



Bridging Specialized **ML** Research and **Systematic Investing**

Transforming the Research Pipeline

— Gautier Marti, Quant @ ADIA —
25/04/2025

The Quant Lab in 2025



The Standard Research Pipeline in Quant Investing

Modular, specialization-heavy, and ripe for agentic transformation?

📁 Data Sourcing → 💡 Idea Generation → ⚙️ Data Processing → 📈 Forecasting Model → 🇮🇹 Portfolio Construction → 📄 Execution

The standard workflow spans three major stages:

🔍 Ideation & Data Collection

- Start from **data sourcing** (market, fundamental, alternative, text, etc.)
- Develop **hypotheses** based on intuition or published research

📁 Feature Engineering & Predictive Modeling

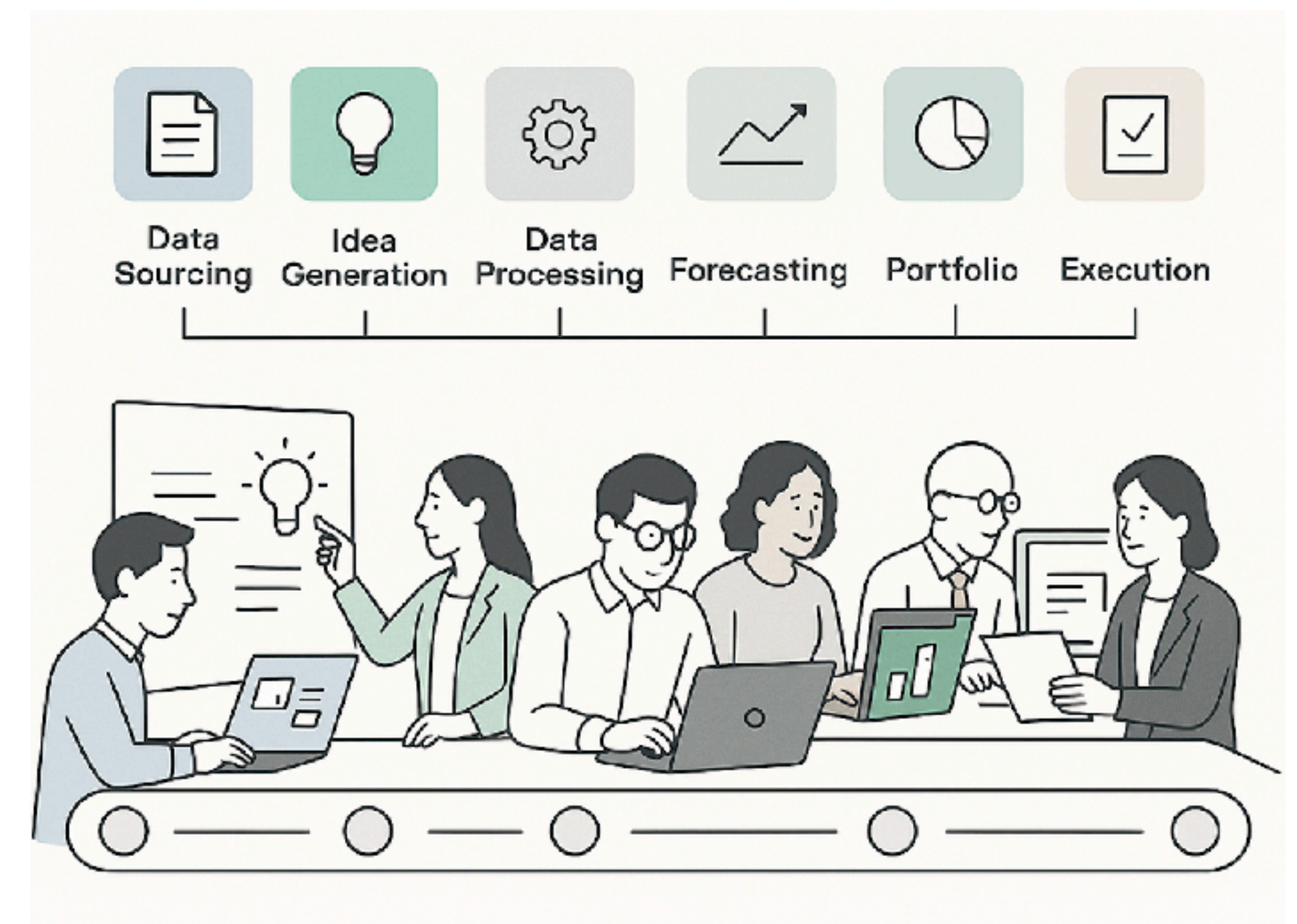
- Build **clean, engineered datasets**
- Fit **predictive models** (linear, tree-based, PCA, etc.)

💰 Monetization

- Convert forecasts into **portfolio weights** under constraints
- Simulate execution, cost, risk, turnover and **optimize trading**

➡ *These steps are typically modular, human-driven, and hand-coded.*

➡ *Researchers often specialize in one or more of these stages, but rarely master them all.*

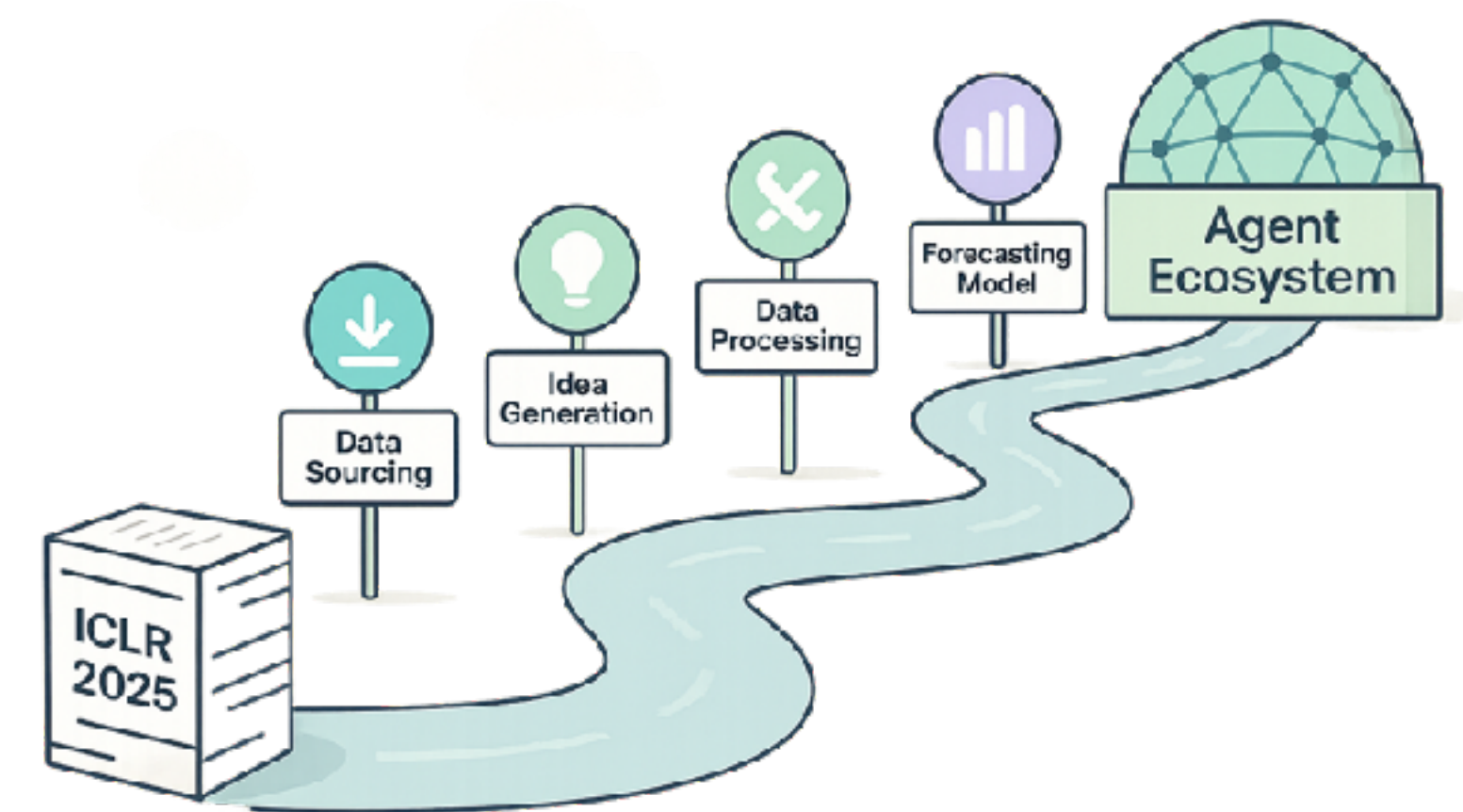


Talk Roadmap

Bridging ML Research and the Quant Workflow — One Step at a Time

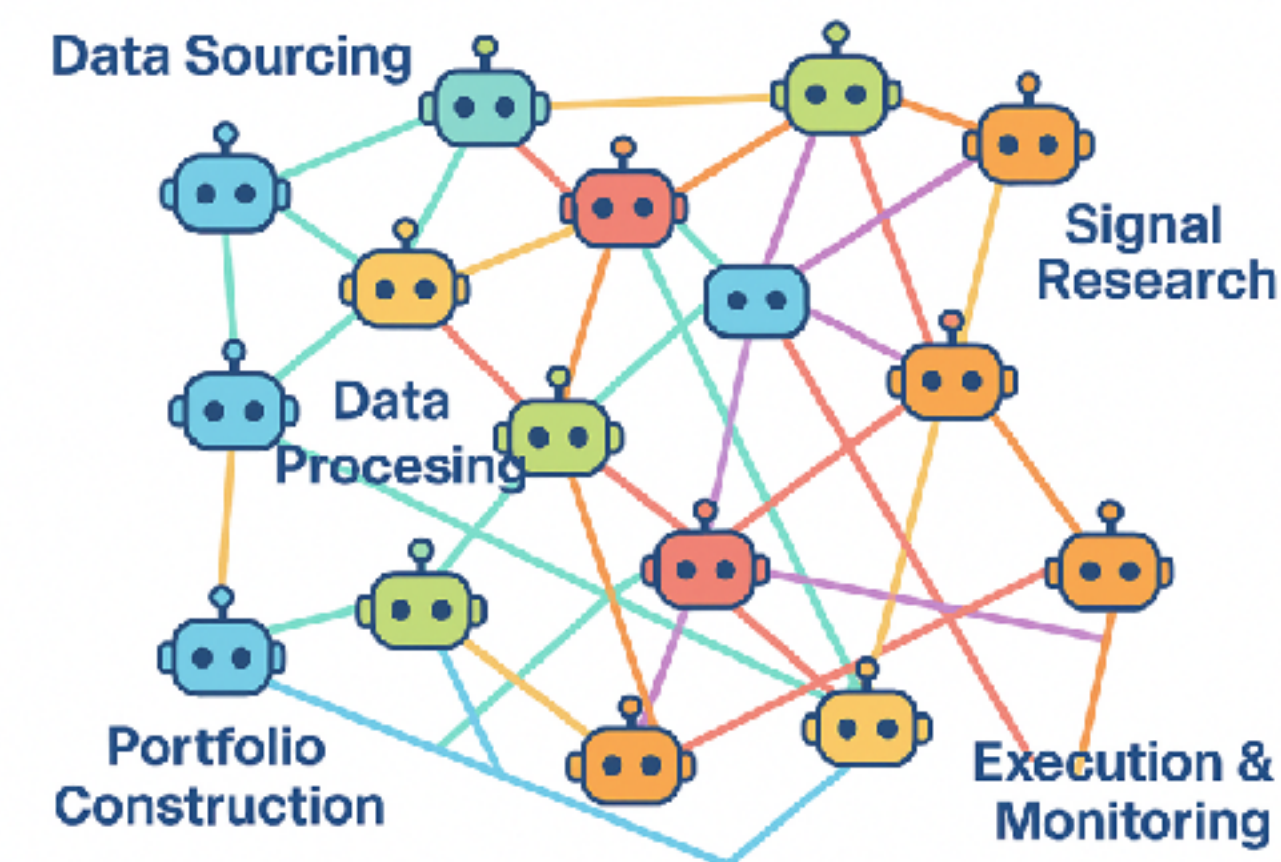
◆ Part 1 — Step-by-Step: Enhancing the Pipeline

- ➡ Walk through the 6 core stages of the quant workflow
- ➡ At each step, we'll highlight:
 - ML techniques from **ICLR 2025**
 - How they can **improve, scale, or automate** the process



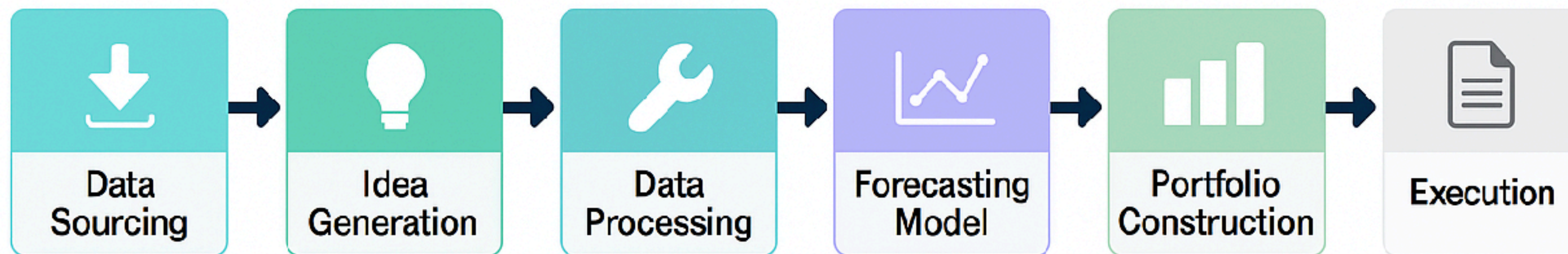
◆ Part 2 — Rethinking the Workflow

- ➡ Envision a next-gen architecture built on:
 - Agentic workflows
 - Feedback loops
 - Semi-autonomous research systems



