Große Sprachmodelle: Von Abstraktion zu Realität

Veröffentlicht am 8. Oktober 2024

Christopher Gannatti, CFA
Global Head of Research

Die wichtigsten Erkenntnisse

- Große Sprachmodelle zeigen reale Auswirkungen und ersparen Unternehmen wie Amazon viele Entwicklerjahre.
- Grundlagenmodelle dienen als wichtige Infrastrukturebene für zukünftige KI-Innovationen.
- LLMs steigern die Produktivität und Effizienz in der Softwareentwicklung und im Geschäftsbetrieb.
- Das wirtschaftliche Potenzial von LLMs wird wachsen, da große Unternehmen weiter stark in Infrastruktur und Talente investieren.
- Die Weiterentwicklung von KI-Modellen kann stärker spezialisierte Schnittstellen hervorbringen, die auf bestimmte Aufgaben zugeschnitten sind.
- Verbundene Produkte WisdomTree Artificial Intelligence UCITS ETF – USD Acc Mehr erfahren

Das Problem bei großen Sprachmodellen (Large Language Models, LLMs) besteht darin, dass die Vorstellung, was sie überhaupt sind, für Menschen, die keine Softwareingenieure sind, schwierig ist. Noch komplizierter wird es, wenn man bedenkt, dass die Entwicklung und das Training der größten LLMs Hunderte von Millionen Dollar kostet.

Dennoch haben die Geschäftsführer einiger der größten Unternehmen der Welt angedeutet, dass sie in den kommenden Jahren insgesamt mehr als 1 Billion US-Dollar aufwenden wollen, um mehr Recheninfrastruktur aufzubauen, damit mehr dieser Modelle praktikabler betrieben werden können1&2.

Wir suchen ständig nach Anwendungsfällen oder Geschichten, mit denen wir die Abstraktion eines LLM in echte geschäftliche Auswirkungen umsetzen können. Wenn wir genügend solcher Fälle finden, könnten wir vielleicht schon bald beobachten, dass diese Auswirkungen in die breiteren Statistiken zur wirtschaftlichen Produktivität einfließen.

Daher waren wir begeistert, als Amazon-CEO Andy Jassy Folgendes über das Q-System von Amazon erklärte, das im Grunde ein LLM ist, das Softwarecode erzeugen kann3:

Die durchschnittliche Zeit für das Upgrade einer Anwendung auf Java 17 verringerte sich von normalerweise 50 Entwicklertagen auf nur wenige Stunden. Unserer Schätzung nach konnten wir dadurch 4.500 Entwicklerjahre an Arbeit einsparen (ja, diese Zahl ist verrückt, aber sie ist real).

In weniger als sechs Monaten konnten wir über 50 % unserer Java-Produktionssysteme mit einem Bruchteil des üblichen Zeit- und Arbeitsaufwands auf modernisierte Java-Versionen umstellen. Und unsere Entwickler haben 79 % der Überprüfungen von automatisch generiertem Code ohne zusätzliche Änderungen übermittelt.

Die Vorteile reichen über den Aufwand hinaus, den wir den Entwicklern erspart haben. Durch die Upgrades wurde die Sicherheit erhöht und die Infrastrukturkosten wurden gesenkt, was jährliche Effizienzgewinne in Höhe von schätzungsweise 260 Millionen US-Dollar ermöglicht.

