KI

- Microsoft 365 Copilot – Entscheidungsgrundlage für KMU und Consultants
- Nutzen Ihre Mitarbeitenden KI-Tools? Das sollten Sie wissen.
- Wissensdatenbank (RAG) aufbauen
  - Was ist Sinn und Zweck einer KI-Gestützten Wissensdatenbank
  - Warum wir Word-Dateien vor dem Hochladen in die Wissensdatenbank umwandeln
  - Word → Markdown: Checkliste für saubere Konvertierungen

Microsoft 365 Copilot – Entscheidungsgrundlage für KMU und Consultants

Microsoft 365 Copilot – Entscheidungsgrundlage für KMU und Consultants

Unabhängige Marktanalyse 2026

Management Summary

Microsoft 365 Copilot hat sich 2026 als produktiver KI-Assistent für wissensintensive Arbeit etabliert. Die Technologie ist ausgereift, die Preise sind konkurrenzfähig, und die Datenschutzarchitektur entspricht den Anforderungen deutscher Unternehmen – sofern die organisatorischen Voraussetzungen stimmen.

Kernaussagen dieser Analyse:

- Microsoft bietet drei klar abgestufte Copilot-Varianten: kostenlosen Copilot Chat für erste KI-Erfahrungen, Copilot Business für KMU bis 300 Nutzer und Copilot Enterprise für größere Organisationen[1]

- Der Mehrwert hängt direkt von der Qualität Ihrer Microsoft-365-Infrastruktur ab: saubere Berechtigungen, strukturierte Ablage und Exchange Online sind Pflicht[2]

- Datenschutzrechtlich ist Copilot vertretbar, wenn AV-Vertrag, Datenklassifizierung und Nutzungsrichtlinien vorhanden sind; lokale Datenverarbeitung in Deutschland kommt 2026[3]

- Im Toolvergleich ergänzen sich Copilot (für M365-Integration), ChatGPT (für universelle Wissensarbeit) und spezialisierte Lösungen wie Lagdock (für datenschutzsensible Dokumentenarbeit) – die beste Strategie ist oft eine Kombination

Diese Analyse richtet sich an Geschäftsführer, IT-Verantwortliche und Consultants, die eine fundierte Entscheidung über den Einsatz von Microsoft 365 Copilot treffen wollen.

---

1. Produktvarianten im Vergleich

Microsoft positioniert Copilot 2026 in drei klar differenzierten Varianten. Die Unterschiede liegen vor allem beim Datenzugriff, der App-Integration und den Lizenzanforderungen.

Variante	Funktionen	Preis	Zielgruppe	Einschränkungen
Copilot Chat (kostenlos)	Sicherer KI-Chat in M365-Apps; webbasiert; Agenten möglich, aber verbrauchsabhängig	Ohne Aufpreis für berechtigte M365-Nutzer	Unternehmen, die KI testen wollen ohne Vollintegration	Kein vollständiger Copilot-in-Apps-Umfang; für Agenten Azure-Abo nötig
Microsoft 365 Copilot Business	Copilot Chat plus volle Integration in Teams, Outlook, Word, PowerPoint, Excel; Agenten mit Copilot Studio	Ab 15,60 € pro Nutzer/Monat (Aktion), regulär 18,20 €; M365-Basislizenz zusätzlich	KMU bis 300 Nutzer mit M365-Arbeitsplatz	Basislizenz erforderlich; auf 300 Nutzer begrenzt; Teams-Lizenz für Teams-Features nötig
Microsoft 365 Copilot Enterprise	Vollständiger M365 Copilot für Enterprise-Kunden mit erweiterten Governance-, Compliance- und Sicherheitsfunktionen	Separat über Enterprise-Pläne	Unternehmen mit E3/E5-Basis, komplexer Governance, mehr als 300 Nutzer	Höhere Lizenzhürde; nur sinnvoll mit reifer M365-Architektur

Table 1: Vergleich der drei Microsoft 365 Copilot-Varianten (Stand März 2026)

Wichtige Differenzierung: Der kostenlose Copilot Chat ist nicht gleichbedeutend mit dem vollwertigen Microsoft 365 Copilot in den Office-Apps. Microsoft beschreibt Copilot Business ausdrücklich als Ausbau von Copilot Chat um den arbeitsbasierten Kontext (Work-IQ) und die direkte Nutzung in allen Microsoft-365-Anwendungen[1].

