

MIT IAP 2026

Agentic AI in Finance

The Quest for an AI Quant

Sri Krishnamurthy, CFA, CAP

sri@quantuniversity.com | QCreate.ai

MIT | Jan 15th 2026

“Agents turn AI from answers into execution.”

McKinsey CEO Bob Sternfels says the firm now has 60,000 employees: 25,000 of them are AI agents

By Lakshmi Varanasi [Follow](#)



McKinsey & Company CEO Bob Sternfels says the firm is ramping up its use of AI agents in the workforce. Screenshot from CNBC

Jan 12, 2025, 5:41 AM ET

[Share](#) [Save](#)

Join Us Online!

Employee Benefit Research Institute & Milken Institute 2025 Retirement Symposium

"The Forces Transforming Retirement"

Date: Tuesday, March 11, 2025

Time: 8:45 AM - 5 PM E.T.

Admission: Complimentary

Visit www.EBRI.org to view the latest event agenda and to register!



Where We Are Going

If you can deploy agents in finance, you can deploy them anywhere.

SESSION MAP

- ▶ Finance is a hostile environment for naive AI
- ▶ Agents: what they are, and what they are not
- ▶ Architectures that survive reality
- ▶ Finance Use cases
- ▶ Case study 1: Regulated sandbox + SAR generator
- ▶ Case study 2: PE AI-Readiness score

What This Talk Is (and Is Not)

In finance, “cool demo” is the opposite of “deployable system.”

× NOT This

- × Prompt-engineering tips
- × “ChatGPT tricks” session
- × Laundry list of use cases

Demos impress. Systems deliver.

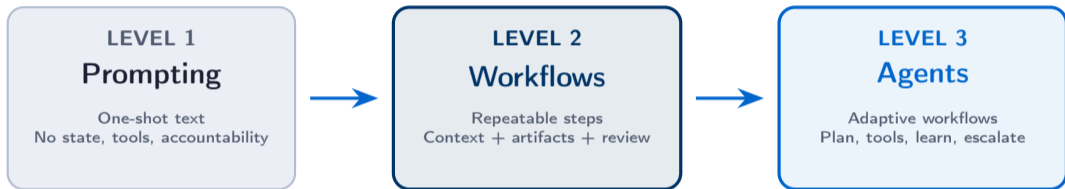
✓ YES This

- ✓ Systems thinking
- ✓ Workflows, tools, controls
- ✓ Deployable capability

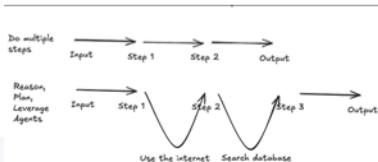
Mechanism + Governance + Value

From Prompting to Decision Systems

Value appears when AI sits inside the decision loop — not next to it.



IN FINANCE
Outputs must be **reproducible**, **reviewable**, and **attributable**



Why Finance Breaks Naive AI

Finance is where “mostly right” becomes “catastrophically wrong.”

1

Non-stationarity

Data generating process moves

2

Feedback loops

Predictions change markets

3

Adversarial incentives

Someone profits from mistakes

4

Regulation

Constraints are part of spec

5

P&L Attribution

Noisy, delayed, path-dependent

Finance punishes silent failure.
Controls matter, not just accuracy.

Why This Favors Agentic Systems

Multi-step workflows need more than single-shot answers.



Why agents fit:

- Complex multi-step workflows need planning & tool use
- Heavy reliance on code, APIs, datasets, backtests
- High demand for reproducibility, controls, model risk
- **Tools matter. Controls matter.**

AI Adoption: Implicit vs Explicit Roles

The winning path is "implicit first, explicit with controls."

IMPLICIT (Behind the Scenes)

- Coding assistance
- Intent detection & routing
- Next-best action suggestions
- Search, summarization, extraction

Lower risk, high adoption

EXPLICIT (Front Stage)

- AI as an assistant
- Automate / simplify workflows
- Replace brittle rule-based steps
- Operate with approvals + logging

Higher value, requires governance

Four Classes of Financial AI Workflows

In finance, the “use case” is a workflow with constraints.

1. Knowledge Synthesis

Research, compliance, disclosures
RAG + templates + approvals

2. Trust & Authentication

Fraud, biometrics, deepfake defense
Real-time scoring + step-up auth

3. Conversational Intelligence

Voice, sentiment, service QA
Privacy constraints + drift monitoring

4. Data Governance

Lineage, PII detection, refactoring
Deterministic rules + validated models

Workflow Class 1: Generative Knowledge

In regulated domains, generation without grounding is just risk.



