

Copilot-Highlights September: Multi-Modell-Ansatz mit Grok- und OpenAI-Modellen

Executive Summary

Microsoft hat im September aus Copilot endgültig ein Buffet gemacht. Statt eines einzigen Modells, das brav alles erledigt, stehen jetzt mehrere Köche in der Küche: **GPT-6 Astra** von OpenAI als Reasoning-Modell für die harten Analysen, **Claude Fable 5.1** von Anthropic in Cowork und Copilot Studio, Microsofts eigene **MAI-Familie** und als neuester Gast **Grok** von SpaceXAI, dem Unternehmen, das früher xAI hieß. Dazu kommen Antwortkarten im Chat, ein Planner Agent, der Statusberichte schreibt, und ein neu sortiertes Copilot mit den Bereichen Home, Code und Autopilot.

Klingt nach Party. Ist es auch, bis die Rechnung kommt. Denn jedes zusätzliche Modell bedeutet zusätzliche Fragen: Wo landen meine Daten? Wer hat das freigegeben? Warum ist das Credit-Budget am 12. des Monats aufgebraucht? Die Kernbotschaft dieses Briefings: **Leg fest, welche Modelle in deinem Tenant freigegeben sind, bevor es deine Anwender für dich tun.** Für deutsche Tenants gibt es eine kleine Pointe obendrauf: Grok ist in der Preview für EU, EFTA und UK schlicht gesperrt. Du musst also erst mal gar nichts abschalten, aber du solltest wissen, warum.

★ Wichtig: Die drei Sätze für den Vorstand

Copilot wird zur Multi-Modell-Plattform, und die Modellwahl ist ab sofort eine Governance-Entscheidung, keine Geschmacksfrage. Neue Funktionen wie Usage-Billing in Copilot Studio machen Kosten variabel. Wer jetzt Freigabe-Regeln, Budgets und Pilotgruppen definiert, spart sich später den peinlichen Termin mit dem Datenschutzbeauftragten.

Worum geht es im Detail?

Fangen wir mit den Hintergründen an, denn ohne die versteht man das September-Update nicht. Microsoft war lange Zeit quasi mit OpenAI verheiratet. Copilot lief auf GPT-Modellen, Punkt. Seit 2025 hat sich die Ehe geöffnet: Erst kamen Anthropic-Modelle in Researcher und Copilot Studio, dann eigene MAI-Modelle, und jetzt Grok. Die Logik dahinter ist simpel und knallhart wirtschaftlich. Wer nur einen Lieferanten hat, wird erpressbar. Wer fünf hat, kann verhandeln und für jede Aufgabe das passende, sprich günstigste oder beste, Modell nehmen. Microsoft positioniert Copilot damit als Orchestrierungsschicht: Die eigentliche Wertschöpfung steckt im Zugriff auf deine Daten über Microsoft Graph und Work IQ, nicht im Sprachmodell darunter. Das Modell wird zur austauschbaren Zutat.

Die neuen Modelle im Überblick

Modell	Anbieter	Wo verfügbar	Besonderheit
GPT-6 Astra	OpenAI	Copilot, Cowork, Copilot Studio (Rollout ab 4. September)	Reasoning-Modell für tiefe Analysen und mehrstufige Aufgaben
Claude Fable 5.1	Anthropic	Cowork, Copilot Studio (Rollout ab 1. September)	Löst Fable 5 ab, stark bei langen Texten und Agenten
Grok	SpaceXAI (ehem. xAI)	Word, Excel, PowerPoint über das Frontier-Programm	Standardmäßig aus, Verarbeitung außerhalb von Microsoft, in EU/EFTA/UK gesperrt
Grok 4.1 Fast	SpaceXAI	Copilot Studio (US-Maker)	Schnell und günstig, laut Microsoft-Sicherheitsbewertung aber weniger „aligned“
MAI-Familie	Microsoft	Diverse Copilot-Funktionen	Hauseigene Modelle, voll im Microsoft-Vertragsrahmen

Der interessanteste Satz in der ganzen Grok-Ankündigung steht im Kleingedruckten: Die Grok-Modelle werden **außerhalb der von Microsoft verwalteten Umgebungen** verarbeitet und unterliegen den Enterprise-Nutzungsbedingungen und dem Auftragsverarbeitungszusatz von SpaceXAI. SpaceXAI steht inzwischen auf Microsofts Liste der Unterauftragsverarbeiter. Übersetzt heißt das: Sobald jemand in Excel Grok wählt, verlässt dein Quartalsforecast das Microsoft-Rechenzentrum und macht einen kleinen Ausflug zu Elon. Was soll da schon schiefgehen?

