

KI in der Finanzfunktion.

Wie Finanzteams KI auswählen, einführen und im Alltag nutzen.

Was in diesem Leitfaden steht.

.	Vorwort	3
01	KI schafft Mehrwert auf zwei Ebenen	4
02	Der richtige Prozess entscheidet	6
03	Daten müssen nicht erst sauber sein	8
04	Selbst bauen oder kaufen	9
05	KI-nativ oder Add-on	12
06	KI-Kosten verhalten sich wie Energiekosten	14
07	Datenschutz und Governance	15
08	Von reaktiver zu proaktiver KI	16
09	Die ersten 90 Tage	17
A1	Über Orcha	18
A2	Sollten Sie selbst bauen?	19
A3	Mini-Glossar & Quellen	20
A4	Prompt-Vorlagen	21

Ein kurzer Leitfaden für den Einstieg.

Der Markt für KI bewegt sich schnell — fast täglich neue Modelle, Tools und Versprechen. In diesem Lärm ist für ein Finanzteam schwer zu erkennen, was wirklich zählt.

Die Chancen sind real und gross. Trotzdem kommen bisher nur wenige Teams über erste Versuche hinaus zu einer Umsetzung, die im Alltag wirklich funktioniert. Woran es meist hakt, sind die Entscheidungen davor: welcher Prozess zuerst, selbst bauen oder kaufen, welcher Anbieter, welche langfristigen Kosten.

Dieser Leitfaden bündelt unsere 35 Beiträge aus der Orcha-Community zu diesen Fragen — kompakt und zum Einstieg gedacht. Er ersetzt kein Beratungsprojekt; er ordnet die Fragen, die Sie ohnehin beantworten müssen — egal, ob Sie sich am Ende für Orcha, einen Wettbewerber oder Eigenentwicklung entscheiden.

Neun Kapitel. Vom Konzept bis zur Einführung. Am Ende jedes Kapitels eine Frage, mit der Sie Ihre Situation prüfen können.

— *Das Orcha-Team*

KI schafft Mehrwert auf zwei Ebenen.

Auf der einen wird jede Person schneller. Auf der anderen übernehmen Agenten ganze Prozesse. Beide lohnen sich, der Hebel ist aber unterschiedlich.

Die erste Ebene ist die individuelle Produktivität: Excel-Formeln, Präsentationen, Vertragsanalysen, Entwürfe für E-Mails. Gemeint sind Werkzeuge wie ChatGPT, Copilot oder Claude, die einzelne Personen bei ihrer Arbeit unterstützen. Diese Ebene ist sofort verfügbar und der natürliche Einstieg — das ganze Team kann innerhalb weniger Tage beginnen — siehe Schnellstart unten.

Die zweite Ebene ist die Prozessautomatisierung: Agenten, die Rechnungseingang, Bankabgleich oder Monatsabschluss durchgängig übernehmen. Hier liegt das grössere Potenzial, weil diese Prozesse weitgehend von selbst laufen und ein ganzes Team entlasten, statt nur eine einzelne Person schneller zu machen. Sie ist zugleich anspruchsvoller: Sie braucht mehr Vorbereitung und Fachverständnis, bevor sie zuverlässig läuft. Dieser Leitfaden behandelt vor allem diese zweite Ebene.

SCHNELLSTART FÜR PERSÖNLICHE PRODUKTIVITÄT

Vier Dinge, mit denen ein Finanzteam diese Woche anfangen kann

- 1 Excel:** Claude for Excel oder der Excel-Agent in Microsoft 365 Copilot arbeiten direkt in der Tabelle — Formeln erklären, Modelle bauen, Abweichungen kommentieren.
- 2 Präsentationen:** Für Foliensätze lohnt sich der agentische Modus (Claude Cowork, PowerPoint-Agent in Copilot) statt des einfachen Chats — er baut ganze Präsentationen aus Ihren Daten.
- 3 Über Apps hinweg:** Desktop-Agenten wie Claude Cowork greifen auf lokale Dateien zu und erledigen mehrstufige Aufgaben über Excel, Word und Outlook hinweg; Microsoft bietet ähnliche Funktionen im Copilot.
- 4 Datenschutz zuerst:** DSGVO-konform sind nur Business- und Enterprise-Accounts. Dort wird nicht auf Ihren Daten trainiert, und Zero Data Retention lässt sich je nach Vertrag aktivieren. Private Accounts eignen sich nicht für Geschäftsdaten (mehr in Kapitel 07).

Aus der Praxis: Wer KI fest in den Arbeitsalltag einbaut, wird deutlich produktiver als jemand, der sie nur gelegentlich für einzelne Fragen nutzt. Bewährt hat sich, zwei oder drei interne Champions zu benennen, die KI intensiv nutzen und ihr Wissen im Team weitergeben — etwa in einem kurzen All-Hands oder einem Hackathon. Mit etwas Anleitung kommt die Akzeptanz meist schnell, und die Produktivität steigt im ganzen Team.

KI ist ein Teammitglied.

Wer KI wie eine Suchmaschine benutzt, bekommt meist oberflächliche Ergebnisse. Wer sie wie ein neues Teammitglied am ersten Arbeitstag behandelt — fähig, aber ohne Kontext — bekommt in der Regel deutlich bessere Ergebnisse.

Moderne KI (LLM) ist ein sehr fähiges Teammitglied am ersten Arbeitstag. Sie hat viel gelesen, denkt schnell, arbeitet rund um die Uhr. Aber sie kennt Ihren Kontextrahmen, Ihre Mandanten und Ihre internen Begriffe nicht. Ohne diesen Kontext fällt das Ergebnis meist unbrauchbar aus.

Eine knappe Anfrage liefert generisches Material. Mit einem Absatz Kontext — wer Sie sind, was Ihr Unternehmen macht, welche Daten vorliegen — wird daraus ein Entwurf, mit dem das Team weiterarbeiten kann. Das Ergebnis hängt weniger vom Modell ab als davon, wie gut Sie die Aufgabe beschreiben.

