

TRAININGSDATEN

Veröffentlicht am 3. Juli 2026 von Marie

Was sind Trainingsdaten?

Trainingsdaten sind alle Texte und Informationen, aus denen ein KI-Modell gelernt hat. Sie bestimmen, welches Wissen die KI hat und wie gut sie in bestimmten Themen und Sprachen ist.

Beispiel: Eine KI, die überwiegend mit englischen Texten trainiert wurde, antwortet auf Deutsch oft schwächer. Was in den Daten steckt, zeigt sich später in den Antworten.

Worauf es ankommt: Entscheidend sind Qualität, Breite, Aktualität und Ausgewogenheit der Daten. Lücken oder Verzerrungen in den Trainingsdaten schlagen direkt auf die Ergebnisse durch.

Wichtig: Eine KI ist immer nur so gut wie ihre Trainingsdaten. Das erklärt, warum Modelle blinde Flecken haben und manches Wissen veraltet ist.