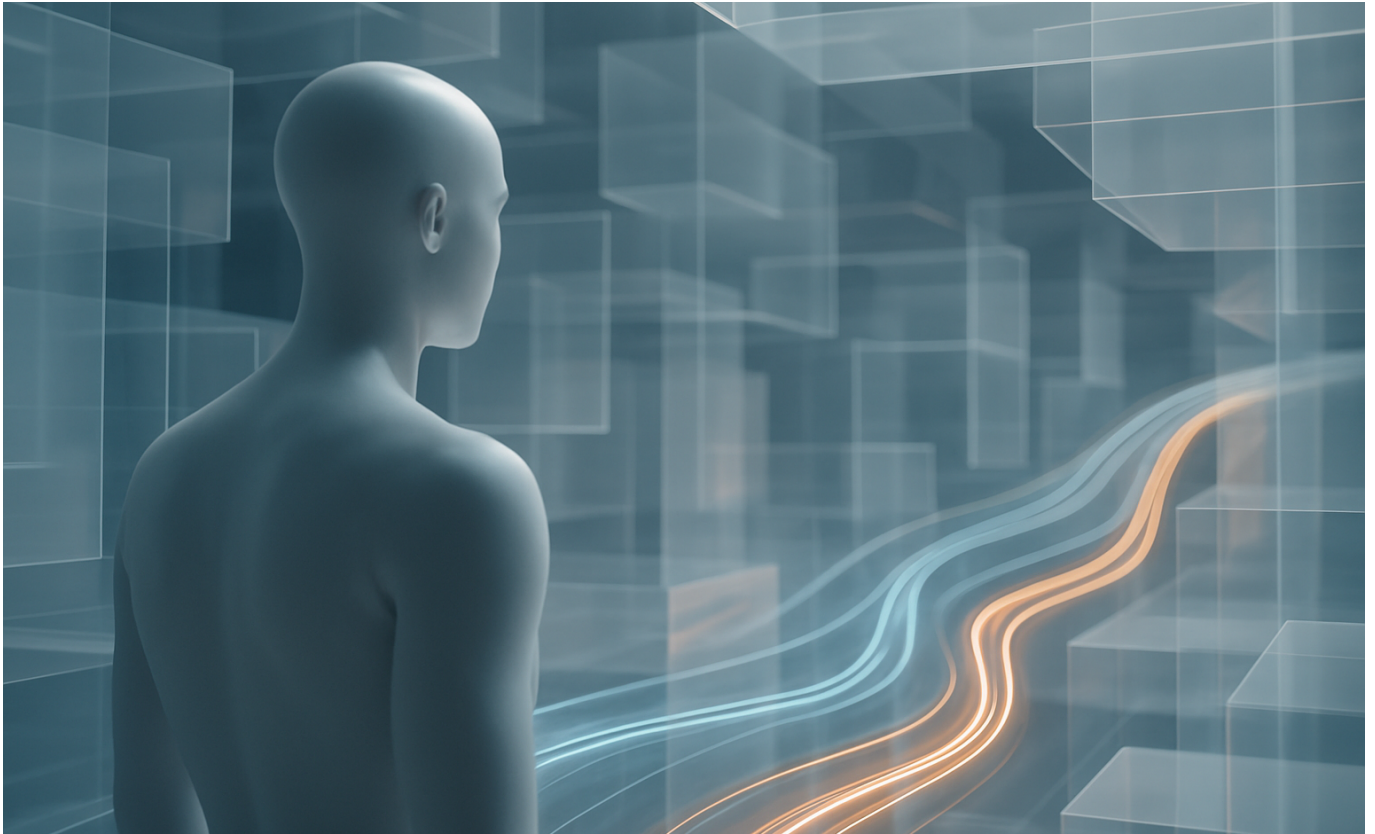


# KI-4-Everyone · Daily News

30. Juli 2026



RES

## Microsoft trainiert KI-Agenten in realistischen Umgebungen

KI-Agenten scheitern oft bei mehrstufigen Aufgaben wie E-Mail oder Kundensupport. Echoverse lässt sie in sich verändernden Umgebungen üben, statt nur mehr Aufgaben zu liefern.

