ZANISTA

From Signals to Decisions: Scaling AI in Finance

Arman Khaledian, PhD
CEO & Co-founder, Zanista AI
arman.khaledian@zanista.ai

10 April 2026 | Virginia Tech
DLQF

Today's Journey

I. The Evolution of Trading & Markets

II. The Rise of Computing & AI - LLMs as Language interface

III. Where AI Performance Breaks Down or Expands! Data is Key

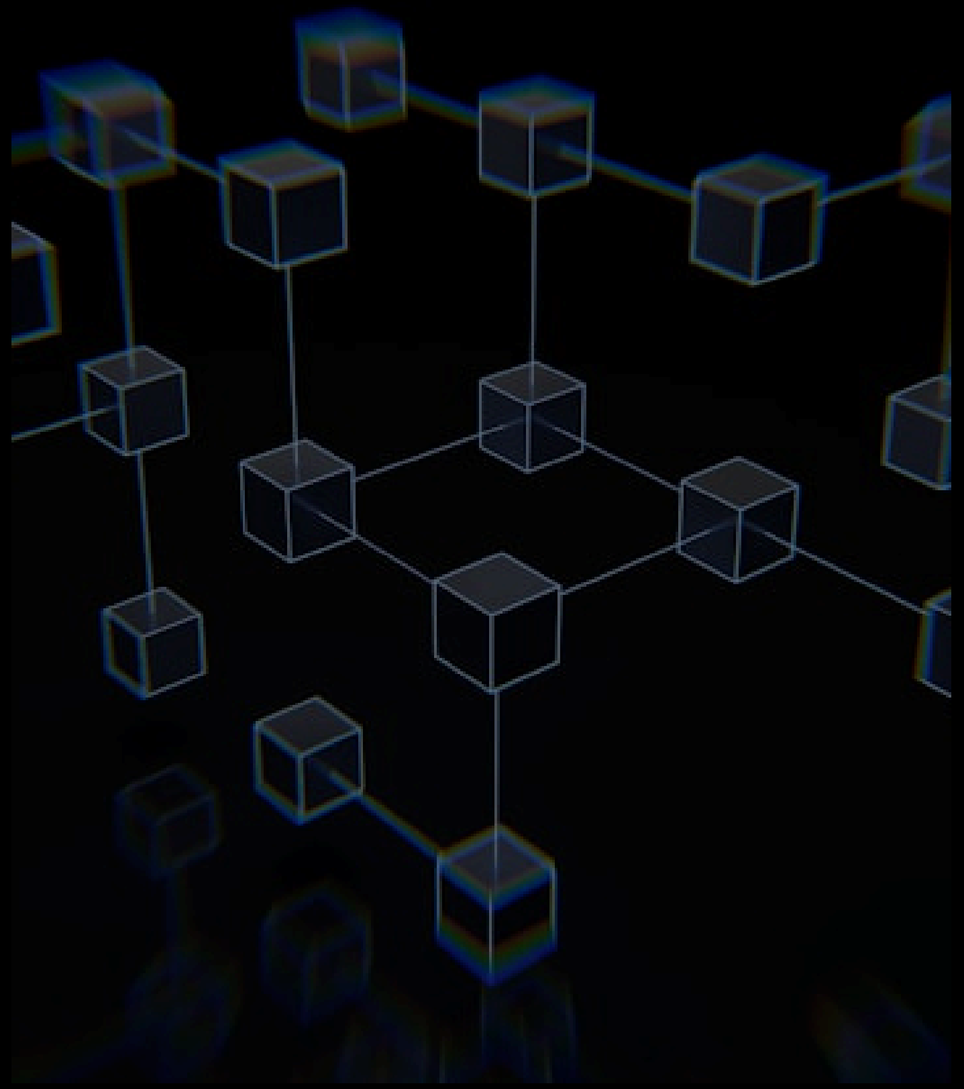
IV. LLMs as Signal Generators

V. From Signals to Decisions

VI. LLM Behavior Under Uncertainty

VII. The Five Sins of Financial LLM Evaluation

VIII. Key Takeaways & What Comes Next



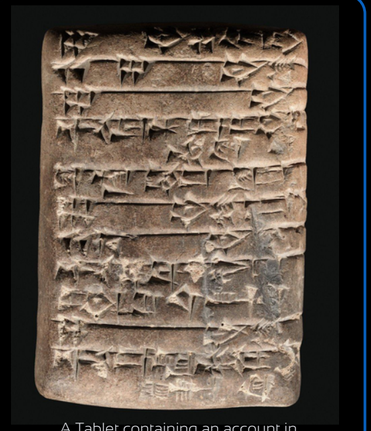
An Introduction to Trading

A Brief History

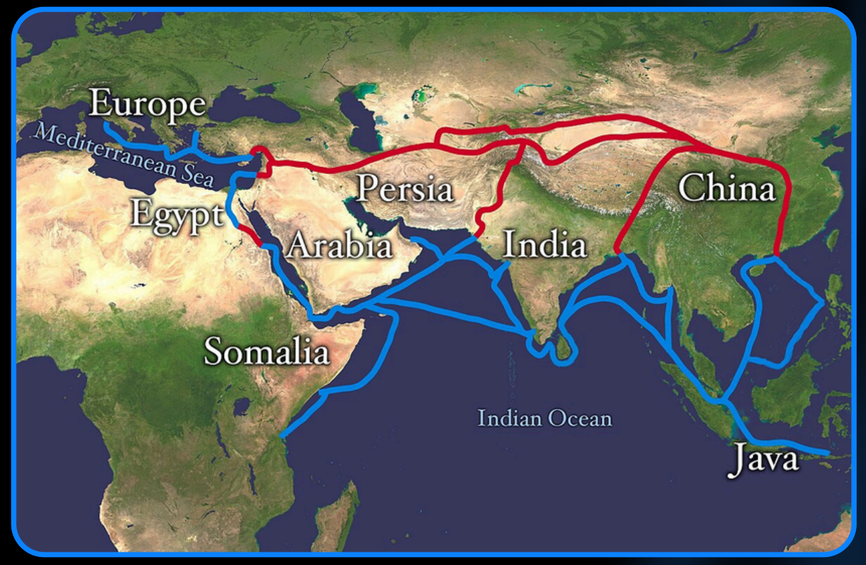
Trading, a form of communication as old as human civilization. Emerged as a by-product of specialization and exclusive access to resources. The oldest evidence of human trade is found in Mesopotamia (modern day Iraq) dating back to the 21st century BC to the Assyrians.

Early Trading Methods include:

- Long-distance trade **on foot**
- Use of domesticated **animals** (horses, camels) for transport
- **Barter systems** evolving into commodity money, shell money, and eventually coins and paper currency



A Tablet containing an account in Sumerian cuneiform describing the receipt of oxen

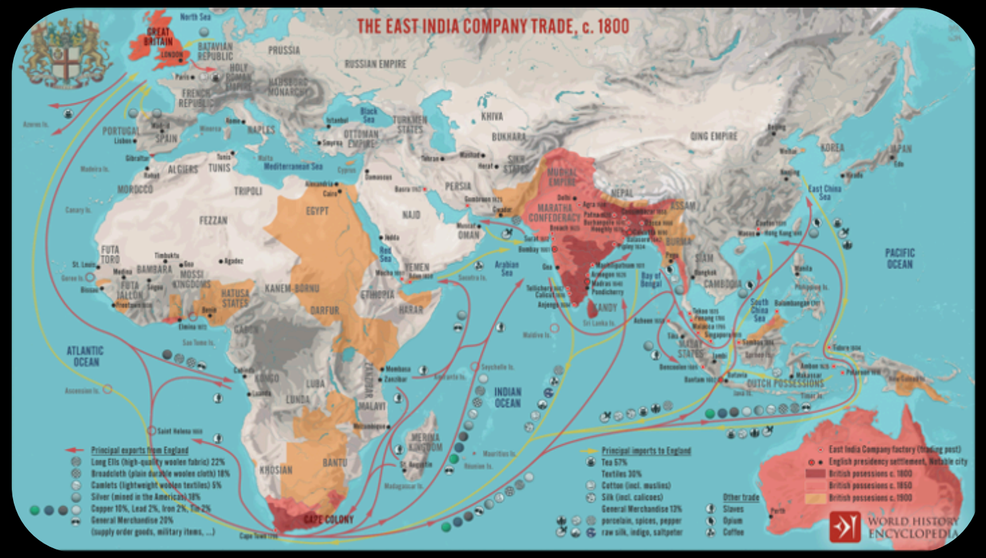


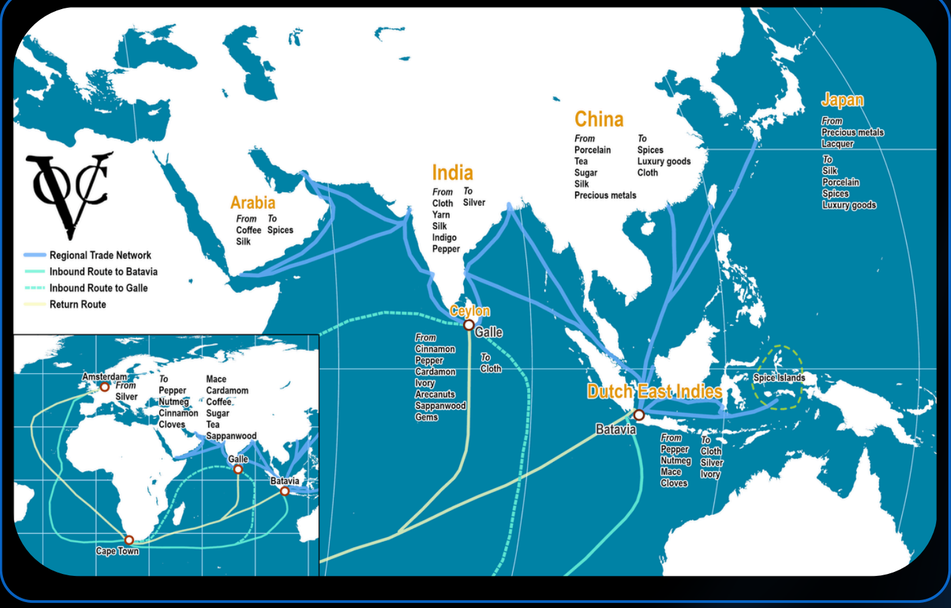
Silk Road

arman.khaledian@zanista.ai

Exchange and Capital Raising Through the Ages

Monopolies and Funds (1600–1780s): Rise of early joint-stock companies such as the Dutch East India Company (VOC), established in 1602, and the British East India Company (1600). These entities pioneered large-scale trade finance, issuing shares to investors and dominating maritime commerce.





arman.khaledian@zanista.ai

Exchange and Capital Raising Through the Ages

Globalization (1780s – Growth Until 1914): International trade expanded rapidly with improved transport and communication networks. In 1851, Paul Reuter (founder of Reuters) used a submarine telegraph cable beneath the English Channel to transmit market data between London and Paris, replacing his earlier network of carrier pigeons.

