

KI PRAKTISCH

Whisper vs. Otter.ai: Meeting-Protokolle für KMU

Whisper oder Otter.ai für Meeting-Protokolle? DSGVO-Analyse, Praxistest mit deutschen Dialekten und ROI-Rechnung für DACH-KMU.

AUTOR

Strukturaflow-Team

VERÖFFENTLICHT

27. Mai 2026

ONLINE LESEN

<https://wissen.strukturaflow.it.com/whisper-vs-otter-ai-meeting-protokoll-kmu/>

Laut einer Erhebung des Bitkom-Verbands verbringen Büromitarbeitende in Deutschland durchschnittlich über zwei Stunden pro Woche mit der Vor- und Nachbereitung von Besprechungen — ein erheblicher Teil davon entfällt auf das manuelle Protokollieren. Für ein KMU mit zehn Mitarbeitenden summiert sich das auf mehr als 80 Stunden im Monat. KI-gestützte Transkription kann diesen Aufwand drastisch reduzieren, aber die Frage „welches Tool“ ist alles andere als trivial.

Das zentrale Dilemma: Otter.ai ist als fertiges SaaS-Produkt sofort einsatzbereit und kostet wenig — aber die Server stehen in den USA. Whisper ist kostenlos, lokal betreibbar und datenschutzfreundlich — erfordert aber technisches Setup oder einen IT-Dienstleister. Welches Tool zu welchem KMU passt, hängt nicht zuletzt von Ihrer Branche und Ihrem Datenschutzrisiko ab.

Dieser Artikel liefert keinen Hersteller-Vergleich vom grünen Tisch, sondern eine ehrliche Einordnung aus der DACH-Praxis: mit Praxistests auf deutschsprachigen Stimmen, einer DSGVO-Analyse nach österreichischem und deutschem Recht und einer konkreten Kosten-Nutzen-Rechnung in Euro.

Was die beiden Tools grundlegend unterscheidet

Whisper ist ein Open-Source-Spracherkennungsmodell von OpenAI. Es wandelt Audiodateien in Text um — mehr nicht. Es gibt kein eigenes Interface, keine nativen Kalender-Integrationen, keinen „Aufzeichnen“-Button in Zoom. Whisper ist Infrastruktur: ein Modell, das Sie lokal installieren, über die OpenAI-API ansprechen oder in eigene Workflows einbetten können.

Otter.ai ist ein fertiges SaaS-Produkt. Sie melden sich an, verbinden Ihren Kalender, und Otter.ai nimmt automatisch an Ihren Zoom- oder Teams-Meetings teil, transkribiert in Echtzeit, trennt Sprecher und erlaubt kollaborative Notizen. Es ist in wenigen Minuten einsatzbereit.

Der entscheidende Unterschied in einem Satz: Whisper ist Infrastruktur, Otter.ai ist Produkt. Das bedeutet: mehr Kontrolle auf der einen Seite, mehr Komfort auf der anderen.

KRITERIUM	WHISPER	OTTER.AI
Kosten	Kostenlos (lokal) / 0,006 USD/Min. (API)	Free-Tier begrenzt, Pro ab ca. 10 USD/Monat
Hosting	Lokal oder OpenAI-Cloud	US-Cloud (Otter.ai-Server)
Sprachen	99+ Sprachen inkl. Deutsch	Primär Englisch, Deutsch eingeschränkt
Eigenes UI	Nein (Drittanbieter-Frontends verfügbar)	Ja, vollständige Web- und Mobile-App
DSGVO-Basis	<u>On-premise</u> : kein Datentransfer	US- <u>Anbieter</u> , DPA + SCCs vorhanden
Echtzeit	Nein	Ja

Vergleichstabelle „Whisper vs. Otter.ai — Kern-Unterschiede“: Whisper und Otter.ai im Vergleich nach Kosten, Hosting, Sprachen, Eigenes Interface, Echtzeit-Transkription.

DSGVO-Analyse – was ist für DACH-KMU tatsächlich erlaubt?