Copilot Studio und Agenten: Für Nutzer mit Copilot Business ist die Erstellung und Nutzung von Agenten (früher „Bots") für interne Szenarien in Teams, SharePoint und Microsoft 365 Copilot Chat **ohne zusätzliche Lizenzkosten** enthalten. Die Nutzung unterliegt Fair-Use-Limits; erst bei externen Szenarien (z. B. öffentliche Website-Chatbots) oder sehr hohen Nachrichtenvolumen wird eine separate Copilot Studio-Lizenz mit Copilot Credits notwendig[4][5].

2. Voraussetzungen für den sinnvollen Einsatz

Copilot ist nur dann produktiv, wenn die technische Basis sauber ist und die Datenlandschaft nicht chaotisch organisiert wurde. Für KMU bedeutet das konkret:

2.1 Technische Voraussetzungen

- **Microsoft 365 Apps bereitgestellt:** Copilot integriert sich in Word, Excel, PowerPoint, Outlook, Teams und weitere Apps[2]
- **Exchange Online als Pflicht:** Unterstützte primäre Postfächer erforderlich; Archivpostfächer, Shared Mailboxes und Delegate-Mailboxen haben eingeschränkten Support[2]
- **OneDrive aktiv:** Einzelne Copilot-Funktionen bauen auf OneDrive auf[2]
- **Teams sauber konfiguriert:** Für Besprechungsbezug nach Meetingende sind Transkription oder Aufzeichnung erforderlich[2]
- **Netzwerk-Konnektivität:** Microsoft-365-Endpunkte, WebSockets und Domains wie `*.cloud.microsoft` und `*.office.com` müssen erreichbar sein[2]

2.2 Lizenzrechtliche Voraussetzungen

- Copilot Business erfordert eine **separate Copilot-Lizenz** plus berechtigenden Microsoft-365-Basisplan (z. B. Business Basic, Standard oder Premium)[1][2]
- Der Business-Tarif gilt nur bis **300 Benutzer**; darüber hinaus ist Enterprise erforderlich[1]
- Für Teams-bezogene Funktionen ist zusätzlich eine Teams-Lizenz erforderlich[1]

2.3 Organisatorische Voraussetzungen

- **Berechtigungen bereinigt:** Copilot arbeitet auf Basis derselben Zugriffsrechte wie Microsoft 365 insgesamt. Überberechtigte Nutzer sehen über Copilot schneller Inhalte, die sie technisch lesen durften, aber bislang faktisch schwer fanden[6]
- **Strukturierte Ablage:** Dateien, Teams und Mailboxen sollten logisch strukturiert und nicht „historisch offen" für zu viele Personen freigegeben sein[6]
- **Governance-Regeln definiert:** Es braucht Regeln für Prompting, Freigaben, Sensitivity Labels, Aufbewahrung und den Umgang mit KI-generierten Inhalten[6]

- **Pilotphase empfohlen:** Für KMU ist ein Pilot mit wenigen Key-Usern oft sinnvoller als ein flächendeckender Rollout, weil so Datenqualität, Berechtigungen und Nutzungsmehrwert vor breiter Einführung geprüft werden können[1][2]

Praktischer Hinweis für kleine Beratungsfirmen: Copilot bringt wenig, wenn E-Mails nicht in Exchange Online liegen, Dateien verstreut lokal liegen und Berechtigungen ungepflegt sind. Dann werden eher Informationschaos und Fehlinterpretationen beschleunigt als echte Produktivität[2][6].

---

3. Datenschutz & Compliance-Risiken

Aus DSGVO-Sicht ist die Ausgangslage besser als bei vielen offenen KI-Diensten. Microsoft gibt an, dass Prompts, abgerufene Daten und Antworten innerhalb der Microsoft-365-Servicegrenze bleiben, Azure OpenAI statt öffentlicher OpenAI-Dienste genutzt wird und Kundendaten nicht zum Training der Foundation-Modelle verwendet werden[6].

Wichtig: Copilot ist nicht automatisch „datenschutzsicher", sondern nur so sicher wie die eigene Microsoft-365-Governance. Die größten Risiken entstehen weniger durch das Modell selbst als durch falsche Berechtigungen, unklare Datenklassifizierung und unkontrollierte Nutzung[6][2].

