

SFI Public Discussion Note

Il ruolo dell'IA nella trasformazione delle pratiche finanziarie



Novembre 2024

Introduzione



Prof. Markus Leippold
*Senior Chair SFI e Professore
di Ingegneria finanziaria
all'Università di Zurigo*

Con la sua serie di *Public Discussion Note*, lo Swiss Finance Institute (SFI) promuove attivamente un dibattito fondato su argomenti rilevanti per il settore finanziario, la politica e il mondo accademico. Inoltre, l'SFI diffonde i suoi risultati attraverso la ricerca, le pubblicazioni, le Master Class e le conferenze.

Questa *Public Discussion Note* si basa sulla ricerca che io e il mio team abbiamo condotto negli ultimi anni nei settori della sostenibilità e dell'IA. Desidero ringraziare i miei coautori, Julia Bingler, Chiara Colesanti-Senni, Glen Gostlow, Jingwei Ni, Tobias Schimanski, Dominik Stammach, Saeid Vaghefi e Tingyu Yu, nonché tutti gli assistenti di ricerca. Questi giovani ricercatori di talento e motivati sono stati fondamentali per il progresso di questo lavoro. Spero che i loro contributi ispirino i futuri studiosi a continuare a progredire in questo importante campo.

Questa versione è una traduzione della versione originale in inglese. La versione originale è disponibile su <https://sfi.ch>.

"Penso che l'intelligenza vecchio stile funzioni abbastanza bene" (Munger, 2023).

Da quando *OpenAI* ha introdotto *ChatGPT* nel novembre 2022, il discorso sull'intelligenza artificiale (IA) ha catturato l'entusiasmo di Wall Street. Tuttavia, anche se l'IA balza in primo piano nelle chiamate di presentazione degli utili aziendali, alcuni stimati investitori rimangono scettici e considerano le sue promesse come semplici esagerazioni.

Warren Buffett della Berkshire Hathaway, una delle principali fonti di competenza per gli investimenti, insieme al suo collaboratore di lunga data, Charles Munger, recentemente scomparso, contrappone l'intelligenza artificiale all'intelligenza "vecchio stile". Mentre altri sono sempre più affascinati dalle capacità dell'IA, dai suoi rapidi progressi e dalla sua integrazione nella sfera aziendale, Buffett sottolinea il valore duraturo dell'intelletto umano. Può l'IA, si chiede, superare davvero la profondità e le sfumature del pensiero e della creatività umana?

Come è stato per la Rivoluzione Industriale, l'attuale era dell'IA è caratterizzata da un rapido progresso tecnologico e dall'ottimismo sulle sue implicazioni per le nostre strutture sociali, l'economia e la vita quotidiana. L'introduzione di metodi di produzione meccanizzati, circa 200 anni fa, ha segnato una svolta significativa nella storia dell'umanità; l'IA ne rappresenta un'altra. Mentre esploriamo la capacità delle macchine di creare, ragionare e interagire in modi che prima si pensava fossero esclusivi per gli esseri umani, dobbiamo anche essere consapevoli delle preoccupazioni etiche, di privacy e di sicurezza che accompagneranno l'adozione diffusa dell'IA.

In questa *Public Discussion Note* spero di trovare un equilibrio tra esuberanza e cautela. Contrapponendo le proiezioni ottimistiche delle grandi aziende tecnologiche, che considerano l'IA un elemento cruciale per la competizione sul mercato, alle prospettive scettiche di investitori come Buffett e Munger, cercherò di svelare le complessità dell'attuale fenomeno dell'IA. Inizierò parlando della tecnologia in sé e del potenziale dell'IA di ridisegnare il panorama economico, per finire con le considerazioni etiche, le implicazioni sociali e le sfide che devono essere affrontate prima di una sua applicazione diffusa in finanza.

Breve storia dell'intelligenza artificiale

Prime visioni (prima del 1950):

Fin dall'antichità, le civiltà hanno immaginato forme di vita artificiale, come testimoniano i miti di Pigmalione e del Golem. Nella storia di Pigmalione, uno scultore si innamora di una statua che ha scolpito nella pietra e che poi prende vita. Il Golem è una creatura fatta di argilla che, allo stesso modo, viene portata in vita con metodi mistici per servire il suo creatore. Inoltre, durante i periodi antico e medievale, inventori provenienti da varie culture crearono dispositivi meccanici che imitavano le azioni umane o animali. Questi automi riflettono anche le prime visioni dell'intelligenza artificiale.

Filosofi come René Descartes (nato nel 1596) e Gottfried Leibniz (nato nel 1646) hanno svolto un ruolo fondamentale nella riflessione sulla mente, sull'intelligenza e sul potenziale delle macchine nell'imitare il ragionamento umano. Cartesio si interrogò sulla natura del pensiero e della coscienza, mentre Leibniz immaginò un linguaggio universale per il ragionamento, noto come *calculus ratiocinator*. Questa prima struttura concettuale, che rappresenta le affermazioni e i ragionamenti logici in una forma matematica precisa, ha influenzato la logica simbolica e la computazione moderne.

Nascita del campo (anni '50):

Il lavoro di Alan Turing negli anni 1940 e 1950, in particolare nello sviluppo del Test di Turing, ha posto le basi per la valutazione dell'intelligenza delle macchine. Il Test di Turing è un metodo per determinare se una macchina può esibire un comportamento intelligente indistinguibile da quello di un essere umano. I lavori precedenti di Turing hanno anche stabilito principi fondamentali che sono alla base dell'informatica moderna. La macchina di Turing, per esempio, è un dispositivo teorico che manipola simboli su una striscia di nastro in base a una serie di regole, fornendo un modello semplice ma potente per capire cosa significa per una funzione l'essere computabile.

La Conferenza di Dartmouth (1956) ha riunito i ricercatori che condividevano la convinzione che l'intelligenza delle macchine potesse essere raggiunta, e ha posto le basi per la ricerca futura. È spesso citata come la nascita formale dell'IA in quanto disciplina scientifica.

