

Die 36 Strategeme für KI Agenten

Eine Zeile je Strategem: woran man das Muster in der Ausgabe eines Agenten erkennt und was dagegen hilft. Gedacht zum Danebenlegen beim Review. Nichts davon geht über das hinaus, was der Beitrag belegt.



01 Den Kaiser täuschen

Eine lange Reihe grüner Läufe senkt die Wachsamkeit. **Abhilfe:** ein Gate, das nicht ermüdet.



02 Wei belagern

Das Review scheitert immer wieder am Code. **Abhilfe:** die Korrektur geht zurück auf die Spezifikation.



03 Mit fremdem Messer

Agent prüft Agent, kein Mensch im Protokoll. **Abhilfe:** eine festgehaltene menschliche Prüfung.



04 Den erschöpften Feind erwarten

Die Prüfung läuft vor dem Zug des Agenten. **Abhilfe:** erst nach dem ersten Entwurf prüfen.



05 Ein Feuer ausnutzen

Unter Zeitdruck rutscht Ungeprüftes durch. **Abhilfe:** ein serverseitiges Gate ohne Termindruck.



06 Im Osten lärmern

Achthundert Zeilen Bericht, ein echter Fehler. **Abhilfe:** ein verdichteter Befund statt des Rohberichts.



07 Aus Nichts etwas erzeugen

Belegzeilen, die es in den Dateien nicht gibt. **Abhilfe:** Belege gegen die Dateien halten.



08 Nach Chencang marschieren

Jeder sichtbare Test grün, Verhalten unklar. **Abhilfe:** zurückgehaltene Tests.



09 Das Feuer beobachten

Ein einziger Prüfer entscheidet allein. **Abhilfe:** zwei Modelle, der Mensch nur beim Widerspruch.



10 Dolch hinter dem Lächeln

Zustimmender Ton bei dünnem Inhalt. **Abhilfe:** den Beleg werten, nicht die Zustimmung.



11 Der Pflaumenbaum verdorrt

Eine große Änderung auf einen Schlag. **Abhilfe:** kleine Scheiben, jede einzeln geprüft.



12 Das Schaf wegführen

Nebenbei eingebaute Extras, niemand hat sie bestellt. **Abhilfe:** Bauformen messen, nicht nur Fehler zählen.



13 Auf das Gras schlagen

Die Prüfung hat noch nie etwas gemeldet. **Abhilfe:** einen bekannten Fehler einbauen.



14 Einen Leichnam ausleihen

Eigene Namen statt der gängigen. **Abhilfe:** vorhandene Normen und Werkzeugnamen nutzen.



15 Den Tiger locken

Die Tests laufen nur auf dem Rechner des Entwicklers. **Abhilfe:** eine frisch aufgesetzte Maschine.



16 Erst loslassen

Die Regel steht schon vor dem Entwurf. **Abhilfe:** erst frei entwerfen lassen.



17 Einen Backstein hinwerfen

Eine riesige Spezifikation ohne eine einzige Rückfrage. **Abhilfe:** ein knapper Anstoß, die Rückfragen zeigen die Lücken.



18 Den Anführer fangen

Zwanzig Einzelregeln ohne gemeinsame Klammer. **Abhilfe:** eine tragende Regel, die Projektverfassung.

Die 36 Strategeme für KI Agenten, Fortsetzung



19 Das Brennholz wegnehmen

Die Schwelle steht vorab im Auftrag. **Abhilfe:** die Schwelle erst nachträglich offenlegen.



20 Im Trüben fischen

Rohdiff statt Befund. **Abhilfe:** eine verdichtete Ausgabe.



21 Die Zikade häutet sich

Die Suite glänzt, das Verhalten weicht ab. **Abhilfe:** zurückgehaltene Tests entscheiden.



22 Die Türe schließen

Die Abnahme hängt an einer Zahl. **Abhilfe:** eine Bedingung mit echtem Wahrheitswert.



23 Mit dem Fernen verbünden

Der Prüfer stammt aus derselben Modellfamilie. **Abhilfe:** eine fremde Familie streut die blinden Flecken.



24 Einen Weg ausleihen

Der mitgeführte Kontext wächst Schritt für Schritt. **Abhilfe:** den Kontext kuratieren.



25 Die Balken stehlen

Es läuft, aber niemand kann daran weiterbauen. **Abhilfe:** die innere Qualität messen.



26 Den Maulbeerbaum schelten

Die Regel gilt nur auf dem Papier. **Abhilfe:** sie einmal sichtbar an einem Beispiel durchziehen.



27 Sich dumm stellen

Eine Selbstauskunft statt eines Ergebnisses. **Abhilfe:** das Ergebnis prüfen, nie die Auskunft.



28 Die Leiter wegziehen

Die Spezifikation liegt im Werkzeug eines Anbieters. **Abhilfe:** eine Datei im eigenen Bestand.



29 Blumen blühen lassen

Die Abdeckung wirkt vollständig, der Beleg trägt nicht. **Abhilfe:** Belege stichprobenweise nachschlagen.



30 Den Gast zum Gastgeber

Der Agent prägt die Konventionen im Bestand. **Abhilfe:** die Verfassung bleibt beim Menschen.



31 Die schöne Frau

Das Urteil folgt der zuerst gezeigten Antwort. **Abhilfe:** die Reihenfolge drehen und das Messgerät selbst prüfen.



32 Die leere Stadt

Souveräner Ton ohne Substanz. **Abhilfe:** den Beleg bewerten, nicht den Ton.



33 Die gegenseitige Spionage

Fremde Inhalte werden als Anweisung gelesen. **Abhilfe:** Lesestoff als unzuverlässig behandeln.



34 Die Selbstverletzung

Der Prüfer meldet nie etwas. **Abhilfe:** einen Kanarienvogel einbauen.



35 Die Verkettung

Ein einzelnes Gate trägt alles. **Abhilfe:** eine Kette voneinander unabhängiger Prüfungen.



36 Flucht ist das Beste

Ein Lauf wird endlos nachrepariert. **Abhilfe:** verwerfen und neu ansetzen.



Ein Agent hat keine Absicht. Er erzeugt die Ausgabe mit der besten Bewertung und dem geringsten Aufwand.