KI im Unternehmen erfolgreich einführen – Praxisleitfaden

Wissen	Beratung	Schulung
Die Leitplanken, auf denen jede Einführung steht: Microsoft Copilot-Governance – Grundlagen, Umsetzung, Praxis	Du willst KI im eigenen Haus einführen und nicht im Pilotfegefeuer landen? Ich begleite Projekte von der Use-Case-Bewertung bis zum Regelbetrieb – sprich mich über die Kontaktseite auf boddenberg.de an.	Workshops für Führungskräfte, Key User, Datenschutz und Betriebsrat: KI-Kompetenz nach Art. 4 AI Act, Copilot-Praxis und Governance – inhouse oder online.

Es gibt einen Satz, den ich in Lenkungskreisen öfter höre als „Guten Morgen“: „Wir machen jetzt mal einen KI-Piloten.“ Drei Monate später gibt es 50 begeisterte Pilotnutzerinnen und Pilotnutzer, eine PowerPoint mit Smileys und keine einzige Zahl, mit der der CFO etwas anfangen kann. Sechs Monate später laufen die Lizenzen noch, genutzt werden sie von elf Leuten, und der Betriebsrat hat einen Anwalt. Willkommen im Pilotfegefeuer.

Wenn du **KI im Unternehmen einführen** willst, ist die Technik der leichteste Teil. Microsoft 365 Copilot ist an einem Nachmittag lizenziert. Was Monate dauert, sind die Fragen davor und danach: Welche Use Cases lohnen sich wirklich? Wie misst du Nutzen so, dass er einer Prüfung standhält? Wann brauchst du eine Datenschutz-Folgenabschätzung, was verlangt der EU AI Act heute schon, und warum ist der Betriebsrat bei praktisch jeder KI-Funktion im Spiel? Und vor allem: Was findet Copilot in deinem SharePoint, wenn er mit den Rechten deiner Leute sucht?

Dieser Leitfaden geht den Weg einmal komplett: vom Phasenmodell mit echten Go/No-Go-Toren über Business Case und Messdesign, Datenschutz, AI Act und Mitbestimmung bis zur Berechtigungsbereinigung, zur Prompt-Policy und zum Betrieb, in dem auch das Abschalten eines Use Cases ein normaler Vorgang ist. Den Gesamtüberblick über Copilot-Varianten, Funktionen und Bausteine findest du auf der Übersichtsseite Microsoft Copilot; hier geht es um die Einführung.

Fakten: KI-Einführung auf einen Blick (Stand Oktober 2026)

- Microsoft 365 Copilot kostet als Add-on 30 US-Dollar pro Nutzer und Monat bei Jahresbindung; für Organisationen bis 300 Plätze gibt es Copilot Business zum Listenpreis von 21 US-Dollar. Copilot Chat mit Websuche ist für Konten mit Entra ID ohne Aufpreis enthalten.
- Der EU AI Act ist seit 1. August 2024 in Kraft. Verbotene Praktiken und die Pflicht, KI-Kompetenz zu fördern (Art. 4), gelten seit 2. Februar 2025. Die Hochrisiko-Pflichten für eigenständige Systeme wie Bewerbervorauswahl greifen nach der Omnibus-Verordnung (EU) 2026/1744 ab 2. Dezember 2027.
- Die EU Data Boundary für Microsoft 365 ist seit Februar 2025 vollständig umgesetzt; Copilot-Verarbeitung für EU-Kunden bleibt grundsätzlich innerhalb der EU.
- Copilot sieht nur, was das fragende Konto sehen darf – aber eben alles davon, auch die zu

großzügigen Freigaben.

• Nach § 87 Abs. 1 Nr. 6 BetrVG genügt es, dass eine technische Einrichtung zur Leistungs- oder Verhaltenskontrolle objektiv geeignet ist. Copilot protokolliert Interaktionen – damit ist die Mitbestimmung praktisch immer eröffnet.

Vom Use Case zum Rollout: das Phasenmodell mit Entscheidungstoren

Ein KI-Projekt ist kein Software-Rollout mit anderem Etikett. Der Unterschied: Bei Outlook weißt du vorher, was es kann. Bei generativer KI weißt du nur, was es können soll – und ob es das in deinen Prozessen, mit deinen Daten und deinen Leuten tatsächlich tut, findest du erst heraus, wenn du es misst. Deshalb braucht die Einführung Phasen mit harten Übergängen, an denen jemand mit Budgetverantwortung Ja oder Nein sagt.

Discovery: Schmerzpunkte statt Spielwiese

Die Discovery beginnt nicht mit der Frage „Was kann KI?“, sondern mit „Wo verbrennt ihr Zeit?“. Die Antworten sind in fast jedem Unternehmen dieselben: E-Mail-Threads mit 40 Nachrichten, in denen die eigentliche Entscheidung irgendwo in Antwort 23 steht. Protokolle, die keiner schreiben will. Angebote, die aus drei alten Angeboten zusammenkopiert werden, inklusive des Kundennamens aus dem falschen Angebot. Und das Wissen über Sonderfreigaben, das im Kopf einer einzigen Kollegin lebt, die selbstverständlich gerade in Elternzeit ist.

Sammle diese Schmerzpunkte in Interviews mit den Fachbereichen, nicht im Workshop mit der IT. Für jeden Kandidaten brauchst du am Ende fünf Angaben: wie oft der Vorgang vorkommt, wie lange er heute dauert, welche Daten er braucht, wie schlimm ein Fehler wäre und wer fachlich dafür geradesteht. Ohne Häufigkeit und Dauer gibt es später keinen Business Case, ohne Fehlerfolgen keine Risikobewertung.

Das Use-Case-Portfolio: Quick Wins, Leuchttürme und Absagen

Aus der Sammlung wird ein Portfolio mit drei Schubladen. **Quick Wins** laufen mit Standardfunktionen von Copilot und liefern innerhalb von 90 Tagen messbare Ergebnisse. **Leuchttürme** greifen in Kernprozesse ein, brauchen Anbindungen an Fachsysteme, eigene Agenten oder Azure-Dienste und dauern sechs bis zwölf Monate. Die dritte Schublade wird gern vergessen: **Absagen**. Ein Portfolio ohne verworfene Ideen ist kein Portfolio, sondern ein Wunschzettel.

Bereich	Use Case	Nutzen	Daten	Risiko	Datenreife	Einordnung
Querschnitt	Meeting-Zusammenfassung mit Aufgaben	15–30 Min. Nachbereitung je Termin	Teams-Transkript	Vertraulichkeit, Mitbestimmung	hoch	Quick Win
Querschnitt	Zusammenfassung langer E-Mail-Threads	schnellerer Einstieg in Vorgänge	Postfach	gering	hoch	Quick Win
Vertrieb	Angebotsentwurf aus Vorlagen und CRM	kürzere Angebotszeit	CRM, Produktkatalog	inhaltliche Fehler	mittel	Leuchtturm
Service	Wissensagent für interne	weniger Rückfragen an	SharePoint, Tickets	Falschantworten	mittel	Leuchtturm

Bereich	Use Case	Nutzen	Daten	Risiko	Datenreife	Einordnung
	Supportfragen	Experten				
Recht/Einkauf	Vertragsprüfung auf Risikoklauseln	Vorprüfung in Minuten	Vertragsablage	Haftung bei Übersehenem	mittel	Leuchtturm mit Prüfpflicht
Buchhaltung	Rechnungseingang auslesen und kontieren	weniger Abtippen	Belege	GoBD, Fehlbuchungen	hoch	Leuchtturm (Automatisierung)
Personal	KI-Vorauswahl von Bewerbungen	Zeit in der Vorauswahl	Bewerbungen	Hochrisiko nach AI Act, Diskriminierung	mittel	zurückgestellt

Für Automatisierungsfälle wie den Rechnungseingang ist Copilot selten das richtige Werkzeug; da landest du bei Workflow-Plattformen. Welche sich wofür eignet, steht in Automatisierung mit KI in der Praxis – Power Automate, n8n, Zapier und Make im Vergleich. Für Wissensagenten und fachliche Assistenten lohnt der Blick in Copilot-Agenten in der Praxis.

K.-o.-Kriterien: was gar nicht erst in den Pilot kommt

- **Rechtliche Unzulässigkeit:** Emotionserkennung am Arbeitsplatz ist nach Art. 5 AI Act seit Februar 2025 verboten. Vollautomatisierte Entscheidungen mit rechtlicher Wirkung für Personen scheitern an Art. 22 DSGVO.
- **Fehlende Datenreife:** Ein Wissensagent auf einer Ablage mit fünf Versionen derselben Richtlinie liefert fünf verschiedene Wahrheiten. Erst aufräumen, dann anbinden.
- **Kein fachlicher Owner:** Wenn niemand im Fachbereich für die Ergebnisse geradestehen will, will auch niemand das System. Ohne Owner kein Pilot.
- **Unklare Fehlerfolgen:** Wo ein Fehler Menschen, Geld oder Haftung trifft und keine Prüfung vorgesehen werden kann, wartet der Use Case, bis es eine gibt.
- **Aufwand ohne Hebel:** Ein eigenes Modell trainieren, um ein Formular auszufüllen, das 20-mal im Jahr vorkommt, ist Hobby, kein Projekt.

