

Universität
Münster

KI-Dienste und –Anwendungen in der Hochschule

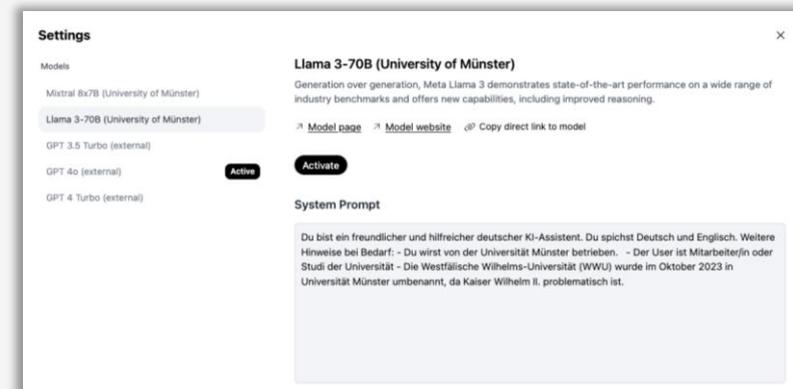
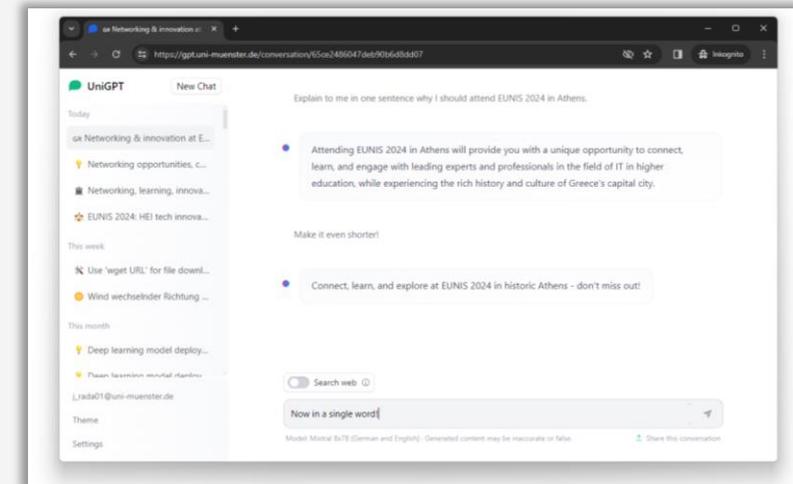


KI als wichtige Werkzeuge für IT- und Informationsversorgung

- Insbesondere im Hinblick auf die Anwendungen von LLMs
- Einige der Schlüsselementen, die Transformationen in der Verwaltung, Lehre und Forschung an Hochschulen vorantreiben können, sind:
 - Chatbot für allgemeine Aufgaben (Chat-Plattform mit unterschiedlichen Modellen)
 - Verbesserte Verarbeitung, Analyse und Abfrage von unstrukturierten Daten (z.B. Texte, Bilder, Audios, Videos)
 - Angepasste Modelle für spezifische Anwendungsfälle / Kontexte