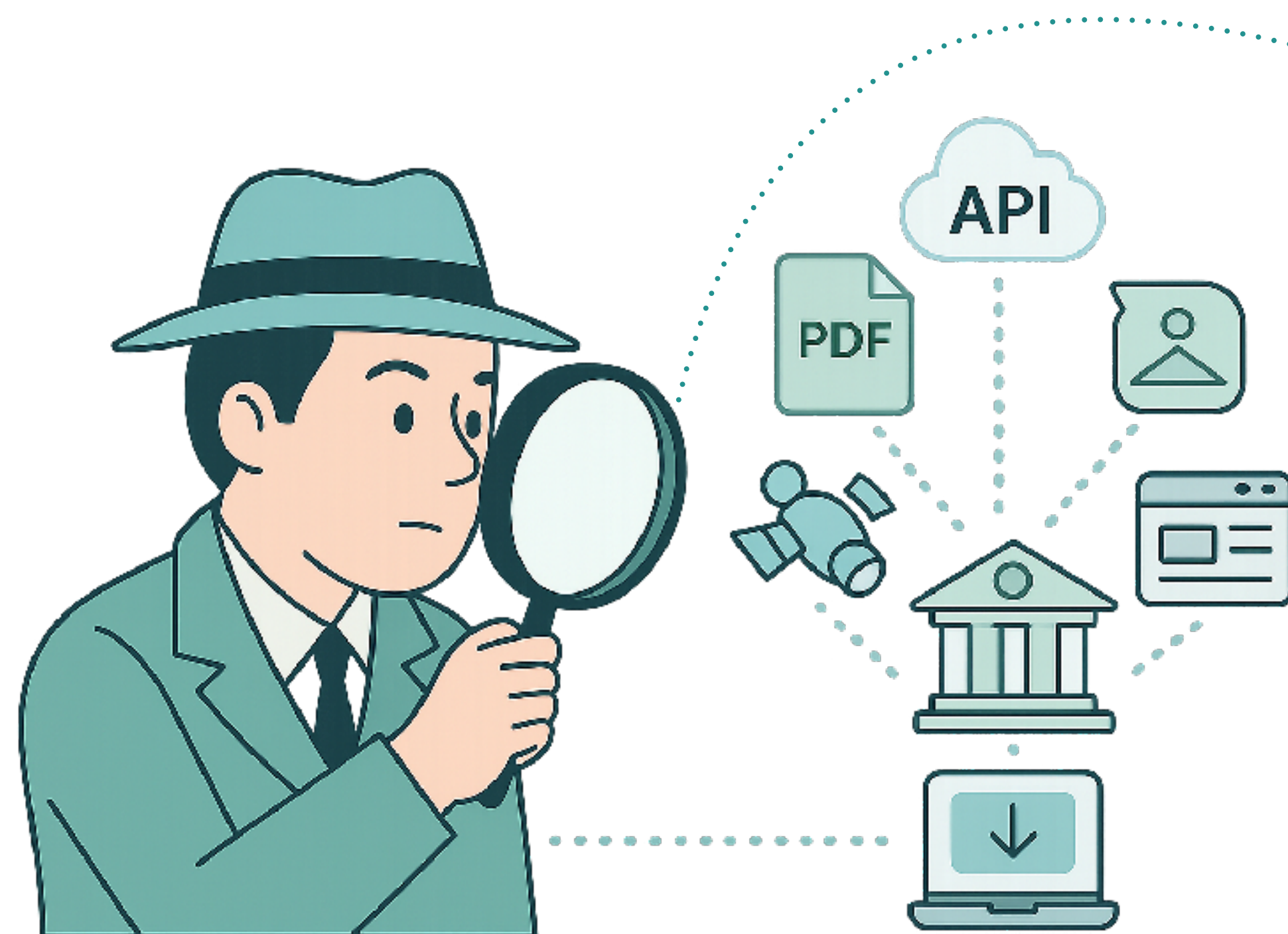
Part 1 — Enhancing the Quant Research Pipeline

Bridging **quant** workflows with specialized **ML** research



How Quants Scout for the Right Data

Data sourcing is where systematic research begins — and it's far from trivial.



1. Start with the Research Questions

- What are we trying to understand or predict?
- Macro trends, firm behavior, sector dynamics, etc.

2. Search for Potential Data Sources

- Government reports, startup APIs, web data, vendors...
- Often siloed, obscure, or unstructured — like investigative journalism

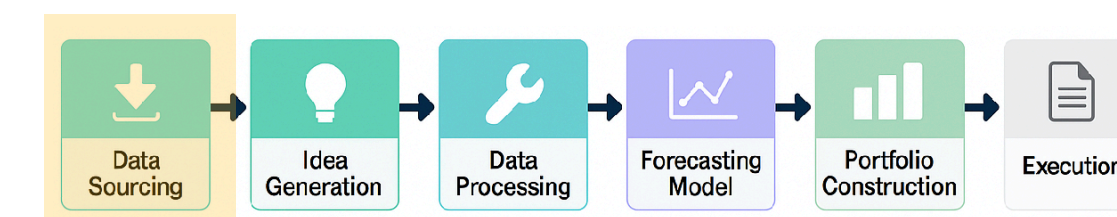
3. Evaluate Relevance

- Does this data hold signal for the research goal?
- Granularity, history, structure, and signal quality matter

4. Check for Quant Constraints

- Can it be used in a live systematic process?
- Needs to be timely, machine-readable, and reliably delivered

 *Much of this scouting work is still manual and labour intensive - ripe for ML automation?*



ML Research Spotlight (ICLR 2025)

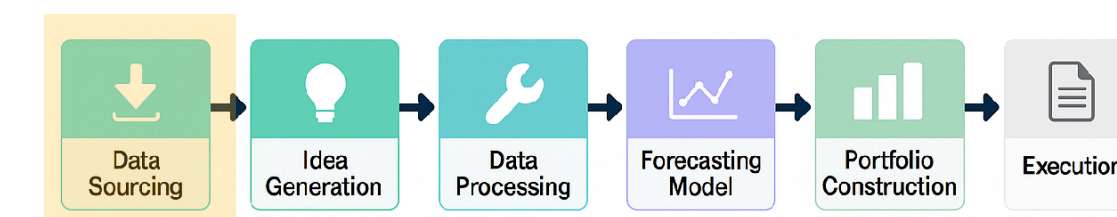
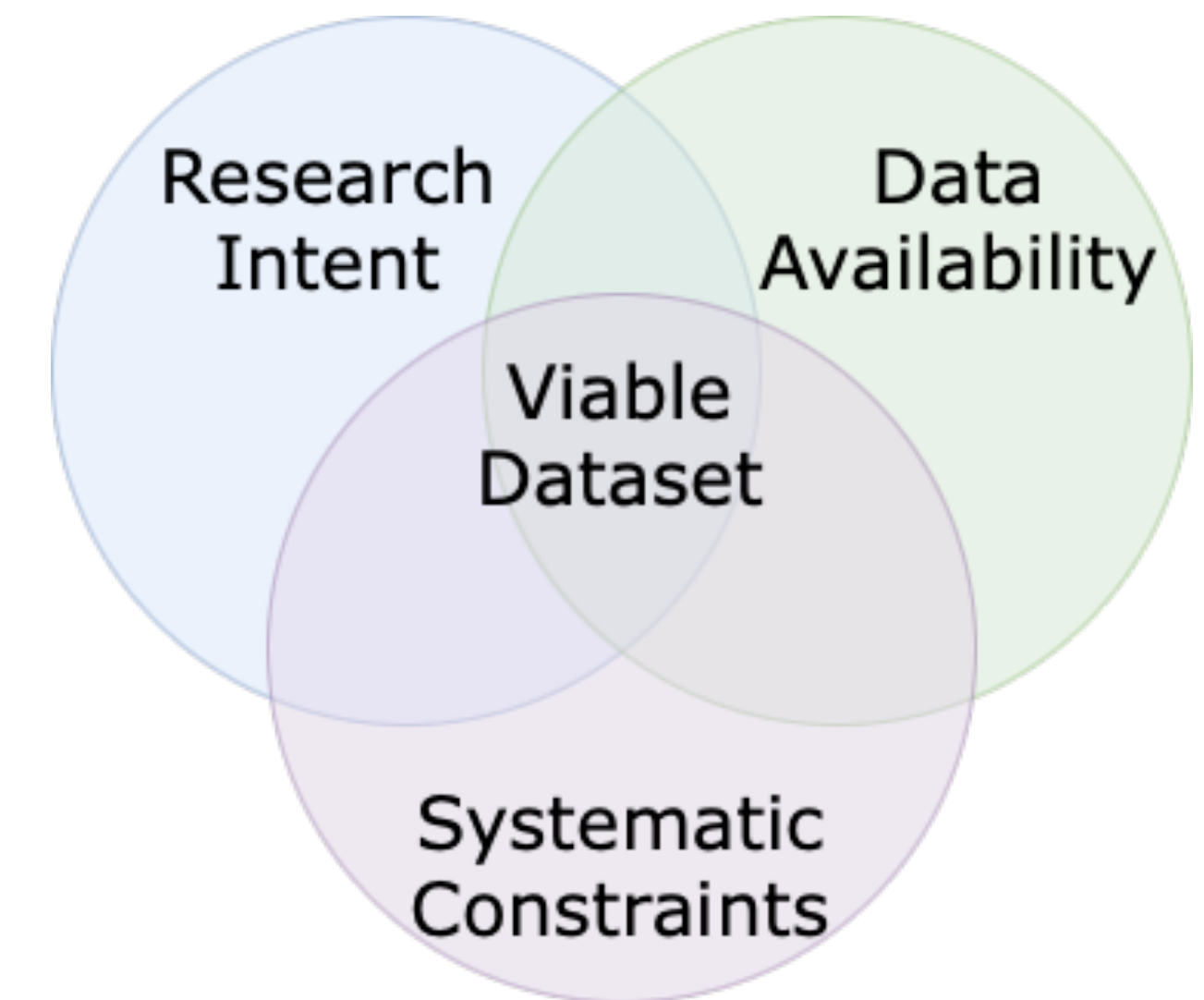
Bridging Data Sourcing and ML Research

Traditional data sourcing is high-touch, labor-intensive, and slow.

But several ICLR 2025 papers offer promising tools — in **retrieval**, **representation learning**, and **cross-modal search** — to help **bridge research questions with real-world data assets**.

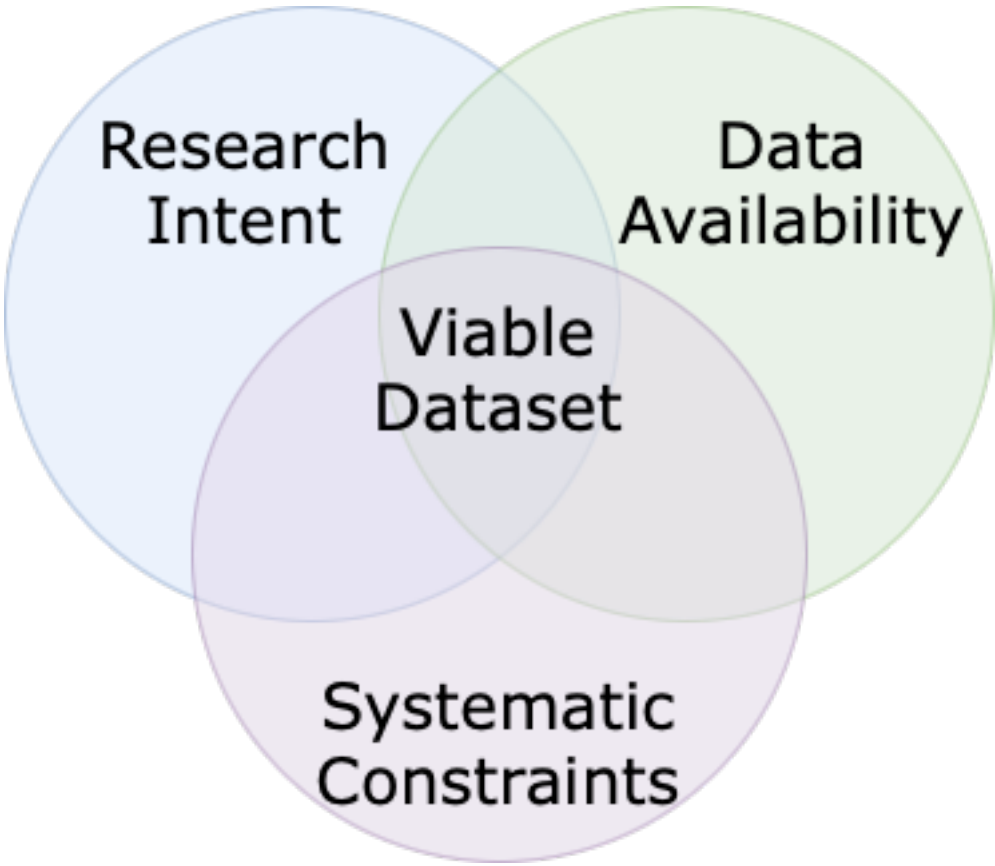
🔍 What ML can help with:

- **Intent clarification** and matching across research queries
- **Search & visualization** via learned **embeddings**
- **Retrieval** from unstructured or semi-structured corpora
- **Constraints checking** via metadata filtering & QA layers



ML Research Spotlight (ICLR 2025)

Bridging Data Sourcing and ML Research



Training Large Language Models for Retrieval-Augmented Question Answering through Backtracking Correction
[Huawen Feng](#), [Zekun Yao](#), [Junhao Zheng](#), [Qianli Ma](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)

The diagram illustrates a neural network architecture for Retrieval-Augmented Question Answering. It shows a sequence of operations: 'Embedding', 'Attention', 'Aggregation', 'Correction', and 'Output'. The network is composed of multiple layers of nodes, with arrows indicating the flow of information. A checkmark is visible in the bottom left corner.

StructRAG: Boosting Knowledge Intensive Reasoning of LLMs via Inference-time Hybrid Information Structurization
[Zhuoqun Li](#), [Xuanang Chen](#), [Haiyang Yu](#), [Hongyu Lin](#), [Yaojie Lu](#), [Qiaoyu Tang](#), [Fei Huang](#), [Xianpei Han](#), [Le Sun](#), [Yongbin Li](#)
Th, Apr 24, 11:00 +04 – [Poster Session 2](#)

The diagram shows the StructRAG architecture. It includes a 'Hybrid Structure Router' that determines the optimal structure type based on the current task. This is followed by a 'Scattered Knowledge Structurizer' that transforms original documents into structured knowledge. Finally, a 'Structured Knowledge Utilizer' decomposes questions and utilizes the structured knowledge. A checkmark is visible in the bottom left corner.

