



Lokales Assistenzsystem im Personalwesen Möglichkeiten und Grenzen



Virginia Giesemann und Claudia Graf-Pfohl

25 | 08 | 2026

Im Gegensatz zu bekannten Sprachmodellen wie ChatGPT oder Mistral verarbeitet ein RAG-System (Retrieval-Augmented Generation) nur vorab festgelegte Informationen. Unter Retrieval versteht man das gezielte Auffinden passender Textstellen in einer bereitgestellten Dokumentensammlung – etwa Betriebsvereinbarungen, Handbüchern oder Onboarding-Unterlagen. Diese Fundstellen werden dem Sprachmodell als Kontext übergeben, welches daraus eine verständliche Antwort formuliert (Generation) und die Quelle mitliefert. Antworten bleiben so nachvollziehbar. Wird das System lokal betrieben, bleiben die Daten im Unternehmen.

Genau darin liegt der Nutzen für das Personalwesen, wo hochsensible personenbezogene Daten DSGVO-konform verarbeitet werden müssen. Im Rahmen des Projekts Perspektive Arbeit Lausitz (PAL) wurde dazu die lokale Assistentin PAULA als Low-Code-Lösung auf Basis geprüfter Open-

Source-Anwendungen entwickelt und erprobt.

Dieses Fact Sheet zeigt, welche Entscheidungen bei der Einführung abzuwägen sind, welche Grenzen ein lokales Assistenzsystem hat und wie eine betriebsspezifische Zielarchitektur gefunden werden kann – wirtschaftlich, technisch, organisatorisch und im Rahmen der KI-Verordnung (EU) 2024/1689.

Einsatz im Personalwesen

Eine personalverantwortliche Person bereitet ein Entwicklungsgespräch vor. Bisher hieß das: Personalakte herausuchen, die letzte Zielvereinbarung finden, Schulungsnachweise abgleichen, Fristen im Personalhandbuch nachlesen. Ein Assistenzsystem wie PAULA soll diesen Weg verkürzen, indem Fragen in normaler Sprache gestellt werden und die Antwort mit Quellenangabe zurückkommt. Im Test funktionierte das dort am besten, wo eine Person eindeutig benannt und ein konkretes

Dokument abgefragt wurde, etwa: „Welche Abteilung, Stellenbezeichnung und Führungskraft sind für eine bestimmte Mitarbeitenden-ID hinterlegt?“

Folgende Daten können hinterlegt werden:

- Personalhandbücher
- Betriebsvereinbarungen
- HR-Richtlinien und Prozessbeschreibungen
- Vorlagen und Formulare
- Rollen- und Kompetenzmodelle
- FAQs der Personalabteilung
- Konsolidierte Personalakten
- Entwicklungspläne und Zielvereinbarungen
- Schulungsnachweise
- Arbeitszeugnisse

Pro Person im Unternehmen sollte genau eine führende Personalakte bestehen, verknüpft über eine eindeutige Mitarbeitenden-ID. Bewerbungsunterlagen werden getrennt geführt. Der Zugriff auf personenbezogene Inhalte bleibt auf autorisierte HR-Nutzende beschränkt, die Verarbeitung erfolgt lokal, Zugriffe werden protokolliert.

Welche Quelle gilt bei Widersprüchen?

Es empfiehlt sich, die Gültigkeit verschiedener Prozessdokumente vorab festzulegen und zu definieren, um Widersprüche zu vermeiden, die die

Antwortqualität beeinträchtigen.

Beispielsweise:

1. Systemprompt (Verhaltensregeln)
2. Datenschutz- und Berechtigungskonzept
3. Betriebsvereinbarung
4. Personalhandbuch
5. HR-Prozessbeschreibung / SOP
6. HR-Richtlinie / Policy
7. FAQ der Personalabteilung
8. Prompting-FAQ
9. Personenbezogene Einzeldokumente (Personalakten, PE-Dokumente)



Foto: Unternehmensvertretende diskutieren die Einsatzfälle

Funktionsweise RAG-System

Der folgende Ablauf zeigt vereinfacht, wie ein lokales RAG-System eine Frage aus dem Personalbereich beantwortet. Die Frage wird zunächst nicht an das Sprachmodell weitergereicht, sondern für eine Suche in der eigenen Wissensbasis genutzt. Erst die gefundenen Textstellen bilden die Grundlage der Antwort.