- Draft market briefs, IC memos, client letters — faster cycle time
- RAG for factual grounding; templates for structure and tone
- Controls: source attribution, redaction, approval gates
- **Best for: analyst productivity + consistent narrative quality**

ChalGPT 4 -

You

Prepare a letter for a potential high net-worth investor who is interested in leveraging your investment advisory services. The investor is planning to retire in 10 years and has two kids ready to go to college in 2 and 3 years. He has a chronic health condition that requires significant medical expenditure on short notice and he needs to plan enough liquidity in his portfolio. List out the data points you would want your client to bring to the meeting. Start the letter by summarizing how the market did in 2023 (research the web for this information)

ChatGPT

I did a quick search for more information and here's what I found:

Dear [Investor's Name],

I trust this letter finds you well. As we look forward to discussing how our investment advisory services can be tailored to your unique financial goals and needs, I would like to provide a brief overview of the market dynamics we witnessed in 2023.

The past year has demonstrated resilience in financial markets, with the S&P 500 exhibiting a positive trend overall. Despite various challenges and geopolitical tensions, the markets have shown an upward trajectory with intermittent fluctuations. For instance, the S&P 500 index saw a rise from the beginning of the year, with periodic setbacks due to economic, political, and technological events. The index experienced notable growth, particularly in the second half of the year, closing at a higher level in December compared to January. Correspondingly, the 10-year U.S. Treasury Yield also saw varied movements, reflecting our changing economic landscape.



Investment Advisory Services: Engagement Letter

Dear [Investor's Name],

I trust this letter finds you well. As we look forward to discussing how our investment advisory services can be tailored to your unique financial goals and needs, I would like to provide a brief overview of the market dynamics we witnessed in 2023.

The past year has shown a positive trend in the S&P 500, despite various challenges. The index rose from the beginning of the year, with significant growth towards the end, and the 10-year U.S. Treasury Yield reflected the changing economic landscape with its varied movements.

For our upcoming meeting, it would be beneficial for you to bring the following data points:

1. A detailed overview of your current investment portfolio.
2. Statements of your investment accounts for at least the last three years.
3. An estimation of future college expenses for your children.
4. Records of your medical expenditures and any forecasts of future healthcare costs.
5. Your latest tax returns to understand your current financial standing better.
6. Any documentation pertaining to your retirement accounts and pensions.
7. Your estate planning documents, if available.

This information will enable us to construct a comprehensive and dynamic investment strategy that considers your upcoming retirement, your children's education, and the liquidity necessary for unforeseen medical expenses.

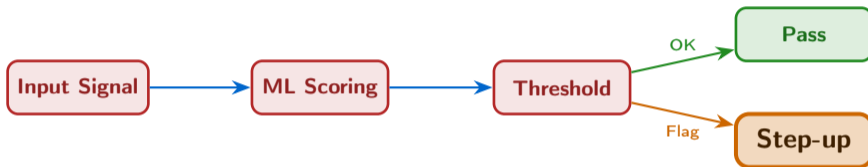
I look forward to our meeting and am committed to helping you navigate these important financial decisions.

Warm regards,

[Your Name]
[Your Title]
[Your Contact Information]

Workflow Class 2: Trust, Fraud, and Deepfake Defense

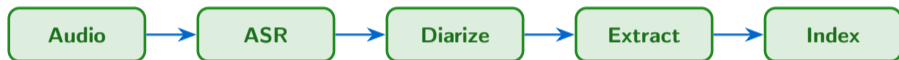
As identity becomes generative, authentication must become probabilistic.



- Voice biometrics: authenticate callers using speaker verification
- Synthetic media detection: flag AI-generated audio/video
- Operational shape: real-time scoring + step-up authentication
- **Controls: false positive management + auditability**

Workflow Class 3: Conversational Intelligence

If it's not searchable, it's not governable.



- ASR + diarization: searchable calls and meetings
- Extract sentiment, intent, and themes — at scale
- Finance use: client service QA, complaints, suitability red flags
- **Controls: privacy constraints + model drift monitoring**

Workflow Class 4: Data Governance + Privacy

Governance is a product feature, not an afterthought.



- Automated code & data lineage mapping for audit readiness
- PII detection and redaction across emails, notes, chat, calls
- Value: reduced manual documentation + lower privacy risk
- **Controls: deterministic rules + validated models + logging**

2025 Ecosystem Shift: Why Agents Are “Now”

The missing ingredient was not intelligence — it was orchestration.

Orchestration

LangGraph
CrewAI
AutoGen
OpenAI Agents SDK

Framework explosion

Secure Execution

E2B Sandboxes
Docling (parsing)
Code Interpreters
Isolated Runtimes

Safe tool use

Model Performance

Better reasoning
Tool calling
Structured output
Context windows

Outperform rules

Models now outperform brittle rule-based pipelines on many tasks



Distributed and intelligent information access in manufacturing enterprises through mobile devices