Chatbot für allgemeine Aufgaben: UniGPT

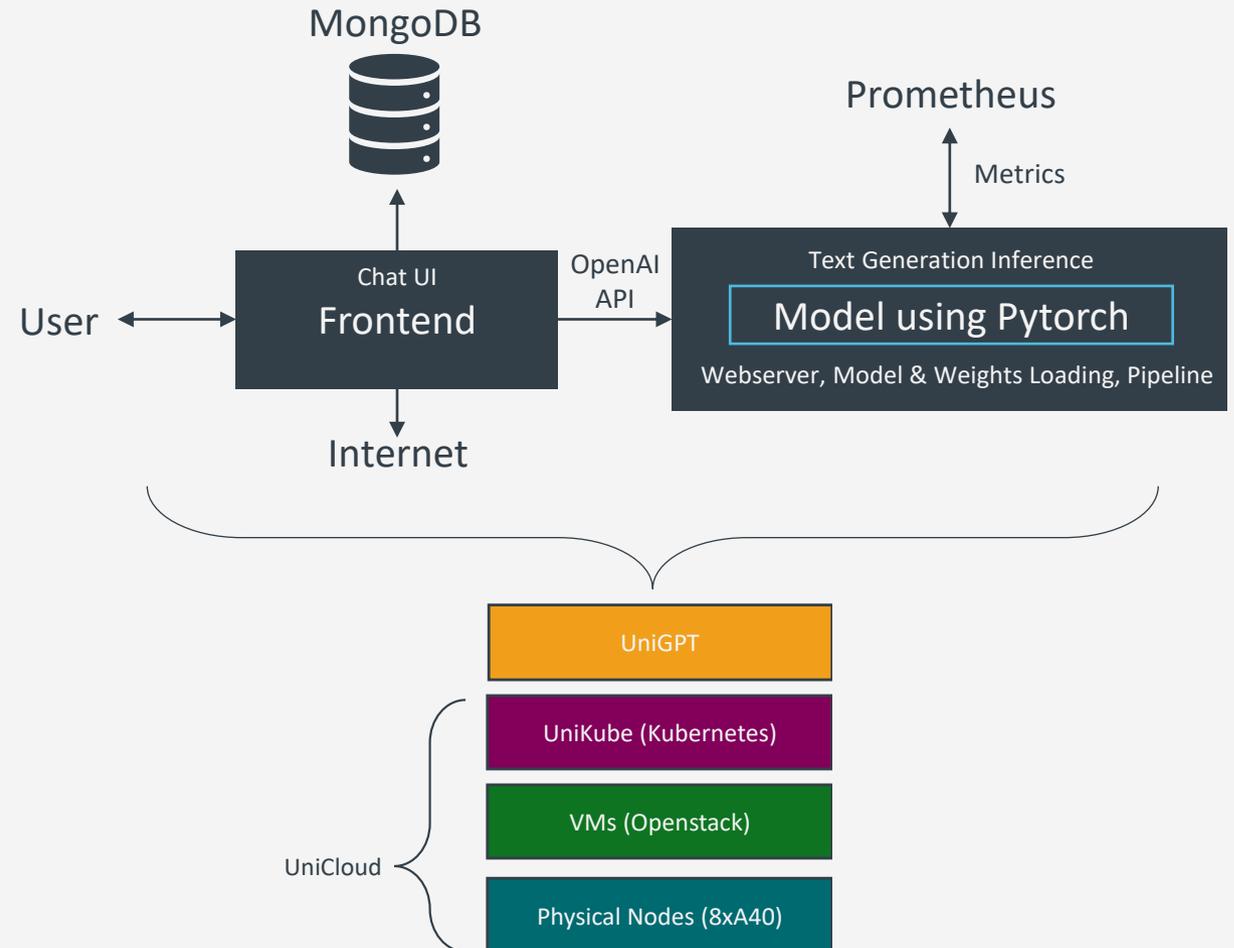
- Motivation:
 - Mehr Kontrolle über die Modelle
 - Weniger Abhängigkeit von proprietären Modellen (z.B. OpenAI)
 - Datenschutzbedenken bei proprietären Modellen
 - Ausgangspunkt für weitere Forschung/Anwendungen
- Unsere Erfahrung:
 - Größeres Interesse als ursprünglich erwartet
 - Lokale Modelle werden überwiegend genutzt, wahrscheinlich aufgrund sensibler Daten
 - Forscher fordern API-Zugriff an



Chatbot für allgemeine Aufgaben: UniGPT

Unterliegende (on-premise) Infrastruktur:

- 8 Nvidia A40, 48 GB VRAM
- Mehrere GPUs hauptsächlich für Verfügbarkeit, nicht Durchsatz
- Läuft auf einem Kubernetes-Cluster in der UniCloud



