

SFI Public Discussion Note

Die Rolle von KI bei der Transformation der Finanzpraxis



November 2024

Einführung



Prof. Markus Leippold
*SFI Senior Chair und Professor
für Financial Engineering an
der Universität Zürich*

Mit den Public Discussion Notes fördert das Swiss Finance Institute (SFI) aktiv eine fundierte Diskussion von Themen, die für die Finanzindustrie, Politik und Wissenschaft relevant sind. Darüber hinaus macht das SFI die gewonnenen Erkenntnisse mittels Forschung, Publikationen, Master Classes und Konferenzen zugänglich.

Diese *Public Discussion Note* basiert auf der Forschung, die mein Team und ich in den vergangenen Jahren in den Bereichen Nachhaltigkeit und Künstliche Intelligenz durchgeführt haben. Ich möchte meinen Mitautorinnen und -autoren Julia Bingler, Chiara Colesanti-Senni, Glen Gostlow, Jingwei Ni, Tobias Schimanski, Dominik Stammbach, Saeid Vaghefi und Tingyu Yu, sowie allen wissenschaftlichen Assistierenden meinen Dank aussprechen. Diese talentierten und motivierten jungen Forschenden waren massgeblich am Fortschritt dieser Arbeit beteiligt. Ich hoffe, dass ihre Beiträge zukünftige Forschende dazu inspirieren werden, in diesem wichtigen Bereich weiter voranzukommen.

Diese Version ist die Übersetzung der Originalversion in englischer Sprache. Die Originalversion ist unter <https://sfi.ch> abrufbar.

"Ich denke, dass altmodische Intelligenz ziemlich gut funktioniert" (Munger, 2023).

Seit OpenAI ChatGPT im November 2022 eingeführt hat, versetzt das Potential der künstlichen Intelligenz (KI) die Wall Street in Begeisterung. Doch selbst wenn das Thema KI in Telefonkonferenzen zu Finanzergebnissen in den Vordergrund rückt, bleiben einige angesehene Investoren skeptisch und betrachten die Versprechen der KI als blossen Hype.

Warren Buffett, nebst seinem langjährigen Geschäftspartner bei Berkshire Hathaway, dem kürzlich verstorbenen Charles Munger eine der führenden Quellen für Investitionsweisheiten, stellt künstliche Intelligenz der "altmodischen" Intelligenz gegenüber. Während andere zunehmend von den Fähigkeiten der KI, ihren rasanten Fortschritten und ihrer Integration in fast allen Unternehmensbereichen fasziniert sind, betont Buffett den dauerhaften Wert des menschlichen Intellekts. Kann die KI, so fragt er, die Tiefe und Nuancen des menschlichen Denkens und der menschlichen Kreativität wirklich übertreffen?

Wie während der industriellen Revolution ist auch die gegenwärtige Ära der KI durch raschen technologischen Fortschritt und Optimismus hinsichtlich ihrer Auswirkungen auf unsere sozialen Strukturen, unsere Wirtschaft und unser tägliches Leben gekennzeichnet. Die Einführung mechanisierter Produktionsmethoden vor etwa 200 Jahren markierte einen bedeutenden Wendepunkt in der Geschichte der Menschheit. KI stellt einen weiteren dar. Während wir die Fähigkeit von Maschinen erforschen, auf eine Art und Weise kreativ zu sein, zu denken und zu interagieren, von der wir bisher dachten, dass sie nur Menschen vorbehalten sei, müssen wir auch die ethischen, datenschutzrechtlichen und sicherheitstechnischen Bedenken berücksichtigen, die mit einer breiten Einführung von KI einhergehen werden.

In dieser Swiss Finance Institute Public Discussion Note möchte ich die richtige Balance zwischen Überschwang und Vorsicht finden. Durch die Gegenüberstellung der optimistischen Prognosen der grossen Technologieunternehmen, die KI als entscheidend für den Wettbewerb auf dem Markt ansehen, und den skeptischen Ausblicken von Investoren wie Buffett und Munger, werde ich versuchen, die Komplexität des heutigen KI-Phänomens zu entwirren. Ich werde zunächst die Technologie selbst erklären und das Potenzial der KI, die Wirtschaft neu zu gestalten, diskutieren und abschliessend auf ethische Überlegungen, gesellschaftliche Auswirkungen und die vor einer breiten Anwendung im Finanzwesen zu lösenden Herausforderungen eingehen.

Eine kurze Geschichte der KI

Frühe Visionen (vor 1950):

Seit dem Altertum haben sich Zivilisationen Formen künstlichen Lebens ausgedacht, wie zum Beispiel in den Mythen von Pygmalion und dem Golem. In der Geschichte von Pygmalion verliebt sich ein Bildhauer in eine von ihm aus Stein gemeisselte Statue, die daraufhin zum Leben erwacht. Der Golem ist eine Kreatur aus Ton, die ebenfalls durch magische Rituale zum Leben erweckt wird, um ihrem Schöpfer zu dienen. Darüber hinaus schufen Erfinderinnen und Erfinder in der Antike und im Mittelalter in vielen verschiedenen Kulturen mechanische Vorrichtungen, die menschliche oder tierische Handlungen nachahmten. Auch diese Automaten spiegeln frühe Visionen von künstlicher Intelligenz wider.

Philosophen wie René Descartes (geb. 1596) und Gottfried Leibniz (geb. 1646) spielten eine entscheidende Rolle bei der Erforschung des Geistes, der Intelligenz und des Potenzials von Maschinen, das menschliche Denken nachzuahmen. Descartes hinterfragte die Natur des Denkens und des Bewusstseins, während Leibniz eine universelle Sprache des Denkens, den so genannten *calculus ratiocinator* entwickelte. Dieser frühe konzeptionelle Rahmen, der logische Aussagen und Schlussfolgerungen in einer präzisen, mathematischen Form darstellt, hat die moderne symbolische Logik und Berechnungen massgeblich beeinflusst.

Die Geburtsstunde des Fachgebietes (1950er Jahre):

Alan Turings Arbeiten in den 1940er und 1950er Jahren, insbesondere die Entwicklung des Turing-Tests, schuf die Grundlagen für die Bewertung der Intelligenz von Maschinen. Der Turing-Test ist eine Methode zur Feststellung, ob eine Maschine ein vom Menschen nicht zu unterscheidendes intelligentes Verhalten zeigen kann. Turings frühere Arbeiten legten auch die Grundsteine, auf denen die moderne Computertechnik beruht. Die Turing-Maschine ist zum Beispiel ein theoretisches Modell, das Symbole auf einem unendlichen Band gemäss einer Reihe von Übergangsregeln manipuliert. Es bietet ein einfaches, aber leistungsfähiges Modell zur Untersuchung, was es bedeutet, dass eine Funktion berechenbar ist.

Auf der sogenannten Dartmouth-Konferenz (1956) trafen sich Forschende, die von der Möglichkeit maschineller Intelligenz überzeugt waren, und legten den Grundstein für die künftige Forschung. Die Dartmouth-Konferenz wird oft als die formale Geburtsstunde der KI als wissenschaftliche Disziplin bezeichnet.

Die Achterbahn des Fortschritts (1960er-1990er Jahre):

Da der Hype um die KI grösser war als der tatsächliche Fortschritt, gab es Zeiten, in denen die Finanzierung und das Interesse an diesem Bereich zurückgingen. Diese so genannten "KI-Winter" fanden vor allem in den späten 1970er bis frühen 1980er Jahren und erneut in den späten 1980er bis frühen 1990er Jahren statt. Sie waren geprägt von Skepsis gegenüber dem Potenzial der KI, ihre Versprechen zu erfüllen, was zu geringeren Investitionen und einer Neubewertung der Ziele führte.

Trotz dieser Herausforderungen war die Entwicklung von Expertensystemen ein grosser Erfolg für die KI in kommerziellen und industriellen Anwendungen. Diese Systeme waren in der Lage, die Entscheidungen menschlicher Expertinnen und Experten in bestimmten Bereichen wie der medizinischen Diagnose und der Mineralienexploration zu imitieren. Im Finanzwesen wurden Expertensysteme zur Schätzung von Ausfallverlusten eingesetzt, wodurch Banken und Finanzinstitute über ausgefeilte Instrumente zur Bewertung des Kreditrisikos und zur Vorhersage der Ausfallwahrscheinlichkeit von Kreditnehmenden verfügten.

