LEITFADEN · KI IM ARBEITSALLTAG

Mit KI wirksam zusammenarbeiten

Was die KI übernimmt, wie du sie beauftragst und wie stark du checken musst, ob das Ergebnis stimmt.

New Learning Lab · Siara KI · Jan Foelsing

FELDEXPERIMENT MIT 758 BERATERINNEN UND BERATERN

Dasselbe Modell hat dieselbe Zielgruppe entweder besser oder schlechter gemacht. Warum?

>40 %

Besser: höhere Ergebnisqualität bei Aufgaben, die das Modell gut versteht und beherrscht

−19

Schlechter: 19 Prozentpunkte seltener richtig, falls Anfrage knapp außerhalb der KI-Stärken liegt

Deine Vorarbeit entscheidet über die Ergebnisqualität: zwei Dinge, die du in der Hand hast:

1 · Welche Aufgabe du der KI gibst. Aufgaben sehen teilweise fast gleich schwer aus. Ob die KI etwas gut kann, merkst du nur, wenn du das Ergebnis selbst beurteilen kannst. Beginne also bei bekannten Themen, um herauszufinden, was gut funktioniert und was nicht.

2 · Wie du mit dem Ergebnis umgehst. Auch bei der falschen Lösung klingt die KI überzeugend. Stelle gezielt Rückfragen. Lass eine andere KI die Ergebnisse challengen. Überprüfe die Ergebnisse je nach Kontext unterschiedlich umfangreich.

Quelle: Dell'Acqua et al., HBS Working Paper 24-013 (2023); die begutachtete Fassung (Organization Science, 2026) nennt leicht andere Werte

Das 4D-Modell von Anthropic

1 DELEGATION

Aufteilen

Was macht die KI, was mache ich?

→

2 DESCRIPTION

Beauftragen

Wie sage ich der KI, was ich brauche?

→

3 DISCERNMENT

Überprüfen

Stimmt das Ergebnis?

→

4 DILIGENCE

Verantworten

Kann ich es vertreten?

Deine Rückfragen und Überarbeitungen verbinden Step 2 und 3: beschreiben, prüfen, nachschärfen, bis der Auftrag erfüllt ist. Danach die KI fragen, was wir dabei gelernt haben und was davon wir uns wie merken sollten.

Modell: Rick Dakan, Joseph Feller und Anthropic, „The AI Fluency Framework“ (2025), CC BY-NC-SA 4.0. Deutsche Begriffe und Ergänzungen: Jan Foelsing

Was die KI übernimmt und wie gründlich du prüfst

	LEICHT ZU PRÜFEN dein Fach, Quelle liegt vor	SCHWER ZU PRÜFEN fremdes Fach, große Mengen
Fehler harmlos intern, leicht zu korrigieren	HIER ANFANGEN **Übergeben, überfliegen** Protokolle, interne Notizen	**Übergeben, Stichproben** Umfrage-Freitexte nach Themen sortieren, Einstieg in ein fremdes Fachgebiet
Fehler folgenreich kaum umkehrbar, für Kunden sichtbar oder mit sensiblen Daten	**Übergeben, alles prüfen** Kundenmail mit Zahlen aus deinem Angebot	Nicht an die KI übergeben Preise, Personal, Recht: Du entscheidest, KI liefert höchstens Ideen und Gegenargumente

Matrix: eigene Darstellung, angelehnt an „Delegation" im AI Fluency Framework (Dakan, Feller und Anthropic, 2025)

Sechs Bausteine für einen klaren Auftrag

Ziel
Was soll entstehen, und wofür?

Adressat
Wer liest es, was weiß die Person oder Zielgruppe?

Material
Originale, Vorlagen, Beispiele, Negativabgrenzung, …

Form
Länge, Aufbau, Ton, gern mit Beispiel

Regeln
Was sie nicht darf, wann sie nachfragt

Fertig, wenn …
Woran du ein gutes Ergebnis erkennst

„Fertig, wenn“ ist dein Prüfmaßstab für Schritt 3. Dazu kannst du Vorgehen (welche Schritte) und Verhalten (kritisch oder bestärkend) beschreiben.

Quellen: Anthropic, Prompting-Dokumentation (lange Dokumente oben); AI Fluency Framework, 2025 (Vorgehen und Verhalten); Bausteine: eigene Darstellung

Derselbe Auftrag, einmal ohne und einmal mit den Bausteinen

VORHER · 4 WÖRTER

„Fass das Protokoll zusammen.“

Zweck, Leser, Umfang und Maßstab muss die KI raten. Aktuelle Modelle setzen genau um, was dasteht. Was fehlt, müssen sie sich zusammenreimen.