Great Reversal: Disruptions caused by WWI, the Great Depression (1929), and WWII halted globalization.

Return of Globalization (1980s–1990s): Deregulation, free-market policies, and technological advancements reignited global capital flows.

Boom in Electronic Markets (2000s): Rise of fully digital trading platforms, high-frequency trading, and algorithmic strategies.

Reversal ? (2025): Rise of Tarrifs and nations shifting to self reliance in the midst of global chaos and uncertainty may cause a reversal

arman.khaledian@zanista.ai

Just as social media democratized the attention economy, AI is now democratizing knowledge and knowledge-based expertise.

Example: From Traders to Machines to Models

Early 2000s: electronic trading crushed manual execution. Costs and commissions collapsed. Execution desks shrank by ~90%. Traders didn't all vanish, but most had to shift. Some moved to quant, risk, or ops. Some stayed in prop or niche trading. Human execution became marginal. Commissions collapsed 10x.

Now: AI is doing the same to quants and developers. Models write code, test ideas, and take over repetitive work.

Pattern:
Routine work gets automated.
Fewer people do more, at higher leverage.

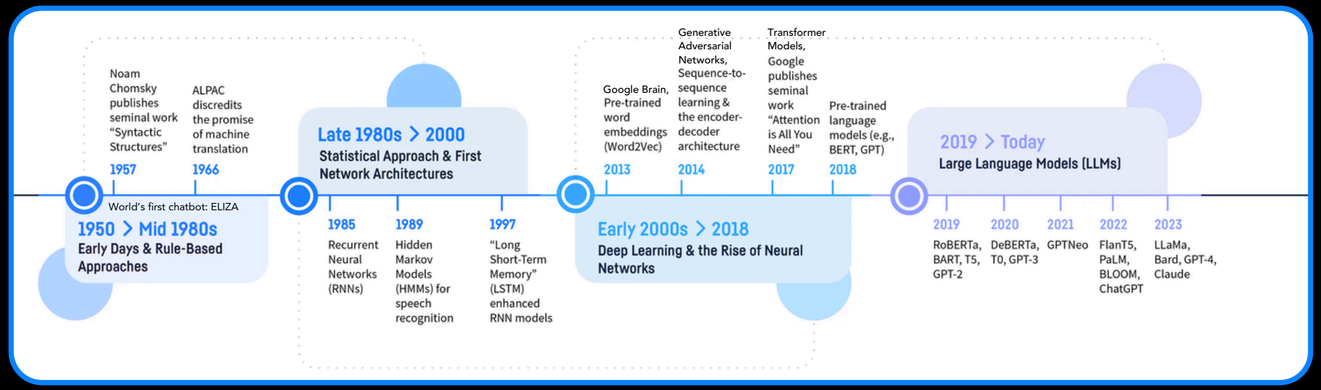
Lesson:
Old edge = doing the work fast.
New edge = building the system that does it.

arman.khaledian@zanista.ai

The rise of Smart Machines

Over the past couple of decades computers have advanced in a way that it has reached the point of rivaling human intelligence. This revolution is called artificial intelligence.

Artificial Intelligence (AI): computer systems dedicated to performing tasks close to human intelligence, such as speech recognition, language translation, and visual perception. It is the intelligence of software as opposed to the intelligence of humans.

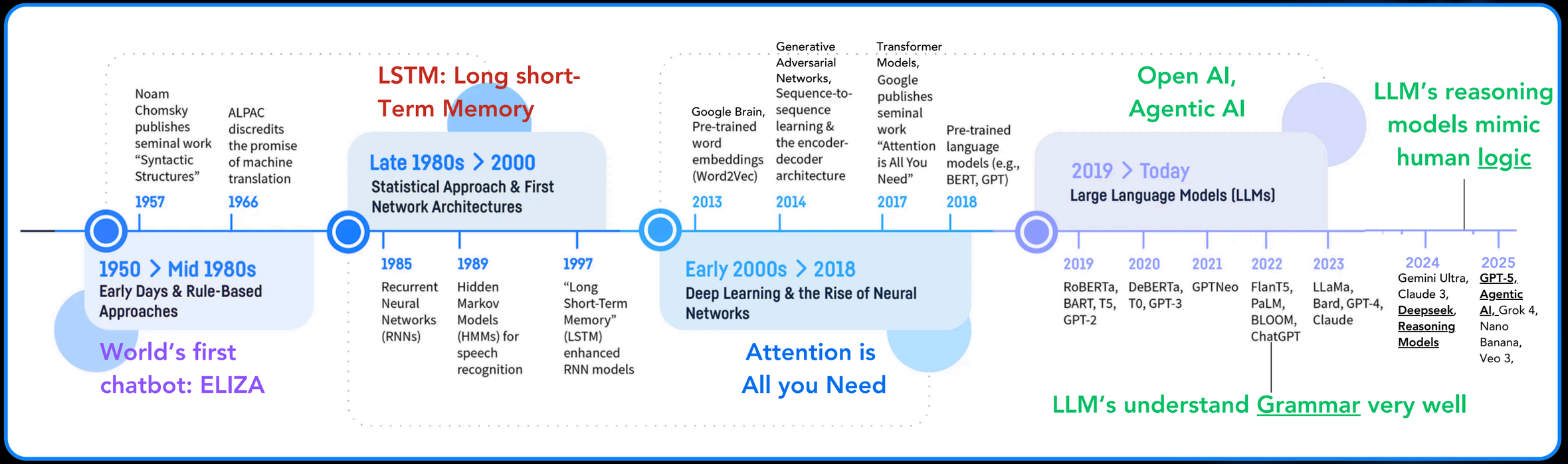


Language AI refers to a subfield of AI that focuses on developing technologies capable of understanding, processing, and generating human language. The term Language AI can often be used interchangeably with natural language processing (NLP) with the continued success of machine learning methods in tackling language processing problems. We use the term Language AI to encompass technologies that technically might not be LLMs but still have a significant impact on the field.

arman.khaledian@zanista.ai

The rise of Smart Machines
LLMs as Language interface - new PC/iphone moment

Over the past couple of decades computers have advanced in a way that it has reached the point of rivaling human intelligence. This revolution is called artificial intelligence.