PROD

## GitHub Copilot kann jetzt mehrere Arbeitsschritte stapeln

Mit gestapelten Sessions und Pull Requests modernisiert GitHub Copilot alten Programmcode in der App. Das zeigt ein konkretes Praxisbeispiel aus dem GitHub-Blog.

# Microsoft Research zeigt Echaverse: KI-Agenten lernen in lebenden Umgebungen statt in starren Aufgabenlisten

*Statt immer neue Trainingsaufgaben zu stapeln, sollen KI-Agenten in einer sich mitentwickelnden Testwelt reifen. Ein Ansatz, der die Schwächen heutiger Büro-Assistenten adressiert.*

**W**er schon einmal versucht hat, einen KI-Assistenten selbstständig E-Mails beantworten oder einen Kundenvorgang von Anfang bis Ende bearbeiten zu lassen, kennt das Muster: Der erste Schritt klappt, der zweite auch noch, spätestens beim dritten verheddert sich das System. Genau an diesem Punkt setzt Microsoft Research mit einem Projekt namens Echaverse an. Die Grundidee ist ein Perspektivwechsel: Nicht der Agent bekommt mehr Aufgaben vorgesetzt, sondern die Umgebung, in der er arbeitet, entwickelt sich mit ihm weiter.

In einem Blogbeitrag beschreibt Microsoft Research Echaverse als eine Trainingsumgebung für sogenannte computer-use agents. Das sind KI-Systeme, die nicht nur Text ausgeben, sondern tatsächlich Programme bedienen, klicken, tippen, Formulare ausfüllen. Laut Microsoft haben diese Agenten heute besondere Probleme mit mehrstufigen Abläufen, konkret genannt werden E-Mail-Bearbeitung und Kundensupport. Echaverse soll dieses Problem lösen, indem es Agenten in realistischen Umgebungen trainiert, statt einfach die Zahl der Trainingsaufgaben zu erhöhen. Aufgaben, Tests und Umgebungen sollen sich dabei parallel weiterentwickeln, sodass der Agent nicht nur eine feste Prüfung besteht, sondern mit steigender Schwierigkeit mitwächst. Der Beitrag ist auf den Microsoft-Research-Kanälen erschienen und beschreibt Echaverse als Forschungsprojekt.

Der Ansatz ist deshalb interessant, weil er auf eine bekannte Grenze aktueller Agenten-Modelle zielt. Viele Anbieter demonstrieren beeindruckende Einzelschritte, doch sobald ein Vorgang mehrere Werkzeuge, Wartezeiten und Rückfragen umfasst, sinkt die Zuverlässigkeit deutlich. Wenn ein Trainingsbenchmark statisch ist, kann ein Modell darauf optimiert werden, ohne wirklich robuster zu werden.

Eine Umgebung, die sich verändert, während der Agent besser wird, ist näher an der Realität eines Büro-Alltags, in dem Software-Updates, neue Kollegen und wechselnde Prozesse ohnehin ständig etwas verschieben. Für Microsoft passt das strategisch zu einem Feld, in dem gerade viele Player - von großen Modellanbietern bis zu spezialisierten Startups - um die Vorherrschaft bei Büro-Automatisierung ringen. Wer die besseren Trainingsumgebungen hat, könnte am Ende die zuverlässigeren Agenten liefern, unabhängig davon, welches Sprachmodell darunter steckt.

