

Neue Dimensionen der Halbleiter-und KI-Innovationen



Timur Sugito
Aktienanalyst

Im Überblick

- **Das Moore'sche Gesetz stößt an seine Grenzen:** Die Transistorgrößen nähern sich dem Sub-Nanometerbereich, sodass eine weitere Skalierung (Scaling) aufgrund technischer und wirtschaftlicher Herausforderungen immer komplexer und kostspieliger wird.
- **Neue Ansätze, immer mehr Transistoren auf einem Chip unterzubringen:** Da sich das Scaling verlangsamt, setzen Unternehmen auf neue Technologien wie High-NA-EUV-Lithographie und Advanced Packaging, um die Leistung zu steigern.
- **Chancen mit Edge-KI:** Edge-KI, die eine Verarbeitung von Daten und Algorithmen direkt von einem Endpunktgerät aus – also nicht über den Umweg der Cloud – ermöglicht, verspricht enormes Wachstum, vor allem beim Einsatz in Smartphones.

Definitionen:

Wafer, Die und Chip

Ein Wafer ist eine dünne Scheibe, die aus einem Siliziumkristall gedreht wird. Ein Die ist ein Schaltkreis, der mechanisch oder chemisch auf einen Bereich dieses Wafers aufgebracht wird. Ein Chip besteht aus einem einzelnen, aus dem Wafer herausgeschnittenen Die und den zugehörigen Schaltelementen (Pufferspeicher/Cache, Speichercontroller).

Greenfield-Finanzierung

Investitionen in Kauf, Bau und Einrichtung neuer Fabriken auf einem noch nicht erschlossenen Grundstück.

MOSFET

Abkürzung für Metall-Oxid-Halbleiter-Feldeffekttransistor (Metal-Oxide-Semiconductor Field-Effect Transistor). Planare MOSFETs zählen zu den frühesten Transistoren, die in integrierten Schaltkreisen verwendet wurden.

FinFET

Ein Fin Field-Effect Transistor (FinFET) ist ein 3D-Transistor, bei dem das stromführende Kanal die Form einer „Finne“ hat. Das Gate umschließt die Finne an drei Seiten. Deshalb bieten FinFETs eine bessere elektrostatische Kontrolle als planare Transistoren.

GAA

Bei einem Gate-All-Around (GAA)-Transistor umschließt die Gate-Elektrode den Kanal von allen vier Seiten. Hier ist die elektrostatische Kontrolle besser als bei älteren Transistoren wie MOSFET oder FinFET.

Die Halbleiterbranche steht an einem spannenden Wendepunkt. Seit Jahrzehnten erfolgten die Fortschritte so wie im Jahr 1965 von Gordon Moore, dem Mitbegründer von Intel, prognostiziert. Er ging davon aus, dass sich die Anzahl der Transistoren in einem integrierten Schaltkreis (IC) etwa alle zwei Jahre verdoppelt. Er behielt recht, sodass immer kleinere, schnellere und kostengünstigere Chips produziert werden konnten, die unsere Infrastruktur, unsere Häuser und Wohnungen und sogar unsere mobilen Endgeräte völlig veränderten.

Bis heute versucht man, Transistoren auf unter einen Nanometer (nm) zu verkleinern und steht dabei vor noch nie da gewesenen Herausforderungen. Aufgrund von Hürden wie explodierenden Produktionskosten ändern die Chiphersteller ihre Strategien. Das Moore'sche Gesetz stößt an seine Grenzen.

