Use of Automated Text Analysis in Political Science and Political Psychology”. In: *Journal of Social and Political Psychology* 7.1, pp. 124–143. doi: [10.5964/jsp.p.v7i1.964](#).

- Stuhler, Oscar, Cat Dang Ton, and Etienne Ollion (2025). “From Codebooks to Promptbooks: Extracting Information from Text with Generative Large Language Models”. In: *Sociological Methods & Research* 54.3, pp. 794–848. doi: [10.1177/00491241251336794](https://doi.org/10.1177/00491241251336794).
- Timoneda, Joan C. and Sebastián Vallejo Vera (2025). “BERT, RoBERTa, or DeBERTa? Comparing Performance Across Transformers Models in Political Science Text”. In: *The Journal of Politics* 87.1, pp. 347–364. doi: [10.1086/730737](https://doi.org/10.1086/730737).
- Tunstall, Lewis, Nils Reimers, Unso Eun Seo Jo, Luke Bates, Daniel Korat, Moshe Wasserblat, and Oren Pereg (2022). *Efficient Few-Shot Learning Without Prompts*. arXiv: [2209.11055](https://arxiv.org/abs/2209.11055) [cs].
- Turney, P. D. and P. Pantel (2010). “From Frequency to Meaning: Vector Space Models of Semantics”. In: *Journal of Artificial Intelligence Research* 37, pp. 141–188. doi: [10.1613/jair.2934](https://doi.org/10.1613/jair.2934).
- Törnberg, Petter (2024). *Best Practices for Text Annotation with Large Language Models*. arXiv: [2402.05129](https://arxiv.org/abs/2402.05129) [cs].
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin (2023). *Attention Is All You Need*. arXiv: [1706.03762](https://arxiv.org/abs/1706.03762) [cs].
- Wang, Yu (2023). “On Finetuning Large Language Models”. In: *Political Analysis*, pp. 1–5. doi: [10.1017/pan.2023.36](https://doi.org/10.1017/pan.2023.36).
- Wankmüller, Sandra (2022). “Introduction to Neural Transfer Learning With Transformers for Social Science Text Analysis”. In: *Sociological Methods & Research*, p. 004912412211345. doi: [10.1177/004912412211345](https://doi.org/10.1177/004912412211345).
- Weber, Maximilian and Merle Reichardt (2023). *Evaluation is all you need. Prompting Generative Large Language Models for Annotation Tasks in the Social Sciences. A Primer using Open Models*. arXiv: [2401.00284](https://arxiv.org/abs/2401.00284) [cs].
- Wilkerson, John and Andreu Casas (2017). “Large-Scale Computerized Text Analysis in Political Science: Opportunities and Challenges”. In: *Annual Review of Political Science* 20.1, pp. 529–544. doi: [10.1146/annurev-polisci-052615-025542](https://doi.org/10.1146/annurev-polisci-052615-025542).
- Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang (2024). “Can Large Language Models Transform Computational Social Science?” In: *Computational Linguistics* 50.1, pp. 237–291. doi: [10.1162/coli_a_00502](https://doi.org/10.1162/coli_a_00502).