Le montagne russe del progresso (1960-1990):

A causa del fatto che il clamore ha superato i progressi effettivi, il campo dell'IA ha conosciuto periodi di riduzione dei finanziamenti e dell'interesse. Questi cosiddetti "inverni dell'IA" si sono verificati principalmente tra la fine degli anni '70 e l'inizio degli anni '80, e di nuovo tra la fine degli anni '80 e l'inizio degli anni '90. Sono stati caratterizzati dallo scetticismo sul potenziale dell'IA di mantenere le sue promesse, portando a una riduzione degli investimenti e a una rivalutazione degli obiettivi.

Nonostante le sfide, lo sviluppo dei sistemi esperti ha rappresentato un successo significativo per l'IA nelle applicazioni commerciali e industriali. Questi sistemi sono stati in grado di imitare le capacità decisionali degli esperti umani in ambiti specifici, come la diagnosi medica e l'esplorazione mineraria. In campo finanziario, i sistemi esperti sono stati utilizzati per stimare le perdite per insolvenza, fornendo a banche e istituzioni finanziarie strumenti sofisticati per valutare il rischio di credito e prevedere la probabilità di insolvenza dei mutuatari.

Rinascita e trasformazione (dagli anni '90 a oggi):

La rinascita dell'IA è stata alimentata in modo significativo dai progressi nell'apprendimento automatico e dalla disponibilità di grandi insiemi di dati noti come *"big data"*. Questi sviluppi hanno permesso ai sistemi di IA di imparare dagli esempi, invece di affidarsi esclusivamente a regole codificate.

L'apprendimento profondo, un tipo di apprendimento automatico che utilizza le reti neurali, ha portato a importanti progressi nella visione artificiale, nel riconoscimento vocale e nell'elaborazione del linguaggio naturale, rendendo possibili applicazioni come la traduzione linguistica in tempo reale, i veicoli autonomi e la ricerca efficace sul web.

Motivazione

All'inizio degli anni Novanta, quando ero ancora giovane, la lettura del romanzo di fantascienza di Stanislaw Lem "Così parlò Golem" ha scatenato la mia fascinazione per le possibilità dell'intelligenza artificiale. La narrazione si svolge attorno a un personaggio chiamato Golem, una macchina dall'intelletto senza precedenti che si impegna in profonde discussioni con i più importanti scienziati del mondo. Se all'epoca il concetto mi sembrava puramente speculativo, il romanzo di Lem, scritto negli anni Ottanta, rispecchia in modo inquietante le interazioni odierne con l'IA in settori come la finanza. L'esplorazione di Lem delle capacità, delle implicazioni etiche e del potenziale di coesistenza dell'IA con gli esseri umani risuona oggi più che mai, dal momento che tecnologie come ChatGPT di OpenAI e *Gemini* di Google entrano a far parte delle nostre attività quotidiane.

Oggi sono attivamente coinvolto negli sviluppi tecnologici che Lem aveva immaginato. Riflettere sul romanzo di Lem sottolinea l'importanza di navigare il ruolo tuttora in espansione dell'IA nel nostro settore con attenzione e lungimiranza. E mi ricorda il delicato equilibrio tra sfruttare i progressi tecnologici e considerare le loro implicazioni più ampie. Questa riflessione diventa particolarmente importante quando osserviamo l'impatto trasformativo dell'IA in medicina, biologia e meteorologia.

Per esempio, *AlphaFold*, sviluppato da *Google DeepMind*, ha rivoluzionato la biologia prevedendo con precisione le strutture in 3D delle proteine basandosi esclusivamente sulle loro sequenze di amminoacidi (Jumper et al., 2021). Questa scoperta risolve l'annoso problema del ripiegamento proteico, accelerando la

scoperta di farmaci e migliorando la comprensione dei processi biologici grazie alla previsione delle strutture di tutti i 200 milioni di proteine note alla scienza in un solo anno. Considerando che, secondo una stima approssimativa, uno studente di dottorato impiega dai quattro ai cinque anni per scoprire la struttura di una singola proteina, questo compito avrebbe altrimenti richiesto centinaia di milioni di anni (!) di lavoro umano. Demis Hassabis e John M. Jumper, entrambi provenienti da *Google DeepMind*, hanno ricevuto il Premio Nobel per la Chimica 2024, per i loro contributi innovativi nel predire il ripiegamento proteico. Analogamente, *GraphCast*, un altro modello innovativo di *Google DeepMind*, ha ridefinito le previsioni meteorologiche superando i metodi tradizionali con una notevole precisione (Cookson, 2023; Lam et al., 2023). Analizzando ampi dati storici con una rete neurale a grafo, *GraphCast* fornisce migliori previsioni delle condizioni meteorologiche, dalla temperatura all'andamento dei venti fino all'approdo degli uragani. Inoltre, riduce significativamente il consumo di energia, abbassando il tempo di calcolo necessario da molte ore a un minuto.

Potremmo assistere a un momento simile a quello di *AlphaFold* o *GraphCast* nel settore finanziario? L'IA generativa, in quanto in grado di individuare modelli e trarre spunti da vasti insiemi di dati, ha il potenziale per trasformare radicalmente l'analisi finanziaria, la valutazione del rischio e le previsioni economiche. Potenti strumenti di IA potrebbero ridefinire il nostro processo decisionale strategico, sottolineando la necessità di prontezza e adattabilità nel settore finanziario per garantire il raggiungimento di risultati significativi e responsabili.



IA generativa nel settore finanziario

Prima di approfondire il suo impatto sul settore finanziario, vorrei chiarire cosa intendo per IA generativa. IA e IA generativa sono concetti correlati ma distinti; mentre tutta l'IA generativa è IA, non tutta l'IA è generativa.

IA è un termine ampio che comprende tutti i sistemi informatici progettati per imitare le funzioni cognitive umane. Questi sistemi possono essere classificati come IA stretta o debole, progettata per svolgere un compito specifico che in precedenza richiedeva l'intelligenza umana, o come IA generale o forte (*Artificial General Intelligence – AGI*), in grado di applicare l'intelligenza a un'ampia gamma di compiti, come fanno gli esseri umani.