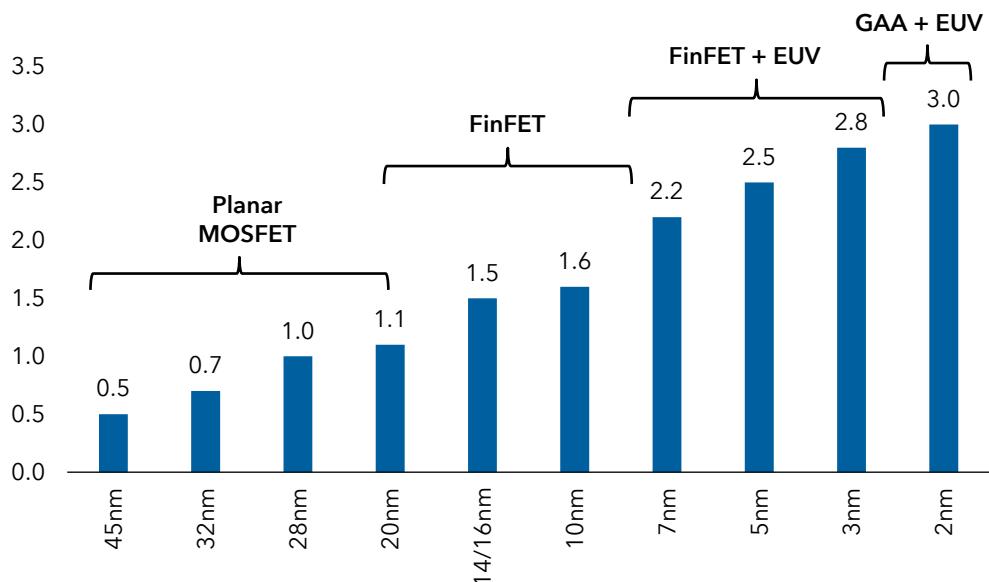
Kleinere Chips, höhere Preise

Eine der bahnbrechendsten Entwicklungen der jüngeren Geschichte der Halbleiterindustrie war die Weiterentwicklung der EUV-Lithografietechnologie (Extreme Ultraviolet Lithography) von ASML. Durch den Einsatz von Licht mit einer Wellenlänge von 13,5 nm, die viel kürzer ist als die 193 nm, die in älteren Systemen für tiefes Ultraviolett (DUV) verwendet werden, haben EUV-Maschinen den Bereich weit vorangebracht, weil sie präzise Strukturen mit einer Größe von unter 7 nm ermöglichen.

Noch präziser wird die Technik High-NA EUV (High Numerical Aperture) sein. Mit ihr könnte sich die Größe auf 2 nm oder noch weniger verkleinern lassen. Seit 2024 wurden sechs solcher Maschinen an Intel, TSMC, Samsung und SK Hynix ausgeliefert. Berichten zufolge beläuft sich ihr Stückpreis auf stolze 400 Millionen US-Dollar.¹

Scaling wird teurer

Greenfield-Investitionsintensität (Mrd. USD je 10.000 Wafer pro Monat)



Stand der Daten 31. Juli 2025. Quelle: Arete Research

1. Stand der Daten 22. Mai 2025. Quelle: CNBC

Definitionen:

Lithografie

Das Aufbringen mehrerer Schichten winziger Strukturen auf Siliziumwafer, die die Transistoren und Verbindungen eines Chips bilden