NACHHER · 45 WÖRTER

Fasse das angehängte Protokoll der Projektsitzung zusammen **[Ziel]**. Es geht an die Geschäftsführung, die nicht dabei war **[Adressat]**. Nutze nur das Protokoll **[Material]**. Höchstens acht Stichpunkte: Entscheidungen, offene Punkte, nächste Schritte mit Namen **[Form]**. Erfinde nichts; fehlt etwas, schreib „offen“ **[Regeln]**. Fertig, wenn jede Entscheidung mit der Protokollstelle belegt ist **[Fertig, wenn]**.

Quelle: Anthropic, „Prompting best practices“, Claude-Dokumentation für die aktuellen Modelle (abgerufen November 2026). Beispielauftrag: eigene Darstellung

Prüfe die Ergebnisse und Quellen regelmäßig!

Lass dir von der KI zeigen

Annahmen: Was hat sie vorausgesetzt?

Belege: Welche Stelle stützt jede Aussage?

Lücken: Was fehlt ihr?

Prüfe selbst

Belegstelle im Original nachlesen

Zahlen nachrechnen

Testfrage oder Stichprobe

Gegen dein „Fertig, wenn“ halten

Nicht belegbar? Als offen markieren

„Bist du sicher?“ ist keine Prüfung. Auf Nachfrage können auch richtige Antworten durch die KI anders bewertet und verändert werden. Ein zweites Modell irrt oft genauso, ist aber eine erste, sinnvolle Prüfschleife vor dem eignen Check.

Quellen: Sharma et al., ICLR 2024 (Nachfragen); Kim et al., ICML 2025 (zweites Modell); Anthropic, „Reduce hallucinations“ (Belegstellen)

Nachbessern, bis das Ergebnis passt

Du hältst das Ergebnis gegen dein „Fertig, wenn …“, sagst der KI genau, was noch fehlt, und lässt sie eine neue Fassung schreiben. Diese Runde wiederholst du, bis alles erfüllt ist.

Entwurf → Prüfen → Feedback → Neue Fassung → Fertig, wenn …

Konkret sagen

Was, wo, wie: „Absatz 2 kürzen, Zahl aus Tabelle 3.“ Gezielte Feedbackrunden brachten im Schnitt wesentlich bessere Ergebnisse.

In Etappen/Iterationen

Erst die Gliederung freigeben, dann Abschnitt für Abschnitt. So fallen Fehler früh auf.

Ggf. neu starten

Dreht sich die KI im Kreis? Werden die Ergebnisse eher schlechter? Neuen Chat starten und vollständigen Auftrag und Ergebnisse übergeben.

Quellen: Madaan et al., „Self-Refine“, NeurIPS 2023 (Feedbackrunden); Laban et al., „LLMs Get Lost in Multi-Turn Conversation“, arXiv 2025 (neu starten)

Übernimm nur, was du verantworten kannst

Freigabe wie in der Matrix

Wie gründlich du vor der Freigabe prüfst, zeigt die Matrix aus Schritt 1: überfliegen, Stichproben, alles prüfen oder **gar nicht an die KI übergeben**.

Offenlegen: Sag, wo KI am Ergebnis beteiligt war.

Je besser die Ergebnisse der KI auf dich wirken, desto eher wird einfach abgenickt. Das ist nachvollziehbar und menschlich, aber ein Fehler. Bleib genau!

Prüfen zur Gewohnheit machen

Das Gefühl täuscht dabei leicht:

+20 % schneller fühlten sich erfahrene Entwickler mit KI

−19 % langsamer waren sie, als man die Zeit maß

Darum nach jedem Check notieren:
Was hast du korrigiert? Wo macht die KI immer wieder Fehler? Das kommt in deinen KI-Memory oder einen Skill. Frag: Wie können wir uns dies hier so merken, dass der Fehler nicht mehr vorkommt?

Quellen: Dell'Acqua, „Falling Asleep at the Wheel", Working Paper, Harvard Business School (2022); Becker et al. (METR), arXiv 2025; Offenlegen: AI Fluency Framework (2025)

ZUM MITNEHMEN

Deine Checkliste für die nächste Übergabe

VORHER: WAS DU ÜBERGIBST

	leicht prüfbar	schwer prüfbar
Fehler harmlos	Grob checken	Stichproben
Fehler folgenreich	Alles prüfen	KI max. als Sparring

Die Matrix zeigt, ob du übergibst und wie gründlich du danach prüfst.

- ✓ **Aufteilen** (Delegation) — Mit der Matrix klären, was die KI übernimmt und was bei dir bleibt
- ✓ **Beauftragen** (Description) — Sechs Bausteine nutzen und sagen, wann es fertig ist
- ✓ **Überprüfen** (Discernment) — Belege anfordern, am Original prüfen
- ✓ **Überarbeiten** (Schleife) — Konkret nachbessern, im Loop? Bei Bedarf neuer Chat
- ✓ **Verantworten** (Diligence) — Erst nach Prüfung freigeben, KI-Anteil offenlegen, Fehler festhalten