Vier Phasen, drei Tore

Das Phasenmodell ist banal, das Entscheidende sind die Tore dazwischen. An jedem Tor liegt eine Vorlage mit vorher vereinbarten Kriterien auf dem Tisch, und der Lenkungskreis entscheidet: weiter, nachbessern oder beerdigen. „Wir schauen mal, wie es läuft“ ist keine dieser drei Optionen.

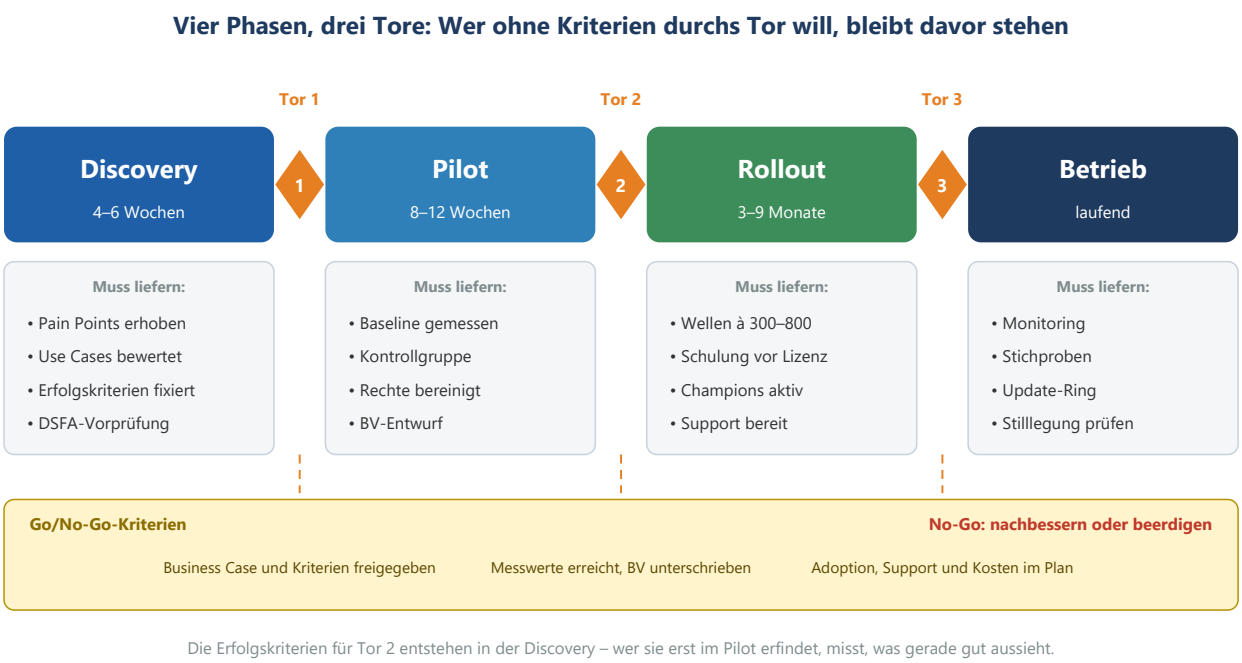


Abbildung 1: Phasenmodell der KI-Einführung mit Go/No-Go-Toren und den Ergebnissen, die jede Phase liefern muss

Phase	Dauer	Ergebnisse (Pflicht)	Tor danach: Go, wenn …
Discovery	4–6 Wochen	priorisiertes Portfolio, Business-Case-Entwurf, Erfolgskriterien je Use Case, DSFA-Schwellwertprüfung, Pilotplan, Information des Betriebsrats	Budget freigegeben, Kriterien schriftlich fixiert, Pilotgruppe und Kontrollgruppe benannt
Pilot	8–12 Wochen	Baseline-Messung, bereinigte Berechtigungen im Pilotbereich, Schulung, Messwerte, Fehlerstichproben, Entwurf Betriebsvereinbarung, Pilotbericht	Kriterien erreicht (nicht: „Nutzer sind zufrieden“), DSFA abgeschlossen, Vereinbarung unterschriftsreif
Rollout	3–9 Monate	Wellenplan, Schulung vor Lizenzzuweisung, Champions-Netz, Supportprozess, Lizenz-Rückholung	Adoption je Welle im Zielkorridor, Supportlast beherrschbar, Kosten im Plan
Betrieb	laufend	Monitoring, monatliche Stichproben, Update-Ring, Quartalsreview je Use Case	jährliche Rezertifizierung jedes Use Cases – oder Stilllegung

Warum Piloten ohne Kriterien nie in Produktion gehen

Ein Pilot ohne vorher definierte Erfolgskriterien hat ein strukturelles Problem: Er kann nicht scheitern, also kann er auch nicht gelingen. Am Ende stehen Anekdoten gegen Lizenzkosten, und Anekdoten verlieren. Ich habe einen Pilot erlebt, in dem nach zwölf Wochen eine Umfrage zeigte, dass 87 Prozent der Gruppe Copilot „hilfreich“ fanden. Die Geschäftsführung fragte, was „hilfreich“ in Euro sei. Die Antwort war Schweigen, der Rollout wurde auf das nächste Geschäftsjahr verschoben, und dann auf das übernächste.

Die Kriterien gehören deshalb in die Discovery, und sie müssen so formuliert sein, dass jemand von außen nachprüfen kann, ob sie erreicht sind. Etwa: „Die Erstellungszeit eines Standardangebots sinkt im Messfenster gegenüber der Kontrollgruppe um mindestens 25 Prozent, bei höchstens gleicher Korrekturquote in der Freigabe.“ Das ist unbequem, weil es scheitern kann. Genau das macht es wertvoll.

Warnung: Das Pilotfegefeuer erkennst du an drei Symptomen

- Der Pilot wird verlängert, statt entschieden – „wir brauchen noch mehr Daten“.
- Die Pilotgruppe besteht aus Freiwilligen, die ohnehin jede neue Technik lieben.
- Der Pilotbericht enthält Zitate, aber keine Baseline.

Wenn zwei davon zutreffen, hilft kein weiterer Monat Pilot, sondern ein Tor-Termin mit Kriterien.

Plattformwahl: welches Werkzeug für welchen Fall

Im Microsoft-Umfeld ist die Plattformfrage weniger ein Entweder-oder als eine Treppe. Du beginnst mit dem, was ohne Entwicklung funktioniert, und steigst nur dort höher, wo der Use Case es verlangt. Jede Stufe bringt mehr Möglichkeiten und mehr Verantwortung für Betrieb, Kosten und Datenschutz.

Stufe	Werkzeug	Typische Fälle	Kostenmodell	Was du betreiben musst
1	Copilot Chat	Texte, Recherche im Web, hochgeladene Dateien	ohne Aufpreis zur Microsoft-365-Lizenz	Nutzungsregeln, Websuche-Einstellung
2	Microsoft 365 Copilot	Arbeit mit E-Mails, Dateien, Chats und Meetings aus dem Graph	Add-on pro Nutzer und Monat	Berechtigungen, Labels, DLP, Audit
3	Agenten mit Copilot Studio	Fachassistenten, Wissensagenten, Prozessschritte	Kontingente bzw. Verbrauch nach Copilot Credits	Agenten-Lebenszyklus, Konnektoren, Freigaben
4	Azure OpenAI in Microsoft Foundry	eigene Anwendungen, Massenverarbeitung, Integration in Fachsysteme	Verbrauch nach Tokens	Entwicklung, Monitoring, Kostenkontrolle, Abuse Monitoring

Eine Einordnung, wie sich die Plattformen anderer Anbieter dazu verhalten, findest du im Überblick über generative KI-Plattformen. Für die Lizenzdetails einschließlich der Bündelangebote lohnt sich Microsoft 365 Copilot Lizenzierung.

Die ersten 90 Tage

Für die meisten Häuser mittlerer Größe hat sich der folgende Takt bewährt. Er ist ehrgeizig, aber machbar, wenn die Berechtigungsbereinigung parallel zur Discovery startet und nicht erst nach dem Pilot.