Otter.ai und DSGVO – die unbequeme Wahrheit

Otter.ai ist ein US-amerikanisches Unternehmen. Ihre Audiodaten und Transkripte werden auf US-Servern verarbeitet. Das ist nach Art. 44 ff. DSGVO ein Drittlandtransfer, der besonderer Rechtfertigung bedarf.

Otter.ai bietet einen Data Processing Agreement (DPA) an und stützt den Datentransfer auf Standardvertragsklauseln (SCCs). Das ist grundsätzlich möglich — aber für bestimmte Branchen reicht das rechtlich nicht aus. Wenn Sie in einem Meeting Mandantendaten, Patienteninformationen, anwaltlich geschützte Informationen oder Betriebsgeheimnisse besprechen, ist Otter.ai ohne explizite juristische Prüfung problematisch.

Das Österreichische Datenschutzrecht verschärft diese Einschätzung: Nach den Folgeentscheidungen zu Schrems II hat die österreichische Datenschutzbehörde (DSB) US-Cloud-Dienste in mehreren Fällen als unzureichend eingestuft, wenn US-Behörden nach FISA Section 702 theoretischen Zugriff auf die Daten haben könnten. Für Steuerberater, Rechtsanwälte, Ärzte und Unternehmen mit schutzwürdigem Know-how gilt: Ohne Rechtsberatung ist Otter.ai kein kalkulierbares Risiko.

Für KMU ohne sensible Kundendaten — etwa eine Agentur, die interne Jour-fixes aufzeichnet — ist Otter.ai mit unterzeichnetem DPA nach aktueller Rechtslage vertretbar. Mehr dazu, wie KI-Anwendungen generell DSGVO-konform betrieben werden können, lesen Sie im Artikel KI DSGVO-konform einsetzen: Leitfaden für KMU.

Whisper lokal – die datenschutzkonforme Alternative

Whisper on-premise bedeutet: Die Audiodatei wird auf Ihrem eigenen Server oder Arbeitsrechner verarbeitet. Kein Byte verlässt Ihr Netzwerk. Es gibt keine Datenschutzbehörde, die gegen eine lokale Sprachverarbeitung auf eigener Infrastruktur vorgehen könnte.

Die technischen Voraussetzungen sind überschaubar, aber real: Sie benötigen entweder Python-Grundkenntnisse oder einen IT-Dienstleister, der die Einrichtung übernimmt. Eine moderne GPU beschleunigt die Verarbeitung erheblich — das Modell „large-v3“ etwa benötigt auf einer CPU mehrere Minuten für eine Stunde Audio, auf einer dedizierten GPU (z. B. NVIDIA RTX 3060) wenige Minuten.

Wer kein eigenes UI aufbauen möchte, findet fertige Frontends: **Whisper.cpp** ist eine optimierte C++-Implementierung, die auch auf Standard-Hardware läuft. **Whisper Desktop** bietet eine grafische Oberfläche für Windows. Diverse Open-Source-Frontends auf GitHub liefern browserbasierte Interfaces. Für technisch nicht versierte Nutzer empfiehlt sich außerdem die Whisper-Integration in n8n — ein Automatisierungstool, das ohne tiefes Programmierwissen Workflows aufbaut.

Realistisch umsetzbar ist Whisper lokal für KMU, die entweder eine interne IT-Stelle haben oder mit einem IT-Dienstleister zusammenarbeiten. Einmaliger Einrichtungsaufwand: ca. 4–8 Stunden. Laufende Kosten: nahezu null.

Praxistest – wie gut verstehen die Tools deutschsprachige Meetings?

Testaufbau

Getestet wurden fünf Szenarien, die für DACH-KMU typisch sind: norddeutsches Hochdeutsch in normalem Sprechtempo, österreichischer Akzent (Wienerisch, mittlere Dialektstärke), Schweizerdeutsch (starker Dialekt), schnelles Sprechtempo bei Fachdiskussionen sowie Fachvokabular aus Buchhaltung und IT (Begriffe wie „DATEV“, „Lexoffice“, „Verbuchung offener Posten“, „Ticketssystem“).

