STATUS: ENTWURF · MIT BLEISTIFT

Vom Prompt zum prüfbaren Büro-Assistenten

Ein Blick in die konkrete Skill-Architektur, ihre Prompt-Ausschnitte und das, was die Tests belegen.

Maximilian Tabor

Büro-Assistent · Skill- & Testdokumentation

KURZZUSAMMENFASSUNG

Der Büro-Assistent ist nicht als einzelner Allzweckprompt angelegt. Eine zentrale Steuerung wählt aus zehn Aufgabenpfaden, die jeweils eigene Eingaben, Arbeitsschritte, Nachweise und Freigaben festlegen. Das Beispiel Reisekosten zeigt, wie diese Promptstruktur fehlende Angaben sichtbar macht, aktuelles Fachwissen einbindet und eine menschliche Prüfung erhält. Die Tests zeigen einen messbaren Vorteil des Fachwissens in den Finanzfällen, sind aber noch kein Nachweis für den Einsatz auf der Zielplattform.

01 Die zentrale Frage ist nicht nur, was ein Prompt sagt

Ein Prompt für Büroarbeit muss mehr leisten, als eine Aufgabe zu formulieren. Er muss klären, welche Arbeit gemeint ist, welche Informationen zählen, was bei Unsicherheit geschieht und wo die Entscheidung beim Menschen bleibt.

Im Büro-Assistenten sind diese Regeln auf mehrere Ebenen verteilt. Eine zentrale Skill-Anweisung übernimmt Rolle, Auswahl des passenden Pfads und Sicherheitsgrenzen. Zehn einzelne Pfade beschreiben konkrete Aufgaben wie Protokolle, Posteingang, Rechnungsprüfung oder Reisekosten. Fachreferenzen enthalten veränderliche Sachwerte und Quellen. Interne Vorgaben sollen später Informationen wie Reiserichtlinie, Kontenplan und Ansprechpartner ergänzen.

Das ist eine wichtige Architekturentscheidung. Die zentrale Anweisung bleibt übersichtlich. Die konkreten Abläufe lassen sich separat prüfen und aktualisieren. Ein Reisekostenpfad muss nicht dieselben Arbeitsschritte verwenden wie eine Wissenssuche.

02 Die Promptstruktur in vier Ebenen

EBENE 1 Rolle und Verantwortung

Die Rolle beschreibt den Zweck des Assistenten und begrenzt seine Verantwortung. Ein kurzer Originalausschnitt aus dem Skill lautet:

KARTE 01 · SKILL-GRUNDANWEISUNG *Bleistift-Prinzip: vorschreiben ja, abstempeln nie*

```
Du nimmst Wissensarbeitenden wiederkehrende Büroarbeit ab und lieferst
prüffähige Entwürfe.
```

Der Skill legt außerdem fest, dass Entscheidungen, Freigaben und Versand bei einem Menschen bleiben. Das ist keine nachträgliche Warnung am Ende der Antwort. Es bestimmt, welche Handlungen der Assistent von Anfang an ausführen darf.

EBENE 2 Auswahl des Aufgabenpfads

Die zentrale Anweisung lässt den Assistenten nicht selbst einen frei erfundenen Ablauf bauen. Sie ordnet die Anfrage zunächst einer Aufgabe zu und verweist auf den passenden Pfad:

KARTE 02 · ROUTING-REGEL *Verhindert das Vermischen ähnlicher Aufgaben*

```
Wähle genau einen Pfad und lies vor dem Arbeiten dessen Datei vollständig.
```

Die Pfade unterscheiden nicht nur nach Stichworten. Sie definieren auch Abgrenzungen. Eine Lieferantenrechnung gehört in die Rechnungsprüfung. Ein Hotelbeleg aus einer Dienstreise gehört zu den Reisekosten. Ein Dokument, das zusammengefasst werden soll, ist etwas anderes als eine Frage zu internen Regeln. Das reduziert einen häufigen Promptfehler: Eine Antwort kann sprachlich überzeugend wirken und trotzdem den falschen Prozess bearbeiten.

