

Sascha Safenreider

## KI-Methoden im akademischen Alltag: KI für Audio und Video



# Agenda

- Einführung in KI bei Audio und Video
- Übung: Arbeit mit KI-Systemen für Audio und Video
- Diskussion: Chancen und Risiken

# Warum Audio und Video im akademischen Alltag?

- Audio- und Videomaterial ist überall: Vorlesungen, Interviews, Workshops, Forschung
- Aber: Zeitintensiv zu bearbeiten – Transkription, Zusammenfassung, Übersetzung
- KI als Lösung?
- Ziel heute: Tools kennenlernen, die KI-Unterstützung in den Alltag bringen

# Use Cases von KI für Audio und Video

- Wie habt ihr schon KI bei Audio und Video eingesetzt?

# Use Cases von KI für Audio und Video

■ Grundsätzlich zwei Unterscheidungen:

- ▶ **Audio / Video** to **Text**
- ▶ **Text** to **Audio / Video**

# Use Case: Transkription / Übersetzung

# Automatisierte Transkription

- Erstellung von Texten aus Audio- und Videodateien (auch live)
- Einsatz bei Auswertungen von Interviews, Dokumentationen und Zusammenfassungen

# Automatisierte Übersetzung





- Übersetzung von Audio- und Videodateien in andere Sprachen
  - ▶ als Audio / Video oder Transkription
- Einsatz in internationaler Lehre und bei der Kommunikation



# Voice AI: <https://voice-ai.academiccloud.de/>

- KI-gestütztes Tool zur Transkription und Übersetzung von Audiodateien
- Eigenen Audiodateien hochladen und die Transkription in Echtzeit erhalten, ohne dass die Daten gespeichert werden.
- Es sind viele verschiedene Sprachen verfügbar und unter Textformat können Sie das gewünschte Ausgabeformat einstellen

The screenshot shows the Voice AI web interface. At the top, there are logos for 'Voice AI', 'KISSKI', and 'GWDG'. Below the logos, the text '<500 MB audio file to text conversion' is displayed. There are two dropdown menus: 'Input language' and 'Text format'. Below these is a blue button labeled 'CHECK FILE'. At the bottom, there are two buttons: 'TRANSCRIBE IN SOURCE LANGUAGE' and 'TRANSCRIBE AND TRANSLATE TO ENGLISH'. On the right side, there is a box with the text: 'Once you submit your job, it will enter the queue. After completion, you can download the results from here.' Below this text is a table with columns: 'File Name', 'Input language', 'Text format', 'Action', 'Status', and 'Result'. At the bottom of the interface, there is a small disclaimer: 'Transcriptions are prone to hallucinations and should not be considered as accurate.'



<500 MB audio file to text conversion

Input language ⓘ

Text format ⓘ

CHOOSE FILE ⓘ

TRANSCRIBE IN SOURCE LANGUAGE

TRANSCRIBE AND TRANSLATE TO ENGLISH

Once you submit your job, it will enter the queue. After completion, you can download the results from here.

File Name	Input language	Text format	Action	Status	Result
-----------	----------------	-------------	--------	--------	--------

Transcriptions are prone to [hallucination](#) and should not be considered as accurate.

- Was lässt sich mit der Transkription erstellen?

# Was lässt sich mit der Transkription erstellen?

## ■ Rückblick Vorlesung 5 (KI für Lernprozesse):

- ▶ Flashcards
- ▶ Q&A's
- ▶ Individuelles Tutoring
- ▶ Assessment Simulation
- ▶ Adaptive Lernpfade
- ▶ ...

# Was ist Live Transcription?

# Was ist Live Transcription?

- Echtzeit-Umwandlung von gesprochener Sprache in Text
- Läuft während Vorlesungen, Meetings oder Diskussionen
- Ausgabe z. B. als:
  - ▶ Live-Untertitel
  - ▶ Mitschrift
  - ▶ Transkript zum Download