Aufschwung und Wandel (1990er-Jahre-Gegenwart):

Das Wiederaufleben der KI wurde massgeblich durch Fortschritte beim maschinellen Lernen und durch die Verfügbarkeit grosser Datensätze, die als "Big Data" bekannt sind, vorangetrieben. Diese Entwicklungen haben es KI-Systemen ermöglicht, aus Beispielen zu lernen, anstatt sich ausschliesslich auf fest kodierte Regeln zu verlassen.

Deep Learning, eine Art des maschinellen Lernens unter Verwendung neuronaler Netze, hat zu bedeutenden Durchbrüchen in den Bereichen Computer Vision, Spracherkennung und Verarbeitung natürlicher Sprache geführt und Anwendungen wie Sprachübersetzung in Echtzeit, autonome Fahrzeuge und effektive Websuche möglich gemacht.

Motivation

Meine Faszination für die Möglichkeiten der KI wurde als junger Mensch in den frühen 1990er Jahren durch die Lektüre von Stanisław Lems Science-Fiction-Roman "So sprach Golem" geweckt. Die Erzählung dreht sich um eine Figur namens Golem, eine Maschine von beispiellosem Intellekt, die mit den führenden Forschenden der Welt tiefe Gespräche führt. Während mir das Konzept damals rein spekulativ erschien, spiegelt Lems Roman, der in den 1980er Jahren geschrieben wurde, auf unheimliche Weise die heutigen Interaktionen mit KI in Bereichen wie der Finanzwirtschaft wider. Lems Studie der Fähigkeiten von KI, ethischen Implikationen und ihres Potenzials, mit Menschen zu koexistieren, ist heute aktueller denn je, da Technologien wie ChatGPT von OpenAI und Gemini von Google in unser tägliches Leben integriert werden.

Heute bin ich aktiv an den technologischen Entwicklungen beteiligt, die Lem einst vorschwebten. Die Reflexion über Lems Roman unterstreicht, wie wichtig es ist, die wachsende Rolle der KI in unserer Branche mit Bedacht und Weitsicht zu steuern. Er erinnert mich an das heikle Gleichgewicht zwischen der Nutzung technologischer Fortschritte und der Berücksichtigung ihrer breiteren Auswirkungen. Diese Überlegung wird besonders deutlich, wenn wir die transformativen Auswirkungen der KI in der Medizin, Biologie und Meteorologie beobachten.

So hat beispielsweise AlphaFold, entwickelt von Google DeepMind, die Biologie revolutioniert, indem es die 3D-Strukturen von Proteinen allein auf der Grundlage ihrer Aminosäuresequenzen genau vorhersagt (Jumper et al., 2021). Dieser Durchbruch löst das seit langem bestehende Problem der Proteinfaltung, beschleunigt die

Entdeckung von Medikamenten und verbessert unser Verständnis von biologischen Prozessen, indem er die Strukturen aller 200 Millionen der Wissenschaft bekannten Proteine in nur einem Jahr vorhersagte. Wenn man bedenkt, dass eine Doktorandin oder ein Doktorand grob geschätzt vier bis fünf Jahre braucht, um eine einzige Proteinstruktur zu verstehen, hätte diese Aufgabe sonst Hunderte von Millionen Jahren (!) menschlicher Arbeit erfordert. Demis Hassabis und John M. Jumper, beide von Google DeepMind, erhielten im Jahr 2024 den Nobelpreis für Chemie für ihre bahnbrechenden Beiträge zur Vorhersage der Proteinfaltung. In ähnlicher Weise hat GraphCast, ein weiteres innovatives Modell von Google DeepMind, die Wettervorhersage revolutioniert, indem es herkömmliche Methoden mit bemerkenswerter Genauigkeit übertroffen hat (Cookson, 2023; Lam et al., 2023). Durch die Analyse umfangreicher historischer Daten mit einem Graph-Neuronalen Netzwerk liefert GraphCast überlegene Vorhersagen von Wetterbedingungen, von der Temperatur über Windmuster bis hin zum Landfall von Hurrikanen. Ausserdem wird der Energieverbrauch erheblich gesenkt, indem die erforderliche Rechenzeit von vielen Stunden auf eine Minute reduziert wird.

Könnten wir einen vergleichbaren AlphaFold- oder GraphCast-Moment im Finanzsektor erleben? Generative KI hat das Potenzial, Finanzanalysen, Risikobewertungen und Wirtschaftsprognosen grundlegend zu verändern, da sie Muster aufdeckt und Erkenntnisse aus riesigen Datensätzen zieht. Leistungsstarke KI-Tools könnten unsere strategischen Entscheidungsprozesse neu definieren. Es ist wichtig, dass die Finanzbranche bereit ist und sich flexibel anpasst, um aussagekräftige und verantwortungsvolle Ergebnisse zu erzielen.



Generative KI im Finanzsektor

Bevor wir uns mit den Auswirkungen auf den Finanzsektor befassen, möchte ich erklären, was ich unter generativer KI verstehe. KI und generative KI sind verwandte, aber unterschiedliche Konzepte; während alle generative KI eine KI ist, ist nicht alle KI generativ.

KI ist ein weit gefasster Begriff. Er umfasst alle Computersysteme, die menschliche kognitive Funktionen imitieren sollen. Diese Systeme können entweder als enge oder schwache KI kategorisiert werden, die nur eine bestimmte Aufgabe erfüllen, für die früher menschliche Intelligenz erforderlich war, oder als allgemeine oder starke KI (Englisch: artificial general intelligence oder AGI), die Intelligenz in einem breiten Spektrum von Aufgaben einsetzen kann, wie es der Mensch tut.

KI-Systeme wie ChatGPT, die derzeit im Rampenlicht der Öffentlichkeit stehen, sind Beispiele für schwache KI. Sie sind hoch entwickelt und in der Lage, menschenähnliche Texte zu erzeugen, aber sie sind auf bestimmte Aufgaben spezialisiert und verfügen nicht über die weitreichenden Fähigkeiten einer starken KI. Viele Forscher und Experten sind sich einig, dass eine umfassende starke KI hypothetisch bleibt, da noch kein System die vereinbarten Kriterien erfüllt hat. Einige glauben jedoch, dass die Verwirklichung von starker KI eine Frage des *Wann* und nicht *des Ob* ist (Fry, 2022).

Generative KI bezieht sich auf ein KI-System – ob schwach oder stark –, das neue Inhalte oder Daten generieren kann, die den Inhalten und Daten, mit denen es trainiert wurde, ähnlich, aber nicht mit ihnen identisch sind. Beim Training eines Algorithmus wird dieser mit einem grossen Datensatz konfrontiert und muss eine Reihe von Aufgaben lösen, die durch einen Belohnungsmechanismus gesteuert werden, der Rückmeldung über seine Leistung gibt. Dieser Prozess ermöglicht es dem Algorithmus, sein Verständnis von Mustern und Entscheidungskriterien zu verfeinern und so seine Fähigkeit zu verbessern, ähnliche Aufgaben in Zukunft zu lösen. Der Datensatz kann Text, Bilder, Musik, Sprache und andere Medien enthalten. Generative KI-Modelle sind darauf ausgelegt, die zugrunde liegenden Muster und Strukturen dieser Eingabedaten zu verstehen, so dass sie auf der Grundlage des Gelernten neue, originelle Ergebnisse produzieren können.

Das entscheidende Merkmal der generativen KI ist ihre Fähigkeit, etwas zu erschaffen, was ein erhebliches Potenzial für die Erstellung von Inhalten, für Design und sogar für die Entdeckung von neuen Medikamenten bietet. Die kognitiven und kreativen Fähigkeiten der generativen KI sind Gegenstand einer anhaltenden philosophischen Debatte (Millière & Buckner, 2024). Sie hat jedoch das Potenzial, nicht nur menschliche Aufgaben zu replizieren, sondern

auch die menschliche Kreativität zu steigern, indem sie neue Ideen, Entwürfe und Perspektiven generiert und damit innovative Anwendungen bietet, die über die derzeitigen Analyse- und Entscheidungsfähigkeiten der KI hinausgehen.