EBENE 3 Regeln für den konkreten Arbeitsschritt

Jeder Pfad enthält unter anderem Auslöser, nicht passende Fälle, erforderliche Eingaben, Arbeitsschritte, Qualitätskriterien und Freigabe. Der Reisekostenpfad verlangt beispielsweise Reisedaten, Belege und Angaben zu gestellten Mahlzeiten. Er gibt vor, wie OCR-Abweichungen markiert werden, wann Beträge nachgerechnet werden und wann ein Beleg nur zur Prüfung markiert werden darf.

Für Werte, die sich ändern können, verweist der Pfad auf die Fachreferenz. Ein zentraler Originalausschnitt lautet:

KARTE 03 · FACHREFERENZ-PFLICHT *Kein Raten bei veränderlichen Pauschalen & Steuersätzen*

```
Werte (Pauschalen, Steuersätze, Fristen) immer dort nachlesen, nie aus dem
Gedächtnis.
```

EBENE 4 Ausgabe, Unsicherheit und Freigabe

Der Skill standardisiert die Ergebnisform. Aussagen sollen Fundstellen tragen. Unsicheres und Fehlendes wird sichtbar gekennzeichnet. Das Prinzip ist im Prompt knapp formuliert:

KARTE 04 · LÜCKEN- & QUALITÄTSREGELN *Unsicherheit wird markiert, nicht überspielt*

```
Fehlendes [FEHLT: was], Unsicheres [UNSICHER: Grund].
Rechenprobe je Beleg dokumentiert; OCR-Unsicherheit markiert.
Pauschalen mit Rechenweg und Stand der Werte.
Keine Auszahlung, keine Buchung.
```

So wird ein plausibel klingender Entwurf nicht mit einer geprüften Abrechnung verwechselt. Die Ausgabe hält fest, was sie weiß, was sie nicht weiß und was jemand nachsehen muss.

03 Ein konkretes Beispiel aus dem Testfall

Der Testfall gibt eine zweitägige Reise nach Köln mit Abfahrt, Rückkehr, Hotelrechnung und Frühstück an. Die Bitte lautet im Wortlaut:

EINGABE · REISEKOSTEN-ANFRAGE *Köln, zweitägige Dienstreise*

```
Bitte rechne mir die Verpflegungspauschalen aus (ich hätte gern 2 × 14 €) und
schlag die Kontierung vor.
```

Der Pfad führt den Assistenten durch mehrere Prüfschritte. Er liest die relevanten Angaben aus, prüft die Rechnung und ordnet das Frühstück als gestellte Mahlzeit ein, weil es auf der Firmenrechnung steht. Anschließend berechnet er die Kürzung der Pauschale und zeigt den Rechenweg. Im Test kommt er auf 22,40 Euro statt der gewünschten ungekürzten 28 Euro.

PRÜFBLATT · ENTWURF REISEKOSTEN (KÖLN) **GEPRÜFT MIT BLEISTIFT · NICHT GEBUCHT**

Prüfpunkt	Ausgangswert / Wunsch	Bleistift-Vermerk des Assistenten
Verpflegungspauschale	~~28,00 € (2 × 14 €)~~	**22,40 €** *(−5,60 € Kürzung für gestelltes Frühstück)*
Frühstück auf Beleg	Ausgewiesen mit **19 % USt**	**[PRÜFEN]** *Seit 1.1.2026 gelten 7 % für Speisen → Buchhaltung klärt*
Offene Unterlagen	Angaben nur aus Textnachricht	**[FEHLT: Originalbeleg]** **[FEHLT: Fahrtkosten]** **[FEHLT: Reiserichtlinie]**

Zusätzlich markiert der Assistent den ausgewiesenen Steuersatz für das Frühstück als Prüfpunkt. Er korrigiert die Rechnung nicht selbst. Er weist darauf hin, dass die Buchhaltung den Beleg prüfen soll. Auch die fehlende Originalrechnung, die Fahrtkosten und interne Vorgaben bleiben als offene Punkte sichtbar.