Vieles bleibt im vorliegenden Material allerdings offen. Der Blogbeitrag beschreibt die Idee und den Anspruch, aber konkrete Messwerte, Vergleiche mit anderen Trainingsansätzen oder eine unabhängige Bewertung sind hier nicht enthalten. Unklar ist auch, in welchem Reifegrad Echaverse steht - reines Forschungsprojekt, interner Prototyp oder Baustein für kommende Produkte. Ebenfalls nicht im Material belegt ist, ob und wann externe Entwickler oder Unternehmen Zugriff bekommen sollen, welche konkreten Aufgabenarten getestet wurden und wie mit Risiken umgegangen wird, wenn ein Agent in einer sich verändernden Umgebung eigenständig Aktionen auslöst. Wer die Ankuendigung ernst nimmt, sollte in den kommenden Wochen darauf achten, ob Microsoft Research nachlegt: mit Zahlen, mit Beispielvideos oder mit einer wissenschaftlichen Veröffentlichung, die den Ansatz vergleichbar macht. Erst dann lässt sich einschätzen, ob Echaverse mehr ist als eine wohlklingende Umbenennung für das, was in der Agenten-Forschung ohnehin gerade viele versuchen.

Für den Alltag heisst das vorerst: Der nächste E-Mail-Assistent, der wirklich zuverlässig einen Vorgang abschliesst, ist noch nicht da - aber die Forschung sortiert sich sichtbar in diese Richtung.

## REG

**Alibaba zeigt auf WAIC 2026 komplette Agenten-Cloud**

Alibaba präsentierte auf der WAIC 2026 eine Agenten-Native-Cloud mit Multi-Agenten-Orchestrierung. Dazu zeigte das Unternehmen Qwen Office und DAMO-Forschungsagenten. Das ist ein klarer Schwenk weg vom reinen Modellgeschäft.

## RES

**Bagel Labs launcht WorldDiT: Weltmodell für Robotik**

Bagel Labs hat WorldDiT veröffentlicht, ein generatives Weltmodell speziell für Robotik-Anwendungen. Weltmodelle simulieren, wie die Umgebung auf Aktionen reagiert. Das hilft Robotern, Situationen vor auszuplanen.

## PROD

**Gemini Robotics ER 2: Roboter verstehen Videos und arbeiten zusammen**

DeepMind hat Gemini Robotics ER 2 vorgestellt. Das Modell verbessert Videoverständnis, Werkzeugsteuerung und Zusammenarbeit mehrerer Roboter gleichzeitig. Damit sollen Roboter reale Aufgaben besser lösen.

## PROD

**Gemini Robotics 2 soll Ganzkörper-Intelligenz in Roboter bringen**

Google bringt mit Gemini Robotics 2 laut Ankündigung Ganzkörper-Intelligenz in Roboter. Details zu Fähigkeiten und Verfügbarkeit sind im vorliegenden Material nicht spezifiziert. Der genaue Umfang bleibt unklar.

## MARKT

**Top-KI-Startups veröffentlichen kaum noch Forschung**

Führende KI-Startups publizieren laut Bericht immer weniger ihrer Forschungsergebnisse öffentlich. Das erschwert unabhängige Überprüfung von Fähigkeiten und Sicherheitseigenschaften. Konkrete Zahlen oder Namen nennt das vorliegende Material nicht.

## REG

**GCC-Steuerungskomitee kündigt KI-Richtlinien an**

Ein GCC-Steuerungskomitee hat eine KI-Politik angekündigt. Inhaltliche Details oder konkrete Maßnahmen sind im vorliegenden Material nicht beschrieben. Der genaue Geltungsbereich bleibt unklar.

## PROD

**GPT-5.5 Pro hilft bei echtem Beweis in hochdimensionaler Geometrie**

Informatiker nutzten GPT-5.5 Pro, um einen neuen mathematischen Beweis in der theoretischen Informatik zu erarbeiten. Das ist laut Material das erste Mal, dass ein KI-Modell so direkt an einem Grundlagenforschungs-Beweis beteiligt war.

## RES

**Microsoft Research: EvoLib lässt KI-Modelle aus Erfahrung lernen**

Microsoft Research hat EvoLib vorgestellt. Das System wandelt Erfahrungen in wiederverwendbare Fähigkeiten und Erkenntnisse um, statt nur mehr Daten zu speichern. So sollen Modelle auch nach dem Training noch dazulernen.