Scaling

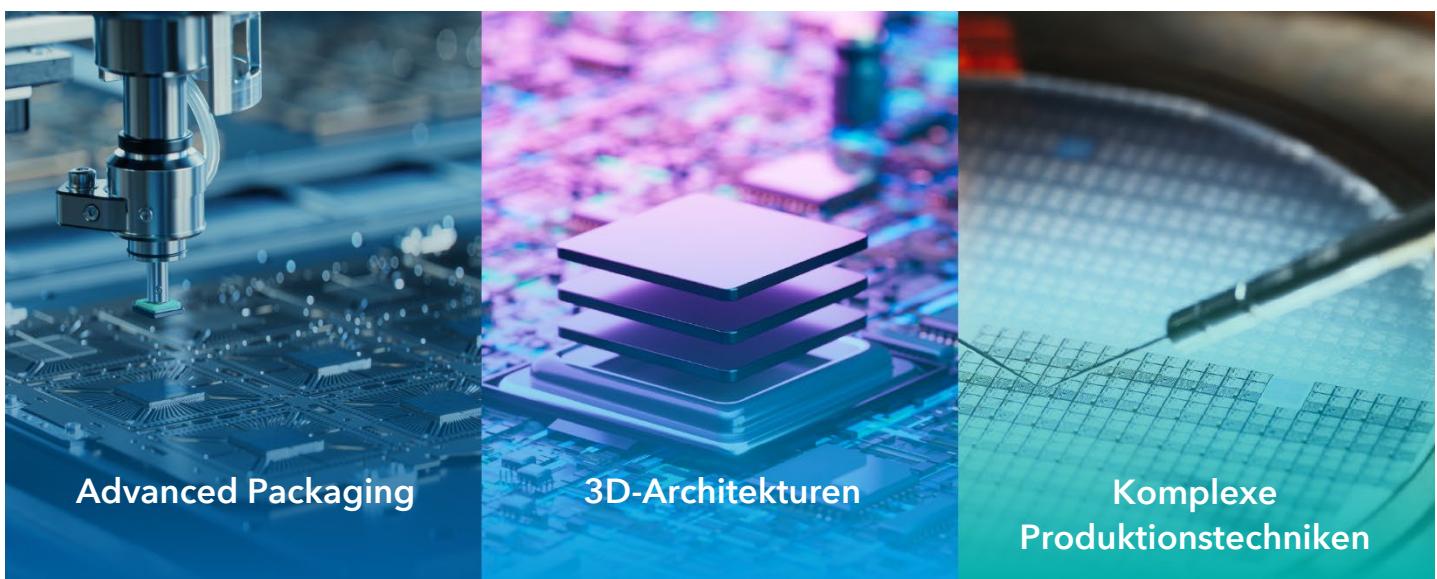
Die Verkleinerung von Transistoren und anderen Komponenten eines Chips, damit mehr davon auf dieselbe Fläche passen

High-NA-EUV-Maschinen, die erheblich teurer sind als herkömmliche EUV-Anlagen, sind ein Beispiel für einen klaren Trend in der Branche: Das weitere Scaling wird immer teurer. Viele Unternehmen aus der Halbleiterlieferkette loten deshalb Innovationen aus, zumal die Nachfrage nach noch leistungsfähigeren und effizienteren Chips weiter steigt, vor allem im Zuge der Fortschritte der generativen KI (GenAI).

„Es steht fest, dass es in den letzten 15 Jahren teurer geworden ist, mit der Verkleinerung von Chips Vorteile zu erzielen“, erklärt Timur Sugito, Investmentanalyst bei Capital Group. „Aber die Tatsache, dass das Moore'sche Gesetz an seine Grenzen stößt, veranlasst die gesamte Branche, nach Wegen außerhalb des traditionellen Scaling-Konzepts zu suchen.“

Neue Dimensionen der Halbleiter-Innovationen

Die Halbleiterbranche sucht Lösungen, um die bisherigen Grenzen der Verkleinerung zu überwinden – mit neuen Technologien und Strategien. Dazu zählen die Analyse verschiedener Architekturen und die Entwicklung neuer Testmethoden, um trotz physikalischer und wirtschaftlicher Einschränkungen mehr Chips auf einem Wafer unterzubringen.



1. Advanced Packaging

Advanced Packaging ist der Einsatz modernster Techniken, mit denen Halbleiterbauelemente so zusammengesetzt werden, dass sich Leistung, Energieeffizienz, Größe und Funktionalität verbessern. Beim traditionellen Packaging wird der Chip lediglich geschützt und mit der Platine verbunden. Advanced Packaging geht weiter: Mehrere Chips (Chiplets) oder Komponenten werden in einem kompakten, leistungsstarken Format zusammengefasst.

Der Markt für Advanced Packaging wird von mehreren wichtigen Faktoren bestimmt, die Einfluss auf die Halbleiter- und Elektronikbranche haben. Durch die Verbreitung von Künstlicher Intelligenz (KI), 5G-Netzen und High-Performance-Computing (HPC)-Anwendungen wächst der Bereich Advanced Packaging.