Auswirkungen der generativen KI im Finanzsektor

Generative KI hat bereits erhebliche Auswirkungen auf den Finanzsektor und wirkt sich auf die Produktivität, die betriebliche Effizienz, das Risikomanagement und die Schaffung neuer Geschäftsmodelle aus. Einem McKinsey-Bericht zufolge könnte generative KI dem Bankensektor jährlich zwischen 200 und 340 Milliarden USD einbringen, was 9 bis 15 Prozent des operativen Ergebnisses des Bankensektors entspricht (McKinsey & Company, 2023). Dieser Wert ergibt sich vor allem aus der gesteigerten Produktivität, wobei die grössten absoluten Gewinne im Bereich des Firmen- und Privatkundengeschäfts erwartet werden.

Durch generative KI können Finanzinstitute ihre Serviceleistungen verbessern und das Kundenerlebnis steigern. Laut McKinsey & Company (2023) verwendet Goldman Sachs beispielsweise ein KI-basiertes Tool zur Automatisierung der zuvor arbeitsintensiven Testerstellung. Die Citigroup hat generative KI eingesetzt, um die Auswirkungen der neuen US-Kapitalvorschriften zu bewerten.

Maufe und Brown (2023) nennen fünf praktische Anwendungen für generative KI in der Finanzdienstleistungsbranche, die alle die Produktivität der Mitarbeitenden steigern. Dazu gehören die Suche und Synthese von Finanzdokumenten, verbesserte virtuelle Assistenzprogramme, Kapitalmarktforschung, Softwareentwicklung und personalisierte Finanzempfehlungen.

Der Internationale Währungsfonds hat jedoch darauf hingewiesen, dass der Einsatz von generativer KI im Finanzsektor mehrere Risiken birgt (International Monetary Fund, 2023). Dazu gehören eingebettete Verzerrungen, Datenschutzmängel, Undurchsichtigkeit im Entscheidungsprozess, Probleme mit der Robustheit, Bedrohungen der Cybersicherheit und mögliche Auswirkungen auf die Finanzstabilität im Allgemeinen. Der Internationale Währungsfonds warnt, dass ein übermässiges Vertrauen in generative KI das Risiko des Übergreifens von Finanzrisiken erhöhen und neue systemische Risiken im Finanzsektor schaffen könnte.

Trotz dieser Bedenken erforscht der Finanzsektor aktiv die Möglichkeiten der generativen KI. Der Wettbewerbsdruck treibt die rasche Einführung von Anwendungen voran, die Effizienzgewinne, Kosteneinsparungen und ein besseres Risikomanagement versprechen.

Die generative KI-Technologie bietet eine Reihe von Einsatzmöglichkeiten im Finanzbereich, darunter:

- **Ausarbeitung von Prognosen und Budgets.** Eine KPMG-Umfrage ergab, dass 83% der Umfrageteilnehmenden KI in der Finanzplanung einsetzen (KPMG, 2024), vor allem für zukunftsorientierte Datenanalysen, die Erstellung von Szenarien und Einblicke in die Budgetplanung.
- **Erstellung von Finanzberichten und Präsentationen.** Generative KI reduziert den Zeit- und Arbeitsaufwand für die Erstellung regelmässiger Finanzberichte erheblich, indem Vorlagen automatisch mit neuen Daten gefüllt werden und Erkenntnisse aus Finanz- und Geschäftsdaten abgeleitet werden.
- **Extrahieren von Marktinformationen.** Mit den umfassenden Sprachmodellen der generativen KI können Unternehmen öffentliche Datensätze nutzen, um Markteinblicke zu gewinnen, Wettbewerbsinformationen zu erstellen und Kundentrends durch massgeschneiderte Analysen aufzudecken, die entweder regional-spezifisch oder personenspezifisch sein können.
- **Strategische Erkenntnisse aus Daten ableiten.** Durch den Einsatz generativer KI zur Analyse von Daten, die für das Kundenbeziehungsmanagement (Englisch: customer relationship management oder CRM) oder die Unternehmensressourcenplanung (Englisch: enterprise resource planning oder ERP) relevant sind, können Unternehmen wertvolle strategische Erkenntnisse gewinnen, um Finanzentscheidungen zu treffen, Herausforderungen bei der Preisgestaltung und Leistung zu bewältigen oder um andere Geschäftsabläufe zu verbessern.
- **Automatisierung der Vertragsverwaltung.** Generative KI vereinfacht die Vertragserstellung und konzentriert sich auf nicht standardisierte Klauseln, die Identifizierung umsatzbezogener Klauseln und die Dokumentation der buchhalterischen Behandlung.
- **Verbesserung der Erkennung von Anomalien und der Betrugsprävention.** Durch die Identifizierung von Datenunregelmässigkeiten und die Überwachung laufender Transaktionen bietet generative KI einen robusten Mechanismus zur Erkennung von Fehlern und zur Betrugsprävention, wodurch die finanzielle Integrität eines Unternehmens geschützt wird.

- **Personalisierte Anlageberatung.** Generative KI kann Anlageempfehlungen auf das Risikoprofil und die finanziellen Ziele einer Person zuschneiden und so die Kundenerfahrung verbessern.
- **Aufdeckung von Betrug.** Durch die Analyse grosser Transaktionsdatensätze kann die generative KI betrügerische Muster erkennen und so die finanzielle Sicherheit verbessern.
- **Marktanalyse und Risikomanagement.** Ebenso kann die generative KI, sofern entsprechende Datensätzen vorhanden sind, Markttrends erkennen und potenzielle Risiken bewerten, um Investitionsentscheidungen zu unterstützen.

Diese breite Palette von Anwendungen im Finanzbereich wirft mehrere wichtige Diskussionspunkte auf. An erster Stelle stehen ethische Überlegungen, insbesondere die Frage, wie Transparenz aufrechterhalten werden kann. Es werden neue Standards und Protokolle benötigt, um sicherzustellen, dass KI-generierte Finanzinhalte nicht zur Verbreitung von Verzerrungen oder zu Fehlinformationen führen. Der zweite Punkt sind die Auswirkungen der Automatisierung auf die Mitarbeitenden im Finanzsektor. Da KI traditionelle Rollen verdrängt, müssen Mitarbeitende neue Fähigkeiten erlernen und auf notwendige Anpassungen vorbereitet sein. Der dritte Punkt ist das regulatorische Umfeld. Durch Fortschritte bei generativen KI-Technologien könnten finanzielle Regulierungen unzureichend oder veraltet werden, und die regulatorischen Rahmenbedingungen müssen angepasst werden, um sowohl Innovation zu ermöglichen als auch den Schutz der Verbrauchenden zu gewährleisten. Der letzte Punkt ist die fortlaufende Entwicklung innovativer Produkte und Dienstleistungen, um neuen Marktanforderungen gerecht zu werden und Finanzinstituten einen Wettbewerbsvorteil zu verschaffen.

Praktische Anwendungen

Lassen Sie uns einige praktische Anwendungen generativer KI betrachten und erkunden, wie diese Innovation im Finanzsektor vorantreiben. Diese Beispiele aus Umwelt-, Sozial- und Governance-Initiativen (Englisch: Environmental, Social, and Governance oder ESG) zeigen, wie generative KI einen konkreten Bereich des Finanzwesens verändern kann, und sie verdeutlichen das Potenzial von KI, die gesamte Branche zu revolutionieren.

Fallstudie 1: Identifizierung grüner Innovationen

In einem kürzlich erschienenen Artikel haben mein Koautorinnen und ich untersucht, wie generative KI eine wichtige Rolle bei der Entschlüsselung der technischen Sprache grüner Patente spielen kann (Leippold und Yu, 2024). Die technischen Feinheiten dieser Patente zu verstehen kann Personen, welche die Nachhaltigkeitsinitiativen eines Unternehmens bewerten möchten, vor grosse Herausforderungen stellen. Generative KI-Modelle meistern diese Herausforderung, indem sie komplexe Beschreibungen und Fachterminologie in prägnante Zusammenfassungen herunterbrechen. Die KI "liest" bei diesem Prozess das Patent systematisch, interpretiert seine entscheidenden Elemente und artikuliert diese in einfachen Begriffen neu. Diese Anwendung der generativen KI erfüllt mehrere Funktionen, die für den Finanzsektor entscheidend sind.

