

DIGITALE PRIVATSPHÄRE & DATENSCHUTZ

Purview für Enterprise AI Datenschutz

Executive Summary

Microsoft hat sich zum Aufsichtsorgan über Werkzeuge gemacht, die Microsoft gar nicht gehören. Seit Mai 2026 liegt im Purview-Portal ein nativer Connector für Anthropic Claude Enterprise, den für ChatGPT Enterprise gibt es schon länger. Beide ziehen Prompts, Antworten und Aktivitätsereignisse über die Compliance-API des jeweiligen Anbieters ab und kippen sie in DSPM for AI, ins Unified Audit Log, in eDiscovery und in die Aufbewahrung. Das ist bemerkenswert: Deine Governance-Konsole für generative KI ist ausgerechnet die von Microsoft, auch wenn kein einziges Token durch einen Microsoft-Dienst gelaufen ist.

Die gute Nachricht für alle, die vor einem Aufsichtsrat sitzen und über KI reden müssen: Du bekommst endlich eine belastbare Antwort auf die Fragen „Wer nutzt was?“, „Welche personenbezogenen Daten sind da hineingelaufen?“ und „Können wir das im Streitfall herausgeben?“. Auskunftersuchen nach Artikel 15 DSGVO, Löschkonzepte, Aufbewahrungsfristen, Rechenschaftspflicht nach Artikel 5 Absatz 2 — dafür ist der Connector genau das richtige Werkzeug.

Die schlechte Nachricht, und die musst du in jedem Kickoff in der ersten Viertelstunde loswerden: Der Connector erzwingt gar nichts. Er ist ein Read-only-Journal mit rund einem Tag Verzögerung. Vertraulichkeitsbezeichnungen, Verschlüsselung und vor allem Data Loss Prevention sind für beide Enterprise-KI-Apps schlicht nicht unterstützt. Wer glaubt, mit der Connector-Aktivierung sei die DSGVO in KI-Workflows durchgesetzt, hat eine Videoüberwachung eingebaut und die Tür offen gelassen. Durchsetzung passiert an der Kante — Endpoint-DLP, Browser-DLP in Edge, Netzwerk-DLP über SASE — und das ist ein komplett eigenes Projekt mit eigenem Budget und eigenem Zeitplan.

FAKTEN · Der Stand der Dinge im September 2026

Für ChatGPT Enterprise und für Claude Enterprise gilt inzwischen dieselbe Funktionsmatrix: DSPM for AI, Auditing, Datenklassifizierung, Insider Risk Management, Communication Compliance, eDiscovery, Data Lifecycle Management und Compliance Manager sind unterstützt. Nicht unterstützt sind Vertraulichkeitsbezeichnungen, Verschlüsselung ohne Bezeichnung und Data Loss Prevention. Wer noch mit der Analyse vom Juni argumentiert, in der Claude nur Reporting und Auditing konnte, verkauft dem Kunden einen veralteten Stand.

Worum geht es im Detail?