Abbildung 1: Ablauf der Funktionsweise eines HR-RAG

Für die Praxis bedeutet das zweierlei: Das System antwortet nur zu Inhalten, die im Unternehmen hinterlegt sind und es benennt die Dokumente, aus denen die Antwort stammt (siehe Abb. 1). Die Qualität der Antwort hängt damit unmittelbar von der Qualität und Struktur der eingespeisten Dokumente ab, nicht allein vom eingesetzten Sprachmodell.

RAG sinnvoll einsetzen in HR

Die Kehrseite frei verfügbarer Sprachmodelle ist die Verarbeitung der Daten, die häufig schwer nachvollziehbar ist. Damit steht die Adaption in Bereichen wie dem Personalwesen vor einer besonderen Herausforderung. Ein RAG-System löst die

Frage der Nachvollziehbarkeit über die Quellenangabe. Wo die Daten verarbeitet werden, entscheidet sich dagegen an der Betriebsart.

Bei einer On-Premise-Lösung auf Low-Code-Basis verbleiben die Daten im Unternehmen, und die Umsetzung gelingt ohne Programmierkenntnisse. Dem stehen Aufwände für Hardware und laufende Pflege gegenüber. Eine Eigenentwicklung passt exakt zu den eigenen Prozessen, bindet aber dauerhaft IT-Fachkräfte. Eine Cloud-Lösung ist sofort verfügbar, allerdings werden die Daten extern verarbeitet. Für personenbezogene HR-Daten ist sie deshalb in der Regel keine Option.

Welcher Weg trägt, entscheidet sich weniger an der Technik als an vier betrieblichen Fragen:

- **Wie groß ist die Datengrundlage?**
Wenige gepflegte Dokumente liefern bessere Antworten als ein großer, ungeordneter Bestand.
- **Ist sie bereits konsistent?**
Uneinheitliche Formate und veraltete Stände verursachen den eigentlichen Aufwand, nicht die Installation.
- **Wer soll Zugriff haben?**
Neueinsteigende, Beschäftigte und Personalverantwortliche benötigen unterschiedliche Wissensbestände. Abbilden lässt sich das über getrennte Workspaces.

▪ Wie wird mit Empfehlungen umgegangen?

Auskünfte sind Arbeitshilfe, nicht Entscheidung. Die menschliche Kontrolle bleibt verbindlich.

Grenzen des lokalen Modells

Ein lokal betriebenes Assistenzsystem ist nicht die kleinere Version einer Cloud-Lösung, sondern eine Lösung mit eigenem Zuschnitt. Die verfügbare Hardware begrenzt, wie leistungsfähig das eingesetzte Sprachmodell sein kann. Antworten fallen kürzer und weniger flüssig aus als bei großen kommerziellen Diensten und die Verarbeitung dauert länger.

Die deutlichste Grenze liegt jedoch nicht im Modell, sondern in den Daten. In unserer Erprobung scheiterten Antworten meist daran, dass die passende Textstelle nicht

gefunden oder mit einer ähnlich aufgebauten verwechselt wurde (siehe Abbildung 2 unten). Personenbezogene Einträge mit gleicher Struktur wurden vermischt, Trefferlisten blieben unvollständig, und fachlich benachbarte Kategorien wurden nicht sauber getrennt. Die durchschnittliche Antwortqualität lag im Test deutlich unter dem, was für einen produktiven Einsatz nötig wäre.

Daraus folgt: **Ein lokales Assistenzsystem ersetzt weder die Personalakte noch die fachliche Prüfung.** Wer es einführen möchte, sollte den Aufwand nicht bei der Software veranschlagen, sondern bei der Pflege und Strukturierung der eigenen Dokumente. Diese Arbeit bleibt dauerhaft bestehen und gehört in die Hand einer verantwortlichen Fachkraft. Es verkürzt den Weg zur richtigen Quelle, trifft aber keine Entscheidungen.

