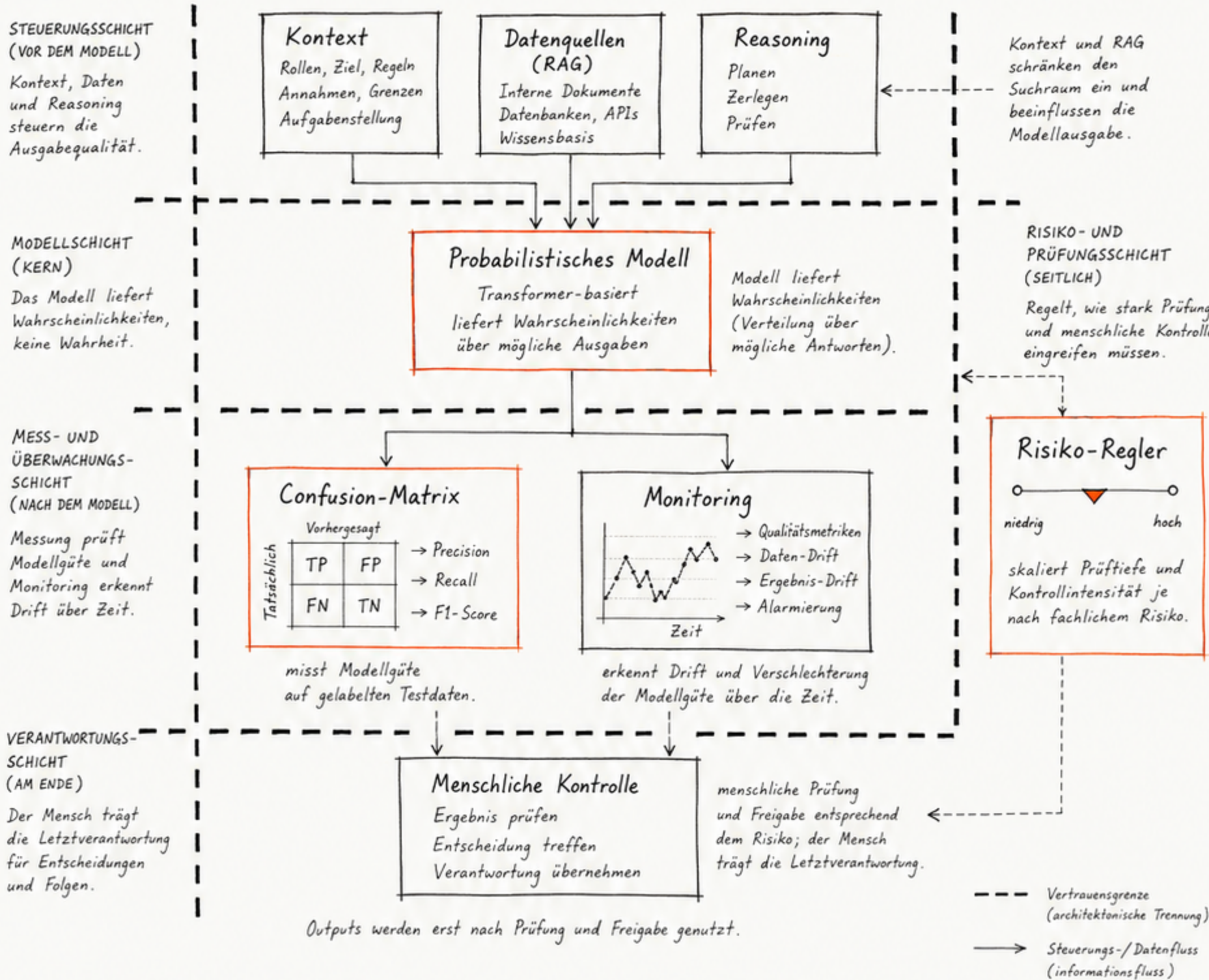


Qualität entsteht nicht im Modell allein

Ein KI-Modell liefert Wahrscheinlichkeiten, keine Wahrheit; Qualität entsteht erst durch Kontext, Daten, statistische Prüfung und Verantwortung, abgestuft nach Risiko.



Wahrscheinlichkeit statt Wahrheit

Ein Sprachmodell berechnet das wahrscheinlichste nächste Wort, nicht die korrekte Antwort. Es kann überzeugend falsch liegen und erfindet Aktenzeichen, Paragraphen oder Urteile, ohne Unsicherheit zu signalisieren.

→ kein bug, sondern bauart

Vom Testen zum Messen

Deterministische Software kennt nur bestanden oder durchgefallen, KI-Qualität dagegen ist eine Kennzahl über viele Beispiele: Precision, Recall und F1 entstehen aus der Confusion-Matrix, nicht aus einem einzelnen Pass/Fail.

→ statistische gütemaße statt grüner haken

Qualität entsteht im Kontext

Nicht das Modell, sondern seine Einbettung bestimmt die Güte. Expliziter Kontext, angebundene Datenquellen über RAG, Rollen-Prompts und Reasoning steuern das Ergebnis; eine perfekte Suche gibt es dabei nie.

→ rag liefert treffer, keine garantie

Risiko skaliert die Kontrolle

Modellgüte driftet über die Zeit, deshalb reicht eine einmalige Abnahme nicht. Je höher das fachliche Risiko, desto stärker müssen statistische Validierung, laufendes Monitoring und menschliche Letztverantwortung greifen.

→ verantwortung lässt sich nicht delegieren

⚠ Keine personenbezogenen Daten in offene Modelle eingeben und ausschließlich von der IT freigegebene Werkzeuge nutzen; sonst verlässt schützenswertes Wissen den kontrollierten Raum.

⚠ Sykophantie und Halluzination machen plausible Antworten gefährlich: Ergebnisse fachlich, rechtlich und ethisch prüfen, statt der überzeugenden Formulierung des Modells zu vertrauen.

⚠ Eine einmalige Abnahme genügt nicht: Ohne kontinuierliches Monitoring bleibt die schleichende Verschlechterung der Modellgüte über die Zeit unbemerkt, bis Entscheidungen darauf beruhen.

→ KERNGEDANKE

Wer KI produktiv einsetzt, baut kein klügeres Orakel, sondern ein geprüftes System: Das Modell schlägt vor, der Kontext schärft, die Messung belegt, der Mensch verantwortet.