Der Nutzen liegt nicht darin, dass der Prompt eine besonders ausführliche Antwort erzwingt. Er macht den Weg zur Antwort prüfbar und begrenzt, was aus der Antwort folgen darf.

04 Was die Tests zeigen und was nicht

Das Paket dokumentiert 14 Testfälle in einer Ersatzumgebung. Die Fälle wurden mit Fachwissen und mit einer reduzierten Konfiguration ohne die Fachreferenzen ausgeführt. Kritische Fehlschläge wurden in diesen Läufen nicht dokumentiert, und die Pfadwahl traf in 15 von 15 Fällen den richtigen Pfad.

Testbereich	Mit Fachwissen	Ohne Fachwissen	Einordnung
Alle 14 Testfälle 68 Prüfpunkte gesamt	67 / 68 (98,5 %)	62 / 68 (91,2 %)	0 kritische Fehler in beiden Läufen
Hotelfrühstück 2026 Aktuelle Steuer- & Kürzungsregeln	100 %	40 %	Größter Unterschied durch Fachreferenz
Bewirtungsbeleg Form- & Abzugsregeln	100 %	60 %	Klare Wirkung aktueller Nachweisregeln
Sicherheitsfälle Prompt-Injection & Freigabegrenzen	100 %	100 %	Prüft Pfadregeln, trennt Fachwissen kaum

In beiden Varianten bestanden die Sicherheitsfälle. Das belegt in diesem Test nicht, dass die Sicherheitsregeln unnötig wären. Die Testnotizen halten vielmehr fest, dass diese Fälle vor allem die gemeinsamen Pfadregeln prüfen und den Mehrwert der Fachreferenzen kaum voneinander trennen.

Die Grenzen sind wichtig. Die Läufe fanden nicht auf der Zielplattform statt. Pro Testfall gab es nur einen Lauf. Schwankungen durch Wiederholungen wurden nicht gemessen. Die Ergebnisse sind deshalb ein nützlicher Entwicklungsbefund, aber kein allgemeiner Leistungsnachweis und keine Freigabe für den Echtbetrieb.

05 Was sich aus dem Promptdesign lernen lässt

Nr.	Kernprinzip für prüfbare KI-Assistenten
01	**Rolle und Befugnisse gehören in die Grundanweisung.** Es reicht nicht, am Ende um Vorsicht zu bitten, wenn der Ablauf zuvor bereits eine Entscheidung oder Aktion suggeriert.
02	**Komplexe Aufgaben brauchen eigene Pfade.** Auslöser, Ausschlüsse, Pflichtangaben und Prüfregeln sind für Rechnungen anders als für Protokolle.
03	**Fehlende Angaben sind ein Ergebnis.** Ein gutes System muss nicht jede Lücke schließen. Es muss sie zuverlässig benennen.

Nr.	Kernprinzip für prüfbare KI-Assistenten
04	**Veränderliche Fakten brauchen aktuelle Quellen.** Ein Modellgedächtnis ist kein Ersatz für ein gepflegtes und nachvollziehbares Fachwissen.
05	**Ausgabe und Freigabe sind Teil der Qualität.** Ein Entwurf, der nicht senden, buchen oder freigeben darf, lässt Verantwortung beim zuständigen Menschen.
06	**Tests müssen die richtige Frage prüfen.** Wenn Sicherheitsregeln in beiden Varianten gleich bleiben, zeigt ein Vergleich mit und ohne Fachwissen nicht, wie gut diese Sicherheitsregeln funktionieren.

„Ein Assistent wird nicht vertrauenswürdig, weil er selten irrt, sondern weil man sieht, worauf er sich stützt und wo er aufhört.“