Simple is Effective: The Roles of Graphs and Large Language Models in Knowledge-Graph-Based Retrieval-Augmented Generation
[Mufei Li](#), [Siqu Miao](#), [Pan Li](#)
Fr, Apr 25, 11:00 +04 – [Poster Session 4](#)

The graph plots Macro-F1 scores (Y-axis, ranging from 54 to 96) against the Number of Triples (X-axis, ranging from 100 to 500). Three data series are shown: 'Upper Bound' (red line with dots), 'GPT4o-mini' (orange line with dots), and 'Llama3.1-8B' (blue line with dots). The 'Upper Bound' starts at approximately 88 and increases to nearly 96. 'GPT4o-mini' starts at approximately 60 and increases to about 68. 'Llama3.1-8B' starts at approximately 54 and increases to about 58. A checkmark is visible in the bottom left corner.

MindSearch: Mimicking Human Minds Elicits Deep AI Searcher
[Zehui Chen](#), [Kuikun Liu](#), [Qiuchen Wang](#), [Jiangning Liu](#), [Wenwei Zhang](#), [Kai Chen](#), [Feng Zhao](#)
Fr, Apr 25, 06:00 +04 – [Poster Session 3](#)

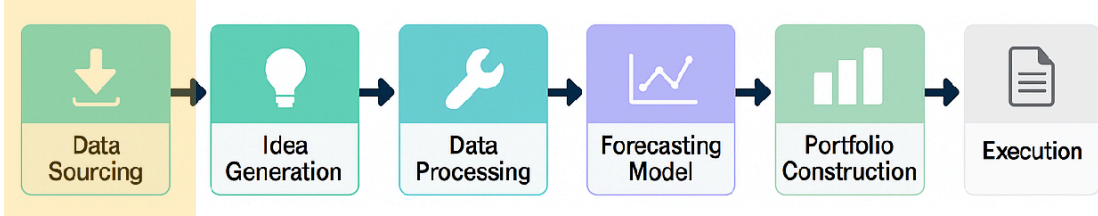
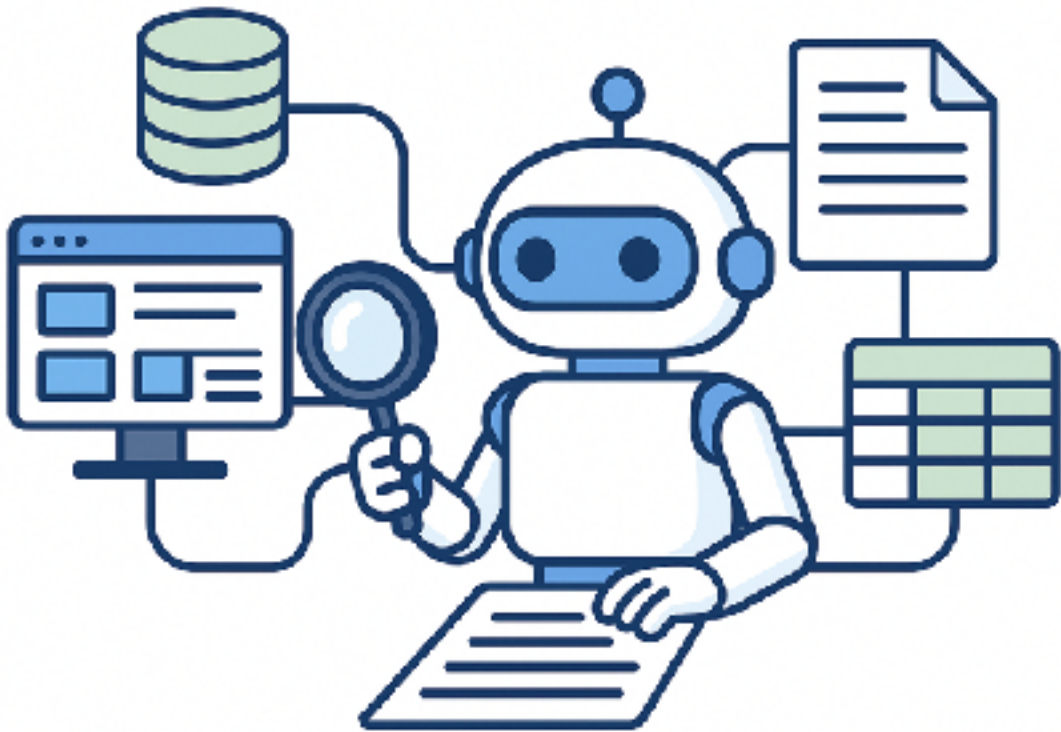
The diagram illustrates the MindSearch architecture. It shows a sequence of operations: 'Embedding', 'Attention', 'Aggregation', 'Correction', and 'Output'. The network is composed of multiple layers of nodes, with arrows indicating the flow of information. A checkmark is visible in the bottom left corner.

Contextual Document Embeddings
[John X. Morris](#), [Alexander Rush](#)
Fr, Apr 25, 11:00 +04 – [Poster Session 4](#)

The diagram shows the Contextual Document Embeddings architecture. It includes a 'Hybrid Structure Router' that determines the optimal structure type based on the current task. This is followed by a 'Scattered Knowledge Structurizer' that transforms original documents into structured knowledge. Finally, a 'Structured Knowledge Utilizer' decomposes questions and utilizes the structured knowledge. A checkmark is visible in the bottom left corner.

Learning to Clarify: Multi-turn Conversations with Action-Based Contrastive Self-Training
[Maximillian Chen](#), [Ruoxi Sun](#), [Tomas Pfister](#), [Sercan Arik](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)

The diagram illustrates the Learning to Clarify architecture. It shows a sequence of operations: 'Embedding', 'Attention', 'Aggregation', 'Correction', and 'Output'. The network is composed of multiple layers of nodes, with arrows indicating the flow of information. A checkmark is visible in the bottom left corner.



What We Need

To Advance Data Sourcing with AI & Quant Skills

Tackling the data sourcing challenge at scale requires a blend of capabilities:

- 🕵️ A **curious, investigative mindset** — like an ML-powered detective
- 🧠 A passion for **organizing and structuring information**, especially **textual data**
- 🔍 Interest in **retrieval systems**, **search infrastructure**, and **representation learning**
- 🌐 Familiarity with **internal/external data catalogs**, **web corpora**, or **vendor APIs**
- 🧭 Skills in **interactive data exploration** and **metadata-driven navigation**
- 📦 Experience with **knowledge bases**, **ontologies**, or **LLMs + knowledge graphs**

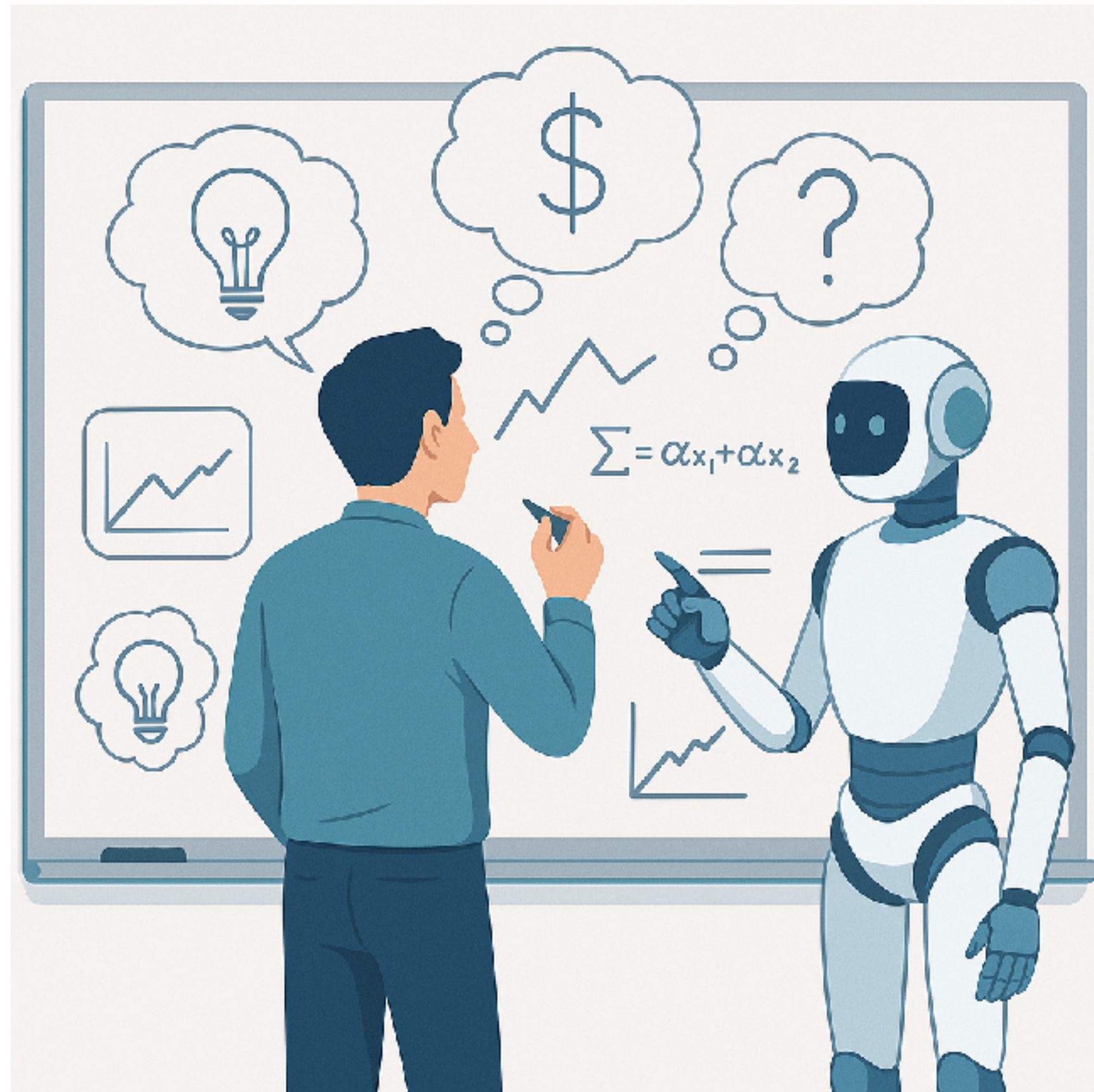
🧭 *Advancing systematic research starts with better data discovery — combining machine learning, structured metadata, and domain expertise to **uncover what matters**.*

🔍 *From investigative search to scalable semantic indexing, this is where technical depth meets applied creativity.*



💡 How Quants Generate Ideas for Signals

Idea generation is where hypotheses take shape. But it's not always easy...



1. Start from Market Intuition or Known Drivers

- What might explain asset returns or risk?
- Think macro trends, firm behavior, valuation, momentum, crowding, etc.

2. Explore Literature, Data and Known Patterns

- Browse anomalies, academic or industry research
- Spot statistical oddities or structure in the data

3. Frame a Hypothesis You Can Test

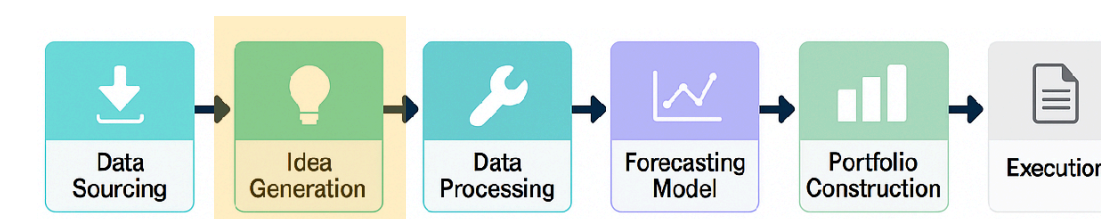
- Can it be expressed as a signal or a model input?
- Would this idea generalize — or is it just noise?

4. Iterate and Refine

- Try, fail, learn, repeat.
- It's often slow, nonlinear, and guided by hard-won experience.

📌 *The challenge? Inspiration doesn't strike on command — and idea quality varies.*

☁️ *Can we build agents that generate thoughtful, feasible signal ideas?*



ML Research Spotlight (ICLR 2025)

Bridging Idea Generation and ML Research

Generating new signals is often slow, nonlinear, and heavily dependent on domain experience. But ICLR 2025 papers offer emerging tools — in generative modeling, program synthesis, and agentic reasoning — that could help **expand the creative space** of signal ideation.

🔍 What ML can help with:

Pattern discovery from raw or structured data

→ Detect latent drivers, anomalies, or higher-order structures

Hypothesis suggestion via prompt-to-signal agents

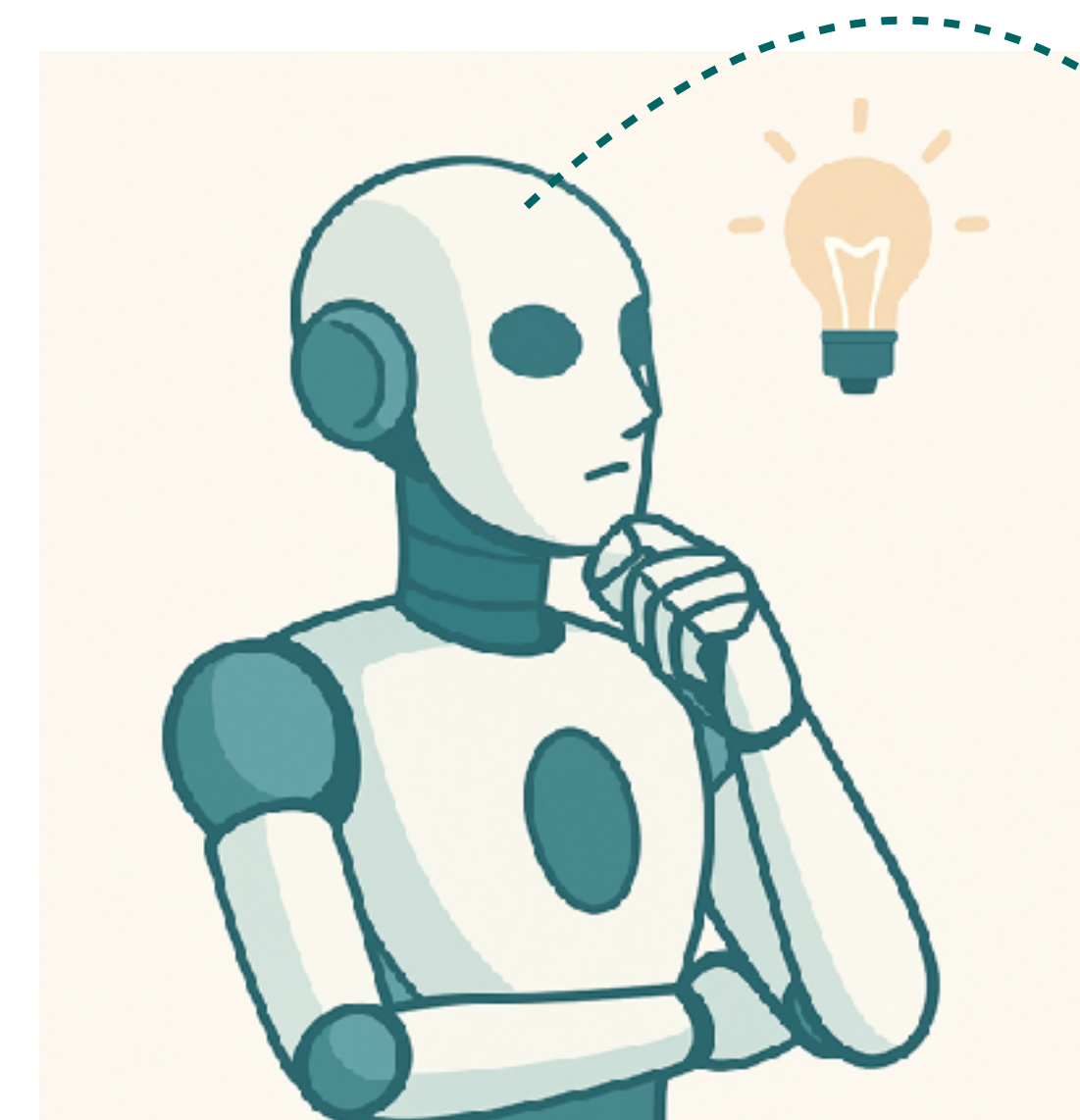
→ Grounded in data availability and quant constraints