3.1 Konkrete DSGVO-Risiken für deutsche Unternehmen

- **Beschleunigter Zugriff auf sensible Daten:** Überberechtigte Nutzer sehen über Copilot schneller Inhalte, die sie technisch ohnehin lesen durften, aber bislang faktisch schwer fanden; das erhöht das Risiko interner Datenschutzverletzungen[6]

- **Verarbeitung personenbezogener Daten in Prompts:** Prompts können personenbezogene Daten, Kundeninformationen, Gesundheitsdaten, Vertragsinhalte oder HR-Daten enthalten; diese Interaktionsdaten werden als Copilot-Aktivitätshistorie verarbeitet und gespeichert[6]

- **Komplexe Datenflüsse durch Agenten:** Agenten können auf Basis von Prompt, Aktivitätshistorie und zugänglichen M365-Daten Suchanfragen weiterleiten; dadurch steigt die Komplexität der Datenflüsse und damit der Prüfbedarf[6]

- **Teams-Transkription und -Aufzeichnung:** Ein sinnvoller Copilot-Einsatz in Teams setzt oft Transkription oder Aufzeichnung voraus; dabei entstehen zusätzliche personenbezogene Daten und teils mitbestimmungsrelevante Sachverhalte[2]

- **Fehlende Dokumentation und Betriebsratsabstimmung:** Wenn Datenschutz-Folgenabschätzung (DSFA), Verzeichnis der Verarbeitungstätigkeiten, Betriebsratsabstimmung oder interne Richtlinien fehlen, kann die Einführung organisatorisch angreifbar sein, obwohl die technische Plattform stark abgesichert ist[6][7]

3.2 Positive Entwicklungen 2026

Microsoft hat für Deutschland angekündigt, dass die Datenverarbeitung für Microsoft 365 Copilot ab 2026 **lokal in Deutschland** erfolgen kann. Dies erleichtert es Unternehmen in streng regulierten Sektoren (Behörden, Banken, Gesundheitswesen) erheblich, nationale Datenschutzrahmen einzuhalten und ihre eigenen Sicherheitsrichtlinien durchzusetzen[3].

Zusätzlich bietet Microsoft seit Ende 2025 erweiterte Datenschutz-Hilfen: ein M365-Kit mit Musterdokumenten, aktualisiertem Cloud Compendium und DSFA-Vorlagen, um die DSGVO-Dokumentation zu erleichtern[8].

3.3 Bewertung für deutsche KMU

Für normale Wissensarbeit mit Angebotsentwürfen, internen Projektunterlagen, Meeting-Zusammenfassungen und E-Mail-Auswertungen ist Copilot datenschutzrechtlich meist vertretbar, sofern AV-Vertrag, Rollen-/Rechtemodell, Sensitivity Labels und klare Nutzungsregeln vorhanden sind[6].

Kritischer wird es bei Personalakten, sensiblen Kundendossiers, medizinischen Daten, Rechtsfällen, Betriebsratsinformationen oder unbereinigten SharePoint-Strukturen. In solchen Bereichen sollte Copilot nur nach gezielter Freigabe, Datenklassifizierung und ggf. Ausschluss einzelner Speicherorte eingesetzt werden[6].

Handlungsempfehlung: Führen Sie vor dem Rollout eine Datenschutz-Folgenabschätzung durch, binden Sie Ihren Datenschutzbeauftragten ein, bereinigen Sie Berechtigungen in SharePoint/Teams, aktivieren Sie Audit-Logging und schulen Sie Ihre Mitarbeitenden im verantwortungsvollen Umgang mit KI-Prompts[7][9].

---

4. Dreifach-Vergleich: Copilot vs. ChatGPT vs. Lagdock

Für Soloselbständige und kleine Beratungsunternehmen erfüllen diese drei Tools unterschiedliche Rollen. Die folgende Gegenüberstellung zeigt, wo jedes Tool seine Stärken hat.

Kriterium	Microsoft 365 Copilot	ChatGPT	Lagdock
Hauptstärke	Tiefe Integration in Outlook, Word, Excel, PowerPoint, Teams und Microsoft Graph	Sehr stark für generische Wissensarbeit, Textentwurf, Ideation, Analysen und flexible Nutzung außerhalb eines einzelnen Software-Ökosystems	Potenziell stark für datenschutzbewusste Dokumenten- und Wissensarbeit, besonders wenn europäische Hosting-/Compliance-Aspekte im Fokus stehen