I sistemi di IA attualmente sotto i riflettori del pubblico, come ChatGPT, sono esempi di IA debole. Sono molto avanzati e capaci di generare testi simili a quelli umani, ma sono specializzati per compiti specifici e non possiedono le capacità ad ampio raggio di un'AGI. Molti ricercatori ed esperti concordano sul fatto che l'AGI completa rimane ipotetica, poiché nessun sistema ha ancora soddisfatto i criteri concordati. Tuttavia, alcuni ritengono che il raggiungimento dell'AGI sia una questione di quando, non di se (Fry, 2022).

L'IA generativa si riferisce a un sistema di IA – debole o forte – in grado di generare nuovi contenuti o dati simili, ma non identici, a quelli per cui è stato addestrato. L'addestramento di un algoritmo consiste nel presentargli un ampio insieme di dati e nell'assegnargli una serie di compiti da risolvere, guidati da un meccanismo di ricompensa che fornisce un feedback sulle sue prestazioni. Questo processo consente all'algoritmo di affinare la comprensione dei modelli e dei criteri decisionali, migliorando la sua capacità di risolvere compiti simili in futuro. L'insieme di dati può includere testo, immagini, musica, voce e altri media. I modelli di intelligenza artificiale generativa sono progettati per comprendere gli schemi e le strutture fondamentali di questi dati in ingresso, consentendo loro di produrre risultati nuovi e originali sulla base di ciò che hanno appreso.

La caratteristica cruciale dell'IA generativa è la sua capacità di creare, offrendo un notevole potenziale nella creazione di contenuti, nel design e persino nella scoperta di nuovi farmaci. Le capacità cognitive e creative dell'IA generativa sono oggetto di un continuo dibattito filosofico (Millière & Buckner, 2024). Tuttavia essa ha il potenziale non solo di replicare i compiti umani, ma anche di aumentare la creatività umana, generando nuove idee, progetti e prospettive, offrendo applicazioni innovative che vanno oltre l'attuale portata delle capacità analitiche e decisionali dell'IA.

Impatto dell'IA generativa nel settore finanziario

L'IA generativa sta già avendo un impatto significativo sul settore finanziario, con implicazioni per la produttività, l'efficienza operativa, la gestione del rischio e la creazione di nuovi modelli di business. Secondo un rapporto di McKinsey, l'IA generativa potrebbe aggiungere tra i 200 e i 340 miliardi di dollari all'anno al settore bancario, pari al 9-15% dei suoi profitti operativi (McKinsey & Company, 2023). Questo valore deriva principalmente dall'aumento della produttività, con i guadagni assoluti più significativi previsti nei settori corporate e retail.

Grazie all'IA generativa, gli istituti finanziari possono potenziare le operazioni di servizio e migliorare l'esperienza dei clienti. Per esempio, secondo McKinsey & Company (2023), Goldman Sachs utilizza uno strumento basato sull'IA per automatizzare la generazione di test, che in precedenza richiedeva una grande quantità di lavoro. Citigroup ha utilizzato l'IA generativa per valutare l'impatto delle nuove norme patrimoniali statunitensi.

Maufe e Brown (2023) identificano cinque usi pratici dell'IA generativa nel settore dei servizi finanziari, e tutti e cinque aumentano la produttività dei dipendenti. Questi includono la ricerca e la sintesi di documenti finanziari, il potenziamento degli assistenti virtuali, la ricerca sui mercati dei capitali, lo sviluppo di software e le raccomandazioni finanziarie personalizzate.

Tuttavia, il Fondo Monetario Internazionale ha osservato che l'impiego dell'IA generativa nel settore finanziario comporta diversi rischi (International Monetary Fund, 2023). Questi includono pregiudizi integrati ("*embedded bias*"), carenze nella privacy, mancanza di trasparenza nel processo decisionale, problemi di robustezza, minacce alla sicurezza informatica e potenziali impatti sulla stabilità finanziaria in generale. Il Fondo Monetario Internazionale avverte che un eccessivo affidamento all'IA generativa potrebbe aumentare il rischio di contagio e creare nuovi rischi sistemici nel settore finanziario.

Nonostante queste preoccupazioni, il settore finanziario sta esplorando attivamente le capacità dell'IA generativa, con pressioni competitive che alimentano la rapida adozione di applicazioni che promettono guadagni in termini di efficienza, risparmi sui costi e migliore gestione del rischio.

La tecnologia dell'IA generativa offre una serie di potenziali utilizzi in ambito finanziario, tra cui:

- **Elaborazione di previsioni e bilanci.** Un'indagine di KPMG ha rilevato che l'83% dei partecipanti ha impiegato l'IA nella pianificazione finanziaria (KPMG, 2024), affidandosi ad essa per l'analitica predittiva, la generazione di scenari e la fornitura di informazioni di bilancio.
- **Creazione di relazioni e presentazioni finanziarie.** L'IA generativa riduce in modo significativo il tempo e l'impegno necessari per la produzione di relazioni finanziarie regolari, popolando automaticamente i modelli con nuovi dati e ricavando approfondimenti dai dati finanziari e aziendali.
- **Estrazione di informazioni sul mercato.** Utilizzando gli ampi modelli linguistici dell'IA generativa, le organizzazioni possono attingere a insiemi di dati pubblici per raccogliere informazioni utili sul mercato, creare informazioni sulla concorrenza e scoprire le tendenze dei clienti attraverso analisi personalizzate che possono essere specifiche per regione o per persona.
- **Ricavare informazioni strategiche dai dati.** Utilizzando l'IA generativa per analizzare i dati relativi alla gestione delle relazioni con i clienti (*Customer Relationship Management – CRM*) o alla pianificazione delle risorse aziendali (*Enterprise Resource Planning – ERP*), le aziende possono ottenere preziose informazioni strategiche per formare le loro decisioni finanziarie, per affrontare le sfide relative ai prezzi e alle prestazioni o per migliorare altre operazioni aziendali.
- **Automatizzare la gestione dei contratti.** L'IA generativa semplifica la creazione dei contratti, concentrandosi sui termini non standardizzati, identificando le clausole relative ai ricavi e documentando i trattamenti contabili.