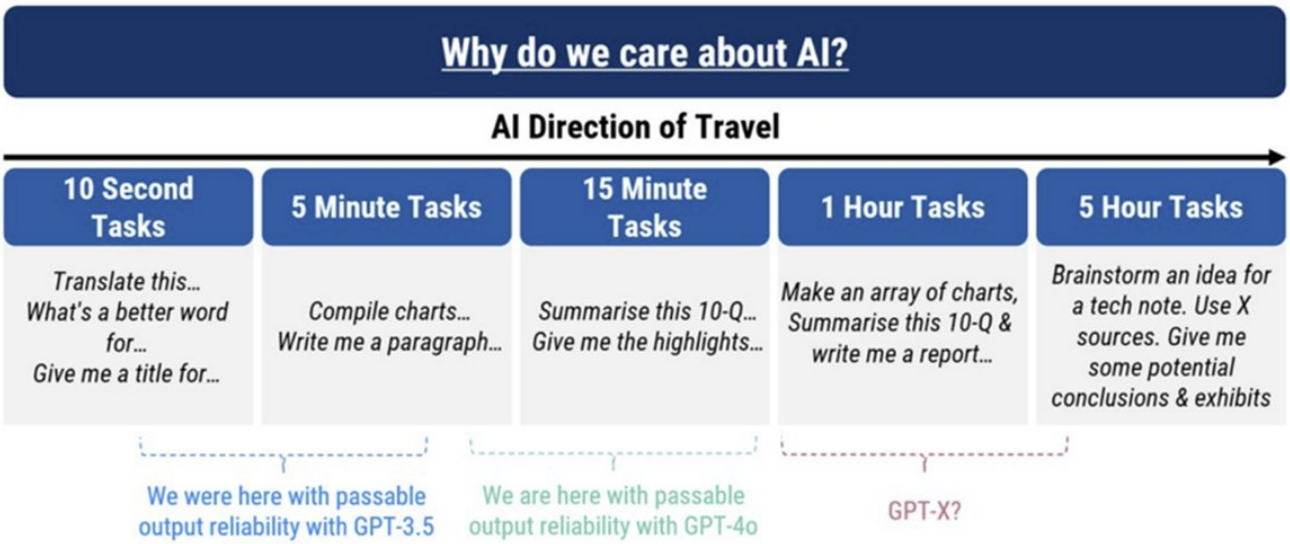
4.500 Entwicklerjahre??? Jährliche Effizienzgewinne von 260 Millionen US-Dollar??? Das sind enorme Zahlen. Wir sind uns bewusst, dass die KI-Revolution noch in den Kinderschuhen steckt. Aber vielleicht inspiriert Jassys Darstellung der Wirkung von KI für die Entwickler seines Teams andere, ähnliche Erfahrungen zu schildern und zu veröffentlichen.

Ein Konzept für KI-Entwicklungen4

Abbildung 1 trägt unseres Erachtens maßgeblich dazu bei, den Zweck dieser LLMs auf andere Weise zu veranschaulichen. Allen Wissensarbeitern ist klar, dass ihre Arbeit aus verschiedenen Aufgabenbereichen besteht und dass jeder Bereich einen sehr unterschiedlichen Zeitaufwand verursachen kann.

Die aktuellen Versionen von LLMs können auf einfache Fragen oder E-Mails antworten, aber es ist weit weniger klar, wie diese Systeme von Grund auf neue Berichte oder originelle Ideen erstellen können. Das heißt nicht, dass sie das nicht können – es heißt nur, dass der Grad der Überprüfung, der für einfache Antworten auf einfache Fragen nötig ist, ein ganz anderer ist als der Grad der Überprüfung, den eine brandneue Präsentation mit möglicherweise 60 Originalfolien, die alle von KI entwickelt wurden, erfordert.

Abbildung 1: Ein Konzept für KI-Entwicklungen



Quelle: Stanley, Edward et al. „Mapping AI's Rate of Change." Morgan Stanley Research. 4. Juni 2024.

Sind LLMs Software, Infrastruktur oder beides?

Manche Leute bezeichnen LLMs auch als „Grundlagenmodelle“. Das Wort „Grundlage“ regt zum Nachdenken an, denn in vielen Zusammenhängen ist eine Grundlage etwas, worauf man aufbauen kann. Wenn wir über die Wertschöpfung in einigen wenigen Ökosystemen nachdenken:

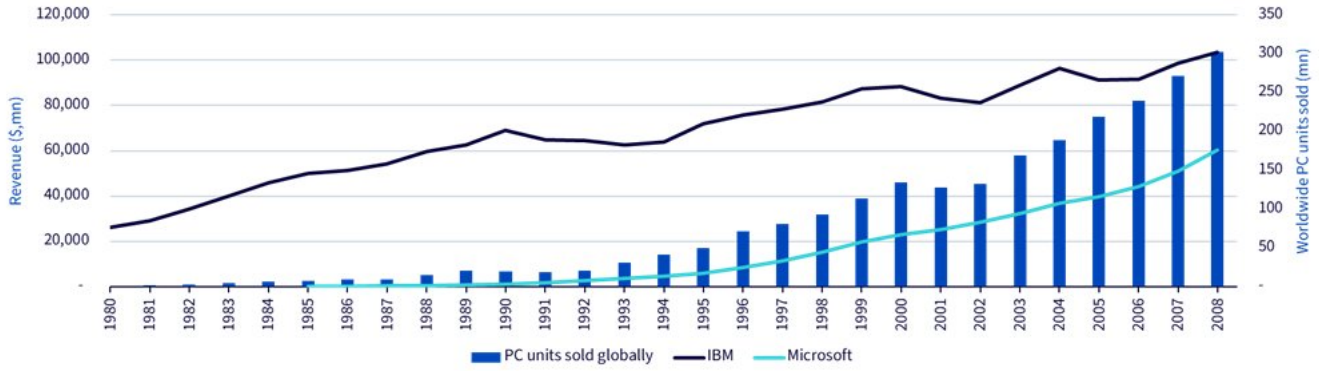
- Warum sind Computer wertvoll? Man könnte über die verschiedenen Hardwarekomponenten nachdenken und den Wert jedes dieser Teile zusammenrechnen, um den Wert eines Computers zu bestimmen. Wenn wir jedoch die Entstehungsgeschichte von Computern betrachten, gab es eine Zeit, in der man dachte, dass die Menschen ihre eigenen Computer bauen wollten. Die Benutzerfreundlichkeit, die Microsoft in den Umgang mit dem Computer einbrachte, war ungemein wertvoll, ebenso wie die Tatsache, dass viele Nutzer diese Softwarepakete nachfragten und lernten, sie alle gleichzeitig zu benutzen.
- Warum ist das Internet wertvoll? Man könnte sagen, dass wir dank des Internets Informationen viel einfacher um die Welt schicken können. Was aber halten wir von Google – einem Unternehmen, das die im Internet verfügbaren Informationen so organisieren wollte, dass man viel schneller das finden kann, was man braucht? Wie würden wir wohl dastehen, wenn wir das Internet ohne Suchmaschinen nutzen müssten?
- Wenn Sie an das iPhone denken: Liegt der Wert in den Komponenten des Telefons und in dem, was die Technologie konkret leisten kann, oder liegt der Wert in all den verschiedenen Anwendungen, die auf der iOS-Software aufbauen?

In jedem dieser Fälle lautet die reale Antwort wahrscheinlich, dass alles einen gewissen Wert hat. Doch der Grund, warum wir uns oft auf Produkte und Dienstleistungen beziehen, die von Unternehmen mit einer Marktkapitalisierung von über 1 Billion US-Dollar angeboten werden, ist, dass verschiedene Effekte auf sie zurückgehen und sich exponentiell vervielfachen. Hinzu kommen gigantische Netzwerkeffekte – nichts führt zu mehr Nutzern und mehr Wachstum als eine große Ausgangsbasis von Nutzern.

Abbildungen 2a und 2b vermitteln uns den Eindruck, dass wir so etwas schon einmal gesehen haben:

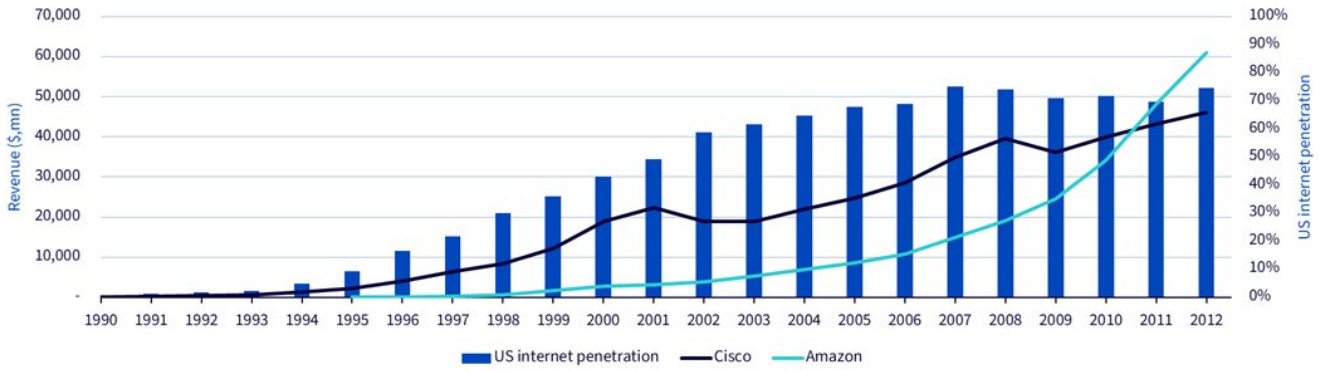
- Abbildung 2a setzt die Einnahmen von IBM und Microsoft ins Verhältnis zur Anzahl der weltweit verkauften PCs. Es leuchtet ein, dass IBM zunächst PCs verkauft oder es anderen Unternehmen ermöglicht, die so genannten IBM-Klone zu bauen und anzubieten. Mit einer wachsenden Zahl von PCs und schließlich einem gewissen Niveau von PC-Nutzern steigen die Einnahmen von Microsoft, weil alle diese PC-Nutzer von den Vorteilen der Microsoft-Betriebssysteme hören und diese nutzen wollen. Das System von Microsoft hat erst dann einen Wert, wenn PCs auf fast jedem Schreibtisch stehen.
- Abbildung 2b vergleicht die Internetverbreitung in den USA mit den Einnahmen von Cisco und Amazon.com. Mit den Geräten von Cisco können mehr Unternehmen und Menschen das Internet nutzen. Mit einer steigenden Zahl von Internetnutzern ist plötzlich der Wert des Angebots von Amazon.com – des E-Commerce – immer offensichtlicher. Es würde nicht einleuchten, dass die E-Commerce-Aktivitäten von Amazon.com wertvoll sind, bevor eine kritische Masse von Anwendern das Internet nutzt.

Abbildung 2a: Verhältnis der weltweit verkauften PCs zu den Einnahmen von IBM und Microsoft im Zeitverlauf



Quelle: WisdomTree, Bloomberg. Die Daten zu den verkauften PCs stammen aus „Total Share: Personal Computer Market Share 1975-2010“, Jeremy Reimer und Gartner. **Die historische Wertentwicklung ist kein Hinweis auf künftige Ergebnisse, und Anlagen können im Wert sinken.**

Abbildung 2b: Verhältnis der Internetverbreitung in den USA zu den Einnahmen von Cisco und Amazon.com im Zeitverlauf



Quelle: WisdomTree, Bloomberg, Weltbank. **Die historische Wertentwicklung ist kein Hinweis auf künftige Ergebnisse, und Anlagen können im Wert sinken.**

Fazit: Was werden wir auf Grundlagenmodellen errichten?

Das Interessante an Grundlagenmodellen, zumindest in der zweiten Hälfte des Jahres 2024, ist, dass nur die größten und rentabelsten Unternehmen der Welt die Ressourcen haben, sie weiterzuentwickeln und voranzutreiben. Auch wenn es den Anschein hat, dass einige davon Teil unabhängiger Unternehmen sind, übernehmen die größten Konzerne der Welt in der Regel erhebliche finanzielle Beteiligungen, die die entsprechenden und ständig steigenden Investitionen in Talente und die benötigte Recheninfrastruktur ermöglichen.

Wir wissen nicht genau, was als Nächstes kommt – und wir sind uns bewusst, dass dies eine Billionen-Dollar-Frage sein kann. Dennoch wissen wir, dass die Menschen tendenziell über diskrete, einzelne

Aufgaben nachdenken und vielleicht nicht immer auf ein Modell zugreifen müssen, das alle wichtigen von uns entwickelten Prüfungen bestehen kann. Uns gefällt aber die Vorstellung von stärker spezialisierten Schnittstellen, die auf spezifischeren Aufgaben basieren und dann Teile der umfassenderen Modelle nutzen können, um Arbeiten zu erledigen.

Wie auch immer sich das Bild entwickelt, unserer Meinung nach werden die größten Unternehmen der Welt, die mit diesen Grundlagenmodellen verbunden sind, in den kommenden Jahren eine wichtige Rolle spielen.