Die Gewohnheit, auf die es ankommt

Der grösste Hebel liegt oft weniger im "richtigen Tool" als in der Arbeitsweise: *vor* jeder Aufgabe einmal fragen, ob KI dabei helfen kann — nicht erst, wenn man feststeckt. Wer das zur Gewohnheit macht, holt einen grossen Teil der möglichen Produktivitätsgewinne. Als Faustregel: zuerst KI, manuelle Arbeit nur dort, wo die KI nicht weiterkommt.

Mit der Zeit verschiebt das auch die Rolle des Teams — weg von der Verarbeitung, hin zum Urteil. Die Arbeit besteht dann weniger darin, Belege zu erfassen, als zu entscheiden, welche Fälle Aufmerksamkeit brauchen und wo die KI korrigiert werden muss.

BEISPIEL · ZWEI ANSÄTZE, GLEICHE FRAGESTELLUNG

Im Controlling fragt jemand: "Schreib mir eine Abweichungsanalyse für Q2." Output: generisch, mit Excel-Hinweisen, ohne Geschäftsbezug.

Zweiter Versuch mit Kontext: "Wir sind ein Maschinenbauer, 80 Mio. € Umsatz, SKR 03. Im Anhang: unser Q2-Plan-Ist und die Abweichungsanalyse des Vorquartals. Vergleiche beide Quartale, erkläre die drei grössten Abweichungen mit ihren wahrscheinlichen Treibern und formuliere sie so, dass sie der Aufsichtsrat ohne Rückfragen versteht." Ergebnis: drei nachvollziehbare Treiber, mit Anschlussfragen an die Kostenstellen-Verantwortlichen.

SELBST-CHECK

Drei Indikatoren für "KI als Default"

- 1 Sie öffnen Claude oder ChatGPT, *bevor* Sie Excel öffnen.
- 2 Sie lassen E-Mails von der KI vorformulieren, statt sie selbst zu tippen.
- 3 Wenn die KI mal nicht erreichbar ist, machen Sie Pause, anstatt manuell zu arbeiten.

Der richtige Prozess entscheidet.

Bevor Sie Tools vergleichen, klären Sie eine andere Frage: Welcher Prozess in Ihrem Finanzteam verdient die Automatisierung überhaupt? Dafür reichen eine 2x2-Matrix und ein 15-Minuten-Test.

Damit wechselt der Blick von der ersten Ebene — Einzelne, die mit KI schneller arbeiten — auf die zweite: Prozesse, die ein Agent für das ganze Team übernimmt. Sie beginnt mit der Auswahl des richtigen Prozesses.

Die Prozessauswahl wiegt schwerer als die Technik. Komplexe Aufgaben mit vielen Sonderfällen und viel Urteilsvermögen eignen sich anfangs selten; ein strukturierter, monatlich wiederkehrender Prozess meist deutlich besser.

Zwei Bedingungen müssen *gleichzeitig* erfüllt sein, damit ein Prozess für KI taugt:

- **Zeitlicher Aufwand:** Mehr als rund 5 Stunden pro Woche im Team.
- **KI-Qualität:** Ein 15-Minuten-Test mit echten Daten und Dokumenten zeigt, wie gut die KI das heute schon kann — am besten mit dem stärksten Modell des Anbieters und einem ausführlichen, gut formulierten Prompt. Wenn Sie die KI mehrfach korrigieren müssen, ist der Task meist nicht geeignet.

Die Auswahl-Matrix

Die schnelle Übung: Notieren Sie zuerst alle Prozesse, die Ihnen einfallen, und ordnen Sie jeden anhand dieser beiden Dimensionen ein. Mit den Prozessen im Feld "Sofort starten" geht es auf der nächsten Seite weiter.

	KI-QUALITÄT NIEDRIG	KI-QUALITÄT HOCH
ZEIT HOCH	<p>Beobachten. Grosser Schmerz, aber KI noch nicht reif. Halbjährlich erneut testen.</p>	<p>Sofort starten. Hoher Schmerz, KI bereit. Hier liegt der grösste Hebel.</p>
ZEIT NIEDRIG	<p>Nicht priorisieren. Kein Hebel, keine Reife. Ignorieren.</p>	<p>Quick Win. Klein, aber verlässlich. Hilft dem Team, Vertrauen aufzubauen.</p>

Farblich hervorgehoben ist das Feld mit dem grössten Hebel: hoher Schmerz, KI bereit — hier zuerst ansetzen.

Sechs Fragen, die zeigen, ob ein Prozess KI-tauglich ist.

Stellen Sie diese sechs Fragen für jeden Prozess, den Sie in der Matrix oben rechts einsortiert haben. Fünf oder mehr "Ja": klarer Kandidat. Drei bis vier: lohnt eine genauere Prüfung. Weniger als drei: lassen Sie ihn manuell.

1. **Zeitlicher Aufwand:** Beansprucht der Prozess im Team mehr als 5 Stunden pro Woche?
2. **Wiederholbarkeit:** Läuft er regelmässig in ähnlicher Form ab (täglich, wöchentlich, monatlich)?
3. **Praxistest bestanden:** Liefert die KI bei echten Dokumenten in 15 Minuten ein brauchbares Ergebnis?
4. **Fehlertoleranz:** KI liefert nicht in 100% der Fälle korrekte Ergebnisse. Geeignet sind Prozesse, bei denen gelegentliche Fehler verkraftbar sind — oder bei denen ohnehin ein Mensch das Ergebnis abschliessend prüft.
5. **Datenverfügbarkeit:** Kommen Sie an die nötigen Daten heran — über eine Integration, eine API oder einfachen Upload? Ob die Daten strukturiert sind, spielt hier keine Rolle; darum kümmert sich die KI (mehr in Kapitel 03).
6. **Datenschutz geklärt:** Ist klar, welche Daten die KI sehen darf und welche nicht?

BEISPIEL · AP-AUTOMATISIERUNG

Rechnungseingang bei einem Mittelständler: 500 Rechnungen im Monat, je 12 bis 15 Minuten manuell — über 100 Stunden im Team. Alle sechs Fragen mit "Ja". Das gehört in die obere rechte Ecke.