- **Migliorare il rilevamento delle anomalie e la prevenzione delle frodi.** Attraverso l'identificazione delle irregolarità dei dati e il monitoraggio delle transazioni in corso, l'IA generativa offre un solido meccanismo per rilevare gli errori e prevenire le frodi, salvaguardando così l'integrità finanziaria di un'organizzazione.
- **Consulenza personalizzata sugli investimenti.** L'IA generativa può adattare le raccomandazioni di investimento al profilo di rischio e agli obiettivi finanziari di un individuo, migliorando l'esperienza del cliente.
- **Rilevamento delle frodi.** Analizzando grandi insiemi di dati di transazioni, l'IA generativa può identificare modelli fraudolenti, migliorando la sicurezza finanziaria.
- **Analisi di mercato e gestione del rischio.** Allo stesso modo, con i dati appropriati, l'IA generativa può identificare le tendenze del mercato e valutare i rischi potenziali, aiutando le decisioni di investimento.

Questa ampia varietà di applicazioni in campo finanziario solleva diversi importanti punti di discussione all'interno della professione. Il primo è costituito dalle considerazioni etiche, in particolare le modalità per mantenere la trasparenza. Sono necessari nuovi standard e protocolli per garantire che i contenuti finanziari generati dall'IA non perpetuino errori o disinformazioni. Il secondo punto è l'impatto dell'automazione sui nostri dipendenti. Poiché l'IA sostituisce i ruoli tradizionali, la nostra forza lavoro deve apprendere nuove competenze ed essere pronta ad adattarsi. Il terzo punto è il panorama normativo. Man mano che le tecnologie di IA generativa avanzano e le normative finanziarie diventano inadeguate o obsolete, i quadri normativi devono adattarsi per garantire sia l'innovazione sia la protezione dei consumatori. L'ultimo punto è lo sviluppo continuo di prodotti e servizi innovativi per soddisfare le richieste dei mercati emergenti e fornire alle istituzioni finanziarie un vantaggio competitivo.

Applicazioni pratiche

Esaminiamo alcune applicazioni pratiche dell'IA generativa ed esploriamo il suo ruolo nel guidare i progressi nel settore finanziario. Questi esempi di iniziative ambientali, sociali e di *governance* (*Environmental, Social, and Governance – ESG*) dimostrano come l'IA generativa possa trasformare un'area specifica della finanza, evidenziando il suo potenziale per rivoluzionare l'intero settore.

Case Study 1: identificare l'innovazione verde

In un recente articolo, il mio coautore e io abbiamo analizzato come l'IA generativa svolga un ruolo cruciale nel decifrare il denso linguaggio dei brevetti verdi (Leippold e Yu, 2024). Comprendere le complessità tecniche di questi brevetti può essere scoraggiante per chi vuole valutare le iniziative di sostenibilità di un'azienda. I modelli di IA generativa superano questa sfida trasformando descrizioni complesse e terminologia specializzata in sintesi concise. Il processo prevede che l'IA "legga" sistematicamente il brevetto, ne interpreti gli elementi cruciali e li riarticoli in termini semplici. Questa applicazione dell'IA generativa svolge molteplici funzioni che sono vitali per il settore finanziario.

In primo luogo, sintetizzando il gergo tecnico in un linguaggio chiaro, l'IA generativa garantisce un accesso più ampio ai dettagli dell'innovazione verde. Estende le informazioni del brevetto a un pubblico eterogeneo, tra cui analisti finanziari, investitori e responsabili politici, che potrebbero non possedere le profonde competenze tecniche altrimenti necessarie per comprenderle. Questa democratizzazione delle informazioni è fondamentale, in quanto consente a una più ampia gamma di *stakeholder* di prendere decisioni informate.

In secondo luogo, le sintesi fornite dall'IA migliorano l'efficienza analitica. Gli analisti finanziari, che operano in ambienti in cui la velocità e l'accuratezza sono fondamentali, traggono immensi vantaggi dalla possibilità di comprendere rapidamente il significato di un brevetto senza impantanarsi nei dettagli. In questo caso, l'IA non solo fa risparmiare tempo, ma migliora anche la qualità dell'analisi, garantendo che le decisioni finanziarie siano basate su una comprensione approfondita dell'impatto potenziale e del valore delle innovazioni di un'azienda.

In terzo luogo, le sintesi dell'IA facilitano l'analisi dei mercati. Gli analisti possono confrontare più facilmente gli sforzi di sostenibilità di varie aziende, settori e industrie quando hanno una chiara comprensione dei miglioramenti principali di ciascun brevetto. Questa capacità è preziosa per i potenziali investitori, in quanto garantisce che le loro decisioni siano basate su una piena comprensione dell'impegno delle diverse aziende verso l'innovazione verde.

Dopo aver utilizzato l'IA generativa per riassumere i complessi testi dei brevetti verdi in sintesi chiare e concise, i miei coautori e io abbiamo creato un modello linguistico basato su *BERT*, chiamato *ClimateBERT*, per affinare ulteriormente la nostra comprensione (Webersinke et al., 2022). *BERT* è l'acronimo di *Bidirectional Encoder Representations from Transformers*. Sviluppato da Google, *BERT* comprende meglio il contesto di una parola considerando sia la parola precedente sia quella successiva in una frase. Abbiamo addestrato *ClimateBERT* a riconoscere le sfumature della terminologia della sostenibilità, consentendo una profonda analisi contestuale dei testi semplificati prodotti dall'IA generativa.

Analizzando questi testi nel tempo, *ClimateBERT* è in grado di identificare le tendenze nel discorso sulla sostenibilità da parte di un'azienda. Questo tipo di analisi evidenzia i cambiamenti nell'enfasi sull'innovazione verde da parte dell'azienda, fornendo segnali precoci di cambiamenti strategici che potrebbero influenzare il suo impegno a lungo termine verso la sostenibilità, nonché la sua performance finanziaria. Combinando l'IA generativa e *ClimateBERT*, riusciamo a comprendere meglio il modo in cui l'azienda si impegna nell'innovazione verde.

