



Talk2UZH – Project Completion

DSI Infrastructures & Labs

Final report / Summer 2025

Project overview

«ChatGPT», «Gemini», «DeepSeek»: powerful Large Language Model (LLM) tools have rapidly seeped into our daily lives. Nevertheless, their productive use in scientific research faces various hurdles. On the one hand, the majority of these programs are not yet well adapted to subject-specific terminology or local languages (e.g., dialects). On the other hand, privacy and data protection concerns remain salient because the data is stored and processed in cloud servers abroad.

The Talk2UZH Lab was established for harnessing LLM tools for research and teaching in a controlled environment within the University of Zurich (UZH). As a starting use case we implemented OpenAI's Whisper for automated speech-to-text transcription on UZH infrastructure and created Application Programming Interfaces (API) and a survey-system frontend. The Talk2UZH system ensures compliance with data protection regulations: all data processing occurs within the UZH, and sound recordings are fully deleted immediately after transcription. Transcripts are relayed back to end-users through encrypted communication channels.

Embedded in large-scale UZH initiatives including the Population Research Center (PRC), the Digital Society Initiative (DSI), and the Linguistic Research Infrastructure Initiative (LiRI), Talk2UZH has also created a collaborative network to explore the application and usefulness of speech recognition in research studies, which demonstrate the versatility of speech in research questions, ranging from interview transcriptions to survey-based research studies that rely on narratives, for example to analyze lived experiences with a chronic disease.

Talk2UZH will continue and remain at disposal for UZH researchers on a cost-based user fee. Future expansion plans include the development of semi-automated data processing pipelines to offer researchers easy-to-use interfaces for basic descriptive analyses of their speech data.