



UNIVERSITY OF  
CAMBRIDGE

# A Context–Engineering Based Knowledge Framework for Quantitative Finance

(work done in University of Cambridge, paper accepted to Neurips 2025)

Haoxue Wang

London, UK





UNIVERSITY OF  
CAMBRIDGE

## Disclaimer

This presentation is for educational purposes only based on published paper and reflects the personal views of the speaker. It does not represent the views of **Millennium Capital Management** or any affiliated entity.

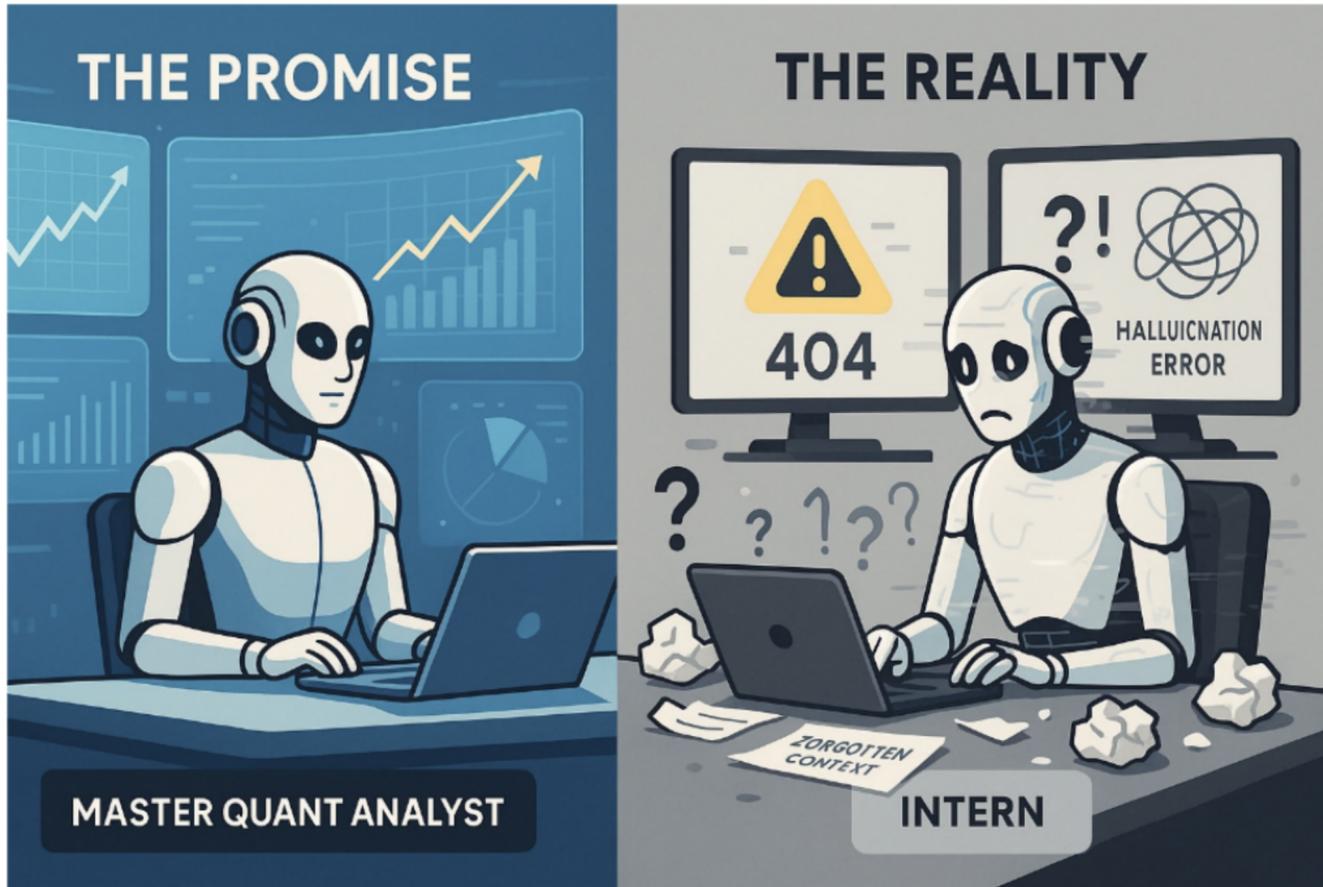
Nothing contained herein constitutes investment advice, a recommendation, or an offer to buy or sell any security or financial instrument.