Infine, per comprendere gli impatti più ampi dell'innovazione verde, utilizziamo i nostri dati per creare metriche che mostrano come le iniziative di sostenibilità abbiano influenzato i rendimenti anomali cumulativi corretti per il mercato in occasione di eventi importanti, come la vittoria di Trump alle elezioni del 2016, la vittoria di Biden nel 2020, lo scoppio della guerra tra Russia e Ucraina e l'annuncio dell'*Inflation Reduction Act*. Il nostro lavoro offre agli analisti finanziari e agli investitori un potente strumento per comprendere meglio e mitigare i rischi legati all'innovazione verde e alla sostenibilità.

Case Study 2: Analisi della sostenibilità aziendale

I miei coautori e io abbiamo sviluppato uno strumento di IA chiamato *CHATREPORT* (Ni et al., 2023), basato sul nostro precedente lavoro su *CHATCLIMATE* (Vaghefi et al., 2023), per affrontare i dati densi e voluminosi che si trovano nei resoconti di sostenibilità aziendali. Questi resoconti descrivono in dettaglio l'impatto ambientale, le strutture di governance e le iniziative sociali di un'azienda e fungono da lente critica attraverso la quale gli investitori, i responsabili politici e il pubblico in generale possono valutare l'impegno dell'azienda nei confronti delle pratiche sostenibili. Tuttavia, la loro complessità e la mole di dati che contengono rendono l'analisi manuale un compito arduo. La sfida è ulteriormente amplificata quando l'analisi deve attenersi a specifici quadri normativi o linee guida. Tradizionalmente, solo

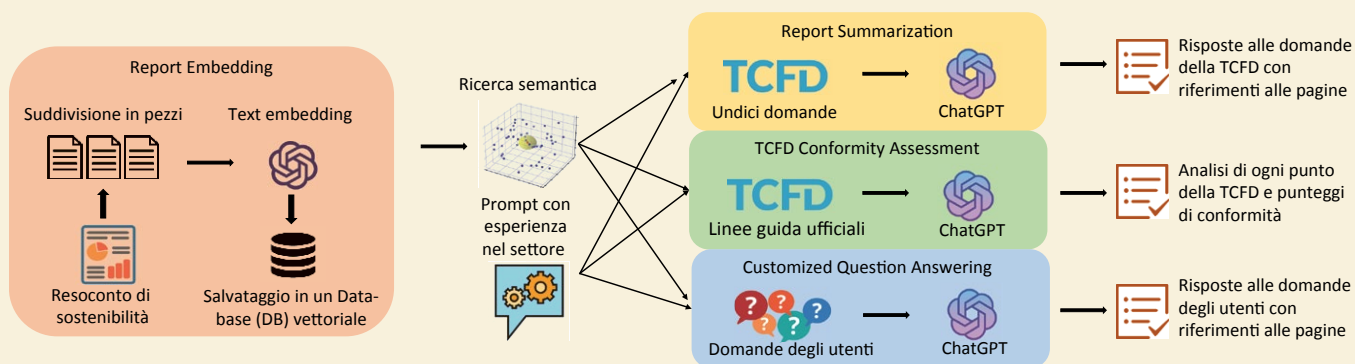
poche entità sono state dotate delle risorse necessarie per un'analisi approfondita e dettagliata di questi resoconti, creando una notevole barriera alla trasparenza. CHATREPORT democratizza il processo, consentendo a un'ampia gamma di stakeholder di entrare in contatto con questi resoconti e di agire di conseguenza.

Il cuore di CHATREPORT è una sofisticata *pipeline*, illustrata nella Figura 1, che trasforma gli intricati dati dei resoconti di sostenibilità aziendali in approfondimenti utilizzabili. Un aspetto fondamentale della nostra metodologia consiste nell'integrare le competenze del settore direttamente nel ciclo di sviluppo dell'IA. Il processo inizia con l'estrazione del testo grezzo di un resoconto, la normalizzazione dei formati dei dati e la rimozione di tutte le informazioni irrilevanti. Per comprendere e analizzare il contenuto, la pipeline utilizza modelli linguistici avanzati di grandi dimensioni (*Large Language Models – LLM*) che vengono messi a punto con conoscenze specifiche del settore. Gli utenti e altri esperti umani creano e perfezionano i prompt analitici, garantendo che i risultati dell'IA siano tecnicamente accurati e contestualmente rilevanti. Lo strumento identifica i temi chiave, valuta la conformità alle linee guida come quelle della *Task Force on Climate-related*

Financial Disclosures (TCFD) e genera approfondimenti in un formato di facile utilizzo. Ogni fase della pipeline di CHATREPORT è progettata per garantire che i dati siano elaborati con grande precisione e allineati con gli standard e le pratiche di sostenibilità più recenti, dando vita a uno strumento potente per gli stakeholder che devono prendere decisioni ben formate.

Una sfida comune con i contenuti generati dall'IA, come discusso di seguito, è il rischio di "allucinazione", in cui l'IA produce informazioni errate o fuorvianti non sostenute dai dati. CHATREPORT affronta questo problema grazie alle sue caratteristiche di tracciabilità. Ogni contenuto generato dall'IA è direttamente riconducibile a sezioni specifiche dei documenti di partenza, consentendo all'utente di verificare manualmente l'accuratezza delle informazioni. Inoltre, CHATREPORT è fortemente incentrato sull'utente e offre funzioni che consentono agli utenti di adattare l'analisi alle loro esigenze. Attraverso un modulo dinamico di risposta alle domande gli utenti possono interrogare l'IA su aspetti particolari di un resoconto. L'IA allora recupera ed elabora i dati pertinenti, fornendo approfondimenti personalizzati.

