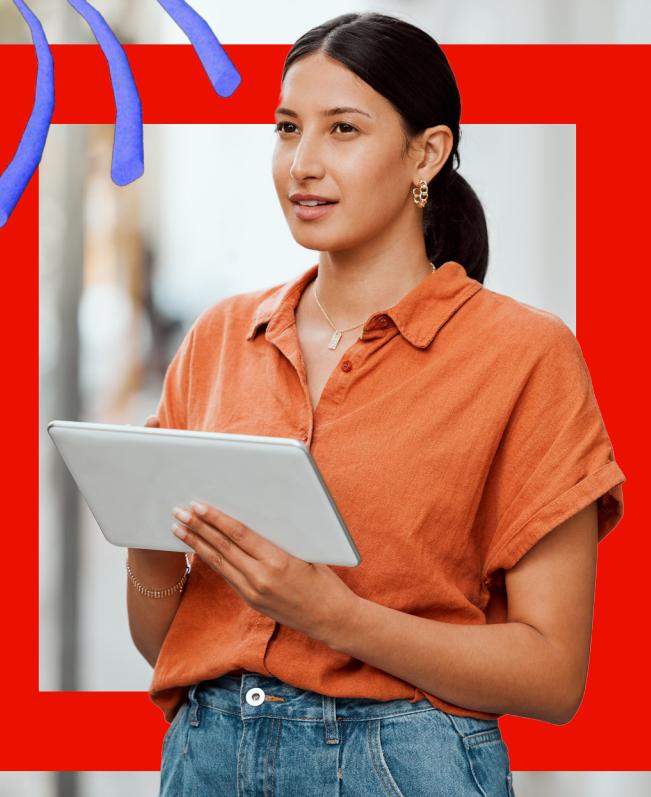


Datenblatt

Adobe Firefly: Ethik und Voreingenommenheit („Bias“).

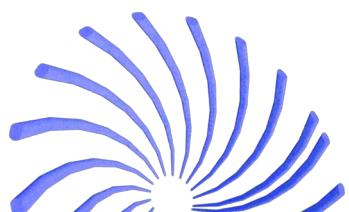


Richtig implementiert, eröffnet generative künstliche Intelligenz (KI) neue Möglichkeiten für bestehende Arbeitsabläufe und steigert die Kreativität. Wie bei den meisten Technologien, die Grenzen verschieben, bringt auch generative KI nicht nur Chancen, sondern auch Herausforderungen und Risiken mit sich – von Urheberrechtsverletzungen bis zu schädlichen Auswirkungen und Verzerrungen. Wir blicken auf über ein Jahrzehnt KI-Innovation in unseren Cloud-Technologien zurück. Auf dieser Grundlage haben wir die generative KI Firefly entwickelt und eingeführt – stets im Einklang mit unseren [KI-Ethikgrundsätzen](#) Rechenschaftspflicht, Verantwortung und Transparenz.

Adobe hat es sich zur Aufgabe gemacht, kreative, generative KI mit Bedacht zu entwickeln, damit ihr euch auf die von euch generierten Inhalte verlassen könnt.

Firefly ist kommerziell nutzbar.

Die proprietären generativen KI-Modelle von Firefly sind für die kommerzielle Nutzung unbedenklich, d. h. Kundinnen und Kunden können ihre Ergebnisse in der Gewissheit veröffentlichen, dass Adobe Lizenz- und Urheberrechtsfragen verantwortungsvoll handhabt. Wir haben die generativen KI-Modelle von Firefly mit einer großen Anzahl von Bildern, Audio- und Videodateien aus lizenzierten Inhalten wie Adobe Stock und gemeinfreien Inhalten, deren Urheberrecht abgelaufen ist, trainiert.



IP-Haftungsschutz.