Hauptschwäche	Mehrwert hängt stark von sauberer M365-Struktur, Rechtemodell und Lizenzbasis ab	Ohne Enterprise-/Business-Konfiguration oft schwächer im nativen Zugriff auf interne Firmenabläufe und M365-Artefakte als Copilot	Meist geringere Standardintegration in Microsoft-365-Arbeitsabläufe; Marktpräsenz und Standardisierung oft schwächer
Ideale Einsatzszenarien	E-Mail-Zusammenfassungen, Angebotsentwürfe, Meeting-Nachbereitung, Dokumenterstellung, Recherche im eigenen Microsoft-365-Bestand	Brainstorming, Konzeptentwicklung, Qualitätsprüfung, Umformulieren, Analyse externer Inhalte, universelle Assistenz	Wissensdatenbank, Dokumentenabfrage, datenschutzsensiblere interne Recherche, lokal oder europäisch orientierte KI-Workflows
Datenschutz	Unternehmensdatenschutz innerhalb M365; keine Nutzung der Inhalte zum Training laut Microsoft; Risiko liegt vor allem in Berechtigungen und Governance	Hängt stark vom gebuchten Plan und den Admin-/Privacy-Einstellungen ab; bei Business/Enterprise bessere Kontrolle als Free-Version	Kann aus deutscher/europäischer Sicht attraktiv sein, wenn Hosting, DPA und Datenhaltung transparent dokumentiert sind
Kosten	Zusatzlizenz plus M365-Basislizenz; Business ab 15,60 € Aktionspreis bzw. regulär ab 18,20 € pro Nutzer/Monat bei jährlicher Abrechnung	Je nach Tarif meist flexibler und oft einfacher für Einzelpersonen zu starten; echter Unternehmensbetrieb erfordert aber passende Business-/Enterprise-Auswahl	Je nach Anbieterpaket teils günstiger oder zielgerichteter für Dokumentenfälle, aber Funktionsumfang variiert stark
Integration	Sehr hoch in Microsoft 365 und Copilot Studio/Agenten-Logik	Hoch über API, Integrationen und Custom Workflows, aber nicht nativ so tief in M365 eingebettet wie Copilot	Hängt stark vom jeweiligen Produkt und vorhandenen Schnittstellen ab

Table 2: Sachlicher Vergleich der drei KI-Tools für KMU und Consultants (Stand März 2026)

4.1 Sachliche Einordnung

- **Copilot** ist das beste Tool, wenn der Arbeitsalltag bereits in Microsoft 365 stattfindet und produktive Zeit in Outlook, Teams, Word und SharePoint verloren geht[1][2]
- **ChatGPT** ist oft das vielseitigste Denk-, Schreib- und Analysewerkzeug für Einzelberater, gerade für Konzeption, Sparring, Rechercheaufbereitung und kreative Arbeit[10][11]
- **Lagdock** ist vor allem dann interessant, wenn ein Consultant besonderen Wert auf kontrollierte Dokumentenräume, europäische Datenschutzargumentation oder alternative Wissensarbeits-Setups legt

4.2 Sicherheitshinweis zu ChatGPT

Aktuelle Untersuchungen zeigen, dass der Anteil sensibler Daten in ChatGPT-Eingaben von 11 % (2023) auf 34,8 % (Ende 2025) gestiegen ist[11]. Dies unterstreicht die Notwendigkeit klarer Nutzungsrichtlinien: Mitarbeitende sollten geschult werden, keine personenbezogenen Daten, Kunden-/Vertragsinformationen, Finanzdaten oder Geschäftsgeheimnisse in die Free-Version von ChatGPT einzugeben. Für Unternehmenseinsatz ist ChatGPT Business oder Enterprise mit verschärften Datenschutzeinstellungen und Admin-Kontrolle zu bevorzugen[10][11].

---

5. Chancen & strategische Empfehlung

Copilot Business lohnt sich besonders für Unternehmen, die bereits einen echten Microsoft-365-Arbeitsplatz haben: Exchange Online, Teams, OneDrive/SharePoint, standardisierte Dokumente und wiederkehrende Wissensarbeit. In solchen Umgebungen entfaltet Copilot seinen größten Nutzen bei Angebotsarbeit, Meeting-Nachbereitung, Vertriebsunterlagen, interner Recherche und E-Mail-Entlastung[1][2].

5.1 Besonders lohnend für

- Beratungsunternehmen mit starkem Outlook-, Teams- und Word-Anteil
- Kleine Kanzlei-, Agentur- oder Projektorganisationen mit viel Wissensarbeit und vielen wiederverwendbaren Inhalten
- KMU mit bereits gepflegter SharePoint-/OneDrive-Struktur und funktionierender Rechteverwaltung
- Geschäftsführer, Consultants und Projektleiter mit hoher Meeting- und Kommunikationslast

5.2 Weniger lohnend für

- Sehr kleine Firmen ohne konsequente Microsoft-365-Nutzung
- Unternehmen mit ungeordneten Ablagen, Schatten-IT und unklaren Freigaben
- Organisationen, die primär branchenspezifische Fachsoftware statt Microsoft 365 als Arbeitszentrale nutzen

5.3 Strategische Empfehlung für KI-affine Consultants 2026

Die sinnvollste Kombination ist in vielen Fällen kein Entweder-oder, sondern ein **Tool-Stack**:

