

KI-Agenten kontrollieren: Warum die Prüfschicht deterministisch sein muss

MANAGEMENT SUMMARY & KERNAUSSAGE

KI-Agenten produzieren immer mehr Arbeit – geprüft werden muss trotzdem deterministisch. Das Zürcher Startup Klarent zeigt mit seiner Seed-Runde, wie 'deterministic autonomy' funktioniert: Die KI schreibt Tests, ausgeführt werden sie deterministisch. Für Schweizer KMU heisst das: feste Regeln, Konfidenzschwellen, menschliche Eskalation – und genau diese Kontrollschicht lässt sich outsourcen.

KONTEXT & RELEVANZ FÜR ENTSCHEIDER

Wenn KI-Agenten Rechnungen, Verträge, Analysen und Entscheidungsvorlagen erstellen, verschiebt sich die knappe Ressource vom Produzieren zum Beurteilen. Vom Ausführen zum Beurteilen beschreibt genau diesen Wechsel: Wer nur auf mehr Output setzt, übersieht, dass jede zusätzliche automatisierte Leistung auch geprüft werden muss. Genau hier setzt das Prinzip der 'deterministic autonomy' an, mit dem das Zürcher Startup Klarent im Oktober 2026 eine Seed-Runde über €7.13 Mio. abgeschlossen hat – Kunden sind unter anderem NZZ, Sixt und JD Sports.

Der gefährlichste Fehler ist, die Prüfung einem zweiten KI-Agenten zu geben, der den ersten 'korrigiert'. Wenn die Prüfschicht selbst halluziniert, vertraust du einer falschen Kontrolle – der schlimmste Ausfallmodus. Deterministisch bedeutet: gleicher Input, gleicher Output, reproduzierbar und auditierbar. Die Frage, wer bei Abweichungen entscheiden darf, klärst du am besten mit einer [Entscheidungsmatrix](/ratgeber/wer-darf-entscheiden-die-ki-entscheidungsmatrix-fur-kmu-fuhrungskrafte).

VERGLEICH & SCHWEIZER MARKTPRAXIS

Nur Agent, keine Prüfschicht	Schnell, aber unkontrolliert. Fehler fallen spät auf, Vertrauen leidet.
KI prüft KI	Skaliert, aber Prüfschicht kann halluzinieren. Kein Audit-Trail.
Deterministische Regeln + Konfidenzschwellen	Reproduzierbar, auditierbar. Regeln sind aber initial aufwendig.
Deterministische Prüfschicht + Human-in-the-Loop (Outsourcing)	Klare Eskalation, kein eigenes Team nötig. Bestes Verhältnis aus Kontrolle und Aufwand.

SCHWEIZER UMSETZUNGSKRITERIEN (REVD SG & IT-ARCHITEKTUR)

- ☑ Vollständige Konformität mit dem neuen Schweizer Datenschutzgesetz (revDSG).
- ☑ Keine Abhängigkeiten: Zero Lock-In auf proprietäre geschlossene Modelle.
- ☑ Schnittstellen zu Schweizer ERPs (Abacus, Bexio, M365) im operativen Betrieb.

Unverbindliche Prozess-Analyse

Lass uns in 20 Minuten prüfen, wo KI-Agenten in deinem Unternehmen sofort messbare Entlastung schaffen.

ki-outsourcing.ch