

Multimedia-gestützte KI

Entscheidungsunterstützung im Wissensmanagement

Hausarbeit / Bachelorarbeit / Masterarbeit

Unternehmen haben zunehmend multimediale Inhalte (Videos, PDFs, 3D-Modelle, Audio) im Unternehmens-Wissensspeicher. Die Suche und Analyse dieser heterogenen Datenquellen ist ressourcenintensiv und fehleranfällig. KI-Techniken können zwar semantische Metadaten erzeugen, aber die Integration von Multimedia und KI bleibt fragmentiert. In dieser Arbeit soll ein Prototyp entwickelt werden, der Audio-/Video-Inhalte automatisch transkribiert, klassifiziert und in ein semantisches Repository einbettet. Als Basis hierfür können auch Graph-Codes eingesetzt werden (siehe <https://www.mdpi.com/2504-2289/5/3/33>).

Mögliche Forschungsfragen:

- Welches Modell liefert die höchste Transkriptions- und Klassifikationsqualität für Unternehmensvideos
- Inwiefern verbessert sich die Effizienz der Wissensabfrage bei Anwendern, wenn Multimedia-Metadaten eingebunden werden
- Wie lassen sich extrahierte Multimedia-Metadaten in RAG-Systeme integrieren

Hintergrund: Speech-To-Text Modelle, wie Whisper (Open AI) und Video-Captioning-Technologien sind mittlerweile in Forschung und Industrie etabliert. Wissensmanagement-Frameworks nutzen häufig Graph-Datenbanken, jedoch fehlt eine standardisierte Schnittstelle für multimodale Embeddings.

Prof. Dr.-Ing. Stefan Wagenpfeil

Wirtschaftsinformatik | Software-Engineering

E-Mail: s.wagenpfeil@pfh.de

Internet: www.stefan-wagenpfeil.de/pfh

Termine: www.stefan-wagenpfeil.de/termine

PFH Private University of Applied Sciences

Weender Landstraße 3-7

37073 Göttingen