TEAMS & TELEFONIE

Teams: Drittanbieter-Erkennung synthetischer Audio- und Videoinhalte in Besprechungen

Consulting Briefing · 18.11.2026 · boddenberg.de

Executive Summary

Dein Finanzvorstand ruft per Teams an, Kamera an, Stimme vertraut, Anliegen dringend: „Die Übernahme ist vertraulich, bitte heute noch 1,8 Millionen auf dieses Konto." Früher hätte man gesagt, so dumm ist niemand. Heute muss man sagen: So gut ist inzwischen die Software. Genau gegen dieses Szenario öffnet Microsoft Teams jetzt eine Schnittstelle, über die zertifizierte Drittanbieter Besprechungen auf Deepfakes untersuchen und verdächtige Teilnehmer direkt in der Teilnehmerliste markieren.

Die drei Dinge, die du wissen musst: Erstens macht Microsoft die Erkennung nicht selbst, sondern liefert nur die Steckdose. Der Strom kommt von einem Fremdanbieter, mit eigenem Vertrag und eigener Cloud. Zweitens passiert nichts automatisch: Ein Administrator muss Anbieter auswählen, App bereitstellen, Berechtigungen erteilen und Besprechungsrichtlinien setzen. Drittens analysiert hier eine KI Stimme und Gesicht jedes Teilnehmers. Damit bist du mitten im Datenschutz und, sobald Beschäftigte betroffen sind, mitten in der Mitbestimmung nach § 87 Abs. 1 Nr. 6 BetrVG.

Laut Microsoft-Ankündigung MC1509787, die Microsoft am 17.11.2026 zum dritten Mal geändert hat, startet der Targeted Release Ende November 2026 statt wie zuvor Mitte November. Einen Termin für die allgemeine Verfügbarkeit nennt Microsoft nicht. Unsere Empfehlung: Jetzt nicht einschalten, aber jetzt anfangen. Die Technik ist in einer Stunde konfiguriert, die Betriebsvereinbarung nicht.

Fakten: Das Wichtigste auf einen Blick

Quelle: Microsoft-Ankündigung MC1509787 (Fassung 3 vom 17.11.2026), Roadmap-ID 573451.

Plattformen laut Roadmap: Teams Desktop (Windows), Mac und Web. Android und iOS fehlen bisher.

Rollout: Targeted Release Ende November 2026, allgemeine Verfügbarkeit noch ohne Datum.

Aktivierung: ausschließlich per Admin-Entscheidung, kein Automatismus.

Worum geht es im Detail?

Der Hintergrund ist so unangenehm wie lehrreich. Anfang 2024 überwies ein Mitarbeiter des Ingenieurbüros Arup in Hongkong rund 25 Millionen US-Dollar an Betrüger, nachdem er in einer Videokonferenz scheinbar mit dem Finanzchef und mehreren Kollegen gesprochen hatte. Alle Gesprächspartner waren synthetisch. Der Mann war nicht naiv, er war in der Unterzahl: allein gegen eine komplette, KI-generierte Führungsetage. Seitdem ist „CEO-Fraud per Video" kein

Konferenzvortrag mehr, sondern ein Geschäftsmodell. Das FBI warnte im Juli 2026 ausdrücklich davor, dass KI-generierte Führungskräfte in Echtzeit-Videoanrufen kaum noch zu erkennen sind.

Und der Mensch als Detektor? Eher ein Münzwurf mit Selbstbewusstsein. Eine Studie des University College London fand 2023, dass Probanden gefälschte Sprache nur in 73 Prozent der Fälle erkannten. Eine Metaanalyse über 56 Studien kam 2024 auf eine menschliche Trefferquote von gut 55 Prozent. Dein Bauchgefühl ist also statistisch nur knapp besser als Raten, aber deutlich teurer, wenn es danebenliegt.