arman.khaledian@zanista.ai

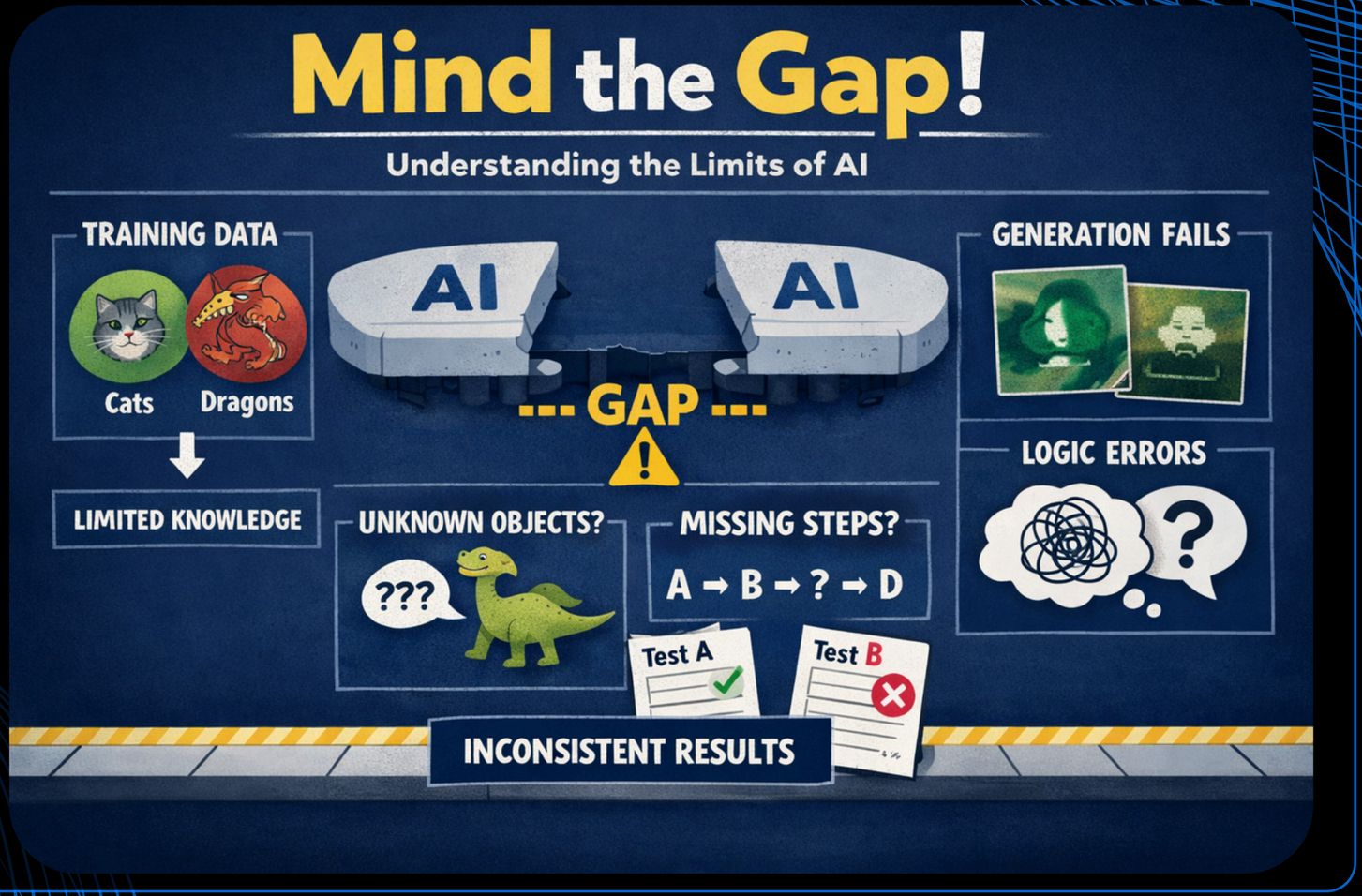
Where AI performance breaks down

AI models do not fail everywhere. They fail in specific gaps.

They perform well on tasks and data they have seen, but struggle when examples, steps, or reasoning patterns are missing. This often appears as **hallucinations**: confident outputs that are incorrect or incomplete.

You find these gaps through experimentation. You fill them by adding targeted examples and by breaking complex tasks into smaller, explicit steps. This reduces missing reasoning and makes the model aware of what it does not know.

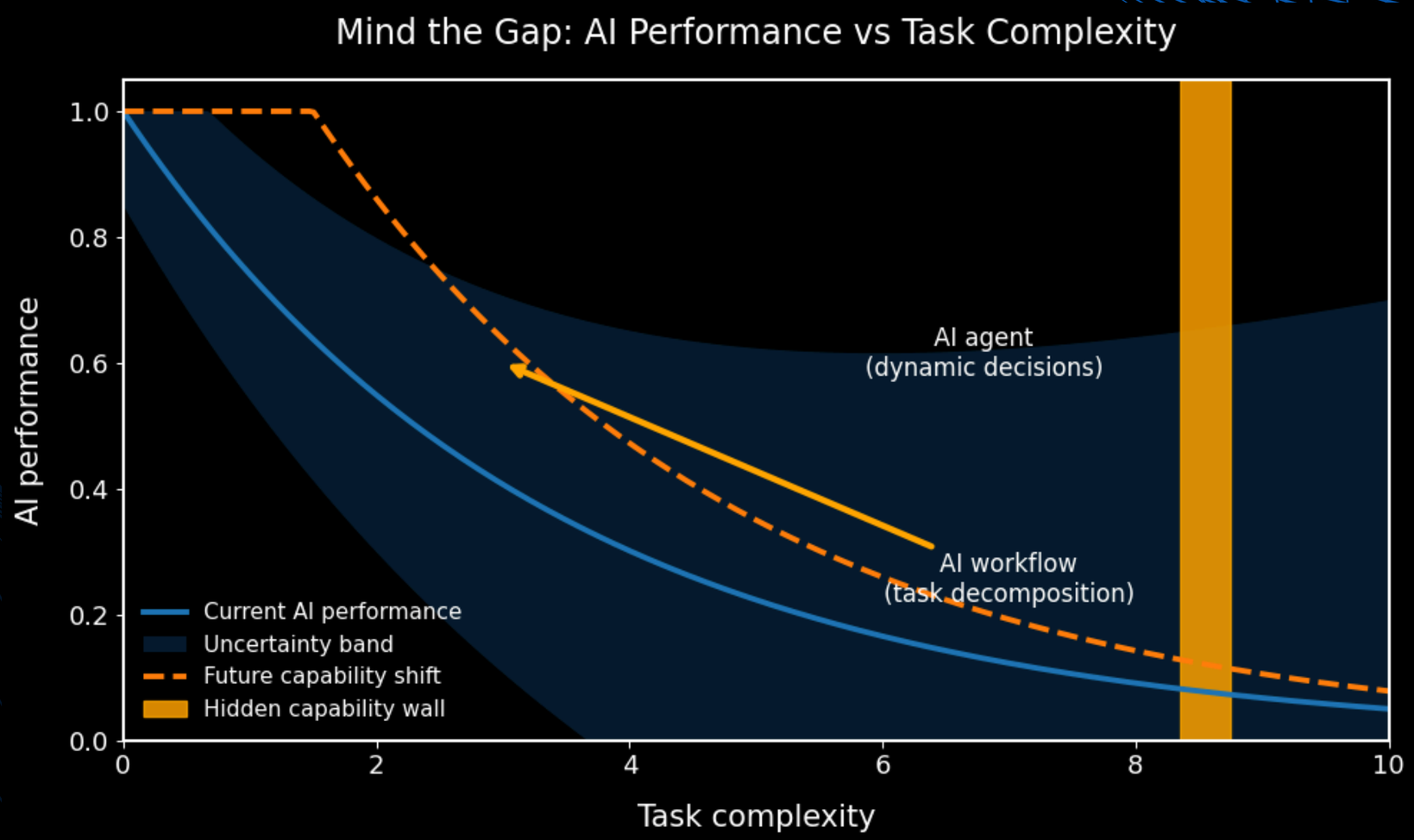
Mind the gap between apparent capability and true understanding.



In the opposite side - The real superpower is this: elite data from elite experts. Example: If we can capture the reasoning patterns, intuition, and niche knowledge of top mathematicians, we can fine tune an LLM into something that has never existed before, a single coherent mathematical intelligence built from many elite minds. Instead of rare expertise being fragmented across individuals, it becomes unified in one system. That is super math. That is a super mathematician. Same in Coding.

Where AI performance breaks down

- AI performance drops as task complexity increases, with growing uncertainty (Gap) and hallucinations.
- Model improvements shift capability to the right, but performance decay remains.
- Some complexity levels form a hidden capability wall that models may not reliably cross.
- **AI workflows (Algorithm)** keep humans in control. Tasks are manually broken into smaller steps with explicit logic to stay in high-reliability regions.
- **AI agents** automate this process. They decide how to decompose tasks, call tools, and act autonomously, but still face uncertainty and limits.

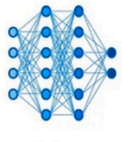


Chaos midst simplicity

After training: We can generate text by predicting one word at a time

A trained language model can

Input

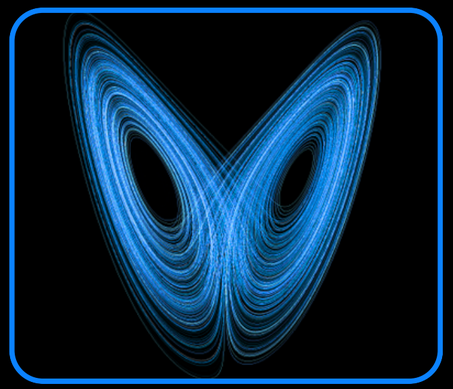


LLM

Word	Probability
speak	0.065
generate	**0.072**
politics	0.001
...	...
walk	0.003

Output at step 1

LLMs are an example of what's called "Generative AI"

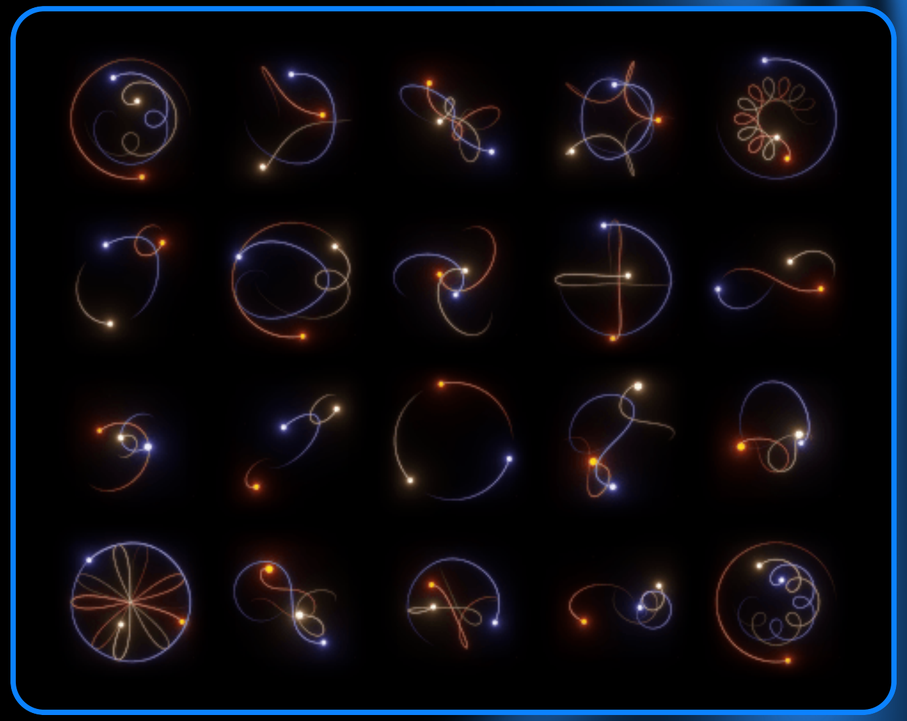


The next token prediction is very simple in principle but the results are too complex to comprehend

Word	Probability
ability	0.002
text	0.084
coherent	**0.085**
...	...
ideas	0.041

Output at step 2

Three body problem rises from very simple rules of gravity



arman.khaledian@zanista.ai

Financial LLM Applications

LINGUISTIC TASKS

Linguistic tasks harness the power of AI to process vast amounts of unstructured text data, enabling better decision-making in finance.