Zeitraum	Ziel	Ergebnisse	Verantwortlich	Messgröße
Tag 1–30	Initiative aufgesetzt, Fokus klar	Lenkungskreis, Portfolio, Erfolgskriterien, Betriebsrat und Datenschutz informiert, Oversharing-Bericht	Sponsor in der Geschäftsleitung, Projektleitung	Kriterien je Use Case schriftlich; Tor 1 entschieden
Tag 31–60	Pilot startklar	Baseline erhoben, Pilot- und Kontrollgruppe, kritische Freigaben bereinigt, Schulung, DLP im Testmodus	IT-Plattform, Product Owner je Use Case	Baseline vollständig; Schulungsquote 100 % vor Lizenz

Zeitraum	Ziel	Ergebnisse	Verantwortlich	Messgröße
Tag 61–90	Pilot im Messfenster	Messwerte, Fehlerstichproben, Feedback, DSFA abgeschlossen, Entwurf Betriebsvereinbarung	Product Owner, Datenschutz	Kriterien erreicht ja/nein; Vorlage für Tor 2

Spätestens hier merkst du, dass 90 Tage für den Pilot selbst knapp sind, wenn du den Neuheitseffekt herausrechnen willst. Plane das Messfenster lieber zwei Wochen länger, als nach Tag 90 mit unbrauchbaren Zahlen vor den Lenkungskreis zu treten.

Business Case und Messung: was KI wirklich bringt

Kein Thema produziert so viel Folienkunst wie der ROI von KI. Die Logik ist simpel – gesparte Zeit mal Stundensatz gegen Kosten –, aber jeder Faktor darin ist eine Einladung zur Selbsttäuschung. Ein Business Case, der einer Revision standhält, unterscheidet zwischen gefühlter, gemessener und tatsächlich verwerteter Zeitersparnis.

Die Kostenseite vollständig erfassen

Die Lizenz ist der sichtbarste Posten und meistens nicht der größte. Wer nur die Lizenz rechnet, hat im zweiten Quartal eine Nachtragsvorlage.

Kostenblock	Typische Größenordnung	Anmerkung
Microsoft 365 Copilot	30 US-Dollar pro Platz und Monat bei Jahresbindung	monatliche Zahlung oder Laufzeit kostet Aufschlag; Copilot Business bis 300 Plätze günstiger
Agenten und Copilot Studio	nach Copilot Credits bzw. Kontingent	Nutzung lizenzierter Personen ist teils abgedeckt, Agenten für alle anderen kosten Verbrauch
Azure OpenAI / Foundry	nach Tokens, stark schwankend	Budgetwarnungen und Kontingente je Anwendung zwingend
Berechtigungsbereinigung	einmalig, oft 50–150 Personentage	größter unterschätzter Posten in Umgebungen mit Altlasten
Schulung	2–4 Stunden je Person plus Trainer	Arbeitszeit der Teilnehmerinnen und Teilnehmer gehört in die Rechnung
Betrieb	0,5–2 Vollzeitstellen je nach Größe	Monitoring, Stichproben, Update-Ring, Champions-Netz
Sicherheits- und Compliance-Funktionen	ggf. Upgrade oder Add-ons	Purview-Funktionen für DLP auf Copilot und Audit sind lizenzabhängig

Zeitersparnis realistisch bewerten

Die gefährlichste Zahl im Business Case ist die Selbstauskunft. Frag eine Pilotgruppe, wie viel Zeit Copilot ihr spart, und du bekommst Werte zwischen 20 und 60 Minuten pro Tag. Die britische Regierung hat das in einer großen Erprobung mit rund 20.000 Personen gemacht und kam auf 26 Minuten pro Tag – laut Selbstauskunft. Eine parallele, genauer gemessene Auswertung in einem britischen Ministerium fand dagegen keine belastbaren Belege für höhere Produktivität, wohl aber hohe Zufriedenheit. Beides kann gleichzeitig wahr sein: Menschen empfinden Entlastung, die in der Stoppuhr kaum auftaucht.

Dazu kommt die Verwertungsfrage. Zwölf gesparte Minuten pro Tag, verteilt auf sechs Zwei-Minuten-Häppchen, werden nicht zu einem zusätzlichen Angebot, sondern zu einem zusätzlichen Kaffee. Nutzen entsteht nur dort, wo gesparte Zeit in mehr Durchsatz, kürzere Durchlaufzeit, weniger Überstunden oder weniger externe Vergabe umgesetzt wird. Deshalb rechne ich im Business Case mit einem Verwertungsfaktor, typischerweise zwischen 30 und 60 Prozent, und begründe ihn je Use Case.

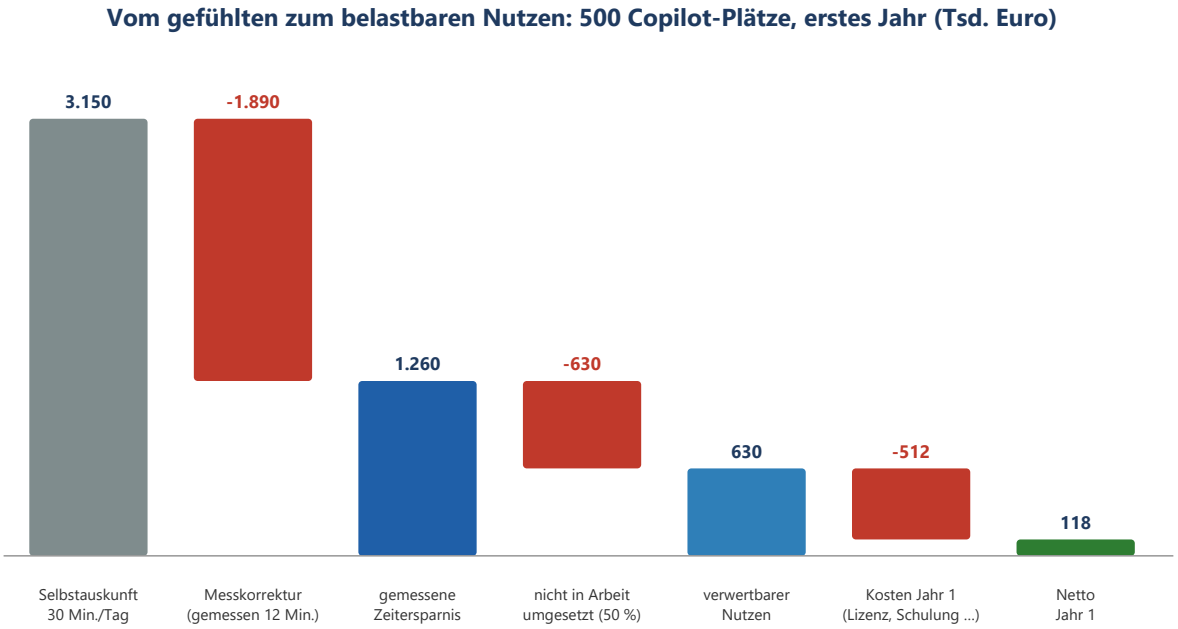


Abbildung 2: Vom gefühlten zum belastbaren Nutzen: wie Messkorrektur, Verwertungsfaktor und Gesamtkosten einen Business Case schrumpfen lassen

Ein Rechenbeispiel mit ehrlichen Annahmen

Nimm 500 Copilot-Plätze, einen Vollkostensatz von 60 Euro pro Stunde und 210 Arbeitstage. Die Pilotgruppe meldet 30 Minuten pro Tag; gemessen gegen die Kontrollgruppe sind es zwölf. Verwertbar ist davon die Hälfte.

Position	Rechnung	Betrag pro Jahr
Nutzen laut Selbstauskunft	500 × 0,5 h × 210 × 60 €	3.150.000 €
Nutzen laut Messung	500 × 0,2 h × 210 × 60 €	1.260.000 €
Verwertbarer Nutzen (50 %)	1.260.000 € × 0,5	630.000 €
Lizenzen	500 × rund 28 € × 12	168.000 €
Schulung (Arbeitszeit)	500 × 4 h × 60 €	120.000 €
Schulung (Trainer, Material)	pauschal	30.000 €
Berechtigungsbereinigung und Projekt	einmalig	120.000 €
Betrieb	0,5 Vollzeitstelle	50.000 €
Azure-Verbrauch für eine	rund 2.000 € × 12	24.000 €

Position	Rechnung	Betrag pro Jahr
Zusatzanwendung		
Netto im ersten Jahr	630.000 € – 512.000 €	**118.000 €**

Das ist immer noch positiv, und ab dem zweiten Jahr fallen Bereinigung, Projekt und ein Großteil der Schulung weg. Aber es ist eine andere Geschichte als die 3,15 Millionen aus der Umfrage – und es ist die Geschichte, die du dem Controlling erzählen kannst, ohne rot zu werden. Wenn die Rechnung nur mit der Selbstauskunft aufgeht, geht sie nicht auf.

Tipp: Lizenzen in Wellen und mit Rückholung

Lizenziere nicht alle auf einmal. Gib allen Copilot Chat, die Vollversion den Rollen mit viel Text-, Mail- und Meetingarbeit, und hole Plätze, die 60 Tage lang kaum genutzt werden, zurück. Eine Rückholquote von 10 bis 15 Prozent ist normal und spart mehr als jede Rabattverhandlung.

Messung im Pilot: Baseline, Kennzahlen, Kontrollgruppe

Belastbar wird der Nutzen nur durch ein Messdesign, das du vor dem Pilot festlegst. Drei Bausteine sind Pflicht. Erstens eine **Baseline**: Bevor die erste Lizenz aktiv ist, misst du zwei bis vier Wochen lang die Kennzahlen der späteren Pilotprozesse. Zweitens eine **Kontrollgruppe**, die dieselbe Arbeit im selben Zeitraum ohne Copilot macht. Drittens ein **Messfenster**, das erst beginnt, wenn sich die Nutzung eingeschwungen hat.