Microsoft reagiert darauf nicht mit einem eigenen Detektor, sondern mit einer Plattform. Technisch funktioniert das so: Der Erkennungsdienst eines zertifizierten Anbieters tritt der Besprechung als Bot bei, greift Audio- und Videoströme ab und analysiert sie in seiner eigenen Infrastruktur. Findet er Anzeichen für Manipulation, meldet er das über eine neue Microsoft-Graph-Schnittstelle namens reportSyntheticMedia (derzeit in der Beta) an Teams zurück. Die Meldung enthält einen Konfidenzwert zwischen 0,0 und 1,0, einen Schweregrad (niedrig, mittel, hoch), eine Einschätzung, ob Böswilligkeit vorliegt, und den Typ der Fälschung, etwa voiceClone für geklonte Stimmen sowie faceSwap und lipSync für Videomanipulation. Ein Bot darf pro Teilnehmer höchstens einmal pro Minute melden.

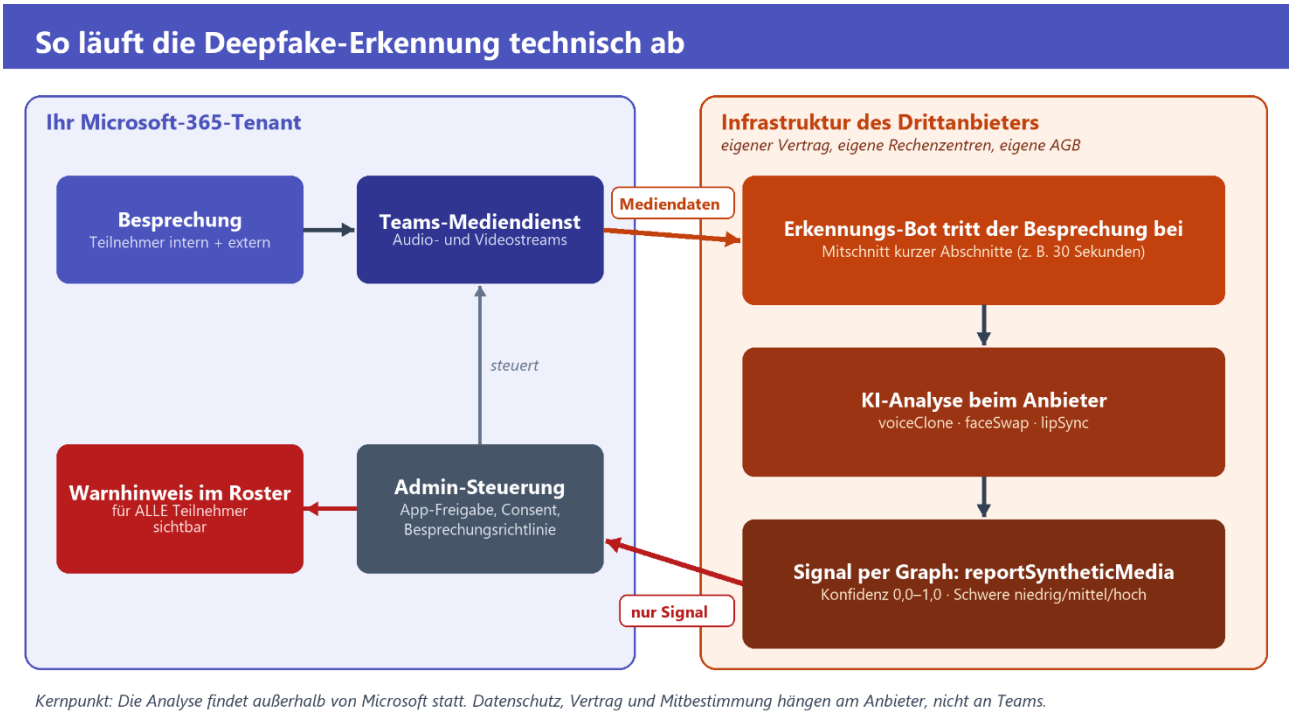


Abbildung 1: Teams liefert die Bühne, der Drittanbieter spielt den Türsteher. Die Medienanalyse läuft außerhalb des Microsoft-Tenants.