Gegenbeispiel: Four-Way-Match mit vielen Sonderfällen oder eine Cashflow-Prognose mit zu vielen Variablen. Auf dem Papier wirken sie wie Kandidaten — in der Praxis frisst der Long Tail der Ausnahmen jeden Effizienzgewinn auf.

FRAGE AN IHR TEAM

Welche drei Prozesse würden Sie heute in die obere rechte Ecke einordnen?

Wenn Ihr Team diese Frage nicht in 15 Minuten beantworten kann, machen Sie eine einwöchige Bestandsaufnahme: 3–4 Mitarbeitende protokollieren ihre wiederkehrenden Aufgaben mit Zeitaufwand.

Daten müssen nicht erst sauber sein.

Diese Anforderung stammt aus der Welt vor LLMs — heute bremst sie Projekte, die längst laufen müssten.

Die alte Logik lautet: erst Datenqualität, dann Automatisierung. Bei klassischen Systemen mit starren Schemata stimmte das. Bei KI-basierten Systemen stimmt es (meist) nicht mehr — und ist oft ein unnötiger Grund, der Innovationsprojekte blockiert.

Moderne LLMs lesen "Reisek.", "Reisekosten" und "Travel Expenses" als dieselbe Kategorie. Sie kommen mit gemischten Sprachen, unterschiedlichen Formaten und unsystematisch gepflegten Stammdaten zurecht. Datenqualität bleibt relevant. Aber sie entsteht durch *die KI selbst*, nicht durch eine vorgelagerte Bereinigung im ERP.

KI als Strukturschicht über bestehender Infrastruktur

1. **Datenquellen anbinden** — ERP, Excel, E-Mail, Bank-Portale. Keine Migration nötig.
2. **KI liest, interpretiert, strukturiert** — format- und sprachunabhängig, mit einem Konfidenzwert je Extraktion.
3. **Saubere Daten landen in einer Datenbank** — abfragbar über SQL, BI-Tools oder weitere KI.
4. **Bestehende Systeme bleiben** — DATEV, SAP, Ihr Kontenrahmen. Nichts wird ersetzt.

BEISPIEL · EIN 780-MIO.-USD-LEHRSTÜCK

Ein US-Einzelhändler aus der Fortune 500 stoppte 2024 eine geplante 780-Mio.-USD-ERP-Migration. Stattdessen legte das Unternehmen eine KI über die bestehenden Systeme (Quelle: McKinsey 2025). Das funktioniert nicht nur bei Fortune-500-Budgets. Auch ein Mittelständler mit DATEV und Excel-Historien kann eine KI über bestehende Systeme legen — ohne Systeme und Prozesse zu ändern.

FRAGE ZUR ALTEN REIHENFOLGE

Welcher KI-Pilot wartet bei Ihnen auf eine Datenmigration?

Fällt Ihnen ein solcher Prozess ein, prüfen Sie ihn als nächsten Pilot-Kandidaten.

Selbst bauen oder kaufen.

Rund 31% der internen Software-Projekte bleiben im Zeitplan, im Budget und im vereinbarten Umfang. Auf jeden Euro Entwicklungskosten kommen vier bis fünf Euro Wartung über fünf Jahre (Gartner).

Eigenentwicklungen für Standardprozesse bringen meist keinen Wettbewerbsvorteil — es ist *Routinearbeit, die jeder im Markt leisten muss*. Werner Vogels (Amazon) nennt das "undifferentiated heavy lifting": Aufwand, der Sie nicht vom Wettbewerb unterscheidet. Für ein Finanzteam heisst das in den meisten Fällen: Kaufen Sie ein fertiges Tool und bauen Sie nur, was Sie vom Wettbewerb unterscheidet.

Das heisst nicht, dass Eigenentwicklung nie Sinn macht. Standardprozesse lassen sich mit eigenem Code oft um 20–40% verbessern, in Einzelfällen mehr. Aber Edge Cases, die Sie heute nicht sehen, kosten schnell mehr als die Software-Lizenz des externen Anbieters. Den Rechnungseingang beispielsweise kann man mit ordentlicher Quote selbst digitalisieren - aber eine 90%+ Dunkelbuchungsziffer bei internationalen Belegen, plus Four-Way-Match, Betrugserkennung, automatische Lieferantenanlage und automatische Lieferantenkommunikation — das ist ein eigenes Software-Projekt.

	SELBST BAUEN	KAUFEN
Time-to-Value	9–18 Monate	Wochen
Erfolgsquote	31%	Bewährtes Produkt am Markt
Wartung pro Jahr	15–20% der Initialkosten	in der Lizenz enthalten
Weiterentwicklung	Jedes neue Feature selbst bauen	Anbieter liefert laufend neue Funktionen
Regulatorik (GoBD, DSGVO)	Eigene Verantwortung	Anbieter
5-Jahres-TCO	4–5× Initialkosten, nicht planbar	Linear, planbar

Vier Fragen, die über Eigenbau entscheiden.

Die ersten beiden sind die Gartner-Bedingungen — Eigenbau lohnt nur, wenn beide zutreffen; die anderen beiden kommen aus der Praxis:

1. Ist es Kern Ihres Wettbewerbsvorteils? (Gartner)

Wenn drei Wettbewerber dasselbe Problem lösen müssen, ist es selten ein Vorteil — eher Betriebsaufwand, den alle tragen.

2. Gibt es keine passende Lösung am Markt? (Gartner)

Erst wenn der Markt nichts Passendes bietet, lohnt der Blick auf Eigenbau.

3. Haben Sie die Kapazität, das System über Jahre zu warten — und regulatorische Änderungen selbst umzusetzen?

Eine GoBD-Anpassung pro Jahr, eine DSGVO-Auditierung, eine API-Migration alle 18 Monate binden dauerhaft Kapazität im Team.

4. Was könnte Ihr Entwicklungsteam stattdessen bauen — etwas, das Umsatz bringt?

Opportunitätskosten sind unsichtbar, aber real: Jede Stunde Entwicklungsarbeit an einem Buchhaltungssystem fehlt einem Produkt, für das Kunden zahlen.