No confidential, proprietary, or non-public information is included in this presentation

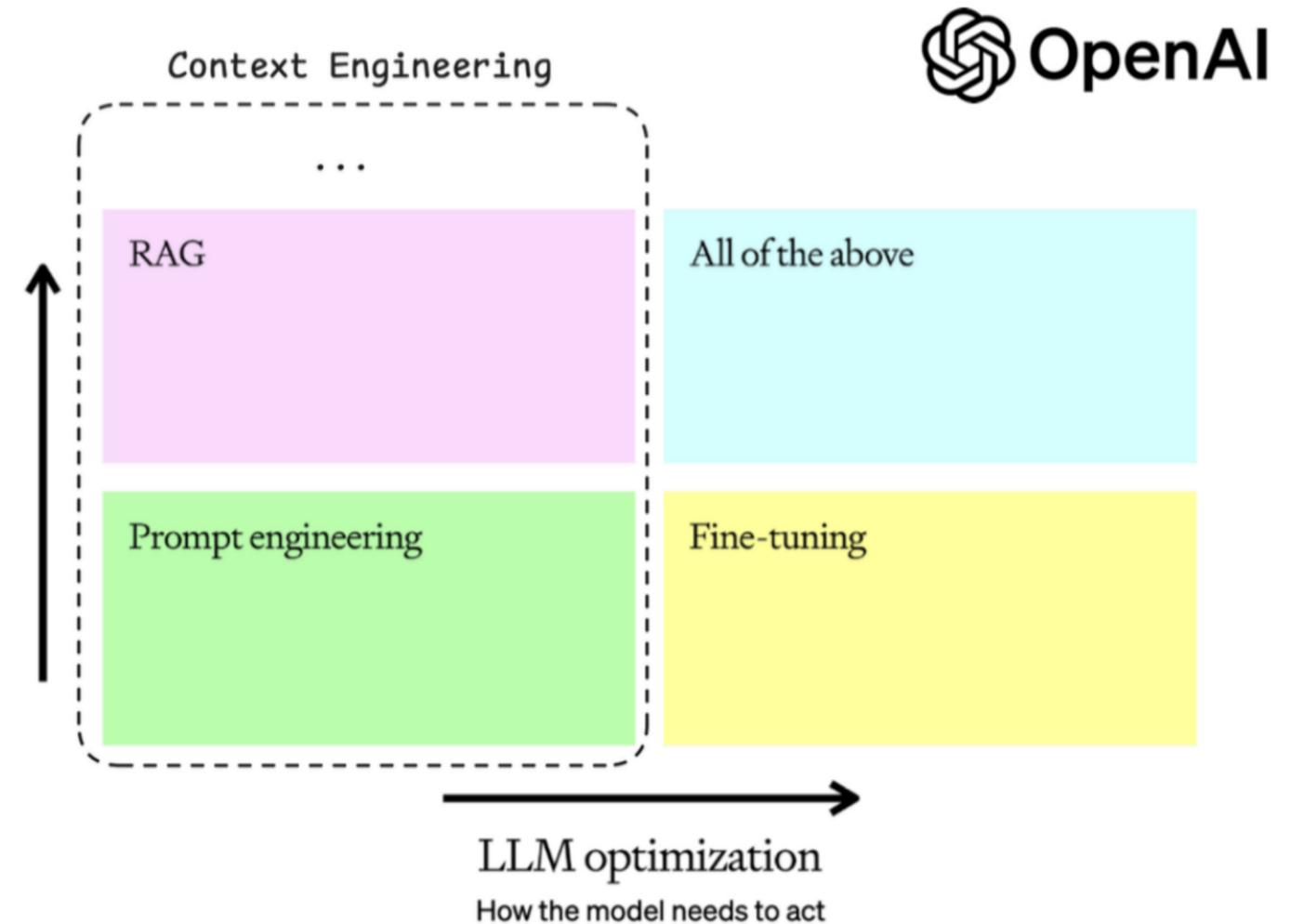




## The Problem of Current Agentic System



Context optimization  
What the model needs to know



Most AI Agent failures are not failures of model capability, but...?