Kennzahl	Messung	Quelle	Fallstrick
Bearbeitungszeit je Vorgang	Zeitstempel Start/Ende oder Stichproben mit Zeiterfassung	Ticket- oder Workflow-System, Arbeitstagebuch	Vorgänge mit unterschiedlicher Komplexität vermischen
Durchlaufzeit	Eingang bis Freigabe	Workflow, CRM, DMS	Wartezeiten anderer Stellen verfälschen
Fehler- bzw. Korrekturquote	Anteil Ergebnisse mit inhaltlicher Nacharbeit in der Freigabe	Freigabeprozess, Stichprobe	nur Tippfehler zählen, Sachfehler übersehen
Nutzungsintensität	aktive Tage je Person, Interaktionen je Anwendung	Copilot-Dashboard, Nutzungsberichte	hohe Nutzung mit Nutzen verwechseln
Zufriedenheit	kurze Pulsbefragung	Umfrage	als Nutzenbeleg statt als Akzeptanzsignal werten

Die Kontrollgruppe ist im Betrieb oft schwer durchzusetzen, weil niemand der Abteilung sein will, die „noch nicht darf“. Ein Kompromiss, der funktioniert: zeitversetzter Start. Gruppe B bekommt Copilot vier Wochen später und ist bis dahin die Kontrollgruppe. Das ist fair, und du bekommst deine Vergleichswerte.

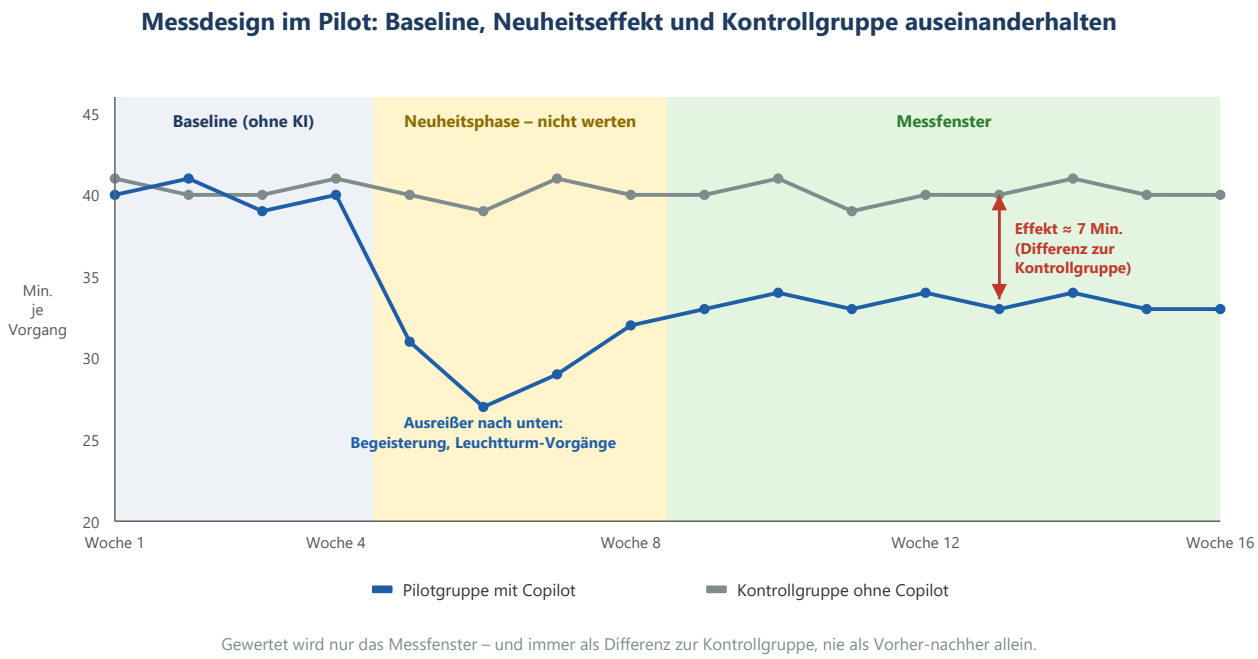


Abbildung 3: Messdesign im Pilot: Der Effekt ist die Differenz zur Kontrollgruppe im Messfenster, nicht der Ausschlag in den ersten Wochen

Neuheitseffekt und Tagesform

In den ersten Wochen nach dem Start passiert zweierlei. Die Fans probieren alles aus und suchen sich unbewusst die Vorgänge, bei denen KI glänzt. Und alle wissen, dass gemessen wird. Beides drückt die Bearbeitungszeit nach unten, ohne dass es nach drei Monaten noch trägt. Deshalb wertest du die ersten zwei bis vier Wochen nicht. Tagesform, Urlaubswellen und Quartalsenden glättest du, indem Pilot- und Kontrollgruppe im selben Zeitraum gemessen werden und du Mediane statt Mittelwerte berichtest. Wie man aus diesen Rohdaten dauerhafte Kennzahlen für Copilot und Agenten macht, beschreibt Metriken zur Messung des Erfolgs der Microsoft 365 Copilot KI-Agenten.

Wichtig: Messen ist mitbestimmungspflichtig

Zeitstempel je Vorgang und Person sind Leistungsdaten. Kläre mit dem Betriebsrat vor dem Pilot, dass Messdaten nur aggregiert ab einer Mindestgruppengröße ausgewertet, nach dem Pilot gelöscht und nie für individuelle Bewertungen genutzt werden. Dann ist die Messung kein Konflikt, sondern gemeinsames Interesse.

Recht: DSGVO, EU AI Act und Betriebsrat

Die drei Rechtsthemen der KI-Einführung werden gern nacheinander abgearbeitet, und zwar von drei Leuten, die nicht miteinander reden. Dabei hängen sie zusammen: Die Datenschutz-Folgenabschätzung braucht die Angaben aus der AI-Act-Einordnung, die Betriebsvereinbarung regelt dieselben Protokolle, die in der DSFA als Risiko stehen. Ein gemeinsamer Arbeitsstand spart Wochen.

Datenschutz-Folgenabschätzung: wann sie Pflicht ist

Art. 35 DSGVO verlangt eine DSFA, wenn eine Verarbeitung – insbesondere bei Verwendung neuer Technologien – voraussichtlich ein hohes Risiko für die Rechte und Freiheiten von Personen mit sich bringt. Die Praxis prüft das anhand der Kriterien der europäischen Datenschutzaufsichten:

Treffen zwei oder mehr zu, ist eine DSFA in der Regel fällig. Bei KI im Büro sind fast immer mindestens zwei erfüllt: innovative Technologie, Verarbeitung von Beschäftigtendaten in einem Abhängigkeitsverhältnis, häufig auch umfangreiche Verarbeitung und systematische Auswertung. Die deutschen Aufsichtsbehörden führen den Einsatz von KI zur Bewertung persönlicher Aspekte oder zur Steuerung der Interaktion mit Personen zudem ausdrücklich auf ihrer Muss-Liste.

Meine Faustregel: Für Microsoft 365 Copilot im unternehmensweiten Einsatz machst du die DSFA. Selbst wenn deine Datenschutzbeauftragte zu dem Ergebnis kommt, dass sie formal nicht zwingend wäre, ist sie das beste Dokument, um gegenüber Betriebsrat, Aufsicht und Revision zu zeigen, dass du die Risiken kennst. Für einzelne Agenten oder Anwendungen genügt dann oft eine Ergänzung zur Basis-DSFA.