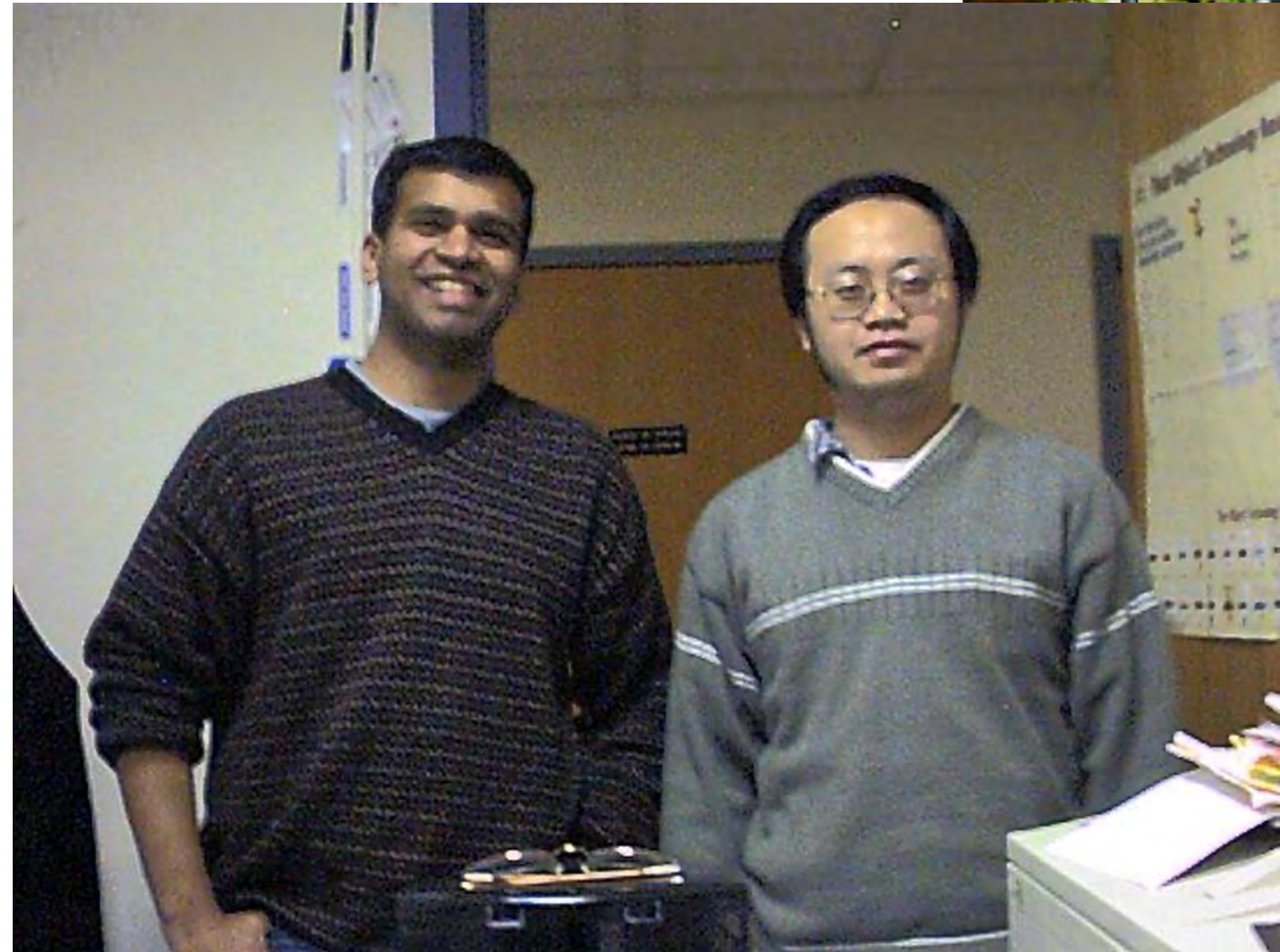
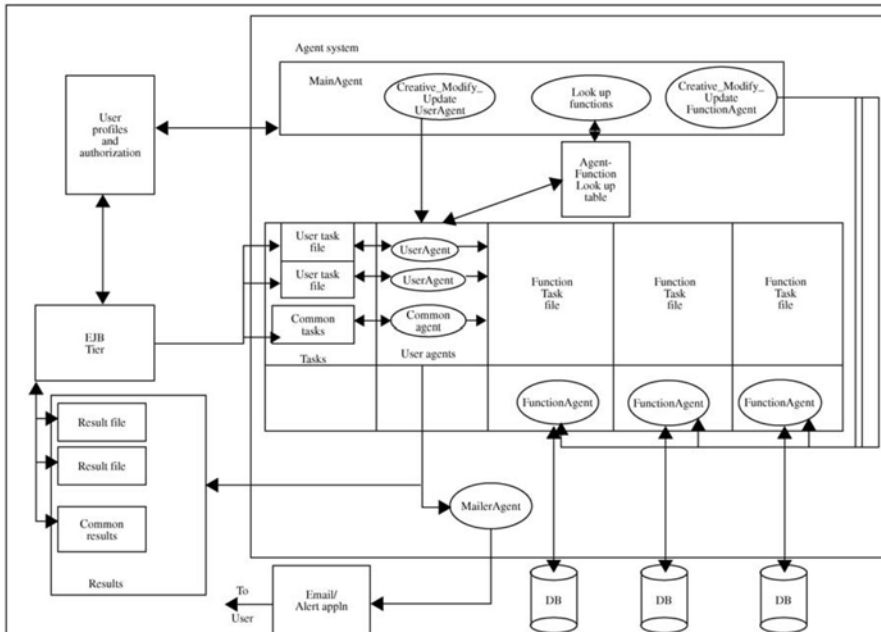
SRIKANTH KRISHNAMURTHY and IBRAHIM ZEID