Programmatic signal generation

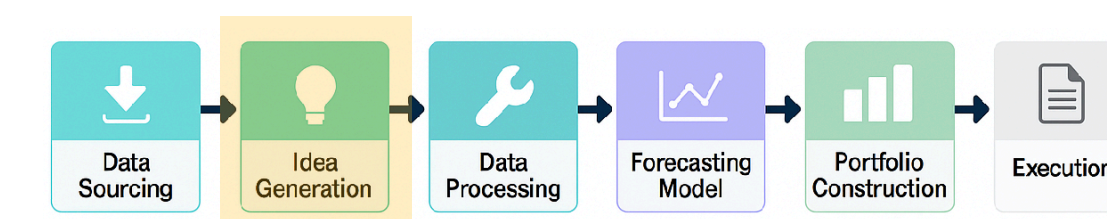
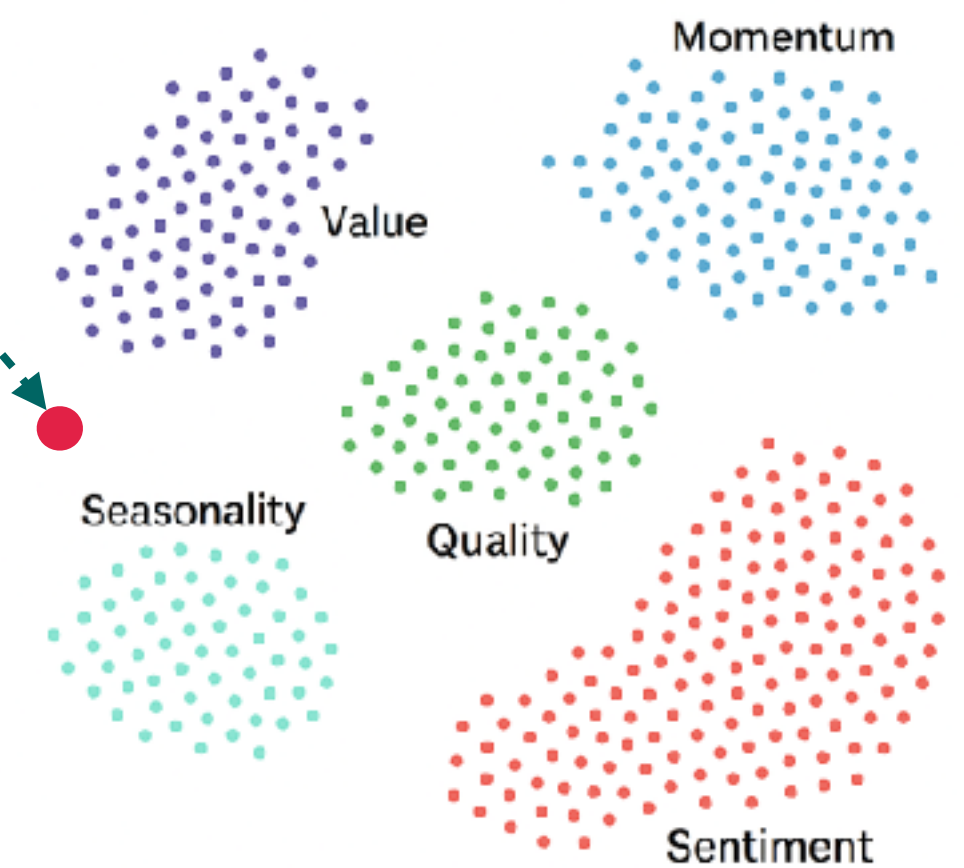
→ Use code-generation models to express testable signal logic

Idea clustering and refinement

→ Group ideas by structure, novelty, or empirical support



Map of Signals



ML Research Spotlight (ICLR 2025)

Bridging Idea Generation and ML Research

Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers
[Chenglei Si](#), [Diyi Yang](#), [Tatsunori Hashimoto](#)
Th, Apr 24, 06:00 +04 – [Poster Session 1](#)



DiscoveryBench: Towards Data-Driven Discovery with Large Language Models
[Bodhisattwa Prasad Majumder](#), [Harshit Surana](#), [Dhruv Agarwal](#), [Bhavana Dalvi Mishra](#), [Abhijeetsingh Meena](#), [Aryan Prakhhar](#), [Tirth Vora](#), [Tushar Khot](#), [Ashish Sabharwal](#), [Peter Clark](#)
Th, Apr 24, 06:00 +04 – [Poster Session 1](#)



ScienceAgentBench: Toward Rigorous Assessment of Language Agents for Data-Driven Scientific Discovery
[Ziru Chen](#), [Shijie Chen](#), [Yuting Ning](#), [Qianheng Zhang](#), [Boshi Wang](#), [Botao Yu](#), [Yifei Li](#), [Zeyi Liao](#), [Chen Wei](#), [Zitong Lu](#), [Vishal Dey](#), [Mingyi Xue](#), [Frazier N. Baker](#), [Benjamin Burns](#), [Daniel Adu-Ampratwum](#), [Xuhui Huang](#), [Xia Ning](#), [Song Gao](#), [Yu Su](#), [Huan Sun](#)
Fr, Apr 25, 11:00 +04 – [Poster Session 4](#)



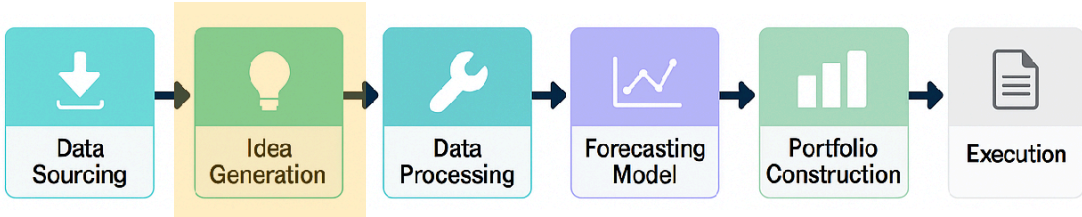
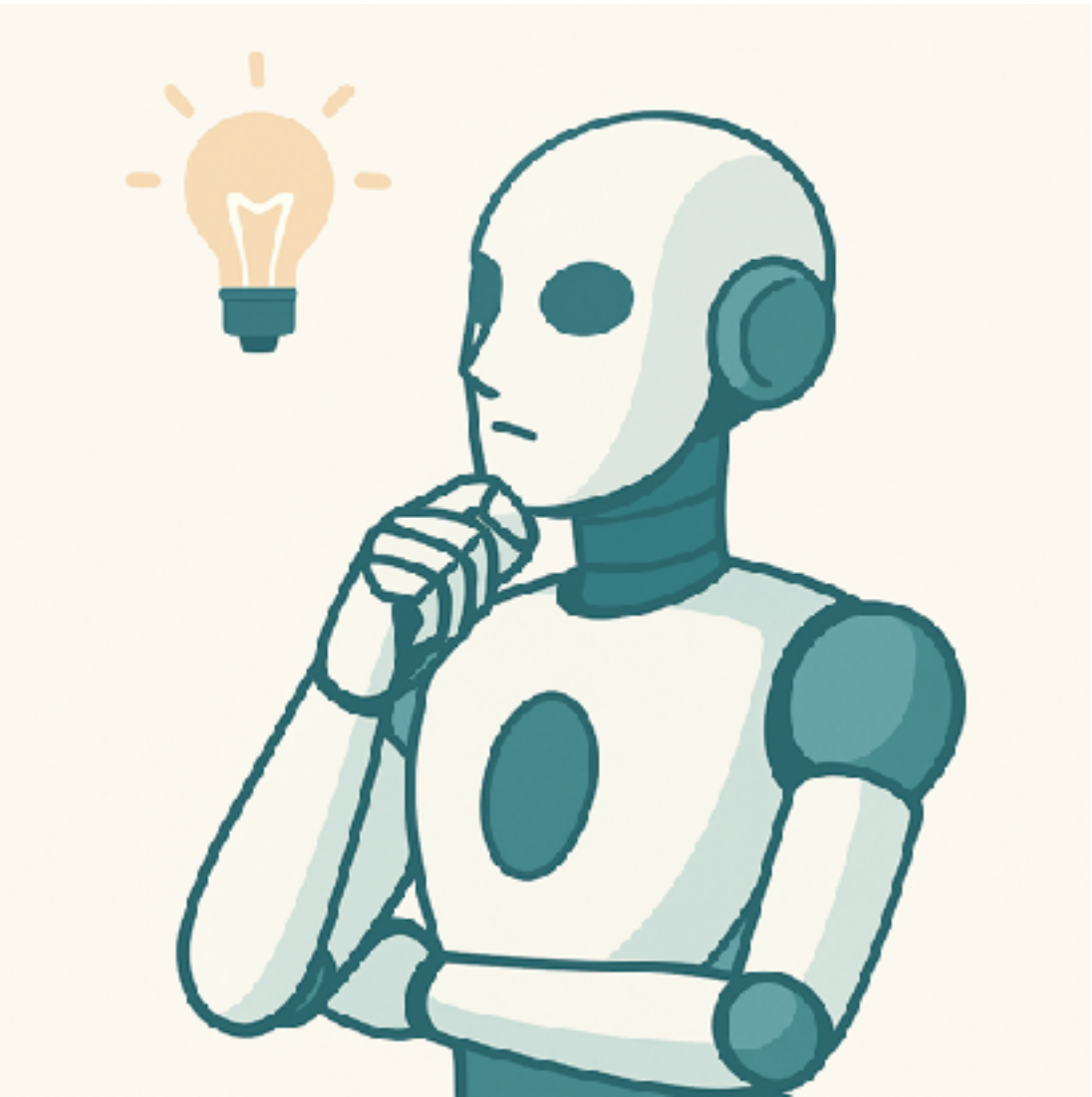
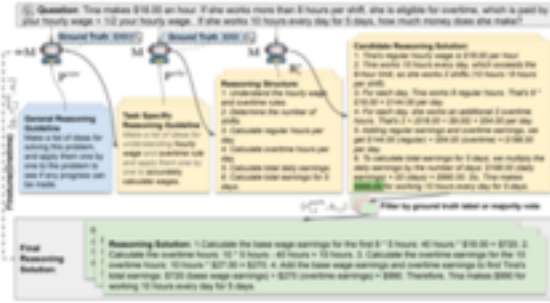
InsightBench: Evaluating Business Analytics Agents Through Multi-Step Insight Generation
[Gaurav Sahu](#), [Abhay Puri](#), [Juan A. Rodriguez](#), [Amirhossein Abaskohi](#), [Mohammad Chegini](#), [Alexandre Drouin](#), [Perouz Taslakian](#), [Valentina Zantedeschi](#), [Alexandre Lacoste](#), [David Vazquez](#), [Nicolas Chapados](#), [Christopher Pal](#), [Sai Rajeswar](#), [Issam Laradji](#)
Fr, Apr 25, 06:00 +04 – [Poster Session 3](#)



Teaching LLMs How to Learn with Contextual Fine-Tuning
[Younwoo Choi](#), [Muhammad Adil Asif](#), [Ziwen Han](#), [John Willes](#), [Rahul G. Krishnan](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)



ReGenesis: LLMs can Grow into Reasoning Generalists via Self-Improvement
[XIANGYU PENG](#), [Congying Xia](#), [Xinyi Yang](#), [Caiming Xiong](#), [Chien-Sheng Wu](#), [Chen Xing](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)
Sa, Apr 26, 12:06 +04 – [Oral Session 6A](#)



What We Need

To Automate Idea Generation for Quant Investing

Advancing idea generation calls for a mix of creative and technical aptitudes:

- 🧠 **Formulate rich, non-trivial hypotheses** grounded in financial & economic data
- 📊 **Test and refine ideas empirically**, using models and structured pipelines
- ⚖️ **Evaluate feasibility** based on data availability, infrastructure, and cost
- 📁 **Clustering and organizing ideas** by structure, theme, or novelty
- 📈 **Rank & prioritize the research agenda** based on expected impact
- 🔍 **Identify gaps** — what's missing to pursue the best ideas?

📌 We want to *automate* the path from inspiration to implementation — not just execution, but *curiosity* itself.