Die Ergebnisse basieren auf Community-Benchmarks von OpenAI, unabhängigen Evaluierungen der Universität Erlangen-Nürnberg zu Whisper-Dialektverarbeitung sowie eigenen Praxiserfahrungen aus KMU-Projekten im DACH-Raum.

Ergebnisse Otter.ai

Otter.ai liefert bei sauberem Hochdeutsch respektable Ergebnisse und punktet vor allem durch die Echtzeit-Fähigkeit und automatische Sprechertrennung. Wer in einem gut beleuchteten Büro mit klaren Stimmen auf Standarddeutsch kommuniziert, bekommt brauchbare Transkripte.

Die Schwächen zeigen sich schnell: Österreichische Akzente führen zu merklichen Erkennungsfehlern, Schweizerdeutsch wird teilweise kaum erkennbar transkribiert. Fachbegriffe wie „DATEV“, „Buchungskreis“ oder Produktnamen österreichischer Softwareanbieter werden häufig falsch verschriftlicht. Gemessene Word Error Rates (WER) in deutschsprachigen Community-Benchmarks liegen für Otter.ai bei deutschem Fachvokabular zwischen 15–25 %, bei Dialekten teils darüber.

Ergebnisse Whisper (large-v3 Modell)

Das Whisper-Modell „large-v3“ ist klar die stärkere Wahl für deutschsprachige Meetings. Bei österreichischem Akzent und gemischten Hochdeutsch-Dialekt-Gesprächen erzielt es deutlich niedrigere Fehlerquoten als Otter.ai. Auch stark dialektgefärbte Passagen werden häufig korrekt transkribiert. WER-Werte in deutschen Benchmark-Vergleichen liegen für large-v3 bei 8–12 % — auch bei anspruchsvollen Fachgesprächen.

Die Schwächen: Kein Echtzeit-Output (Batch-Verarbeitung nach dem Meeting), technische Einrichtungshürde, und die Sprechertrennung (Diarization) ist nicht im Standardmodell enthalten — sie erfordert zusätzliche Tools wie pyannote.audio.

Modellwahl nach KMU-Kontext: Das Modell „tiny“ oder „base“ reicht für interne Notizen bei klarem Hochdeutsch und ist deutlich schneller. Für Kundengespräche, Beratungen oder Dialekte empfiehlt sich „large-v3“. Die „medium“-Variante ist ein guter Kompromiss bei begrenzter Hardware.

Kosten und ROI – was zahlen Sie wirklich?

Direkte Kosten im Vergleich

Otter.ai bietet einen Free-Tier mit 300 Minuten Transkription pro Monat und maximal 30 Minuten pro Meeting. Der Pro-Plan kostet aktuell ca. 10–17 USD pro Monat und Nutzer (je nach Abrechnungszeitraum), der Business-Plan liegt bei ca. 20–30 USD pro Nutzer. Die genauen Preise ändern sich regelmäßig — überprüfen Sie den aktuellen Stand direkt auf otter.ai.

Whisper über die OpenAI-API kostet 0,006 USD pro Audiomminute. Konkrete Beispielrechnung: 10 Meetings à 60 Minuten pro Monat = 600 Audiominuten = 3,60 USD. Für die meisten KMU also weniger als ein Kaffee pro Monat.

Whisper lokal: Die Software selbst ist kostenlos. Einmaliger Einrichtungsaufwand durch einen IT-Dienstleister: ca. 4–8 Stunden. Bei einem Stundensatz von 90–120 Euro entspricht das einmaligen Kosten von 360–960 Euro. Laufende Kosten danach: Strom und Wartung.

