

KI-Atlas: Welche KI passt zu welcher Aufgabe?

45 Kategorien, 95 Produkte, 123 Bewertungen — Entscheidungsgrundlage für Werkzeugauswahl.

Stand 24.09.2026 · Gestaltung nach CI-Bibel KI-Agenten.shop

Was dieser Vergleich ist — und was nicht

Bewertet wurde die dokumentierte Eignung aus offiziellen Herstellerangaben: Funktionspassung 50 Prozent, Steuerbarkeit 25 Prozent, Workflow-Anschluss 25 Prozent, je 1 bis 5 Punkte. **Es ist kein eigener Produkttest** und keine Messung von Ausgabequalität, Geschwindigkeit oder Kosten je brauchbarem Ergebnis.

Kein universeller KI-Sieger, keine Vollerhebung aller weltweit existierenden Produkte, keine eigenen Produkttests oder kontospezifische Preis-/Lizenzprüfung. Medizin, Finanzen, Personal, Architektur und industrielle Anwendungen benötigen eigene Fachabnahmen. Weitere Kategorieübersicht und konkrete Projektkombinationen stehen in der HTML-Fassung.

Zahlen der Erhebung

- **45** Kategorien
- **95** Produkte und Modellfamilien
- **123** Einzelbewertungen
- Erhebungsstand: 19.09.2026

01 · Allgemeine KI-Assistenten & LLMs

ChatGPT und Claude gemeinsam in die engere Wahl nehmen; Gemini besonders bei vorhandenen Google-Abläufen. Kein hier gemessener Intelligenzsieger.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
ChatGPT	★★★★★ 5	Gemischte Wissensarbeit und fertige Arbeitsdateien	Modus und gebuchte Werkzeuge beeinflussen das Ergebnis.	Tarif und Nutzungslimits
Claude	★★★★★ 5	Lange Texte, Dokumente und mehrstufige Wissensarbeit	Quellen, Berechtigungen und Resultate müssen kontrolliert werden.	Tarif und Nutzungslimits
Gemini / Gemini API	★★★★½ 4.5	Multimodale Anwendungen und Google-naher Einsatz	Konkretes Modell, Modalitäten und Zugang getrennt prüfen.	Modellabhängige API-Nutzung / App-Tarif
Mistral Vibe, vormal's Le Chat	★★★★ 4	Alternative Chat-/Coding-Umgebung	Nicht ungeprüft mit früheren Le-Chat-Tarifen gleichsetzen.	Tarif / API separat
Grok	★★★★ 4	Dialog und zeitnahe Themen als Recherche-Einstieg	Aktuelle Informationen sind nicht automatisch verlässlich.	Tarif / API separat
DeepSeek API	★★★★ 4	Eigene API-Anwendungen	Selbst zu bauende Anwendung; Qualität und Datenweg separat prüfen.	Nutzungsabhängige API

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Qwen-Modellfamilie	★★★★ 4	Eigene Auswahl eines passenden offenen Modells	Gewichte, Lizenz und Hardwarebedarf pro konkretem Modell prüfen.	Hosting und Rechenleistung / API

Vergleichstest: Ein identisches Briefing mit Quellen, Widersprüchen und gewünschter Datei an alle Kandidaten geben.

02 · Webrecherche & Quellenarbeit

Perplexity für die Suche im offenen Web; Gemini Notebook für einen definierten eigenen Quellenbestand. Die Empfehlungen betreffen unterschiedliche Aufgaben.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Perplexity	★★★★½ 4.5	Schneller Einstieg mit verlinkten Originalquellen	Zitate müssen die konkrete Aussage tatsächlich tragen.	Tarif und Recherchekontingent
Gemini Notebook / NotebookLM	★★★★★ 5	Fragen zu einem abgegrenzten Quellenpaket	Vollständigkeit und Qualität hängen vom Quellenbestand ab.	Kontingent / Workspace-Zugang
ChatGPT	★★★★ 4	Recherche mit anschließender Ausarbeitung	Modus und gebuchte Werkzeuge beeinflussen das Ergebnis.	Tarif und Nutzungslimits

Vergleichstest: Zehn überprüfbare Aussagen verlangen; jede Originalquelle öffnen und richtige Belegstelle zählen.

03 · Wissenschaft & Literatúrauswertung

Elicit für strukturierte Extraktion und Review-Abläufe; Consensus für Literaturfragen und Zitationswege.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Elicit	★★★★★ 5	Studien systematisch sichten und tabellarisch extrahieren	Studienauswahl und Extraktionen benötigen fachliche Prüfung.	Tarif / Nutzungspool
Consensus	★★★★½ 4.5	Literatur finden und Zusammenhänge nachverfolgen	Eine Zusammenfassung ersetzt keine Bewertung der Studienmethodik.	Tarif / Suchkontingente

Vergleichstest: Eine bekannte Studienmenge vorgeben; fehlende Studien, falsche Zahlen und unpassende Schlussfolgerungen zählen.

04 · Programmierung & Repository-Arbeit

Codex und Claude Code für abgegrenzte Projektaufgaben testen; Cursor für Arbeit im KI-Editor; Copilot bei GitHub-zentriertem Team.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
OpenAI Codex	★★★★★ 5	Repo-Aufgaben bis zu überprüften Änderungen	Produktionsreife entsteht erst durch echte Tests und Review.	Tarif / Nutzungsgrenzen
Claude Code	★★★★★ 5	Terminal- und Repository-Workflows	Umgebungszugriff, Tests und Freigaben müssen eingerichtet sein.	Tarif / API-Nutzung
Cursor	★★★★★ 5	Interaktive Entwicklung im KI-Editor	Modellwechsel und Agentenschleifen verändern Kosten und Verhalten.	Tarif / Modellverbrauch

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
GitHub Copilot	★★★★★ 5	Bestehende Entwicklerteams im GitHub-Workflow	Repository-Regeln und menschliches Code-Review bleiben erforderlich.	Tarif / Premium-Nutzung

Vergleichstest: Drei echte Issues mit versteckten Regressionstests lösen lassen; nur geprüfte Änderungen zählen.

