ZAHER ALKAEI

Computational linguist & composer · Arabic discourse analysis · NLP and LLM engineering

Berlin, Germany · zaheralkaei@gmail.com · alkaei.com · discursa.io · github.com/zaheralkaei

PROFILE

Researcher and AI specialist analysing online political discourse in Arabic and German. Founder of Discursa, an Arabic discourse intelligence practice for foundations, NGOs, broadcasters and research institutions. I combine native, in-language reading of Levantine Arabic sources with computational methods (stance and framing analysis, annotation protocols with agreement measurement, transformer and LLM-based classification) and build privacy-preserving, on-premise LLM systems. M.Sc. Cognitive Systems, University of Potsdam (grade 1.4). Also a composer and violinist. Working languages: Arabic, German, English.

EXPERIENCE

Founder & Lead AI Architect · Discursa, Berlin — Apr 2026 – present

- Arabic discourse research across Facebook, YouTube, TikTok and the web: stance and sentiment distribution, framing and narrative mapping, volume timelines, actor and influence mapping, risk assessment.
- Dialect-segmented analysis (Gulf, Levantine, Egyptian, Maghrebi), with diaspora audiences tracked as their own population; written briefs and interactive dashboards for non-technical readers.
- *Damascus alcohol restriction* (Mar 2026): same-day analysis of 880 comments from 534 voices on Syrian government platforms, showing the debate as a proxy contest over the character of the post-transition state.
- *DW Arabic, Jafar Talk* (Mar 2026): reception analysis of 3,868 comments across 100 videos; sentiment, dialect segmentation, seven recurring themes.
- Recent project: Arabic-language discourse on the German far right around the September 2026 state elections (stance classification, inductive framing taxonomy, dialect segmentation).

AI Specialist / IT Consultant · Scitech IT Solutions GmbH — Sep 2025 – Mar 2026

- RAG pipelines with FAISS and open embedding models for an internal FAQ chatbot answering from company documents.
- On-premise LLM infrastructure with Ollama, vLLM and Docker on Ubuntu, with no cloud dependency.
- Multimodal pipelines (document OCR with Qwen-VL, speech-to-text with Voxtral) and a multi-agent system in which a chat agent directs an OCR/vision agent.
- Assistant interfaces with FastAPI, Gradio and Streamlit; workflow automation with n8n; client advice, documentation and user training in German.

Student Assistant · University of Potsdam, Chair of Semantics and IECL programme — Oct 2021 – Sep 2025

Research support in formal and computational semantics (literature review, manuscript review); German-language tutorials in linguistics and semantics; student advising.

Community Interpreter, Arabic–German · PädaLogik GmbH, Berlin — Feb 2018 – Sep 2025

Interpreting in medical and educational settings for families, many of them refugees, under strict confidentiality.

TEACHING

Lecturer · University of Potsdam, Department of Linguistics — Oct 2026 – Feb 2027

Seminar *Classifying Disinformation Narratives and Entities* (in English), co-taught with Arthur Hilbert (TU Berlin). Students replicate published narrative-classification studies with fine-tuned and prompt-based models, then extend them through error analysis, annotation and agreement measurement.

Lecturer · SRH University Berlin — May – Aug 2026

Designed and taught two undergraduate courses, *Global Music History* and *Trans-Traditional Music Pedagogy*, including curriculum, materials and assessment.

PUBLICATIONS

- Alkaei, Z., Stede, M., & Abels, C. M. (2026). Climate and Migration on German Political Party YouTube: A Corpus with Stance, Propaganda, and Divisiveness Annotation. *Proceedings of the 3rd Workshop on NLP Meets Climate Change (ClimateNLP 2026)*, 282–297. Budapest: ACL.
- Alkaei, Z., Hilbert, A., & Memminger, R. (2026). BundesTube: A German Political YouTube Corpus with Zero-Shot Transcript Restoration. *KONVENS 2026 Workshops*, Hamburg. To appear. Code.
- Alkaei, Z. (2024). Syrian Rap: Political, Social, Gender, Artistic, and Linguistic Aspects. Funded research report commissioned by Ettijahat (written 2022). In *A Decade of Arts*, Vol. 2, Mamdouh Adwan Publishing House.
- Alkaei, Z. (2021). Classical Music and Modern Syria: Education, Politics, War, and Migration. Research article.
- Alkaei, Z. (2025). Arabic Music and the Challenges of Artificial Intelligence. *Arab Music Magazine* (in Arabic).