Erstens ermöglicht die generative KI hier einen breiteren Zugang zu den Details grüner Innovationen, indem sie Fachjargon in eine einfache Sprache übersetzt. Sie macht die Informationen des Patents einem breiten Publikum zugänglich, darunter Personen in der Finanzanalyse, Investierende und politischen Entscheidungstragende, die möglicherweise nicht über das zum Verständnis erforderliche technische Fachwissen verfügen. Diese Demokratisierung von Informationen ist wichtig, da sie einem breiteren Spektrum von Interessengruppen ermöglicht, fundierte Entscheidungen zu treffen.

Zweitens verbessern die von der KI bereitgestellten Zusammenfassungen die analytische Effizienz. Personen in der Finanzanalyse arbeiten in einem Umfeld, in dem Schnelligkeit und Genauigkeit von grösster Bedeutung sind. Sie profitieren immens davon, dass sie, ohne sich in Details zu verlieren, die Bedeutung eines Patents schnell erfassen können. Hier spart KI nicht nur Zeit, sondern verbessert auch die Qualität der Analyse. Sie stellt sicher, dass finanzielle Entscheidungen auf einem gründlichen Verständnis der potenziellen Auswirkungen und des Werts der Innovationen eines Unternehmens beruhen.

Drittens erleichtern die KI-Zusammenfassungen die marktweite Analyse. Personen in der Finanzanalyse können die Nachhaltigkeitsbemühungen verschiedener Unternehmen, Sektoren und Branchen leichter vergleichen, wenn sie ein klares Verständnis der wichtigsten

Innovationen der einzelnen Patente haben. Diese Möglichkeit ist für potenzielle Investierende von unschätzbarem Wert, da ihre Entscheidungen so auf einem umfassenden Verständnis des Bekenntnisses zu grüner Innovation verschiedener Unternehmen beruhen.

Nachdem wir generative KI eingesetzt hatten, um komplexe grüne Patenttexte in klare, prägnante Zusammenfassungen zu destillieren, erstellten meine Koautorinnen, Koautoren und ich ein BERT-basiertes Sprachmodell namens ClimateBERT, um unser Verständnis weiter zu verfeinern (Webersinke et al., 2022). BERT steht für Bidirectional Encoder Representations from Transformers. BERT wurde von Google entwickelt und versteht den Kontext eines Wortes besser, indem es sowohl die vorangehenden als auch die nachfolgenden Wörter im Satz berücksichtigt. Wir haben ClimateBERT darauf trainiert, die Nuancen der Nachhaltigkeitsterminologie zu erkennen, was eine tiefgreifende kontextuelle Analyse der von der generativen KI erzeugten vereinfachten Texte ermöglicht.

Durch die Analyse dieser Texte im Zeitverlauf kann ClimateBERT Trends im Nachhaltigkeitsdiskurs eines Unternehmens erkennen. Diese Art der Analyse hebt Verschiebungen in der Betonung grüner Innovationen im Unternehmen hervor und liefert frühzeitig Signale für strategische Veränderungen, die sich auf das langfristige Engagement für Nachhaltigkeit sowie auf die Finanzergebnisse auswirken könnten. Durch die Kombination von generativer KI und ClimateBERT erreichen wir so ein tieferes Verständnis dafür, wie sich das Unternehmen mit grüner Innovation auseinandersetzt.

Um schliesslich die breiteren Auswirkungen grüner Innovationen zu verstehen, erstellen wir mit unseren Daten Kennzahlen, die zeigen, wie Nachhaltigkeitsinitiativen die marktbereinigten kumulativen abnormalen Renditen bei wichtigen Ereignissen wie Trumps Wahlsieg 2016, Bidens Sieg 2020, dem Ausbruch des Russland-Ukraine-Kriegs und der Ankündigung des Inflation Reduction Act beeinflusst haben. Unsere Arbeit bietet Personen in der Finanzanalyse und Anlegenden ein leistungsfähiges Instrument, um Risiken im Zusammenhang mit grünen Innovationen und Nachhaltigkeit besser zu verstehen und zu mindern.

Fallstudie 2: Analyse der unternehmerischen Nachhaltigkeit

Meine Koautorinnen, Koautoren und ich haben ein KI-Tool namens CHATREPORT (Ni et al., 2023) entwickelt, das auf unserer früheren Arbeit an CHATCLIMATE (Vaghefi et al., 2023) basiert. CHATREPORT verarbeitet die komplexen und umfangreichen Daten aus den Nachhaltigkeitsberichten von Unternehmen. In diesen Berichten werden die Umweltauswirkungen, die Governance-Strukturen und die sozialen Initiativen eines Unternehmens detailliert beschrieben. Sie dienen als wichtiges Instrument, durch das Investierende,

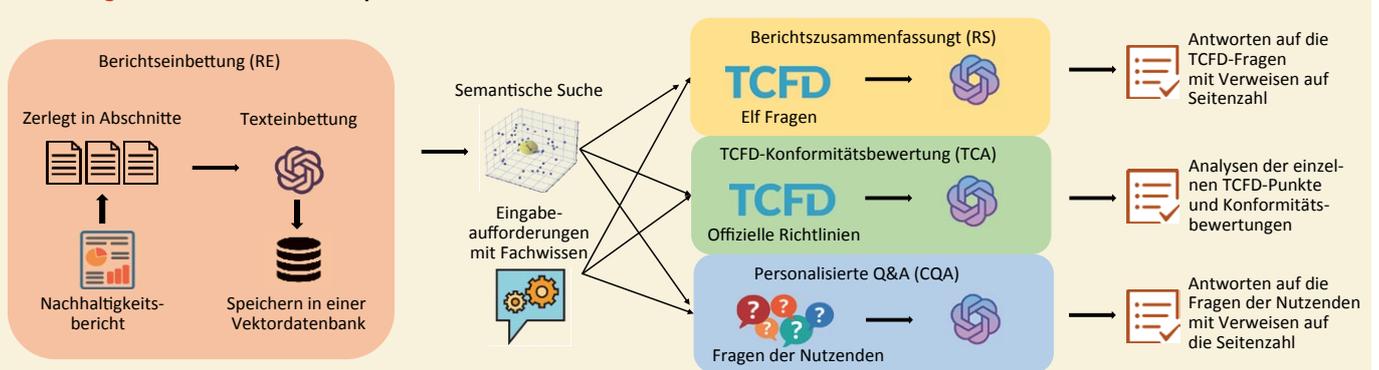
politische Entscheidungstragende und die breite Öffentlichkeit die Verpflichtung eines Unternehmens zu nachhaltigen Praktiken beurteilen können. Ihre Komplexität und die schiere Menge an Daten, die diese Berichte enthalten, machen die manuelle Analyse jedoch zu einer fast unmöglichen Aufgabe. Die Herausforderung wird noch grösser, wenn sich die Analyse an bestimmte gesetzliche Rahmenbedingungen oder Richtlinien halten muss. Bisher verfügten nur wenige Institutionen über die Ressourcen, die für eine gründliche, detaillierte Analyse dieser Berichte erforderlich sind, was ein erhebliches Hindernis für Transparenz darstellt. CHATREPORT demokratisiert den Prozess, indem es einer Vielzahl von Interessengruppen die Möglichkeit gibt, sich mit diesen Berichten zu befassen und darauf zu reagieren.