MIME Department, Northeastern University, Boston, MA 02115
 E-mail: zeid@coe.neu.edu

Received April 2002 and accepted June 2003

Due to the distributed nature of today's enterprises and the growing number of employees who are mobile, newer models of communication are required to cater to the information needs of manufacturing personnel. Mobile applications such as information alerts, disconnected information processing and background processing are applications gaining significant attention in manufacturing companies. In this paper, we describe the framework and architecture of a mobile-agent-based system, which provides mobile information services to the workforce in motion in manufacturing enterprises. We emphasize on background and disconnected processing and demonstrate the use of a Mobile-agent-based system to achieve it. Recognizing the need to exchange information between various distributed information sources, we explore the use of XML-based agent communication and discuss how XML can be used for inter-agent communication. We describe the design and implementation of a prototype system and discuss application scenarios.

Keywords: Manufacturing, wireless communication, mobile agents, disconnected and background processing



manus Comprehensive Tesla Stock Analysis and Investment Insights [Log in](#)

I'd like a thorough analysis of Tesla stock, including:

Summary: Company overview, key metrics, performance data and investment recommendations
 Financial Data: Revenue trends, profit margins, balance sheet and cash flow analysis
 Market Sentiment: Analyst ratings, sentiment indicators and news impact
 Technical Analysis: Price trends, technical indicators and support/resistance levels
 Compare Assets: Market share and financial metrics vs. key competitors
 Value Investor: Intrinsic value, growth potential and risk factors
 Investment Thesis: SWOT analysis and recommendations for different investor types

Manus

I'll help you create a comprehensive analysis of Tesla stock. I'll gather the latest financial data, market sentiment, technical analysis, competitive comparisons, and develop investment recommendations. This will take some time to research thoroughly, but I'll work on it right away and provide you with a detailed report.

- Connected to datasource(s)
 - Get stock chart
 - Get stock holders
 - Get stock insights
 - Get stock profile
 - Get stock SEC filing
 - Get what analysts are saying of a stock

Create comprehensive Tesla stock analysis task list

```

Manus is using Terminal
Executing command: cd /home/ubuntu/tesla_analysis/data && python3 tesla_technical_analysis.py
shell3
ubuntu@sandbox:~/tesla_analysis/data $ cd /home/ubuntu/tesla_analysis/data && cd /home/ubuntu/tesla_analysis/data && python3 tesla_technical_analysis.py
Bad key font.manager.use_fontconfig in file /home/ubuntu/.config/matplotlib/matplotlibrc, line 33 ('font.manager.use_fontconfig : True')
You probably need to get an updated matplotlibrc file from https://github.com/matplotlib/matplotlib/blob/v3.10.1/lib/matplotlib/mpl-data/matplotlibrc or from the matplotlib source distribution
Performing technical analysis on Tesla stock...
Loaded historical price data with 178 records
Data preparation complete
Calculating moving averages...
Calculating MACD...
Calculating RSI...
Calculating Bollinger Bands...
Identifying support and resistance levels...
Current price: $284.65
Key resistance levels: ['$381.59']
Key support levels: ['$175.79', '$123.18', '$12.52']
Analyzing volume patterns...
Analyzing price trends...
Calculating momentum indicators...
RSI: 59.31 - Neutral
  
```

I'd like a thorough analysis of Tesla stock, including:

Summary: Company overview, key metrics, performance data and investment recommendations

Financial Data: Revenue trends, profit margins, balance sheet and cash flow analysis

Market Sentiment: Analyst ratings, sentiment indicators and news impact

Technical Analysis: Price trends, technical indicators and support/resistance levels

Compare Assets: Market share and financial metrics vs. key competitors

Value Investor: Intrinsic value, growth potential and risk factors

Investment Thesis: SWOT analysis and recommendations for different investor types

Tesla Stock Analysis Dashboard

Comprehensive analysis of Tesla's financial performance, market position, and investment outlook

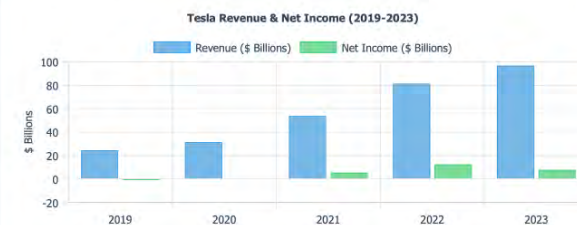
Executive Summary

Tesla maintains a dominant position in the EV market with significant technological advantages, strong brand recognition, and expanding manufacturing capacity. While the company has demonstrated impressive revenue growth (42.3% CAGR over 5 years) and achieved consistent profitability, our discounted cash flow analysis suggests the current stock price exceeds our calculated intrinsic values across multiple scenarios.