Bestimmten Kundinnen und Kunden steht der Erwerb einer Berechtigung offen, die vertraglichen IP-Haftungsschutz bietet, sollte es wider Erwarten zu Ansprüchen im Zusammenhang mit von Firefly generiertem Output kommen. Dies bedeutet, dass Adobe die Verantwortung für die Rechtsverteidigung und den Schadensersatz bei Ansprüchen übernimmt, denen zufolge Firefly-Output unmittelbar die Patent-, Urheber-, Marken-, Veröffentlichungsrechte oder Datenschutzrechte Dritter verletzt oder beeinträchtigt. Adobes Haftung für Ansprüche aus Firefly-Output entspricht den gleichen Beschränkungen für technologiebasierte IP-Ansprüche, die in den bestehenden Verträgen mit Adobe festgelegt sind. Der Haftungsschutz von Adobe erstreckt sich nicht auf Ansprüche, die auf Handlungen oder Inhalten von Nutzenden beruhen oder daraus resultieren, einschließlich jeglicher Modifikation des Firefly-Outputs, des Kontexts, in dem der Firefly-Output verwendet wird, der für benutzerdefiniertes Training bereitgestellten Inhalten oder jeglicher Verwendung von Firefly-Output, die gegen die Kundenvereinbarung verstößt.

[Erfahrt mehr über die sichere kommerzielle Nutzung und den Haftungsschutz von Firefly.](#)



Für kommerzielle Nutzung konzipiert.

Vor der Freigabe von KI-Funktionen befolgen unsere Test-Teams einen detaillierten Prozess zur Entschärfung schädlicher und unerwünschter Ausgaben, zur Priorisierung von Sprache, die klassifiziert und gefiltert werden muss, zur Erstellung von sofortigen Blockierungs- und Ablehnungslisten für Prompts und zur Reduzierung von NSFW-Inhalten (Not Safe for Work).

[Erfahrt mehr über Content-Moderation von Firefly.](#)

Rechenschaftspflicht.

Bei Adobe übernehmen wir die Verantwortung für die Ergebnisse unserer KI-gestützten Tools und verfügen über standardisierte Prozesse zur Einhaltung unserer KI-Ethikverpflichtungen, darunter:



Etablierung von Governance-Verfahren zur Nachverfolgung von Trainingsdaten und KI-Modellen



Weiterbildung Entwicklungs- und Produkt-Management-Personal über ethische Fragen im Zusammenhang mit KI



Einsetzung eines Gremiums für KI (AI Ethics Review Board), das die Verbreitung der KI-Entwicklung überwacht und als Resonanzboden für ethische Bedenken in Bezug auf KI dient



Erfordernis einer KI-Folgenabschätzung (AI Impact Assessment) während des Entwicklungsprozesses unserer Dienstleistungen zur Sicherstellung, dass eine KI-Ethikprüfung stattfindet, bevor wir neue KI-Technologien einsetzen



Entwicklung von Prozessen zur Sicherstellung von Rückmeldungen und Abhilfemaßnahmen im Falle negativer Auswirkungen von KI nach dem Einsatz

Wir verfügen über Verfahren zur Entgegennahme von und Reaktion auf Bedenken bezüglich unserer KI sowohl intern als auch extern und stellen sicher, dass gegebenenfalls Korrekturmaßnahmen ergriffen werden. Zusätzlich zu unserem internen AI Ethics Review Board verfügt Firefly über einen eingebauten Feedback-Mechanismus, sodass Nutzende melden können, wenn eine Funktion ein Ergebnis liefert, das beispielsweise als voreingenommen oder inkorrekt empfunden wird. Diese Feedback-Schleife mit unserer Community ist ein wichtiger Weg, um sicherzustellen, dass unsere Tools Schaden auf ein Minimum begrenzen und verantwortungsbewusst bleiben.

[Erfahrt mehr über Adobes Rechenschaftsverpflichtung.](#)

Verantwortung.