Refs: Lee et al. (2025) FinLLMs; Bridging Finance and AI Survey (2025); From Deep Learning to LLMs (2025)

SENTIMENT ANALYSIS

Sentiment analysis utilizes AI to gauge market sentiment from news articles and social media, providing valuable insights for traders and investors.

TIME SERIES REASONING

Time series reasoning involves analyzing historical financial data, helping predict future trends and inform investment strategies through advanced AI models.

Harnessing LLMs for Sentiment Analysis

SCALING INSIGHTS

Large Language Models (LLMs) enhance sentiment analysis by generating nuanced insights from unstructured data, transforming how financial narratives are interpreted and acted upon.

Refs: FinSentLLM arxiv.org/abs/2509.12638 | ChatGPT and the Stock Market (2024) | What is Financial Sentiment? (2025)

Beyond Sentiment Analysis

STRUCTURED EVENTS

Structured Event Representation focuses on extracting critical details about who did what to whom in financial contexts.

KNOWLEDGE GRAPHS

Knowledge graphs leverage LLMs to extract enterprise relations, enhancing our ability to visualize complex interconnections in finance.

REASONING MODELS

Fin R1 and Fino1 are advanced reasoning models that implement a chain of thought specifically tailored for finance-related predictions and analyses.

Refs: Fin-R1 (2025) | Fino1 arxiv.org/abs/2502.08127 | Knowledge Graph-Driven Forecasting (2025) | Structured Event Representation (2025)

Integrating AI Insights into Portfolio Management

AI-DRIVEN DECISION MAKING

Leveraging LLM insights enhances portfolio optimization, allowing for smarter investment strategies that adapt to market changes and investor goals.

Refs: Black-Litterman + LLM Views (2025); HARLF (2025); Decision-informed NNs (2025); Sentiment-augmented RL (2025)

Advancements in Multi-Agent Trading Systems

TRADING AGENTS COLLABORATION

Multi-agent systems enhance trading by allowing agents to collaborate, analyze market conditions, and execute trades based on shared intelligence and strategies.

Refs: TradingAgents arxiv.org/abs/2412.20138 | MASS (2025) | MASFIN arxiv.org/abs/2512.21878 | Trading-R1 arxiv.org/abs/2509.11420

Foundation Models Transforming Financial Time Series

REVOLUTIONIZING DATA ANALYSIS

Financial time series data requires **specialized models** to address its unique characteristics, including noise, non-stationarity, and regime changes.

Refs: Kronos arxiv.org/abs/2508.02739 | FinCast (2025) | Re(Visiting) TSFMs (2025) | Dual Adaptation ICML 2025

Integrating News and Price Data for Insights

MULTIMODAL DATA FUSION

This approach enhances forecasting by combining textual news data with price movements, enabling better volatility predictions and informed trading decisions.

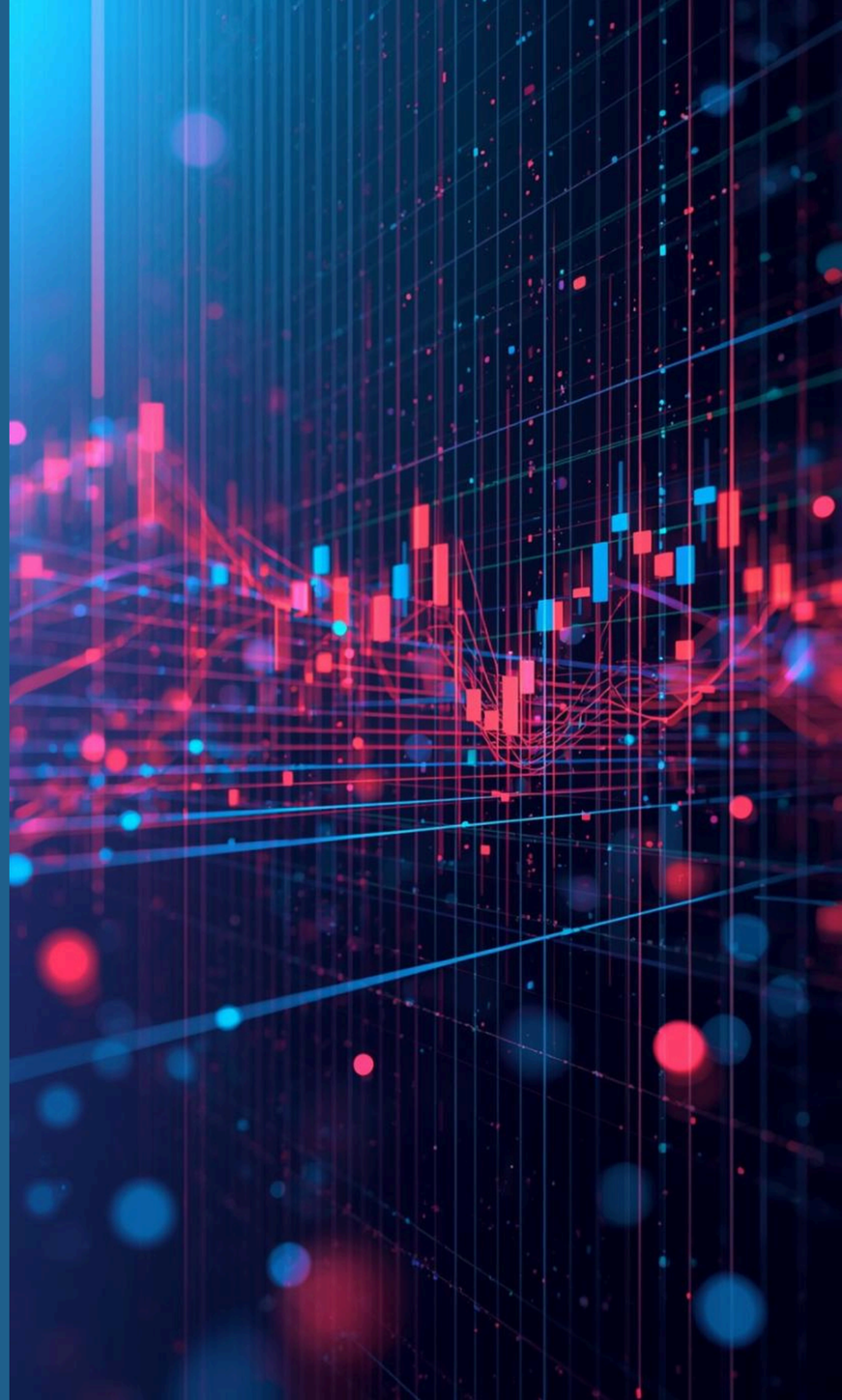
Refs: Cross-Modal Temporal Fusion (2025); Multimodal Language Models (2025); FinMultiTime (2025)

Understanding Market Microstructure Through AI

VISUALIZING LIMIT ORDER BOOKS

AI models analyze **market microstructure**, optimizing trade execution by interpreting limit order books, enhancing liquidity and pricing efficiency in real-time markets.

Refs: TradeFM from JPMorgan (2026) | LOBERT arxiv.org/abs/2511.12563 | TLOB arxiv.org/abs/2502.15757 | Deep LOB Briola et al. (2025)



Understanding LLM Behavior Under Uncertainty

EVALUATION FRAMEWORK

This framework characterizes LLM performance across risk, ambiguity, and paradox, identifying crucial dimensions for effective deployment in finance.

Refs: Financial Brain Scan arxiv.org/abs/2508.21285 | LLM Knows Geometry (2025) | Kong et al. arxiv.org/abs/2602.14233 | Risk, Ambiguity, and Infinity: Behavioral Signatures of Modern Large Language Models (2025)

Ambiguity and Paradox in LLMs

ELLSBERG'S URN

GPT-4 consistently avoids selecting from an ambiguous urn, showcasing its **risk-averse behavior** in uncertain scenarios and decision-making processes.

ST. PETERSBURG GAME

In the St. Petersburg paradox, GPT-4 aligns closely with human baselines, demonstrating its ability to engage effectively in complex decision-making challenges.

Refs: Risk, Ambiguity, and Infinity: Behavioral Signatures of Modern Large Language Models (2025); A Financial Brain Scan of the LLM (2025); When LLMs Go Abroad: Foreign Bias (2025)

The Five Sins

Kong et al. (2026): 164 papers reviewed. No single bias in >28% of studies. Biases compound to create an illusion of validity.

LOOK AHEAD

Look Ahead Bias occurs when future information inadvertently influences model predictions, leading to overly optimistic results and flawed evaluations.

SURVIVORSHIP

Survivorship Bias excludes failed firms from analysis, creating a misleading view of performance by only considering successful entities in the dataset.