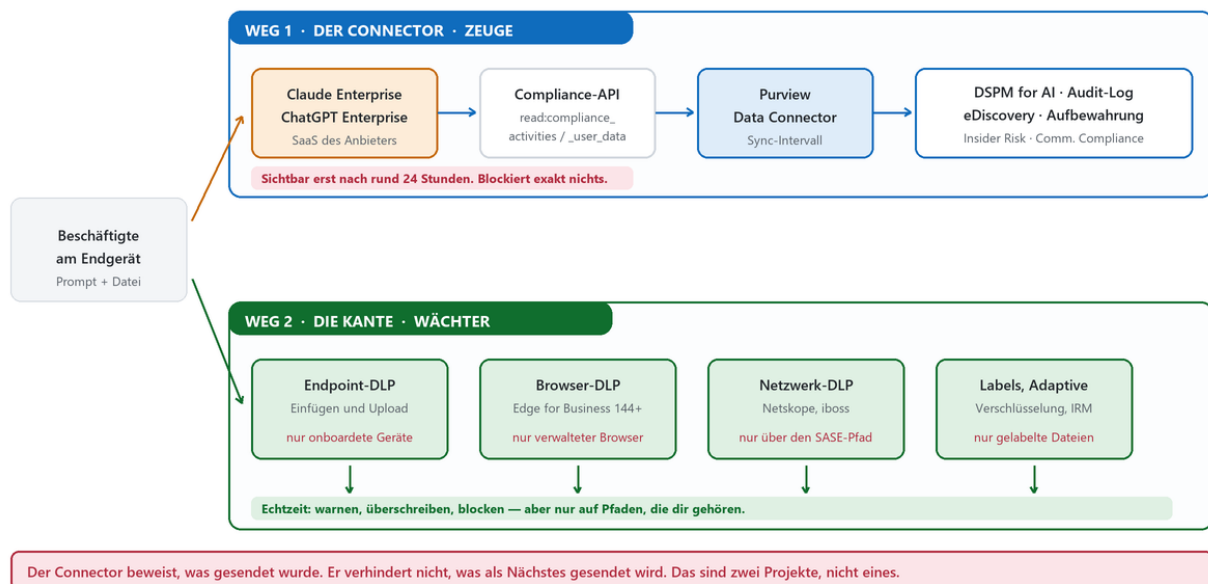
Purview sortiert KI in drei Schubladen, und diese Sortierung entscheidet über alles, was du danach an Kontrollen bekommst. Erstens die Copilot-Erfahrungen: Microsoft 365 Copilot, Security Copilot. Hier sitzt Microsoft im Datenpfad und kann inline eingreifen, inklusive echtem DLP. Zweitens die Enterprise-KI-Apps: ChatGPT Enterprise, Claude Enterprise, Microsoft Foundry. Hier sitzt Microsoft nicht im Datenpfad, sondern klopft per API hinterher an und fragt höflich, was

denn so passiert ist. Drittens die anderen KI-Apps: das private ChatGPT-Konto, Gemini, DeepSeek und was das Team sonst noch im Browser offen hat. Da bleibt nur die Erkennung über Browser und Netzwerk.

Der technische Trick hinter Schublade zwei ist elegant und heißt Compliance-API. Anthropic hat sie für Claude Enterprise und Claude Platform freigeschaltet, OpenAI hat ein Gegenstück für ChatGPT Enterprise. Purview authentifiziert sich mit einem Zugriffsschlüssel und holt sich die Daten im gewählten Intervall ab. Bei Anthropic gibt es zwei Datenbereiche, die du getrennt anhaken kannst: den Activity Feed mit Anmeldungen, Administrationsaktionen und Konfigurationsänderungen (Scope read:compliance_activities) und die Chat Conversations mit dem vollständigen Nachrichtenverlauf inklusive Prompts und Antworten (Scope read:compliance_user_data). Erfasst werden dabei auch Projekte, hochgeladene Dateien und Sitzungsinhalte aus Cowork und Claude Code — das ist deutlich mehr, als die meisten beim Setzen des Hakens vermuten.

Zwei Wege, zwei Wahrheiten

Wie Purview an KI-Interaktionen kommt — und wo tatsächlich blockiert wird



Der Connector protokolliert, die Kante entscheidet. Wer beides in ein Projekt packt, liefert am Ende keines von beiden.

Genau hier liegt der Denkfehler, der in jedem zweiten Workshop auftaucht. Die Compliance-API liefert Daten nach der Interaktion. Bei ChatGPT Enterprise dokumentiert Microsoft eine Verzögerung von rund 24 Stunden, bei Claude erscheinen die Daten ebenfalls erst nach etwa einem Tag in den DSPM-Reports. Ein Mitarbeiter, der um neun Uhr morgens die vollständige Inhaltsliste in einen Prompt kopiert, taucht am nächsten Vormittag im Activity Explorer auf. Da hat der Anwalt der Gegenseite schon zu Ende gefrühstückt. Der Connector ist ein hervorragendes Beweismittel und ein miserabler Türsteher.