UNIVERSITY OF  
CAMBRIDGE

# Model Optimization or Context Optimization

《Are ChatGPT and GPT-4 General-Purpose Solvers for Financial Text Analytics? A Study on Several Typical Tasks》

J.P.Morgan



**Comparison over LLMs.** We are able to benchmark the performance of ChatGPT and GPT-4 with four other LLMs on five tasks with eight datasets. ChatGPT and GPT-4 significantly outperforms others in almost all datasets except the NER task. It is interesting to observe that both models perform better on financial NLP tasks than BloombergGPT, which was specifically trained on financial corpora. This might be due to the larger model size of the two models. Finally, GPT-4 constantly shows 10+% boost over ChatGPT in straightforward tasks such as Headlines and FiQA SA. For challenging tasks like RE and QA, GPT-4 can introduce 20-100% performance growth. This indicates that GPT-4 could be the first choice for financial NLP tasks before a more powerful LLM emerges.

💡 a generalist model + the right context often outperforms a niche model, making “ what you feed it now” more decisive than “ what it was trained on.”



UNIVERSITY OF  
CAMBRIDGE

## The Problem: Context Engineering



**Uncertain assumption**

**Hallucination**

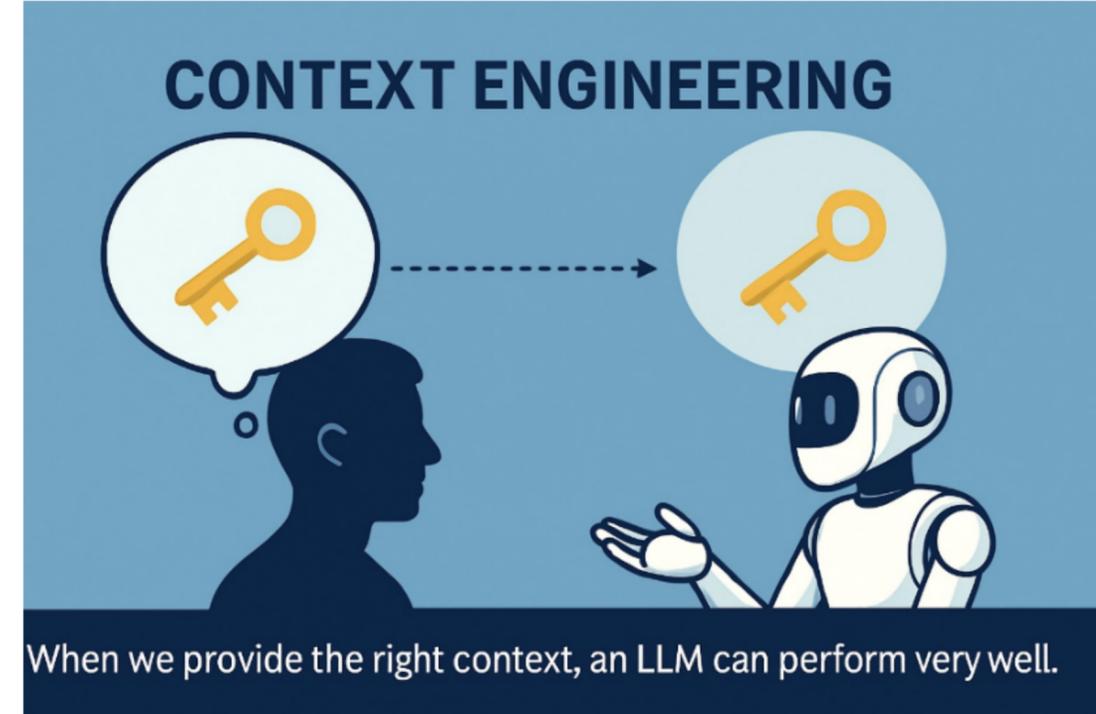
LLMs cannot read minds && Bad Context



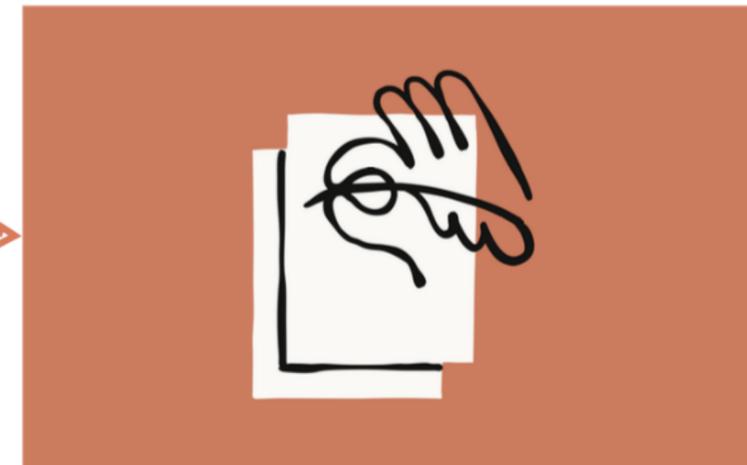
# The Problem: Context Engineering



LLMs cannot read minds && Bad Context



Need



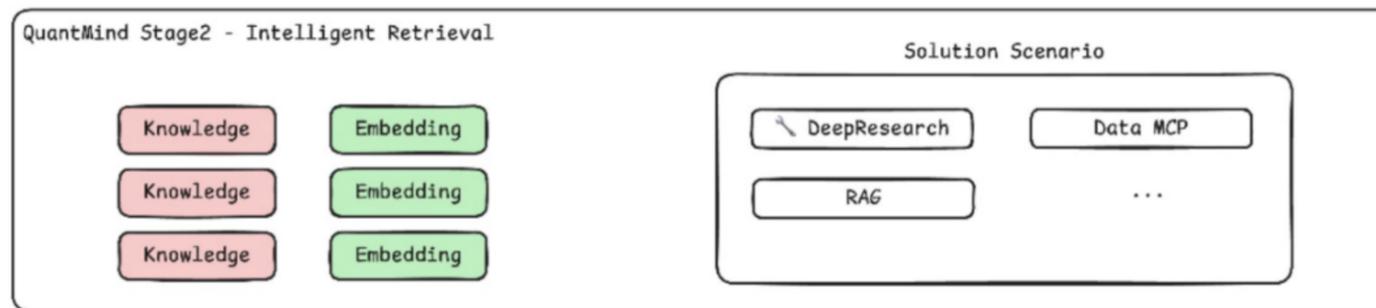
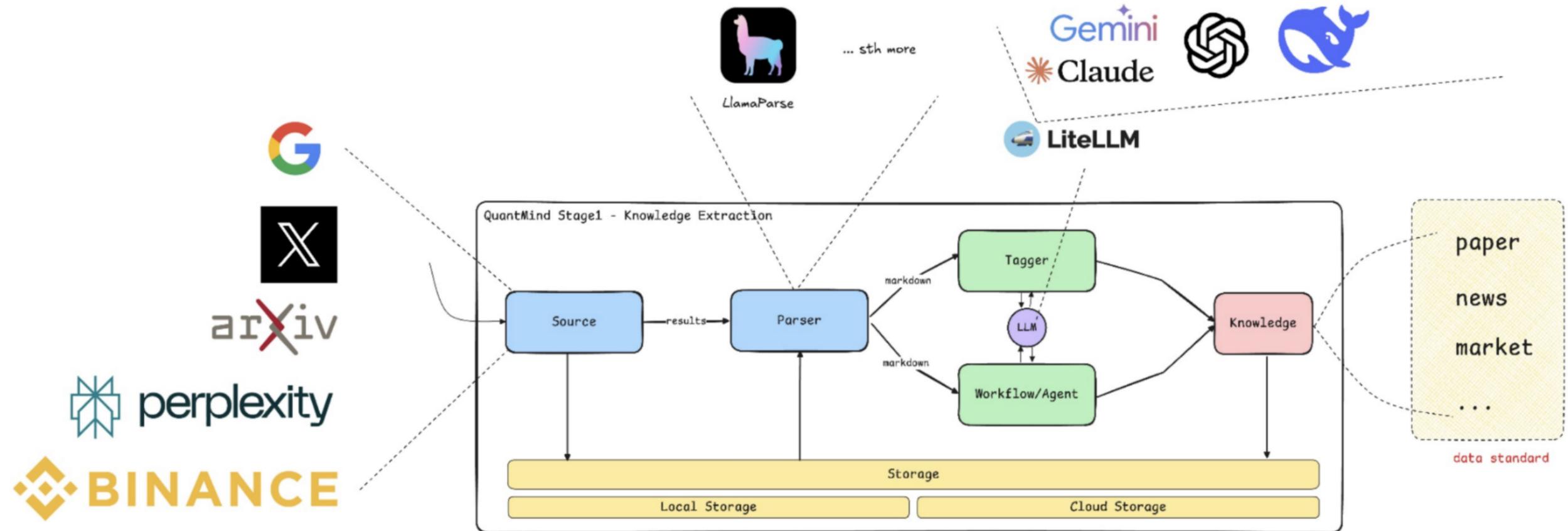
**Context Engineering**