Ein zusätzliches Risiko liegt im Wissen einzelner Personen: Geht die Schlüsselperson, fängt der Wissensaufbau oft bei null an — anders als bei einem Anbieter, der das System dauerhaft pflegt.

WENN ES SCHIEFGEHT

Lidl stoppte 2018 ein eigenes SAP-Ersatzprojekt nach rund **500 Mio. €** — ohne lauffähiges System. **Nike** verlor 2001 **100 Mio. USD** Umsatz, nachdem ein hausgebautes Forecast-System Lagerbestände falsch berechnet hatte.

SCHRIFTLICHER BELEGTEST

Können Sie beide Gartner-Bedingungen für jeden geplanten Eigenbau schriftlich belegen?

Fünf Fragen, die jeder Einkaufs-Plan beantworten muss.

Auch "kaufen" verlangt eine eigene Due Diligence. Die Lizenzgebühr sagt wenig über die Folgekosten, und eine Demo zeigt selten, was wirklich funktioniert.

1. Passt sich der Anbieter an Ihre Welt an — oder umgekehrt?

DATEV-Belegtransfer und Buchungsdatenservice im Standard (DATEVconnect für tiefere Integration), offene API zu Ihrem ERP (SAP, proALPHA, abas, Sage, Microsoft Dynamics), Anbindung an Ihren Kontenrahmen, Zusammenarbeit mit Ihrer Steuerberatung. Wenn die Antwort "Sie müssen Ihre Prozesse an unser Tool anpassen" lautet, sind die Change-Management-Kosten meist ein Vielfaches der Software-Lizenzkosten.

2. Gehören die Daten Ihnen — auch nach einem Wechsel?

Wo liegen die Daten (EU vs. USA), wer sind die Sub-Prozessoren, gibt es einen dokumentierten Exportpfad in einem nutzbaren Format? Ein schriftlich zugesicherter Exit-Prozess ist nötig.

3. Ist das System API-first?

Gerade weil sich die KI-Werkzeuge derzeit rasant weiterentwickeln, darf kein Lock-in entstehen — der Fokus liegt auf offenen Schnittstellen. Immer mehr KI-Tools brauchen direkten Zugriff auf Ihre Daten; eine offene, dokumentierte API erleichtert spätere Erweiterungen und verringert die Bindung an einen Anbieter.

4. Ist das System auditierbar — und der Anbieter belastbar?

Nachvollziehbare Entscheidungen mit Audit-Log und Konfidenzwert, Verfahrensdokumentation für die Betriebsprüfung, AVV nach DSGVO, EU-AI-Act-konforme Risikoeinstufung.

5. Was kostet der Betrieb über die Lizenz hinaus?

ERP-Anbindung, Schulung im Team, Umbau interner Prozesse, Anbieter-Risiko (Übernahme, Insolvenz, Produkteinstellung) — und der teuerste Posten: Lock-in. Was kostet ein Wechsel in fünf Jahren?

BEISPIEL · GESCHLOSSENES VS. OFFENES SYSTEM

Anbieter A ist ein geschlossenes System: Daten gelangen nur über die Oberfläche hinein und hinaus, ein vollständiger Export ist bestenfalls "auf Anfrage" möglich. Jede Erweiterung hängt davon ab, was der Anbieter freischaltet.

Anbieter B ist API-first: Alle Daten lassen sich über eine dokumentierte API lesen und schreiben, dazu ein MCP-Server (Model Context Protocol) für den direkten Zugriff durch KI-Werkzeuge. Beide kosten gleich viel — aber B fügt sich in Ihre Systemlandschaft ein und lässt einen späteren Wechsel offen.

FRAGE AN DIE ANBIETER AUF IHRER SHORTLIST

"Was passiert mit unseren Daten und unserer Integration, wenn wir in fünf Jahren wechseln?"

Bitten Sie zugleich um eine ausführliche API-Dokumentation — und um eine klare Liste, was sich nur über die Nutzeroberfläche erledigen lässt, aber nicht über die API. Eine konkrete schriftliche Antwort hilft hier am meisten.

KI-nativ oder Add-on.

KI entwickelt sich rasant weiter. Wer kauft, braucht Software, die für die Zukunft mit KI gebaut ist. Von aussen ist oft schwer zu erkennen, wie stark ein Anbieter im KI-Bereich wirklich ist. Im Kern gibt es zwei Muster: einzelne KI-Funktionen als Add-on — oder eine KI-native Anwendung.

Etablierte Anbieter haben ihre Plattformen über Jahre verfeinert und in den letzten Monaten KI nachgerüstet. KI-native Anbieter starten mit den KI-Modellen als Fundament. Auf den ersten Blick sehen beide oft gleich aus. Der Unterschied ist grundlegend — und prägt, wie gut ein Anbieter mit der KI-Entwicklung mithält.

Ein einfacher Test zeigt, was Sie vor sich haben: Entfernen Sie gedanklich alle KI-Funktionen. Läuft das System mit alter Template-Logik weiter, sitzt KI nur obendrauf; bleibt nichts übrig, trägt KI das System. Im Alltag macht sich der Unterschied so bemerkbar:

	TEMPLATE-BASIERT (KLASSISCH)	KI-NATIV (LLM-ZENTRIERT)
Einführung	Mehrere Monate Lernphase	Funktioniert sofort
Neuer Lieferant	Neues Template anlegen	Funktioniert sofort
Layout-Änderung	Template anpassen	Modell erkennt es
Fremdsprache	Eigenes Template pro Sprache	Modell versteht es
Freitext-Felder	Mühsam, fehleranfällig	Stärke des Modells
Kontierung	Regelwerke pflegen	Modell schlägt vor
Wartung über Zeit	Steigt mit Lieferanten	Sinkt mit Volumen

Ein zweites Kriterium ist der Modell-Lock-in. Das Modell selbst wird zunehmend zur Commodity — der Vorsprung liegt im Harness um das Modell herum: Workflows, Integrationen, Schnittstellen, Kontrollen. Die nützliche Frage an einen Anbieter lautet darum: Lässt sich das zugrunde liegende Modell austauschen, ohne den ganzen Workflow neu zu bauen?