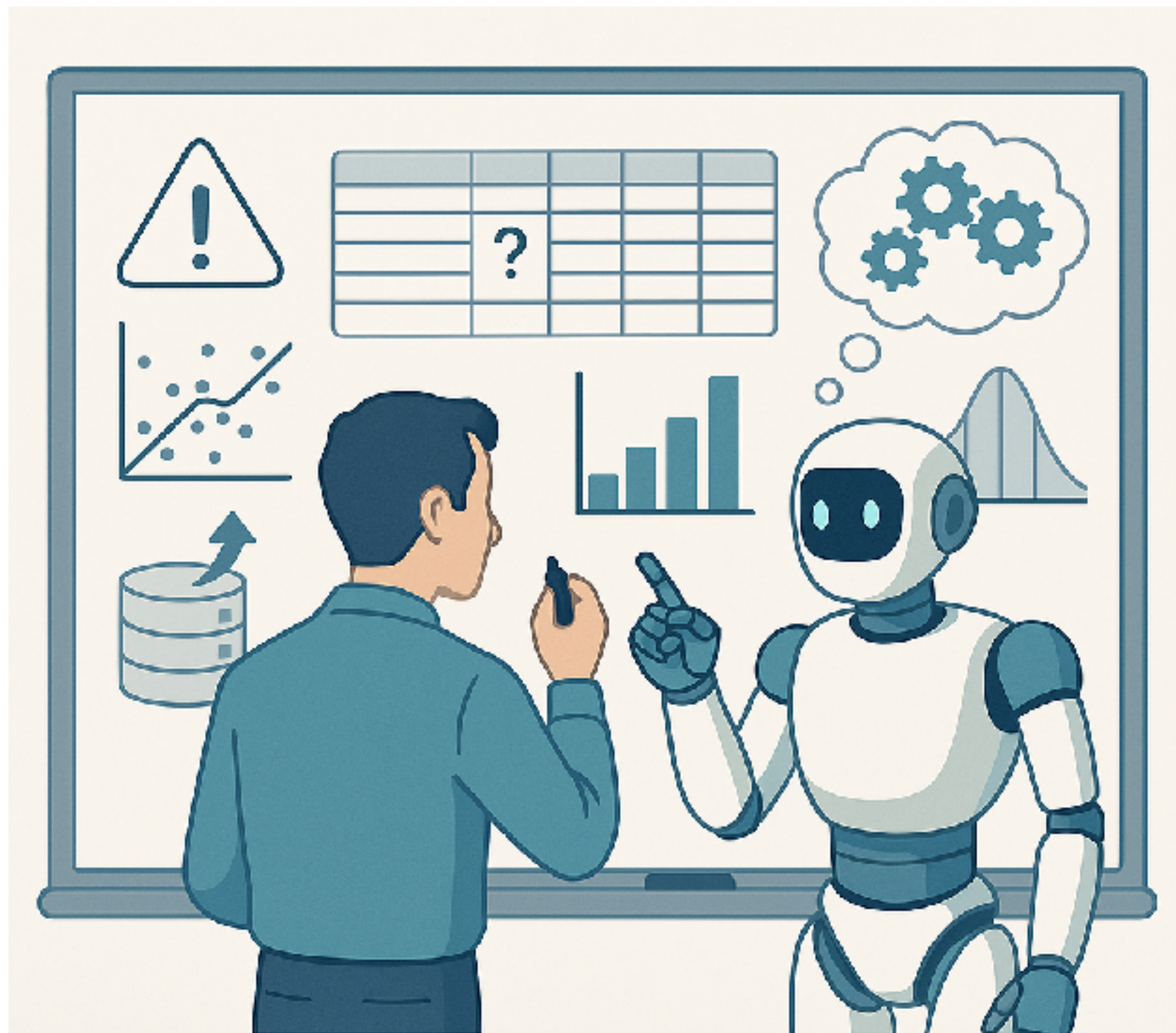
🔗 If you're excited by reasoning, discovery, and scientific automation — *let's talk*.



To Generate Quant Research Ideas with AI

🔧 How Quants Process Data for Research

Data processing is where raw data becomes signal-ready — and the details really matter.



1. Clean & Normalize the Raw Data

- Handle missing values, outliers, inconsistent formats
- Normalize time series (e.g., returns, volatility)

2. Align Across Sources & Frequencies

- Merge datasets with different timeframes or identifiers
- Sync economic indicators with asset-level data, news, etc.

3. Engineer Features & Transformations

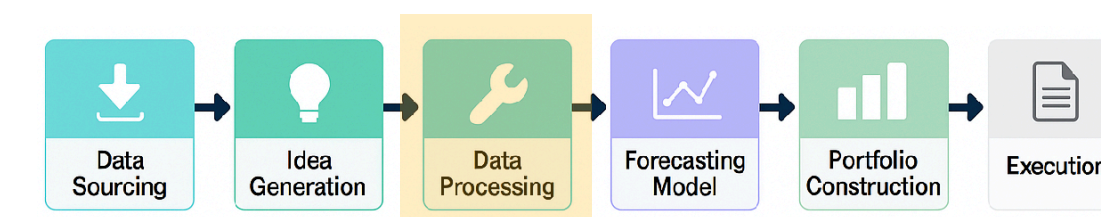
- Create lagged features, growth rates, decays, rolling statistics
- Apply transformations (log, winsorize) to make data model-friendly

4. Filter, Flag, and QA

- Identify leaks, suspicious jumps, or stale values
- Tag anomalies, and build QA dashboards or automated checks

📌 *The challenge? Precision is key — subtle bugs or misalignments can ruin research results.*

☁️ *Can we build models that prepare high-quality research data — reproducibly and at scale?*



ML Research Spotlight (ICLR 2025)

Bridging Data Processing and ML Research

Systematic research depends on clean, aligned, and structured data — but real-world datasets are messy, fragmented, and multimodal.

ICLR 2025 papers offer promising approaches in representation learning, multi-modal modeling, and automated cleaning — to help **scale robust data preparation** for quant workflows.

🔍 What ML can help with:

Outlier detection

→ Spot anomalies using statistical or embedding-based methods

Missing data imputation

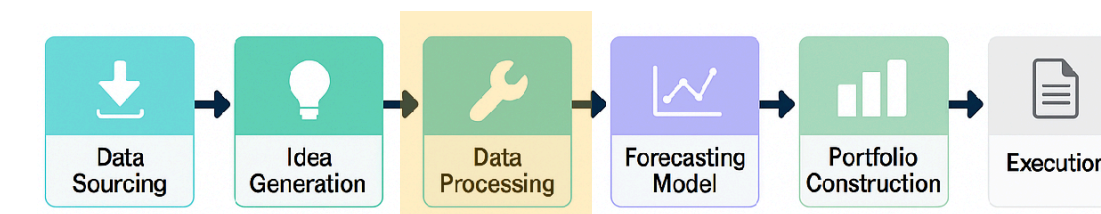
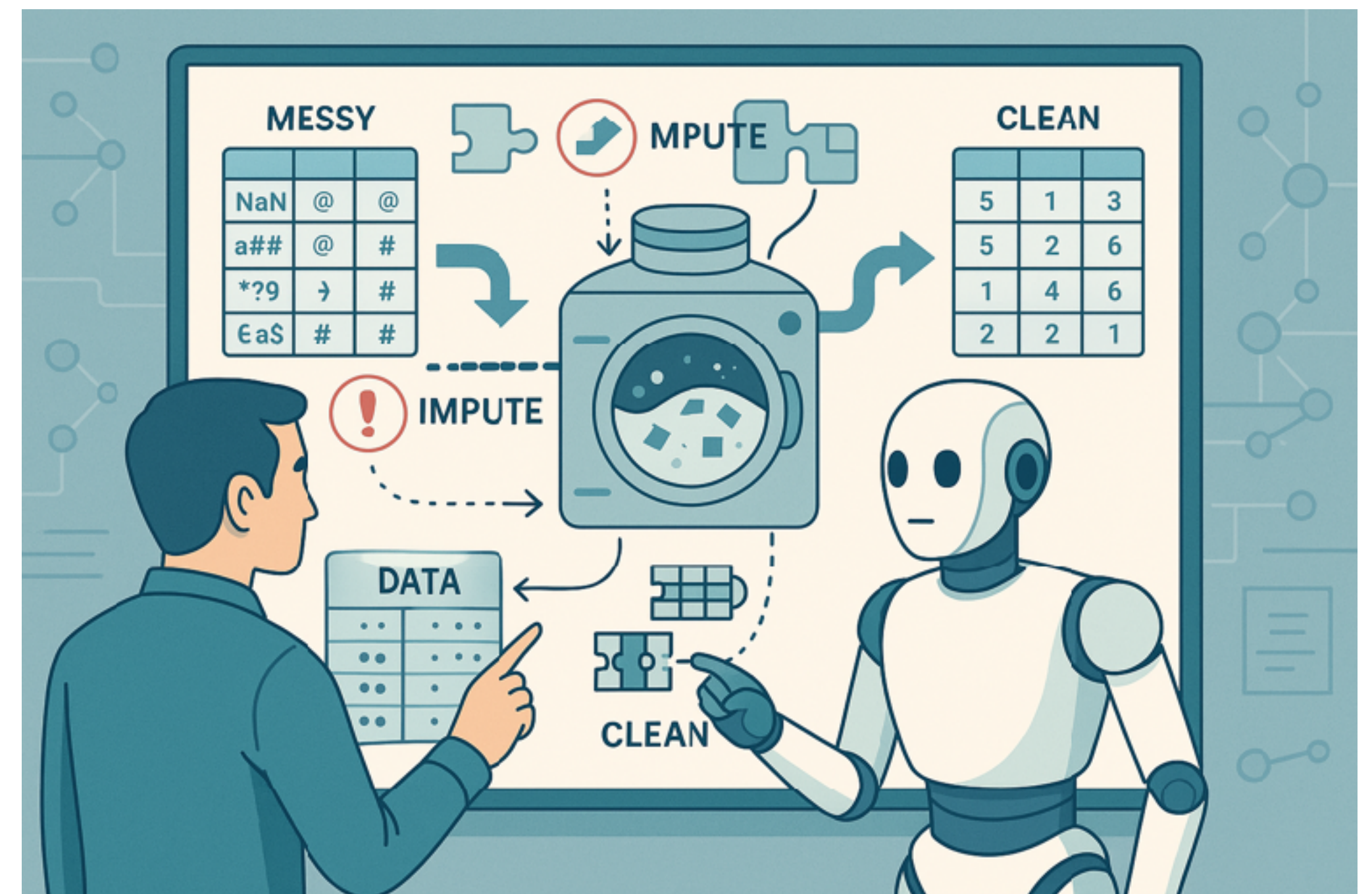
→ Fill gaps with generative models or structured priors

Smart joins and entity alignment

→ Align datasets via metadata or semantic similarity

Multi-modal processing

→ Integrate tabular, text, audio, or image inputs



ML Research Spotlight (ICLR 2025)

Bridging Data Processing and ML Research

Multi-Resolution Decomposable Diffusion Model for Non-Stationary Time Series Anomaly Detection
[Guojin Zhong](#), [pan wang](#), [Jin Yuan](#), [Zhiyong Li](#), [Long Chen](#)
Fr, Apr 25, 06:00 +04 – [Poster Session 3](#)

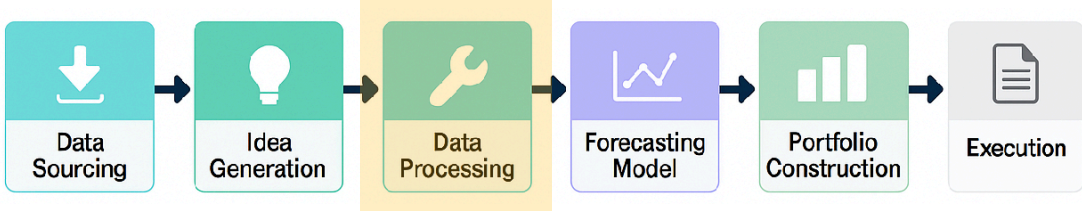
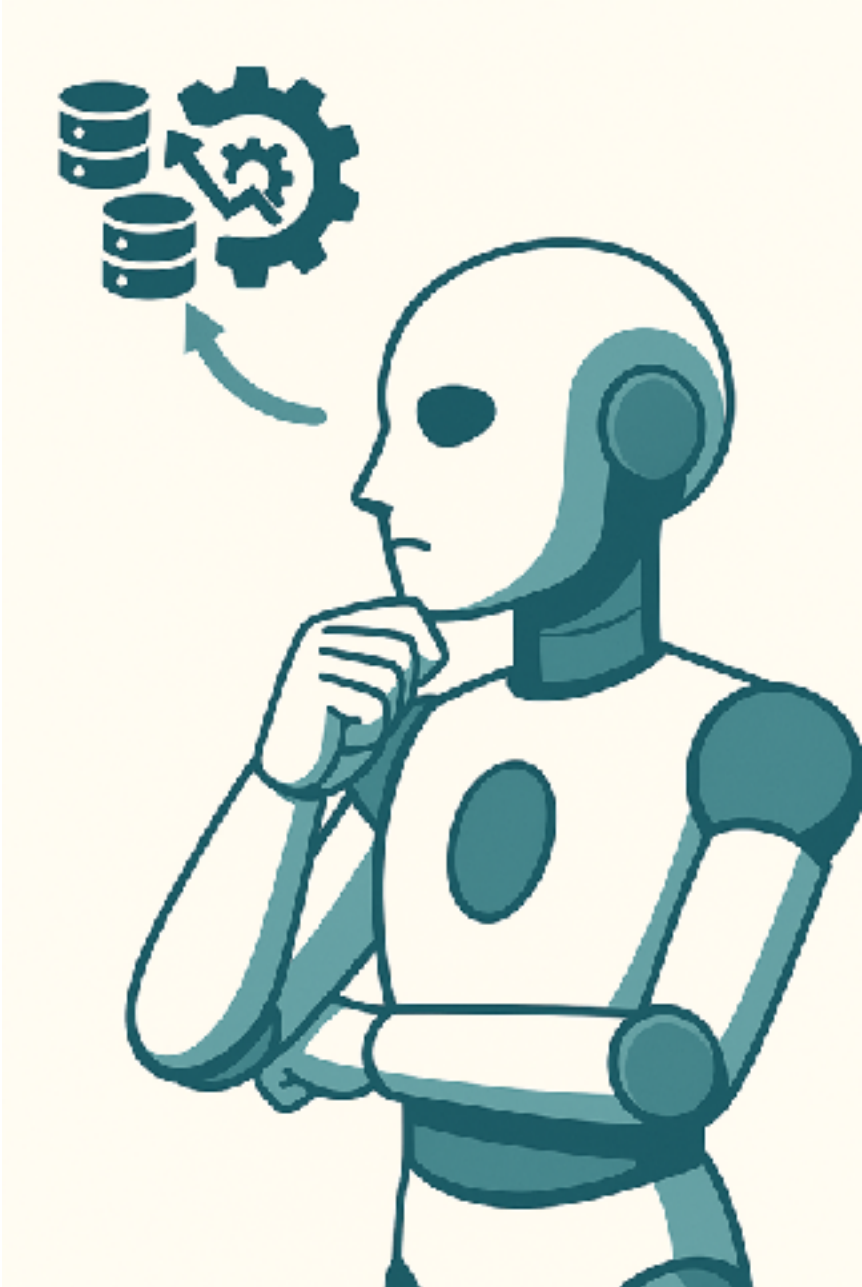
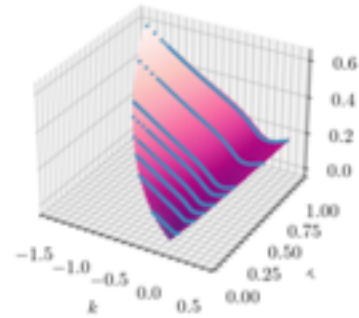
Imputation for prediction: beware of diminishing returns.
[Marine Le Morvan](#), [Gael Varoquaux](#)
Th, Apr 24, 11:00 +04 – [Poster Session 2](#)