Das Herzstück von CHATREPORT ist eine ausgeklügelte Pipeline (siehe Abbildung 1), die die komplexen Daten der Nachhaltigkeitsberichte von Unternehmen in handlungsrelevante Erkenntnisse umwandelt. Ein entscheidender Aspekt unserer Methodik ist die Integration von Fachwissen direkt in die KI-Entwicklungsschleife. Der Prozess beginnt mit der Extraktion des Rohtextes eines Berichts, der Normalisierung der Datenformate und der Entfernung aller irrelevanten Informationen. Um den Inhalt zu verstehen und zu analysieren, verwendet die Pipeline fortschrittliche grosse Sprachmodelle (Englisch: large language models oder LLMs), die mit bereichsspezifischem Wissen abgestimmt werden. Benutzende und andere menschliche Expertinnen und Experten erstellen und

verfeinern die analytischen Eingabeaufforderungen, um sicherzustellen, dass die Ergebnisse der KI technisch korrekt und kontextbezogen relevant sind. Das Tool identifiziert Schlüsselthemen, bewertet die Einhaltung von Richtlinien wie denen der Task Force on Climate-related Financial Disclosures (TCFD) und generiert Erkenntnisse in einem benutzerfreundlichen Format. Jeder Schritt in der CHATREPORT-Pipeline soll sicherstellen, dass die Daten mit hoher Präzision verarbeitet werden und den neuesten Nachhaltigkeitsstandards und -praktiken angeglichen sind. Dies führt zu einem leistungsstarken Instrument für Stakeholder, die fundierte Entscheidungen treffen müssen.

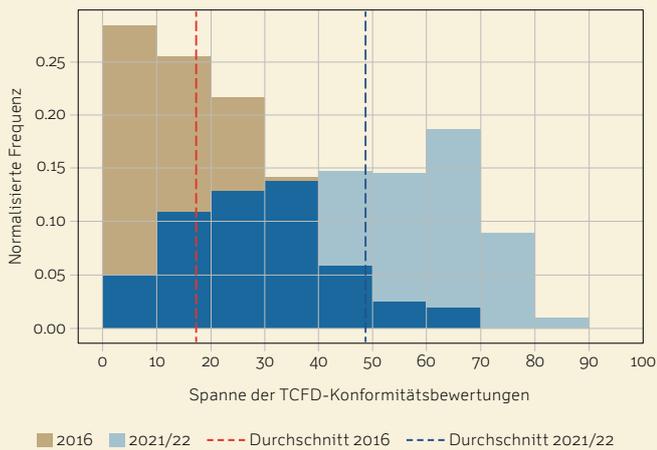
Eine häufige Herausforderung bei KI-generierten Inhalten ist, wie weiter unten erläutert, das Risiko von "Halluzinationen", bei denen die KI falsche oder irreführende Informationen produziert, die nicht durch Daten gestützt werden. CHATREPORT begegnet diesem Problem durch seine Rückverfolgbarkeitsfunktionen. Jeder von der KI generierte Inhalt lässt sich direkt zu bestimmten Abschnitten der Quelldokumente zurückverfolgen, so dass die Benutzenden die Richtigkeit der Informationen manuell überprüfen können. Darüber hinaus ist CHATREPORT in hohem Masse benutzerorientiert und ermöglicht es dem Benutzer, massgeschneiderte Analysen zu erstellen. Über ein dynamisches Frage-Antwort-Modul können Nutzende die KI zu bestimmten Aspekten eines Berichts befragen. Die KI ruft dann die relevanten Daten ab, verarbeitet sie und liefert massgeschneiderte Erkenntnisse.

Abbildung 1: Die CHATREPORT-Pipeline



Hinweis: Die Pipeline für CHATREPORT automatisiert die Analyse von Nachhaltigkeitsberichten von Unternehmen durch vier Hauptmodule. Das Modul Report Embedding (RE) zerlegt den Bericht in Textabschnitte und bettet sie in einen Vektorraum für die semantische Suche auf der Grundlage der TCFD-Empfehlungen ein. Das Modul Report Summarization (RS) ruft relevante Berichtsabschnitte ab und veranlasst ein grosses Sprachmodell (LLM), die Angaben gemäss den elf empfohlenen Aspekten der TCFD zusammenzufassen. Das Modul TCFD Conformity Assessment (TCA) bewertet die Einhaltung der TCFD-Richtlinien durch Analyse des Berichts und generiert Konformitätsbewertungen und Analyseabschnitte. Das Modul Customized Question Answering (CQA) ermöglicht es Nutzenden, spezifische Fragen zu stellen, auf die das LLM auf die entsprechenden Abschnitte des Berichts verweisende Antworten liefert. Diese umfassende Pipeline erhöht die Transparenz und Effizienz der Analyse von Nachhaltigkeitsberichten.

Quelle: Ni et al. (2023)

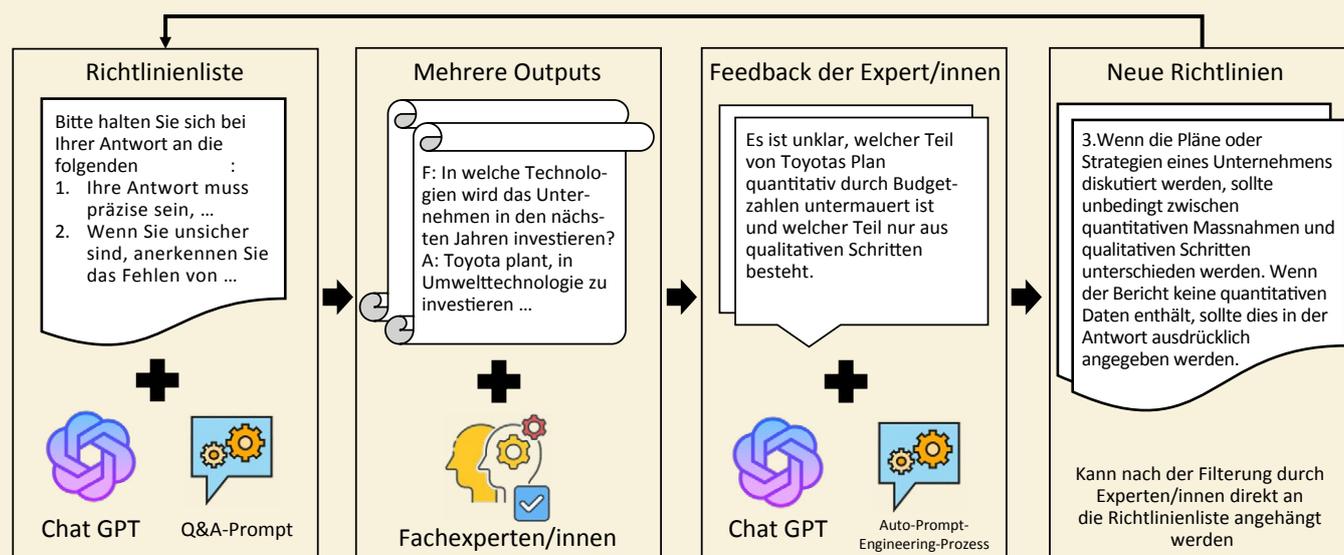
Abbildung 2: Dichteplot der TCFD-Konformitätsbewertungen

Hinweis: Diese Abbildung zeigt die Verteilung der TCFD-Konformitätsbewertungen für eine Stichprobe von Nachhaltigkeitsberichten von an der NYSE kotierten Unternehmen für die Berichtsjahre 2016 und 2021/22. Die TCFD-Konformitätswerte werden danach berechnet, wie gut die Berichte die TCFD-Richtlinien einhalten. Das Dichte-Diagramm zeigt eine deutliche Verbesserung der durchschnittlichen Konformitätswerte zwischen den Jahren 2016 und 2021/22. Diese deutet darauf hin, dass Unternehmen ihre Nachhaltigkeitsangaben zunehmend an den TCFD-Empfehlungen ausrichten.
Quelle: Ni et al. (2023)

In unserer Studie haben wir zum Beispiel über 10'000 Nachhaltigkeitsberichte von Unternehmen auf ihre Konformität mit den TCFD-Berichtsanforderungen geprüft. Abbildung 2 zeigt die Ergebnisse für zwei Berichtszeiträume: vor der Einführung der TCFD (braun) und (einige Jahre) nach der Einführung der TCFD (hellblau). Unsere Daten zeigen, dass das Niveau der TCFD-Konformität zwischen den beiden Zeiträumen erheblich gestiegen ist. Vor der offiziellen Einführung der TCFD, im Jahr 2016, lag der durchschnittliche Konformitätswert bei 17%. Im Jahr 2021/22 lag er bei 49%. Dieser Trend unterstreicht die wachsende Bedeutung von transparenter und umfassender Offenlegung klimabezogener Finanzinformationen. In weiteren Analysen haben wir Trends und Ausreisser identifiziert, die eine Diskussion über die Effektivität der aktuellen Praktiken und über potenzielle Verbesserungsbereiche ermöglichen.



