

# Wettbewerbsvorteil durch datenschutzkonforme KI- Infrastruktur

*Wie KMUs internes Wissen nutzbar machen – bei voller Datensouveränität*

---

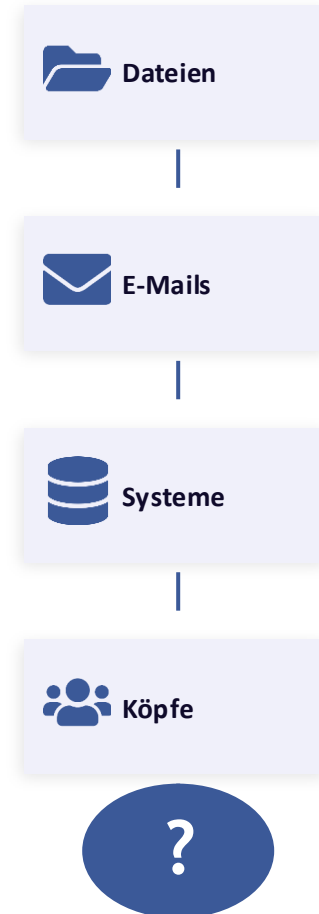
Sarah Gottwald · COO deAI-Labs



# Globale KI-Innovation zwingt Unternehmen zur Anpassung

Europäische Unternehmen sammeln Wissen in Datensilos und im Gedächtnis der Mitarbeiter – dadurch gehen im Alltag Kontext und Produktivität verloren.

Die Herausforderung besteht darin, dieses Wissen für KI produktiv nutzbar zu machen – ohne dabei die Kontrolle zu verlieren.



# Deutsche KMUs haben große Probleme, vorhandenes Wissen produktiv einzusetzen

60%

der Arbeitszeit verbringen  
Wissensarbeiter mit Koordination,  
Statusabfragen und  
Informationssuche.

1 Tag

pro Woche verbringt ein  
Wissensarbeiter damit, nach  
Informationen zu suchen, die  
bereits vorhanden sind.

71%

der Unternehmens-  
anwendungen sind nicht  
integriert. Daten sind  
vorhanden – aber nicht  
vernetzt.

***Wenn Mitarbeiter im Urlaub sind oder in den Ruhestand gehen, sinkt die Produktivität – weil essentielles Wissen nicht mehr vorhanden ist.***

# Der Wunsch: Eine einzige Quelle für das gesamte Unternehmenswissen



**Wissen bleibt, Expertise wird reproduzierbar.**

Neue Mitarbeiter produktiv ab Tag 1.



**Abruf in natürlicher Sprache.**

Nutzer-basierte Antworten in Sekunden.



**Datensilos lösen sich auf.**

Bottlenecks werden sichtbar. Produktivität steigt.

# Warum kann KI Ihre Firmendaten nicht einfach lesen?

## Unstrukturierte Daten

*= für Maschinen nicht direkt lesbar*



**PDFs**

Fließtext, gescannte Bilder, Tabellen in Bildern



**E-Mails**

Anhänge, Threads, verschiedene Formate



**Word/PPT**

Formatierungen, eingebettete Objekte



**Bilder/Scans**

Keine Textinformation ohne OCR

## Strukturierte Daten

*= maschinenlesbar, durchsuchbar*



**Datenbanken**

Zeilen, Spalten, klare Felder



**CSV/JSON**

Standardisierte Austauschformate



**APIs**

Strukturierte Schnittstellen



**Vektoren**

Embeddings – das Zielformat für KI

# Vom Datenchaos zur KI-fähigen Wissensbasis

1

## Extraktion

Text aus Dokumenten  
herauslösen: OCR für Scans,  
Parser für PDFs/Word,  
Konnektoren für E-Mails

2

## Bereinigung

Duplikate entfernen, Formate  
vereinheitlichen

3

## Chunking

Texte in sinnvolle Abschnitte  
aufteilen (500–1000 Tokens)

4

## Embedding

Abschnitte in Vektoren  
umwandeln, in  
Vektordatenbank speichern

### Was ist Extraktion?

Extraktion bedeutet, den lesbaren Text aus verschiedenen Dateiformaten herauszulösen. Ein gescanntes PDF ist für einen Computer nur ein Bild – OCR (optische Zeichenerkennung) wandelt es in durchsuchbaren Text um.

### Was ist ein Embedding?

Ein Embedding wandelt Text in einen Zahlenvektor um. Dadurch kann die KI Ähnlichkeiten erkennen – z.B. dass „Urlaub beantragen“ und „Freistellung nehmen“ das Gleiche meinen.

# Der Paradigmenwechsel: KI zu den Daten bringen

## BISHER: Daten zur KI

Mitarbeiter kopiert Vertrag  
in ChatGPT / Copilot



Daten gehen an US-Server



Kein Einfluss auf Speicherung,  
Verarbeitung, Weitergabe



## Option: KI zu den Daten

Open-Source-KI-Modell wird  
in eigener Infrastruktur in/ EU Hosting installiert



Daten bleiben im Unternehmen  
oder in EU-Rechenzentrum



Volle Kontrolle, DSGVO-konform,  
kein Vendor Lock-in

# Open Source: Die strategische Antwort

Die Lücke zwischen Open Source und proprietären Modellen schließt sich rasant. Fine-Tuning macht offene Modelle oft besser als Generalisten-Modelle – bei voller Kontrolle.