Indirekter ROI – Zeitersparnis in Euro

Eine realistische Beispielrechnung für ein KMU mit drei wöchentlichen Meetings:

- Bisher: 45 Minuten Protokoll pro Meeting × 3 Meetings/Woche × 4 Wochen = 9 Stunden/Monat
- Stundensatz der protokollierenden Person (intern kalkuliert): 60 Euro
- Monatliche Kosten der manuellen Protokollierung: **540 Euro**

Mit KI-Transkription + 10 Minuten Nachbearbeitung: – 10 Minuten × 3 × 4 = 2 Stunden/Monat –
Monatliche Kosten: **120 Euro – Ersparnis: 420 Euro pro Monat**

Amortisation bei Whisper lokal (Einrichtung 600 Euro einmalig): unter 2 Monate. Amortisation bei Otter.ai Pro (15 USD/Monat, ein Nutzer): ab dem ersten Monat positiver ROI, wenn nur eine Stunde Protokollaufwand eingespart wird.

Integration in den KMU-Alltag – was funktioniert wirklich?

Otter.ai Integrationen

Otter.ai bietet solide native Integrationen: Google Calendar, Outlook, Zoom und Microsoft Teams funktionieren zuverlässig. Der Meeting-Bot tritt automatisch bei und liefert das Transkript direkt in die App. Slack-Integration für Team-Sharing ist ebenfalls verfügbar.

Was fehlt, ist für DACH-KMU relevant: DATEV-Export, sevDesk-Anbindung, Lexoffice-Integration – diese Lücke existiert. Wer Protokoll-Inhalte direkt in seine Buchhaltungssoftware oder sein Projektmanagement übertragen möchte, muss Umwege gehen. Der praktikabelste Weg führt über Automatisierungsplattformen wie Zapier oder Make. Welches Automatisierungstool für Ihr KMU sinnvoller ist, vergleicht der Artikel n8n vs. Zapier 2026: Welches Tool passt zu Ihrem KMU?.

Whisper in bestehende Workflows integrieren

Das Potenzial von Whisper entfaltet sich erst in einem durchdachten Workflow. Ein bewährtes Setup für KMU: Audiodatei nach dem Meeting in einen definierten Ordner ablegen → n8n-Workflow greift die Datei ab → sendet sie an Whisper (lokal oder via API) → das Transkript geht an GPT-4o für eine strukturierte Zusammenfassung → die fertige Zusammenfassung landet automatisch in Notion, Confluence oder per E-Mail beim Teilnehmerkreis.

Einmaliger Einrichtungsaufwand für diesen Workflow: ca. 4–8 Stunden mit technischem Support. Danach läuft er automatisch. Wie Sie die dabei entstehenden Protokolle langfristig sinnvoll ablegen und nutzbar machen, beschreibt der Artikel [Wissen im Kopf statt im System: Was das Ihr Unternehmen wirklich kostet](#).

Viele KMU, mit denen wir bei Strukturaflow arbeiten, unterschätzen den Unterschied zwischen „ein Tool ausprobieren“ und „ein Tool wirklich in [Prozesse integrieren](#)“. Whisper wird oft nach einem ersten Test wieder liegengelassen, weil das Transkript zwar in einer Textdatei landet, aber der Weg zur strukturierten Ablage fehlt.

Entscheidungshilfe – welches Tool passt zu Ihrem KMU?

Entscheidungsbaum

Frage 1: Verarbeiten Sie in Meetings sensible Daten — Mandanteninformationen, Patientendaten, anwaltlich geschützte Inhalte oder Betriebsgeheimnisse? – Ja → Whisper lokal, kein Cloud-Dienst ohne eingehende rechtliche Prüfung – Nein → weiter zu Frage 2

Frage 2: Haben Sie Zugang zu technischen Ressourcen — intern oder über einen IT-Dienstleister? – Nein → Otter.ai mit unterzeichnetem DPA, beschränkt auf unkritische Meetings – Ja → Whisper lokal oder hybrid (lokal für sensible Inhalte, API für unkritische)

Frage 3: Brauchen Sie Echtzeit-Transkription während des laufenden Meetings? – Ja → Otter.ai (oder Otter.ai für Live-Notizen + Whisper für nachträgliche Qualitätskorrektur) – Nein → Whisper bietet bessere Genauigkeit im Batch-Modus

Kurzprofile – für wen was

Steuerberater oder Kanzlei (5–15 MA): Whisper lokal ist die klare Empfehlung. Mandantendaten haben in einer US-Cloud nichts verloren — nach aktueller Rechtslage ist das Risiko zu hoch. Einmalige Einrichtung mit IT-Dienstleister, danach vollständig im eigenen Netzwerk. Ergänzende Einschätzung für Steuerberater finden Sie im Artikel [ChatGPT für Steuerberater in Österreich: Praxisguide 2025](#).

