Toward a Humanities-Oriented Methodology for Topic Modeling

Johan Cuda

Topic modeling is a polarizing method for text analysis in the digital humanities. This skepticism partly stems from the lack of methodological transparency, evaluation procedures, and reproducibility in many studies (Da 2019). Building on my master's thesis, I propose in this contribution to reexamine the use of topic modeling in the digital humanities, by providing a new methodological framework more aligned with humanities research practices, that emphasizes interpretability and a reproducible evaluation pipeline, so that topic modeling can be better understood and used by researchers in the field.

Topic modeling encompasses a wide range of algorithms designed to uncover latent topics (themes) within a corpus. These topics take the form of a lexicon of predefined size that represents a semantic concept (Abdelrazek et al. 2023). The topic model selected for this study is BERTopic (Grootendorst 2022). Because of its modularity and ease of implementation, BERTopic is a particularly suitable tool for this framework.

When applied rigorously, topic modeling requires knowledge in various domains, such as parameter optimization for model implementation or statistical analysis for result evaluation. These conditions cannot reasonably be expected of all researchers in Digital Humanities (Viola and Spence 2023). Moreover, numerous precautions must be taken in order to avoid producing trivial or misleading results (Da 2019). To address these limitations, I propose a hybrid methodological framework centered around four main axes, namely :

1. Hyperparameter exploration
2. Evaluation of model quality
3. Topic selection
4. Topic interpretation and labeling

Specifically, I first approach parameter selection by going through an iterative process more suited to research in the humanities, instead of the classical automated parameter optimization approach, it being hardly accessible to researchers without advanced expertise in computational methods who nevertheless wish to experiment with topic modeling. Secondly, I select evaluation metrics for the results and the model that are easily implementable and understandable by researchers. To date, there is no commonly accepted metric for evaluating topic modeling (Rüdiger et al. 2022). The selection of evaluation metrics and the justification for their use constitute an important task in any topic modeling study (Rüdiger et al. 2022), as this step enables others to assess the robustness of the model and to partially replicate and verify the results. For our particular framework, I propose to use topic coherence through human evaluation and large language model–assisted scoring (Babalola, Ojokoh, and Boyinbode 2024), topic stability and topic diversity (Dieng, Ruiz, and Blei 2020). Thirdly, I approach topic selection by coupling human inspection of the topics created by the model with the coherence metrics introduced above. This approach allows me both to retain only the topics that achieved a satisfactory coherence score and to leverage the stochastic variability inherent in the model. Lastly, I tackle topic labeling, a process that serves for the analysis of the resulting topics, through a hybrid approach building on our knowledge of the data used, by following the LLM-based topic naming approach proposed by Babalola, Ojokoh, and Boyinbode (2024) and human corrections that better fit our data.

This study seeks to make topic modeling more accessible to humanities researchers by proposing a hybrid evaluation approach that combines quantitative metrics with human interpretation. It highlights the challenges involved in the careful and contextually appropriate application of topic modeling to humanities textual corpora. By formalizing these methodological steps, this contribution aims to provide a transparent and reproducible workflow for topic modeling in digital humanities research, bridging computational rigor and interpretive practices.

References

Abdelrazek, Aly et al. (Feb. 2023). "Topic Modeling Algorithms and Applications: A Survey". In: *Information Systems* 112, p. 102131. ISSN: 0306-4379. DOI: 10.1016/j.is.2022.102131. (Visited on 04/22/2025).

Babalola, Olusola, Bolanle Ojokoh, and Olutayo Boyinbode (Nov. 23, 2024). "Comprehensive Evaluation of LDA, NMF, and BERTopic's Performance on News Headline Topic Modeling". In: *Journal of Computing Theories and Applications* 2, pp. 268–289. DOI: 10.62411/jcta.11635.

Da, Nan Z. (Mar. 2019). "The Computational Case against Computational Literary Studies". In: *Critical Inquiry* 45.3, pp. 601–639. ISSN: 0093-1896. DOI: 10.1086/702594. (Visited on 02/05/2025).

Dieng, Adji B., Francisco J. R. Ruiz, and David M. Blei (2020). "Topic Modeling in Embedding Spaces". In: *Transactions of the Association for Computational Linguistics* 8. Ed. by Mark Johnson, Brian Roark, and Ani Nenkova, pp. 439–453. DOI: 10.1162/tacl_a_00325. URL: https://aclanthology.org/2020.tacl-1.29/ (visited on 03/31/2025).

Grootendorst, Maarten (Mar. 2022). *BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure*. DOI: 10.48550/arXiv.2203.05794. arXiv: 2203.05794. (Visited on 12/06/2024).

Rüdiger, Matthias et al. (Apr. 28, 2022). "Topic Modeling Revisited: New Evidence on Algorithm Performance and Quality Metrics". In: *PLOS ONE* 17.4, e0266325. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0266325. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0266325 (visited on 03/18/2025).

Viola, Lorella and Paul Spence (Nov. 30, 2023). *Multilingual Digital Humanities*. 1st ed. London: Routledge. ISBN: 978-1-003-39369-6. DOI: 10.4324/9781003393696. URL: https://www.taylorfrancis.com/books/9781003393696 (visited on 02/04/2025).