## Proprietäre Modelle

- OpenAI (GPT), Anthropic (Claude), Google (Gemini)
- Höchste Leistung out-of-the-box
- US CLOUD Act: Behörden können Zugriff erzwingen – auch auf EU-Server

## Open-Source-Modelle

- Llama (Meta), Mistral (FR), Kimi K2 (CN), Qwen (CN)
- Frei verfügbar, selbst hostbar in DE/EU
- Kosten: oft geringer
- Leistungslücke schließt sich
- 80% der Y-Combinator-Startups setzen auf open source Modelle

## Fine-Tuned für Unternehmen

- Open-Source-Basis + Training auf Ihren Daten
- Oft besser als Generalisten für Ihren Use Case
- Volle Kontrolle, kein Vendor Lock-in
- Hosting auf deutschen/EU-Servern
- AVV & DSGVO vollständig erfüllbar

# Zwei sichere Wege für Ihr Unternehmen



## On-Premise

- Maximale Datenkontrolle
- Keine Daten verlassen das Haus
- Ideal für Kanzleien, Gesundheitswesen, Verteidigung
- Eigene GPU-Hardware oder Appliance nötig
- Wartung & Updates intern



## Sichere EU-Cloud

- DSGVO-konform (Daten in der EU)
- Kein eigener Server nötig
- Schneller Start in Tagen, nicht Wochen
- Skalierbar nach Bedarf
- Managed Service – Updates inklusive

*z.B. IONOS, Hetzner, Open Telekom Cloud, OVHcloud, STACKIT*

# Ihre Daten bleiben, wo sie sind - die KI dockt an

Ein Konnektor ist wie ein Adapter: Er verbindet Ihre bestehenden Systeme mit der KI, ohne dass Sie etwas ändern oder Daten umziehen müssen. Die KI liest mit – Ihre Systeme laufen weiter wie bisher.



Netzlaufwerke



E-Mails



Datenbanken



Cloud-Speicher



Dokumente



## Ihre KI-Wissensbasis

- Liest automatisch alle Quellen ein
- Strukturiert Inhalte für die KI-Suche
- Aktualisiert sich laufend
- Kein Umbau Ihrer IT nötig

# RAG: Retrieval-Augmented Generation im Detail

RAG = Die KI sucht erst relevante Dokumente, bevor sie antwortet. Keine Daten fließen ins Modelltraining– das Modell bleibt unverändert.



Vorteil: Kein Fine-Tuning nötig, Daten bleiben in der Vektordatenbank, Antworten immer mit Quellenangabe nachvollziehbar.

# Praxis: Internes Wissensmanagement

## Szenario

Ein neuer Mitarbeiter sucht die Reisekostenrichtlinie. Statt 30 Min. im Intranet zu suchen, fragt er die interne KI: „Wie reiche ich Reisekosten ein?“ – Antwort in Sekunden, mit Quellenangabe auf das richtige Dokument.



**70-90 %**

**Zeitersparnis**

bei Informationssuche



**Mit Quelle**

**Genauigkeit**

immer nachprüfbar



**3× schneller**

**Onboarding**

produktiv ab Tag 1

# Das unsichtbare Gold: Implizites Expertenwissen

Implizites Wissen = Erfahrungswissen, das in den Köpfen Ihrer Experten steckt. Es ist nicht dokumentiert, schwer zu verbalisieren – und geht verloren, wenn Mitarbeitende das Unternehmen verlassen.

## Das Problem

- Produktionsleiter weiß, welche Maschine bei Luftfeuchtigkeit über 70% Probleme macht
- Vertriebsleiterin kennt die Eigenheiten jedes Großkunden aus 20 Jahren Erfahrung
- IT-Admin ist der Einzige, der das Legacy-System versteht
- **Dieses Wissen steht in keinem Handbuch**

## Die Lösung: Wissen digitalisieren

- Strukturierte Wissensinterviews mit Experten führen
- Entscheidungslogik sichtbar machen: „Wenn X und Y, dann Z“
- Erfahrungswissen in die KI-Wissensbasis einpflegen
- **Digitaler Wissens-Zwilling des Experten entsteht**

# Praxis: Prozessautomatisierung



## Dokumenten- analyse

Verträge prüfen, Klauseln  
extrahieren,  
Zusammenfassungen erstellen



## E-Mail- Verarbeitung

Anfragen automatisch  
klassifizieren und Antworten  
vorformulieren



## Report- Generierung

Berichte aus verschiedenen  
Quellen automatisch  
zusammenstellen



## Workflow- Automation

Wiederkehrende Aufgaben  
identifizieren und  
automatisieren

# Der Business Case für KMUs

**8-15 h**

Zeitersparnis pro  
Mitarbeiter / Woche

**< 4 Wo.**

Bis zum  
ersten Ergebnis

**100 %**

Kontrolle über  
Ihre Daten

## Was bringt es konkret?

- Schnellere Einarbeitung neuer Mitarbeiter
- Weniger Zeitaufwand für Dokumentensuche und Routineaufgaben
- Bessere Entscheidungen durch schnellen Zugriff auf internes Wissen
- Wettbewerbsvorteil durch Datenschutz-Compliance als Verkaufsargument

# Zusammenfassung



80% des Firmenwissens steckt in unstrukturierten Daten – das ist Ihr ungenutztes Potenzial



KI-Nutzung ohne Datenkontrolle ist riskant – entscheidend ist, wo die Daten verarbeitet werden



Lösung: KI-Modelle zu den Daten bringen – On-Premise oder sichere EU-Cloud



RAG macht internes Wissen sofort nutzbar – ohne Daten ins Modell zu geben



EU AI Act & DSGVO sind kein Hindernis, sondern ein Wettbewerbsvorteil

# Vielen Dank!

*Fragen & Diskussion*

---

## Sarah Gottwald

COO deAI-Labs



[www.mira-os.com/enterprise](http://www.mira-os.com/enterprise)



[sarah.gottwald@deai-labs.com](mailto:sarah.gottwald@deai-labs.com)



[linkedin.com/in/sarah-gottwald](https://linkedin.com/in/sarah-gottwald)



SCAN ME

