

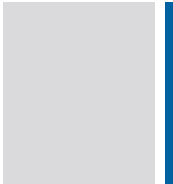
IFC-Bank of Italy Workshop on "Data science in central banking"

18-20 February 2025

From text to insights: AI-based topic modeling for central banking applications¹

Marlena Pietsch, Diana Sola, Felix Marx, Michaela Vollmer and
Babak Zandi,
Deutsche Bundesbank

¹ This contribution was prepared for the workshop. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.



From Text to Insights

AI-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 AI Services

Which topics can frequently
be found in these documents?



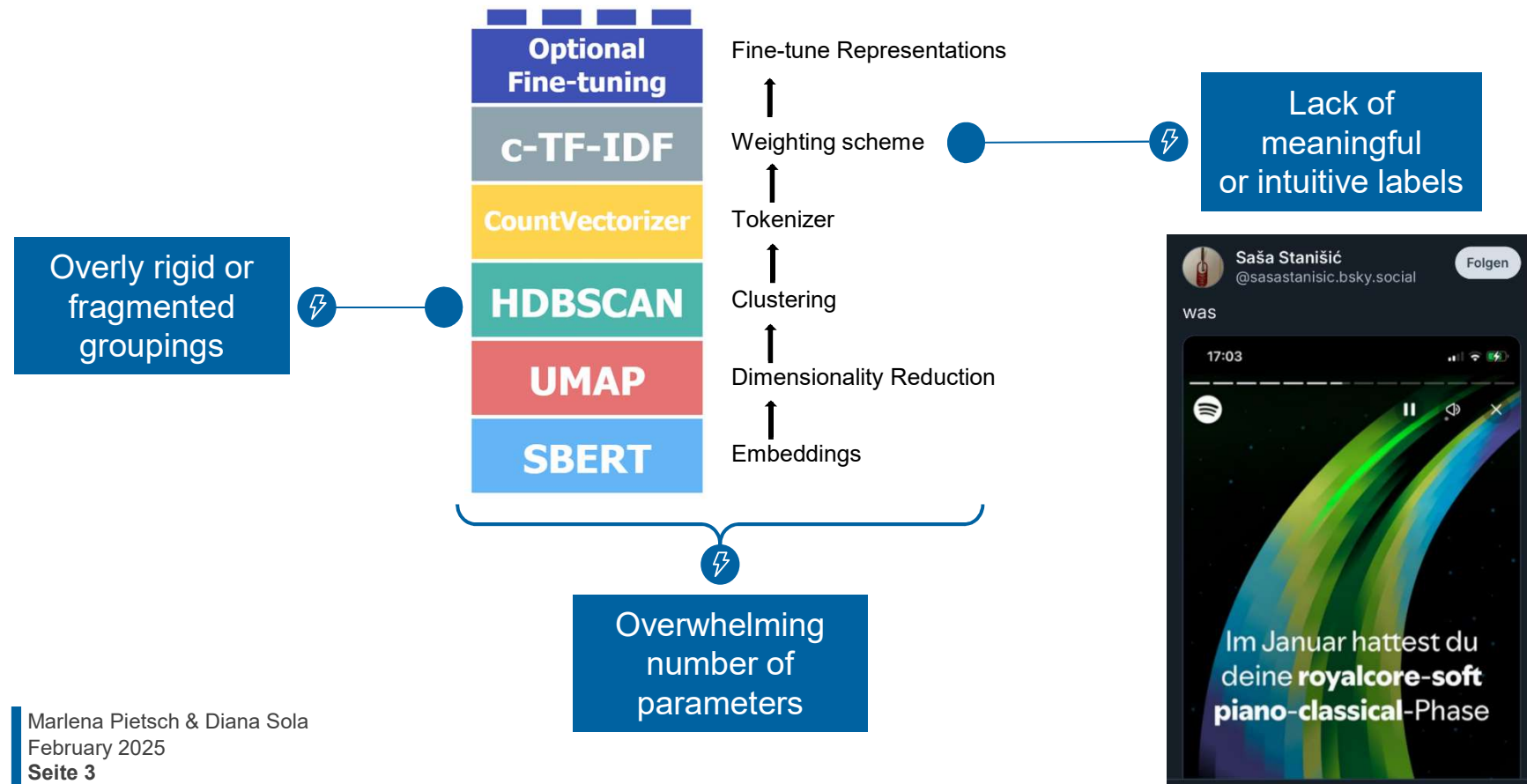
How can we extract
structured information from
tables in documents?