Der entscheidende Satz in der Meldung ist fast versteckt: Die Erkennung erfolgt durch den Drittanbieter, Teams stellt nur die Integration bereit. Übersetzt heißt das: Alles, was Microsoft dir zu Datenschutz, Datenresidenz und „Ihre Daten trainieren keine Modelle“ für die eigene KI zusagt, gilt für den externen Bot nicht automatisch. Du holst dir einen neuen Auftragsverarbeiter mit Live-Zugriff auf jede Stimme und jedes Gesicht in deinen Besprechungen ins Haus.

Für die Einrichtung nennt Microsoft vier Schritte: Anbieter evaluieren und auswählen, dessen App im Teams Admin Center unter „Teams-Apps“ › „Apps verwalten“ bereitstellen, die nötigen Berechtigungen per Admin-Consent erteilen und die passenden Besprechungsrichtlinien konfigurieren. In Teams PowerShell tauchen dafür die Richtlinienparameter SyntheticMediaDetection und SyntheticMediaDetectionAppId auf; bis zur Freigabe haben sie keine

Wirkung. Der Bot benötigt neben den üblichen Calls-Berechtigungen zum Beitritt eine eigene Anwendungsberechtigung zum Melden der Signale.

> **Wichtig: Wer sieht die Warnung?**
>
> Die Signale werden über die Teilnehmerliste an alle Teilnehmer verteilt, nicht nur an den Organisator. Ein Fehlalarm ist also kein stilles Ticket im SOC, sondern ein öffentlicher Vermerk neben dem Namen einer realen Person, gern auch vor Kunden. Microsoft hat bislang weder Trefferquoten noch Fehlalarmraten veröffentlicht.

Welche Anbieter zertifiziert sind, verrät Microsoft bisher nicht. Am Markt bieten unter anderem Sensity, Reality Defender, Pindrop und Resemble AI bereits Teams-Integrationen an; ob diese das Zertifikat für die neue Schnittstelle tragen, ist offen. Sensity beschreibt zum Beispiel, dass sein Bot für jede Analyse rund 30 Sekunden Audio und Video aufnimmt. Das ist für die Datenschutzbewertung ein entscheidendes Detail, denn „wir schauen kurz rein“ und „wir zeichnen auf“ sind juristisch zwei verschiedene Welten.

Was sind Chancen? Was sind Risiken?

Fangen wir mit der guten Nachricht an: Zum ersten Mal bekommst du ein technisches Signal genau an der Stelle, an der der Betrug passiert, nämlich mitten in der Besprechung. Bisher lag die gesamte Abwehrlast auf Schulungen und Rückrufprozessen. Die funktionieren, solange niemand unter Zeitdruck steht. Betrüger sorgen zuverlässig dafür, dass jemand unter Zeitdruck steht.

Aspekt	Chance	Risiko
Sicherheit	Echtzeitwarnung bei Stimmklonen und Gesichtstausch, genau im Angriffsmoment	Falsches Sicherheitsgefühl: kein Warnhinweis heißt nicht „echt“
Betrieb	Gezielte Aktivierung per Richtlinie, etwa nur für Finanzen und Vorstand	Zusätzlicher Bot in jeder Besprechung, neue Abhängigkeit vom Anbieter
Datenschutz	Belegbare technische Maßnahme nach Art. 32 DSGVO	Neuer Auftragsverarbeiter mit Zugriff auf Stimme und Gesicht, oft Drittlandbezug
Mitbestimmung	Gemeinsam vereinbarte Regeln schaffen Akzeptanz	Technische Einrichtung im Sinne des § 87 Abs. 1 Nr. 6 BetrVG: ohne Zustimmung kein Einsatz
Reputation	Signal an Kunden: Hier wird gegen Betrug geprüft	Fehlalarm brandmarkt echte Teilnehmer öffentlich
Kosten	Günstiger als ein einziger erfolgreicher CEO-Fraud	Lizenzmodell offen, Anbieterkosten kommen obendrauf