05 · Websites, Apps & Prototypen

v0 für visuell kontrollierte Weboberflächen; Lovable oder Replit für einen integrierten App-Einstieg. Kein automatisches Produktionssiegel.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
v0	★★★★★ 5	Web-UI mit sichtbarer Bearbeitung und Deployment	Authentifizierung, Datenzugriff und Zahlungslogik separat testen.	Generierung plus Hosting
Lovable	★★★★½ 4.5	Schneller Full-Stack-Prototyp	Eine funktionierende Vorschau beweist keine sichere Produktion.	Generierung plus Cloud-Dienste
Replit Agent	★★★★★ 5	App entwickeln und betreiben in einer Umgebung	Laufende Hostingkosten und Zugriffsregeln sind Teil des Projekts.	Agentennutzung plus Hosting

Vergleichstest: Login, Rollentrennung, mobile Ansicht, Export und Wiederherstellung desselben Mini-Projekts testen.

06 · Bildgenerierung & visuelle Konzepte

Midjourney für Stilfindung vormerken; FLUX für integrierbare Bildabläufe; Firefly bei bestehender Adobe-Arbeit. Ästhetischer Sieger bleibt offen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Midjourney	★★★★½ 4.5	Moodboards und visuelle Konzeptsuche	Bildästhetik wurde hier nicht in einem Blindtest verglichen.	Tarif / Rechenkontingent
FLUX / Black Forest Labs	★★★★★ 5	Bildgenerierung als eigener technischer Workflow	Lizenz und lokale Nutzbarkeit unterscheiden sich je Variante.	API / Hosting / Lizenz
Adobe Firefly	★★★★★ 5	Generierung mit anschließender Designbearbeitung	Bedingungen von Partnermodellen nicht mit Adobe-Modellen gleichsetzen.	Tarif / Generierungskredite
ChatGPT	★★★★ 4	Bilder im dialogischen Arbeitsablauf	Modus und gebuchte Werkzeuge beeinflussen das Ergebnis.	Tarif und Nutzungslimits
Ideogram	★★★★★ 5	Motive mit bearbeitbarer Typografie	Typografie und Markenmaterial vor Ausgabe kontrollieren.	Tarif / API
Google Images	★★★★½ 4.5	Bildgenerierung innerhalb eines Google-Workflows	Zugang und Bearbeitungsfunktionen vom verwendeten Produkt abhängig.	Produktzugang / API

Vergleichstest: Gleiche sechs Motive mehrfach erzeugen; Details, Identität, Text und erforderliche Nacharbeit beurteilen.

07 · Design, Werbung & Präsentationen

Canva für wiederkehrende Markenformate, Gamma für einen ersten strukturierten Foliensatz. Die Endabnahme erfolgt am exportierten Dokument.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Canva AI	★★★★★ 5	Social Designs und wiederverwendbare Vorlagen	Individuelle Gestaltung braucht weiterhin redaktionelle Entscheidungen.	Tarif / KI-Kontingente
Gamma	★★★★½ 4.5	Präsentationen aus einem strukturierten Briefing	Exportierte Folien auf Umbruch, Schrift und Bearbeitbarkeit prüfen.	Tarif / Generierungskontingent
Adobe Firefly	★★★★½ 4.5	Individuelle Assets für Adobe-Designabläufe	Bedingungen von Partnermodellen nicht mit Adobe-Modellen gleichsetzen.	Tarif / Generierungskredite

Vergleichstest: Ein gemeinsames Branding-Briefing in zehn Folien und drei Social-Formate übersetzen.

08 · Cinematic Video & generierte Szenen

Für unsere Produktion Higgsfield als Arbeitsoberfläche plus gezielte Modellwahl testen. Runway als alternative Produktionsumgebung; Veo für Szenen mit Audio; Seedance für referenzreiche Vorgaben. Kein Blindtest-Sieger.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Higgsfield	★★★★★ 5	Gebündelte kreative Werkzeuge und filmische Gestaltung	Plattform und darunter verwendetes Modell getrennt vergleichen.	Tarif / Credits / AP
Runway	★★★★★ 5	Generierung und Weiterbearbeitung im Produktionsablauf	Ein brauchbarer Langfilm benötigt Auswahl, Schnitt und Kontinuitätskontrolle.	Tarif / Generierungskredi

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Google Veo	★★★★½ 4.5	Szenen mit generiertem Bild und Ton	Zugang, Version, Clipgrenzen und Steuerung hängen vom Produktweg ab.	Produktzugang / API-Verbrauch
Dreamina Seedance	★★★★★ 5	Referenzgestützte Szenen und Figurenführung	Anbieter-Demos belegen keine konstante Figurenqualität im eigenen Film.	API / Plattform-Credits
Kling AI	★★★★★ 4	Bewegungsreferenzen; aktuelle Oberfläche vor Einsatz prüfen	Produktseite nicht lesbar; Funktionsbeleg aus März 2026, heutige Tarife und Zugang offen.	Aktueller Tarif nicht verifiziert
Luma	★★★★★ 4	Alternative Text-/Bild-zu-Video-Pipeline	Referenztreue und Weiterbearbeitung im konkreten Workflow testen.	Tarif / Credits
Sora (Auslaufprodukt)	★ 1	Nur bestehende Inhalte und Migration prüfen	Keine Empfehlung für einen neu aufzubauenden Produktionsworkflow.	Keinen neuen Workflow darauf aufbauen

Vergleichstest: Drei zusammenhängende Einstellungen derselben Person erzeugen:
Close-up, Gehbewegung, Interaktion. Kontinuität und Kosten aller Fehlversuche
erfassen.