# UNIVERSITY OF CAMBRIDGE

## A Context-Engineering Based Knowledge Framework for Quantitative Finance





UNIVERSITY OF  
CAMBRIDGE

A Context-Engineering Based Knowledge  
Framework for Quantitative Finance

1. **Long-running, background** raw information processing workflows to obtain various contexts
2. Provide various **vertical** tools to support Context Engineering for Quant/Finance Agentic Systems. (Universal Context Engineering Solution for Quant/Finance & LLM Scenarios)



**Results.** Table 4 reports descriptive statistics and pairwise differences. Each treatment condition includes approximately equal numbers of observations (66–68), and the variance structure indicates heterogeneity across both participants and papers. For *Quality*, QuantMind significantly outperforms both baselines: a 1.14-point gain over the no-AI condition ( $p < 0.001$ ) and a 0.43-point gain over the generic AI assistant ( $p = 0.001$ ). For *UX Rating*, while the overall treatment effect was not significant at the conventional level ( $p = 0.108$ ), QuantMind improved average ratings by 0.38 points and yielded a statistically significant enhancement in perceived helpfulness ( $p = 0.003$ ). These findings suggest that QuantMind not only enhances the accuracy and depth of research outputs but also provides a more supportive user experience compared to both unaided reading and a generic AI assistant. Complete statistical outputs, including confidence intervals and additional pairwise contrasts, are provided in Appendix A.4.

| Metric          | Treatment        | $n$ | Mean | SD   | Median | Min–Max |
|-----------------|------------------|-----|------|------|--------|---------|
| Accuracy (0–5)  | Without AI       | 66  | 3.11 | 0.68 | 3.0    | 1.0–4.5 |
|                 | AI Assistant     | 68  | 3.82 | 0.64 | 4.0    | 2.0–5.0 |
|                 | <i>QuantMind</i> | 66  | 4.25 | 0.55 | 4.5    | 3.0–5.0 |
| UX Rating (1–5) | Without AI       | 66  | —    | —    | —      | —       |
|                 | AI Assistant     | 68  | 3.78 | 0.62 | 4.0    | 2.0–5.0 |
|                 | <i>QuantMind</i> | 68  | 4.21 | 0.55 | 4.0    | 3.0–5.0 |



UNIVERSITY OF  
CAMBRIDGE

Q&A



QUANT  
STRATS