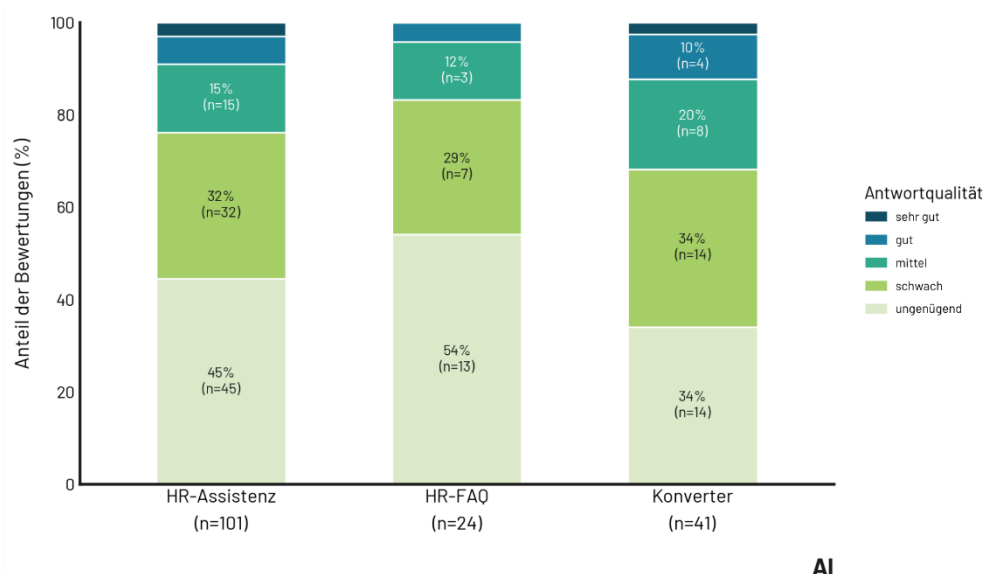


Abbildung 2: Stresstest des HR-RAG für drei Zielgruppen

Einführungs- und Umsetzungsprozess

Ob ein Assistenzsystem im Personalwesen Sinn macht, entscheidet sich nicht an der Software, sondern an der Passung von Technik, Organisation und Qualifizierung. Für die betriebsspezifische Zielarchitektur hat sich folgende Reihenfolge bewährt:

1. Anwendungsfall eingrenzen:

Mit einem klar umrissenen Bereich beginnen, etwa Onboarding, statt mit dem gesamten Dokumentenbestand.

2. Datengrundlage sichten und aufbereiten:

Bestand prüfen, Dubletten und veraltete Stände entfernen, einheitliche Formate und eine eindeutige Kennzeichnung je Person festlegen. Hier entsteht der eigentliche Aufwand des Projekts.

3. Betriebsart und Zielarchitektur wählen:

On-Premise-Lösung, Eigenentwicklung oder Cloud anhand von Datenschutzerfordernungen, vorhandenen IT-Ressourcen und Budget abwägen.

4. Zugriffsrechte und Quellenhierarchie festlegen:

Klären, welche Zielgruppe welche Inhalte sieht und welche Quelle bei Widersprüchen gilt.

5. Rechtliche Anforderungen prüfen:

Der Einsatz von Assistenzsystemen im Personalbereich ist zulässig, aber reglementiert. Zu prüfen sind insbesondere Diskriminierungsrisiken und die Verarbeitung personenbezogener Daten; die menschliche Kontrolle bleibt verbindlich (KI-Verordnung (EU) 2024/1689).

6. Beteiligte einbeziehen und qualifizieren:

Interessenvertretung und Datenschutzbeauftragte früh einbinden. Beschäftigte benötigen tätigkeitsbezogene Kompetenz im Umgang mit dem System. Angrenzende Unternehmensbereiche sowie Außenstehende, etwa Bewerbende, sind verständlich über den Einsatz zu informieren; diese Hinweise sind mit den bestehenden DSGVO-Informationen abzustimmen.

7. Pilotieren, messen und nachsteuern:

Mit wenigen Dokumenten und festen Testfragen starten, die Antwortqualität dokumentieren und die Datenstruktur schrittweise nachschärfen.

Vor der Einführung empfiehlt sich eine Investitionsrechnung, die neben Hardware- und Lizenzkosten vor allem den dauerhaften Aufwand für Datenpflege berücksichtigt.