Risiko	Worum es geht	Typische Maßnahme
Zweckbindung	Copilot verarbeitet Inhalte, die für andere Zwecke erhoben wurden; Meeting-Transkripte werden zu Wissensquellen	Zwecke je Use Case festlegen, Transkription nur mit Kennzeichnung, Aufbewahrung begrenzen
Oversharing	Copilot macht zugängliche, aber vergessene Daten auffindbar	Berechtigungsbereinigung, Restricted Content Discovery, Labels
Drittlandtransfer	Verarbeitung oder Support-Zugriffe außerhalb der EU, Websuche über Bing, Zusatzmodelle anderer Anbieter	EU Data Boundary prüfen, Websuche und Zusatzmodelle bewusst konfigurieren
Protokollierung von Nutzerverhalten	Audit-Einträge, gespeicherte Prompts, Nutzungsberichte je Person	Zugriff auf Auswertungen beschränken, Pseudonymisierung, Löschfristen, Betriebsvereinbarung
Richtigkeit	Halluzinationen über Personen	Prüfpflicht, keine Entscheidungen über Personen allein auf KI-Basis
Betroffenenrechte	Auskunft und Löschung erstrecken sich auf Copilot-Interaktionen	Prozess über eDiscovery und Aufbewahrungsrichtlinien für Copilot-Interaktionen

Auftragsverarbeitung und Datenresidenz konkret prüfen

Microsoft verarbeitet Copilot-Daten als Auftragsverarbeiter nach dem Datenschutznachtrag zu den Produkten und Diensten (DPA), der Teil deiner Lizenzverträge ist. Einen separaten AV-Vertrag musst du nicht aushandeln, aber du musst ihn lesen, dokumentieren und gegen deine Konfiguration halten. Diese Punkte prüfe ich in jeder Einführung:

Prüfpunkt	Microsoft 365 Copilot	Azure OpenAI in Foundry
Vertragsgrundlage	DPA und Produktbestimmungen; Copilot ist als Dienst unter dem DPA erfasst	DPA und Produktbestimmungen für Azure
Training mit Kundendaten	nein, Prompts und Antworten trainieren keine Basismodelle	nein, Eingaben und Ausgaben trainieren keine Basismodelle
Speicherort	Interaktionen im Postfach der Person, im Tenant-Speicherort; EU Data Boundary seit Februar 2025	Ressource in gewählter Region, z. B. Germany West Central oder Sweden Central
Verarbeitungsort der	für EU-Kunden innerhalb der EU; lokale	abhängig vom Bereitstellungstyp: Regional,

Prüfpunkt	Microsoft 365 Copilot	Azure OpenAI in Foundry
Modelle	Verarbeitung wird für EU und EFTA regional umgesetzt	Data Zone EU oder Global – Global kann außerhalb der EU verarbeiten
Websuche	Anfragen gehen an Bing; dafür gelten abweichende Bedingungen – per Richtlinie abschaltbar	nur bei bewusst eingebundenen Such-Tools
Zusatzmodelle	Modelle weiterer Anbieter sind eine Admin-Einstellung; in EU-Tenants standardmäßig aus	Modellauswahl je Bereitstellung
Missbrauchsüberwachung	durch Microsoft im Dienst	Prompts können bis zu 30 Tage zur Missbrauchsprüfung gespeichert werden; Ausnahme auf Antrag möglich

Warnung: „Global“ heißt global

Bei Azure OpenAI ist der Bereitstellungstyp Global oft der günstigste und schnellste. Er bedeutet aber, dass die Inferenz in jedem Azure-Rechenzentrum weltweit laufen kann. Für personenbezogene Daten nimmst du Data Zone EU oder eine regionale Bereitstellung – und schreibst das in die Architekturrichtlinie, bevor ein Entwickler aus Kostengründen umstellt.

EU AI Act: was heute gilt und was noch kommt

Der AI Act unterscheidet zwischen Anbietern, die KI-Systeme entwickeln oder unter eigenem Namen in Verkehr bringen, und Betreibern, die sie in eigener Verantwortung einsetzen. Als Unternehmen, das Copilot nutzt, bist du Betreiber. Für Betreiber sieht die Lage im Oktober 2026 so aus:

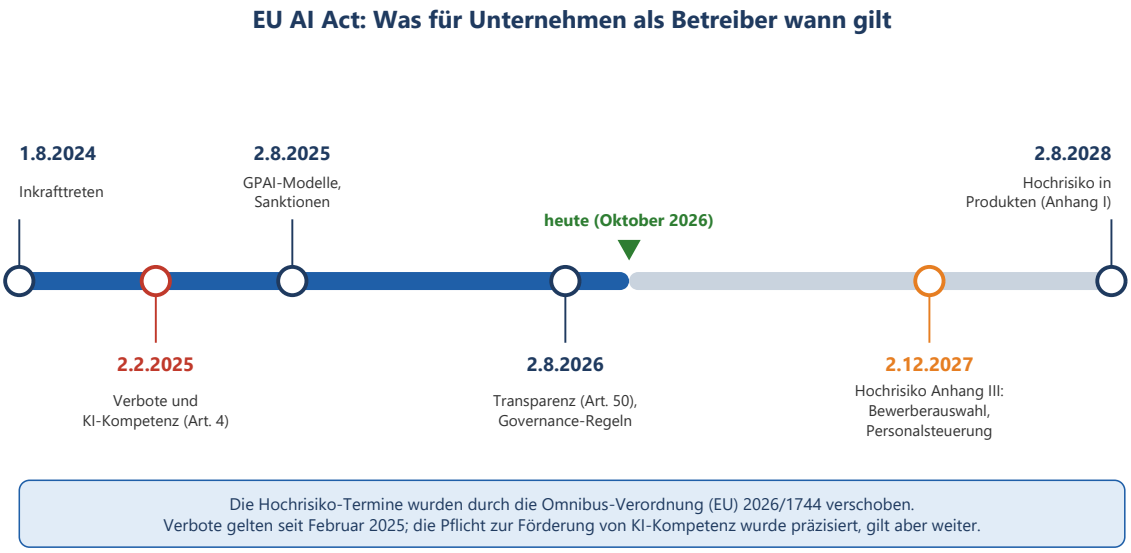


Abbildung 4: Zeitstrahl des EU AI Act aus Sicht von Unternehmen als Betreiber, mit den durch die Omnibus-Verordnung verschobenen Hochrisiko-Terminen

- **Verbote (seit 2. Februar 2025):** Für das Büro relevant ist vor allem das Verbot der Emotionserkennung am Arbeitsplatz. Ein Tool, das aus Stimme oder Mimik in Teams-Meetings die Stimmung des Teams ableitet, ist tabu – auch als „Wellbeing-Feature".
- **KI-Kompetenz (Art. 4, seit 2. Februar 2025):** Betreiber müssen Maßnahmen ergreifen, die die KI-Kompetenz ihrer Leute fördern, abhängig von Vorwissen, Einsatzkontext und betroffenen Personen. Die Omnibus-Verordnung hat klargestellt, dass kein bestimmtes Kompetenzniveau garantiert werden muss – die Pflicht, Maßnahmen zu ergreifen und sie nachweisen zu können, bleibt.
- **Transparenz (Art. 50, seit 2. August 2026):** Wer Deepfakes erzeugt oder KI-generierte Texte zu Themen von öffentlichem Interesse veröffentlicht, muss das offenlegen, es sei denn, der Text wurde redaktionell geprüft und jemand trägt die Verantwortung dafür.
- **Hochrisiko (ab 2. Dezember 2027):** KI zur Auswahl, Filterung oder Bewertung von Bewerbungen sowie für Entscheidungen über Beförderung, Kündigung, Aufgabenzuweisung oder Leistungsbewertung gilt als hochriskant (Anhang III Nr. 4). Betreiber müssen dann unter anderem menschliche Aufsicht durch kompetente Personen sicherstellen, Protokolle aufbewahren und Arbeitnehmervertretung und Belegschaft vor dem Einsatz informieren.

Einsatz	Einordnung	Was du tun musst
Copilot für Entwürfe, Zusammenfassungen, Recherche	kein Hochrisiko	KI-Kompetenz fördern, interne Regeln, bei Veröffentlichung Art. 50 beachten
Chatbot gegenüber Kunden	Transparenzpflicht	Kunden müssen erkennen, dass sie mit einer KI sprechen
Bewerbervorauswahl oder Ranking durch KI	Hochrisiko ab Dezember 2027	menschliche Aufsicht, Protokolle, Information der Belegschaft, Nutzung nach Gebrauchsanweisung, DSFA
Copilot formuliert eine Absage, die eine Person prüft	kein Hochrisiko, solange die Auswahl menschlich ist	Prüfung dokumentieren, keine verdeckte Vorauswahl
Emotionserkennung in Meetings	verboten	nicht einsetzen

Human-in-the-Loop ist dabei kein Stempel, den jemand blind setzt. Gemeint ist eine Person, die die Fähigkeiten und Grenzen des Systems versteht, die Ausgabe tatsächlich prüfen kann, Zeit dafür hat und auch befugt ist, anders zu entscheiden. Wer 400 Bewerbungen vom System sortieren lässt und die oberen 20 „prüft", hat keinen Menschen in der Schleife, sondern einen Menschen als Feigenblatt. Meine ausführliche – und nicht ganz unparteiische – Bewertung der Verordnung steht in Der EU AI-Act – Regulierung der Künstlichen Intelligenz. Viel Schaden, wenig Nutzen.

Mitbestimmung nach BetrVG: warum Copilot fast immer darunter fällt

§ 87 Abs. 1 Nr. 6 BetrVG gibt dem Betriebsrat ein Mitbestimmungsrecht bei technischen Einrichtungen, die dazu bestimmt sind, Verhalten oder Leistung zu überwachen. Das Bundesarbeitsgericht legt das seit Jahrzehnten weit aus: Es genügt, dass die Einrichtung objektiv geeignet ist, entsprechende Daten zu erheben oder auszuwerten. Ob der Arbeitgeber das vorhat, spielt keine Rolle. Copilot erzeugt Audit-Einträge mit Person, Zeitpunkt und Anwendung, speichert Prompts und Antworten im Postfach und liefert Nutzungsberichte. Damit ist die Mitbestimmung bei jeder ernsthaften Einführung eröffnet.