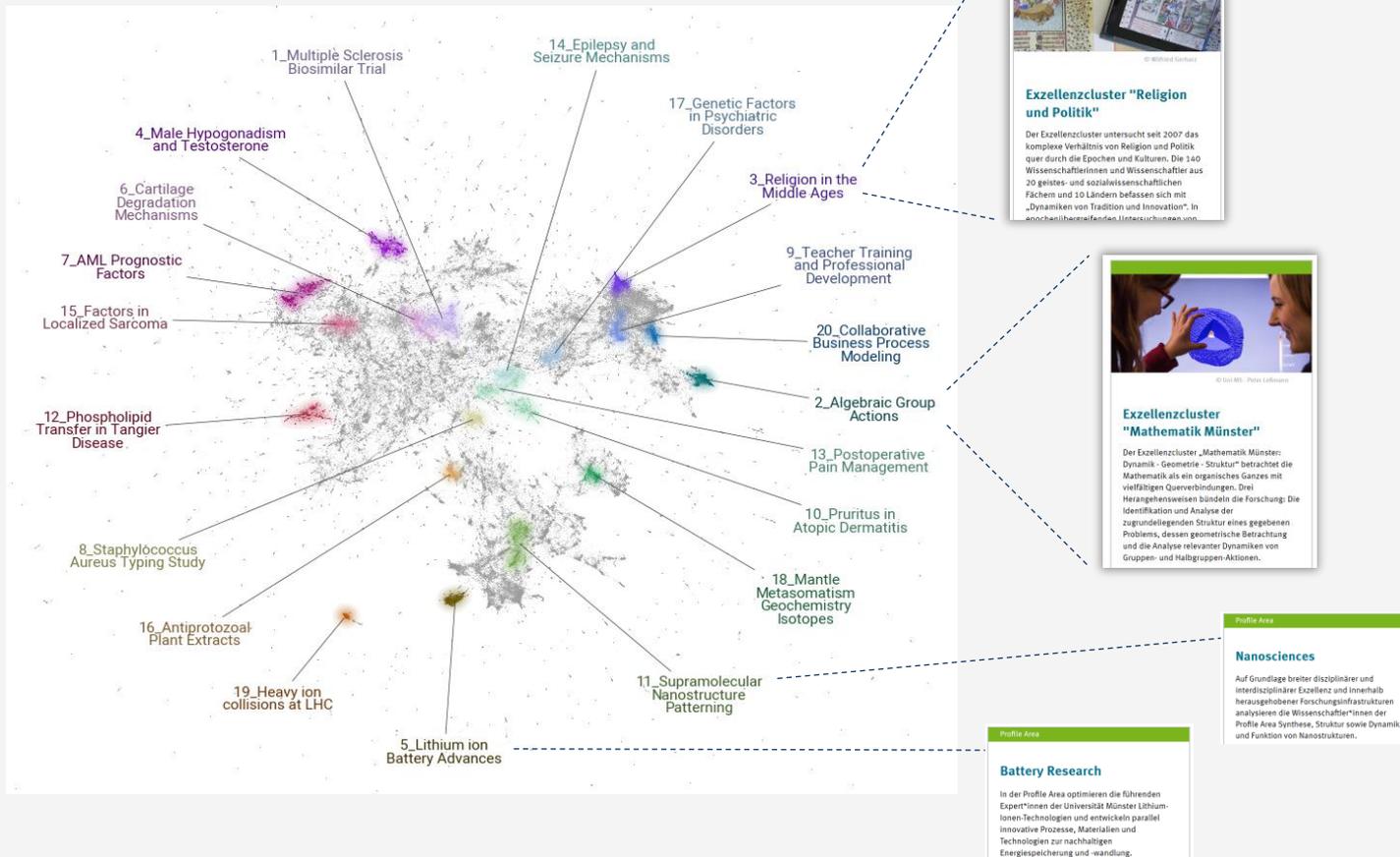
Chatbot für allgemeine Aufgaben: UniGPT

Herausforderungen:

- Nutzungsbedingungen/Datenschutz
- Wartung und Aktualisierung der Modelle und Infrastruktur
 - Neue Modelle erscheinen schnell (sehr schnell): die meisten Modelle in der OpenLLM-Leaderboard sind weniger als einen Monat alt
 - Neue Funktionen erscheinen schnell. Multimodalität und Funktionsaufrufunterstützung kamen in diesem Jahr