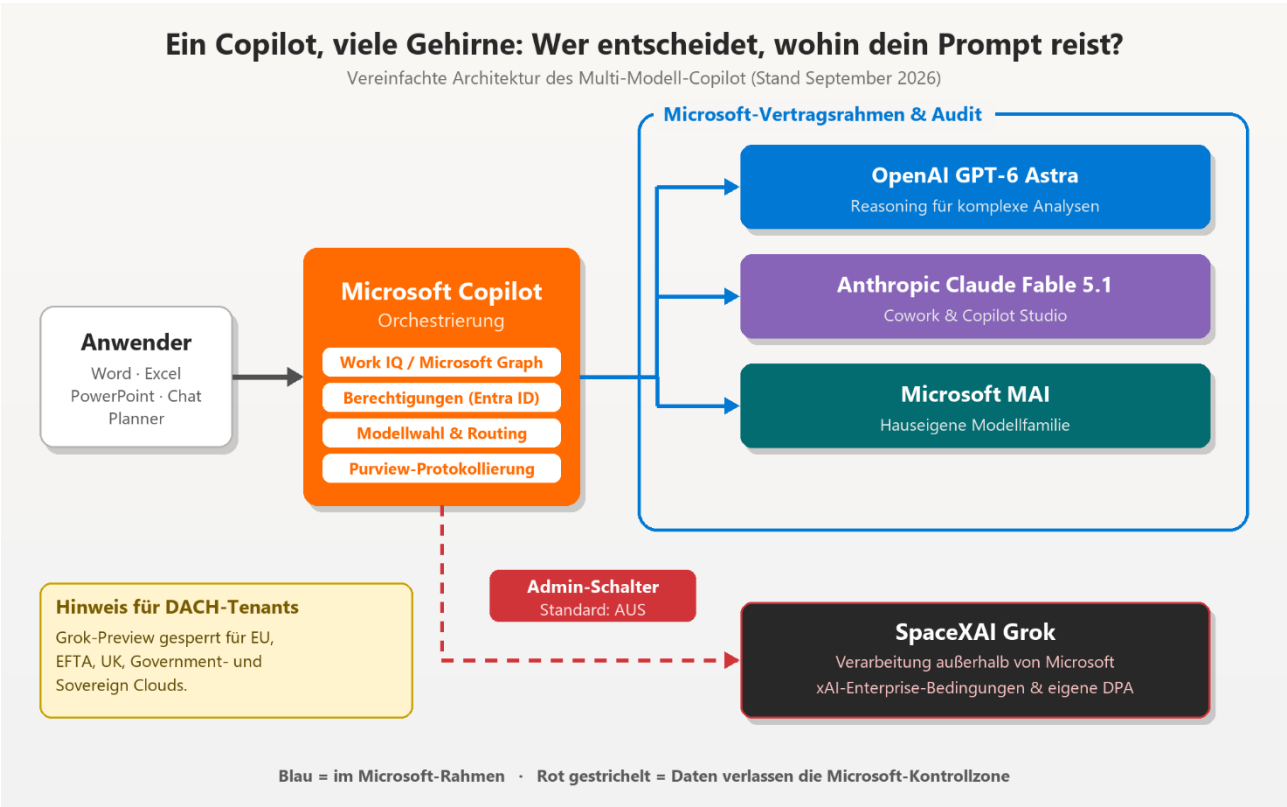


Abbildung 1: Der Prompt nimmt je nach Modell einen anderen Weg. Bei Grok verlässt er den Microsoft-Rahmen.