Diese Technologien erfordern Chips, die Daten sehr schnell verarbeiten können, wobei die räumliche Nähe von Rechenleistung und Speicher immer wichtiger wird. Advanced Packaging ermöglicht die Zusammenfassung mehrerer Chips zu einem „Paket“, mit der Folge einer kürzeren Signalverzögerung und eines niedrigeren Stromverbrauchs. So nutzen Chiplet-Architekturen, die in Rechenzentren zum Einsatz kommen, moderne Gehäuselösungen, um eine effiziente Verbindung der Chips zu ermöglichen und so die Skalierbarkeit und Leistung für KI- und HPC-Workloads zu verbessern.

Die Anfänge des Advanced Packaging reichen nur bis in die frühen 2000er-Jahre zurück, aber 2024 hatte der Markt bereits etwa 40 Milliarden US-Dollar Volumen. Bis 2030 werden es Schätzungen zufolge 55 Milliarden sein. Das macht Advanced Packaging zu einem dynamischen und wachstumsstarken Bereich innerhalb der Halbleiterbranche.²

2. 3D-Architekturen

Weil eine Verkleinerung immer schwieriger und teurer wird, geht das Design von Halbleiterchips mittlerweile in Richtung 3D. Dieser Ansatz ermöglicht eine höhere Dichte und einen geringeren Stromverbrauch pro Flächeneinheit des Siliziumwafers, wodurch sich Rechenleistung und Energieeffizienz verbessern.

3D-Architekturen finden sich immer häufiger in mehreren Segmenten des Halbleiterbereichs. Beispiele sind Speicherchips (NAND)³ und Logikchips. Bei 3D-NAND werden die Speicherchips vertikal in bis zu mehreren hundert Schichten aufgebracht, was für eine immense Steigerung der Speicherdichte sorgt und die Kosten je Bit sinken lässt.⁴ 2024 hatte der Markt etwa 28 Milliarden US-Dollar Volumen. Bis 2034 sollen es dank der Nachfrage für den Einsatz in Datenzentren, Smartphones und Automobilsystemen 150 Milliarden sein.⁵

Unterdessen befindet sich 3D-DRAM noch am Anfang der Entwicklung. Unternehmen wie Micron, Samsung Electronics und SK Hynix arbeiten aktiv an der Weiterentwicklung, aber bis zur Massenherstellung wird es noch einige Jahre dauern.

„Hier kann die Möglichkeit, langfristig zu investieren, ein großer Vorteil sein“ sagt Sugito. „Investoren, die sich Zeit für Treffen mit Unternehmen, Branchenexperten und Zulieferern nehmen, vor allem in zyklischen Branchen, zu denen auch Halbleiter zählen, können früh langfristige Trends und Unternehmen finden, die davon profitieren dürften.“

2. Daten aus dem Advanced Packaging Market Size Industry Report 2030. Quelle: Grand View Research
3. DRAM ist ideal für den zeitlich begrenzten High-Speed-Zugriff auf Daten, während sich NAND besser für die langfristige Datenspeicherung eignet.

4. „Kosten je Bit“ ist das Maß für die Kosten der Speicherung oder Übertragung eines einzelnen Datenbits.
5. Stand der Daten 26. April 2024. Quelle: Future Market Insights

Definitionen:

Halbleiterdeposition

Der Prozess, bei dem dünne Schichten von Materialien auf einen Wafer aufgebracht werden, um die verschiedenen Komponenten eines Chips zu erstellen - sehr präzise und gleichmäßig

Ätzen

Der Prozess, bestimmte Teile eines Halbleiter-Wafers zu entfernen, um die komplexen Muster und Strukturen zu schaffen, die für Schaltkreise benötigt werden