1 Goldman Sachs, https://www.datacenterdynamics.com/en/news/goldman-sachs-1tn-to-be-spent-on-ai-data-centers-chips-and-utility-upgrades-with-little-to-show-for-it-so-far/
2 Nvidia, https://www.linkedin.com/posts/leeps_nvidia-ceo-predicts-1-trillion-will-be-spent-activity-7101349410281836544-hB72/
3 Quelle: Auszug aus einem LinkedIn-Post von Amazon-CEO Andy Jassy, bezogen auf https://nextbigteng.substack.com/p/hello-ai-world-evolution-of-developer-economy-in-the-age-of-ai.
4 Quelle: Stanley, Edward et al. „Mapping AI's Rate of Change." Morgan Stanley Research. 4. Juni 2024.

Important Risks Related to this Article

Wichtige Informationen

Im Europäischen Wirtschaffsraum („EWR“) herausgegebene Marketingkommunikation: Dieses Dokument wurde von WisdomTree Ireland Limited, einer von der Central Bank of Ireland zugelassenen und regulierten Gesellschaff, herausgegeben und genehmigt.

In Ländern außerhalb des EWR herausgegebene Marketingkommunikation: Dieses Dokument wurde von WisdomTree UK Limited, einer von der United Kingdom Financial Conduct Authority zugelassenen und regulierten Gesellschaff, herausgegeben und genehmigt.

WisdomTree Ireland Limited und WisdomTree UK Limited werden jeweils als „WisdomTree“ bezeichnet. Unsere Richtlinie über Interessenkonflikte und unser Verzeichnis sind auf Anfrage erhältlich.

Nur für professionelle Kunden. Die in diesem Dokument enthaltenen Informationen dienen ausschließlich Ihrer Information und stellen weder ein Angebot zum Verkauf bzw. eine Aufforderung oder ein Angebot zum Kauf von Wertpapieren oder Anteilen dar. Dieses Dokument sollte nicht als Basis für eine Anlageentscheidung verwendet werden. Anlagen können an Wert zunehmen oder verlieren und Sie können einen Teil oder den gesamten Betrag der Anlage verlieren. Die Wertentwicklung in der Vergangenheit ist nicht notwendigerweise ein Hinweis auf zukünffige Ergebnisse. Anlageentscheidungen sollten auf den Angaben im entsprechenden Prospekt sowie auf unabhängiger Anlage-, Steuer- und Rechtsberatung basieren.

Die Anwendung von Verordnungen und Steuergesetzen kann off zu unterschiedlichen Interpretationen führen. Alle in dieser Mitteilung dargestellten Ansichten oder Meinungen spiegeln die Auffassung von WisdomTree wider und sollten nicht als aufsichtsrechtliche, steuerliche oder rechtliche Beratung ausgelegt werden. WisdomTree übernimmt keine Garantie oder Zusicherung hinsichtlich der Richtigkeit der in dieser Mitteilung geäußerten Ansichten oder Meinungen. Anlageentscheidungen sollten auf den Angaben im entsprechenden Prospekt sowie auf unabhängiger Anlage-, Steuer- und Rechtsberatung basieren.

Bei diesem Dokument handelt es sich nicht um Werbung bzw. eine Maßnahme zum öffentlichen Angebot von Anteilen oder Wertpapieren in den USA oder einer zugehörigen Provinz bzw. einem zugehörigen Territorium der USA, und es darf unter keinen Umständen als solche verstanden werden. Weder dieses Dokument noch etwaige Kopien dieses Dokuments sollten in die USA mitgenommen, (direkt oder indirekt) übermittelt oder verteilt werden.

Obwohl WisdomTree bestrebt ist, die Richtigkeit des Inhalts dieses Dokuments sicherzustellen, übernimmt WisdomTree keine Gewährleistung oder Garantie für seine Richtigkeit oder Genauigkeit. Die Drittanbieter, deren Dienste in Anspruch genommen werden, um die in diesem Dokument enthaltenen Informationen zu beziehen, übernehmen keine Gewährleistung oder Garantie jeglicher Art bezüglich dieser Daten. Dort, wo WisdomTree seine eigenen Ansichten in Bezug auf Produkte oder Marktaktivitäten äußert, können sich diese Auffassungen ändern. Weder WisdomTree, noch eines seiner verbundenen Unternehmen oder einer seiner jeweiligen leitenden Angestellten, Verwaltungsratsmitglieder, Partner oder Mitarbeiter übernimmt

irgendeine Haffung für direkte Schäden oder Folgeschäden, die durch die Verwendung dieses Dokuments oder seines Inhalts entstehen.