Investment Recommendations

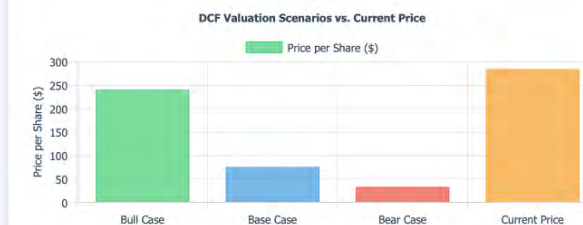
Growth	Hold/Selective Buy
Value	Avoid
Income	Avoid
Speculative	Selective Buy/Trade
ESG	Buy

Financial Performance



Revenue CAGR (5Y)	Gross Margin (2023)	Operating Margin (2023)	Net Margin (2023)
42.3%	19.0%	9.6%	8.2%

Valuation Analysis



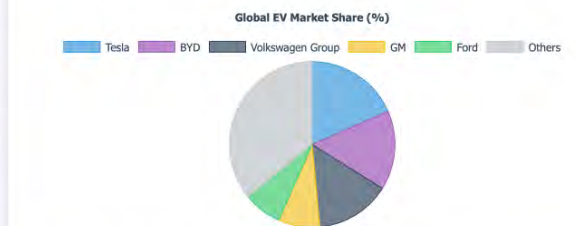
P/E Ratio	P/S Ratio	EV/EBITDA
80.2x (Industry: 10.3x)	9.3x (Industry: 0.8x)	44.5x (Industry: 8.2x)

Technical Analysis



RSI (14)	MACD	Key Support	Key Resistance
59.31 Neutral	26.02 Bullish	\$175.79	\$381.59

Market Position



EV Market Share	Market Cap	Annual Deliveries
18.4% Market Leader	\$900.5B (Toyota: \$285.6B)	1.81M

Why Quant Finance Is Ideal for Agents

Quant is “workflow-intensive” — which is exactly what agents are for.



Why agents fit quant:

1. Complex multi-step workflows: data → hypothesis → model → validation → action
2. Heavy reliance on tools: code, APIs, libraries, datasets, backtests
3. High demand for transparency: reproducibility, controls, model risk

Agents vs Rule-Based Automation

Agents are flexible — your controls must be non-negotiable.

Rule Engines

- × Deterministic and brittle
- × Static logic
- × Hard to adapt to new regimes
- × Fails silently when assumptions drift

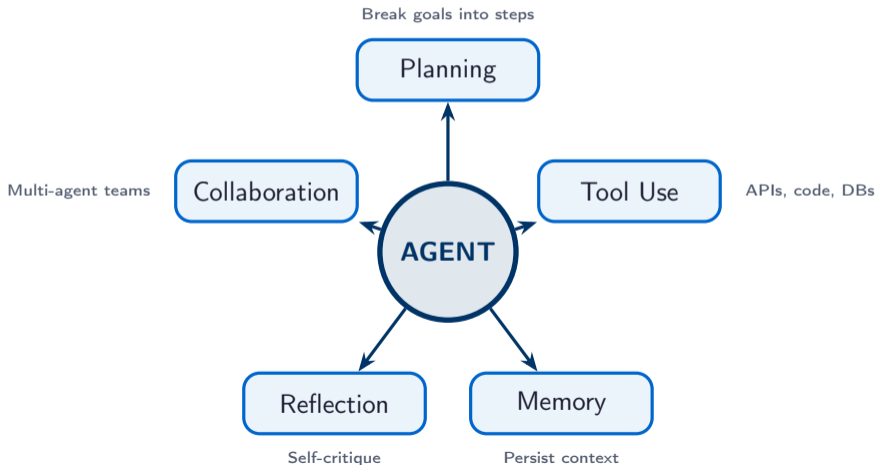
Agents

- ✓ Context-sensitive execution
- ✓ Tool-using and stateful
- ✓ Can plan and recover from errors
- ✓ Still requires hard guardrails

“Adaptable” is not “uncontrolled.” The design target is **bounded flexibility**.

The Five Capabilities That Make an Agent

A model answers. An agent executes.



Planning

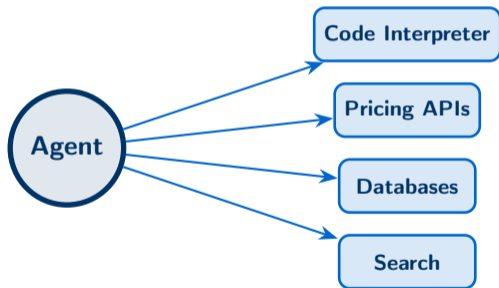
Planning turns “open-ended” into “auditable.”



- Treat goals as projects: tasks, dependencies, checkpoints
- Finance use: research plans, IC memo outlines, validation plans
- **Key control: stop conditions and explicit acceptance criteria**

Tool Use

Tools are power — permissions are safety.