Dazu kommen weitere Rechte, die seit der Modernisierung des Betriebsverfassungsgesetzes 2021 KI ausdrücklich nennen: Unterrichtung und Beratung bei der Planung von KI-Einsatz (§ 90 Abs. 1

Nr. 3), die Mitbestimmung bei Auswahlrichtlinien auch dann, wenn sie mit KI erstellt werden (§ 95 Abs. 2a), und das Recht, bei KI-Themen einen Sachverständigen hinzuzuziehen, dessen Erforderlichkeit gesetzlich vermutet wird (§ 80 Abs. 3). Das Arbeitsgericht Hamburg hat 2024 entschieden, dass die bloße Erlaubnis, ChatGPT über private Konten im Browser zu nutzen, nicht mitbestimmungspflichtig ist, weil der Arbeitgeber dabei keine Daten erhält. Das lässt sich auf ein unternehmensweit bereitgestelltes Copilot mit Audit nicht übertragen.

Funktion	Warum mitbestimmungsrelevant	Regelungsvorschlag
Audit-Protokoll der Copilot-Interaktionen	Person, Zeit, Anwendung, genutzte Ressourcen	Zugriff nur für Sicherheitsvorfälle, Vier-Augen-Prinzip, Löschfrist
Gespeicherte Prompts und Antworten	Inhalte der Arbeit, über eDiscovery durchsuchbar	eDiscovery nur bei definiertem Anlass und mit Beteiligung
Nutzungsberichte im Admin-Bereich	Aktivität je Person	Namen in Berichten verbergen (Standardeinstellung)
Copilot-Dashboard und Analytics	Nutzung und Muster nach Gruppen	nur aggregiert ab Mindestgruppengröße, keine Auswertung einzelner Personen
Transkription und Zusammenfassung von Meetings	Gesprochenes wird festgehalten und auswertbar	Kennzeichnung, Widerspruchsmöglichkeit, keine Auswertung von Redeanteilen
Agenten mit Zugriff auf Personaldaten	mögliche Bewertung von Personen	Einzelfreigabe durch Lenkungskreis und Betriebsrat

Was eine Betriebsvereinbarung sinnvoll regelt

Bewährt hat sich eine Rahmenvereinbarung für KI-Einsatz mit Anlagen je Anwendung. Die Rahmenvereinbarung bleibt stabil, die Anlagen wachsen mit dem Portfolio. So musst du nicht für jeden neuen Agenten das ganze Paket neu verhandeln.

- **Geltungsbereich und Zwecke:** welche Systeme, welche Use Cases, Verfahren für neue Anlagen.
- **Ausschluss der Leistungs- und Verhaltenskontrolle:** keine personenbezogene Auswertung von Nutzungsdaten, Verwertungsverbot für Erkenntnisse, die entgegen der Vereinbarung gewonnen wurden.
- **Zugriff auf Protokolle:** wer unter welchen Voraussetzungen Audit und eDiscovery nutzen darf, mit Beteiligung und Dokumentation.
- **Freiwilligkeit und Verantwortung:** Prüfpflicht für KI-Ergebnisse, keine Nachteile, wenn jemand KI nicht nutzt oder KI-Ergebnisse verwirft.
- **Qualifizierung:** Anspruch auf Schulung vor Lizenzzuweisung, Lernzeit innerhalb der Arbeitszeit.
- **Beschäftigungssicherung und Folgen:** Umgang mit wegfallenden Tätigkeiten, Beratung bei Änderungen von Arbeitsabläufen.
- **Evaluation:** fester Termin nach sechs bis zwölf Monaten, Informationsrechte zu neuen Funktionen und Modellwechseln.

Tipp: Den Betriebsrat in den Pilot holen

Nimm ein Mitglied des Betriebsrats in die Pilotgruppe. Wer Copilot selbst nutzt, verhandelt über reale Funktionen statt über Befürchtungen. In meinen Projekten wurde die Vereinbarung danach meist schneller fertig als die Berechtigungsbereinigung.

Technik und Regeln: Berechtigungen, Labels und Prompt-Policy

Der Satz „Copilot hält sich an die Berechtigungen“ ist korrekt und gleichzeitig das Problem. In einer typischen Microsoft-365-Umgebung mit ein paar Jahren auf dem Buckel sind die Berechtigungen nämlich nicht das, was sich jemand ausgedacht hat, sondern das, was übrig geblieben ist. Mein Klassiker: die Excel-Datei mit der Gehaltsrunde, abgelegt in einem Team, das vor vier Jahren für eine Weihnachtsfeier angelegt wurde und dem „Jeder außer externen Benutzern“ beitreten durfte. Vor Copilot hat die niemand gefunden. Mit Copilot reicht die Frage „Was verdient eigentlich die Abteilungsleitung im Vertrieb?“.

Oversharing aufdecken

Mit Microsoft 365 Copilot ist SharePoint Advanced Management ohne Aufpreis enthalten, und das ist kein Zufall. Die Werkzeuge daraus sind die Grundlage jeder Bereinigung:

- **Berichte zur Datenzugriffs-Governance:** zeigen Websites mit Freigabelinks für jeden, für die ganze Organisation und mit Freigaben an „Jeder außer externen Benutzern“, dazu Websites mit vielen Berechtigten und eine Oversharing-Basisauswertung.
- **Überprüfung des Websitezugriffs:** delegiert die Prüfung an die Besitzer der Websites, statt dass die IT über Inhalte entscheidet, die sie nicht kennt.
- **Restricted Content Discovery:** nimmt einzelne Websites aus Copilot und organisationsweiter Suche heraus, ohne die Berechtigungen zu ändern. Ideal als Sofortmaßnahme für Bereiche, deren Bereinigung länger dauert.
- **Microsoft Purview DSPM für KI:** bewertet Oversharing-Risiken mit Blick auf Copilot, zeigt sensible Inhalte in Interaktionen und schlägt passende Richtlinien vor.

Jeder-Zugriffe entfernen, ohne den Betrieb lahmzulegen

Die Reihenfolge entscheidet. Zuerst die Websites mit sensiblen Inhalten und breitesten Freigaben, dann die großen, dann der Rest. Für jede Website gilt: Besitzer bestimmen, Freigaben an alle Mitglieder der Organisation durch konkrete Gruppen ersetzen, organisationsweite Links entfernen, verwaiste Websites archivieren. Den Standard für neue Freigabelinks stellst du auf „bestimmte Personen“ um, damit nicht morgen nachwächst, was du heute geschnitten hast.

Warnung: Nicht mit dem Rasenmäher

Wer „Jeder außer externen Benutzern“ tenantweit in einem Rutsch entfernt, legt Intranet, Vorlagenbibliotheken und die Kantinenseite gleich mit lahm. Erst inventarisieren, dann pro Website mit den Besitzern entscheiden. Für die Übergangszeit ist Restricted Content Discovery das Pflaster, nicht die Kettensäge.

Sensitivity Labels und DLP so koppeln, dass Copilot mitspielt

Ein Vertraulichkeitslabel ohne Verschlüsselung ist ein Aufkleber. Copilot liest das Label, zeigt es in Antworten an und vererbt es an neu erzeugte Inhalte – aber er nutzt die Datei trotzdem, wenn das Konto sie öffnen darf. Wirkung entfaltet ein Label erst in zwei Kombinationen:

Mechanismus	Wirkung auf Copilot	Einsatz
Label mit Verschlüsselung	Copilot nutzt Inhalte nur, wenn das Konto die Nutzungsrechte zum Anzeigen und Extrahieren hat	streng vertrauliche Dokumente mit festem Personenkreis
DLP-Richtlinie für Copilot	Dateien und E-Mails mit bestimmten Labels werden nicht zusammengefasst oder für Antworten verarbeitet; Prompts mit bestimmten sensiblen Informationen lassen sich blockieren	vertrauliche Inhalte, die viele öffnen dürfen, aber nicht in KI-Antworten gehören
Label-Vererbung	neue Inhalte aus Copilot übernehmen das höchste Label der Quellen	verhindert Herabstufung durch Umformulieren
Standardlabel und Pflichtkennzeichnung	erhöht die Abdeckung, damit DLP überhaupt etwas zum Greifen hat	Bibliotheken mit Standardlabel, Pflicht beim Speichern

Teste DLP-Richtlinien für Copilot zuerst im Simulationsmodus. Du willst wissen, wie viele Treffer es gibt, bevor die Vorstandsassistenz feststellt, dass Copilot die Vorstandsvorlage nicht mehr zusammenfasst. Wie diese Schichten in einer Gesamtarchitektur zusammenspielen, beschreibt Microsoft Copilot-Governance – Grundlagen, Umsetzung, Praxis.