09 · Avatare, Sprecher & Schulungsvideos

HeyGen für mehrsprachige Presenter; Synthesia für standardisierte Unternehmensschulung.
Keines ersetzt automatisch eine szenisch inszenierte Filmproduktion.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
HeyGen	★★★★★ 5	Presenter und mehrsprachige Videoerklärungen	Sprachvariante, Aussprache und Identitätsfreigaben prüfen.	Tarif / Minuten / API separat
Synthesia	★★★★★ 5	Wiederholbare Trainings- und Unternehmensvideos	Szenischer Film und standardisiertes Training sind verschiedene Aufgaben.	Tarif / Videominuten / Teamumfang

Vergleichstest: Dasselbe 60-Sekunden-Skript mit Namen, Zahlen und Gesten in Deutsch und Vietnamesisch abnehmen.

10 · Videoschnitt & komplette Filmfertigstellung

Resolve oder Premiere für den vollständigen Film; CapCut für kurze Social-Formate; Descript für sprachdominierte Inhalte.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
DaVinci Resolve	★★★★★ 5	Gesamter Film mit Farbe, Ton und Effekten	KI-Funktionsumfang der Free- und Studio-Ausgabe unterscheiden.	Free / Studio / geeignete Hardware
Adobe Premiere	★★★★★ 5	Professioneller Schnitt im Adobe-Umfeld	Einarbeitung und Projektorganisation bleiben Teil der Arbeit.	Abonnement / Zusatzverbrauch
CapCut	★★★★½ 4.5	Schnelle Social-Video-Bearbeitung	Funktionen, Export und Verfügbarkeit je Plattform und Tarif prüfen.	Tarif / funktionsabhängige Credits
Descript	★★★★½ 4.5	Interview, Podcast und textbasierter Schnitt	Fachwörter und automatisch entfernte Passagen kontrollieren.	Tarif / Medienkontingente

Vergleichstest: Ein Projekt mit zehn Clips, Musik, Untertiteln und zwei Formaten vollständig exportieren.

11 · Shorts, Untertitel & Wiederverwertung

OpusClip für Vorschläge aus langen Aufnahmen; CapCut für gezielte Nachbearbeitung. Ein Virality-Score ist keine Reichweitengarantie.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
OpusClip	★★★★½ 4.5	Automatische Clipauswahl aus Langmaterial	Automatische Auswahl kann Aussagen aus dem Zusammenhang reißen.	Tarif / Verarbeitungsminuten
CapCut	★★★★★ 5	Manueller Feinschnitt und Social-Aufbereitung	Funktionen, Export und Verfügbarkeit je Plattform und Tarif prüfen.	Tarif / funktionsabhängige Credits
Descript	★★★★★ 4	Sprachbasierte Auswahl und Untertitel	Fachwörter und automatisch entfernte Passagen kontrollieren.	Tarif / Medienkontingente

Vergleichstest: Drei Clips aus einem Interview gewinnen; Kontexttreue, erste Sekunden, Untertitel und Safe Areas prüfen.

12 · Video-Restauration & Upscaling

Topaz als Spezialkandidat, Resolve als Alternative im vorhandenen Schnittprojekt. Nicht jedes Motiv profitiert von generierten Details.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Topaz Video	★★★★★ 5	Gezielte Verbesserung vorhandener Videodateien	Hinzugerechnete Details sind keine wiedergewonnenen Tatsachen.	Lizenz / Cloud-Verbrauch / Hardware
DaVinci Resolve	★★★★½ 4.5	Verbesserung direkt im Finishing-Workflow	KI-Funktionsumfang der Free- und Studio-Ausgabe unterscheiden.	Free / Studio / geeignete Hardware

Vergleichstest: Gesichter, Text und schnelle Bewegung vor/nach Bearbeitung bei normaler Wiedergabe vergleichen.

13 · Songs, Rap & Musik mit Gesang

Suno ist unsere erste Auswahl für einen vollständigen Rap mit vorgegebenen Lyrics. Google Lyria und Eleven Music als Gegenkandidaten testen. Udio erhält wegen des dokumentierten Exporthindernisses für diesen Zweck nur einen Stern.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Suno	★★★★★ 5	Vollständiger Song mit eigenen Lyrics und Weiterbearbeitung	Aussprache, musikalische Struktur und Nutzung des konkreten Tarifs prüfen.	Tarif / Songkontingent / Studio-Zugang
Eleven Music	★★★★ 4	Musikproduktion als Alternative im Audio-Ökosystem	Stems, Export, Sprache und Lizenz im gewünschten Tarif nachprüfen.	Tarif / Generierungsverbrauch
Google Lyria	★★★★½ 4.5	Alternative Musikgenerierung im Google-Umfeld	Zugang, Export und Nutzungsbedingungen der konkreten Oberfläche prüfen.	Produktzugang / Kontingent
Udio	★ 1	Nur Erkundung innerhalb des Dienstes; externe Lieferung blockiert	Für extern abzuliefernde Musik aktuell Ausschlusskriterium bis Export bestätigt ist.	Exportfreigabe zuerst klären

Vergleichstest: Gleiche acht Zeilen und Stilvorgaben verwenden; Texttreue, Aussprache, Hook, Songende und herunterladbare Dateien prüfen.