Jetzt der Teil mit dem schwarzen Humor. Stell dir vor, dein Vertriebsleiter sitzt im ICE, Bluetooth-Headset mit Funkloch-Kompression, Bild eingefroren im ungünstigsten Gesichtsausdruck. Der Bot meldet „voiceClone, hoch“, und beim Kunden steht ab sofort ein rotes Warnsymbol neben dem

Mann, der gerade den Jahresvertrag verlängern wollte. Glückwunsch: Die Technik hat einen Betrug verhindert, nur leider den eigenen Umsatz. Fehlalarme sind kein Randproblem, sie entscheiden über die Akzeptanz.

Warnung: Der Deepfake-Detektor ist kein Freifahrtschein

Angreifer testen ihre Fälschungen genau gegen die Detektoren, die am Markt sind. Ein grünes Roster ist ein Indiz, kein Beweis. Wer deshalb den Rückruf über eine bekannte Nummer abschafft, hat das Prinzip nicht verstanden und verdient den Betrüger, den er bekommt.

Dazu kommt das Mitbestimmungsthema, und das ist in Deutschland kein Detail, sondern ein Stoppschild. Nach ständiger Rechtsprechung des Bundesarbeitsgerichts genügt es für § 87 Abs. 1 Nr. 6 BetrVG, dass eine technische Einrichtung objektiv geeignet ist, Verhalten oder Leistung zu überwachen. Ob der Arbeitgeber das will, ist egal. Ein Bot, der jede Stimme und jedes Gesicht analysiert und Ergebnisse pro Person ausgibt, erfüllt das mühelos. Bei öffentlichen Auftraggebern greift § 80 Abs. 1 Nr. 21 BPersVG beziehungsweise die entsprechende Vorschrift des jeweiligen Landespersonalvertretungsgesetzes. Wer ohne Zustimmung aktiviert, riskiert eine einstweilige Verfügung und darf das Feature dann mit dem Betriebsrat in der Einigungsstelle diskutieren. Das ist zeitlich ungefähr so effizient wie ein Faxgerät mit Wartungsvertrag.

Datenschutzrechtlich solltest du von einer Datenschutz-Folgenabschätzung nach Art. 35 DSGVO ausgehen: systematische, KI-gestützte Analyse von Sprach- und Videodaten, viele Betroffene, darunter Externe, die keinerlei Beziehung zu deinem Anbieter haben. Ob die Analyse biometrische Daten im Sinne von Art. 9 DSGVO verarbeitet, hängt davon ab, ob der Anbieter Personen identifiziert oder nur Artefakte erkennt. Genau das musst du schriftlich klären, nicht im Vertriebsgespräch.

Was müssen wir jetzt schon vorbereiten?

Die gute Nachricht: Du hast Zeit. Die schlechte: weniger, als du denkst. Microsoft hat den Targeted Release in Fassung 3 der Meldung nur um ein paar Tage auf Ende November verschoben, der GA-Termin folgt per Message Center. Erfahrungsgemäß liegen zwischen Targeted Release und allgemeiner Verfügbarkeit wenige Wochen bis Monate. Eine Betriebsvereinbarung mit Datenschutz-Folgenabschätzung braucht im Mittelstand realistisch drei bis sechs Monate, in Behörden gern auch länger.

Wann	Wer	Was konkret
Bis 30.11.2026	Teams-Admin	Prüfen, dass keine Besprechungsrichtlinie die Funktion vorschnell aktiviert. App-Freigaberichtlinien so setzen, dass Erkennungs-Apps nicht von Anwendern selbst hinzugefügt werden. Admin-Consent für Anwendungsberechtigungen nur über einen dokumentierten Prozess.
Bis 15.12.2026	IT-Sicherheit + Fachbereich	Risikoanalyse: Welche Rollen sind Ziel von CEO-Fraud (Finanzen, Einkauf, Geschäftsführung, Personal)? Daraus den Pilotkreis ableiten.