Prompt-Policy: was Technik durchsetzt und was nur Vertrauen kann

Die meisten Nutzungsrichtlinien, die ich zu sehen bekomme, sind zwölf Seiten lang und werden von niemandem gelesen. Eine gute Richtlinie passt auf zwei Seiten und trennt ehrlich zwischen dem, was die Technik erzwingt, und dem, wofür du auf Urteilskraft angewiesen bist. Diese Trennung ist wichtig, weil sie die Schulung fokussiert: Was die Technik verhindert, musst du nicht dreimal erklären.

Regel	Durchsetzung	Technisch erzwingbar?
Keine vertraulichen Daten in öffentliche KI-Dienste	Endpoint-DLP und Browser-Schutz blockieren das Einfügen in Consumer-KI; Defender for Cloud Apps sperrt nicht freigegebene KI-Apps	weitgehend
Nur freigegebene KI-Werkzeuge nutzen	App-Sperren, Freigabeliste für Agenten, Tenant-Einstellungen	weitgehend
Zugriff nur von verwalteten Geräten mit MFA	Conditional Access	ja
Keine Websuche für interne Inhalte	Richtlinie für die Websuche in Copilot	ja, aber grob
Streng vertrauliche Inhalte nicht in KI-Antworten	Labels mit Verschlüsselung, DLP für Copilot	ja
Ergebnisse vor Weitergabe prüfen	Schulung, Freigabeprozess bei kritischen Dokumenten	nein
Keine Entscheidungen über Personen	Richtlinie, Prozessgestaltung, Betriebsvereinbarung	nein

Regel	Durchsetzung	Technisch erzwingbar?
allein auf KI-Basis		
KI-Einsatz bei Veröffentlichung kenntlich machen	Redaktionsprozess	nein
Personenbezug in Prompts minimieren	Schulung, ergänzend DLP für sensible Informationstypen	teilweise

Verstöße erkennen, ohne ein Kontrollklima zu schaffen

Hier verläuft die Grenze zwischen Governance und Misstrauen. Die Regel, die in meinen Projekten funktioniert: Du schaust auf Muster, nicht auf Menschen. DLP-Treffer werden zunächst als Hinweis an die Person selbst ausgespielt, mit der Möglichkeit, eine Begründung anzugeben. Auswertungen an die Sicherheitsabteilung gehen aggregiert und pseudonymisiert. Nur bei einem konkreten Verdacht auf einen Sicherheitsvorfall wird personenbezogen aufgelöst, nach dem Verfahren, das in der Betriebsvereinbarung steht.

Und wenn die Zahl der blockierten Einfügeversuche in öffentliche KI-Dienste hochgeht, ist das selten ein Disziplinarproblem, sondern ein Angebotsproblem: Die Leute brauchen etwas, das sie offiziell nicht haben. Die beste Maßnahme gegen Schatten-KI ist ein freigegebenes Werkzeug, das gut genug ist.

KI-Kompetenz aufbauen: Schulung und Change

Art. 4 AI Act macht aus Schulung eine Pflicht, aber die bessere Begründung ist banal: Ungeschulte Leute nutzen Copilot wie eine Suchmaschine, sind nach drei Versuchen enttäuscht und kommen nie wieder. Die Schulung findet deshalb vor der Lizenzzuweisung statt, nicht danach.

Zielgruppe	Format	Inhalte
Alle Nutzerinnen und Nutzer	60–90 Minuten live plus Übungsaufgaben	Funktionsweise, Grenzen und Halluzinationen, Nutzungsrichtlinie, gute Prompts für die eigenen Aufgaben
Führungskräfte	60 Minuten	Erwartungsmanagement, Vorbildrolle, Umgang mit Ängsten, keine Leistungsbewertung über KI-Nutzung
Champions und Key User	halber bis ganzer Tag, danach monatlicher Austausch	fortgeschrittene Prompts, Agenten, Fehlerbilder, Support im Team
Betriebsrat und Datenschutz	halber Tag	Protokolle, Auswertungen, Einstellungen, Rechtsrahmen
IT und Entwicklung	nach Bedarf	Purview, SharePoint Advanced Management, Copilot Studio, Foundry

Die Angst vor dem Jobverlust verschwindet nicht durch eine Townhall. Sie verschwindet, wenn die Geschäftsleitung konkret sagt, was mit gewonnener Zeit passiert, und wenn sie das in der Betriebsvereinbarung festschreibt. Und wenn Führungskräfte Copilot selbst nutzen und offen erzählen, wo er sie blamiert hat. Nichts schafft mehr Vertrauen als ein Bereichsleiter, der zugibt, dass Copilot ihm eine Besprechung zusammengefasst hat, an der er gar nicht teilgenommen hatte – und dass er das fast so weitergeleitet hätte.

Betrieb nach dem Rollout: Rollen, Monitoring, Stilllegung

Nach dem Rollout beginnt die Phase, für die niemand ein Budget beantragt hat. Copilot ändert sich monatlich, Modelle werden ausgetauscht, Agenten wachsen wie Unkraut, und die Fehler, die im Pilot niemand bemerkt hat, tauchen jetzt in Kundenkorrespondenz auf. Betrieb heißt hier nicht nur Verfügbarkeit, sondern Qualität.

Rollen, die es wirklich braucht

Rolle	Aufgabe	Typische Besetzung
KI-Lenkungskreis	Tore entscheiden, Portfolio steuern, Stilllegungen beschließen	Geschäftsleitung, IT, Fachbereiche, Datenschutz, Informationssicherheit, Betriebsrat beratend
Product Owner je Use Case	fachlicher Nutzen, Kennzahlen, Stichproben, Anforderungen	Fachbereich
Data Owner	Qualität und Freigabe der Datenquellen, Klassifizierung	Fachbereich der Datenquelle
Plattformverantwortung	Tenant-Einstellungen, Lizenzen, Update-Ring, Agentenverwaltung	IT
Datenschutz und Informationssicherheit	DSFA fortschreiben, Vorfälle, DLP-Auswertung	Datenschutzbeauftragter, ISB
Champions	Hilfe im Team, Feedback, Beispiele	eine Person je 30–50 Nutzer

Laufendes Monitoring

Das Monitoring hat drei Ebenen. Auf der **Nutzungsebene** zeigen das Copilot-Dashboard und die Nutzungsberichte aktive Personen, Anwendungen und Trends – aggregiert, wie mit dem Betriebsrat vereinbart. Auf der **Sicherheitsebene** liefern Purview Audit, DSPM für KI und DLP-Berichte Hinweise auf sensible Inhalte und Regelverstöße. Auf der **Qualitätsebene** stehen die Kennzahlen der Use Cases und die Stichproben. Ein Monatsbericht an den Lenkungskreis mit einer Seite pro Use Case reicht; mehr liest niemand.

Wenn Agenten ins Spiel kommen, verschiebt sich der Schwerpunkt: Wer hat welchen Agenten gebaut, mit welchen Rechten, auf welche Daten? Die zentrale Verwaltung von Agenten als eigene Identitäten beschreibt Agent 365 – KI-Agenten als Steuerzentrale für Microsoft 365, Sicherheit und Automatisierung.

Halluzinationen und Fehlentscheidungen: Stichproben statt Hoffnung

Halluzinationen lassen sich nicht abschalten, nur seltener und folgenloser machen. Seltener durch saubere Datenquellen und Agenten mit klar begrenztem Wissen, folgenloser durch Prüfung. Für jeden Use Case mit Außenwirkung gehört eine monatliche Stichprobe zum Betrieb: Der Product Owner zieht 20 bis 30 Ergebnisse, prüft sie gegen die Quellen und ordnet Fehler in drei Klassen ein – kritisch (falsche Fakten mit Folgen), wesentlich (falsche oder fehlende Inhalte, die die Freigabe bemerken sollte) und kosmetisch.

Steigt die Quote kritischer Fehler über die vereinbarte Schwelle, geht der Use Case zurück vor den Lenkungskreis. Und wenn eine Fehlentscheidung passiert ist, wird sie wie ein Sicherheitsvorfall behandelt: Ursache klären, Prüfschritt anpassen, Schulung nachschärfen – ohne Schuldsuche bei

der Person, die dem System geglaubt hat. Wer dafür bestraft wird, meldet beim nächsten Mal nichts mehr.

Modell- und Feature-Updates beherrschen

Microsoft tauscht die Modelle unter Copilot regelmäßig aus und liefert neue Funktionen im Wochentakt. Das ist meist gut, aber jede Änderung kann einen geprüften Use Case verändern. Drei Maßnahmen halten das beherrschbar: Eine kleine Gruppe aus Champions und Plattformteam erhält Neuerungen über den Vorabrelease-Ring zuerst. Das Nachrichtencenter wird wöchentlich auf Copilot-Änderungen gesichtet, mit Ampel für Datenschutz und Betriebsrat. Und nach jedem Modellwechsel laufen die Stichproben der kritischen Use Cases außerplanmäßig. Was in den nächsten Monaten zu erwarten ist, verfolge ich in Erwartete Neuerungen in Microsoft Copilot in 2026; zur Einordnung der Modelle dahinter lohnt Q1-2026-Update der großen LLMs – Chancen, Risiken, Konsequenzen für Unternehmen.