14 · Hintergrundmusik, Instrumental & Sounddesign

SOUNDRAW für anpassbare Hintergründe; AIVA für kompositorische Arbeit und MIDI; Stable Audio für eigene Audiopipelines.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
SOUNDRAW	★★★★★ 5	Video-Unterlegung mit steuerbarer Länge und Energie	Hintergrundnutzung und Songdistribution tarifbezogen unterscheiden.	Creator-/Artist-/Enterprise Tarif
AIVA	★★★★★ 5	Komposition und anschließende musikalische Bearbeitung	Für fertige Rap-Vocals nicht dieselbe Aufgabe wie Suno.	Tarif / Lizenzumfang / Downloads
Stable Audio	★★★★ 4	Technische Audioexperimente und eigene Integration	Modelllizenzen, Hardware und Produktoberfläche getrennt prüfen.	Hosting / API / Produktzugang
Eleven Music	★★★★ 4	Musikalische Varianten für Medieninhalte	Stems, Export, Sprache und Lizenz im gewünschten Tarif nachprüfen.	Tarif / Generierungsverbrauch

Vergleichstest: Eine 30-Sekunden-Werbung vertonen: gewünschter Spannungsbogen, sauberes Ende und Platz für Sprache.

15 · Voiceover & synthetische Stimmen

ElevenLabs zuerst für isoliertes Voiceover testen; HeyGen, wenn zugleich ein Presenter gebraucht wird.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
ElevenLabs Text to Speech	★★★★★ 5	Mehrsprachiges Voiceover mit Sprechregie	Gute Aussprache in einer Sprache belegt keine Qualität in allen Sprachen.	Tarif / Zeichen oder Credits
HeyGen	★★★★ 4	Stimme als Teil eines Presenter-Videos	Sprachvariante, Aussprache und Identitätsfreigaben prüfen.	Tarif / Minuten / API separat

Vergleichstest: Gleicher Text mit Eigennamen, Zahlen und Emotion; Muttersprachler beurteilt Verständlichkeit und Natürlichkeit.

16 · Videoübersetzung & Dubbing

HeyGen bei gewünschter Lippensynchronisation, Rask als spezialisierte Alternative. Quelle und Übersetzung müssen dieselbe Aussage transportieren.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
HeyGen	★★★★★ 5	Übersetztes Presenter-Video	Sprachvariante, Aussprache und Identitätsfreigaben prüfen.	Tarif / Minuten / API separat
Rask AI	★★★★½ 4.5	Lokalisierung bestehender Videoaufnahmen	Namen, Fachbegriffe und Lippenbewegung im Zielmarkt abnehmen.	Tarif / Minuten

Vergleichstest: Ein Zwei-Sprecher-Video mit Fachbegriffen lokalisieren; Timing, Namen und Sprecherwechsel prüfen.

17 · Textübersetzung & internationale Inhalte

DeepL für fokussierte Übersetzung; Claude für redaktionelle Anpassung und Erklärung von Alternativen. Übersetzung und freie Werbeadaptation unterscheiden.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
DeepL	★★★★★ 5	Dokument- und Textübersetzung	Zielsprachliche Redaktion bleibt bei Werbung und Fachsprache nötig.	Tarif / API / Dokumentgrenzen
Claude	★★★★½ 4.5	Transkreation und begründete Formulierungsvarianten	Quellen, Berechtigungen und Resultate müssen kontrolliert werden.	Tarif und Nutzungslimits

Vergleichstest: Deutsch, Englisch und Vietnamesisch mit Zahlen, Höflichkeitsformen und einem verbindlichen Glossar vergleichen.

18 · Meetings, Transkription & Aufgaben

Fathom für kompakte Meetingnotizen, Fireflies für nachgelagerte Team-/CRM-Arbeit, Otter als weiteren Kandidaten. Die konkrete Meetingsprache entscheidet mit.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Fathom	★★★★½ 4.5	Notizen und Zusammenfassung nach Gesprächen	Privatsphäre und Genauigkeit nicht aus der Produktdemo ableiten.	Tarif / Teamfunktionen
Fireflies	★★★★★ 5	Meetingwissen mit nachgelagerten Teamprozessen	Automatisch erkannte Aufgaben auf richtige Verantwortliche prüfen.	Tarif / Plätze / KI-Nutzung
Otter	★★★★½ 4.5	Transkription und Meetingauswertung	Unterstützte Sprache, Sprecher und Zustimmung pro Meeting prüfen.	Tarif / Minuten / Plätze

Vergleichstest: Eine freigegebene Aufnahme mit drei Sprechern prüfen: Namen, Zahlen, Entscheidungen und Aufgaben.

19 · Office, E-Mail & internes Wissen

Die bestehende Arbeitsumgebung gibt den Ausschlag: Microsoft 365, Google Workspace oder Notion. Ein Wechsel nur wegen KI kann mehr Aufwand als Nutzen erzeugen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Microsoft 365 Copilot	★★★★★ 5	Arbeit im Microsoft-365-Unternehmenskontext	Vorhandene Zugriffsrechte und passende Lizenzen entscheiden.	Plätze / Lizenz / Zusatzfunktionen
Gemini in Google Workspace	★★★★★ 5	Gmail-, Docs- und Drive-basierte Zusammenarbeit	Produktfunktionen und Administratorfreigaben unterscheiden sich.	Workspace-Tarif / Kontingente

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Notion AI	★★★★★ 5	Wissens- und Projektarbeit in Notion	Veraltete interne Quellen liefern weiterhin veraltete Antworten.	Tarif / Agentennutzung

Vergleichstest: Eine Aufgabe mit Mail, Dokument und Termin durchführen; Rechte und Quellenalter gezielt variieren.