40% erkannt heisst nicht 40% weniger Arbeit.

OCR ist eine ältere Generation von KI: Sie liest Zeichen, stützt sich aber auf Regeln und Vorlagen, die jemand pflegen muss. Sie sagt Ihnen, was sie gelesen hat — aber nicht, wie sicher sie sich ist. Also müssen Sie jeden Beleg prüfen. Bei 40% korrekter Erkennung sparen Sie real vielleicht 20% Zeit.

Moderne LLMs sind einen grossen Schritt weiter. Sie erkennen auch Fälle ohne festes Regelwerk und liefern zu jeder Extraktion einen Konfidenzwert; geprüft wird nur noch, wo das System unsicher ist. So sparen Sie real bis zu 90% der Zeit — statt 20%.

Mehr als nur Lesen

Moderne KI liest nicht nur — sie prüft Lieferanten, gleicht Rechnung, Bestellung und Wareneingang ab (Four-Way-Match) und schlägt die Kontierung vor. Ein guter Anhaltspunkt: Übernimmt die KI Aufgaben, die sonst erfahrene Fachkräfte erledigen, steckt mehr dahinter als reine Texterkennung.

BEISPIEL · EIN KI-NATIVES AP-SYSTEM

Orcha besteht den Test mit "Nein": ohne LLM gäbe es das Produkt nicht. Nicht der einzige KI-native Anbieter am Markt, aber einer, an dem sich das Prinzip ablesen lässt. 80–90% der Rechnungen laufen vollkontiert und gebucht durch — ohne menschliche Nachbearbeitung. Der Branchenschnitt liegt bei rund 33% (Ardent Partners 2025).

FRAGE AN DIE ANBIETER AUF IHRER SHORTLIST

Was passiert, wenn ein neuer Lieferant ein unbekanntes Rechnungsformat schickt?

Muss jemand ein Template anlegen — oder erkennt das System es von selbst und legt den neuen Lieferanten gleich automatisch an? Die Antwort zeigt, ob KI das Fundament ist oder nur punktuell unterstützt.

KI-Kosten verhalten sich wie Energiekosten.

Zwanzig Jahre SaaS haben uns an feste Lizenzgebühren gewöhnt. Bei KI ändert sich das: Die Kosten steigen mit der Nutzung.

Ein Power-User kostet schnell ein Vielfaches eines Light-Users. Eine einheitliche Flatrate rechnet sich für Anbieter darum oft nicht mehr. Drei Marktbewegungen folgen daraus:

1. **Flatrates verschwinden** oder werden gedeckelt; echtes "Unlimited" gibt es bei KI-nativen Tools 2026 kaum noch.
2. **Hybrid-Modelle werden Standard** — Basisgebühr plus verbrauchsabhängiger Anteil.
3. **Verbrauchstransparenz wird Verhandlungssache** — verlangen Sie ausdrücklich vertraglichen Einblick in den tatsächlichen Verbrauch.

Die neue Kennzahl: Kosten je Transaktion

Monatssummen sagen wenig aus, sobald die Kostenbasis variabel wird. Entscheidend wird: Was kostet die einzelne KI-Aktion — ein Beleg, eine Analyse, ein Dokument? Vereinbaren Sie die Abrechnung möglichst nach einer Grösse, die Sie selbst steuern und vorhersehen können — etwa pro Beleg oder pro Dokument, nicht pro Token. Wie viele Belege Sie im Monat haben, wissen Sie; wie viele Tokens dabei anfallen, lässt sich kaum vorhersagen.

Für die Budgetierung heisst das: Der laufende KI-Betrieb ist Betriebsaufwand (Opex), keine einmalige Investition. Den Wechsel von einmaligen Anschaffungen zu laufenden Gebühren kennen Finanzteams schon aus der SaaS-Welt; neu ist vor allem, dass die Kosten zusätzlich mit der Nutzung schwanken.

BEISPIEL · POWER-USER VS. LIGHT-USER

Power User: Eine Person im Controlling, die über eine Woche mit Hilfe von KI eine Varianzanalyse in Excel aufbaut, verbraucht viel — intensive, kontextreiche Sitzungen summieren sich.

Light User: Wer KI nur für einzelne Fragen und Entwürfe nutzt, bleibt deutlich darunter — zwischen beiden liegt schnell das Drei- bis Vierfache. Dasselbe Tool, dieselbe Lizenz, sehr unterschiedlicher Verbrauch.

VIER PUNKTE FÜR JEDEN KI-VERTRAG

Was sich in der nächsten Verhandlung festhalten lässt

1. **Klare Preismetrik:** pro Dokument oder pro Outcome — eine Grösse, die Sie verstehen und selbst steuern, nicht pro Token.
2. **Kostenbremsen und Alarmschwellen:** Ab wann werden Sie benachrichtigt? Ab wann greift eine automatische Sperre?
3. **Verbrauchseinblick in Echtzeit:** Tages- oder Stundenauflösung — keine Monatsrechnung.
4. **Hochrechnung auf echten Daten:** Vor Unterschrift anhand realer Volumina kalkulieren und die erwartete Kostenspanne schriftlich festhalten — das beugt späteren Überraschungen vor.

Datenschutz und Governance.

Mit dem richtigen Business-Account, signiertem AVV, EU-Datenresidenz und einer internen Policy lässt sich DSGVO-konformer KI-Einsatz in der Regel in Tagen aufsetzen.

Kostenlose und private KI-Accounts sind für ein Finanzteam meist die falsche Wahl.

Free- und Privat-Accounts (OpenAI Free, ChatGPT Plus, Claude Free) trainieren standardmässig auf Ihren Eingaben. Business-Accounts (OpenAI Enterprise, Claude Team/Enterprise, Microsoft 365 Copilot) sind vertraglich davon ausgenommen. Der Wechsel auf einen Business-Account schliesst dieses Risiko aus; die Mehrkosten fallen dabei kaum ins Gewicht.