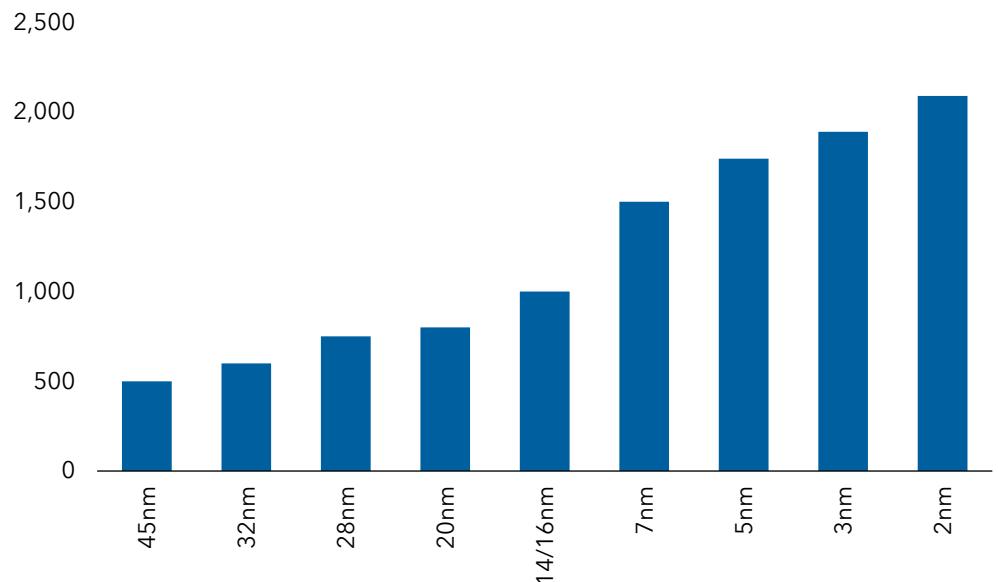
3. Komplexe Produktionstechniken

Zugleich erfordern die Umstellung auf 3D-Architekturen und unterschiedliche Integrationsanforderungen Durchbrüche in der Halbleiterfertigungstechnik. Beim Aufbau vertikaler Strukturen stößt man auf völlig neue Probleme. Beispielsweise müssen ultradünnen Schichten aufgetragen werden, damit die gesamte Struktur nicht zu dick wird. Dies erfordert neuartige Depositionstechnologien, die aufs Atom präzise sein müssen. Eine weitere Herausforderung sind die notwendigen, sehr schmalen, langen Verbindungslöcher (sogenannte VIAs), die nur mit Ätztechniken mit einem hohen Aspektverhältnis (Tiefe/Durchmesser) gelingen.

Diese Innovationen ermöglichen nicht nur den Fortschritt. Sie werden zur Schicksalsfrage für Unternehmen wie Applied Materials, Lam Research Corp. und Tokyo Electron, weil durch sie Dutzende, wenn nicht Hunderte neuer Prozessschritte notwendig werden. Und sie werden in den nächsten Jahren den Wettbewerb in der Halbleiterbranche prägen. Außerdem wird es immer schwerer, die Funktionalität der Produkte zu messen, je komplexer und kompakter sie werden. Die Umstellung auf 3D-Architekturen und die unterschiedliche Integration mehrerer Chips in einem „Paket“ verkomplizieren die Beurteilung zusätzlich.

„Mit den steigenden Herstellungskosten werden Fehler immer teurer. Deshalb gewinnen Strategien zur Fehlervermeidung und Qualitätsmessung an Bedeutung“, erklärt Sugito. „Aus diesem Grund spielen auch Geräte zur Kontrolle und Evaluierung dieser Prozesse eine immer wichtigere Rolle in der Halbleiterproduktion.“

Die Chipherstellung wird immer komplexer (Zahl der Prozessschritte je Node)



Quellen: Applied Materials, Arete Research, Omdia

ASML und KLA Corporation haben viel in diesen Bereich investiert. Das E-Beam-System von ASML ist extrem wichtig für das Erkennen winziger, zufälliger Defekte, die von optischen Geräten während des Lithografieprozesses übersehen werden können. KLA hat sich auf Prozesskontrolllösungen in der

gesamten Halbleiterherstellung spezialisiert. Einige seiner Geräte sind in der Lage, auf einem einzigen 12-Zoll-Wafer etwa acht Billionen Transistoren zu prüfen, von denen jeder nur fünf Nanometer breit ist.⁶

Die Hersteller von Prozesskontrollanlagen entsenden in der Regel auch Servicetechniker zu den Kunden, um die Ingenieurteams der Chiphersteller zu unterstützen. Sie kalibrieren Geräte in der Forschungs- und Entwicklungsphase, kümmern sich um Probleme während der Produktion und helfen, die an die Fertigung gestellten Anforderungen zu erfüllen. Außerdem aktualisieren sie die Anlagen und programmieren sie neu, wenn sie für andere Aufgaben eingesetzt werden.