Spider 2.0: Evaluating Language Models on Real-World Enterprise Text-to-SQL Workflows
[Fangyu Lei](#), [Jixuan Chen](#), [Yuxiao Ye](#), [Ruisheng Cao](#), [Dongchan Shin](#), [Hongjin SU](#), [Zhaoqing Suo](#), [Hongcheng Gao](#), [Wenjing Hu](#), [Pengcheng Yin](#), [Victor Zhong](#), [Caiming Xiong](#), [Ruoxi Sun](#), [Qian Liu](#), [Sida Wang](#), [Tao Yu](#)
Fr, Apr 25, 06:00 +04 – [Poster Session 3](#)
Fr, Apr 25, 11:42 +04 – [Oral Session 4B](#)

DiffPuter: An EM-Driven Diffusion Model for Missing Data Imputation
[Hengrui Zhang](#), [Liancheng Fang](#), [Qitian Wu](#), [Philip Yu](#)
Fr, Apr 25, 11:00 +04 – [Poster Session 4](#)

Optimal Transport for Time Series Imputation
[Hao Wang](#), [zhengnan li](#), [Haoxuan Li](#), [Xu Chen](#), [Mingming Gong](#), [BinChen](#), [Zhichao Chen](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)

Operator Deep Smoothing for Implied Volatility
[Ruben Wiedemann](#), [Antoine \(Jack\) Jacquier](#), [Lukas Gonon](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)



What We Need

To Advance Data Processing for Quant Research

The challenges:

- 🧹 **Clean and sanitize messy datasets** — outliers, anomalies, inconsistencies
- 🧠 **Impute missing data** — smart fills using priors, context, or learned structure
- 🕒 **Align and merge datasets** — via semantic similarity, metadata, or latent space
- 🧬 **Extract and standardize latent signals** — normalize across time, sources, assets
- 🎧 **Process multi-modal inputs** — tabular, text, audio, image, video
- 📐 **Ensure modeling consistency** — reproducible inputs, versioned, and robust

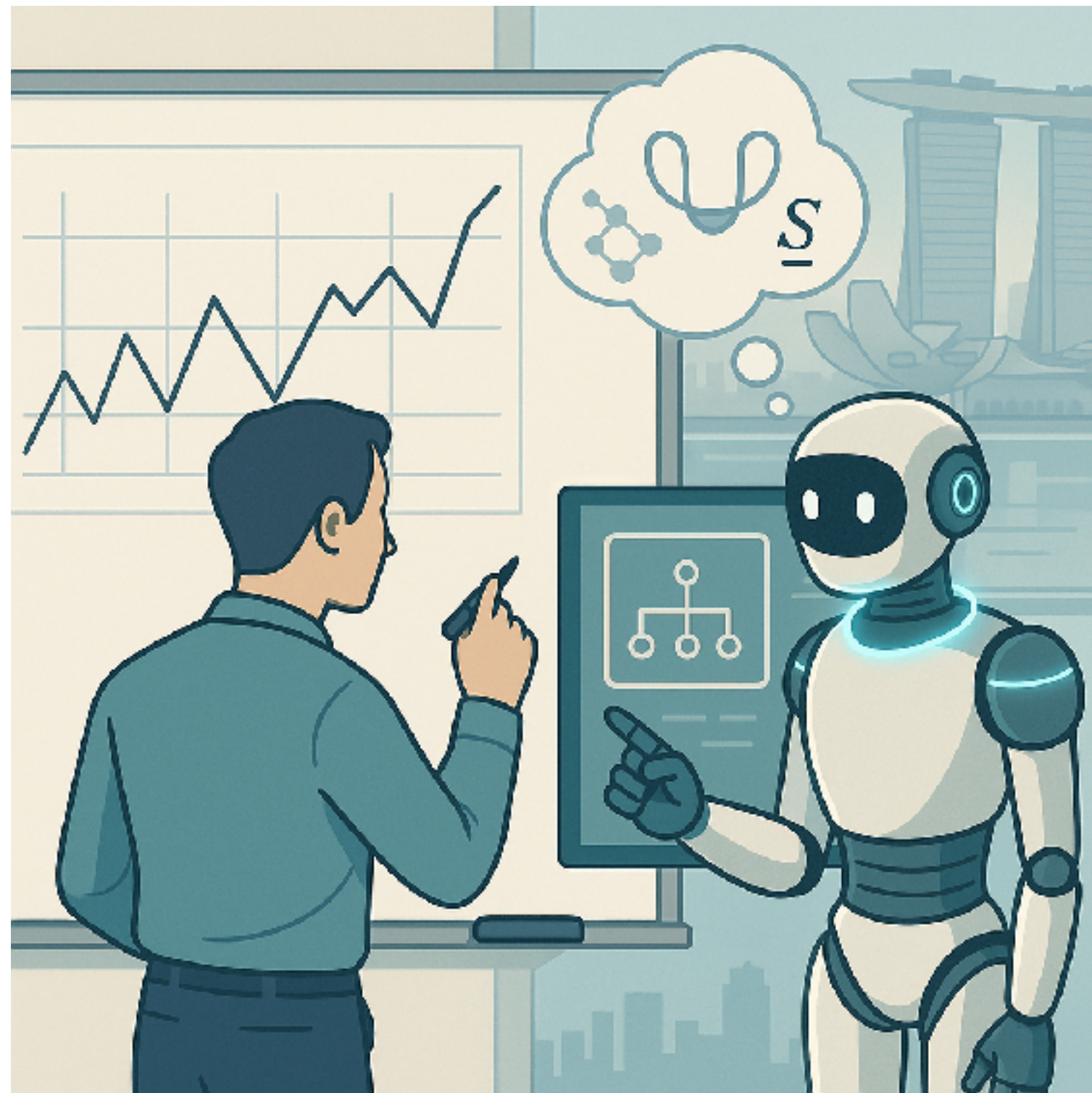
📌 *We're treating **data processing** as a **modeling challenge** — not just mundane cleaning, but extracting structure with precision.*

🔍 *Structuring **messy, diverse data** into reliable inputs is one of the central challenges in making systematic research scalable, reproducible, and impactful.*



🔮 How Quants Build Forecasting Models

Forecasting models turn cleaned data and ideas into signals — and ultimately, predictions.



1. Select Model Type(s)

- Choose among linear models, tree ensembles, neural nets, etc.
- Balance interpretability, generalization, and training cost

2. Choose Meaningful Inputs

- Select features that align with the target and your research hypothesis
- Favor features with intuitive and causal relevance

3. Train & Validate with Care

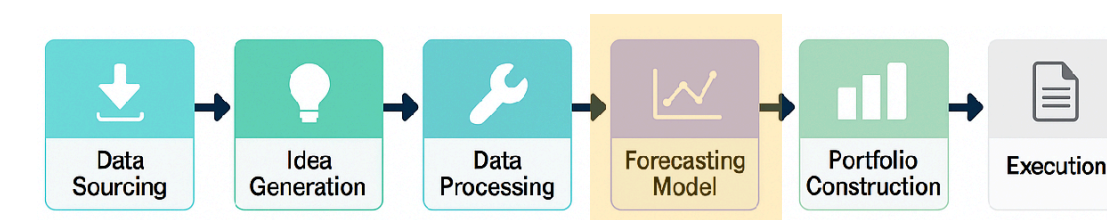
- Guard against lookahead bias, leakage, and hindsight traps
- Monitor drift, track performance decay, and re-train as needed

4. Evaluate Model & Signal Quality

- Use XAI tools to probe model behavior and attribution
- Run diagnostic tests to assess predictive strength and robustness

📌 *The challenge? Weak, fleeting signals in a highly competitive landscape.*

☁️ *Can we build models that react fast enough to stay ahead of the crowd?*



ML Research Spotlight (ICLR 2025)

Bridging Quant Forecasting Models and ML Research

Building forecasting models for quant investing is tough — signals are noisy, short-lived, and fiercely competed over.

ICLR 2025 papers offer new directions in **time-series modeling**, **uncertainty quantification**, and agent-based approaches — to help make forecasts more adaptive, expressive, and robust for systematic strategies.

🔍 What ML can help with:

Modeling weak, noisy signals

→ Learn from sparse, volatile data with limited signal-to-noise ratio

Adapting to changing regimes

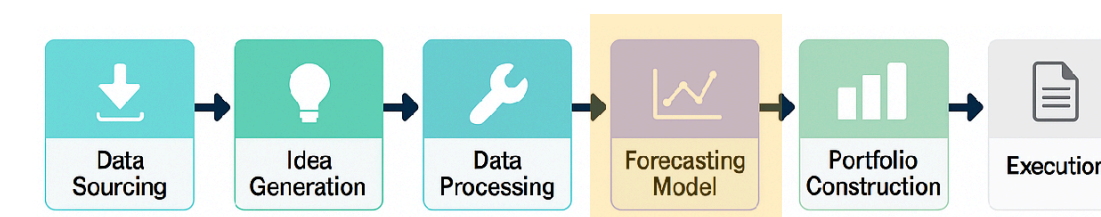
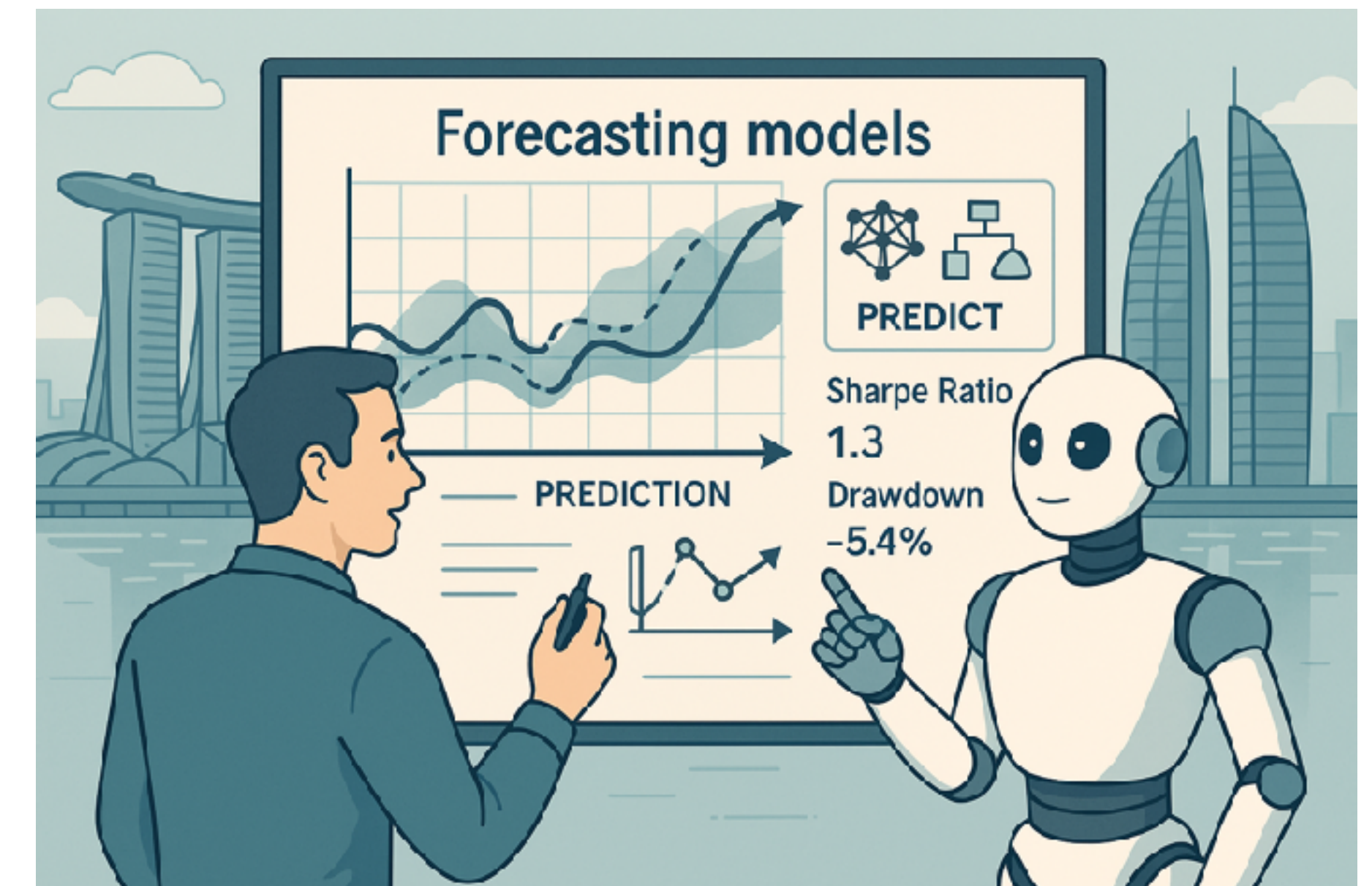
→ Detect shifts in patterns and evolving relationships over time

Quantifying uncertainty

→ Forecast with confidence intervals to guide risk-aware decisions

Integrating diverse inputs

→ Fuse fundamentals, prices, news, or alternative data into unified models



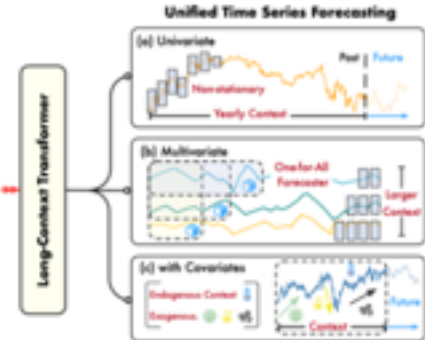
ML Research Spotlight (ICLR 2025)