Abbildung 3: Automatische Prompt-Engineering-Pipeline



Hinweis: CHATREPORT bezieht das Feedback von Nutzenden und anderen Domänenexpertinnen und -experten ein, um die vom grossen Sprachmodell (LLM) verwendeten Prompts zu verfeinern und zu verbessern. Zunächst geben die Domänenexpertinnen und -experten Feedback zu spezifischen Antworten, die vom LLM generiert wurden. Dieses Feedback wird dann durch einen automatisierten Prompt-Engineering-Prozess in allgemeine Richtlinien umgewandelt. Die KI integriert die neuen Richtlinien in ihre Vorlagen und verbessert so die Qualität und Präzision ihrer zukünftigen Ausgaben. Dieser iterative Prozess führt zu einer genaueren, detaillierteren und kritischeren Analyse von Nachhaltigkeitsberichten.

Quelle: Ni et al. (2023)

Schliesslich, und wie in Abbildung 3 gezeigt, ist CHATREPORT darauf ausgelegt, mit der Zeit zu lernen und sich anzupassen. Das System bezieht kontinuierlich Rückmeldungen von Nutzenden und anderen menschlichen Expertinnen und Experten ein, um seine Analysemodelle und Eingabeaufforderungen zu verfeinern. Dieser iterative Prozess stellt sicher, dass das System in Bezug auf Genauigkeit, Zuverlässigkeit und Relevanz ständig verbessert wird. Um die Transparenz zu betonen, haben meine Koautorinnen, Koautoren und ich CHATREPORT als Open-Source-Projekt entwickelt. Es kann unter <https://reports.chatclimate.ai> eingesehen werden. Indem wir unsere Methoden, Datensätze und Tools öffentlich zugänglich machen, fördern wir die Zusammenarbeit und laden Entwickelnde und Forschende aus der ganzen Welt ein, zur Weiterentwicklung der nachhaltigen Finanzanalyse beizutragen.

Die Stärke von CHATREPORT liegt in seiner Fähigkeit, den Analyseprozess drastisch zu beschleunigen. Mit diesem KI-Tool können relevante Informationen aus Tausenden von Nachhaltigkeitsberichten von Unternehmen – bis zu 10'000 – in nur wenigen Stunden extrahiert werden. CHATREPORT verdichtet riesige Datenmengen zu prägnanten Zusammenfassungen und ermöglicht so eine schnellere, datengestützte Entscheidungsfindung in einem bisher unerreichten Umfang.

Ein Blick in die (sehr) nahe Zukunft

Generative KI könnte schon bald die Art und Weise drastisch verändern, wie Investitionen und Finanzentscheidungen analysiert werden. Kürzlich habe ich einen Ansatz für die automatische Überprüfung von Fakten getestet, der an die sokratische Methode erinnert. Im antiken Griechenland stellte Sokrates seinen Gesprächspartnern unerbittlich Fragen, um Widersprüche in ihren Gedanken aufzudecken und eine tiefere Erforschung der zugrunde liegenden Probleme zu fördern. Dieses philosophische Erbe spiegelt sich in dem von Leippold et al. (2024) zur Überprüfung von Klimabehauptungen verwendeten Mediator/in-Advokaten/innen-Rahmen wider.

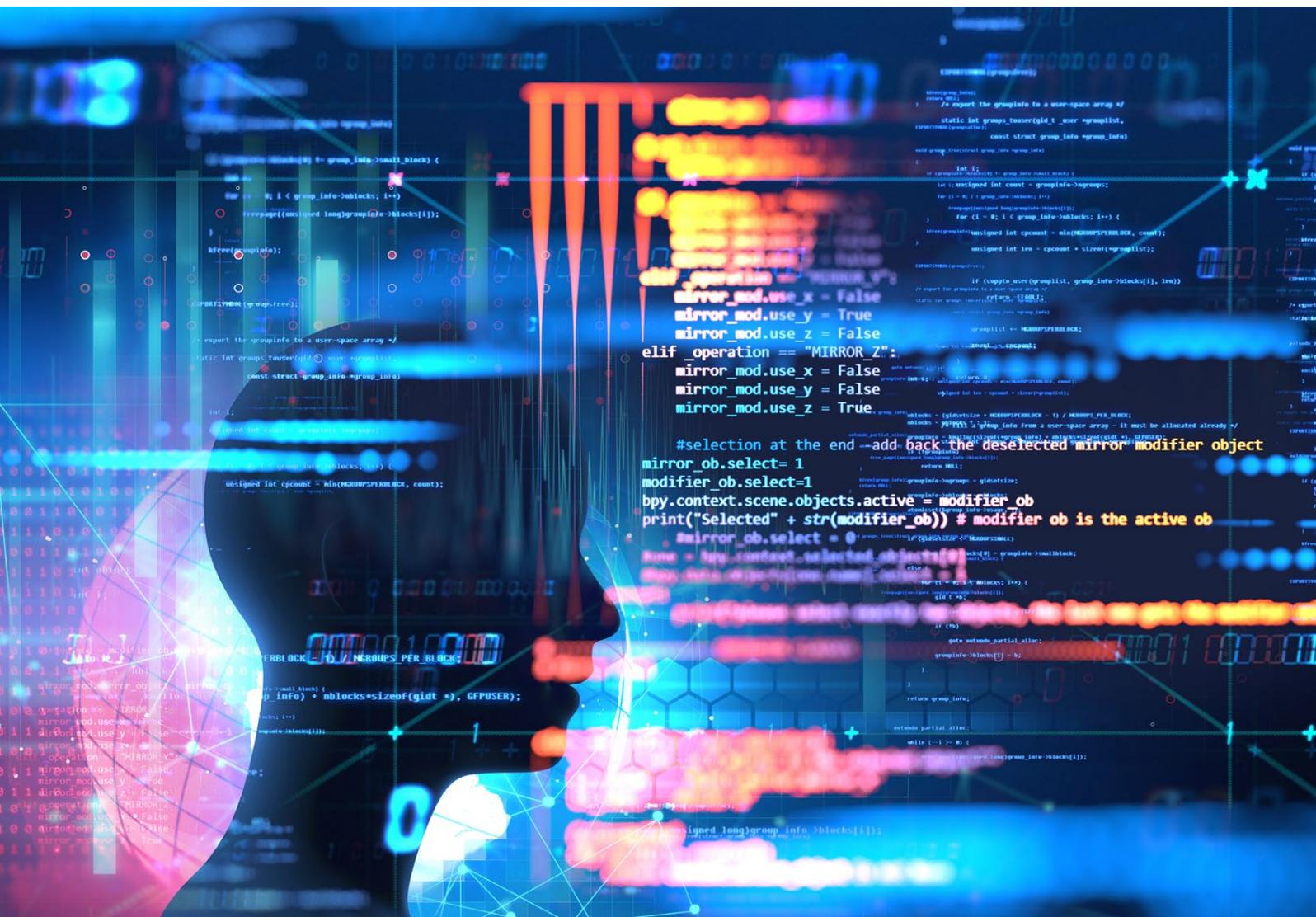
Überträgt man dieses Instrument auf den Finanzbereich, so hätten wir mehrere LLMs, die als "KI-Maschinenanalysten" bezeichnet werden, und die jeweils eine bestimmte Facette von Finanzdaten – wie Markttrends, regulatorische Auswirkungen, Wirtschaftsindikatoren oder Leistungskennzahlen von Unternehmen – analysieren und potenzielle Investitionsstrategien vorschlagen. Diese Argumente würden dann von einem "Chief AI-Investment Manager Maschine" zusammengefasst. Um eine kohärente Bewertung zu erstellen, würde dieser Manager den KI-Analysten

Folgefragen stellen, ähnlich wie Sokrates seine Schülerinnen und Schüler durch Dialektik zu einem tieferen Verständnis geführt hat. Dieser Prozess stellt sicher, dass viele verschiedene Perspektiven berücksichtigt werden, die sich gegenseitig in Frage stellen und ergänzen und letztlich zu einer solideren Schlussfolgerung führen.