NARRATIVE, OBJECTIVE & COST

3. Narrative Bias: LLM generates plausible but ungrounded causal stories.

4. Objective Bias: Training rewards confidence over calibrated uncertainty.

5. Cost Bias: Ignores latency, inference cost, and trading friction.

Refs: Kong et al. arxiv.org/abs/2602.14233

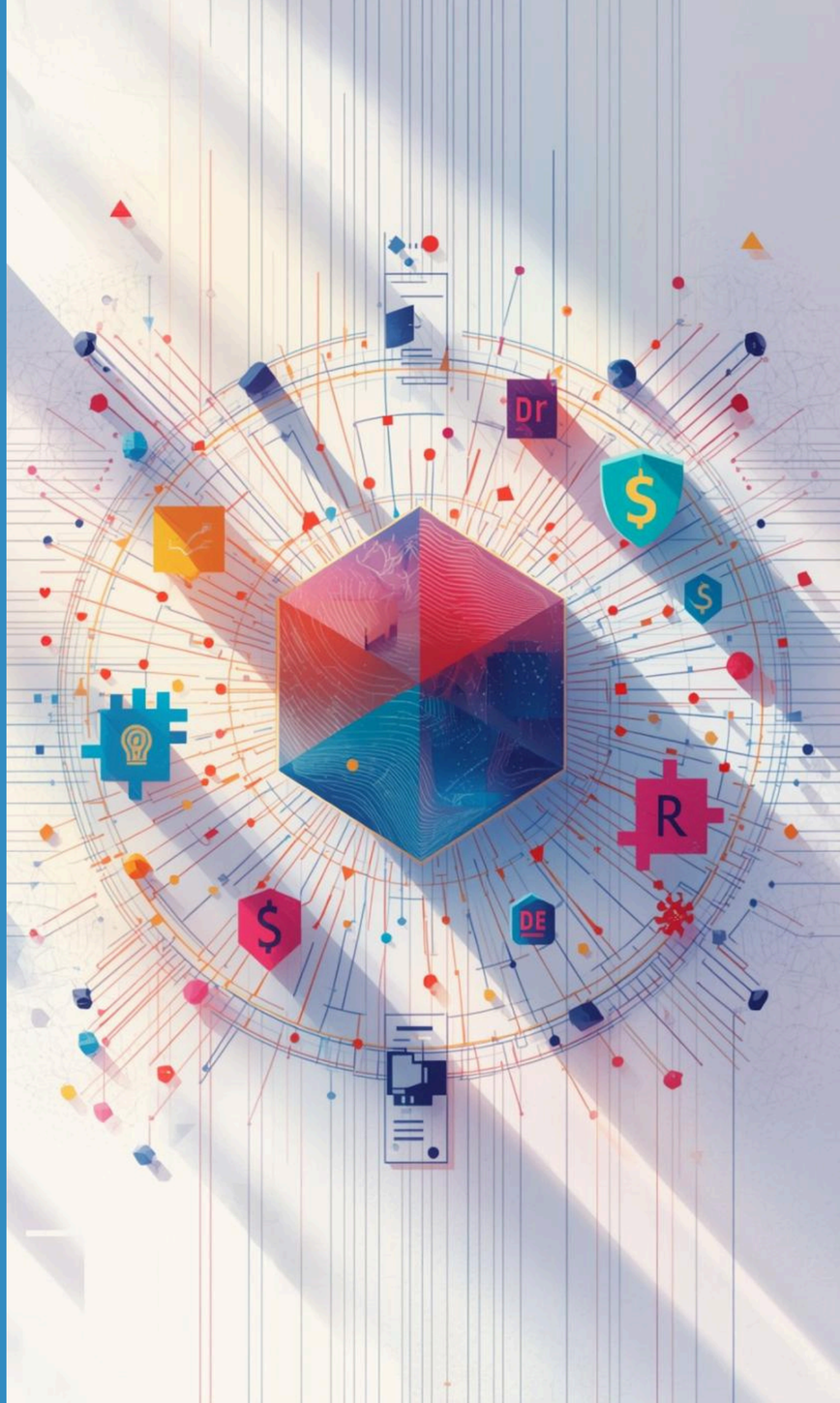
STRUCTURAL VALIDITY FRAMEWORK: FIVE PILLARS FOR DEPLOYMENT READINESS. 76% OF PRACTITIONERS SAY EVALUATION TOOLS ARE SCARCE.
REFS: KONG ET AL. ARXIV.ORG/ABS/2602.14233 | AI AGENT MISINFORMATION (2025) | SOCIAL GROUP BIAS (2025) | FROM TASKS TO TEAMS (2025)

Trust and Regulation in Financial AI

ENSURING ETHICAL PRACTICES

Building trust in AI for finance requires **robust regulations** and a solid framework that prioritizes ethical considerations and transparency in decision-making.

Refs: The Agentic Regulator (2025); CFA Explainable AI Report (2025); BIS AI Explainability Paper (2025); US Treasury AI Risk Framework (2026)