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



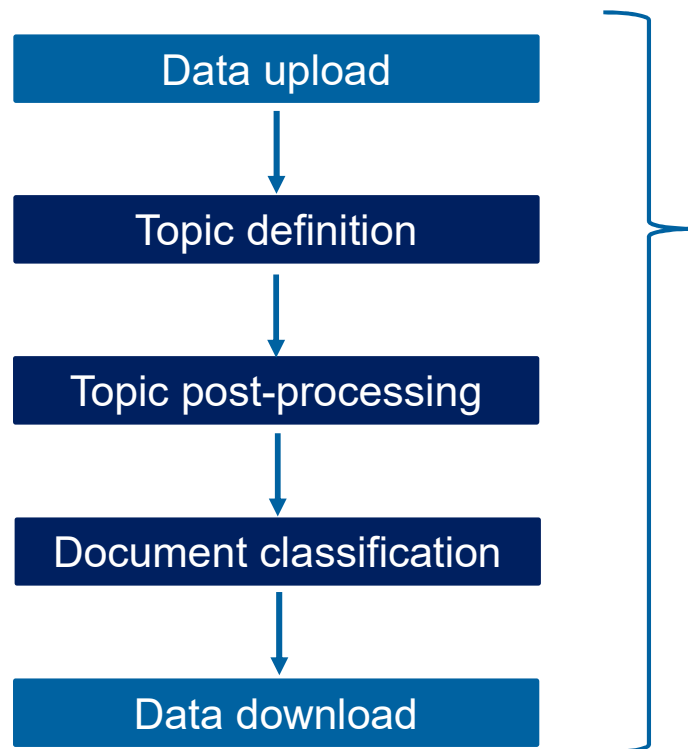
Aligned
and
human-like topics



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 AI's work in every step manually



transparency



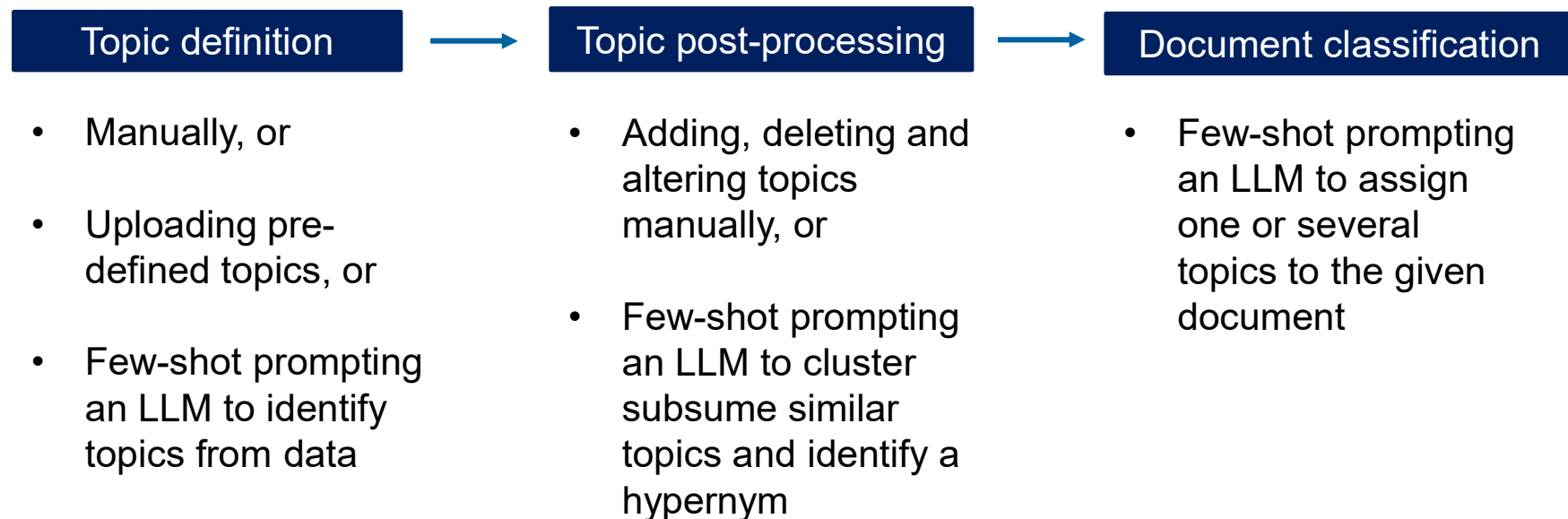
debuggability



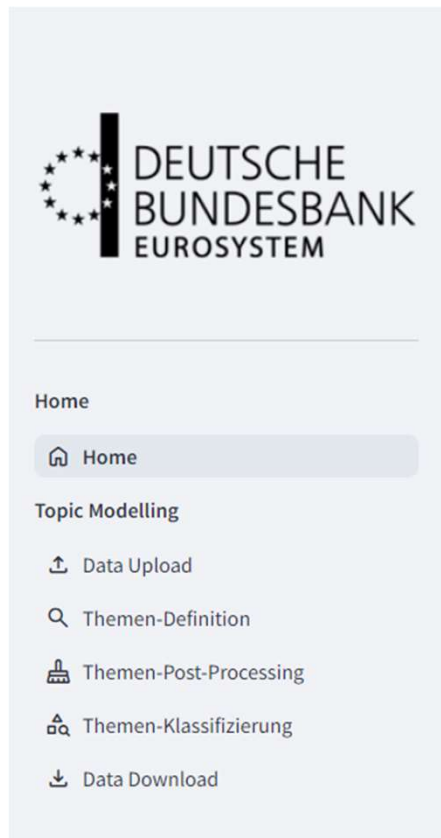
controllability

The Topic Modeling Process

From Data to Insights in Action



Topic Model App Demo



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: ☐, Done: ☒) begleiten Sie durch den Prozess des Topic Modellings:

- ☐ Data Upload
- ☐ Themen-Definition
- ☐ Themen-Post-Processing
- ☐ Themen-Klassifizierung
- ☐ 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

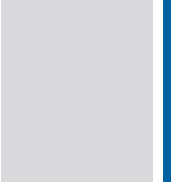
- Importance of interdisciplinary collaboration in app development and understanding of business needs
- Test users preferred pre-written prompts and limited tuning options to the high flexibility in hyperparameter tuning

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