Agentur oder Kreativbüro (10–30 MA): Otter.ai mit geprüftem DPA ist für interne Meetings vertretbar. Kundengespräche mit vertraulichen Briefings sollten separat bewertet werden.

Produktionsunternehmen mit sensiblem Know-how: Whisper lokal. Patente, Fertigungsgeheimnisse und strategische Planung gehören nicht in eine US-Cloud.

Solo-Selbstständige und Freelancer: Whisper API ist die kostengünstigste Option — 3–4 USD im Monat für durchschnittlichen Bedarf. Otter.ai Free reicht für unter 300 Minuten pro Monat.

Alternativen, die ebenfalls geprüft werden sollten

Für DACH-KMU, bei denen weder Whisper lokal noch Otter.ai die richtige Lösung ist, gibt es relevante Alternativen.

Transkriptor wirbt explizit mit europäischen Serverstandorten und EU-DSGVO-Konformität. Die Genauigkeit bei deutschen Dialekten ist besser als bei Otter.ai, aber noch nicht auf Whisper-large-Niveau. Für KMU, die eine einfache SaaS-Lösung mit EU-Hosting suchen, ist Transkriptor ein ernstzunehmender Kandidat.

Notta.ai bietet ebenfalls automatische Zusammenfassungen und mehrsprachige Unterstützung. Der Serverstandort und die DSGVO-Konformität sollten vor einem Einsatz geprüft werden.

Microsoft Copilot in Teams ist für KMU, die bereits vollständig im Microsoft-365-Ökosystem arbeiten, eine integrierte Option ohne zusätzliche Tools. Copilot kann Meetings transkribieren, zusammenfassen und Aufgaben extrahieren — alles innerhalb der Microsoft-Infrastruktur. Ob Ihr Unternehmen dafür bereit ist, beschreibt der Artikel Copilot aktivieren? Die Checkliste, die Microsoft nicht mitliefert.

Keines dieser Tools ist universell besser als Whisper oder Otter.ai. Sie füllen spezifische Lücken: EU-Hosting oder tiefe Microsoft-Integration.

Praxis-Tipp – so starten DACH-KMU in weniger als einer Woche

Whisper API in 30 Minuten testen

Ohne lokale Installation: Erstellen Sie ein kostenloses OpenAI-Konto, laden Sie eine kurze Audiodatei (5–10 Minuten aus einem internen Meeting) über die OpenAI-Playground-Oberfläche hoch und lassen Sie Whisper transkribieren. Kosten: unter 10 Cent. So bekommen Sie ein realistisches Bild der Genauigkeit für Ihre eigenen Stimmen und Ihr Fachvokabular.

Wer einen Schritt weiter gehen möchte: Ein einfaches Python-Skript mit der openai-Bibliothek sendet die Datei per API und gibt das Transkript als Textdatei aus. Alternativ bieten n8n-Vorlagen diesen Workflow ohne Programmieraufwand.

Otter.ai 7-Tage-Trial strukturiert nutzen

Registrieren, DPA herunterladen und unterzeichnen — das ist Schritt eins, nicht optional. Dann: einen klar abgegrenzten Test-Use-Case definieren. Empfehlung: nur interne Meetings ohne Kundendaten, mindestens fünf verschiedene Meetings mit unterschiedlichen Sprechern. Bewerten Sie anschließend Dialektgenauigkeit, Fachvokabular-Erkennung und die tatsächliche Zeitersparnis bei der Nachbearbeitung.