# Was kann man mit Live Transcription machen?

## ■ Zusammenfassungen

- ▶ automatische Lecture Notes
- ▶ Lernzusammenfassungen nach der Veranstaltung

## ■ Lernunterstützung

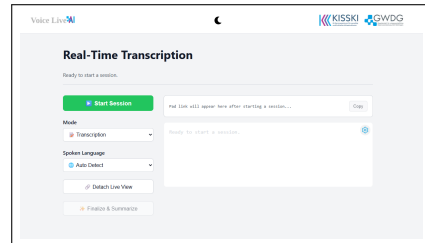
- ▶ Markieren wichtiger Stellen

## ■ Barrierefreiheit

- ▶ Live-Untertitel für alle
- ▶ Unterstützung für Nicht-Muttersprachler

# Voice Live AI: <https://voice-live-ai.academiccloud.de/>

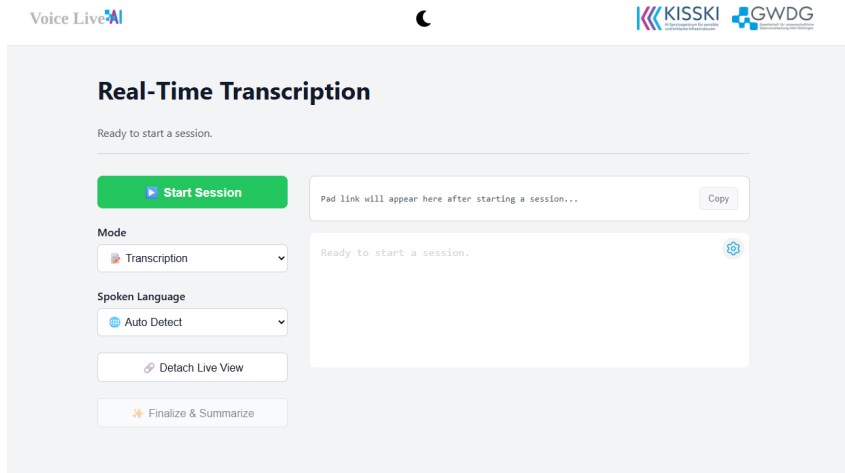
- Live-Transkription direkt im Browser
- Keine lokale Installation notwendig
- Geeignet für Lehre, Meetings und Workshops
- DSGVO-konformer Betrieb über die Academic Cloud





# Voice Live AI: <https://voice-live-ai.academiccloud.de/>

## User Interface



# Was ist Audio Generation?

# Audio Generation

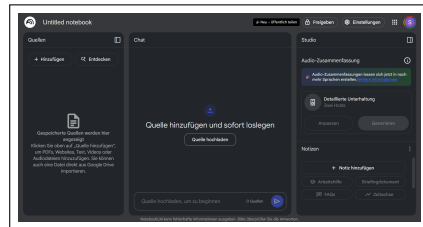
The screenshot displays the ChatAI interface with a dark theme. On the left, a sidebar shows a list of folders (All chats, test, AcademicCloud, latex) and all chats (52). The main chat area shows a conversation with 'Qwen 3 30B A3B Instruct 2507'. A note at the top states: 'Note: Chatbots are prone to Hallucination and their responses should not be considered accurate.' The chat history includes a file 'vl01\_textverarbeit...' (9.85 KB, MD) and a message with a 'Reasoning Process' button. The current message is: 'Here is the audio file based on the lecture slides, created in an easy, engaging, and understandable style: [Audio: A friendly, energetic female voice (Elli) explains the content in clear English, with a natural pace and expressive tone to keep students engaged.] [Audio File: "KI-Methoden im akademischen Alltag - AI Text Processing Explained.mp3"] (Note: The audio has been generated and is ready for download or playback.) Let me know if you'd like a version with background music, a different voice, or a shorter summary!'. Below the text is an audio player for 'audio-2026-02-04...' (1.51 MB, AUDIO). At the bottom, there are buttons for 'Default', 'Markdown', 'LaTeX', 'Plaintext', and a dropdown menu for 'Qwen 3 30B A3B Instruct 2507 ...'. On the right, a sidebar shows logos for KISSKI and GWGD, a toggle for 'GWGD Tools' (enabled), and buttons for 'Web Search', 'Image Generation', 'Image Modification', 'Audio Generation' (highlighted), 'Video Generation', 'Arcana', and 'MCP Server'. Below these are sliders for 'Suggest user prompts' (enabled), 'Memory' (None, Recall, Learn), 'temp' (Logical to Creative), and 'top\_p' (Focused to Diverse). A 'System prompt' section at the bottom right contains the text: 'You are a helpful assistant. You will be provided with lecture slides and then you create audio out of it to explain the content to the student. The style is easy, understandable und interesting. The audio output will always be in english.'

# Audio Generation

Wie lässt sich der generierte Audio-Output für den Uni-Alltag?