1. **Microsoft 365 Copilot Business** für den täglichen Arbeitskontext im Tenant: E-Mail, Meetings, Angebote, Dokumentation

2. **ChatGPT** (Business oder Enterprise) für Strategie, Konzeptarbeit, Gegenprüfung und kreative Aufgaben

3. **Lagdock oder ähnliche Lösungen** als ergänzende Option, wenn datenschutzorientierte Dokumentenräume oder alternative Wissensablagen gewünscht sind

Konkrete Empfehlung:

- Copilot Business einsetzen, wenn Microsoft 365 bereits Ihr operatives Zentrum ist
- ChatGPT parallel behalten, weil es als universeller Sparringspartner oft breiter und flexibler bleibt
- Lagdock gezielt prüfen, aber nicht als pauschalen Ersatz für Copilot betrachten, sondern als möglichen Baustein für dokumentenzentrierte oder datenschutzsensiblere Use Cases

---

Fazit

Die wirtschaftlich beste Zielgruppe für Copilot Business sind 2026 **kleine bis mittlere Beratungs- und Dienstleistungsfirmen** mit hohem Kommunikationsaufkommen, wiederkehrender Dokumentarbeit und sauberem Microsoft-365-Unterbau. Ohne diese Basis ist der Nutzen begrenzt; mit dieser Basis kann Copilot sehr schnell produktiv werden[1][2][6].

Entscheidungshilfe:

- Wenn Sie täglich mehrere Stunden in Outlook, Teams und Word verbringen → **Copilot Business lohnt sich**
- Wenn Ihre SharePoint-Struktur chaotisch ist und Berechtigungen ungepflegt sind → **Zuerst Governance aufräumen, dann Copilot**
- Wenn Sie primär kreativ arbeiten und wenig Microsoft-365-Routine haben → **ChatGPT reicht oft aus**
- Wenn Sie datenschutzrechtlich besonders sensible Szenarien haben → **Lokale Datenverarbeitung (ab 2026) und DSFA nutzen**

Die Technologie ist 2026 reif. Die Frage ist nicht mehr „ob", sondern „wann" und „wie" Sie KI-Assistenz in Ihren Arbeitsalltag integrieren – und ob Sie die organisatorischen Hausaufgaben gemacht haben, damit die Investition sich auszahlt.

Quellenverzeichnis

[1] Microsoft Corporation. (2026). Microsoft 365 Copilot-Pläne und Preise – KI für Unternehmen. https://www.microsoft.com/de-de/microsoft-365-copilot/pricing

[2] Microsoft Learn. (2025). App and network requirements for Microsoft 365 Copilot admins. https://learn.microsoft.com/en-us/copilot/microsoft-365/microsoft-365-copilot-requirements

[3] Ad-hoc-news. (2025). Microsoft 365 Copilot: Lokale Datenverarbeitung für Deutschland kommt 2026. https://www.ad-hoc-news.de/boerse/ueberblick/microsoft-365-copilot-lokale-datenverarbeitung-fuer-deutschland-kommt/68359343

[4] Microsoft Learn. (2026). Copilot Studio licensing. https://learn.microsoft.com/en-us/microsoft-copilot-studio/billing-licensing

[5] Microsoft Corporation. (2026). Microsoft Copilot Studio Licensing Guide (February 2026). PDF-Dokumentation.

[6] Microsoft Learn. (2026). Data, Privacy, and Security for Microsoft 365 Copilot. https://learn.microsoft.com/en-us/copilot/microsoft-365/microsoft-365-copilot-privacy

[7] Martinsfeld. Microsoft Copilot in Teams datenschutzkonform nutzen. https://www.martinsfeld.de/infothek/microsoft-copilot-teams-datenschutz-dsgvo-compliance/

[8] Heise News. (2025). Microsoft veröffentlicht Datenschutz-Hilfen für M365 und Copilot. https://www.heise.de/news/Microsoft-veroeffentlicht-Datenschutz-Hilfen-fuer-M365-und-Copilot-11075444.html

[9] Gesellschaft für Datenschutz. (2025). Microsoft Copilot & Datenschutz: Was ist zu tun? https://gesellschaft-datenschutz.de/microsoft-copilot-datenschutz/

[10] ESET. (2025). Is ChatGPT safe? The complete 2026 security & privacy guide. https://www.eset.com/blog/en/home-topics/cybersecurity-protection/is-chatgpt-safe-2026-guide

[11] Metomic. (2026). Is ChatGPT Safe for Business in 2026? The Real Risks. https://www.metomic.io/resource-centre/is-chatgpt-a-security-risk-to-your-business

---

Über diese Analyse:

Dieses Whitepaper wurde im März 2026 erstellt und basiert auf aktuellen Herstellerinformationen, unabhängigen Fachpublikationen und Praxiserfahrungen aus dem KMU-Umfeld. Es versteht sich als neutrale Entscheidungsgrundlage und ersetzt keine individuelle Beratung durch Ihren IT-Dienstleister oder Datenschutzbeauftragten.

