



Was diese Systeme technisch tun — und was **nicht**.

Notizen eines Ingenieurs für Kanzleien. **Die rechtliche Einordnung bleibt bei Ihnen** — ich beschreibe nur, was in der Maschine passiert. Die Pflichten kennen Sie besser als ich; was ein Sprachmodell tatsächlich tut, steht in keinem Kommentar.

1 Wie ein Sprachmodell arbeitet

Es sagt das wahrscheinlichste nächste Textstück voraus, Wort für Wort, auf Basis von Mustern aus Trainingstexten. Es **schlägt nichts nach**, sondern erzeugt. Deshalb entstehen Fundstellen, die plausibel aussehen und nicht existieren: Ein Aktenzeichen im richtigen Format ist für das Modell ein gutes Ergebnis — unabhängig davon, ob es die Entscheidung gibt. Kein Bug, den ein Update behebt, sondern die Funktionsweise. **Zwei Dinge senken die Fehlerquote messbar**: Anbindung an eine echte Datenbank und die Pflicht zur Belegstelle. Beides ersetzt die Prüfung nicht.

2 Wo es trägt, wo es versagt

BELASTBAR

- Lange Dokumente lesen, sortieren, zusammenfassen
- Bestimmte Angaben aus vielen Seiten herausziehen
- Struktur in unstrukturierten Text bringen
- Entwürfe gliedern, Textbausteine vorbereiten
- Posteingang vorsortieren und zuordnen

UNZUVERLÄSSIG

- Rechnen — jede Zahl ist geraten, nicht berechnet
- Fundstellen ohne angebundene Datenbank
- Fristen selbstständig ermitteln
- Vollständigkeit zusichern („alles berücksichtigt“)
- Aussagen zu allem nach dem Trainingsschluss

Merksatz: **Lesen und Ordnen ja, Rechnen und Behaupten nein**. Zweihundert Seiten Kontoauszüge nach Positionen durchgehen: gern. Die Summe darunter bildet ein Rechner.

3 Wohin Ihre Daten gehen — drei Betriebsarten

BETRIEBSART	WO DIE DATEN LIEGEN	PRAKTISCH
Lokal auf eigener Hardware	Verlassen das Haus nicht. Kein Anbieterzugriff möglich.	Höchste Kontrolle, begrenzte Modellstärke, eigene Hardware nötig
Self-hosted auf gemietetem Server	Auf einem Server, den Sie kontrollieren. Zugriff hängt am Hoster.	Guter Kompromiss; Standort und Vertrag entscheiden
Anbieter-Cloud (API)	Beim Anbieter. Technisch besteht Zugriffsmöglichkeit.	Stärkste Modelle; Vertragsgestaltung ist hier alles

Wichtig: „Wir trainieren nicht auf Ihren Daten“ und „wir können nicht darauf zugreifen“ sind zwei verschiedene Zusagen. Die erste ist üblich, die zweite selten. Ob die technische Zugriffsmöglichkeit für Ihre Zwecke ausreicht, beantworten Sie — nicht ich. Kostenlose Endkunden-Tools sind eine eigene Kategorie: Dort ist die Weiterverwendung von Eingaben oft Teil des Geschäftsmodells.

4 Fünf Fragen an jeden Anbieter

1. **Wo genau werden meine Eingaben verarbeitet** — auf welchem Server, in welchem Land?
2. **Wer kann technisch darauf zugreifen**, auch zu Wartungszwecken — und wird das protokolliert?
3. **Wie lange werden Eingaben gespeichert**, und wie lösche ich sie nachweislich?
4. **Woher stammen die Fundstellen** — aus einer angebundene Datenbank oder aus dem Modell selbst?
5. **Was passiert mit meinen Daten bei einem Anbieterwechsel** oder bei dessen Insolvenz?

Zur Einordnung: Ich bin Ingenieur, kein Rechtsanwalt. Diese Seite beschreibt Technik und ist keine Rechtsberatung; sie prüft keinen Einzelfall und bewertet kein Produkt. Berufsrechtliche, datenschutz- und haftungsrechtliche Fragen beurteilen Sie. Ohne Gewähr für Vollständigkeit; technischer Stand Juli 2026.