Microsoft hat dafür einen Admin-Schalter eingebaut, und der steht **standardmäßig auf Aus**. Es braucht also eine bewusste Entscheidung der IT, um Grok freizuschalten. Dass Microsoft selbst in einer Sicherheitsbewertung Grok 4.1 Fast als weniger gut ausgerichtet als die anderen Copilot-Modelle einstuft, mit höherem Risiko für schädliche Inhalte, und es trotzdem hinter einer Admin-Schranke ausliefert, ist ehrlich gesagt die eleganteste Form von „Wir haben euch gewarnt“, die ich seit Langem gesehen habe.

Was sonst noch im September-Paket steckt

- **Neue Struktur:** Copilot gliedert sich in **Home** (Chat plus Schnellzugriff auf Word, Excel, PowerPoint), **Code** (Apps und Dashboards per natürlicher Sprache) und **Autopilot** (ein dauerhaft laufender Agent mit eigener Identität und eigenen Berechtigungen für langlaufende Aufgaben).
- **Antwortkarten im Chat:** Fragen zu Wetter, Finanzen, Sport, Orten oder News liefern formatierte Karten statt Textwüsten. Dazu kommen Karten mit Zusammenfassungen, nächsten Schritten und Dateiausschnitten.
- **Planner Agent:** Er erstellt Statusberichte mit Fortschritt, Risiken, Gesamtstatus und nächsten Schritten, auf Wunsch im Ton für die jeweilige Zielgruppe. Außerdem steckt er jetzt direkt in den Basisplänen.
- **Office-Feinschliff:** PowerPoint liefert Review-Kommentare zu Struktur und Stil, Excel verknüpft Änderungsverfolgung mit Autoren und Versionsverlauf.
- **Umbenennung:** Die Microsoft 365 Copilot App heißt jetzt nur noch Microsoft Copilot, die URL wandert von m365.cloud.microsoft nach copilot.cloud.microsoft. Im Oktober kommt die Zwangsumleitung.
- **Usage-Billing in Copilot Studio:** Agenten auf dem neuen GitHub-Copilot-Harness rechnen nutzungsbasiert ab, und zwar unabhängig von der Copilot-Lizenz. Credits werden auch beim Bauen, Testen und Evaluieren verbraucht.

Noch ein Wort zu den Antwortkarten, weil sie gern unterschätzt werden. Sie wirken wie Kosmetik, sind aber ein Hinweis auf die Richtung: Copilot soll nicht mehr nur Text ausspucken, sondern strukturierte, klickbare Ergebnisse liefern, die direkt in den nächsten Arbeitsschritt führen. Für dich als IT heißt das, dass Copilot vom Spielzeug zur Arbeitsoberfläche wird. Und alles, was Arbeitsoberfläche ist, braucht Support, Schulung und ein Konzept für den Tag, an dem es ausfällt. Genau deshalb ist ein Big Bang hier die schlechteste Idee. Besser sind gestaffelte Piloten mit begleitendem Change-Management.

Auch der neue Autopilot verdient einen zweiten Blick. Ein Agent mit eigener Identität und eigenen Berechtigungen ist aus Sicht von Entra ID nichts anderes als ein neuer Mitarbeiter, der nie schläft, nie Urlaub nimmt und nie fragt, ob er etwas darf. Wer bei Dienstkonten schon heute den Überblick verloren hat, sollte den Autopilot erst einschalten, wenn klar ist, wer für ihn haftet, wer seine Rechte prüft und wer ihn im Zweifel feuert.

i Fakten: Credit-Governance in Zahlen

Pro Umgebung kannst du bis zu **vier Durchsetzungsregeln** definieren, Limits auf einzelne Agenten setzen und den Credit-Verbrauch tenantweit einsehen. Gemessen wird zunehmend in „Arbeitsstunden“ statt in Prompts. Das klingt nach Fortschritt und ist es auch, macht aber Budgetplanung zur Disziplin mit Kristallkugel.

Was sind Chancen? Was sind Risiken?

Die ehrliche Antwort lautet: beides, und zwar gleichzeitig. Der Multi-Modell-Ansatz ist technisch sinnvoll. Kein Modell ist überall das beste. Ein Reasoning-Modell wie GPT-6 Astra glänzt bei der Frage, warum der Deckungsbeitrag in Region Süd eingebrochen ist, ist aber für „Formuliere diese Mail freundlicher“ ungefähr so angemessen wie ein Bagger zum Blumengießen. Wer das richtige Modell für die richtige Aufgabe nimmt, bekommt bessere Ergebnisse und, zumindest theoretisch, niedrigere Kosten.

