WELT 9: MASCHINENRAUM · MISSION 9.5

Das Qualitäts-Tandem

Erst die Kriterien, dann das Duell: Zwei KI-Werkzeuge bauen und prüfen um die Wette, vier Lockvögel testen die Prüfer – und am Schluss geben Sie frei, niemand sonst.

Zur Mission

DAUER 75 Min. KOSTEN Gratis mit Konto WERKZEUGE OpenAI Codex, Claude Code, Claude (Web), Google Gemini, ChatGPT

DATENAMPEL Grün – unbedenklich

Das Wichtigste

- Reihenfolge: erst Kriterien festlegen, dann Werkzeuge vergleichen, dann entscheidet der Mensch.
- Zwei K.O.-Kriterien immer: fachlich falsch und Personendaten.
- Zwei verschiedene Anbieter, exakt derselbe Prompt – sonst vergleichen Sie Prompts statt Werkzeuge.
- Prüferinnen liefern Befunde mit Fundstelle, sie ändern nichts.
- Lockvögel (bekannte Fehler) messen, wie zuverlässig eine Prüfung ist; falsche Alarme mitzählen.
- Wer baut, wer prüft: aus eigenen Messungen ableiten, je Art von Aufgabe – nicht fix.
- KI-Prüfungen sind Zusatzchecks. Die Freigabe gibt die Lehrperson.

Ihr Produkt

Ein ausgefülltes **Tandem-Prüfraster mit Freigabeprotokoll**: Ihre Kriterien mit K.O.-Punkten, der Lockvogel-Test beider Werkzeuge, Ihre Entscheidung – und eine **begründete Rollenverteilung** («wer baut, wer prüft») für diese Art Aufgabe.

Abnahme: Die Kriterien standen fest, bevor Sie ein Ergebnis sahen; der Lockvogel-Test zeigt für jedes Werkzeug, welche der vier Fehler es fand; und das Freigabeprotokoll nennt Ihre Entscheidung, Ihre eigenen Korrekturen und die künftige Rollenverteilung mit Begründung.

So gehen Sie vor

1. Ihre Wette – und die Kriterien zuerst
2. Der Prüfling mit den Lockvögeln
3. Duell 1: Zwei Prüferinnen, ein Prüfling
4. Duell 2: Bauen und kreuzweise prüfen
5. Sie entscheiden – und verteilen die Rollen

Start-Prompt

Du bist eine strenge Qualitätsprüferin für Lernmaterial an einer Schweizer Berufsfachschule.
Prüfe das folgende Arbeitsblatt (Zielgruppe: [BERUF, z. B. Detailhandel EFZ], [LEHRJAHR, z. B. 1. Lehrjahr]) nach diesen Kriterien:
[IHRE KRITERIEN AUS DEM RASTER, z. B. K.O.: fachlich falsch; K.O.: Personendaten; Lernzielbezug; Sprachniveau; Barrierefreiheit; Schweizer Rechtschreibung; Lösung vollständig]
Regeln:
- Ändere nichts am Arbeitsblatt. Liefere nur eine nummerierte Liste von Befunden.
- Jeder Befund nennt: Fundstelle, Kriterium, warum es ein Problem ist, Verbesserungsvorschlag.
- Trenne klar zwischen "Fehler" und "Geschmackssache".
- Wenn du dir bei einer Fachaussage unsicher bist, schreibe [UNSICHER] statt zu raten.
Schweizer Rechtschreibung (immer «ss», kein Eszett).
Arbeitsblatt:
[TEXT DES ARBEITSBLATTS]

Qualitätscheck

- Die Kriterien und Gewichte standen fest, bevor Sie ein Ergebnis gesehen haben.
- Beide Werkzeuge stammen von verschiedenen Anbietern und erhielten exakt denselben Prompt.
- Für jeden Lockvogel ist eingetragen, wer ihn gefunden hat – inklusive falscher Alarme.
- Die Freigabe nennt Ihre eigenen Prüfungen an einer Quelle und Ihre Korrekturen.
- Die künftige Rollenverteilung ist mit Messwerten begründet und gilt ausdrücklich nur für diese Art Aufgabe.
- An keines der Werkzeuge gingen echte Personendaten.

Hängen Sie fest? Frag die KI

Ich bin Lehrperson an einer Schweizer Berufsfachschule und will zwei KI-Werkzeuge vergleichen, die Lernmaterial prüfen. Hilf mir, 6 überprüfbare Kriterien für ein Arbeitsblatt in [IHR FACH, z. B. Detailhandel] festzulegen, darunter zwei K.O.-Kriterien. Erkläre mir auch, warum ich die Kriterien vor dem Vergleich festlegen sollte und welche Denkfehler sonst drohen.

Klappt nicht? Gratis-Tandem: Codex mit Gratis-ChatGPT-Konto als erstes Werkzeug, als zweites ein Gratis-Chat eines anderen Anbieters (Claude, Gemini oder Copilot), dem Sie den Text per Kopieren geben. Wer Claude Code im Abo hat, kann es als zweites Werkzeug einsetzen. Wichtig ist nur: zwei verschiedene Anbieter.

Stand 19.11.2026 · CC BY 4.0 Alpasana GmbH · Werkzeuge ändern sich – prüfen Sie selbst.