Beispiel: Bibliographische Daten

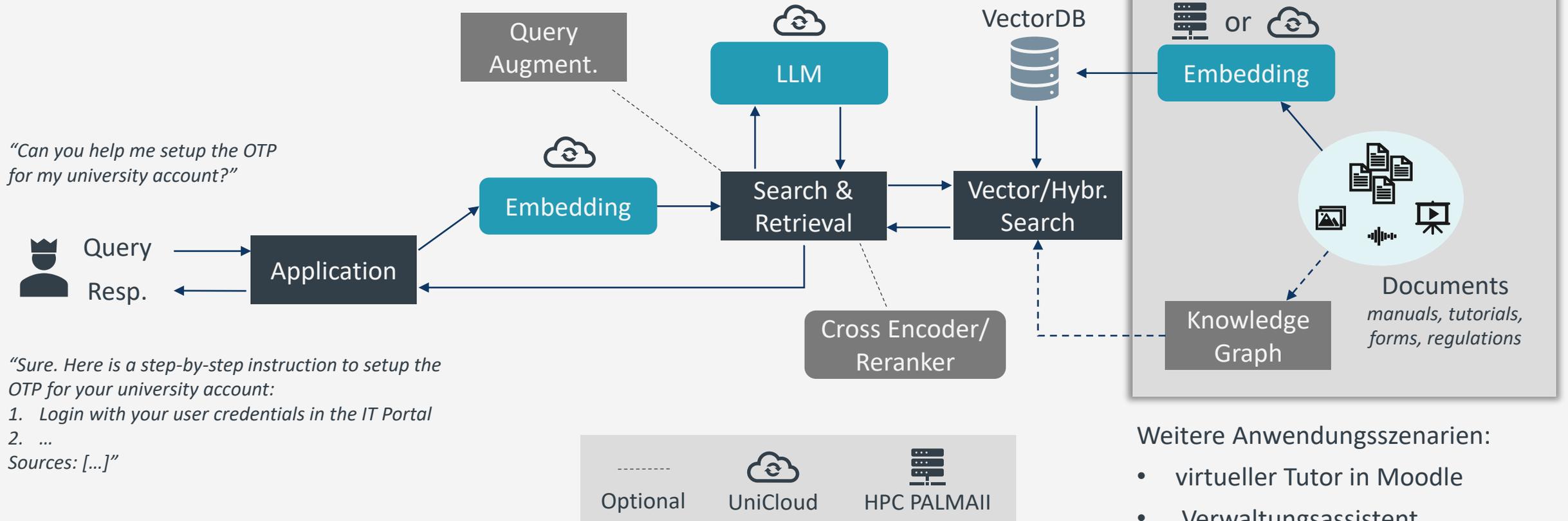
Die 20 größten Cluster an der Universität Münster



- Dokumente: Publikationsdaten
- Embeddings: Titel und Abstract
- Analytik: Clusteranalyse
- LLM: Cluster-Beschriftung
- HPC PALMA-II: Dokumente einbetten
- UniCloud: LLM für Cluster-Beschriftung
- Mögliche nachgelagerte Aufgaben:
 - Forschung zu <...> zusammenfassen
 - Die spezifischen Forschungsschwerpunkte der 10 produktivsten Autoren zu Thema <...> zusammenfassen

Beispiel: Chat-Bot für Kundenkontakt

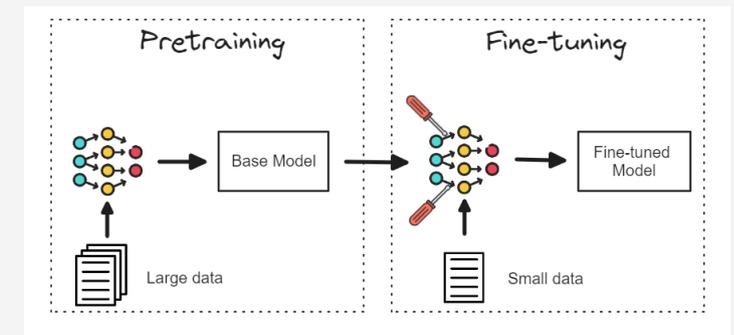
Ein mögliches RAG:



- Weitere Anwendungsszenarien:
- virtueller Tutor in Moodle
 - Verwaltungsassistent

Anpassung der Modelle für spezifische Anwendungsfälle / Kontexte

- Finetuning von LLMs (z. B. Chat-Modellen) und Embeddingsmodellen:
 - Kann das Modell an den Inhalt und die kontextuellen Nuancen der spezifischen Daten/Anwendungsfälle anpassen
- Unterstützt von unserem HPC-Cluster, der mehrere leistungsfähige GPU-Knoten für rechenintensive ML-Workloads enthält
- Herausforderung:
 - Datenverfügbarkeit



Quelle

Abschlussbemerkungen

- KI-basierte Systeme finden viele Anwendungen und transformieren verschiedene Prozesse im Kontext der Hochschule (wir haben in dieser Präsentation nur einige Beispiele behandelt)
- Die Umsetzung erfordert eine geeignete IT-Infrastruktur und Anwendungsdesign:
 - Die Nutzung von On-Premise-Infrastruktur ist entscheidend, um viele Probleme im Zusammenhang mit dem Datenschutz zu umgehen
 - Die On-Premise-Infrastruktur sollte in der Lage sein, KI-basierte Dienste zu skalieren (Inferenz) und verteilte Batch-Verarbeitung zu bewältigen (z. B. Training/Finetuning)
 - Die Anforderungen und Ziele der Anwendungsfälle müssen sich im Anwendungsdesign und dem zugrunde liegenden Technologie Stack/Workflow widerspiegeln