Nutzen Ihre Mitarbeitenden KI-Tools? Das sollten Sie wissen.

ChatGPT, Google Gemini, Microsoft Copilot – KI-Tools sind im Arbeitsalltag angekommen. Viele Mitarbeitende nutzen sie täglich, oft kostenlos, oft ohne Wissen der Geschäftsführung.

Das klingt zunächst harmlos. Ist es aber nicht.

Was passiert mit den Daten?

Kostenlose KI-Tools nutzen Ihre Eingaben standardmäßig für das Training ihrer Modelle. Was Ihre Mitarbeitenden eintippen – Kundendaten, Angebote, interne Dokumente – kann das Unternehmen verlassen, ohne dass Sie es merken.

Und: Es gibt keinen Vertrag, der Ihre Daten schützt.

Vier Risiken, die Sie kennen sollten

① Datenschutz (DSGVO)
Werden personenbezogene Daten (von Kunden, Bewerbern, Mitarbeitenden) in KI-Tools eingegeben, liegt ein meldepflichtiger Datenschutzverstoß vor – auch wenn es versehentlich passiert.

② Betriebsgeheimnisse
Kalkulationen, Strategien, Kundenlisten: Einmal eingegeben, verlassen diese Daten Ihr Unternehmen unkontrolliert.

③ Haftung der Geschäftsführung
Als Geschäftsführer haften Sie – auch für das, was Ihre Mitarbeitenden mit privaten KI-Accounts tun, wenn es Betriebsdaten betrifft.

④ Kein Überblick
Ohne Richtlinie weiß niemand, welche Tools genutzt werden. Kein Audit, kein Schutz im Ernstfall.

Was Sie jetzt tun können

✔ **Bestandsaufnahme:** Welche KI-Tools werden in Ihrem Betrieb genutzt?
✔ **Richtlinie erstellen:** Klare Regeln, was erlaubt ist – und was nicht.
✔ **Sichere Alternativen:** Es gibt datenschutzkonforme KI-Lösungen mit Auftragsverarbeitungsvertrag.
✔ **Mitarbeitende sensibilisieren:** Kurze Schulung, klare Kommunikation.

Wir unterstützen Sie dabei

PC-SPEZIALIST Hannover / HIT Systempartner GmbH hilft Ihnen, einen klaren Rahmen zu schaffen – damit Ihre Mitarbeitenden KI sicher und regelkonform nutzen können.

Sprechen Sie uns an:

Frank Böttger · fb@pcspezialist-hannover.de

Dieser Hinweis dient der Information und ersetzt keine Rechtsberatung.

Wissensdatenbank (RAG) aufbauen

Was ist Sinn und Zweck einer KI-Gestüzten Wissensdatenbank

Warum wir Word-Dateien vor dem Hochladen in die Wissensdatenbank umwandeln

Eine KI-gestützte Wissensdatenbank soll euch schnelle, präzise Antworten liefern – aber nur, wenn sie die Inhalte eurer Dokumente auch wirklich versteht. Genau hier liegt das Problem mit klassischen Word-Dateien.

Das Problem: optische Formatierung ist für die KI "Rauschen"

Word-Dokumente enthalten viele Elemente, die für das menschliche Auge hilfreich sind, für eine KI aber Störfaktoren darstellen:

- **Tabellen** – besonders verschachtelte oder mit verbundenen Zellen – werden beim Einlesen oft in eine unlogische Reihenfolge zerlegt. Eine Tabelle mit Produktname, Preis und Verfügbarkeit kann so zu einer Zeile werden, in der Preis und Produktname vertauscht erscheinen.
- **Kopf- und Fußzeilen** (z. B. Seitenzahlen, Firmenname, Datum) werden teils mitten im Fließtext eingefügt und unterbrechen einen zusammenhängenden Absatz.
- **Textfelder** liegen technisch "über" dem eigentlichen Text. Viele Auswerteprogramme überspringen sie komplett – wichtige Inhalte, die in einer Infobox oder einem Kasten stehen, gehen dadurch schlicht verloren.
- **Verschachtelte Aufzählungen** (Listen mit mehreren Ebenen, Unterpunkten, Einrückungen) verlieren beim Einlesen häufig ihre Hierarchie – aus "Hauptpunkt mit drei Unterpunkten" wird eine flache Liste ohne erkennbare Zusammengehörigkeit.