Controls

- Allow-lists
- Scoped permissions
- Deterministic wrappers
- Audit logging

Memory

Memory is leverage — and also a new attack surface.

Vector Memory

Retrieval of prior context
Papers, policies, history
Semantic search

“What have we seen before?”

Structured Memory

State machine for workflow
Track progress
Enforce constraints

“Where are we now?”

Key control: **provenance**, **retention**, and **contamination checks**

Reflection

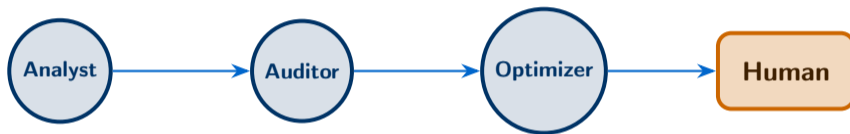
Reflection reduces errors — tests prevent them.



- Self-critique: check reasoning consistency and missing evidence
- Finance use: model sanity checks, narrative coherence checks
- **Key control: explicit tests and constraints beat “self-judgment”**

Collaboration

In regulated work, collaboration is your accountability mechanism.



- Multi-agent teams: analyst, auditor, optimizer, reviewer
- Human collaboration: clarify uncertainty, seek approvals, propose options
- **Key control: clear responsibility boundaries and audit trails**

Agent Design Patterns (Practical Taxonomy)

Patterns are how you make agents predictable.

Planner–Executor

Plan tasks, execute tools, report

Hierarchical

Supervisor coordinates specialists

Collaborative

Peer review and consensus

Recursive

Decompose sub-problems recursively

Self-healing

Error capture and retry

Patterns = reusable governance structures

Planner–Executor Architecture

This pattern scales because it separates “thinking” from “doing.”

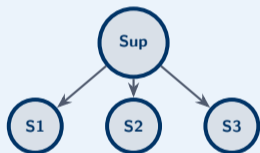


- Planner produces a task list with explicit checkpoints
- Executor runs tools and returns artifacts (code, tables, citations)
- Validator enforces constraints before delivery

Hierarchical vs Collaborative Agents

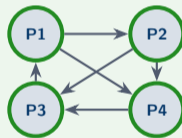
Use hierarchy for control; use collaboration for quality.

Hierarchical



Best when tasks are modular
Clear accountability
Good for regulated workflows

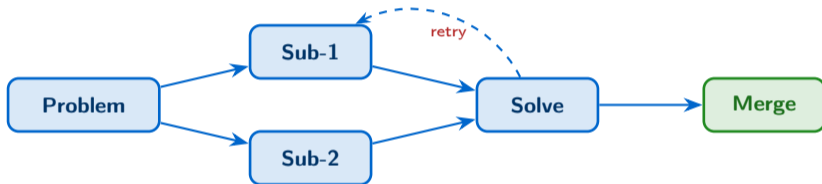
Collaborative



Best when uncertainty is high
Peer review reduces hallucinations
Voting for recommendations

Recursive + Self-Healing Agents

Resilience is a feature — not an incident response plan.



- Recursive decomposition: split hard problems into solvable units
- Self-healing: error capture → retry trees → fallback tools
- **Control: bounded retries + deterministic fallback order**

Canonical Agent Architecture (Finance-Grade)

Bounded autonomy + hard controls = deployable agents.



Control Surfaces

- Tool allow-lists & permissions
- Policy/threshold rules
- Audit trails & versioned artifacts
- Human approvals for regulated outputs

Stochastic OK: drafting, hypothesis
Never: unchecked execution

Case Study: FOMC Research Agent

The edge is not prediction — it's disciplined synthesis at speed.

GOAL

Turn transcripts into tradable, auditable macro signals

INPUTS

Statement + Minutes + Press Q&A + History

OUTPUTS

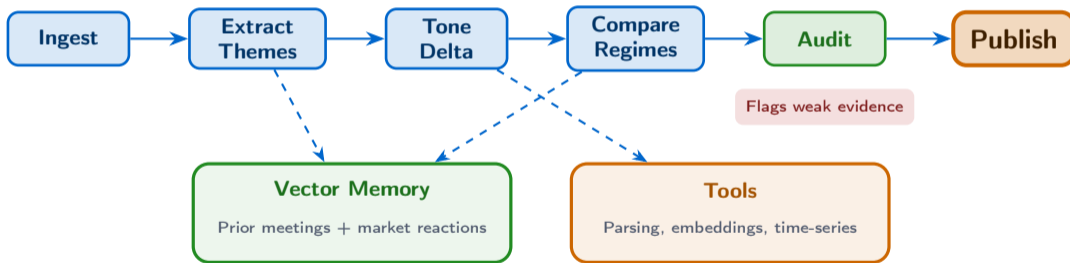
Themes, tone shifts, surprises, confidence + citations

CONSTRAINT

Speed matters, but “wrong fast” is still wrong

FOMC Agent: System Architecture

Architecture beats prompts when the task is multi-step.



