



From Text to Insights Al-Based Topic Modeling for Central Banking Applications

Marlena Pietsch, Diana Sola, Felix Marx, Michaela Vollmer, Babak Zandi 4th IFC Workshop on Data Science in Central Banking Rome, 18-20 February 2025

Motivation Al Services

Which topics can frequently be found in these documents?



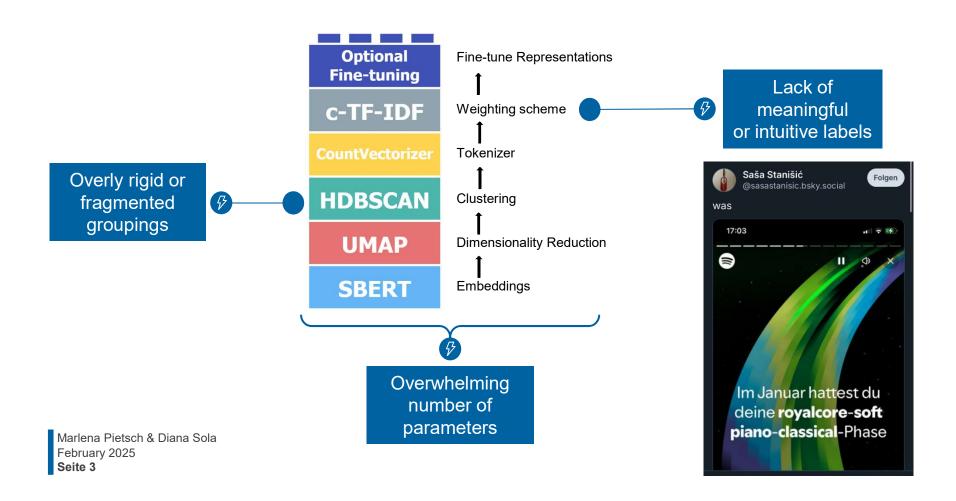
Marlena Pietsch & Diana Sola February 2025 **Seite 2** How can we extract structured information from tables in documents?



How can we quickly summarize key points from dozens of pages?



Early steps Topic modeling with BERTopic



A step beyond traditional methods Topic Modeling with LLMs

Key benefits of using LLMs for topic modeling:



Contextual understanding of the models



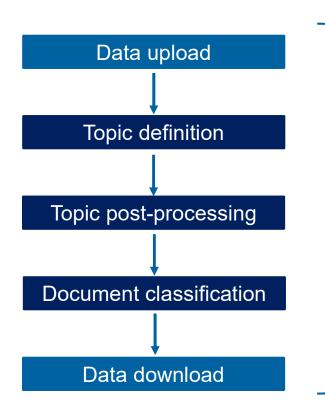
Aligned and human-like topics



Adaptability without hyperparameter tuning

Our Topic Model App bridges LLM capabilities with a user-friendly interface

Topic Model App From Data to Insights in Action



Emphasis on user-centric design for non-technical audiences

- Step-by-step guidance through the topic modelling process
- Option of adapting the Al's work in every step manually



Marlena Pietsch & Diana Sola February 2025 **Seite 5**

The Topic Modeling Process From Data to Insights in Action

Topic definition

- Manually, or
- Uploading predefined topics, or
- Few-shot prompting an LLM to identify topics from data

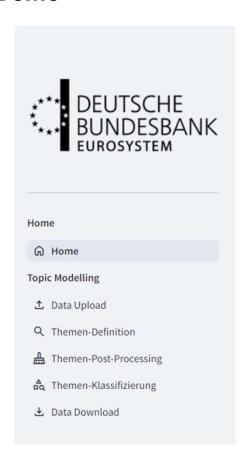
Topic post-processing

- Adding, deleting and altering topics manually, or
- Few-shot prompting an LLM to cluster subsume similar topics and identify a hypernym

Document classification

 Few-shot prompting an LLM to assign one or several topics to the given document

Topic Model App Demo



Marlena Pietsch & Diana Sola February 2025 **Seite 7**

Topic Modelling

Mit dieser App lassen sich Texte nach bekannten Themen klassifizieren. Sollten unbekannte Themen enthalten sein werden diese auch identifiziert.

Die folgenden Schritte (Todo: O, Done: ☑) begleiten Sie durch den Prozess des Topic Modellings:

- O Data Upload
- O Themen-Definition
- O Themen-Post-Processing
- O Themen-Klassifizierung
- O Data Download

:

Lessons learned Challenges & Experiences

Challenges

- Difficulties in evaluating the quality of generated topics
- Scaling to large data inputs
- Aligning model architecture and outputs with domain-specific needs
 - How specific or general do topics need to be to be meaningful?

Experiences

- Importance of interdisciplinary collaboration in app development and understanding of business needs
- Test users preferred pre-written prompts and limited tuning options to the high flexibility in hyperparameter tuning

Outlook Future Directions for the App

Planned Enhancements

- Exploration of automatic and meaningful evaluation methods
- Enabling hierarchical topic modelling for more complex topic structures
- Integrating the app in existing workflows
- Adding statistics and visualization of model outputs





Thank you!

Contact:

Marlena Pietsch

marlena.pietsch@bundesbank.de

Diana Sola

diana.sola@bundesbank.de