WARNUNG · Ein zweites Datenschutzproblem, das du selbst erzeugst

Wenn du den Bereich Chat Conversations aktivierst, kopierst du vollständige Mitarbeitergespräche in eine zweite Plattform. Das ist eine eigenständige Verarbeitung mit eigener Rechtsgrundlage, eigener Zweckbindung und eigenem Eintrag im Verzeichnis von Verarbeitungstätigkeiten. In Deutschland kommt die Mitbestimmung nach Paragraph 87 Absatz 1 Nummer 6

Betriebsverfassungsgesetz dazu, weil eine technische Einrichtung entsteht, die zur Verhaltens- und Leistungskontrolle geeignet ist. Wer den Haken ohne Betriebsrat setzt, hat kein Compliance-Projekt, sondern einen Termin bei der Einigungsstelle.

Die Funktionsmatrix hat sich im August 2026 spürbar bewegt. Claude Enterprise stand zunächst mit zwei Häkchen und neun Kreuzen da und wurde entsprechend als Sichtbarkeitswerkzeug abgestempelt. Inzwischen zieht die Dokumentation mit ChatGPT Enterprise gleich. Was in beiden Fällen fehlt, ist identisch — und es ist ausgerechnet der Teil, den die Geschäftsführung meint, wenn sie das Wort Schutz benutzt.

Purview-Funktion	M365 Copilot	ChatGPT Enterprise	Claude Enterprise
DSPM for AI und Reporting	ja	ja	ja
Auditing im Unified Audit Log	ja	ja	ja
Datenklassifizierung (SIT, trainierbar)	ja	ja	ja, seit 08/2026
Insider Risk Management	ja	ja	ja, seit 08/2026
Communication Compliance	ja	ja	ja, seit 08/2026
eDiscovery	ja	ja	ja, seit 08/2026
Data Lifecycle Management	ja	ja	ja, seit 08/2026
Compliance Manager	ja	ja	ja, seit 08/2026
Vertraulichkeitsbezeichnungen	ja	nein	nein
Verschlüsselung ohne Bezeichnung	ja	nein	nein
Data Loss Prevention, präventiv	ja	nein	nein
Wirkzeitpunkt	inline, Echtzeit	danach, ca. 24 h	danach, ca. 1 Tag

Ein Detail, das reihenweise Projekte gegen die Wand fährt: Der Inhaltsbereich gilt nur für Claude Enterprise, nicht für Claude Platform. Wer über die Entwickler-API arbeitet, bekommt den Activity Feed, aber keine Prompts. Wenn deine Data-Science-Abteilung also über die API mit Kundendaten hantiert, ist sie im Connector nur als Statistik sichtbar. Der zweite Klassiker ist die Abrechnung: Ohne erteilte Pay-as-you-go-Zustimmung und ohne verknüpft Azure-Abonnement passiert schlicht nichts. Das ist der mit Abstand häufigste Blocker beim Rollout, und er wird zuverlässig erst im Termin mit dem Vorstand entdeckt.

Voraussetzung	ChatGPT Enterprise	Anthropic Claude Enterprise
Plan beim Anbieter	ChatGPT Enterprise inklusive Workspace-ID	Anthropic Enterprise, nicht Free, Pro oder Max
Schlüssel	API-Key, Scoping per E-Mail an den OpenAI-Support	Compliance Access Key, erzeugt vom Primary Owner

Ort beim Anbieter	Workspace-Einstellungen	Organization Settings, Data and Privacy
Rolle in Purview	Data Source Administrator, Data Reader	Data Connector Admin
Ablage des Geheimnisses	Azure Key Vault	direkt im Connector-Assistenten
Datenbereiche	Interaktionen und Aktivitäten	Activity Feed, Chat Conversations, einzeln wählbar
Abrechnung	Pay-as-you-go, Azure-Abo zwingend	Pay-as-you-go, Azure-Abo zwingend
Reifegrad	allgemein verfügbar	Preview