Das Problem ist das Wort „theoretisch“. Jedes Modell bringt eigene Datenflüsse, eigene Vertragsbedingungen, eigene Schwächen beim Halluzinieren und eigene Preise mit. Und in der Praxis wählt nicht der Architekt das Modell, sondern Kollege Schulze aus dem Vertrieb, weil ihm der Name gefällt.

Aspekt	Chance	Risiko
Qualität	Bestes Modell pro Aufgabe, spürbar bessere Analysen	Unterschiedliche Antworten auf dieselbe Frage, Vertrauensverlust
Kosten	Günstige Modelle für Routine, Premium nur bei Bedarf	Usage-Billing macht Budgets unberechenbar, auch Testläufe kosten
Datenschutz	Modelle im Microsoft-Rahmen bleiben voll auditierbar	Grok verarbeitet außerhalb von Microsoft, eigener Vertrag, eigene DPA
Betrieb	Weniger Abhängigkeit von einem Anbieter	Mehr Schalter, mehr Doku, mehr Schulung, mehr Supporttickets
Agenten	Autopilot und Planner Agent nehmen echte Arbeit ab	Agenten mit eigener Identität brauchen Berechtigungs- und Audit-Konzept

Ein Beispiel aus dem echten Leben, leicht anonymisiert: Ein Maschinenbauer aus dem Sauerland hatte in Copilot Studio einen Angebotsagenten gebaut. Drei Maker, zwei Wochen Feintuning, gefühlt viertausend Testläufe. Niemand hatte auf dem Schirm, dass das Testen selbst Credits frisst. Die Monatsrechnung landete beim CFO, der daraufhin eine Frage stellte, die in keinem Projektplan stand: „Wer hat eigentlich gesagt, dass wir das dürfen?“ Der Agent funktioniert übrigens hervorragend. Er hat nur das Budget für das gesamte Quartal in elf Tagen gegessen.

⚠️ Warnung: Schatten-KI 2.0

Wenn du keine Modelle freigibst, suchen sich deine Leute eben selbst welche. Die Grok-Add-ins für Word, PowerPoint, Excel und Outlook gab es schon seit Sommer im Store. Wer Add-ins nicht zentral steuert, hat die Datenschutzdiskussion längst, er weiß es nur noch nicht. Das ist wie mit Schimmel: Wenn du ihn siehst, ist es zu spät.