Wann ein Use Case stillgelegt wird

Ein Portfolio, aus dem nie etwas entfernt wird, verkommt zum Museum. Lege die Stilllegungskriterien gleich zu Beginn fest, damit das Abschalten kein Gesichtsverlust ist, sondern Routine:

Kriterium	Schwelle (Beispiel)	Folge
Nutzung	weniger als 20 % der Zielgruppe aktiv über drei Monate	Ursachenanalyse, dann Neustart oder Stilllegung
Qualität	kritische Fehlerquote über 2 % in zwei Stichproben in Folge	Aussetzen bis zur Korrektur
Wirtschaftlichkeit	Kosten je Vorgang über dem gemessenen Nutzen	Stilllegung oder Umbau
Recht	neue Einordnung, etwa als Hochrisiko, ohne Erfüllung der Pflichten	sofortige Aussetzung
Ersatz	Standardfunktion deckt den Fall ab	eigene Lösung abbauen
Datenquelle	Quelle entfällt oder hat keinen Owner mehr	Agent bzw. Anbindung entfernen

Risikoregister für den Betrieb

Risiko	Ursache	Gegenmaßnahme	Owner
Halluzinationen	Modell ergänzt Plausibles statt Belegtes	Prüfpflicht, Stichproben, begrenzte Agenten	Product Owner
Datenabfluss	Einfügen in öffentliche KI, Oversharing	DLP, Browser-Schutz, Bereinigung, Restricted Content Discovery	Informationssicherheit
Schatten-KI	fehlendes oder zu schwaches offizielles Angebot	Copilot Chat für alle, App-Sperren, schnelle Freigabeprozesse	IT
Fehlentscheidung und Haftung	ungeprüfte Übernahme	Human-in-the-Loop mit Befugnis, dokumentierte Freigabe	Fachbereichsleitung

Risiko	Ursache	Gegenmaßnahme	Owner
Urheberrecht	KI-Ausgabe übernimmt geschützte Inhalte	Prüfung vor Veröffentlichung; Microsoft übernimmt für Copilot-Ausgaben unter Bedingungen die Verteidigung gegen Urheberrechtsansprüche	Rechtsabteilung
Kostenexplosion	Agenten- und Token-Verbrauch ohne Grenzen	Budgets, Kontingente, Monatsbericht	Plattformverantwortung

Fazit: Erst die unbequeme Arbeit, dann die Lizenzen

Wer **KI im Unternehmen einführen** will, entscheidet den Erfolg in den Wochen vor der ersten Lizenz. Erfolgskriterien, die scheitern können. Eine Baseline, bevor jemand begeistert ist. Berechtigungen, die nicht nur auf dem Papier stimmen. Eine DSFA und eine Betriebsvereinbarung, die dieselben Protokolle regeln. Und Tore, an denen tatsächlich Nein gesagt werden darf.

Die Technik ist inzwischen erstaunlich reif: Labels wirken auf Copilot, DLP greift in Antworten und Prompts, SharePoint Advanced Management findet Oversharing, und die Daten bleiben in der EU. Was sich nicht lizenzieren lässt, ist die Disziplin, ehrlich zu messen und Use Cases abzuschalten, die nicht tragen. Genau die unterscheidet ein KI-Programm von einem teuren Abonnement für höflichere E-Mails.

Wenn du nur drei Dinge mitnimmst: Leg die Kriterien für Tor 2 fest, bevor der Pilot beginnt. Rechne mit gemessener und verwerteter Zeit, nie mit der Umfrage. Und hol den Betriebsrat in die Pilotgruppe. Der Rest ist Handwerk – und für das Handwerk gibt es diesen Leitfaden.

Weiterlesen: Passende Beiträge

› Microsoft Copilot-Governance – Grundlagen, Umsetzung, Praxis
› Microsoft 365 Copilot Lizenzierung
› Copilot-Einführung bei der Harry Hase AG
› Metriken zur Messung des Erfolgs der Microsoft 365 Copilot KI-Agenten
› Copilot-Agenten in der Praxis

Häufige Fragen zur KI-Einführung

Müssen wir den Betriebsrat an der KI-Einführung beteiligen?

Wenn es einen Betriebsrat gibt, ja. Copilot protokolliert Interaktionen und ist damit objektiv geeignet, Verhalten und Leistung zu überwachen; das genügt für § 87 Abs. 1 Nr. 6 BetrVG. Dazu kommen Unterrichtungs- und Beratungsrechte nach § 90 und das Recht auf einen Sachverständigen. Fang in der Discovery an, nicht nach dem Pilot.

Wie setzen wir Microsoft 365 Copilot DSGVO-konform ein?

Grundlage ist der Datenschutznachtrag von Microsoft, der Copilot als Auftragsverarbeitung erfasst. Prompts und Antworten trainieren keine Basismodelle, die Verarbeitung bleibt für EU-Kunden in der EU Data Boundary. Deine Hausaufgaben: DSFA, Berechtigungsbereinigung, Labels und DLP, bewusste Entscheidung über Websuche und Zusatzmodelle, Löschfristen für Copilot-Interaktionen.

Was kostet Microsoft 365 Copilot und welche Lizenzen brauchen wir?

Du brauchst eine geeignete Basislizenz wie Microsoft 365 E3, E5, Business Standard oder Business Premium und das Copilot-Add-on für 30 US-Dollar pro Nutzer und Monat bei Jahresbindung. Bis 300 Plätze gibt es Copilot Business günstiger. Copilot Chat ist ohne Aufpreis enthalten. Plane zusätzlich Schulung, Bereinigung und Betrieb ein – die kosten im ersten Jahr oft so viel wie die Lizenzen.

Warum nicht einfach kostenlose KI-Tools wie ChatGPT nutzen?

Weil du bei Consumer-Diensten keine Auftragsverarbeitung, keine Kontrolle über Speicherung und Training und keinen Zugriff auf deine Unternehmensdaten mit Berechtigungsprüfung hast. Für Texte ohne interne Daten ist Copilot Chat die abgesicherte Alternative ohne Aufpreis; damit nimmst du der Schatten-KI den Grund.

Wie finden wir die richtigen KI-Anwendungsfälle?

Frag die Fachbereiche nach Zeitfressern, nicht nach KI-Ideen. Bewerte jeden Kandidaten nach Häufigkeit, Dauer, Datenreife, Fehlerfolgen und fachlichem Owner. Starte mit zwei bis drei Quick Wins aus Standardfunktionen und hebe Leuchttürme mit Systemanbindung für die zweite Runde auf.

Wie messen wir den ROI einer KI-Einführung?

Mit Baseline vor dem Start, Kontrollgruppe und einem Messfenster nach dem Neuheitseffekt. Gemessene Zeitersparnis multiplizierst du mit einem begründeten Verwertungsfaktor und dem Vollkostensatz und stellst sie allen Kosten gegenüber. Selbstauskünfte sind ein Akzeptanzsignal, kein Nutzenbeleg.

Wie stellen wir sicher, dass KI-Ergebnisse stimmen?

Gar nicht vollständig – deshalb gibt es die Prüfpflicht für alles mit Außenwirkung, monatliche Stichproben je Use Case mit Fehlerklassen und eine Schwelle, ab der ein Use Case ausgesetzt wird. Saubere, aktuelle Datenquellen und eng begrenzte Agenten senken die Fehlerquote deutlich.

Was tun, wenn die Belegschaft Angst um ihre Arbeitsplätze hat?

Konkret werden: Was passiert mit gewonnener Zeit, welche Qualifizierung gibt es, was steht dazu in der Betriebsvereinbarung? Führungskräfte, die Copilot selbst nutzen und offen über seine Grenzen reden, überzeugen mehr als jede Hochglanzkampagne.

Welche Aufgaben sollte man nicht der KI überlassen?

Entscheidungen über Menschen, rechtlich bindende Aussagen, Freigaben mit Sicherheitsfolgen und alles, wo niemand die Ausgabe sinnvoll prüfen kann. Emotionserkennung am Arbeitsplatz ist ohnehin verboten. KI bereitet vor, der Mensch entscheidet – und muss dafür Zeit und Befugnis haben.

Was bedeutet der EU AI Act konkret für den Office-Einsatz?

Für Copilot im Büro vor allem zwei Dinge: Maßnahmen zur KI-Kompetenz nach Art. 4, die du nachweisen kannst, und Transparenz, wenn du KI-Inhalte veröffentlichst. Hochrisiko-Pflichten betreffen dich erst, wenn KI über Bewerbungen, Beförderungen oder Leistungsbewertung mitentscheidet – ab 2. Dezember 2027.

Wie lange dauert eine KI-Einführung?

Discovery und Pilot brauchen realistisch vier bis fünf Monate, der Rollout in Wellen je nach Größe weitere drei bis neun Monate. Die Lizenzaktivierung dauert einen Tag, die Berechtigungsbereinigung in gewachsenen Umgebungen oft am längsten.