FOMC Agent: Execution Trace

Traceability is how you earn trust in stochastic systems.

- **Step 1**
Planner creates checklist
- **Step 2**
Tools parse transcript
- **Step 3**
Retriever pulls history
- **Step 4**
Validator checks citations
- **Step 5**
Human reviews deltas

themes, tone, deltas, surprises, evidence

structured sections + speaker turns

comparable historical episodes

citations, deltas computed, uncertainty labeled

trade-relevant deltas before distribution

FOMC Agent: Failure Modes & Escalation

The right question is not “did it hallucinate?” — it’s “did it stop?”

Failure Modes

- × Hallucinated regime narratives
- × Over-weighting single quote
- × Confusing tone with policy
- × Ignoring microstructure

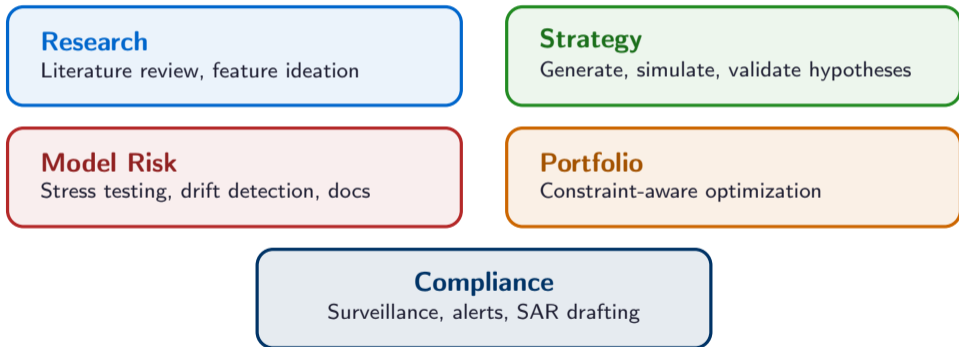
Escalation Protocol

- Low evidence → analyst review
- Conflicts → multi-agent adjudication
- High impact → citations + approval

NEVER autonomous trading

Where Agents Show Up in Finance (Beyond Macro)

Agents are best where work is repeatable but not fully specifiable.



Trading Strategy Generation Agent

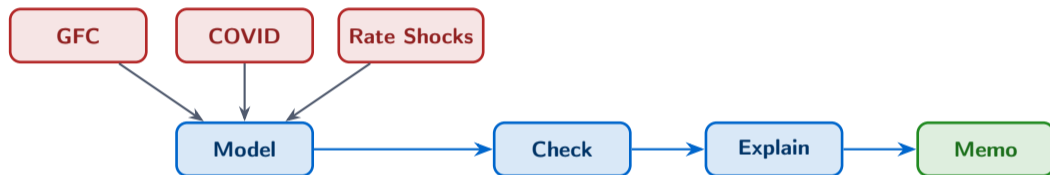
Backtest everything. Trust nothing without validation.



- Inputs: historical data + narratives + templates
- Backtesting: Pandas + Matplotlib
- **Controls: train/test separation, leakage checks**

Model Risk Validation Agent

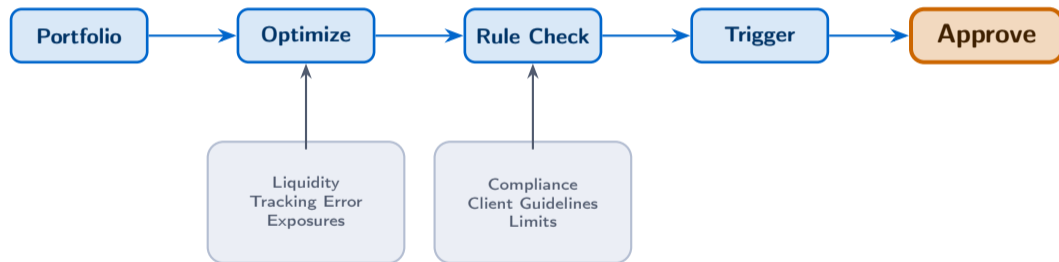
Simulate edge cases. Document everything.



- Assumptions: regime stability, factor crowding, drift
- Explainability: SHAP/LIME
- **Outputs: validation memo + artifacts + sign-offs**

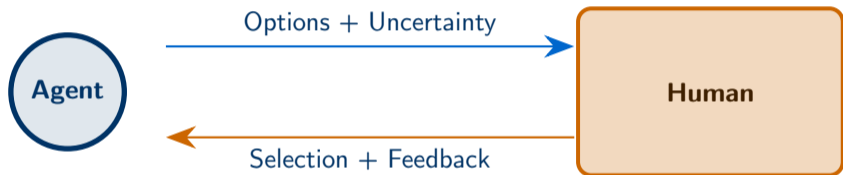
Portfolio Management Agent

Optimize under constraints. Never execute without approval.