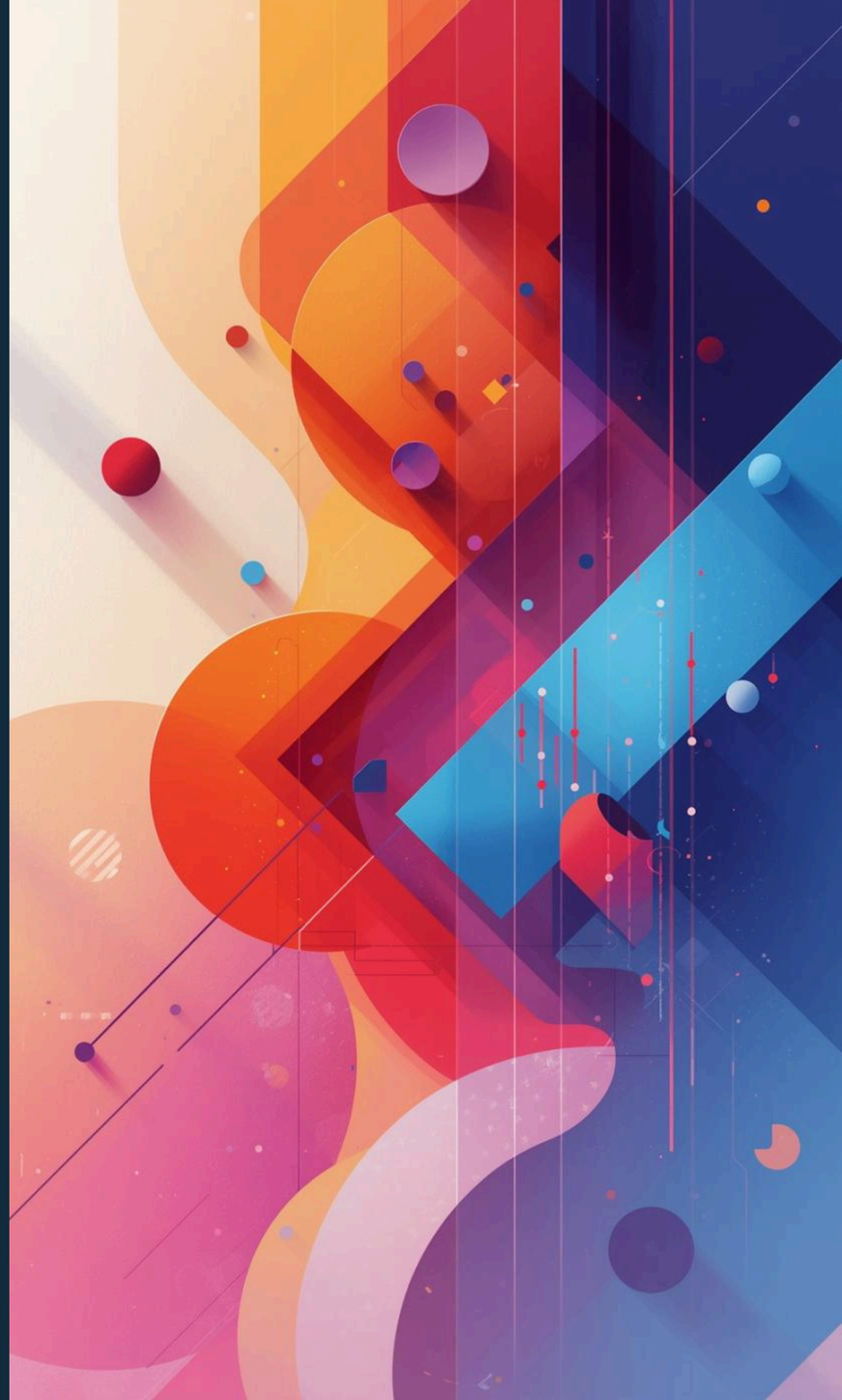


The Complexity of Prediction in Finance

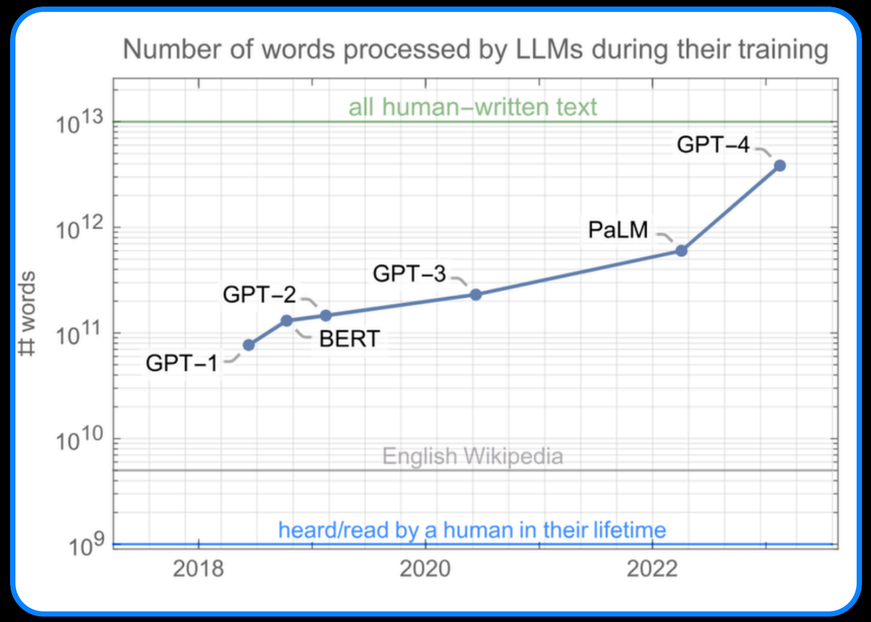
ACCURACY VS. PROFIT

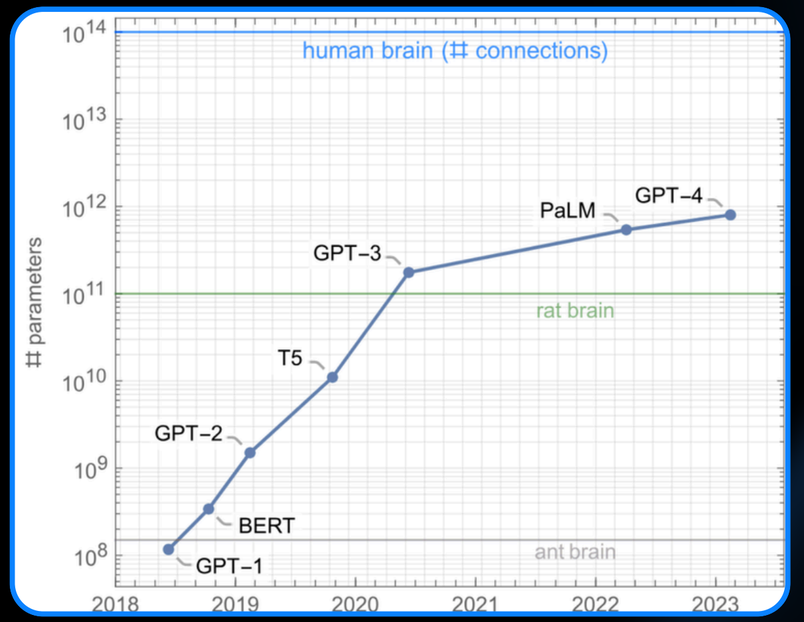
Understanding that **prediction accuracy does not guarantee profitability** is crucial in finance, as chaotic market behavior often leads to unexpected outcomes.

Refs: LiveTradeBench (2025); TradeTrap (2025); Can LLM-based Strategies Outperform? (2025)



LLM Behavior Under Uncertainty





arman.khaledian@zanista.ai

NewsWitch from Zanista

80% of financial data is unstructured: news, filings, calls, social media. LLMs unlock this layer.

Evolution: dictionary methods → Vector Embeddings → GPT sentiment. Dense embeddings capture entity relationships, narrative framing, semantic drift.

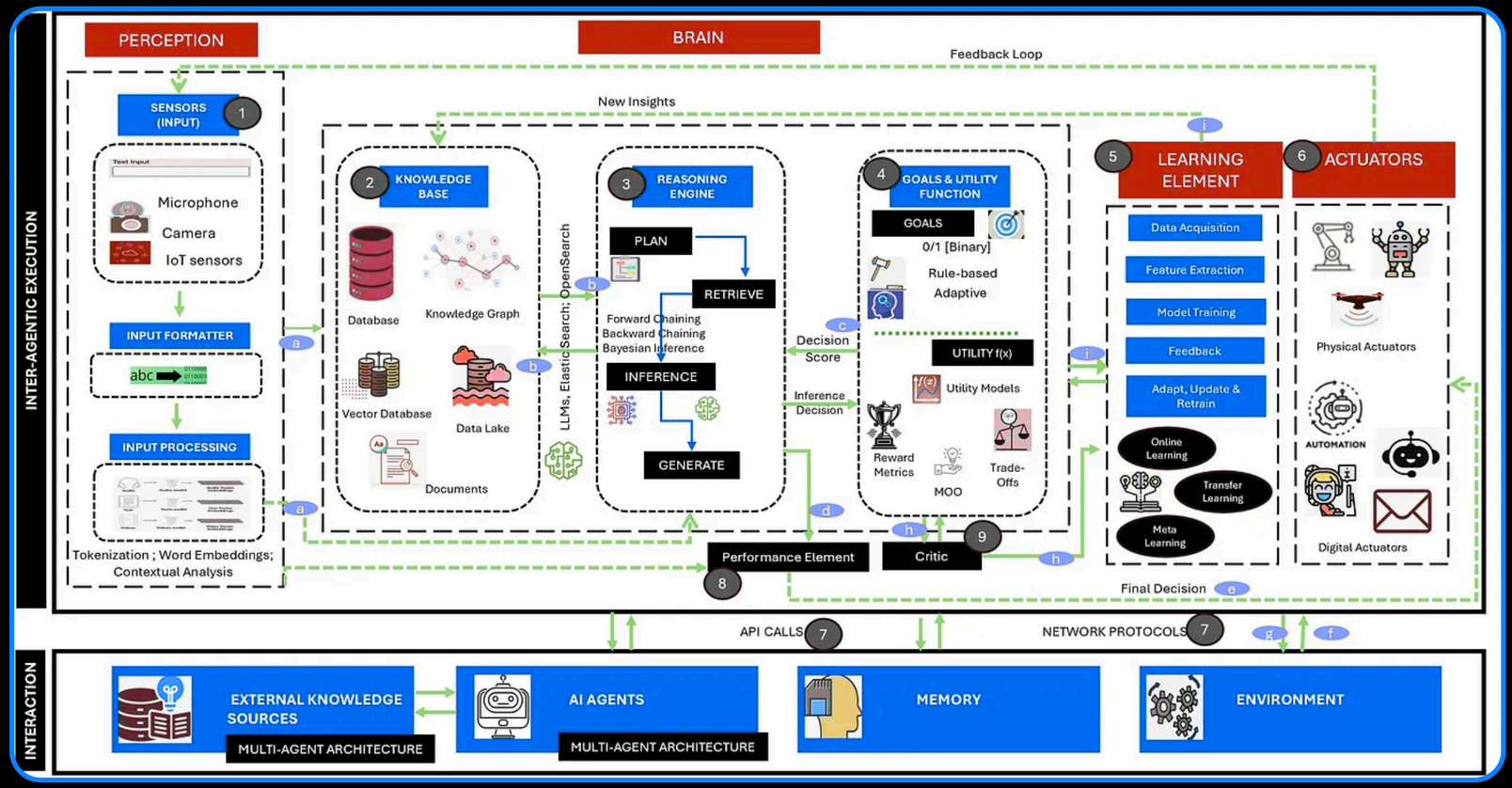
NewsWitch: ~1M articles/day, RAG pipelines + PCA for scalability.

Refs: https://zanista.ai/products/newswitch

arman.khaledian@zanista.ai

From Signals to Decisions: Multi-Agent Trading

TradingAgents (UCLA/MIT 2025): analyst agents collaborate via structured debate



Multi-Agent Systems
MASS, MASFIN: Decomposed financial reasoning, each agent specializes.

Signals to Portfolios
Black-Litterman + LLM views. HARLF: Hierarchical RL with sentiment. Hybrid: sentiment + macro + technical.

Prediction accuracy ≠ trading profit. The gap is non-trivial.

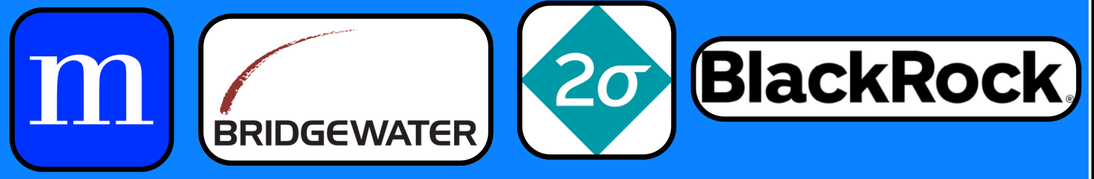
Refs: TradingAgents arxiv.org/abs/2412.20138 | MASS (2025) | MASFIN arxiv.org/abs/2512.21878 | BL+LLM (2025) | HARLF (2025) | Trading-RI arxiv.org/abs/2509.11420

arman.khaledian@zanista.ai

Market Participants

Who's Who in Finance

Buy Side: These are the people and firms who buy investments to grow their money.



Long-Only Investors: Mutual funds, pension funds, Vanguard, Fidelity. They pool money from regular people and invest it.

Wealth Managers / Institutions: Big banks like UBS or Merrill Lynch. They help wealthy families and individuals manage and grow their wealth.

Hedge Funds (Long-Short): Like Citadel or Point72. They use more complex strategies, betting both on prices going up and down.

Sell Side: These are the middlemen who help buy-side clients trade.

- **Broker**: Trades on behalf of clients.
- **Dealer**: Trades with their own money, providing liquidity (used to be a major revenue source before regulations like Dodd-Frank limited them).

arman.khaledian@zanista.ai

Market Participants

Who's Who in Finance

Market Makers & High-Frequency Traders

These players make markets work smoothly by ensuring there's always someone to buy or sell.

- **Market Makers**: Provide liquidity (always ready to buy or sell). Sometimes they also trade for profit.
- **High-Frequency Traders:** Use super-fast computers and networks to trade in fractions of a second.