Vier Datenklassen, vier Freigaben

Ordnen Sie Ihre Daten vier Stufen zu — jede Stufe bestimmt, welcher Account sie sehen darf:

- **Öffentlich** — jeder Account, auch private.
- **Intern** — nur Business-Accounts mit signiertem AVV.
- **Vertraulich** — Business-Accounts mit EU-Datenresidenz.
- **Streng vertraulich** — nur in einer selbst gehosteten Umgebung. Ein eigener API-Schlüssel (Bring-Your-Own-Key) hilft hier nicht: Die Daten laufen weiterhin über die API des Anbieters.

Dazu kommt der Zugriffsweg — einzelner Upload, dauerhaftes Projekt oder fester Connector. Je sensibler die Daten, desto enger der Zugriff.

10-PUNKTE-CHECK FÜR GO-LIVE

Bevor das erste Dokument durch die KI läuft

- Signierter Auftragsverarbeitungsvertrag (AVV)
- Kein Training auf Ihren Daten — schriftlich bestätigt oder bei den grossen Cloud-Anbietern im Admin-Tool aktiviert
- EU-Datenresidenz aktiviert (wo verfügbar)
- Datenschutz-Folgenabschätzung dokumentiert (DSFA)
- Interne KI-Policy verabschiedet
- Subprozessoren-Liste regelmässig geprüft
- KI-Kompetenzschulung durchgeführt (EU AI Act, seit Feb. 2025 Pflicht)
- Betriebsrat informiert, falls vorhanden (Mitbestimmung nach §87 BetrVG / ArbVG bei Leistungsbewertung)
- GoBD-Verfahrensdokumentation aktualisiert (für Betriebsprüfung)
- AT: DSB-Meldung geprüft, BAO §§131/132 abgedeckt

SHADOW AI · INOFFIZIELLE KI-NUTZUNG IM TEAM

Ist der offizielle Weg zu restriktiv oder zu umständlich, greifen Mitarbeitende zum privaten Account. Ein Verbot löst das selten — was hilft, ist ein offizieller Weg, der mindestens so gut funktioniert wie der inoffizielle.

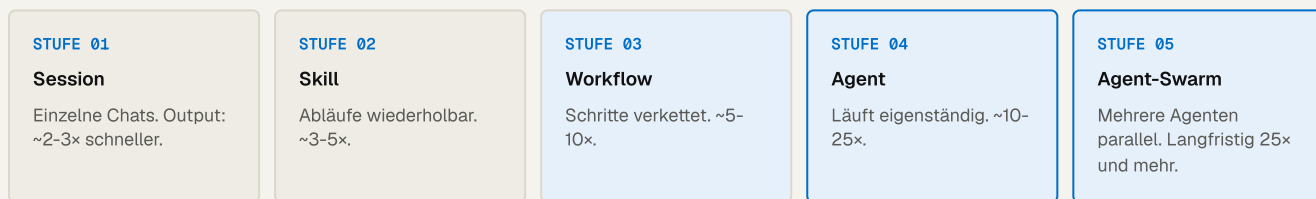
Von reaktiver zu proaktiver KI.

Reaktive KI wartet auf Prompts. Proaktive KI beobachtet laufend die Datenströme und meldet sich, bevor jemand fragt.

Was das konkret heisst, zeigen drei typische Prozesse:

PROZESS	REAKTIV	PROAKTIV
Kreditoren (AP)	Rechnung wird erfasst und gebucht.	KI meldet Dubletten und Abweichungen zur Bestellung, bevor gebucht wird.
Bestellung & Budget	Überschreitung fällt im Monatsabschluss auf.	KI warnt schon bei der Bestellanforderung, wenn sie das Budget übersteigt.
Lieferanten & Einkauf	Konditionen werden bei Vertragsverlängerung geprüft.	KI erkennt laufend Einsparpotenziale und Konzentrationsrisiken.

DIE 5 REIFEGRADE · VOM CHAT ZUM AUTOPILOT



BEISPIEL · WAS WIR BEI ORCHA BAUEN

Bei Orcha laufen Agenten rund um die Uhr: Belege erfassen, Abweichungen melden, Cashflow-Signale früh anzeigen. Stufe 4, teilweise Stufe 5 — eines von mehreren Systemen am Markt, das produktiv im Einsatz ist.

LETZTE FRAGE

Auf welcher Stufe steht Ihr Team heute — und was ist der nächste konkrete Schritt?

Zwei Muster sehen wir oft: Wer Stufe 5 sofort erzwingt, explodiert in Komplexität; wer auf Stufe 1 bleibt, übersieht, dass Stufe 2 nur einen Skill entfernt liegt.

Die ersten 90 Tage.

Die vorigen Kapitel liefern die Grundlagen — dieser Fahrplan macht daraus ein konkretes Vorgehen, mit dem aus dem Mandat ein greifbarer Projektplan entsteht.

Woche 1 — Das Mandat definieren

Ein einseitiges Mandatspapier: Welcher Prozess (genau einer) ist der Kandidat? Was wird in 90 Tagen gemessen? Wer ist die treibende Kraft aus Ihrem Team, wer aus IT, wer aus Audit/Recht? Was ist explizit nicht im Scope? Aus dieser Seite wird später die Zwischenstandsmeldung an die Geschäftsführung. Unabhängig vom Piloten lohnt sich sofort der grösste Schnellgewinn: dem ganzen Team Zugriff auf einen KI-Assistenten geben und es mehrfach darauf schulen.

Woche 2 — Pilot aufsetzen

Setzen Sie den Piloten an einem Prozess auf: 30 Beispielbelege, Datenschutz-Check signiert vor dem ersten Datensatz. Erfolgsmetrik vor dem ersten Beleg festlegen: Wie viele Belege laufen ohne Hand durch, wie lange dauert die Nachbearbeitung? Aussagekräftig sind dabei vor allem Zahlen aus den eigenen Daten, weniger die Demo.