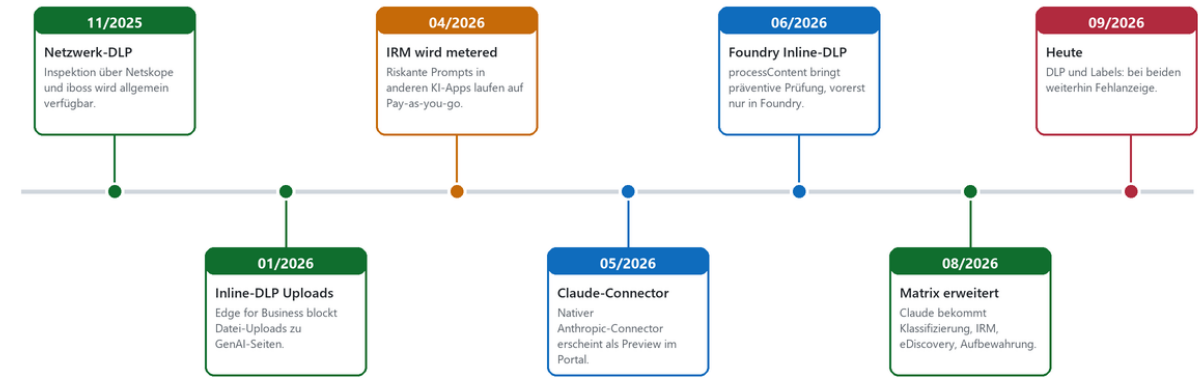
Und noch ein Punkt, der in der Angebotsphase gern unterschlagen wird: Die Erfassung von Prompts und Antworten für Insider Risk Management, Communication Compliance, eDiscovery und Aufbewahrung setzt eine Sammlungsrichtlinie voraus. Die heißt „DSPM for AI - Capture interactions for enterprise AI apps“ und entsteht aus der Empfehlung „Secure interactions from enterprise apps“. Ohne sie hast du Metadaten und sonst nichts. Für die Aufbewahrung wählst du im Übrigen den Speicherort „Enterprise AI apps“, nicht Exchange und nicht SharePoint. Wer das verwechselt, wundert sich drei Monate lang über eine leere Richtlinie.

Was sind Chancen? Was sind Risiken?

Die Chance ist strategisch größer, als sie auf den ersten Blick aussieht. Microsoft baut hier eine Kontrollebene für fremde KI-Systeme, und das Muster aus anbieterseitiger Compliance-API und Purview-Connector skaliert. Anthropic hat seine Schnittstelle bereits für über sechzig Sicherheitsplattformen geöffnet, von CrowdStrike über Splunk bis ServiceNow. Für dich heißt das: eine Konsole, ein Auditorentermin, ein Rechenschaftsnachweis — statt drei Anbieterportale mit drei Exportformaten und drei genervten Administratoren. Der Markt der unabhängigen KI-Governance-Startups mit ihren hübschen Standalone-Dashboards wird gerade sehr still.

Die Ausbaustufen — und warum der Stand von Juni 2026 schon veraltet ist

Purview-Kontrollen für KI-Anwendungen, die Microsoft nicht gehören



Merkatz für die Planung: Die Funktionsmatrix ändert sich schneller, als dein Change-Advisory-Board tagt. Prüfe sie vor jedem Steering.

Zehn Monate, sieben Änderungen. Jede Funktionsmatrix in deinem Foliensatz hat ein Verfallsdatum.

Das Risiko hat drei Gesichter. Erstens die Lücke zwischen Wahrnehmung und Wirkung. Der Vorstand hört „DSGVO in KI-Workflows durchgesetzt“ und denkt an Verhinderung, du lieferst Dokumentation. Wenn du das nicht selbst offenlegst, legt es der erste Vorfall offen — und dann bist du derjenige, der es hätte wissen müssen.

