

# Das *LLM-Wiki.*

Eine KI, die Unternehmenswissen beantwortet — DSGVO-konform, datensouverän, mit lückenlosem Audit-Trail. Dein Bachelor- oder Masterthema bei dataso.

Regulierte Branchen und der Mittelstand wollen „ChatGPT auf den eigenen Daten“ — dürfen ihre Dokumente aber nicht an OpenAI & Co. geben. Genau diese Lücke schließen wir: ein Wissenssystem, das auf firmeneigenen Daten antwortet, ohne dass die Daten je das Haus verlassen.

*Berater reden über KI. Wir programmieren sie.*

## 01 DAS THEMA

### Ein RAG-System, das sich selbst kuratiert.

Du baust und erforschst ein Retrieval-Augmented-Generation-System auf souveräner Infrastruktur: Es beantwortet Fragen auf Basis firmeneigener Dokumente, hält seine Wissensbasis selbst aktuell (erkennt veraltete und widersprüchliche Inhalte) und protokolliert jede Antwort revisionssicher — nachvollziehbar genug für BaFin, DORA und NIS-2. Open-Source-LLMs, gehostet auf STACKIT, keine Daten an Dritte.

TypeScript

Node.js

RAG / LLMs

PostgreSQL · pgvector

STACKIT

## 02 WORAN DU FORSCHST

### Vier Schwerpunkte, ein System.

**01 Retrieval & Self-Curation**  
Wie findet das System relevantes Wissen — und erkennt, wenn etwas veraltet oder widersprüchlich ist, und schlägt Updates vor?

**02 Auditierbarkeit**  
Append-only-Trails: Jede Antwort bleibt nachvollziehbar — welche Quelle, welcher Stand. Pflicht für BaFin, DORA, NIS-2.

**03 Datensouveränität**  
Open-Source-LLMs auf STACKIT statt Cloud-Blackbox. Der spannende Trade-off: Qualität gegen Souveränität.

**04 Evaluation**  
Antwortqualität und Halluzinationsrate messbar machen — eine belastbare Methodik, kein Bauchgefühl.

Der Umfang lässt sich auf **Bachelor- oder Masterniveau** zuschneiden — ein Schwerpunkt in der Tiefe oder das System in der Breite.

## 03 DAS BRINGST DU MIT

## Neugier auf KI, Substanz im Code.

- Solide Programmierkenntnisse in TypeScript / Node.js
- Echtes Interesse an LLMs, RAG und NLP — über den ChatGPT-Hype hinaus
- Saubere, **eigenständige Arbeitsweise** und Lust, ein Thema zu Ende zu denken
- Immatrikuliert, Abschlussarbeit in Informatik o. Ä. geplant

## 04 DAS BEKOMMST DU

## Ein echtes Produkt, kein Trockenthema.

- + **Ein Thema mit Produktbezug** — wird wirklich gebaut, nicht für die Schublade geschrieben.
- + **Betreuung bei uns** durch einen erfahrenen Entwickler und Architekten, abgestimmt mit deinem Erstprüfer an der Hochschule.
- + **Reale Daten und souveräne Infrastruktur** (STACKIT) statt Spielzeug-Setup.
- + **Apple MacBook Pro als Arbeitsgerät** — auch privat nutzbar.
- + **Vergütung oder Kombi mit einer Werkstudententätigkeit** — reden wir drüber.

## ECKDATEN

ART	Bachelor / Master	BEREICH	Applied AI · RAG
BETREUUNG	dataso + Hochschule	START	nach Absprache

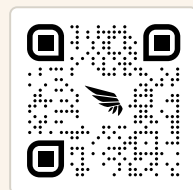
## INTERESSE?

## Mach's zu deiner *Arbeit*.

Schick uns kurz deinen Studiengang, deinen Wunsch-Schwerpunkt und — wenn du magst — einen Link zu Code, an dem du zuletzt gebaut hast. Kein langes Exposé nötig.

[dawid@dataso.de](mailto:dawid@dataso.de)

DAWID SOCHACKI · GRÜNDER & GESCHÄFTSFÜHRER



[dataso.de/jobs](https://dataso.de/jobs)

→ Lieber erst als **Werkstudent** einsteigen? Auch das geht — beides lässt sich gut kombinieren.