20 · Datenanalyse, Tabellen & Berichte

ChatGPT für kombinierte Analyse und Bericht; Julius als spezialisierte Alternative; Copilot bei Excel-zentriertem Team. Rechenrichtigkeit ist noch nicht gemessen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
ChatGPT	★★★★★ 5	Datei auswerten und Bericht mit Grafiken erstellen	Modus und gebuchte Werkzeuge beeinflussen das Ergebnis.	Tarif und Nutzungslimits
Julius AI	★★★★★ 4	Tabellenbezogene Analyseaufgaben	Keine geprüfte statistische Überlegenheit; Rechenweg nachprüfen.	Tarif / Nutzung
Microsoft 365 Copilot	★★★★★ 5	Analyse in bestehender Microsoft-Arbeit	Vorhandene Zugriffsrechte und passende Lizenzen entscheiden.	Plätze / Lizenz / Zusatzfunktionen

Vergleichstest: Eine Tabelle mit fehlenden Werten, Dubletten und absichtlich irreführender Korrelation auswerten.

21 · Automatisierung & systemübergreifende Agenten

n8n für technisch betreute individuelle Abläufe, Make für visuelle Integration, Zapier Agents für angebundene SaaS-Aufgaben. Das Modell ist nur ein Bauteil.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
n8n	★★★★★ 5	Individuelle Prozesse mit Kontrolle über Ausführungen	Selbsthosting verschiebt Betrieb und Wartung zu uns.	Hosting / Ausführungen / Modellkosten
Make	★★★★★ 5	Visuell aufgebaute Integrationsabläufe	Fehlerpfade, Schleifen und Operationen müssen geplant werden.	Tarif / Credits / Modellkosten
Zapier Agents	★★★★★ 5	Agenten mit vorhandenen SaaS-Verbindungen	Autonome Aktionen brauchen klar begrenzte Rechte.	Tarif / Aktivitäten / weitere Dienste

Vergleichstest: Anfrage zu CRM-Entwurf verarbeiten; Dublette, API-Ausfall, Wiederholung und menschliche Freigabe testen.

22 · Kundenservice, Chatbots & Dialogagenten

Fin bei vorhandener Supportorganisation; Voiceflow für individuell entworfene Dialoge; Copilot Studio im Microsoft-Unternehmen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Fin / Intercom	★★★★★ 5	Service-Agent in bestehenden Supportprozessen	Lösungsquote ist ohne eigene Ticketprüfung nicht übertragbar.	Vertrags- und nutzungsabhängig
Voiceflow	★★★★★ 5	Individuelle Dialoge mit API-Anbindung	Wissenspflege und Eskalation im echten Kundenfall testen.	Tarif / Credits / Modellkosten
Microsoft Copilot Studio	★★★★★ 5	Unternehmensagenten im Microsoft-Umfeld	Lizenzierung und Unternehmensumgebung beeinflussen den Aufwand.	Kapazität / Verbrauch / Lizenzen

Vergleichstest: 30 Anfragen einschließlich unbekannter Antwort, Beschwerde und Datenänderung; korrekte Eskalation zählen.

23 · Vertrieb, CRM & Kundenprozesse

HubSpot bevorzugt prüfen, wenn es bereits das CRM ist; Agentforce bei Salesforce. Kein CRM-Wechsel allein wegen eines Agentennamens.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
HubSpot Agent Hub	★★★★★ 5	KI-Arbeit auf bestehenden HubSpot-Kundendaten	Datenqualität und benötigte Hubs bestimmen den Nutzen.	Hub-Lizenzen / Plätze / Credits
Salesforce Agentforce	★★★★★ 5	Komplexere Salesforce-gestützte Kundenprozesse	Für kleine neue Teams kann die Einführung zu aufwendig sein.	Vertrag / Daten / Nutzung

Vergleichstest: Echte anonymisierte Beispieldaten: Bedarf zusammenfassen, nächsten Schritt vorschlagen, Dubletten vermeiden.

24 · Marketingtexte, Content & SEO/GEO

Jasper für markengesteuerte Teamproduktion; Surfer für Inhaltsplanung und Diagnosen; Claude als flexibler Redaktionsassistent. Kein Tool ersetzt eigene Recherche.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Jasper	★★★★★ 5	Wiederkehrende Markenkommunikation im Team	Markenkonsistenz ersetzt weder Faktenprüfung noch gute Strategie.	Tarif / Plätze / Vertragsumfang
Surfer	★★★★½ 4.5	SEO-Content-Audits und Themenplanung	Ein hoher Toolscore garantiert weder Google-Ranking noch Zitate.	Tarif / Projekte / Kontingente
Claude	★★★★½ 4.5	Redaktion, Struktur und Gegenargumente	Quellen, Berechtigungen und Resultate müssen kontrolliert werden.	Tarif und Nutzungslimits

Vergleichstest: Einen belegten Artikel produzieren; Fakten, Originalität, Suchintention und manuelle Nacharbeit bewerten.

25 · E-Commerce & Produktbilder

Photoroom für produktorientierte Bildabläufe; Canva für die daraus gebauten Anzeigenformate.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Photoroom	★★★★★ 5	Produktfreistellung und konsistente Bildvarianten	Produktdetails dürfen gegenüber dem echten Produkt nicht verändert werden.	Tarif / Batch- oder API-Verbrauch
Canva AI	★★★★½ 4.5	Produktkampagnen und Formatvarianten	Individuelle Gestaltung braucht weiterhin redaktionelle Entscheidungen.	Tarif / KI-Kontingente

Vergleichstest: Zehn echte Produkte: gleiche Farben, Form und Beschriftung, verschiedene Hintergründe und Formate.