Bei Adobe legen wir großen Wert darauf, Verantwortung für die Auswirkungen unseres Unternehmens und die Innovationen, die wir für die Welt bereitstellen, zu übernehmen. Wir sind davon überzeugt, dass durchdachte Schutzmechanismen für unsere KI-Entwicklung und -Nutzung dazu beitragen werden, das volle Potenzial von KI zum Nutzen der Gesellschaft auszuschöpfen und gleichzeitig potenziell schädliche Auswirkungen abzumildern. Unser Ansatz für eine verantwortungsvolle KI-Entwicklung umfasst Folgendes:



Durchdachtes Design eines KI-Systems



Bewertung der Interaktion mit
Endnutzenden



Beachtung der Sorgfaltspflicht, um unerwünschte schädliche Verzerrungen („Bias“) zu vermeiden



Bewertung der menschlichen Auswirkungen
von KI-Technologie



Minderung von Schaden und Voreingenommenheit.

KI ist nur so gut wie die Daten, mit denen sie trainiert wird. Wir wissen sehr wohl, dass Daten nicht frei von Verzerrungen sind. Eine verantwortungsvolle Entwicklung von KI erfordert die Vorwegnahme potenzieller Verzerrungen und daraus resultierender möglicher Schäden, die Ergreifung präventiver Maßnahmen zur Entschärfung solcher Probleme, die Messung und Dokumentation der Existenz von Schäden und die Einrichtung von Systemen zur Überwachung und Reaktion auf unerwartete schädliche Ergebnisse. Auch wenn wir unsere Systeme inklusiv gestalten, prüfen wir ständig die Auswirkungen unserer KI-Systeme auf potenziell unfaire, diskriminierende oder ungenaue Ergebnisse, die schädliche Vorurteile oder Stereotypen in Bezug auf menschliche Attribute (z. B. Herkunft, Geschlecht, Einkommen, Religion, Behinderung usw.) verbreiten könnten. Um sicherzustellen, dass unsere Verpflichtungen in Bezug auf Verantwortung und Ethik eingehalten werden, haben wir außerdem spezielle Ressourcen und standardisierte Prozesse eingeführt, um die Auswirkungen der generativen KI Firefly kontinuierlich zu analysieren:

- AI Ethics Committee: Arbeitet mit unseren Entwicklungs-Teams zusammen, um unsere KI-Ethikgrundsätze in jeder Phase der Entwicklung einzuhalten und potenziell schädliche Auswirkungen oder Voreingenommenheit proaktiv zu mindern
- AI Ethics Review Board: Strenge Prüfung aller KI-gestützten Funktionen und Produkte von Adobe vor ihrer Veröffentlichung

Transparenz.

Wir glauben, dass Transparenz Vertrauen schafft, und wir nehmen das Vertrauen ernst, das unsere Kundschaft, Partnerunternehmen und Peers in unsere Produkte setzen. Wir sind bestrebt, diese Transparenz auch auf KI-generierten Output auszudehnen, indem wir Content Credentials auf alle von Adobe Firefly generierten Bilder anwenden. Content Credentials gehören zu einem wachsenden Ökosystem von Technologien, die über die [Content Authenticity Initiative \(CAI\)](#) bereitgestellt werden. Adobe ist Gründungsmitglied dieser Initiative. Diese neue Art von fälschungssicheren Metadaten stärkt die Glaubwürdigkeit und das Vertrauen, während sie gleichzeitig der Verbreitung von Falschinformationen im Internet entgegenwirkt. Dadurch wird nicht nur sichergestellt, dass die Urheber für ihre Arbeit anerkannt werden, sondern es wird auch angezeigt, wenn generative KI in einem Bild verwendet wurde, sodass andere Nutzende oder Verbrauchende darüber informiert werden, wie das Bild entstanden ist.

Erfahrt mehr über das [Transparenzprinzip von Adobe](#).

Wendet euch an unser Team, um zu erfahren, wie Adobe sich für die verantwortungsvolle Entwicklung von KI-Technologien einsetzt, während wir weiter innovieren, lernen und wachsen.