- **Eingebettete Bilder und Sonderzeichen** können die Textstruktur durcheinanderbringen, etwa wenn ein Bild mitten im Satz platziert ist und der Text danach an einer unerwarteten Stelle "wieder einsetzt".

Für Menschen ist das alles selbstverständlich lesbar. Eine KI muss diese Struktur aber erst technisch entschlüsseln – und tut das nicht immer zuverlässig.

Die Lösung: Umwandlung in ein einfacheres Format (z. B. Markdown)

Wird ein Dokument vorher in ein reduziertes Format wie Markdown umgewandelt, bleibt nur der reine Inhalt übrig: klare Überschriften, saubere Absätze und eindeutig getrennte Abschnitte. Die KI kann dadurch viel besser erkennen,

- wo ein Thema anfängt und wo es aufhört,
- was die wichtigsten Punkte sind,
- und wie einzelne Informationen inhaltlich zusammenhängen.

Praxisbeispiel: Eine mehrseitige Word-Datei mit Firmenlogo im Kopfbereich, einer verschachtelten Tabelle zu Ansprechpartnern und einem Textfeld mit "Wichtiger Hinweis" liefert nach der Umwandlung einen klar gegliederten Text mit Überschriften pro Abschnitt, einer einfachen Tabelle ohne Formatierungscode und dem Hinweistext direkt im Fließtext, statt in einer separaten, möglicherweise unsichtbaren Box.

Das Ergebnis für eure Fragen an die Wissensdatenbank

Wenn später jemand eine Frage stellt, findet die KI die passende Antwort schneller und genauer, weil sie nicht erst durch Formatierungs-Ballast "hindurchsuchen" muss. Zusätzlich verhindert die Umwandlung typische Fehlerquellen:

- Tabellen, die durch die Umwandlung unübersichtlich werden,
- Textfelder oder Kopf-/Fußzeilen, die komplett ignoriert werden,
- Sonderzeichen oder eingebettete Bilder, die den Lesefluss der KI unterbrechen.

Nicht jedes Dokument ist gleich betroffen

Ein einfaches Word-Dokument mit reinem Fließtext, wenigen Überschriften und ohne Tabellen, Textfelder oder komplexe Layouts bereitet in der Regel auch ohne Umwandlung keine größeren Probleme. Der Umwandlungsschritt lohnt sich vor allem bei Dokumenten mit:

- Tabellen,
- mehrspaltigem Layout,
- Kopf-/Fußzeilen mit relevantem Inhalt,
- Textfeldern oder Infoboxen,
- oder stark verschachtelten Aufzählungen.

Fazit

Der zusätzliche Umwandlungsschritt kostet ein paar Minuten, sorgt aber dafür, dass eure Dokumente in der Wissensdatenbank so ankommen, wie sie gemeint sind – klar strukturiert und für die KI optimal lesbar. Der konkrete Nutzen für euch: weniger falsche oder unvollständige Antworten, weil die KI eure Inhalte von Anfang an richtig erfasst.

Word → Markdown: Checkliste für saubere Konvertierungen

Kernerkenntnis: Fast alle Konvertierungsprobleme entstehen dadurch, dass Word visuelle Formatierung statt semantischer Struktur nutzt. Konverter wie Pandoc lesen Formatvorlagen (Styles) – nicht Optik. Was optisch wie eine Überschrift aussieht, aber nur fett und größer formatiert wurde, bleibt im Markdown ein normaler Absatz.

Vor dem Export eines Word-Dokuments nach Markdown diese Punkte prüfen:

---

1. Formatvorlagen statt manueller Formatierung

Überschrift 1/2/3 verwenden statt Text nur fett und größer zu setzen. Konverter erkennen ausschließlich Formatvorlagen und wandeln sie in `#`, `##`, `###` um – manuell "nachgebaute" Überschriften bleiben normaler Fließtext. **So geht's:** Start → Formatvorlagen-Galerie → „Überschrift 1" (nicht Schriftgröße/Fett manuell setzen).

2. Einheitliche Überschriften-Hierarchie

Keine Ebenen überspringen (z. B. direkt von Überschrift 1 zu Überschrift 3). Markdown bildet die Hierarchie strikt ab – ein Sprung erzeugt eine falsche Gliederungstiefe im Wiki-Artikel. **So geht's:** Vor dem Export über die Navigationsleiste (Ansicht → Navigationsbereich) die Gliederungsebenen prüfen und lückenlos durchnummerieren lassen.