Zweitens die Kosten. Alles hier läuft über verbrauchsabhängige Zähler. Audit rechnet nach verarbeiteten Datensätzen ab, Communication Compliance nach Textdatensätzen zu je 1.000 Zeichen, Data Lifecycle Management nach Interaktionen unter Aufbewahrungsrichtlinie, Insider Risk Management nach Verarbeitungseinheiten zu je 10.000 Benutzeraktivitäten, Netzwerk- und Browser-DLP nach Anfragen. Seit April 2026 ist auch die Erkennung riskanter Prompts in anderen KI-Apps verbrauchsabhängig. Drei Prompts pro Person und Tag mal viertausend Beschäftigte mal zwei Richtungen ergibt eine Zahl, die dein Controlling nicht mag und die im Angebot meistens fehlt.

WICHTIG · Rechne vorher, nicht in der zweiten Rechnungsperiode

Ein Prompt und eine Antwort sind zwei getrennte Interaktionen. Eine ausführliche Antwort mit 7.500 Zeichen zählt in Communication Compliance als acht Textdatensätze. Wer Chat Conversations für die gesamte Belegschaft aktiviert und gleichzeitig eine Aufbewahrung über zehn Jahre darüberlegt, erzeugt eine Kostenkurve, die jeden Monat steigt und die man nachträglich nicht mehr abstellen kann, ohne das eigene Löschkonzept zu brechen. Modelliere die Zähler im Azure-Preisrechner, bevor du die Zustimmung erteilst.

Drittens das Risiko der Scheinabdeckung. Der Connector sieht Claude Enterprise überall, geräte- und browserunabhängig, weil er an der Quelle abholt statt Netzwerkverkehr abzufangen. Das ist echt stark und schlägt jedes Proxy-Konstrukt. Aber er sieht eben nur, was in eurer Enterprise-Organisation passiert. Der Kollege, der abends privat weitermacht, weil das Firmenkonto ein Nutzungslimit hat, taucht dort nie auf. Für den brauchst du Netzwerk- und Browser-DLP, und dort gibt es harte Grenzen: Endpoint-DLP wertet nur Typen vertraulicher Informationen aus, ignoriert Inhalte unter dreißig Zeichen und deckelt bei vier Megabyte. Vertraulichkeitsbezeichnungen selbst werden beim Einfügen nicht ausgewertet. Und im Edge-Katalog für die Echtzeit-Prompt-Inspektion sind längst nicht alle KI-Anbieter gleich gut vertreten.

Dimension	Chance	Risiko, wenn du nichts tust
Rechenschaftspflicht	Lückenloser Nachweis, wer welche Daten in welche KI gegeben hat	Auskunftersuchen nach Artikel 15 DSGVO nicht beantwortbar
Löschkonzept	Aufbewahrung und Löschung über den Speicherort Enterprise AI apps	KI-Verläufe altern unbegrenzt in der Cloud des Anbieters
Rechtsstreit	eDiscovery-Suche über ItemClass, Export in Prüfmengen	Beweismittel liegen außerhalb deines Zugriffs
Innentäter	Vorlage Risky AI usage, Signale laufen in Defender XDR zusammen	Datenabfluss über Prompts bleibt komplett unsichtbar
Prävention	Kante über Endpoint-, Browser- und Netzwerk-DLP absichern	Nur Protokoll, kein Schutz — das teuerste Missverständnis
Kosten	Verbrauch je Zähler über den Geltungsbereich steuerbar	Ungedeckelte Zähler auf der gesamten Belegschaft

Mitbestimmung	Saubere Betriebsvereinbarung als Rollout-Beschleuniger	Nachträgliche Abschaltung durch die Einigungsstelle
---------------	--	---

Was müssen wir jetzt schon vorbereiten?