EDUCATION

M.Sc. Cognitive Systems: Language, Learning and Reasoning · University of Potsdam 2023 – Jul 2026
Final grade 1.4 (German scale, 1.0 = best). NLP, machine learning, formal logic, Answer Set Programming. Thesis: *Automated Analysis of Stance, Rhetorical Techniques, and Divisiveness in YouTube Comments on Political Party Videos*, an open-source pipeline from corpus construction and annotation with inter-annotator agreement to fine-tuned transformers, evaluation and explainability.

B.Sc. Linguistics · University of Potsdam 2020 – 2023
Thesis: *Projective Content in Syrian Arabic*, applying projection diagnostics (Tonhauser et al. 2013) to Syrian Arabic triggers.

M.A. Musicology · Humboldt University of Berlin 2017 – 2020
Thesis: *Birdsong as a Memeplex*.

B.A. Music Education · University of Homs, Syria 2009 – 2013

Also: TESOL certificate, Arizona State University.

SELECTED SOFTWARE

Thesis pipeline	Collection, LLM-assisted annotation, BERT/LoRA fine-tuning, stance classification and explainability for German political comment threads.
PrivateGPT Zero	LLM chat that runs open-weight models entirely in the browser (WebLLM, WebGPU); no server, no data leaves the device.
Multimodal analysis	Seven synchronised streams extracted from music videos into structured datasets and an interactive dashboard.
Maqamatic	Generates Arabic maqam melodies as MusicXML from per-maqam transition matrices and sayr rules.

19 open-source tools, most with browser demos: alkaei.com/projects.html

SKILLS

Methods	Multi-platform collection · corpus construction · annotation protocols and inter-annotator agreement · stance, sentiment and framing analysis · narrative and actor mapping · model evaluation (precision, recall, F1)
LLMs	On-premise deployment (Ollama, vLLM; Qwen, Llama) · RAG with FAISS · LangChain · agent orchestration · function calling and structured output · OpenAI and Anthropic APIs
ML / NLP	PyTorch · Hugging Face Transformers · scikit-learn · spaCy · NLTK · BERT fine-tuning · LoRA · sentence-transformers · Answer Set Programming (Clingo)
Engineering	Python · JavaScript · R · SQL · HTML/CSS · FastAPI · Flask · Gradio · Streamlit · Docker · Linux · n8n · Git · LaTeX

TALKS & WORKSHOPS

- Workshop: *Turn Up the Volume – Using Music as a Creative Tool in Your Work with Migrants*. R-Lab Festival, IBZ Berlin, Jun 2026.
- Workshop: *Introduction to Arabic Music History and Fundamentals for Beginners*. Berliner Orchestertreff, Landesmusikakademie Berlin, Sep 2024.
- Talk: *An Evolutionary Teleosemantic Framework of Musical Meaning*. Gesellschaft für Musikforschung, Annual Meeting, Berlin, Sep 2022.

MUSIC

Instruments	Violin, electric violin, synthesizer; Arabic classical, contemporary, cross-over and improvised music
Selected works	*Bayat* for Arabic ensemble (2026) · *Irrecoverable Condition*, string quartet (2024) · *Exile in Degrees* for eight players (2021) · *They Are Still There* for clarinet and string trio (2018)
Performances	*Friends of the Wolf Tone*, Sophiensæle Berlin (2026) · Felix Komoll & Ensemble (2021–2026) · RAN, poetry and music (2024–present) · Babylon Orchestra (2016–2021) · member of AMIBerlin (since 2021)

LANGUAGES & SERVICE

Arabic: native (Levantine; analytical coverage of Gulf, Egyptian and Maghrebi varieties) · German: C1 · English: C1 · French: A2

Mentor, German-Syrian Research Foundation (since 2022).