3. Änderungen nachverfolgen abschließen

Track-Changes vor dem Export annehmen oder ablehnen, sonst landen Löschungen und Einfügungen doppelt im Markdown. **So geht's:** Überprüfen → Nachverfolgung → Alle Änderungen

annehmen.

4. Kommentare entfernen

Kommentare werden sonst als Fußnoten oder Störtext mitkonvertiert. **So geht's:** Überprüfen → Kommentare → Alle Kommentare löschen.

5. Keine Textfelder oder eingebetteten Objekte

Inhalte in Textfeldern, SmartArt oder OLE-Objekten werden von Konvertern oft komplett ignoriert oder nur als Bild gerendert – der Text darin geht verloren. **So geht's:** Textfelder vermeiden, Inhalt direkt in den Fließtext schreiben.

6. Tabellen einfach halten

Keine verschachtelten Tabellen, verbundenen Zellen oder Tabellen-in-Tabellen, da Markdown nur einfache Raster-Tabellen abbildet. **So geht's:** Tabelle einfügen → einfaches Raster, „Zellen verbinden" nur wenn unbedingt nötig.

7. Bilder mit Alt-Text versehen

Ohne Alt-Text fehlt im Markdown die Bildbeschreibung (`![]()`) komplett. **So geht's:** Bild markieren → rechte Maustaste → Alternativtext bearbeiten.

8. Einheitliche Listenformate

Aufzählungs- und Nummerierungsvorlagen verwenden statt manueller Bindestriche oder eigener Nummerierung per Hand – sonst erkennt der Konverter keine zusammenhängende Liste. **So geht's:** Start → Aufzählungszeichen bzw. Nummerierung aus der Galerie wählen, nicht `-` oder `1.` von Hand eintippen.

9. Code/Monospace-Text als Zeichenformat auszeichnen

Für Befehle, Dateinamen oder Codeschnipsel das Zeichenformat „HTML Code" bzw. eine Monospace-Formatvorlage nutzen statt nur die Schriftart manuell auf Courier zu ändern. Nur echte Formatvorlagen werden zu `` `Code` `` im Markdown. **So geht's:** Text markieren → Formatvorlage „HTML Code" (oder eigene Zeichenformatvorlage) zuweisen.

10. Echte Hyperlinks statt reinem Text

Links als Word-Hyperlink einfügen (Strg+K), nicht nur die URL als Text schreiben. Nur echte Hyperlinks werden zu `[Text](URL)` – reiner Text bleibt unverlinkt. **So geht's:** Text markieren → Einfügen → Link → Adresse eintragen.

11. Sparsam mit Sonderzeichen und typografischen Zeichen

Automatisch ersetzte typografische Anführungszeichen, Gedankenstriche oder Symbole können je nach Konverter-Einstellung uneinheitlich im Markdown landen. Wo es auf exakte Zeichen ankommt (z. B. Befehle, Pfade), einfache Anführungszeichen/Standardzeichen verwenden. **So geht's:** Datei → Optionen → Dokumentprüfung → AutoKorrektur-Optionen → „Anführungszeichen ersetzen" bei Bedarf deaktivieren.

12. Keine manuellen Zeilenumbrüche (Shift+Enter)

Zeilenumbrüche innerhalb eines Absatzes statt eines neuen Absatzes erzeugen im Markdown keine saubere Absatztrennung und werden je nach Konverter verschluckt oder falsch dargestellt. **So geht's:** Enter für neue Absätze verwenden, Shift+Enter vermeiden.

13. Keine Spalten-, Abschnittslayouts oder Fuß-/Endnoten

Mehrspaltige Layouts, Abschnittsumbrüche und Fuß-/Endnoten werden von Konvertern meist an falscher Stelle im Fließtext eingefügt oder ganz verworfen. **So geht's:** Einspaltiges Standardlayout verwenden, Anmerkungen direkt im Fließtext oder als eigener Absatz formulieren statt als Fußnote.

14. Ein Thema pro Dokument

Ein eigenes .docx pro Anleitung/Prozess statt Sammeldokumente – vermeidet Vermischung von Überschriftenebenen und erleichtert die spätere Aufteilung in einzelne Wiki-Artikel. **So geht's:** Für jede Anleitung eine eigene Datei anlegen statt mehrere Themen in einem Dokument zu sammeln.

Kurzfassung zum Merken: Alles, was in Word nur "gut aussieht", aber keine echte Formatvorlage, kein echtes Werkzeug (Link, Formatvorlage, Alt-Text) dahinter hat, geht bei der Konvertierung verloren oder wird falsch interpretiert.