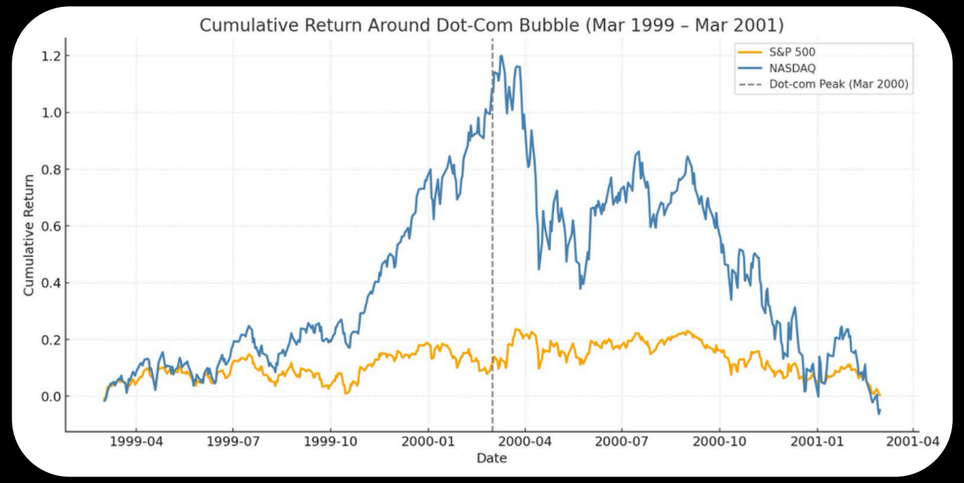
Can also connect with retail trading apps like Robinhood or E*Trade to aggregate liquidity.

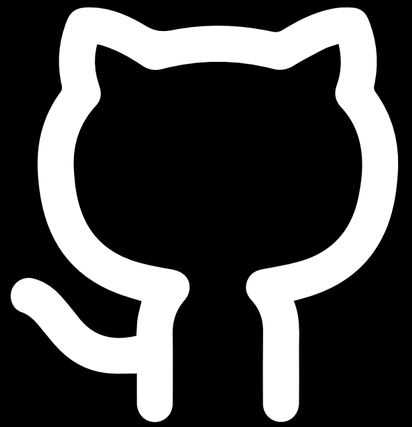
How It All Connects:

- Corporations issue shares or bonds (they need money).
- Sell Side (brokers/dealers) helps move those securities to investors.
- Buy Side (investors) provide the money in exchange for shares/bonds.
- Market Makers/HFTs keep the system running smoothly by making sure there's always someone to buy or sell at any moment.

arman.khaledian@zanista.ai

AI bubble?





arman.khaledian@zanista.ai

Key Takeaways

DATA ACCESS

LLMs enable access to **80 percent** of unstructured financial data, enhancing insights and decision-making capabilities in the finance sector.

Refs: FinSentLLM arxiv.org/abs/2509.12638 | LLMs in Equity Markets doi.org/10.3389/frai.2025.1608365

ENGINEERING CHALLENGES

Transitioning from signals to decisions requires **careful engineering** to ensure effective integration of AI models into traditional financial workflows.

Refs: Bridging Finance and AI Survey (2025); From Deep Learning to LLMs (2025)

DOMAIN SPECIFICITY

Foundation models need **domain specific training** to accurately reflect the unique characteristics and dynamics of financial data for optimal performance.

Key Takeaways

BIASES

LLMs exhibit measurable **behavioral biases** that can impact decision-making and reliability in financial contexts, leading to suboptimal outcomes.

Refs: Financial Brain Scan arxiv.org/abs/2508.21285 | When LLMs Go Abroad (2025) Kong et al. arxiv.org/abs/2602.14233

RESEARCH VALIDITY

The five sins of evaluation **undermine research validity**, affecting the credibility of results and findings within the financial AI landscape.

PROFITABILITY

Prediction accuracy **does not guarantee profit**, highlighting the complexity of market behavior and the challenges of translating insights into successful trades.

Key Takeaways

OVERSIGHT

Human oversight is **essential** in AI-driven finance to ensure ethical decisions and mitigate risks associated with automated trading systems.

Refs: The Agentic Regulator (2025); US Treasury AI Risk Framework (2026)

PREDICTION

Prediction accuracy does not **automatically** translate to trading profit; success requires a deep understanding of market dynamics and behavior.

CHAOS

Simple rules can produce **chaotic behavior** in markets and LLMs, complicating the relationship between signals and actionable trading decisions.

Selected References

LLM Signals, Sentiment & Trading

OUR WORK
Khaledian et al. Increase Alpha: AI-Driven Trading arxiv.org/abs/2509.16707
Khaledian et al. PCA-RAG: Efficient RAG arxiv.org/abs/2504.08386
Kargarzadeh (2024) LLM Trading Strategy (Imperial College London)
Ghafouri, Khaledian et al. Risk, Ambiguity, and Infinity (ACM ICAIF 2025)
ssrn.com/abstract=5688064

LANDMARK PAPERS
Chen, Kelly & Xiu. Expected Returns and LLMs (Chicago/Yale, RFS 2026)
ssrn.com/abstract=4416687
Kim, Muhn & Nikolaev. Financial Statement Analysis with LLMs (Chicago Booth)
arxiv.org/abs/2407.17866
Lopez-Lira & Tang. Can ChatGPT Forecast Stock Prices? (2023) arxiv.org/abs/2304.07619
Glasserman & Lin. Look-Ahead Bias in GPT (Columbia) arxiv.org/abs/2309.17322
Lopez-Lira. Can LLMs Trade? Market Simulations (2025) arxiv.org/abs/2504.10789

SENTIMENT & RETURN PREDICTION
FinSentLLM: Multi-LLM Sentiment Forecasting arxiv.org/abs/2509.12638
Kirtac & Germano. Sentiment Trading with LLMs (LSE, FRL 2024)
arxiv.org/abs/2412.19245
ChatGPT and the Stock Market (2024)
One News Two Markets (2025)
Enhancing Trading Performance Through Sentiment Analysis (2025)
What is Financial Sentiment? (2025)

ALPHA MINING & STOCK ANALYSIS
Alpha-GPT: Human-AI Interactive Alpha Mining (HKUST/IDEA) arxiv.org/abs/2308.00016
MarketSenseAI 2.0: LLM Agents for Stock Analysis (2025) arxiv.org/abs/2502.00415

MULTI-AGENT & PORTFOLIO
TradingAgents, Xiao et al. UCLA/MIT arxiv.org/abs/2412.20138
Trading-R1: LLM Reasoning via RL arxiv.org/abs/2509.11420
MASFIN arxiv.org/abs/2512.21878
MASS: Multi-Agent Simulation for Portfolio (2025)
FinRobot: Open-Source AI Agent Platform arxiv.org/abs/2405.14767
FinCon: Multi-Agent with Verbal Reinforcement arxiv.org/abs/2407.06567
FinMem: LLM Agent with Layered Memory (AAAI 2024) arxiv.org/abs/2311.13743
TwinMarket: Multi-Agent Market Simulation arxiv.org/abs/2502.01506
FCLAgent: Agent-Based Financial Market Simulation (2025) arxiv.org/abs/2510.12189
LLM Agents Do Not Replicate Human Traders (2025) arxiv.org/abs/2502.15800
Black-Litterman + LLM Views (2025)
HARLF: Hierarchical RL with Sentiment (2025)

Selected References

Foundation Models, Microstructure & Reasoning

DOMAIN-SPECIFIC LLMs
BloombergGPT: 50B param model for finance (Bloomberg) arxiv.org/abs/2303.17564
FinGPT: Open-Source Financial LLMs arxiv.org/abs/2306.06031
FinBERT: Extracting Info from Financial Text (Huang et al. CAR 2023)
Open-FinLLMs: Multimodal Financial LLMs (Huang et al. 2024) arxiv.org/abs/2408.11878
FinTral: GPT-4-Level Multimodal Financial LLMs (Bhatia et al. 2024) arxiv.org/abs/2402.10986
InvestLM: LLM for Investment (Yang et al. 2023) arxiv.org/abs/2309.13064
PIXIU/FinMA: Financial LLM & Benchmark (Xie et al. NeurIPS 2023) arxiv.org/abs/2306.05443