Wann	Wer	Was konkret
Dezember 2026	Datenschutz	Anbieter-Shortlist mit Fragenkatalog: Speicherort, Speicherdauer, Mitschnittlänge, Training mit Kundendaten, Unterauftragnehmer, Drittlandtransfer. DSFA beginnen.
Dezember 2026	HR + Betriebs- oder Personalrat	Frühzeitig unterrichten (§ 90 BetrVG), Eckpunkte einer Betriebsvereinbarung entwerfen: Zweck nur Betrugsabwehr, keine Leistungsauswertung, Löschfristen, Umgang mit Fehlalarmen.
Januar 2027	Teams-Admin	Testtenant oder Pilotgruppe im Targeted Release, eigene Richtlinie nur für den Pilotkreis, Fehlalarme dokumentieren.
Bis 15.02.2027	Kommunikation	Belegschaft informieren, Einladungsvorlagen für Externe um Datenschutzhinweis ergänzen, Verhaltensregel bei Warnhinweis festlegen.
Nach GA	Entscheider	Go oder No-Go auf Basis der Pilotdaten, schrittweiser Rollout per Richtlinie, nie pauschal für alle.

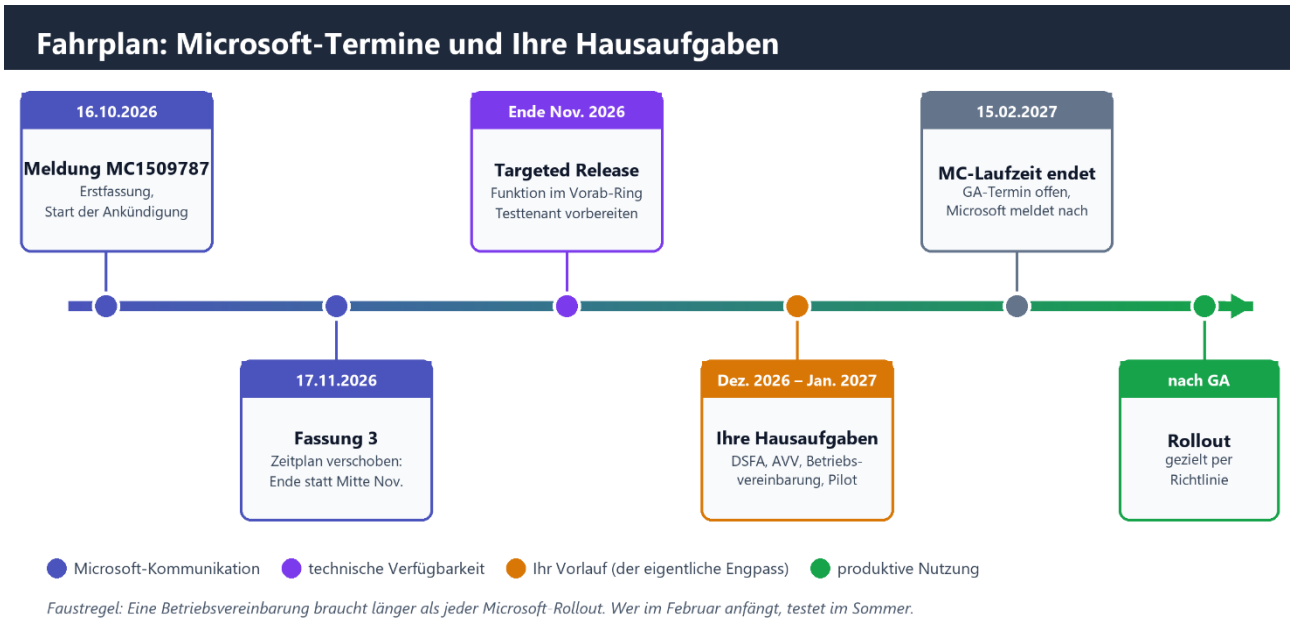


Abbildung 2: Microsofts Termine sind der kleinere Teil des Projekts. Der Engpass liegt in Datenschutz und Mitbestimmung.