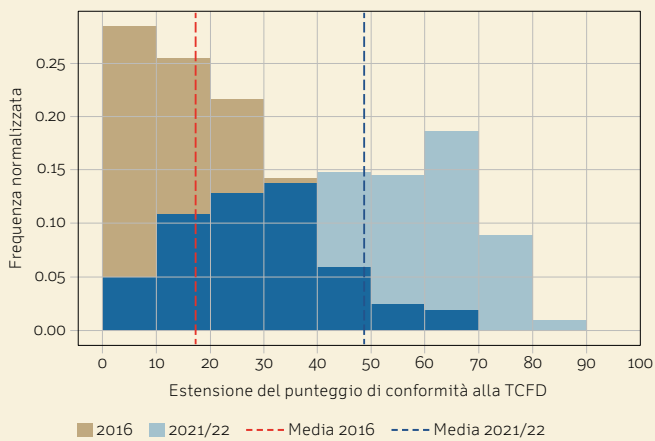
Figura 1: La pipeline di CHATREPORT



Nota: La pipeline di CHATREPORT automatizza l'analisi dei resoconti di sostenibilità aziendali attraverso quattro moduli principali. Il modulo *Report Embedding (RE)* divide il resoconto in pezzi di testo, incorporando questi ultimi in uno spazio vettoriale per la ricerca semantica basata sulle raccomandazioni della TCFD. Il modulo *Report Summarization (RS)* recupera le sezioni rilevanti del resoconto e chiede a un modello linguistico di grandi dimensioni (LLM) di riassumere le informazioni in base agli undici aspetti raccomandati dalla TCFD. Il modulo *TCFD Conformity Assessment (TCA)* valuta l'aderenza del resoconto alle linee guida TCFD, generando punteggi di conformità e paragrafi di analisi. Il modulo *Customized Question Answering (CQA)* consente agli utenti di porre domande specifiche, mentre il LLM fornisce le risposte facendo riferimento alle sezioni appropriate del resoconto. Questa pipeline completa migliora la trasparenza e l'efficienza dell'analisi dei resoconti di sostenibilità.

Fonte: Ni et al. (2023)

Figura 2: Grafico di densità dei punteggi di conformità alla TCFD



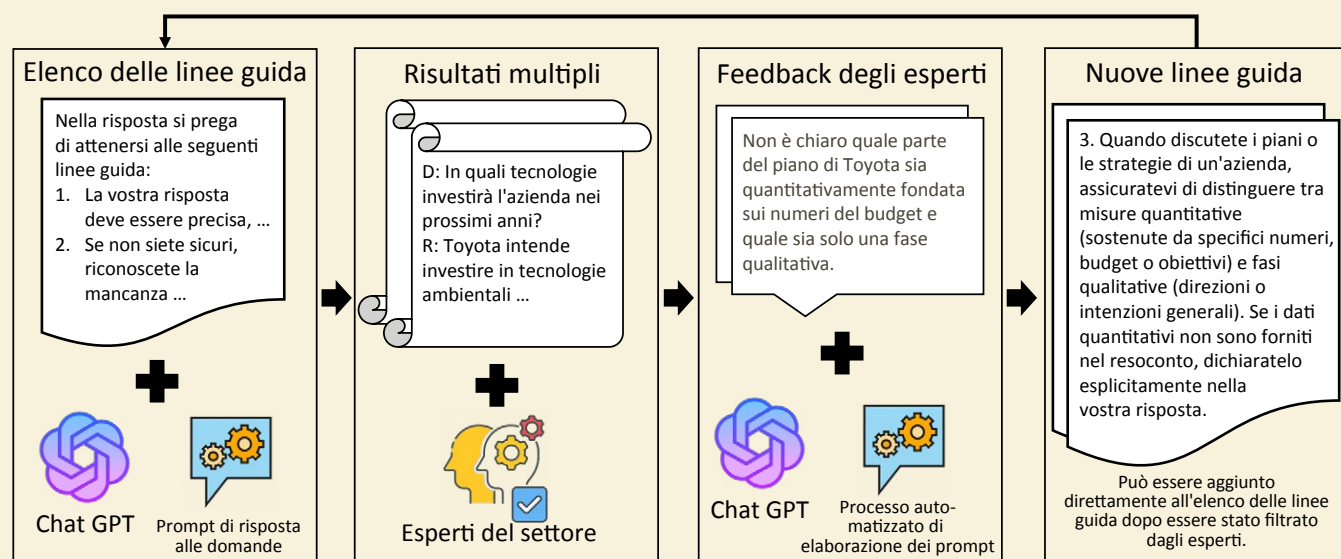
Nota: Questa figura illustra la distribuzione dei punteggi di conformità alla TCFD per un campione di resoconti di sostenibilità aziendali di società quotate al NYSE, per gli anni di riferimento 2016 e 2021/22. I punteggi di conformità alla TCFD sono calcolati in base all'aderenza dei resoconti alle linee guida della TCFD. Il grafico di densità mostra un notevole miglioramento dei punteggi medi di conformità tra il 2016 e il 2021/22, a indicare che le aziende stanno allineando sempre più le loro informazioni sulla sostenibilità alle raccomandazioni della TCFD.

Fonte: Ni et al. (2023)

Nel nostro studio, per esempio, abbiamo valutato oltre 10.000 resoconti di sostenibilità aziendali per verificare la conformità ai requisiti di reporting della TCFD. La Figura 2 presenta visivamente i punteggi per due periodi di relazione: prima della TCFD, in marrone, e (alcuni anni) dopo la TCFD, in blu chiaro. I nostri dati rivelano che il livello di conformità alla TCFD è aumentato sostanzialmente tra i due periodi. Prima dell'inaugurazione formale della TCFD, nel 2016, il punteggio medio di conformità era del 17%. Nel 2021/22 era del 49%. Questa tendenza evidenzia la crescente importanza di un'informativa finanziaria trasparente e completa sul clima. In ulteriori analisi, abbiamo identificato tendenze e valori anomali, facilitando una discussione sull'efficacia delle pratiche attuali e sulle potenziali aree di miglioramento.



Figura 3: Pipeline di elaborazione automatica dei prompt



Nota: CHATREPORT incorpora il feedback degli utenti e di altri esperti del settore per perfezionare e migliorare i prompt utilizzati dal modello linguistico di grandi dimensioni (LLM). Inizialmente, gli esperti del settore forniscono un feedback sulle risposte specifiche generate dall'LLM. Questo feedback viene poi trasformato in linee guida generali attraverso un processo automatizzato di elaborazione dei prompt. L'IA integra le nuove linee guida nei suoi modelli, migliorando la qualità e la precisione dei suoi risultati futuri. Questo processo iterativo porta a un'analisi più accurata, dettagliata e fondamentale dei resoconti di sostenibilità.