Trenne von Anfang an zwei Arbeitspakete: Sichtbarkeit und Durchsetzung. Sichtbarkeit ist der Connector, das geht in Tagen. Durchsetzung ist die Kante, das geht in Quartalen. Wer beides in ein Projekt wirft, liefert nach sechs Monaten nichts Vorzeigbares und darf das dann im Lenkungsausschuss erklären.

Wo greift welche Kontrolle — und wo klappt das Loch

Fünf Schichten, ehrlich bewertet nach Wirkzeitpunkt und Reichweite

SCHICHT	WAS SIE WIRKLICH TUT	WANN	REICHWEITE
Vertraulichkeitsbezeichnung	Verschlüsselung reist mit der Datei. Ohne EXTRACT-Recht kein Kopieren.	vorher	90%
Endpoint-DLP	Einfügen in unterstützte Browser, Upload auf gesperrte Domänen.	Echtzeit	65%
Browser-DLP in Edge	Prompt-Inspektion beim Tippen, katalogabhängig je KI-Anbieter.	Echtzeit	55%
Netzwerk-DLP über SASE	Erkennung vertraulicher Informationstypen auf der Leitung.	Echtzeit	45%
Purview-Connector	Protokoll, Klassifizierung, Aufbewahrung, eDiscovery. Kein Eingriff.	danach	100%

Die Reichweite ist keine Produktnote, sondern deine Realität: Sie fällt mit jedem nicht verwalteten Gerät, jedem privaten Browser und jedem Beratervertrag ohne Endpoint-Onboarding.

Die Prozentwerte sind bewusst keine Herstellerangaben, sondern eine ehrliche Schätzung der Reichweite in einer typischen Umgebung.

Fang bei den Hausaufgaben an, die keinen Lizenzvertrag brauchen. Verknüpfe den Mandanten mit einem Azure-Abonnement und kläre die Pay-as-you-go-Zustimmung — das dauert in großen Konzernen erfahrungsgemäß länger als die gesamte technische Einrichtung, weil plötzlich Einkauf, Konzernrevision und ein sehr interessierter Bereichsleiter mitreden wollen. Parallel dazu: Verzeichnis von Verarbeitungstätigkeiten ergänzen, Datenschutz-Folgenabschätzung anstoßen, Betriebsrat einbinden. Und bevor du eine Erkennungsregel auf einen vermuteten Feldnamen baust, schau dir mit einer Abfrage über OfficeActivity an, welche Datensatztypen und Operationen in deinem Mandanten tatsächlich auflaufen. Vermutete Schemata sind die zweithäufigste Ursache für stille Fehlalarme, gleich nach zu großzügigen Geltungsbereichen.

TIPP • Der Reihenfolge-trick, der dir Wochen spart

Aktiviere zuerst nur den Activity Feed, ohne Chat Conversations. Du bekommst binnen 24 Stunden belastbare Nutzungszahlen, kannst die Kosten hochrechnen und hast eine Datengrundlage für das Gespräch mit Betriebsrat und Datenschutzbeauftragtem — ohne einen einzigen Prompt-Inhalt zu kopieren. Den Inhaltsbereich schaltest du erst dazu, wenn die Betriebsvereinbarung unterschrieben ist. Das kostet zwei Wochen und spart dir eine Eskalation.

Zeitraum	Was du tust	Woran du merkst, dass es funktioniert
Woche 1 bis 2	Azure-Abo verknüpfen, Pay-as-you-go-Zustimmung erteilen, Rollen vergeben, Compliance-API beim Anbieter freischalten	Der Connector validiert den Schlüssel, der Sync-Status steht auf grün
Woche 3 bis 4	Nur den Activity Feed einlesen, DSPM-Report auswerten, Kostenmodell im Azure-Preisrechner rechnen	Belastbare Zahlen zu Nutzern, Interaktionen und Monatskosten
Woche 5 bis 8	Betriebsvereinbarung, Datenschutz-Folgenabschätzung, Verzeichnis der Verarbeitungstätigkeiten, Löschfristen festlegen	Unterschrift liegt vor, Zweckbindung und Fristen sind dokumentiert
Woche 9 bis 12	Chat Conversations aktivieren, Sammlungsrichtlinie setzen, Aufbewahrung auf Enterprise AI apps, IRM-Vorlage Risky AI usage	Prompts erscheinen im Activity Explorer, Vorfälle sind zuordenbar
ab Monat 4	Kante aufbauen: Endpoint-DLP, sensible Dienstdomänengruppen, Edge for Business, Netzwerk-DLP über SASE	Erster echter Block eines Uploads, dokumentiert und begründet
laufend	Funktionsmatrix und Roadmap vor jedem Steering prüfen, Zählerverbrauch monatlich reviewen	Keine Überraschung auf der Azure-Rechnung, keine im Audit