Tipp: Der Rückruf bleibt König

Formuliere jetzt schon eine Regel, die unabhängig von jeder Technik gilt: Zahlungsfreigaben, Kontoänderungen und Passwort-Resets werden nie allein in einer Besprechung entschieden, sondern über einen zweiten, bekannten Kanal bestätigt. Der Detektor ergänzt diese Regel, er ersetzt sie nicht. Ein Deepfake, der den Rückruf auf die hinterlegte Festnetznummer übersteht, muss erst noch erfunden werden.

Ein Wort zu den Lizenzen: Microsoft hat bislang nicht gesagt, ob die Integration an eine bestimmte Teams- oder Microsoft-365-Lizenz gebunden ist. Sicher ist nur, dass der Drittanbieter Geld kostet, meist pro Nutzer oder pro analysierter Besprechungsminute. Plane das Budget deshalb nicht für die ganze Belegschaft, sondern für den Risikokreis. Die Buchhaltung im Gesamtvorstandsmeeting zu schützen ist sinnvoll, das wöchentliche Kantinenplanungs-Meeting eher nicht. Deepfakes von Kantinenchefs sind, Stand heute, kein relevantes Geschäftsmodell.

Und schließlich die Kommunikation an die Belegschaft. Erkläre offen, was der Bot tut und was nicht: Er sucht nach Fälschungen, er bewertet niemanden. Lege fest, was passiert, wenn der Warnhinweis erscheint. Unser Vorschlag: Besprechung freundlich unterbrechen, Identität über einen zweiten Kanal prüfen, keine Entscheidung treffen, Vorfall an die IT-Sicherheit melden. Wer das vorher geübt hat, reagiert ruhig. Wer es zum ersten Mal live vor dem Kunden erlebt, reagiert wie ein Reh im Scheinwerferlicht.

Den aktuellen Stand des Features verfolgst du im öffentlichen Microsoft-365-Roadmap-Eintrag 573451. Neue Fassungen von MC1509787 erscheinen wie gewohnt im Message Center.

Häufig gestellte Fragen

Erkennt Microsoft Teams Deepfakes in Besprechungen selbst?

Nein, die Analyse übernimmt ein zertifizierter Drittanbieter, dessen Bot der Besprechung beitritt. Teams stellt laut Microsoft-Ankündigung MC1509787 nur die Schnittstelle bereit und zeigt die gemeldeten Warnsignale in der Teilnehmerliste an.

Muss der Betriebsrat der Deepfake-Erkennung in Teams zustimmen?

In aller Regel ja, denn ein System, das Stimme und Bild jedes Teilnehmers analysiert und personenbezogene Ergebnisse ausgibt, ist objektiv zur Überwachung geeignet und fällt damit unter § 87 Abs. 1 Nr. 6 BetrVG. In Behörden gilt Entsprechendes nach dem Personalvertretungsrecht des Bundes oder des Landes.

Brauche ich für die Deepfake-Erkennung in Teams eine Datenschutz-Folgenabschätzung?

Sehr wahrscheinlich, weil ein externer Anbieter Audio- und Videodaten vieler Personen systematisch mit KI auswertet, darunter auch Externe. Zusätzlich ist ein Auftragsverarbeitungsvertrag mit dem Anbieter nötig, und ein möglicher Drittlandtransfer muss geprüft werden.

Ab wann ist die Deepfake-Erkennung in Teams verfügbar?

Der Targeted Release beginnt laut Microsoft Ende November 2026 für Teams auf Windows, Mac und im Web. Einen Termin für die allgemeine Verfügbarkeit hat Microsoft noch nicht genannt und kündigt ihn über das Message Center an.

Wird die Deepfake-Erkennung in Teams automatisch eingeschaltet?

Nein, ohne aktives Handeln eines Administrators passiert nichts. Er muss einen Anbieter auswählen, dessen App bereitstellen, Berechtigungen erteilen und die Besprechungsrichtlinien entsprechend konfigurieren.