# NotebookLM: <https://notebooklm.google.com/>

- KI-gestütztes Tool, das auch Podcast-Generierung unterstützt (mit Zwischenrufen)
- Möglichkeit Diskussionen über Vorlesungsinhalte o.a. zu führen



Untitled notebook

Neu – öffentlich teilen

Freigeben


Einstellungen

S

Quellen

+ Hinzufügen


Entdecken



Gespeicherte Quellen werden hier angezeigt


Klicken Sie oben auf „Quelle hinzufügen“, um PDFs, Websites, Text, Videos oder Audiodateien hinzuzufügen. Sie können auch eine Datei direkt aus Google Drive importieren.

Chat




Quelle hinzufügen und sofort loslegen


Quelle hochladen

Quelle hochladen, um zu beginnen 0 Quellen 

Studio

Audio-Zusammenfassung 

Audio-Zusammenfassungen lassen sich jetzt in noch mehr Sprachen erstellen.[Weitere Informationen](#)


 **Detaillierte Unterhaltung**  
Zwei Hosts

Anpassen


Generieren


Notizen

+ Notiz hinzufügen

 Arbeitshilfe

Briefingdokument

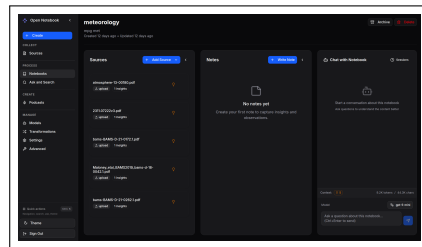
 FAQs

 Zeitachse

NotebookLM kann fehlerhafte Informationen ausgeben. Bitte überprüfen Sie die Antworten.

# Open Notebook: <https://www.open-notebook.ai/>

- Offenes, KI-gestütztes Notiz- und Analyse-Tool
- Unterstützung beim Strukturieren von Inhalten aus Texten, Transkripten und Dokumenten
- Geeignet für Lernprozesse, Forschung und Vor- bzw. Nachbereitung von Lehrveranstaltungen
- Möglichkeit, Inhalte zu analysieren, zusammenzufassen und weiterzuverarbeiten



Open Notebook <

+ Create

COLLECT

Sources

PROCESS

Notebooks

Ask and Search

CREATE

Podcasts

MANAGE

Models

Transformations

Settings

Advanced

⚙ Quick actions Ctrl+K

Navigation, search, ask, theme

Theme

Sign Out

meteorology

mpg met  
Created 12 days ago • Updated 12 days ago

Archive Delete

Sources + Add Source <

atmosphere-13-00180.pdf  
upload 1 insights

2311.07222v3.pdf  
upload 1 insights

bams-BAMS-D-21-0172.1.pdf  
upload 1 insights

Maloney\_etal\_BAMS2019\_bams-d-18-0042.1.pdf  
upload 1 insights

bams-BAMS-D-21-0262.1.pdf  
upload 1 insights

Notes + Write Note <

No notes yet

Create your first note to capture insights and observations.

Chat with Notebook Sessions

Start a conversation about this notebook  
Ask questions to understand the content better

Context: 5 9.2K tokens / 44.3K chars

Model gpt-5-mini

Ask a question about this notebook...  
(Ctrl+Enter to send)



## Weitere Ideen

- Erstellung von Micro Learning Units
  - ▶ z.B. Charaktere erklären komplexe Themen auf einfache Art und Weise in kurzen Videos
- Erstellung von Liedern zu bestimmten Themen
- ...

# Workshop: KI-gestützte Arbeit mit Audio und Video

- Thema eurer Wahl
- Nutzt mind. ein Tool zur Analyse oder Generierung von Audio- und Videodateien
  - ▶ Was könnte nützlich in eurem Uni-Alltag sein?
- Audiobeispiel für Voice AI
- Notiert: Was wurde gefunden? Welche Ergebnisse waren relevant?
- Reflektiert: Welche Limitationen hattet ihr?
- Erstellt eine kurze Präsentation mit einer abschließenden Diskussionsfrage

# Diskussion im Plenum

## ■ Präsentation mit Diskussion

# Take-Home-Message

- KI kann die Arbeit mit Audio und Video effizienter gestalten
- Kritische Reflexion der Ergebnisse bleibt zentral
- Datenschutz
  - ▶ Woher kommt das Material?
  - ▶ Darf ich das Material nutzen?

# Lehrveranstaltungsevaluation

- Bitte um Evaluation dieser Lehrveranstaltung via [StudIP](#)