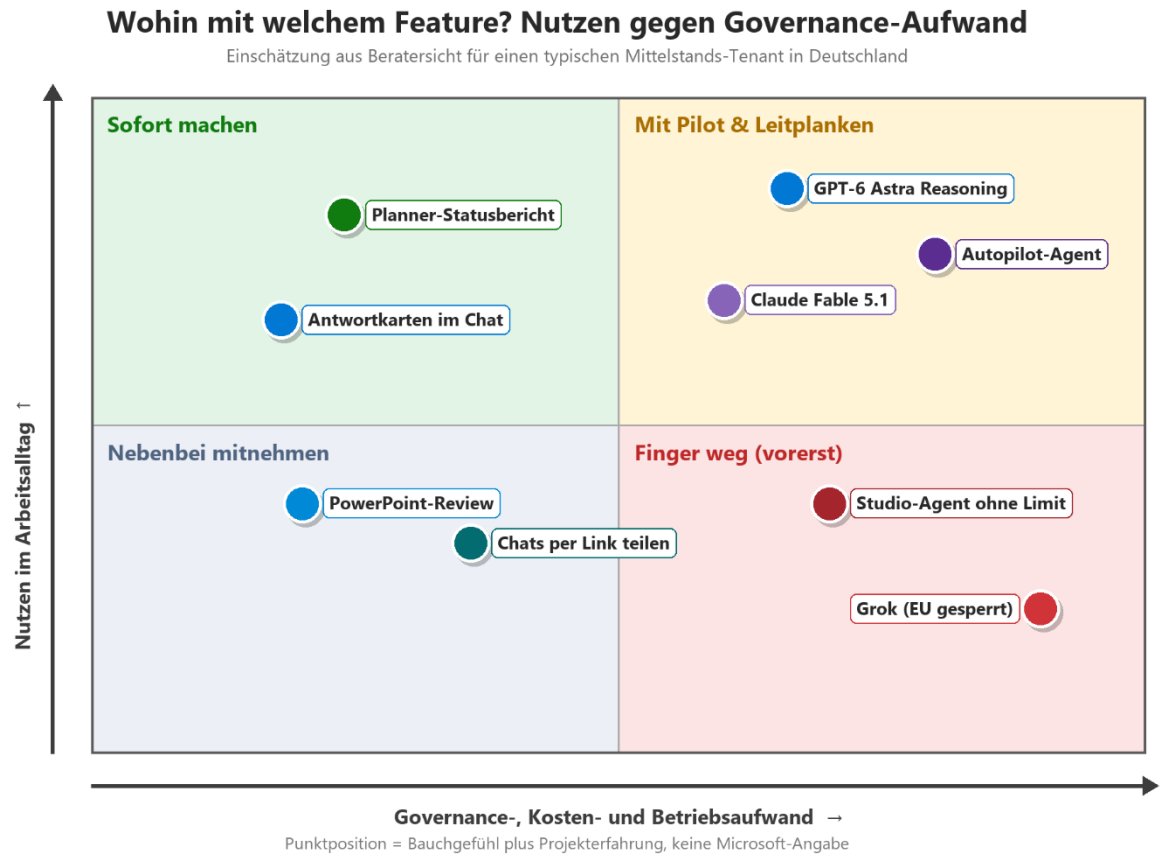


Abbildung 2: Nicht jedes neue Feature verdient denselben Governance-Aufwand. Die Matrix hilft bei der Priorisierung.

Die Matrix zeigt das Muster: Antwortkarten und der Planner-Statusbericht sind „Low Hanging Fruits“, PowerPoint-Review und das Teilen von Chats nimmst du einfach mit. Sie arbeiten im Microsoft-Rahmen, nutzen vorhandene Berechtigungen und machen Anwender sofort glücklich. GPT-6 Astra, Claude und der Autopilot-Agent bringen viel, brauchen aber Leitplanken. Ein Copilot-Studio-Agent ohne Credit-Limit gehört ebenfalls nach rechts unten, denn er ist die zuverlässigste Methode, ein Jahresbudget in Rekordzeit zu verbrennen. Grok liegt für deutsche Tenants rechts unten, nicht weil das Modell schlecht wäre, sondern weil es aktuell weder verfügbar noch vertraglich in deinen Datenschutzrahmen eingebettet ist.

Was müssen wir jetzt schon vorbereiten?

Gute Nachricht: Du musst nicht alles auf einmal machen. Schlechte Nachricht: Ein paar Dinge solltest du wirklich diesen Monat erledigen, weil sie dir sonst im Oktober auf die Füße fallen.

Aufgabe	Warum	Bis wann
Modellfreigabe-Richtlinie schreiben	Welche Modelle, für wen, für welche Datenklassen	Oktober 2026
Admin-Schalter für Grok dokumentieren	Bleibt aus, aber bewusst und begründet	Sofort
copilot.cloud.microsoft freigeben	Zwangsumleitung im Oktober, sonst geht die Web-App nicht	Vor der Umleitung
Credit-Limits in Copilot Studio setzen	Usage-Billing gilt auch für Test und Evaluierung	Vor dem nächsten Agentenprojekt
Add-in-Steuerung prüfen	Drittanbieter-KI-Add-ins umgehen die Copilot-Governance	Sofort
Pilotgruppe für GPT-6 Astra	Nutzen messen, bevor alle es nutzen	Q4 2026
Berechtigungskonzept für Autopilot	Agent mit eigener Identität ist ein neuer Kontotyp	Vor jedem Einsatz

Hintergrund zur Modellfreigabe-Richtlinie: Die Frage ist nicht „Welches Modell ist das beste?“, sondern „Welche Daten dürfen zu welchem Anbieter?“. Verknüpfe die Freigabe daher mit deinen Purview-Vertraulichkeitsbezeichnungen. Alles, was im Microsoft-Rahmen verarbeitet wird, ist vergleichsweise einfach, denn deine bestehende Auftragsverarbeitung mit Microsoft greift. Alles, was hinausgeht, braucht eine eigene Prüfung mit Datenschutz, Betriebsrat und, falls vorhanden, Informationssicherheit. Das ist kein Bürokratiefetisch, sondern schlicht die Folge aus DSGVO und Betriebsverfassungsgesetz.