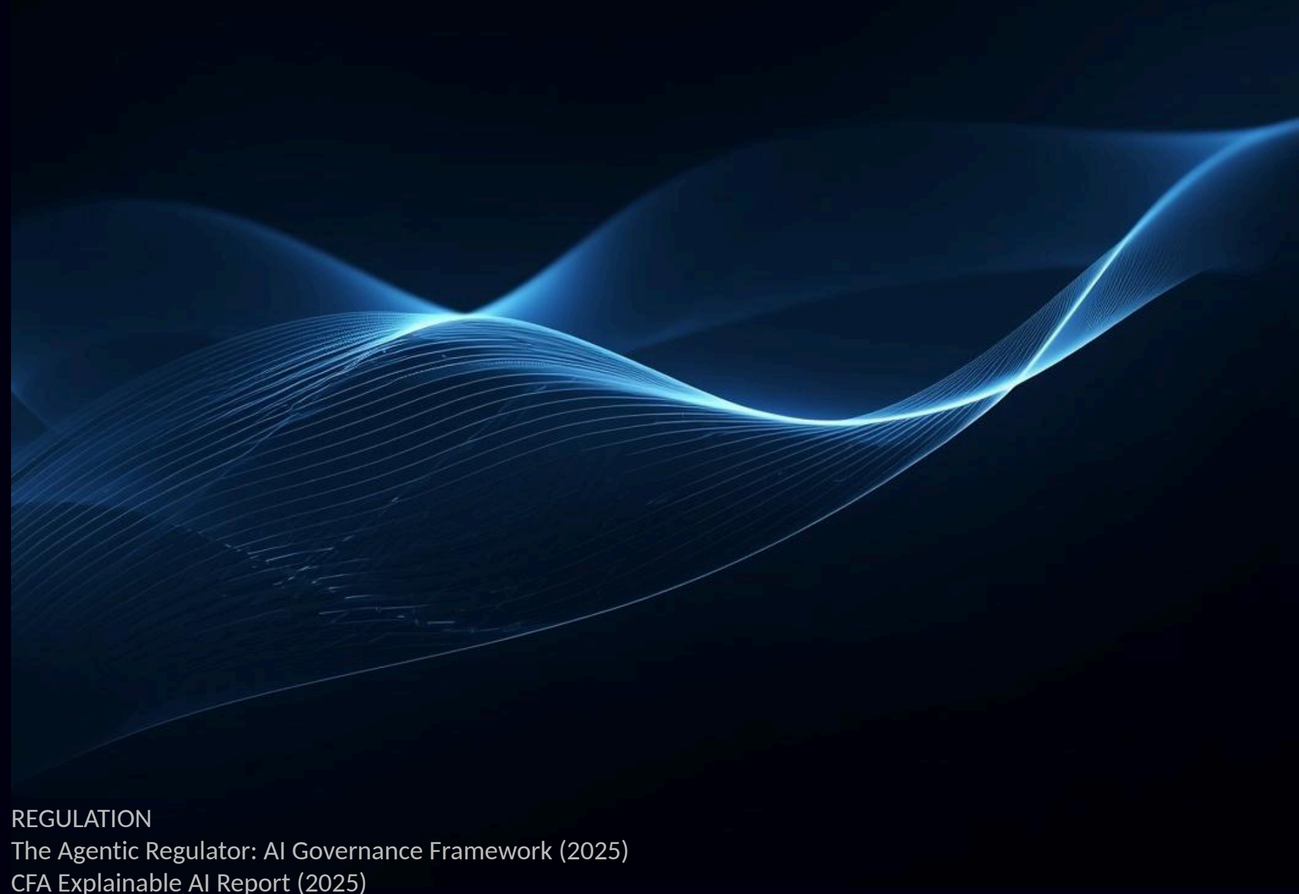
FOUNDATION MODELS FOR FINANCIAL TIME SERIES
Kronos: Foundation Model for Financial Markets arxiv.org/abs/2508.02739
FinCast: Foundation Model for Financial Forecasting (2025)
Re(Visiting) Time Series Foundation Models in Finance (2025)
Dual Adaptation of TSFMs, FMSD @ ICML 2025
Advancing Financial Engineering with Foundation Models (2025)
Bybee, Kelly, Manela & Xiu. Business News and Business Cycles (JoF 2024)
Temporal Data Meets LLM: Explainable TS Forecasting (Yu et al. 2023) arxiv.org/abs/2306.11025
Time Machine GPT: Nonprognosticative LLMs (Oxford, 2024) arxiv.org/abs/2404.18543
FinSeer: RAG for Financial Time Series (2025) arxiv.org/abs/2502.05878

DMARKET MICROSTRUCTURE
TradeFM: Foundation Model for Trade Flow (2025)
LOBERT: Foundation Model for LOB Messages arxiv.org/abs/2511.12563
TLOB: Transformer with Dual Attention for LOB arxiv.org/abs/2502.15757
Deep LOB Forecasting, Briola et al. (Quantitative Finance, 2025)
HLOB: Info Persistence in LOBs (Expert Systems, 2024)

REASONING MODELS
Fin-R1: Financial Reasoning through RL (2025)
Fino1: Transferability of Reasoning LLMs arxiv.org/abs/2502.08127
Reasoning or Overthinking: LLMs on Sentiment (ACM ICAIF 2025)
Knowledge Graph-Driven Financial Forecasting (2025)
Structured Event Representation for Stock Predictability (2025)

Evaluation, Bias & Regulation

Evaluation, Bias, Regulation & Surveys



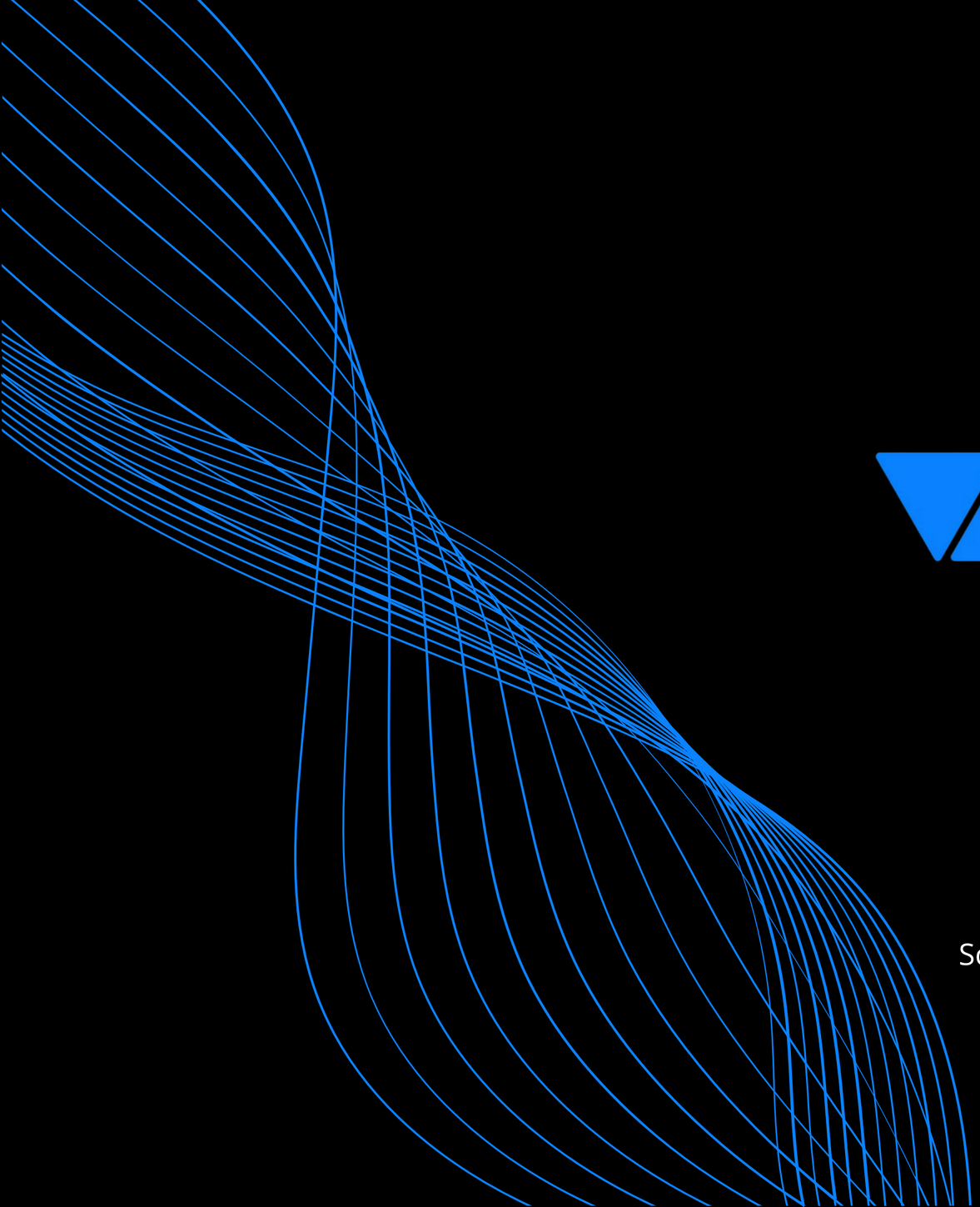
THE FIVE SINS (Kong et al. 2026) arxiv.org/abs/2602.14233
164 papers reviewed. Five compounding biases: Look-Ahead, Survivorship, Narrative, Objective, Cost.

EXPLAINABILITY & BIAS
A Financial Brain Scan of the LLM arxiv.org/abs/2508.21285
Glasserman & Lin. Look-Ahead Bias (Columbia) arxiv.org/abs/2309.17322
Gao et al. A Test of Lookahead Bias in LLM Forecasts (2025) arxiv.org/abs/2512.23847
Look-Ahead-Bench: Standardized Benchmark (2026) arxiv.org/abs/2601.13770
Chronologically Consistent LLMs, He et al. (2025) arxiv.org/abs/2510.11677
When LLMs Go Abroad: Foreign Bias (2025)
AI Agent Misinformation in Financial Decision-Making (2025)
Social Group Bias in AI Finance (2025)
LLM Knows Geometry Better than Algebra (2025)

BENCHMARKS
FinBen: Holistic Financial Benchmark (Xie et al. NeurIPS 2024) arxiv.org/abs/2402.12659
InvestorBench: Financial Decision-Making Benchmark (ACL 2025) arxiv.org/abs/2412.18174
Open FinLLM Leaderboard (Lin et al. 2025) arxiv.org/abs/2501.10963

REGULATION
The Agentic Regulator: AI Governance Framework (2025)
CFA Explainable AI Report (2025)
BIS: AI Explainability in Finance (2025)
US Treasury AI Risk Management Framework (2026)

SURVEYS
Bridging Finance and AI: Comprehensive Survey (2025)
From Deep Learning to LLMs: AI in Quant Investment (2025)
The New Quant: LLMs in Financial Prediction (Fu, 2025) arxiv.org/abs/2510.05533
A Survey of Financial AI (Liu, 2024) arxiv.org/abs/2411.12747
LLMs in Equity Markets (Jadhav & Mirza, Frontiers in AI 2025)
FinLLMs, Lee et al. (Neural Computing & Applications, 2025)
LLM Agent in Financial Trading: A Survey, Ding et al. (2026)
LLM Agents for Investment Management (ACM ICAIF 2025)
Bybee, Kelly & Su. Narrative Asset Pricing (Yale, RFS 2023)
A Review of LLM Agent Apps in Finance (Batra et al. 2025)
LLMs for Financial Apps: Progress & Challenges (Nie et al. 2024) arxiv.org/abs/2406.11903



ZANISTA

Thank you!

arman.khaledian@zanista.ai
https://zanista.ai/

https://www.linkedin.com/in/armankhaledian/
Schedule a call: https://calendly.com/arman-khaledian-zanista/30min