Die Komplexität dieser Geräte und die Kompetenz ihrer Servicetechniker sorgen dafür, dass die Anbieter voll in die Fertigungsprozesse ihrer Kunden integriert sind. Ein Austausch von Geräten und Anlagen erfordert immense Investitionen sowie eine Umstellung der Prozesse, neue Schnittstellen und eine Umschulung durch die verschiedenen Serviceteams. Deshalb sind die Kunden meist sehr treu. Unterstützt wird diese Treue durch eine Fehlerdatenbank, die Daten einer großen Plattform nutzt. Das bringt Vorteile für die Kunden, sodass sie eng an einen Anbieter gebunden sind.

Von der Cloud-KI zur Edge-KI

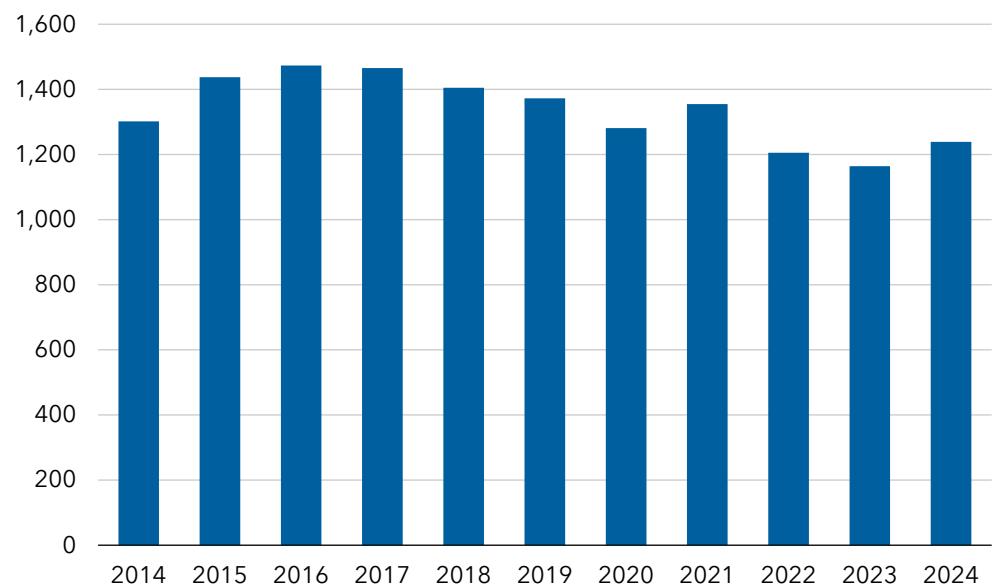
Bei Halbleiterinnovationen wird Durchbrüchen in der Fertigung und im High-Performance-Computing die meiste Aufmerksamkeit geschenkt, aber auch auf der Anwendungsebene gibt es immer mehr Neuerungen, vor allem im Zusammenhang damit, wie und wo KI eingesetzt wird. Innovationen entstehen zunehmend nicht mehr nur dadurch, wie die Chips hergestellt werden, sondern durch die Art ihrer Anwendung. Hier steht Edge-KI ganz vorne. Die Verschiebung ist die Folge eines allgemeineren Trends: Da die Endgeräte immer intelligenter und autonomer werden, steigt die Nachfrage nach lokalisierter Datenverarbeitung in Echtzeit. Dadurch entstehen neue Möglichkeiten zur Leistungssteigerung, höherer Effizienz und besseren Nutzererfahrungen – nicht nur in der Cloud.