26 · 3D, Spiele-Assets & Animation

Meshy und Tripo mit denselben Referenzen vergleichen. Blender bleibt ein separates Werkzeug für Bereinigung und Inszenierung, kein gleichartiger KI-Modellkandidat.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Meshy	★★★★★ 5	3D-Asset mit Textur und Anschlussbearbeitung	Topologie, Maßhaltigkeit und Animation vor Produktion prüfen.	Tarif / Generierungskredite
Tripo	★★★★½ 4.5	Alternative Bild-/Text-zu-3D-Erstellung	Produktionsfähige Geometrie ist nicht allein durch Vorschau belegt.	Tarif / Credits

Vergleichstest: Dasselbe Objekt generieren; Geometrie, Textur, Export, Rig und Verhalten im Zielformat prüfen.

27 · Lokale KI & eigene Modellumgebung

LM Studio für einen zugänglichen lokalen Einstieg; Ollama für eine technisch integrierte Modelllaufzeit. Beide sind nicht das Sprachmodell selbst.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
LM Studio	★★★★★ 5	Lokale Modellarbeit in einer Desktopumgebung	Nur lokaler Betrieb bedeutet nicht automatisch einen vollständig isolierten Ablauf.	Hardware / gewählter Dienst
Ollama	★★★★★ 5	Lokale Modelle als Bestandteil eigener Anwendungen	Lokalen und Cloud-Modus unterscheiden; Modelllizenz separat.	Hardware / Hosting / Cloud je Wahl

Vergleichstest: Gleiches konkretes Modell auf gleicher Hardware testen; Speicher, Zeit, Datenverkehr und Ergebnis vergleichen.

28 · OCR, Rechnungen & Dokumentenprozesse

Document AI und Textract nach vorhandener Cloud und Dokumentart auswählen. Ein extrahierter Wert ist noch keine freigegebene Buchung.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Google Document AI	★★★★★ 5	Strukturierte Dokumentenerfassung in Google Cloud	Feldgenauigkeit pro Dokumentart und Sprache prüfen.	Seiten / Prozessor / Cloud-Nutzung
Amazon Textract	★★★★★ 5	Dokumentenerfassung in einer AWS-Umgebung	Keine automatische Freigabe von Zahlungs- oder Antragsdaten.	Seiten / gewählte Analysefunktionen

Vergleichstest: 100 freigegebene Beispielseiten: Felder gegen Sollwerte prüfen; niedrige Sicherheit an Menschen geben.

29 · Lernen, Training & Wissensvermittlung

Khanmigo für pädagogisch angelegte Unterstützung; Gemini Notebook für das Lernen aus eigenen Unterlagen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Khanmigo	★★★★½ 4.5	Unterrichts- und Lernbegleitung	Verfügbarkeit, Sprache und Lehrplanbezug vor Einführung prüfen.	Programm / Tarif / Region
Gemini Notebook / NotebookLM	★★★★★ 5	Lernmaterial aus einem kontrollierten Quellenpaket	Vollständigkeit und Qualität hängen vom Quellenbestand ab.	Kontingent / Workspace-Zugang

Vergleichstest: Lernenden erst erklären lassen, dann Hinweis geben; Lernerfolg anhand neuer Aufgaben prüfen.

30 · Juristische Facharbeit

Harvey ist ein Fachsoftware-Kandidat für professionell betreute Abläufe. Die Sterne bewerten nur den dokumentierten Workflow, keine Rechtsrichtigkeit.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Harvey	★★★★ 4	Professionelle juristische Arbeitsabläufe	Fachliche und jurisdiktionsbezogene Abnahme erforderlich.	Unternehmensangebot

Vergleichstest: Jurist prüft identische Fälle, Fundstellen, Aktualität und fehlende Gegenargumente.

31 · Sicherheit & Medienauthentizität

Security Copilot für Sicherheitsarbeit im passenden Unternehmensstack; Resemble für Medienauthentizität. Die unterschiedlichen Zwecke nicht als direkten Zweikampf lesen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Microsoft Security Copilot	★★★★½ 4.5	Analyseunterstützung im Sicherheitsbetrieb	Eignung hängt von Telemetrie und bestehendem Sicherheitsstack ab.	Kapazität / Lizenzen
Resemble AI	★★★★ 4	Prüfung synthetischer beziehungsweise manipulierter Medien	Erkennungsergebnisse benötigen Fehlalarm- und Auslassungstests.	Produktbezogener Tarif / Angebot

Vergleichstest: Bekannte echte und manipulierte Fälle verwenden; Fehlalarme und übersehene Fälle getrennt zählen.

32 · Industrie, Simulation & Physical AI

Omniverse als Plattformkandidat; Auswahl erfordert einen eigenen technischen Proof of Concept, nicht eine allgemeine Chatbot-Rangliste.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
NVIDIA Omniverse	★★★★½ 4.5	Simulation als Grundlage industrieller KI-Anwendungen	Kein direkt vergleichbares Endnutzer-Chatprodukt.	Infrastruktur / Engineering / Lizenz

Vergleichstest: Eine reale Anlagen- oder Robotikaufgabe mit definierten Simulations- und Realitätstests abgrenzen.

33 · Podcasts & Audio-Restaurierung

Adobe Podcast für Sprachbereinigung; Descript für redaktionellen Schnitt. Gleiche Aufnahme vergleichen, nicht Demos.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Adobe Podcast	★★★★½ 4.5	Bereinigung von gesprochenem Audio	Starke Bereinigung kann Natürlichkeit und Sprachdetails verändern.	Tarif / Upload- und Längengrenzen

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Descript	★★★★★ 5	Textbasierter Schnitt einer Podcastfolge	Fachwörter und automatisch entfernte Passagen kontrollieren.	Tarif / Medienkontingente

Vergleichstest: Eine eigene Aufnahme mit Störgeräuschen bearbeiten; Verständlichkeit, Artefakte und Natürlichkeit blind bewerten.