Ein solches Instrument würde die Entscheidungsfindung im Finanzbereich verbessern, indem es die Rechengeschwindigkeit der KI nutzt und Verzerrungen zugunsten einer bestimmten Art von Finanzanalyse verringert. Da die Finanzmärkte immer komplexer werden, würde die Skalierbarkeit von KI-Systemen ausserdem sicherstellen, dass Entscheidungstragende den wachsenden Datenbedarf effizient bewältigen können.

Ein zentrales Merkmal eines solchen Systems wäre die Fähigkeit, den Verlauf der Argumente zwischen den KI-Analysten zu verfolgen und aufzuzeichnen und so einen Audit-Pfad für die Entscheidungsfindung zu liefern. Eine solche Nachvollziehbarkeit ist von entscheidender Bedeutung für die Stärkung des Vertrauens zwischen den Beteiligten und für die Erfüllung der regulatorischen Anforderungen der Finanzbranche. Diese verlangen eine gründliche Dokumentation der Entscheidungsprozesse, insbesondere in Bereichen wie Risikobewertung und Compliance.

Die Integration von KI-Systemen in das Finanzwesen verspricht zwar eine Verlagerung hin zu dynamischeren Entscheidungen in Echtzeit, doch dieses Potenzial ist mit vielen theoretischen und praktischen Herausforderungen verbunden, z. B. der Gewährleistung der Datengenauigkeit, der Wahrung der Quellenglaubwürdigkeit und der Abschwächung der Risiken algorithmischer Verzerrungen.



Derzeitige Grenzen der generativen KI

Trotz ihres Potenzials ist die generative KI-Technologie mit erheblichen Einschränkungen behaftet, die dringend behoben werden müssen. Bevor sie den Finanzsektor revolutionieren kann, müssen Probleme wie Halluzinationen, mangelnde Überprüfbarkeit und Anfälligkeit für Angriffe durch feindliche Eingaben behoben werden.

Halluzinationen und Ungenauigkeit

In Schimanski et al. (2024) analysieren meine Koautoren und ich kritisch die Genauigkeit und Überprüfbarkeit der von LLMs generierten Antworten. Bei komplexen Frage-Antwort Aufgaben und bei der Generierung von Inhalten ist das als "Halluzination" bekannte Phänomen häufig anzutreffen. Unter "Halluzination" versteht man in der künstlichen Intelligenz, dass das Modell Informationen generiert, die zwar plausibel sind, aber nichts mit dem tatsächlichen Quellenmaterial zu tun haben und gänzlich unbegründet sind. Dieses ist ein zentrales Problem in Sektoren wie dem Finanz-, Gesundheits- und Rechtswesen, wo die Genauigkeit und Zuverlässigkeit von Informationen von grösster Bedeutung sind. Unsere Studie zeigt, dass LLMs häufig Quellen falsch zuordnen oder Inhalte erfinden können, was das Vertrauen der Nutzer in solche automatisierten Systeme stark untergräbt.

Um diese Herausforderungen zu bewältigen, schlagen wir vor, den Prozess der Textextraktion aus den relevanten Quellen zu verfeinern und die Zuordnung von KI-generierten Inhalten zu verbessern. Solche technischen Korrekturen sind jedoch nur ein Teil der Lösung. Die Folgen unüberprüfter Ungenauigkeit sind enorm und können erhebliche Auswirkungen auf finanzielle Entscheidungen, die Versorgung von Patientinnen und Patienten sowie rechtliche Verfahren haben. Es besteht dringender Forschungsbedarf, um diese Modelle zu verfeinern und einen soliden ethischen Rahmen für ihre Verwendung zu entwickeln.

Mangelnde Überprüfbarkeit

Generative Suchmaschinen, darunter bekannte Plattformen wie *Bing Chat*, *NeevaAI*, *perplexity.ai* und *YouChat*, sind so konzipiert, dass sie auf die Anfragen von Nutzern mit Zitaten versehenen Antworten generieren. Liu et al. (2023) stellen jedoch fest, dass nur etwa 52% der von diesen Suchmaschinen erzeugten Sätze angemessen durch Zitate unterstützt werden. Diese Zahl ist alarmierend niedrig, wenn man bedenkt, wie wichtig zuverlässige Informationen im heutigen digitalen Zeitalter sind. Darüber hinaus zeigt die Studie, dass nur 75% der Zitate die Aussagen, die sie zu bestätigen scheinen, tatsächlich unterstützen. Diese Diskrepanz zwischen Schein und Wirklichkeit führt Nutzende in die Irre, die sich auf scheinbar glaubwürdige, aber im Grunde nicht belegte oder falsch zitierte Informationen verlassen.

Diese Forschung unterstreicht einen entscheidenden Aspekt der KI-Entwicklung: die Vertrauenswürdigkeit. Trotz der ausgefeilten Antworten dieser generativen Suchmaschinen wird ihre Zuverlässigkeit durch die Häufigkeit unbelegter Behauptungen und ungenauer Zitate erheblich beeinträchtigt. Das KI-Training muss verfeinert und es müssen Strategien entwickelt werden, um die Überprüfbarkeit der von diesen Systemen gelieferten Informationen zu verbessern.

Anfälligkeit für Angriffe durch feindliche Eingaben

Generative KI ist auch anfällig für feindliche Angriffe, die KI-Systeme manipulieren, um Daten absichtlich falsch zu interpretieren oder darzustellen. Diese Schwäche ist besonders kritisch bei einigen Arten von Finanzanalysen, bei denen solche Manipulationen das Analyseergebnis stark verfälschen können. In Leippold (2023) habe ich beispielsweise "Sentiment Spin"-Angriffe untersucht, bei denen ein komplexes Modell wie GPT-3 verwendet werden kann, um die Stimmung von Finanztexten zu verändern, indem es beispielsweise die negative Stimmung einer Reihe von Sätzen in eine neutrale oder positive Stimmung umwandelt. Solche Angriffe stellen ein grosses Risiko für die Zuverlässigkeit automatisierter Systeme dar.

Hu et al. (2024) beschreiben verschiedene andere Angriffsstrategien, die die Schwachstellen generativer Suchmaschinen ausnutzen. Ihre Forschung zeigt, wie generative Suchmaschinen trotz ihrer erweiterten Funktionalität durch gezielt gestaltete Anfragen zur Hervorrufung von Fehlern in die Irre geführt werden können. Diese Anfälligkeit unterstreicht die dringende Notwendigkeit einer Verbesserung der KI-Sicherheitsmassnahmen.

Die Auswirkungen feindlicher Angriffe beschränken sich nicht auf den Finanzsektor. KI-Systeme können auch so manipuliert werden, dass sie in anderen Sektoren wie dem Gesundheitswesen, der Politik und der Pharmazie verzerrte oder irreführende Ergebnisse liefern. Die Regulierung der generativen KI sollte sich auf die Schaffung von Rahmenbedingungen konzentrieren, die Transparenz, Verantwortlichkeit und Fairness gewährleisten. In dem Masse, in dem sich der Einsatz von KI ausweitet, müssen die Regulierungs-

behörden strenge Richtlinien einführen, um ethische Praktiken zu fördern und Missbrauch zu verhindern. Solche Regelungen werden Nutzende vor den Auswirkungen feindlicher Angriffe schützen und das Vertrauen in KI-Systeme fördern, was wiederum den Weg für ihre breitere Akzeptanz und Integration in die Gesellschaft ebnet wird.

Zwar gibt es bereits einige Rahmenwerke, wie z. B. das EU-Gesetz zur künstlichen Intelligenz, doch ist es noch ein weiter Weg bis zur Ausarbeitung umfassender und wirksamer Vorschriften. Die Formulierung solcher Vorschriften erfordert einen differenzierten Ansatz, der das rasche Tempo des technologischen Fortschritts und die vielfältigen und wachsenden Anwendungen von KI berücksichtigt. Es ist wichtig, einen ständigen Dialog zwischen Forschenden, führenden Branchenkennerinnen und -kennern – insbesondere aus dem Technologiesektor – und politischen Entscheidungstragenden zu führen.