Bridging Quant Forecasting Models and ML Research

Streamlining Prediction in Bayesian Deep Learning
[Rui Li](#), [Marcus Klasson](#), [Arno Solin](#), [Martin Trapp](#)
Fr, Apr 25, 06:00 +04 -- [Poster Session 3](#)

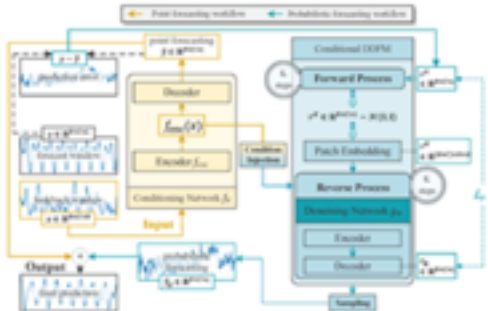
✓

Timer-XL: Long-Context Transformers for Unified Time Series Forecasting
[Yong Liu](#), [Guo Qin](#), [Xiangdong Huang](#), [Jianmin Wang](#), [Mingsheng Long](#)
Sa, Apr 26, 06:00 +04 -- [Poster Session 5](#)



✓

Diffusion-based Decoupled Deterministic and Uncertain Framework for Probabilistic Multivariate Time Series Forecasting
[Qi Li](#), [Zhenyu Zhang](#), [Lei Yao](#), [Zhaoxia Li](#), [Tianyi Zhong](#), [Yong Zhang](#)
Th, Apr 24, 06:00 +04 -- [Poster Session 1](#)



✓

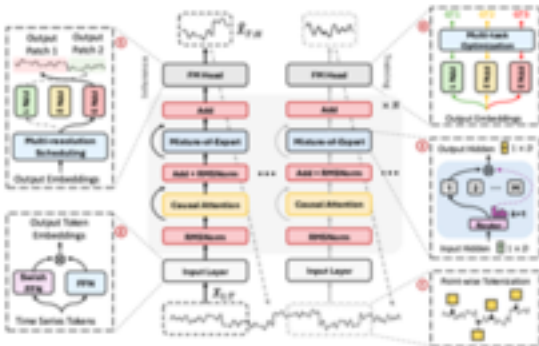
S4M: S4 for multivariate time series forecasting with Missing values
[Jing Peng](#), [Meiqi Yang](#), [Qiong Zhang](#), [Saxon \(Xiaoxiao\) Li](#)
Th, Apr 24, 06:00 +04 -- [Poster Session 1](#)

✓

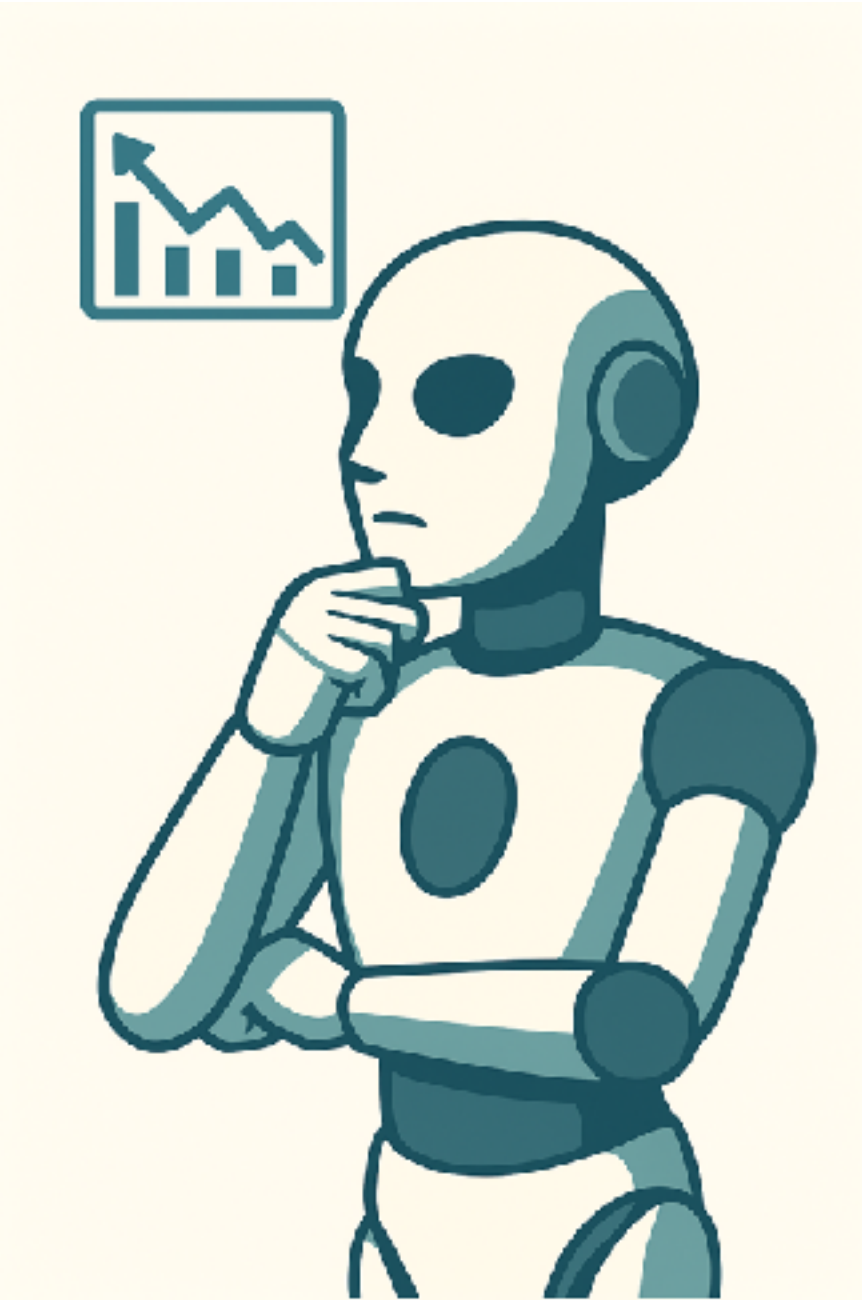
DyCAST: Learning Dynamic Causal Structure from Time Series
[Yue Cheng](#), [Bochen Lyu](#), [Weiwei Xing](#), [Zhanxing Zhu](#)
Fr, Apr 25, 06:00 +04 -- [Poster Session 3](#)

✓

Time-MoE: Billion-Scale Time Series Foundation Models with Mixture of Experts
[Xiaoming Shi](#), [Shiyu Wang](#), [Yuqi Nie](#), [Dianqi Li](#), [Zhou Ye](#), [Qingsong Wen](#), [Ming Jin](#)
Sa, Apr 26, 11:00 +04 -- [Poster Session 6](#)



✓



What We Need

To Advance Forecasting Models for Quant Research

The challenges:

-  **Model weak, noisy signals** — learn from short-lived, low-SNR patterns
-  **Adapt to changing regimes** — stay robust to shifts, rotations, breakdowns
-  **Quantify uncertainty** — forecast with confidence, not just point predictions
-  **Fuse diverse inputs** — model across price, fundamentals, news, alternative data
-  **Build bespoke models** — beyond off-the-shelf templates
-  **Validate signal robustness** — test across assets, horizons, and market regimes

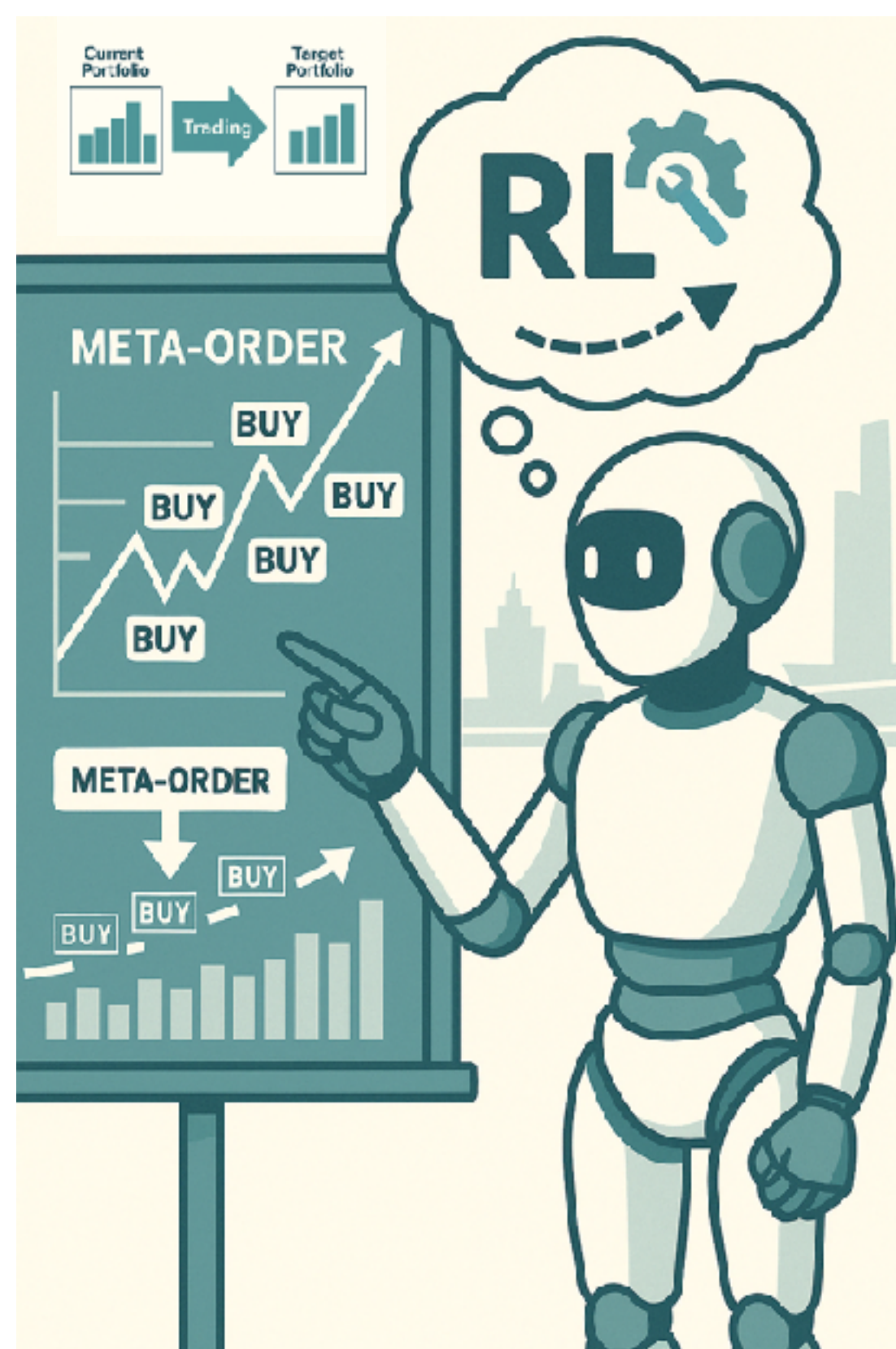
 *We consider forecasting as more than fitting a model — it's a dynamic, adversarial, high-stakes challenge.*

 *Tackling forecasting means learning under noise, drift, and competition — challenges quant researchers face on a daily basis.*



How Quants Turn Predictions into Portfolios

Portfolio construction & execution turn forecasts into positions — and real-world returns.



1. Construct Portfolios

- Translate model outputs into positions via optimization or trading rules
- Manage constraints: risk, turnover, liquidity, exposure

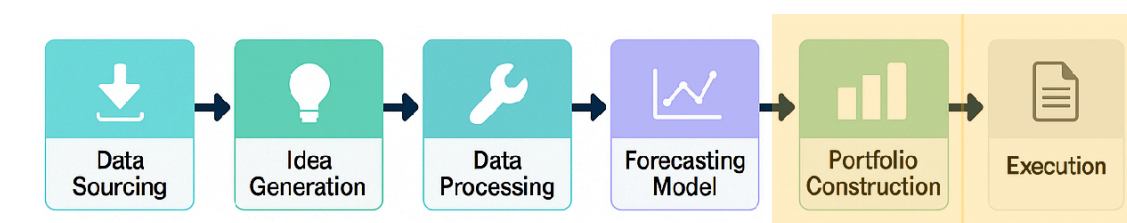
2. Execute Efficiently

- Minimize cost and slippage via smart scheduling or RL agents
- Adapt execution to market conditions in real time

*ICLR 2025 features several papers on **optimization** and **reinforcement learning** — key tools for translating forecasts into **optimal portfolios** and **executing trades** efficiently under real-world constraints.*

 *The challenge? Good predictions mean little if you can't turn them into profitable trades.*

 *Can we align forecasts with the real-world mechanics of trading?*



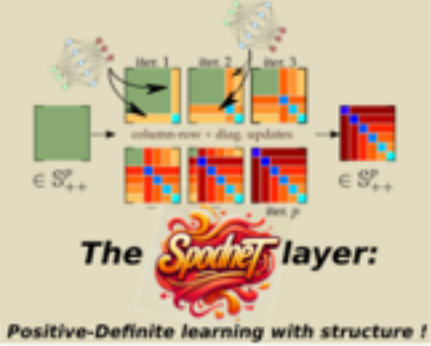
ML Research Spotlight (ICLR 2025)

Bridging Portfolio Construction & Execution and ML Research

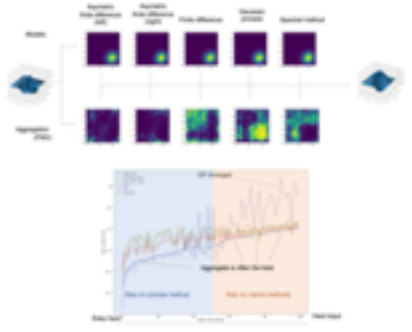
Deep Distributed Optimization for Large-Scale Quadratic Programming
[Augustinos Saravanos](#), [Hunter Kuperman](#), [Alex Oshin](#), [Arshiya Taj Abdul](#), [Vincent Pacelli](#), [Evangelos Theodorou](#)
Th, Apr 24, 06:00 +04 – [Poster Session 1](#)



Schur's Positive-Definite Network: Deep Learning in the SPD cone with structure
[Can Pouliquen](#), [Mathurin Massias](#), [Titouan Vayer](#)
Sa, Apr 26, 06:00 +04 – [Poster Session 5](#)