Wochen 3–4 — Entscheiden, anbinden, schulen

Anbieter anhand der Pilot-Zahlen wählen, nicht anhand der Präsentation. Parallel: AVV unterschreiben, EU-Datenresidenz aktivieren, Team schulen, interne Policy verabschieden, Verfahrensdokumentation an die Steuerberatung versenden. Der Go-Live-Check aus Kapitel 7 gehört vor den ersten produktiven Beleg, nicht danach.

Wochen 5–13 — Produktiv, messen, ausweiten

Erster Prozess produktiv, mit Human-in-the-Loop bei niedriger Konfidenz. Im wöchentlichen Review:

- Welche Ausnahmen tauchen auf?
- Welche dieser Ausnahmen lassen sich als Skill abbilden?

Nach 90 Tagen folgt der Zwischenbericht — Kosten je Transaktion, gewonnene Zeit im Team, der nächste priorisierte Prozess.

BEISPIEL · 90 TAGE IN EINEM MITTELSTANDBETRIEB

Ein Maschinenbauer mit 80 Mitarbeitenden startete im Februar mit AP. Tag 14: Mandatspapier unterschrieben. Tag 30: Pilot mit echten Belegen getestet. Tag 60: produktiv mit Human-in-the-Loop. Tag 90: erster Zwischenbericht. 78% der Rechnungen laufen ohne Nachbearbeitung durch. Das Controlling gewinnt zwei Tage pro Monat zurück.

DREI DINGE NACH 90 TAGEN

Was Sie der Geschäftsführung zeigen können

- 1 Ein Prozess produktiv, mit konfidenzbasiertem Workflow.
- 2 Drei Kennzahlen mit Vergleichswerten — vorher / nachher.
- 3 Ein priorisierter Fahrplan für die nächsten drei Prozesse.

Orcha — KI-native Finanz-Software für den Mittelstand.

Orcha baut Finanz-Software, die ohne Large Language Models nicht möglich war. Belege, Rechnungen, Kontenabgleiche und Auswertungen laufen weitgehend automatisch — über bestehende ERP-Systeme hinweg, ohne Migration.

Die Plattform deckt den Tagesbetrieb der Finanzabteilung ab: Kreditoren- und Debitorenbuchhaltung, Dokumentenmanagement, Freigabe-Workflows sowie Analyse und Reporting. Jeder Beleg wird bis auf die Positionsebene erfasst — mit Sachkonto, Kostenstelle, Steuersatz und einem Konfidenzwert, der zeigt, wo jemand noch einmal hinsehen sollte. Die Zahlen stehen in Echtzeit zur Verfügung, nicht erst zum Monatsabschluss.

Orcha bindet sich an die vorhandene Systemlandschaft an: DATEV, ERP- und SAP-Umgebungen, Excel, BI-Werkzeuge, E-Mail sowie Teams und Slack, dazu eine offene API. Datenschutz und Nachvollziehbarkeit sind eingebaut — DSGVO-konform, GoBD-gerecht, mit durchgängigem Audit-Trail.

Diesen Leitfaden gibt es, weil viele Finanzteams die passenden Werkzeuge kennen, bei der Entscheidung aber oft allein dastehen. Die Werkzeuge in diesem Leitfaden funktionieren unabhängig davon, ob Sie sich am Ende für Orcha entscheiden.

Mehr erfahren

getorcha.com — Produkt, Demo, Case Studies, weitere Beiträge der Orcha-Community.

Drei KI-Assistenten, drei Anwendungsbereiche. Code ist oft nicht nötig.

Die Geschäftsführung und die kaufmännische Leitung müssen nicht programmieren. Hilfreich ist aber, zu erkennen, *wann* ein Prozess mit einem KI-Assistenten wie Claude direkt erledigt ist, wann ein No-Code-Skill reicht und wann sich ein Coding-Assistent, z. B. Claude Code, lohnt.

"Selbst entwickeln" heisst 2026 selten noch "Software bauen". Die drei realistischen Stufen sind:

- Einen KI-Assistenten wie Claude direkt fragen.
- Einen Skill bauen, der den Prozess wiederholbar macht.
- Einen Coding-Assistenten wie Claude Code einsetzen — für etwas, das eigenständig läuft.

Code ist die komplexeste der drei Stufen und wird zugleich oft überschätzt. Der Grund: Ein lokales Tool ist schnell gebaut, der Aufwand kommt danach. Sobald ein Team damit arbeitet, kommen Zugriffsrechte, Audit-Trail bei Änderungen, Versionsstände und Datenschutz dazu — aus dem schnellen Prototyp wird ein Software- und Wartungsprojekt. Gerade ohne technischen Hintergrund unterschätzt man diesen Teil leicht, und deshalb ist Code für die meisten Standardfälle selten die richtige Wahl.

STUFE	WANN SINNVOLL	BEISPIEL
Claude direkt Chat oder Excel-Add-in	Einmalige oder seltene Aufgaben, Ad-hoc-Analysen, Briefings, Dokumentenprüfung.	Vertragsanalyse mit IFRS-16-Check, Abweichungskommentar im Quartalsbericht.
Skill / Projekt No-Code SOP	Wiederkehrende Aufgaben, die das ganze Team identisch erledigen soll.	Monatlicher Reporting-Kommentar nach festem Schema, standardisierte Lieferanten-Onboarding-Checkliste.
Claude Code "Vibe Coding"	Wenn ein eigenständiges Artefakt entstehen soll: Dashboard, Skript, kleines Tool.	Budget-Ist-Dashboard mit Filter, Abgleichsskript für 200 Buchungen.

ENTSCHEIDUNGS-REGEL

Ohne technischen Hintergrund lohnt sich eigener Code fast nie.

Wer dieselbe Aufgabe wiederholt erledigt, spart mit einem Skill Zeit — und No-Code-Skills decken heute viel ab, bis hin zu Zeitplänen und Connectoren in der Cloud. Eigener Code lohnt sich dagegen selten, solange ihn niemand im Team dauerhaft pflegen kann.

Mini-Glossar & Quellen.

Elf Begriffe, die in diesem Leitfaden vorkommen

LLM (Large Language Model)

Das Modell hinter Claude, ChatGPT oder Gemini. Erzeugt Text anhand statistischer Muster — nicht durch Datenbankabfragen.