Fonte: Ni et al. (2023)

Infine, come mostra la Figura 3, CHATREPORT è costruito per imparare e adattarsi nel tempo. Il sistema incorpora continuamente il feedback degli utenti e di altri esperti umani per perfezionare i modelli analitici e i prompt. Questo processo iterativo garantisce che il sistema migliori costantemente in termini di accuratezza, affidabilità e pertinenza. Per enfatizzare la trasparenza, i miei coautori e io abbiamo sviluppato CHATREPORT come progetto *open-source*; è possibile accedervi su <https://reports.chatclimate.ai/>. Rendendo pubbliche le nostre metodologie, le raccolte di dati e gli strumenti, incoraggiamo la collaborazione e invitiamo i programmatori e i ricercatori di tutto il mondo a contribuire al progresso dell'analisi della finanza sostenibile.

La potenza di CHATREPORT risiede nella sua capacità di accelerare drasticamente il processo di analisi. Con questo strumento di IA, è possibile estrarre informazioni rilevanti da migliaia di resoconti di sostenibilità aziendali (*Corporate Sustainability Reports – CSR*), fino a 10.000, in poche ore. CHATREPORT condensa rapidamente enormi quantità di dati in concisi riassunti esecutivi, consentendo un processo decisionale più rapido e basato sui dati a una portata precedentemente irraggiungibile.

Uno sguardo al futuro (molto) prossimo

L'IA generativa potrebbe presto cambiare radicalmente il modo in cui vengono analizzati gli investimenti e le decisioni finanziarie. Recentemente ho testato una strategia per il controllo dei fatti automatizzato (*fact-checking*) che ricorda il metodo socratico. Nell'antica Grecia, Socrate impegnava i suoi interlocutori in domande incessanti per mettere in luce le contraddizioni dei loro pensieri, incoraggiando un'esplorazione più profonda dei principali problemi. Questa eredità filosofica riecheggia nel modello del Mediatore-Avvocato utilizzato da Leippold et al. (2024) per verificare le affermazioni sul clima.

Trasferendo questo strumento nella finanza, avremmo LLM multipli, definiti "*AI Analysts*", ognuno dei quali analizzerebbe una specifica sfaccettatura dei dati finanziari – come le tendenze del mercato, gli impatti normativi, gli indicatori economici o le metriche di performance aziendale – e proporrebbe potenziali strategie di investimento. Questi argomenti verrebbero poi integrati da un "*Chief AI-Investment Manager*". Per formare una valutazione logica, questo manager porrebbe ulteriori domande agli AI Analysts, nello

stesso modo in cui Socrate guidava i suoi allievi verso una comprensione più profonda della realtà attraverso la dialettica. Questo processo garantisce che vengano prese in considerazione molte prospettive diverse, che si confrontano e si completano a vicenda e che, in ultima analisi, producono una conclusione più solida.

Uno strumento di questo tipo migliorerebbe il processo decisionale finanziario sfruttando la velocità di calcolo dell'IA e riducendo i pregiudizi a favore di un tipo di analisi finanziaria rispetto a un altro. Inoltre, con l'aumento della complessità dei mercati finanziari, l'adattabilità dei sistemi di IA garantirebbe ai decisori una gestione efficiente della crescente domanda di dati.

Una caratteristica fondamentale di tale sistema sarebbe la capacità di monitorare e registrare la progressione delle argomentazioni tra gli Analisti IA, fornendo una traccia di controllo su come sono state raggiunte le loro decisioni. Tale tracciabilità è fondamentale per rafforzare la fiducia tra gli stakeholder e per soddisfare i requisiti normativi del settore finanziario, che richiedono una documentazione approfondita dei processi decisionali, soprattutto in aree come la valutazione del rischio e la conformità.

Se da un lato l'integrazione dei sistemi di IA nella finanza promette un cambiamento verso un processo decisionale più dinamico e in tempo reale, dall'altro questo potenziale comporta molte sfide teoriche e pratiche, come garantire l'accuratezza dei dati, mantenere la credibilità delle fonti e mitigare i rischi di errori algoritmici.



Limiti attuali dell'IA generativa

Nonostante il suo potenziale, la tecnologia dell'IA generativa è afflitta da importanti limitazioni che richiedono un'attenzione urgente. Prima che possa rivoluzionare il settore finanziario, è necessario affrontare questioni come l'allucinazione, la mancanza di verificabilità e la vulnerabilità agli attacchi avversari.

Allucinazione e imprecisione

In Schimanski et al. (2024), i miei coautori e io analizziamo criticamente l'accuratezza e la verificabilità delle risposte generate dai LLM. Nei compiti complessi di risposta alle domande e di generazione di contenuti, il fenomeno noto come "allucinazione" è comune. L'allucinazione nell'IA si riferisce alla generazione, da parte del modello, di informazioni che, sebbene plausibili, si distaccano dal materiale della fonte reale e sono del tutto infondate. Questo problema è critico in settori come la finanza, il sistema sanitario e la giurisprudenza, dove l'accuratezza e l'affidabilità delle informazioni sono fondamentali. Il nostro studio evidenzia come i LLM possano spesso attribuire in modo errato le fonti o creare contenuti, minando gravemente la fiducia degli utenti in questi sistemi automatizzati.

Per affrontare queste sfide, proponiamo di migliorare il processo di estrazione del testo dalle fonti pertinenti e di migliorare l'attribuzione dei contenuti generati dall'IA. Tuttavia, queste soluzioni tecniche sono solo una parte della soluzione. Le implicazioni di un'imprecisione non controllata sono vaste, con possibili importanti ripercussioni nel processo decisionale finanziario, nella cura dei pazienti e nei procedimenti legali. Nuovi studi sono urgentemente necessari per perfezionare questi modelli e sviluppare un solido quadro etico per il loro utilizzo.