34 · Architektur, Immobilien & Bauplanung

Forma für frühe Standort- und Gebäudeplanung. Bildgeneratoren eignen sich für Stimmung, nicht für belastbare Maße.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Autodesk Forma	★★★★½ 4.5	Standortanalyse und frühe Planung	Standortdaten und Übergabe an BIM fachlich prüfen; kein fertiger Bauantrag.	Lizenz / Projektumgebung
Adobe Firefly	★★★★½ 3.5	Unverbindliche gestalterische Visualisierung	Bedingungen von Partnermodellen nicht mit Adobe-Modellen gleichsetzen.	Tarif / Generierungskredite

Vergleichstest: Ein Grundstück mit echten Randbedingungen prüfen; Abweichungen, Datengrundlage und BIM-Übergabe protokollieren.

35 · Medizinische Dokumentation

Dragon Copilot ist ein Kandidat für professionell kontrollierte Dokumentationsabläufe. Kein Diagnostik-Ranking.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Microsoft Dragon Copilot	★★★★ 4	Klinische Notizen und Dokumentationsworkflow	Hier nur Dokumentationsworkflow bewertet, keine Diagnostik oder Behandlungsqualität.	Unternehmensangebot / Integration

Vergleichstest: Fachpersonal vergleicht erlaubte Testgespräche mit den Notizen: Auslassungen, falsche Angaben und Korrekturzeit.

36 · Personal, HR & Recruiting-Prozesse

Workday AI im Workday-Umfeld; Joule bei SAP. Systeme nach bestehender Prozesslandschaft auswählen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Workday AI	★★★★★ 5	HR-Arbeitsabläufe im Workday-System	Anwendungsfall und bestehende Unternehmensplattform sind ausschlaggebend.	Unternehmensangebot / Module
SAP Joule Assistants	★★★★★ 5	HR-Administration in passender SAP-Umgebung	Modul, Datenzugang und konkrete Rolle müssen zum vorhandenen SAP passen.	SAP-Lizenzen / Vertragsumfang

Vergleichstest: Test einer administrativen Anfrage vom Eingang bis zur richtigen Auskunft; keine automatische Bewerberentscheidung bewerten.

37 · Finanzen, Einkauf & ERP

SAP Joule und Workday nach vorhandenen Modulen vergleichen. Extraktion, Analyse und Freigabe bleiben getrennte Schritte.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
SAP Joule Assistants	★★★★★ 5	Einkaufs- und Finanzassistenz im SAP-Kontext	Modul, Datenzugang und konkrete Rolle müssen zum vorhandenen SAP passen.	SAP-Lizenzen / Vertragsumfang
Workday AI	★★★★★ 5	Finanzprozesse im Workday-Kontext	Anwendungsfall und bestehende Unternehmensplattform sind ausschlaggebend.	Unternehmensangebot / Module
Amazon Textract	★★★★½ 3.5	Dokumentenextraktion als vorgeschalteter Baustein	Keine automatische Freigabe von Zahlungs- oder Antragsdaten.	Seiten / gewählte Analysefunktionen

Vergleichstest: Einen Einkaufs- oder Finanzfall gegen Freigaberegeln und Sollwerte testen; falsche Buchungen gesondert zählen.

38 · Unternehmenssuche & internes Wissen

Glean für eine dedizierte Unternehmenssuche; Notion und Microsoft bei passender bestehender Wissenslandschaft.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Glean	★★★★★ 5	Wissen aus mehreren Unternehmensquellen	Quellenrechte, gelöschte Dokumente und Aktualisierung im Test prüfen.	Unternehmensangebot / Integrationen
Notion AI	★★★★★ 5	Wissen rund um Notion und angebundene Quellen	Veraltete interne Quellen liefern weiterhin veraltete Antworten.	Tarif / Agentennutzung
Microsoft 365 Copilot	★★★★★ 5	Wissensarbeit in Microsoft 365	Vorhandene Zugriffsrechte und passende Lizenzen entscheiden.	Plätze / Lizenz / Zusatzfunktionen

Vergleichstest: Zehn Fragen mit erlaubten und gesperrten Quellen: richtige Fundstelle, Aktualität und Berechtigungsgrenzen prüfen.

39 · Shopsuche, Retrieval & Empfehlungen

Algolia als Suchplattform; Cohere als Reranking-Baustein. Für unterschiedliche Architekturen, nicht eins zu eins austauschbar.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Algolia AI Search	★★★★★ 5	Suchfunktion für Produkte und Inhalte	Ranking muss mit eigenen Suchanfragen und Geschäftszielen geprüft werden.	Anfragen / Datensätze / Tarif
Cohere Rerank	★★★★½ 4.5	Relevanzsortierung in einer eigenen Retrieval-Kette	Kein vollständiges Wissenssystem; Retrieval und Quellenzugriff zusätzlich nötig.	API / Deployment

Vergleichstest: 50 echte Suchanfragen bewerten: richtige Treffer unter den ersten fünf, Nulltreffer und Antwortzeit.

40 · Prognosen & eigene ML-Anwendungen

Databricks und DataRobot nach Datenplattform und Betrieb auswählen. Keine Prognosegüte aus Produktbeschreibungen ableiten.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Databricks AI	★★★★★ 5	Eigene KI-Anwendungen auf Unternehmensdaten	Datenengineering und fachliche Evaluation bleiben eigene Arbeit.	Cloud / Rechenleistung / Plattform
DataRobot	★★★★½ 4.5	Aufbau und Betrieb eigener KI-Lösungen	Prognosegüte lässt sich nur mit eigenen, zeitlich getrennten Testdaten belegen.	Unternehmensangebot / Infrastruktur

Vergleichstest: Zeitlich getrennte Trainings- und Testdaten verwenden; gegen einfache historische Basisprognose vergleichen.

