

KI-SICHERHEIT VS. UNTERNEHMENSINTERESSEN

Sieben dokumentierte Konflikte um Ethik, Unternehmensführung und Verantwortung bei KI-Entwicklung |
Stand: 15. Oktober 2026

01 | Oktober 2026 — Tomek Korbak, Jasmine Wang und Mikita Balesni · OpenAI

Drei Sicherheitsforscher wurden entlassen. OpenAI beruft sich auf den Umgang mit vertraulichen Informationen; die Betroffenen sehen einen Zusammenhang mit Sicherheitskritik. Die Vorwürfe sind umstritten; kein freiwilliger Rücktritt.

Quelle: [Associated Press, 9.10.2026](#) ■

02 | März 2026 — Caitlin Kalinowski · OpenAI

Die Leiterin des Hardwarebereichs kündigte nach einer Vereinbarung mit dem US-Verteidigungsministerium. Sie kritisierte Entscheidungsprozess und Schutzvorkehrungen bei Überwachung und autonomen Waffen; OpenAI verweist auf bestehende Grenzen.

Quelle: [Reuters, 7.3.2026](#) ■

03 | Februar 2026 — Zoë Hitzig · OpenAI

Die Forscherin verließ OpenAI im Zusammenhang mit der Einführung von ChatGPT-Werbung. Sie warnte vor Interessenkonflikten bei der wirtschaftlichen Nutzung sensibler Nutzerinteraktionen.

Quelle: [The New York Times, Gastbeitrag von Zoë Hitzig \(via NYT-LinkedIn\)](#) ■

04 | Februar 2026 — Mrinank Sharma · Anthropic

Der Leiter eines Forschungsteams für Schutzmechanismen trat zurück und sprach in einem Abschiedsbrief von Druck, unter dem Werte zurückstehen können. Er benannte keinen konkret nachgewiesenen Regelverstoß.

Quelle: [Business Insider Deutschland, 12.2.2026](#) ■

05 | Oktober 2024 — Miles Brundage · OpenAI

Der frühere Berater für AGI-Vorbereitung verließ OpenAI, um branchenübergreifend unabhängiger zu forschen und freier veröffentlichen zu können. Kein ausdrücklich belegter Gewissensprotest.

Quelle: [Miles Brundage, eigene Erklärung, 23.10.2024](#) ■

06 | Mai 2024 — Jan Leike · OpenAI

Der ehemalige Co-Leiter des Superalignment-Teams erklärte, Sicherheitskultur und langfristige Risikoforschung seien gegenüber Produkten zurückgefallen; er sah einen grundlegenden Konflikt mit der Unternehmensführung.

Quelle: [The Guardian, 18.5.2024](#) ■

07 | Mai 2023 — Geoffrey Hinton · Google

Der KI-Pionier verließ Google, um frei über Risiken der Technologie sprechen zu können. Er stellte ausdrücklich klar, dass er Google nicht pauschal Fehlverhalten vorwerfe.

Quelle: [Reuters, 9.5.2023 \(Wiedergabe\)](#) ■

Einordnung: Die Fälle sind nicht gleichzusetzen: freiwilliger Rücktritt, öffentliche Kritik und umstrittene Entlassung sind unterschiedliche Sachverhalte. Sie veranschaulichen Spannungen zwischen KI-Sicherheit, wirtschaftlichen Zielen, unabhängiger Forschung und institutioneller Kontrolle. Quellen führen zu journalistischen Berichten bzw. Eigenerklärungen; ihre Aussagen spiegeln den jeweiligen Veröffentlichungsstand.