Schlussfolgerung

Generative KI bietet sowohl beispiellose Chancen als auch grosse Herausforderungen. In dieser Public Discussion Note habe ich ihr transformatives Potenzial für die Finanzbranche skizziert, von der Verbesserung der Entscheidungsfindung und Risikobewertung bis hin zur Revolutionierung der Nachhaltigkeitspraktiken von Unternehmen. Der Weg zur vollständigen Integration von KI in den Finanzsektor ist jedoch voller Komplexitäten, darunter die Probleme der Halluzination, der Überprüfbarkeit und der Anfälligkeit für feindliche Angriffe, die das Vertrauen in und die Zuverlässigkeit von KI-Systemen untergraben könnten.

Die beträchtlichen Investitionen, die bereits in die KI getätigt wurden, machen deutlich, wie viel auf dem Spiel steht. Laut Goldman Sachs (2024) liegt das Potenzial für erhebliche Produktivitätssteigerungen und betriebliche Effizienzgewinne zwar auf der Hand, doch gibt es konträre Ansichten über die gesamtwirtschaftlichen Auswirkungen der KI. Expertinnen und Experten wie Daron Acemoglu vom MIT und andere äussern sich skeptisch und geben zu bedenken, dass

die Technologie die hohen Kosten nicht rechtfertigt, und dass die Kosten möglicherweise nicht wie erwartet sinken (Acemoglu, 2023). Auf der anderen Seite gibt es viele optimistische Stimmen, die glauben, dass die KI letztendlich beträchtliche Erträge und wirtschaftliche Vorteile bringen wird und dass ihre transformativsten Anwendungen erst noch entwickelt werden müssen.

Als Forscher auf diesem sich entwickelnden Gebiet sehe ich sowohl das Potenzial als auch die Herausforderungen, die für ihre Anwendung im Finanzwesen noch vor uns liegen. Die Zukunft der generativen KI im Finanzwesen ist nicht vorherbestimmt, sondern wird durch unser Handeln gestaltet werden. Wir haben die einmalige Chance zur Nutzung von KI, um Innovationen voranzutreiben, die Nachhaltigkeit zu erhöhen und die Effizienz im Finanzsektor zu verbessern. Das Erreichen dieser Ziele erfordert jedoch eine sorgfältige und durchdachte Umsetzung, um die Risiken, die durch ungebremste Überschwänglichkeit entstehen, zu mindern.



Referenzen

- Acemoglu, D. (2023). Harms of AI. In *The Oxford Handbook of AI Governance* (pp. 660–706). Oxford University Press eBooks. <https://doi.org/10.1093/oxfordhb/9780197579329.013.65>
- Cookson, C. (2023, November 14). AI outperforms conventional weather forecasting methods for first time. *Financial Times*. <https://www.ft.com/content/ca5d655f-d684-4dec-8daa-1c58b0674be1>
- Fry, H. [Host] (2022). *DeepMind: The podcast – The road to AGI* [Audio podcast]. Google. <https://deepmind.google/discover/the-podcast/the-road-to-agi/>
- Golman Sachs (2024). *Gen AI: Too much spend, too little benefit?* <https://www.goldmansachs.com/intelligence/pages/gs-research/gen-ai-too-much-spend-too-little-benefit/report.pdf>
- Hu, X., Li, X., Chen, J., Li, Y., Li, Y., Li, X., Wang, Y., Liu, Q., Wen, L., Yu, P. S., & Guo, Z. (2024). Evaluating robustness of generative search engine on adversarial factual questions. arXiv.org. <https://arxiv.org/abs/2403.12077>
- International Monetary Fund (2023). *Generative artificial intelligence in finance: Risk considerations*. <https://www.imf.org/-/media/Files/Publications/FTN063/2023/English/FTNEA2023006.ashx>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. a. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- KPMG (2024). *Why finance should lead the adoption of generative AI*. <https://kpmg.com/kpmg-us/content/dam/kpmg/pdf/2024/generative-ai-finance.pdf>
- Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., & Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. *Science* 382(6677), 1416–1421. <https://doi.org/10.1126/science.adi2336>
- Leippold, M. (2023). Sentiment spin: Attacking financial sentiment with GPT-3. *Finance Research Letters* 55(B), 103957. <https://doi.org/10.1016/j.frl.2023.103957>
- Leippold, M., Vaghefi, S. A., Stambach, D., Muccione, V., Bingler, J., Ni, J., Colesanti-Senni, C., Wekhof, T., Schimanski, T., Gostlow, G., Yu, T., Luterbacher, J., & Huggel, C. (2024). *Automated fact-checking of climate change claims with large language models*. arXiv.org. <https://arxiv.org/abs/2401.12566>
- Leippold, M., & Yu, T. (2024). *Are green innovations priced? Evidence beyond patents*. Swiss Finance Institute Research Paper. <http://dx.doi.org/10.2139/ssrn.4391444>
- Liu, N. F., Zhang, T., & Liang, P. (2023). Evaluating verifiability in generative search engines. arXiv.org. <https://arxiv.org/abs/2304.09848>
- Maufe, Z., & Brown, T. (2023, October 3). *Five generative AI use cases for the financial services industry*. Google Cloud. <https://cloud.google.com/blog/topics/financial-services/five-generative-ai-use-cases-financial-services-industry>
- McKinsey & Company (2023). *Capturing the full value of generative AI in banking*. <https://www.mckinsey.com/industries/financial-services/our-insights/capturing-the-full-value-of-generative-ai-in-banking#/>
- Millière, R., & Buckner, C. (2024). *A philosophical introduction to language models – Part I: Continuity with classic debates*. arXiv.org. <https://arxiv.org/abs/2401.03910>
- Munger, C. (2023, May 6). Answer [Q&A Session]. Berkshire Hathaway Annual Meeting, Omaha, NE, United States.
- Ni, J., Bingler, J., Colesanti-Senni, C., Kraus, M., Gostlow, G., Schimanski, T., Stambach, D., Vaghefi, S. A., Wang, Q., Webersinke, N., Wekhof, T., Yu, T., & Leippold, M. (2023). *CHATREPORT: Democratizing sustainability disclosure analysis through LLM-based tools* [Conference proceedings]. 2023 Conference on Empirical Methods in Natural Language Processing, Singapore. <https://doi.org/10.18653/v1/2023.emnlp-demo.3>
- Schimanski, T., Ni, J., Kraus, M., Ash, E., & Leippold, M. (2024). Towards faithful and robust LLM specialists for evidence-based question-answering. arXiv.org. <https://arxiv.org/abs/2402.08277>
- Vaghefi, S. A., Stambach, D., Muccione, V., Bingler, J., Ni, J., Kraus, M., Allen, S., Colesanti-Senni, C., Wekhof, T., Schimanski, T., Gostlow, G., Yu, T., Wang, Q., Webersinke, N., Huggel, C., & Leippold, M. (2023). ChatClimate: Grounding conversational AI in climate science. *Communications Earth & Environment* 4(1). <https://doi.org/10.1038/s43247-023-01084-x>
- Webersinke, N., Kraus, M., Bingler, J. A., & Leippold, M. (2022). *ClimateBert: A pretrained language model for climate-related text*. arXiv.org. <https://doi.org/10.48550/arXiv.2110.12010>

Swiss Finance Institute

Mit Unterstützung seiner Gründer – der Schweizer Bankenbranche, der Schweizerischen Eidgenossenschaft sowie führender Schweizer Universitäten – fördert das Swiss Finance Institute (SFI) aktiv Forschung und Lehre auf Weltniveau im Bereich Banking und Finance in der Schweiz. Durch die Verbindung von akademischer Exzellenz mit Praxiserfahrung trägt das SFI zur Stärkung des Schweizer Finanzplatzes bei.

Herausgeber und Kontakt

Dr. Cyril Pasche
Senior Director Publications and Topic Development
+41 22 379 88 25
cyril.pasche@sfi.ch

swiss:finance:institute

Walchestr. 9, CH-8006 Zurich, T +41 44 254 30 80
info@sfi.ch, www.sfi.ch