Mancanza di verificabilità

I motori di ricerca generativi, tra cui piattaforme ben note come *Bing Chat*, *NeevaAI*, *perplexity.ai* e *YouChat*, sono progettati per rispondere alle domande degli utenti generando risposte complete di citazioni. Tuttavia, Liu et al. (2023) hanno scoperto che solo il 52% circa delle frasi prodotte da questi motori di ricerca è adeguatamente sostenute da citazioni. Questo dato è allarmante, se si considera l'importanza di informazioni affidabili nell'odierna era digitale. Inoltre lo studio mostra che solo il 75% delle citazioni sostiene effettivamente le affermazioni che sembra convalidare. Questo divario tra apparenza e realtà trae in inganno gli utenti, che si affidano a informazioni apparentemente credibili ma fondamentalmente non sostenute o citate in modo errato.

Questa ricerca mette in luce un aspetto cruciale dello sviluppo dell'IA: l'affidabilità. Nonostante i sofisticati risultati di questi motori di ricerca generativi, la prevalenza di affermazioni non sostenute e di imprecisioni nelle citazioni ne pregiudica significativamente l'affidabilità. L'addestramento nella padronanza dell'IA deve essere perfezionato e devono essere sviluppate strategie per migliorare la verificabilità delle informazioni fornite da questi sistemi.

Vulnerabilità agli attacchi avversari

L'IA generativa è anche suscettibile di attacchi avversari che manipolano i sistemi di IA per interpretare o rappresentare erroneamente i dati in maniera intenzionale. Questa debolezza è particolarmente critica in alcuni tipi di analisi finanziaria, dove tali manipolazioni possono alterare gravemente il risultato analitico. In Leippold (2023), per esempio, ho studiato gli attacchi di "*sentiment spin*", in cui un modello sofisticato come GPT-3 può essere usato per modificare il sentiment dei testi finanziari, cambiando il sentiment negativo di una serie di frasi, per esempio, in neutro o positivo. Tali attacchi rappresentano un grave rischio per l'affidabilità dei sistemi automatizzati.

Hu et al. (2024) descrivono in dettaglio diverse altre strategie avversarie che sfruttano le vulnerabilità dei motori di ricerca generativi. La loro ricerca mostra come i motori di ricerca generativi, nonostante le loro capacità avanzate, possano essere ingannati da domande create appositamente per causare errori. Questa suscettibilità evidenzia l'urgente necessità di migliorare le misure di sicurezza dell'IA.

Le implicazioni degli attacchi avversari non si limitano alla finanza. I sistemi di IA possono essere manipolati per produrre risultati distorti o fuorvianti anche in altri settori, come quello sanitario, politico e farmaceutico. La regolamentazione dell'IA generativa dovrebbe concentrarsi sulla creazione di strutture che garantiscano trasparenza, responsabilità e correttezza. Man mano che l'uso dell'IA si espande, le autorità di regolamentazione

devono adottare linee guida rigorose per promuovere pratiche etiche e prevenire abusi. Tali norme proteggeranno gli utenti dagli effetti degli attacchi avversari e promuoveranno la fiducia nei sistemi di IA, aprendo la strada a una loro più ampia accettazione e integrazione nella società.

Sebbene esistano alcuni quadri di riferimento, come la legge europea sull'IA, c'è ancora molta strada da fare per elaborare normative complete ed efficaci. La formulazione di tali norme richiede un approccio pieno di sfumature, che tenga conto del ritmo rapido dei cambiamenti tecnologici e delle diverse applicazioni sempre in espansione dell'IA. È essenziale mantenere un dialogo costante tra accademici, leader dell'industria – in particolare dal settore tecnologico – e responsabili politici.



Conclusione

L'IA generativa presenta sia opportunità senza precedenti sia grandi sfide. In questa Public Discussion Note, ho delineato il suo potenziale di trasformazione per il settore finanziario, dal miglioramento del processo decisionale e della valutazione del rischio alla rivoluzione delle pratiche di sostenibilità aziendale. Tuttavia, il cammino verso la piena integrazione dell'IA nella finanza è irto di complessità, quali i problemi di allucinazione, verificabilità e vulnerabilità agli attacchi avversari, che potrebbero minare la fiducia e l'affidabilità dei sistemi di IA.

Gli ingenti investimenti già effettuati nell'IA sottolineano l'elevata posta in gioco. Secondo Goldman Sachs (2024), se da un lato è evidente il potenziale di significativi aumenti di produttività e di efficienza operativa, dall'altro vi sono opinioni contrastanti sull'impatto economico complessivo dell'IA. Esperti come Daron Acemoglu del MIT, tra gli altri, esprimono scetticismo, citando

timori secondo i quali la tecnologia potrebbe non giustificare i costi elevati, che potrebbero non diminuire come previsto (Acemoglu, 2023). D'altro canto, molte voci ottimiste ritengono che l'IA finirà per fornire rendimenti e benefici economici sostanziali e che le sue applicazioni più trasformative devono ancora emergere.

Come ricercatore in questo campo in evoluzione, vedo sia il potenziale sia le sfide che si prospettano per la sua applicazione in finanza. Il futuro dell'IA generativa nella finanza non è predeterminato, ma sarà plasmato dalle nostre azioni. Abbiamo un'opportunità senza precedenti di sfruttare l'IA per promuovere l'innovazione, aumentare la sostenibilità e migliorare l'efficienza del settore finanziario. Tuttavia, il raggiungimento di questi obiettivi richiede una realizzazione attenta e ponderata per mitigare i rischi posti da un'esuberanza incontrollata.



Bibliografia

- 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

Con il sostegno dei suoi fondatori – il settore bancario svizzero, la Confederazione svizzera e le principali università svizzere – lo Swiss Finance Institute (SFI) promuove in modo competitivo la ricerca e l'insegnamento di prim'ordine in ambito bancario e finanziario in Svizzera. Unendo l'eccellenza accademica all'esperienza pratica, l'SFI contribuisce al potenziamento della piazza finanziaria svizzera.

Editori e contatto

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