Unsere Expertise kann Sie in diesem Prozess unterstützen (siehe Abb. 3).

Zielsetzung und Vorgehensweise

Sprachmodelle und mit ihnen die RAG-Technologie werden mittlerweile in unterschiedlichen Arbeitsbereichen adaptiert, um Informationsabfragen verlässlich zu gestalten (Wahl, 2026). Das Projekt PerspektiveArbeit Lausitz (PAL) untersucht wie digitale Assistenzsysteme in KMU die Personalarbeit unterstützen können. Dafür wurde PAULA entwickelt, eine lokale Assistentin für das Personalwesen. Ziel war eine kostenfreie, lokale und datenschutzkonforme Lösung, die sich ohne umfangreiche Programmierkenntnisse umsetzen lässt.

die den Betrieb unterschiedlicher Sprachmodelle lokal ermöglichen, sowie die Abstimmung der Hardware auf deren Anforderungen. Die Erprobung erfolgte in einer Testumgebung anhand eines synthetischen Datensatzes mit einem fiktiven Unternehmen, Mitarbeitenden und verschiedenen HR-Themenbereichen. Der Datensatz wurde in drei Workspaces differenziert, die nach den Bedarfen der Zielgruppen skaliert sind, von Neueinsteigenden bis hin zu Personalverantwortlichen. Die Testungen zeigten, dass Struktur und thematische Einteilung der Daten die Antwortqualität wesentlich beeinflussen.



Abbildung 3: Investitionsrechnung der RAG-Lösungen in HR

AI

Grundlage war die Recherche zu kostenfreien Open-Source-Anwendungen, Ein daraufhin entwickelter Konverter überträgt die Datensätze in ein einfaches

Text- oder Markdown-Format und verbesserte die Ergebnisse, wobei weiterer Feinjustierungsbedarf besteht. Eine regelmäßige Strukturierung der Daten durch eine Fachkraft ist somit unerlässlich für den Betrieb eines eigenständigen Assistenzsystems im Unternehmen.

Der Einsatz von Assistenzsystemen im Personalbereich ist zulässig, aber reglementiert. Zu prüfen sind insbesondere Diskriminierungsrisiken und die Verarbeitung personenbezogener Daten. Beschäftigte benötigen tätigkeitsbezogene Kompetenz im Umgang mit dem System; angrenzende Unternehmensbereiche sowie Außenstehende, etwa Bewerbende, sind verständlich über den Einsatz zu informieren. Diese Hinweise sind mit den bestehenden DSGVO-Informationen abzustimmen.

Literatur

- Freier, A., Junghans, S., & Dietrich, F. (2026). *Rechtlicher Leitfaden zur Nutzung von KI-Systemen in Unternehmen* (1. Aufl.). Westsächsische Hochschule Zwickau; PÜHN Rechtsanwälte
- Merla, L., & Meeser, C. (2026, 4. August). *Neue KI-Pflichten ab August 2026: Handlungsbedarf für Arbeitgeber und HR* [Webinar-Präsentation]. mmp Akademie, Berlin, Deutschland
- Wahl, B. (2026). KI: Semantik stark, Generierung schwach: Warum RAG-Systeme eine informationswissenschaftliche Architektur brauchen. *Information – Wissenschaft & Praxis*, 77(2–3), 90–102. <https://doi.org/10.1515/iwp-2026-2001>

Förderhinweis

Dieses Forschungs- und Entwicklungsprojekt wird durch das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR), Förderrichtlinie „Zukunft der Arbeit: Regionale Kompetenzzentren der Arbeitsforschung. Erste Wettbewerbsrunde: Gestaltung neuer Arbeitsformen durch Künstliche Intelligenz“ im Programm „Innovationen für die Produktion, Dienstleistung und Arbeit von morgen“ (Förderkennzeichen: 02L19C306) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt dieser Publikation liegt bei den Autor:innen.

Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt



Kontaktdaten

ATB Arbeit, Technik und Bildung gGmbH
Neefestraße 76 09119 Chemnitz
Ansprechperson:
Claudia Graf-Pfohl
Telefon: 0371- 3 69 58 17
Email: graf-pfohl@atb-chemnitz.de