Token

Worteinheit, die KI-Kosten bestimmt. 200 Seiten ≈ 80 000 Tokens.

Kontextfenster

Wie viel Text die KI "auf einmal" lesen kann. Reicht heute für ganze Verträge oder Quartalsberichte.

Prompt

Die Anweisung an die KI. Bessere Prompts bringen meist bessere Ergebnisse — oft wichtiger als das gewählte Modell.

Halluzination

Wenn die KI eine plausibel klingende falsche Antwort gibt. Keine Fehlfunktion, sondern eine Eigenschaft des Modells. Kritische Angaben immer gegenprüfen.

RAG (Retrieval-Augmented Generation)

Methode, mit der KI auf Ihre internen Dokumente zugreift, bevor sie antwortet.

Agent

KI, die eigenständig mehrere Schritte ausführt — nicht nur antwortet, sondern handelt.

Konfidenzwert

Masszahl dafür, wie sicher die KI eine Extraktion einstuft. Grundlage für Workflows mit menschlicher Prüfinstanz. Schwelle pro Use Case konfigurierbar.

MCP (Model Context Protocol)

Offener Standard, über den KI-Werkzeuge kontrolliert auf Ihre Systeme zugreifen — Daten lesen und schreiben.

Zero Data Retention

Vertragliche Zusage, dass der Anbieter Ihre Eingaben nicht speichert. Auf Business- und Enterprise-Accounts je nach Vertrag aktivierbar.

Vibe Coding

Software entsteht, indem man sie in normaler Sprache beschreibt; ein KI-Coding-Assistent baut sie. Gut für kleine Tools, nicht für tragende Systeme.

Quellen

Standish Group (Erfolgsquote interner IT-Projekte) · Gartner (TCO, Build-vs-Buy-Regel) · Ardent Partners 2025 (AP-Automatisierungsgrade) · McKinsey 2025 (ERP-Migration vs KI-Layer, Monatsabschluss) · APQC, Billentis, PYMNTS, IOFM (AP-Benchmarks) · Bessemer (KI-Margen) · Anthropic Wall Street Prep Benchmark (Excel-Modellierung) · BaFin / FINRA 2026 (Governance) · EU AI Act (Feb 2025 in Kraft, Aug 2026 High-Risk-Enforcement).

Alle 35 Originalbeiträge der Orcha-Community: getorcha.com/de/community

Sechs Vorlagen, die ein Finanzteam morgen nutzen kann.

Diese Vorlagen lassen sich direkt übernehmen: eigene Daten und Kontext einfügen, [Platzhalter] ersetzen. Sie funktionieren in jedem gängigen KI-Assistenten und in der Excel-Integration.

01 · MAHNUNG MIT SKONTO-HINWEIS

"Du übernimmst die Kreditorenbuchhaltung (AP). Verfasse eine freundliche, aber bestimmte E-Mail an [Lieferant] zur [X] Tage überfälligen Rechnung [Nr.] über [Betrag]. Weise auf das Zahlungsziel [Datum] und den Skonto-Vorteil bei Zahlung bis [Datum] hin. Ton: respektvoll und lösungsorientiert, die Geschäftsbeziehung bleibt erhalten. Max. 120 Wörter, mit Betreffzeile und klarer Handlungsaufforderung."

02 · SKONTO-OPTIMIERUNG

"Du übernimmst die Treasury-Analyse. Anhang: Liste der offenen Verbindlichkeiten mit Skonto-Bedingungen und Fälligkeiten. Ermittle bei einem Hausbankzins von [X]%, welche Rechnungen sich mit Skonto zu zahlen lohnen. Gib eine nach annualisierter Rendite sortierte Tabelle aus (Rechnung, Skontosatz, Rendite, Empfehlung) und nenne den gesamten Skonto-Gewinn."

03 · ABWEICHUNGSANALYSE

"Du übernimmst die FP&A-Analyse. Anhang: Plan-Ist-Vergleich Q[N] je Kostenstelle. Identifiziere die fünf grössten Abweichungen (absolut und in %), erkläre je Abweichung die wahrscheinlichen Treiber und markiere, was einmalig und was strukturell ist. Schliesse mit drei gezielten Rückfragen an die Kostenstellen-Verantwortlichen."

04 · VERTRAGSANALYSE · IFRS 16

"Du übernimmst die bilanzielle Beurteilung. Anhang: Mietvertrag. Prüfe, ob er nach IFRS 16 als Right-of-Use-Asset zu aktivieren ist. Gib aus: relevante Klauseln mit Fundstelle, Laufzeit und Verlängerungsoptionen, deine Einstufung mit Begründung sowie offene Punkte zur Klärung. Kennzeichne jede Annahme ausdrücklich."

05 · LIEFERANTEN-RISIKOSCORE

"Du übernimmst die Procurement-Analyse. Anhang: Top-20-Lieferanten mit Umsatzanteilen und Kategorien. Bewerte je Lieferant Konzentrationsrisiko, Ersetzbarkeit am Markt und geografisches Risiko auf einer Skala von 1 (niedrig) bis 5 (kritisch), mit kurzer Begründung. Gib eine sortierte Tabelle aus und hebe die drei kritischsten Abhängigkeiten hervor."

06 · AUFSICHTSRATS-VORLAGE

"Du übernimmst die Rolle der Stabsstelle. Anhang: Monatsabschluss [Monat]. Verfasse eine zweiseitige Aufsichtsrats-Vorlage in der Reihenfolge GuV → Bilanz → Cashflow → die drei wichtigsten Themen des Monats. Ton sachlich-knapp, je Abschnitt die Abweichung zum Plan in einem Satz. Erklärende Sätze statt Aufzählungsketten."

Ein guter Prompt bringt oft mehr als ein stärkeres Modell. Die Vorlagen funktionieren mit Claude, ChatGPT, Gemini oder Copilot — aber nur mit echtem Kontext und auf eigenen Daten. Mehr Vorlagen im Beitrag "KI-Shortcuts & Tipps" auf getorcha.com/de/community.