Smartphone-Hersteller loten diesen Bereich aktiv aus. Einige wichtige Akteure wie Apple, Samsung und mehrere chinesische Zulieferer (OEMs) wie Oppo, Vivo, Xiaomi und Huawei vermarkten bereits ihre Fortschritte in Edge-KI-Technologie. Dadurch entstehen zahlreiche Investmentchancen, vor allem angesichts der Stagnation des Smartphone-Absatzes in den letzten Jahren, die auf eine Kombination aus Marketsättigung, unsicherer Wirtschaftslage und längeren Weiterentwicklungszyklen zurückzuführen ist. Die Einführung modernster KI-Funktionen könnte den Verkauf wieder in Schwung bringen – sowohl die Absatzzahlen als auch technologische Innovationen. Daher ist davon auszugehen, dass der Weltmarkt für Edge-KI von 20,4 Milliarden US-Dollar im Jahr 2023 bis 2032 auf 270 Milliarden US-Dollar wachsen wird.⁷

6. Stand der Daten 1. August 2025. Quelle: Morningstar

7. Stand der Daten 26. April 2024. Quelle: Future Market Insights

Kann Edge-KI den internationalen Smartphone-Markt beleben? Smartphone-Auslieferungen weltweit von 2004 bis 2014 (Mio. Stück)



Quelle: Statista

Nach Ansicht von Sugito ist das Fehlen einer App, die alles andere in den Schatten stellt, eine der größten Herausforderungen für die Edge-KI. Zwar haben Smartphone-Hersteller nützliche Anwendungen eingeführt, wie Circle to Search von Samsung und Echtzeitübersetzungen, aber bislang hat noch niemand eine „Star-App“ entwickelt, die auf eine breite Akzeptanz stößt. In Bereichen wie persönliche Assistenz und Echtzeitübersetzung gibt es zweifellos Potenzial, aber eine „Must have“-Funktion, die das KI-Erlebnis neu definieren würde, hat sich noch nicht herauskristallisiert.

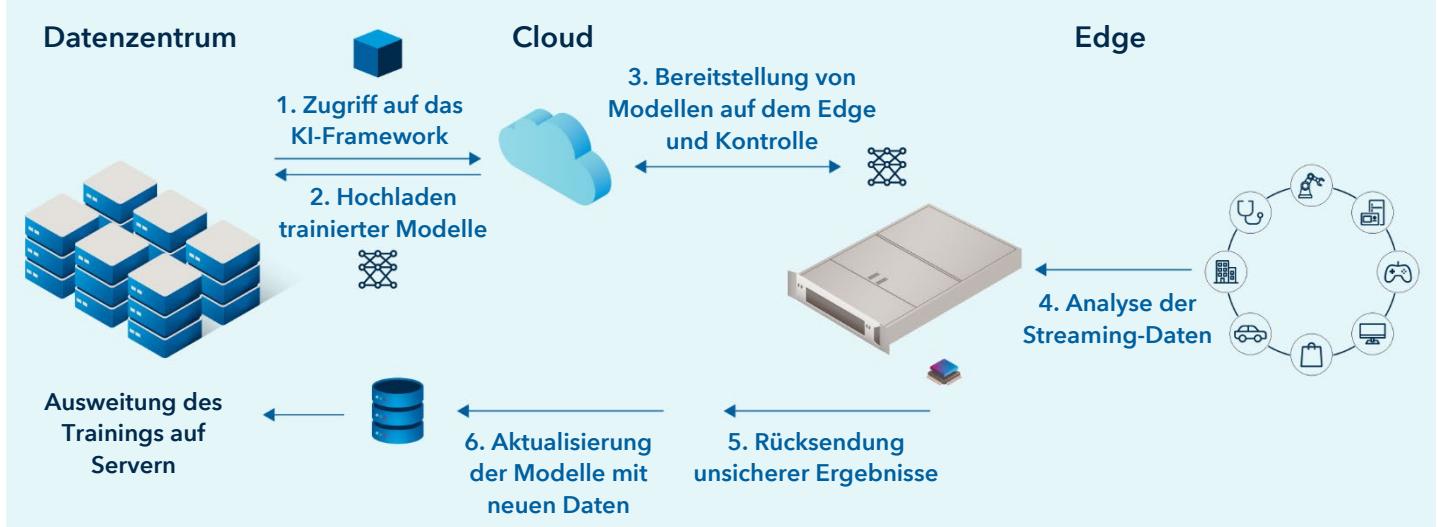
„Wenn ich beispielsweise nach Singapur reise und mein Smartphone frage, welche Aktivitäten es mir für den Nachmittag empfehlen kann, erhalte ich Tipps für Abendessen, Sightseeing und werde an die entsprechenden Orte navigiert“, erklärt Sugito. „Aber Sie und ich würden ähnliche Vorschläge erhalten, weil sie aus allgemeinen Quellen stammen und nicht personalisiert sind.“

