

AI - RAG

Enterprise RAG Engineering

Ein RAG-Demo ist ein Nachmittag, ein RAG-System mit Berechtigungen und messbarer Qualität ist Engineering. Hier bauen Sie das zweite.

DAUER

3 Tage

GRUPPE

bis 12 Personen

FORMAT

Inhouse · Offen · Im Lab · Deutsch oder Englisch

DAS EXISTIERT DANACH

- ▶ Ein lauffähiger RAG-Prototyp auf der Lab-Umgebung, gebaut vom eigenen Team, mit nachvollziehbarer Architekturentscheidung je Komponente
- ▶ Eine Ingestion-Pipeline mit dokumentierten Chunking-Strategien für Fließtext, Tabellen und strukturierte Dokumente
- ▶ Hybrid-Retrieval aus Vektor- und Volltextsuche mit Reranking, im direkten Messvergleich zur naiven Vektorsuche
- ▶ Ein Evaluationsset mit 30-50 Frage-Antwort-Paaren und ein erster gemessener Qualitätsstand als Baseline
- ▶ Ein implementiertes Berechtigungsmodell mit document-level security und ein Grenzdokument, das ehrlich festhält, was der Prototyp nicht kann

AGENDA

Tag 1

Warum das Tutorial-RAG an Ihren Dokumenten scheitert · Ingestion: aus Dokumenten werden verwertbare Einheiten · Chunking ist eine Entscheidung, kein Default · Embeddings und Index in Betrieb nehmen · Tagesergebnis: die Pipeline läuft

Tag 2

Die Grenzen der reinen Vektorsuche · Hybrid-Retrieval bauen · Reranking: Präzision auf den letzten Metern · Das Evaluationsset entsteht · Tagesergebnis: gemessene Qualität statt Bauchgefühl

Tag 3

Berechtigungen: die Frage, die Projekte stoppt · Document-level security implementieren · Antwortqualität: Grounding, Zitate, Ich-weiß-es-nicht · Grenzdokument und Betriebsübergabe · Tagesergebnis: der Prototyp im Review



IHR TRAINER

Dino Bordonaro

8x Microsoft MVP in Folge (Microsoft Azure · Cloud and Datacenter Management), Microsoft Certified Trainer, fast 30 Jahre IT. Baut souveräne KI-Systeme auf eigener Hardware (Azure Local, Foundry Local, NVIDIA-GPUs) und schult genau darauf. Über 800 Professionals ausgebildet, jedes Konzept im eigenen Enterprise-Lab im TÜV-geprüften Rechenzentrum erprobt.

23.700 €

 pauschal, zzgl. MwSt.

Der Track baut einen Prototyp mit Ihrem Team und vermittelt die Engineering-Entscheidungen dahinter. Er ist keine Produkteinführung und kein Implementierungsprojekt in Ihrer Produktionsumgebung, dafür sprechen wir über ein Projekt oder den Sovereign Assistant.

AI - RAG