Ein letzter Blick nach vorn, weil danach garantiert gefragt wird: Microsoft hat mit der processContent-Schnittstelle bereits das Muster für präventives Inline-DLP gebaut, bisher aber nur in Microsoft Foundry ausgerollt. Ob OpenAI und Anthropic diesen Aufruf in ihre Produkte einbauen, ist die eigentliche strategische Frage der nächsten zwölf Monate. Passiert es, wird aus dem Zeugen ein Wächter und du kannst die halbe Kanten-Architektur wieder abräumen. Passiert es nicht, bleibt Durchsetzung dauerhaft dein Problem am Endgerät. Plane für den zweiten Fall und freu dich über den ersten.

Häufig gestellte Fragen

Blockiert der Purview-Connector, wenn jemand Kundendaten in Claude oder ChatGPT einfügt?

Nein. Der Connector liest über die Compliance-API des Anbieters mit rund einem Tag Verzögerung mit und protokolliert, klassifiziert und archiviert die Interaktionen. Blockieren kannst du nur über Endpoint-DLP, Browser-DLP in Edge for Business oder Netzwerk-DLP auf der SASE-Ebene.

Welche Lizenzen und welche Kosten fallen für den Connector an?

Du brauchst einen Enterprise-Plan beim KI-Anbieter, einen mit einem Azure-Abonnement verknüpften Microsoft-365-Mandanten und die erteilte Pay-as-you-go-Zustimmung in Purview.

Abgerechnet wird verbrauchsabhängig über mehrere Zähler, unter anderem verarbeitete Audit-Datensätze, Textdatensätze zu je 1.000 Zeichen und aufbewahrte Interaktionen.

Sieht Purview auch, was über die Claude-Entwickler-API läuft?

Nur eingeschränkt. Für Claude Platform liefert die Compliance-API ausschließlich Activity-Feed-Ereignisse, also Anmeldungen, Administrationsaktionen und Konfigurationsänderungen. Prompts und Antworten sind nur für Claude Enterprise verfügbar.

Brauche ich für die Erfassung von Prompts eine Betriebsvereinbarung?

In Deutschland praktisch immer. Sobald du Chat-Inhalte einliest, entsteht eine technische Einrichtung, die zur Verhaltens- und Leistungskontrolle geeignet ist, und damit greift die Mitbestimmung nach Paragraf 87 Absatz 1 Nummer 6 Betriebsverfassungsgesetz. Zusätzlich brauchst du einen Eintrag im Verzeichnis von Verarbeitungstätigkeiten und in aller Regel eine Datenschutz-Folgenabschätzung.

Wie finde ich KI-Interaktionen später in einem eDiscovery-Fall wieder?

Prompts und Antworten liegen im Postfach des jeweiligen Benutzers und lassen sich über die Eigenschaft ItemClass suchen, für ChatGPT Enterprise etwa über `IPM.SkypeTeams.Message.ConnectedAIApp.Connector` bei Aktivitäten vom Gerät und `IPM.SkypeTeams.Message.CloudAIApp.SaaS` bei browserbasierten Aktivitäten. Voraussetzung ist, dass zum Zeitpunkt der Interaktion bereits eine Sammlungsrichtlinie für Enterprise-KI-Apps aktiv war.