Minimal Variance Model Aggregation: A principled, non-intrusive, and versatile integration of black box models
[Theo Bourdais](#), [Houman Owhadi](#)
Fr, Apr 25, 06:00 +04 – [Poster Session 3](#)



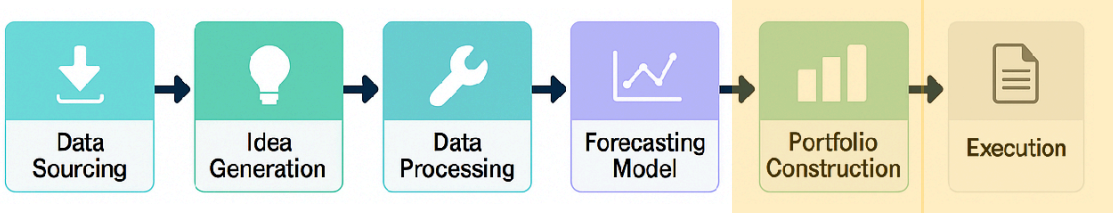
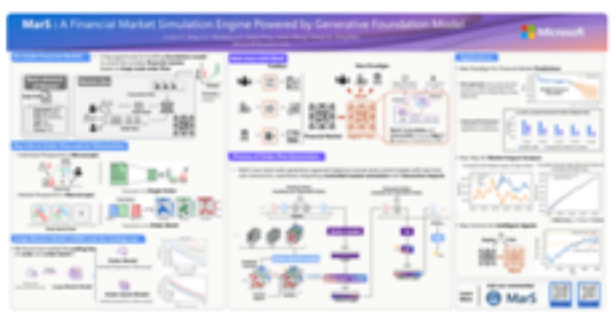
Online Reinforcement Learning in Non-Stationary Context-Driven Environments
[Pouya Hamadani](#), [Arash Nasr-Esfahany](#), [Malte Schwarzkopf](#), [Siddhartha Sen](#), [Mohammad Alizadeh](#)
Fr, Apr 25, 11:00 +04 – [Poster Session 4](#)



Safety Representations for Safer Policy Learning
[Kaustubh Mani](#), [Vincent Mai](#), [Charlie Gauthier](#), [Annie Chen](#), [Samer Nashed](#), [Liam Paull](#)
Fr, Apr 25, 06:00 +04 – [Poster Session 3](#)



MarS: a Financial Market Simulation Engine Powered by Generative Foundation Model
[Junjie Li](#), [Yang Liu](#), [Weiqing Liu](#), [Shikai Fang](#), [Lewen Wang](#), [Chang XU](#), [Jiang Bian](#)
Sa, Apr 26, 11:00 +04 – [Poster Session 6](#)



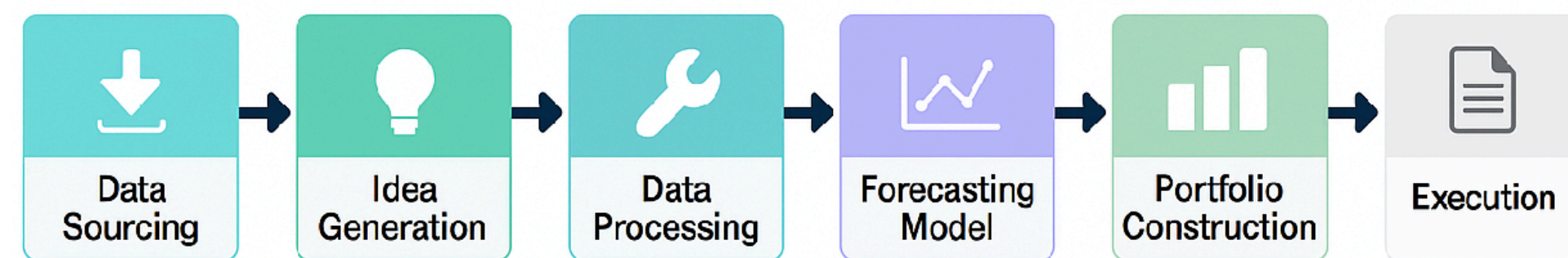
Part 2 — Rethinking the Research Workflow

The Vision: Toward an Autonomous Quant Research Lab

Rethinking the Quant Workflow with AI Agents and Modular Tools

Traditional:

Linear pipeline with minimal feedback

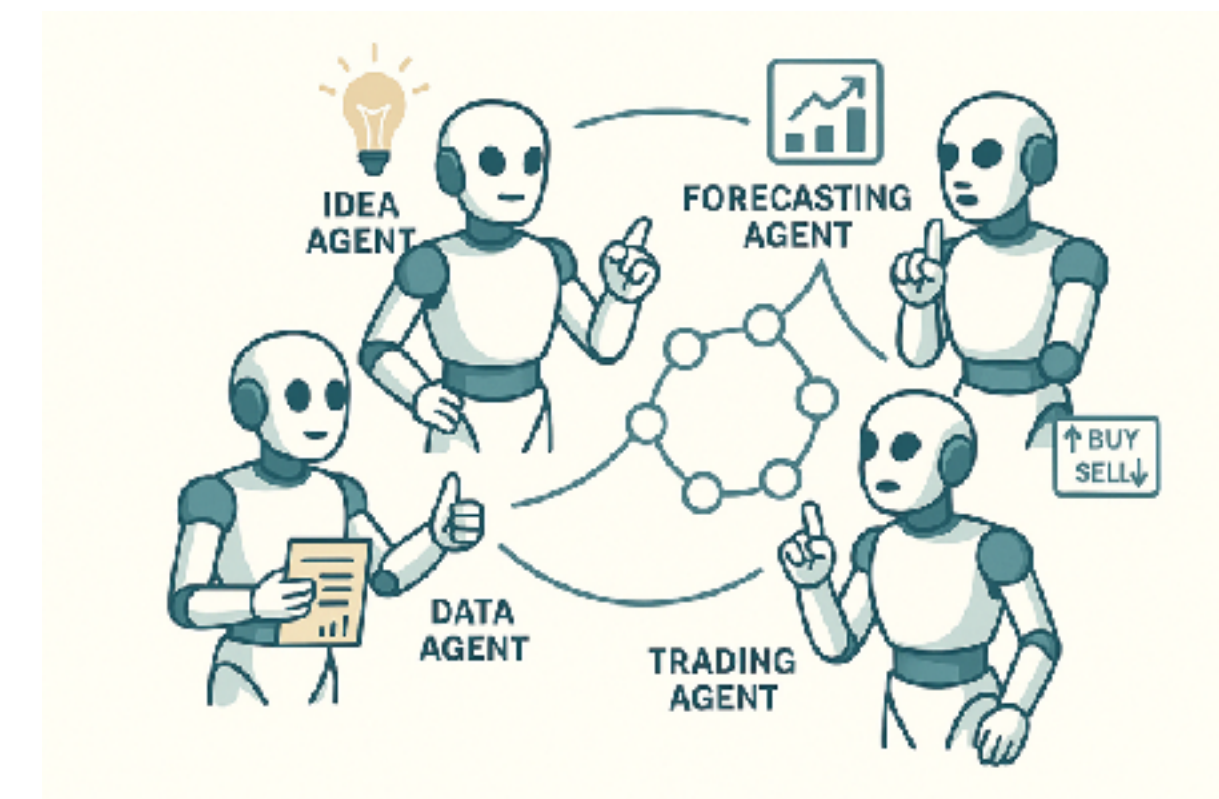


Vision:

We shift from a manual, linear workflow to an agent-driven, exploratory environment, where **tools, data, and hypotheses** can be **recombined autonomously** to search for systematic investment strategies — and implement them with precision and efficiency.

Vision:

Collaborative network of specialized agents



ML Research Spotlight (ICLR 2025)

Rethinking the Quant Workflow with AI Agents and Modular Tools

Benchmarking Agentic Workflow Generation

[Shuofei Qiao](#), [Runnan Fang](#), [Zhisong Qiu](#), [Xiaobin Wang](#), [Ningyu Zhang](#), [Yong Jiang](#), [Pengjun Xie](#), [Fei Huang](#), [Huajun Chen](#)

Th, Apr 24, 11:00 +04 -- [Poster Session 2](#)



WorkflowLLM: Enhancing Workflow Orchestration Capability of Large Language Models

[Shengda Fan](#), [Xin Cong](#), [Yuepeng Fu](#), [Zhong Zhang](#), [Shuyan Zhang](#), [Yuanwei Liu](#), [Yesai Wu](#), [Yankai Lin](#), [Zhiyuan Liu](#), [Maosong Sun](#)

Th, Apr 24, 11:00 +04 -- [Poster Session 2](#)



Flow: Modularized Agentic Workflow Automation

[Boye Niu](#), [Yiliao Song](#), [Kai Lian](#), [Yifan Shen](#), [Yu Yao](#), [Kun Zhang](#), [Tongliang Liu](#)

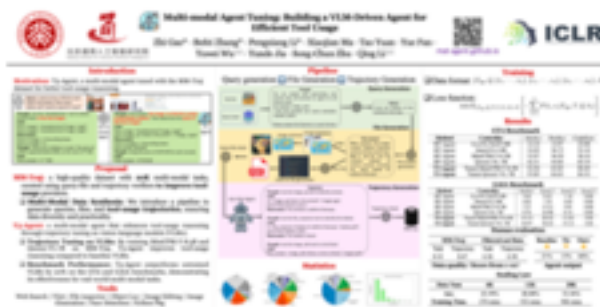
Sa, Apr 26, 06:00 +04 -- [Poster Session 5](#)



Multi-modal Agent Tuning: Building a VLM-Driven Agent for Efficient Tool Usage

[Zhi Gao](#), [Bofei Zhang](#), [Pengxiang Li](#), [Xiaojuan Ma](#), [Tao Yuan](#), [Yue Fan](#), [Yuwei Wu](#), [Yunde Jia](#), [Song-Chun Zhu](#), [Qing Li](#)

Th, Apr 24, 06:00 +04 -- [Poster Session 1](#)



ADAM: An Embodied Causal Agent in Open-World Environments

[Shu Yu](#), [Chaochao Lu](#)

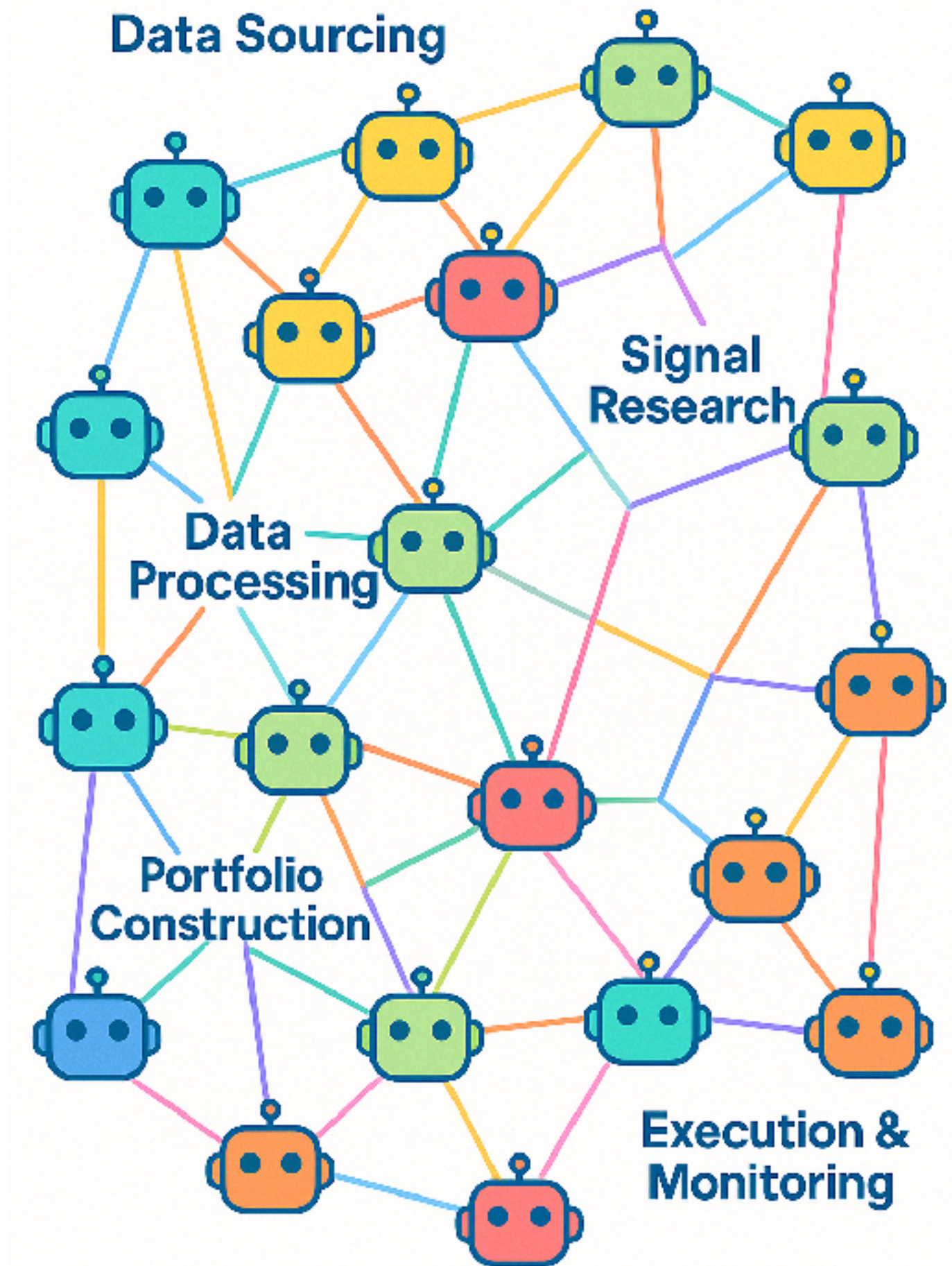
Th, Apr 24, 06:00 +04 -- [Poster Session 1](#)



Agent-Oriented Planning in Multi-Agent Systems

[Ao Li](#), [Yuexiang Xie](#), [Songze Li](#), [Fugee Tsung](#), [Bolin Ding](#), [Yaliang Li](#)

Fr, Apr 25, 06:00 +04 -- [Poster Session 3](#)



Thanks for your attention!

- Want to learn more or chat with the team?
→ Visit us at **Booth J09**
- Stay in touch or explore opportunities
→ Scan the QR code to get more info
and share your contact