Wenn ich diesen Modellen Zugriff auf persönliche Daten auf meinem Telefon gebe – beispielsweise auf das von mir bevorzugte Fitnesstraining, kulinarische Vorlieben, mein Budget fürs Essengehen und besondere Interessen – werden die Empfehlungen individueller und damit wertvoller für mich.“

Aufgrund der Fähigkeit, kontextabhängig zu reagieren, bieten diese Modelle erheblichen Mehrwert, weil sie dem Nutzer das Leben einfacher machen. Deshalb wollen viele Unternehmen jetzt eine Kombination aus geräteinternen und Cloud-Lösungen entwickeln und/oder in ihre Smartphones einbauen.

„Der Übergang von 3G auf 4G dauerte sechs Jahre, während der von 4G auf 5G schon nach vier Jahren erledigt war“, betont Sugito. „Angesichts des großen Interesses von Smartphone-Herstellern und Kunden ist völlig klar, wohin die Reise für Edge-KI in den nächsten Jahren geht.“

Wie funktioniert Edge-KI?



Timur Sugito ist Aktienanalyst bei Capital Group und verantwortlich für asiatische Halbleiterunternehmen, Videospiele und Medienfirmen sowie Branchengeneralist für kleine und mittelgroße Unternehmen. Er hat 18 Jahre Investmentserfahrung und ist seit neun Jahren im Unternehmen. Er hat einen Master in Ingenieurwissenschaften vom Imperial College der University of London und ist Chartered Financial Analyst®. Sugito arbeitet in Hongkong.

Die Aussagen einer bestimmten Person geben deren persönliche Einschätzung wieder. Sie entspricht möglicherweise nicht der Meinung anderer Mitarbeiter der Capital Group oder ihrer Tochtergesellschaften. Die zur Verfügung gestellten Informationen erheben keinen Anspruch auf Vollständigkeit und stellen keine Beratung dar. Diese Information ist für den internen und vertraulichen Gebrauch des Empfängers bestimmt und sollte nicht an Dritte weitergegeben werden. Bei dem Dokument handelt es sich um allgemeine Informationen und weder um Anlage-, Steuer- oder sonstige Beratung noch um eine Aufforderung, irgendein Wertpapier zu kaufen oder zu verkaufen. Stand aller Informationen und Einschätzungen ist das angegebene Datum. Quelle: Capital Group (falls nicht anders angegeben). Die Capital Group trifft angemessene Maßnahmen, um Informationen von Dritten zu erhalten, die sie für korrekt hält. Dies kann jedoch nicht garantiert werden.

Falls nicht anders angegeben, wird dieses Dokument von Capital International Management Company Sàrl (CIMC) herausgegeben. CIMC ist von der Luxemburger Commission de Surveillance du Secteur Financier (CSSF) zugelassen und wird von ihr beaufsichtigt.

In der Schweiz wird dieses Dokument von Capital International Sàrl herausgegeben, einem von der Schweizer Finanzmarktaufsicht (FINMA) genehmigten und regulierten Unternehmen.

Alle Handelsmarken von Capital Group sind Eigentum von The Capital Group Companies, Inc. oder einer Tochtergesellschaft. Alle anderen genannten Unternehmensnamen sind Eigentum der jeweiligen Unternehmen.

© 2025 Capital Group. Alle Rechte vorbehalten. WF8915166 DE ALL