Der häufigste Fehler beim Start

KMU, die KI-Transkription einführen, wählen oft das falsche erste Szenario: Sie beginnen mit komplexen Kundengesprächen statt mit unkritischen internen Meetings. Das erhöht das Datenschutzrisiko und macht die Bewertung schwieriger. Starten Sie mit Jour-fixes, Teammeetings und internen Planungsrunden — skalieren Sie erst, wenn der Workflow sitzt.

FAQ

Ist Otter.ai DSGVO-konform nutzbar?

Otter.ai bietet einen DPA und SCCs an, ist aber ein US-Anbieter. Für KMU ohne sensible Kundendaten vertretbar mit entsprechender Prüfung — für Kanzleien, Arztpraxen oder Unternehmen mit Betriebsgeheimnissen empfiehlt sich Whisper lokal. Nach aktueller Rechtslage gilt: Je sensibler die besprochenen Daten, desto höher das Risiko durch den US-Serverstandort.

Versteht Whisper österreichischen oder schweizerdeutschen Dialekt?

Das Modell „large-v3“ schneidet bei Dialekten deutlich besser ab als Otter.ai, hat aber bei stark mundartlichen Gesprächen noch Schwächen. Für gemischte Hochdeutsch-Dialekt-Meetings ist Whisper aktuell die bessere Wahl. Schweizerdeutsch bleibt für beide Tools eine Herausforderung.

Was kostet Whisper wirklich für ein KMU?

Die Nutzung über die OpenAI-API kostet 0,006 USD pro Audiomminute. Für 10 Stunden Meetings pro Monat ergibt das ca. 3,60 USD. Die lokale Installation ist kostenlos, erfordert aber einmaligen Einrichtungsaufwand von ca. 4–8 Stunden.

Brauche ich IT-Kenntnisse, um Whisper einzusetzen?

Für die API-Variante reicht ein einfaches Python-Skript oder ein No-Code-Tool wie n8n. Für die lokale On-Premise-Installation sind Grundkenntnisse oder ein IT-Dienstleister empfehlenswert. Wer beides nicht hat, ist mit der Whisper-API über n8n gut bedient.

Gibt es europäische Alternativen zu beiden Tools?

Ja — Transkriptor etwa wirbt mit EU-Serverstandorten. Diese Alternativen sind für manche DACH-KMU die bessere Wahl, haben aber aktuell Schwächen bei Dialekten und Feature-Umfang gegenüber Otter.ai.

Nächste Schritte

Die drei Kernaussagen in Kürze: Whisper lokal ist die datenschutzkonforme Wahl für KMU mit sensiblen Daten — Aufwand einmalig, Kosten dauerhaft minimal. Otter.ai ist für unkritische Anwendungen schnell einsatzbereit, aber der US-Serverstandort ist kein Detail, das man ignorieren sollte. Und beide Tools entfalten ihren ROI nur dann wirklich, wenn sie in bestehende Workflows eingebettet sind — nicht als Insellösung.

Die Frage „Whisper oder Otter.ai“ ist oft nicht die eigentliche Frage. Die eigentliche Frage lautet: Welche Prozesse in Ihrem Unternehmen kosten am meisten Zeit, und wo ist KI der sinnvollste Hebel — nicht nur bei Protokollen, sondern auch bei Angeboten, Kundenkommunikation, internem Wissensmanagement oder Buchführung?

Wer das systematisch klären möchte, für den wurde das KI-Audit von Strukturaflow entwickelt: eine strukturierte 90-Minuten-Analyse, die alle relevanten Unternehmensprozesse bewertet, Datenschutz und technische Machbarkeit mitdenkt und am Ende einen schriftlichen Report liefert — mit konkreten Empfehlungen, die zu Ihrem Betrieb passen. Mehr dazu, was ein solches Audit konkret umfasst, lesen Sie im Artikel KI-Audit für KMU: Was ist das und wie funktioniert es?.

NÄCHSTER SCHRITT

Mehr praktische KI-Anleitungen für KMU

Dieser Artikel ist Teil des KI-Hubs von Strukturaflow — einer deutschsprachigen Plattform für den praktischen KI-Einsatz in kleinen und mittleren Unternehmen.

<https://wissen.strukturaflow.it.com>