✓ Tipp: Der Zehn-Minuten-Quick-Win

Leg im Admin Center eine kurze interne Seite „Welche KI darf ich nutzen?“ an und verlinke sie im Copilot-Onboarding. Drei Spalten: freigegeben, im Pilot, verboten. Das beantwortet 80 Prozent der Tickets, bevor sie entstehen, und du hast etwas, worauf du zeigen kannst, wenn es knallt.

Und dann ist da noch der Planner Agent. Teste ihn mit einem echten Projekt. Eine IT-Leiterin bei einem Logistiker erzählte mir neulich, dass ihr Team den automatischen Statusbericht zum ersten Mal ungefiltert an die Geschäftsführung geschickt hat. Der Agent hatte unter „Risiken“ nüchtern vermerkt, dass zwölf von vierzig Aufgaben seit drei Wochen überfällig sind und dem verantwortlichen Bereichsleiter zugeordnet. Der Bericht war korrekt. Die Stimmung im nächsten Jour fixe weniger. Lektion: Prüfe, was die KI schreibt, bevor es der Vorstand liest, denn sie ist ehrlicher als dein Projektcontrolling.

⚠ Warnung: URL-Umzug nicht verschlafen

Wer copilot.cloud.microsoft in Proxy oder Firewall blockiert, erlebt nach der Zwangsumleitung im Oktober, dass niemand mehr die Copilot-Web-App nutzen kann. Die Supporthotline freut

sich schon. Also: Allowlist jetzt anpassen, nicht am Montagmorgen nach dem Ausfall.

Häufig gestellte Fragen

Kann ich Grok in Microsoft 365 Copilot in Deutschland nutzen?

Derzeit nicht. Die Grok-Preview im Frontier-Programm ist für Tenants in der EU, der EFTA und Großbritannien sowie für Government- und Sovereign Clouds gesperrt und auch anderswo standardmäßig deaktiviert.

Wo werden meine Daten verarbeitet, wenn jemand Grok in Copilot verwendet?

Grok-Anfragen werden außerhalb der von Microsoft verwalteten Umgebungen verarbeitet und unterliegen den Enterprise-Bedingungen und dem Auftragsverarbeitungszusatz von SpaceXAI. Damit gelten nicht die üblichen Microsoft-Zusagen zu Speicherort und Audit, was eine eigene Datenschutzprüfung erfordert.

Wie schalte ich einzelne KI-Modelle in Copilot frei oder ab?

Die Freigabe von Drittanbietermodellen erfolgt über Admin-Einstellungen im Microsoft 365 Admin Center beziehungsweise für Agenten in Copilot Studio und im Power Platform Admin Center. Grok ist standardmäßig ausgeschaltet und muss aktiv aktiviert werden.

Kostet die Nutzung von GPT-6 Astra oder anderen Modellen in Copilot extra?

In Microsoft 365 Copilot sind die Modelle grundsätzlich durch die Lizenz abgedeckt, in Copilot Studio fallen bei Agenten auf dem neuen GitHub-Copilot-Harness jedoch nutzungsbasierte Credits an. Diese werden auch beim Bauen und Testen verbraucht, deshalb solltest du Limits pro Umgebung und Agent setzen.

Was kann der Planner Agent in Microsoft 365 Copilot?

Der Planner Agent erstellt Statusberichte mit Fortschritt, Risiken und nächsten Schritten, legt Aufgaben an, beantwortet Fragen zum Plan und funktioniert jetzt auch in gruppenbasierten Basisplänen. Ton und Detailtiefe des Berichts lassen sich an die Zielgruppe anpassen.