41 · Wissenschaft, Klima & biologische Modelle

AlphaFold für molekulare Strukturen; WeatherNext für Wetter. Das sind Fachgebiete mit völlig unterschiedlichen Zielgrößen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
AlphaFold	★★★★½ 4.5	Strukturvorhersage in der Forschung	Eine Strukturvorhersage belegt keine Wirkung eines Medikaments; Lizenz prüfen.	Forschungszugang / Rechenleistung
WeatherNext	★★★★ 4	Wettervorhersage als spezialisierter Datenbaustein	Lokale Güte, Zugriff und Einsatzzeitraum gesondert prüfen.	Datenzugang / Integration

Vergleichstest: Biologie mit unabhängig geprüften Strukturen; Wetter mit zurückgehaltenen Zeiträumen und regionalen Referenzdaten vergleichen.

42 · Robotik, Logistik & visuelle Inspektion

Gemini Robotics für modellbasierte Robotik, Omniverse für Simulation und Roboflow für visuelle Erkennung. Bausteine statt Universalroboter.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Gemini Robotics	★★★★ 4	Robotik-Entwicklung auf geeigneter Hardware	Forschungs-/Entwicklungszugang und reale Hardware begrenzen den Einsatz.	Zugang / Hardware / Entwicklung
NVIDIA Omniverse	★★★★★ 5	Simulation und digitale Umgebungen	Kein direkt vergleichbares Endnutzer-Chatprodukt.	Infrastruktur / Engineering / Lizenz
Roboflow	★★★★★ 5	Visuelle Inspektion und Objekterkennung	Fehler bei Beleuchtung, Perspektive und neuen Objekten separat messen.	Tarif / Inferenz Hardware

Vergleichstest: Definierte reale Aufgabe mit unbekannten Fällen testen;
Fehlhandlungen, Ausfälle und Wiederherstellung messen.

43 · Telefonie & interaktive Sprachagenten

ElevenLabs Agents für Sprachdialoge; Voiceflow für gestaltete Dialoge mit Anbindungen.

Erst den vollständigen Kundenfall abnehmen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
ElevenLabs Agents	★★★★½ 4.5	Gesprochene interaktive Kundenabläufe	Unterbrechungen, Eskalation und falsche Zusagen an echten Fällen testen.	Minuten / Modell- und Telefoniekosten
Voiceflow	★★★★★ 5	Gestaltete Dialoge mit eigener Prozesslogik	Wissenspflege und Eskalation im echten Kundenfall testen.	Tarif / Credits / Modellkosten

Vergleichstest: Anruf mit Unterbrechung, Akzent, falscher Kundennummer und menschlicher Übergabe; keine erfundenen Zusagen zulassen.

44 · Agenten testen, überwachen & betreiben

LangSmith für Beobachtbarkeit; Databricks für KI innerhalb der Datenplattform. Beide brauchen eigene Sollfälle und Fehlergrenzen.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
LangSmith	★★★★★ 5	Traces und Untersuchung von Agentenabläufen	Protokolle allein beweisen keine Qualität; Sollfälle und Abnahmen zusätzlich nötig.	Traces / Speicher / Tarif
Databricks AI	★★★★½ 4.5	KI-Betrieb in bestehender Datenplattform	Datenengineering und fachliche Evaluation bleiben eigene Arbeit.	Cloud / Rechenleistung / Plattform

Vergleichstest: Bekannte Fehler einspielen: falsche Quelle, Timeout, doppelter Auftrag, unerlaubtes Werkzeug; Erkennung und Wiederanlauf prüfen.

45 · Reise, Visa & mehrsprachiger Service

Dokumentenerfassung, Übersetzung und Workflow verbinden. Für Visa Vietnam zählen richtige Daten und verlässliche Übergaben.

PRODUKT	BEWERTUNG	BESONDERS GEEIGNET	GRENZE	KOSTENFAKTOR
Google Document AI	★★★★★ 5	Extraktion und Strukturierung von Antragsunterlagen	Feldgenauigkeit pro Dokumentart und Sprache prüfen.	Seiten / Prozessor / Cloud-Nutzung
DeepL	★★★★½ 4.5	Mehrsprachige Kundenkommunikation	Zielsprachliche Redaktion bleibt bei Werbung und Fachsprache nötig.	Tarif / API / Dokumentgrenzen
n8n	★★★★★ 5	Statusabläufe und geprüfte Übergaben	Selbsthosting verschiebt Betrieb und Wartung zu uns.	Hosting / Ausführungen / Modellkosten

Vergleichstest: Freigegebene Musterfälle mit absichtlich widersprüchlichen Daten durchspielen; amtliche Anforderungen separat aktualisieren.

Eigener Testplan

Gleiches Briefing, gleiche Referenzen, dokumentierte Modellversion/Modus/Tarif/Region, mindestens fünf Wiederholungen pro Kandidat, Blindbewertung soweit möglich, Fehlversuche mitzählen. Gesamtkosten einschließlich manueller Nacharbeit pro akzeptiertem Ergebnis berechnen. Zuerst Song mit Lyrics, filmische Kontinuität, Anfrage-zu-CRM und belegte Newsredaktion testen.

Quelle: interner Kategorien- und Sternevergleich vom 19.09.2026. Sterne stehen für dokumentierte Eignung je Aufgabe, nicht für objektive Bestleistung. Gleiche Sternzahl ist kein Fehler.