Human-Agent Collaboration

Agents propose. Humans choose.



- Surface uncertainty + ask questions
- “What-if” scenarios with justification
- **Approvals as first-class feature**

Why AI Pilots Stall: Integration Beats Intelligence

Scaling is an operating model problem, not a model problem.

95%

of GenAI pilots fail

MIT NANDA REPORT 2025

Common Causes

- × Weak problem framing
- × No governance
- × No workflow integration

“Pilot success” ≠ “production readiness”

Agent-Specific Failure Modes

Agents fail as systems — so you must validate systems.

Interaction Risk

Errors compound across steps

Memory Contamination

Bad context becomes durable

Tool Misuse

Unsafe actions via broad permissions

Recursive Amplification

Retries “invent” confidence

Mitigation

Bounded retries + deterministic checks + human gates

Underwriting AI Execution Risk

Execution risk is the real “model risk” in enterprise GenAI.



Is there a governed sandbox to turn ideas into production?



Are approvals and audit trails built into the workflow?



Can the organization validate agentic systems end-to-end?



Are there metrics for drift, failure modes, reviewer effort?

Governance Frameworks You Will Be Asked About

Compliance is not a blocker — it's your architecture spec.

SR 11-7

Fed Model Risk Management

- Model development & implementation
- Independent validation
- Governance, policies, controls
- Documentation & monitoring

NIST AI RMF 1.0



Trustworthy AI Framework

- Govern — accountability
- Map — risk identification
- Measure — quantification
- Manage — mitigation

Regulated Production



QCreate & SAR Case Study

“The product is not generation — it’s defensible generation.”

Boris Cherny  @bcherny · Dec 27  ...







Correct. In the last thirty days, 100% of my contributions to Claude Code were written by Claude Code

 62  336  2K  522K  

Rich Hickson  @RichHickson · Dec 27  ...

Isn't that the start of skynet? Claude code building claude code? How long until it decides we as humans are no longer needed?



 6  1  85  23K  

Boris Cherny  @bcherny  ...

Possibly. This is why alignment science and mechanistic interpretability are so important to get right.



Citizens

Who We Are

Citizens Financial Group, Inc. is one of the oldest and largest financial services firms in the United States. Citizens offers retail and commercial banking products and services to individuals, small businesses, middle-market companies, large corporations, and institutions. With 1,000 branches and total assets totaling \$220 billion, Citizens is continually innovating and partnering with the world's most promising FinTechs.

Our Challenges

GenAI Sandbox for Ideation in Regulated Environments

Citizens is exploring the development of a secure GenAI sandbox platform to accelerate experimentation with AI-powered tools in a regulated banking environment. The solution should support configurable access, audit-ready logging, and prompt versioning, while enabling users to upload and process documents and datasets. It should also offer tailored user experiences — from onboarding guides for business users, to agentic workflow support for advanced developers — ensuring safe, scalable innovation across the enterprise. By enabling responsible experimentation, this platform will help Citizens unlock new efficiencies, foster cross-functional collaboration, and accelerate GenAI solution development.

Why Sandboxes Fail

The usual failure: too permissive OR too restrictive.

Too Permissive

Compliance breach

Too Restrictive

No adoption

No Lineage

Can't graduate to production

No Audit Trail

Can't pass model risk review

Regulated Sandbox Requirements

Build governance into the architecture, not the policy.

- 1 Isolated execution environments
- 2 Versioned artifacts (prompts, sources, outputs)
- 3 Built-in validation (rules, tests, redaction)
- 4 Approval workflows
- 5 Immutable audit logs

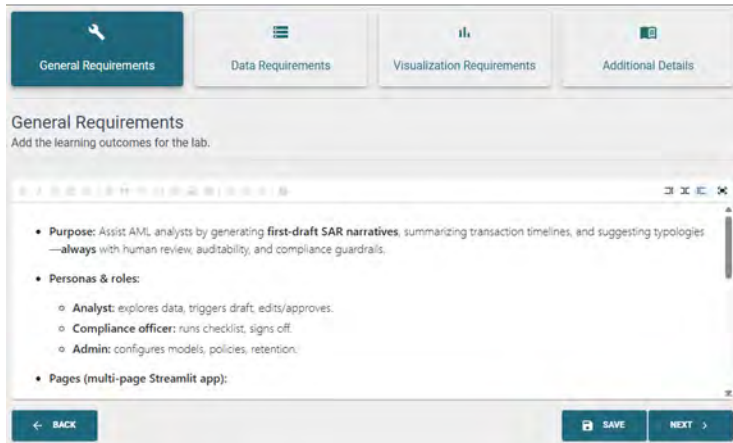
Workflow



Design your objective with constraints

Generate Specs

Create Code



Notebook Overview

This Jupyter Notebook demonstrates an AI-assisted Suspicious Activity Report (SAR) **narrative generator** for AML analysts—covering case intake, data exploration, LLM-powered drafting, review, compliance checks, and export. Human oversight, auditability, and compliance remain central.

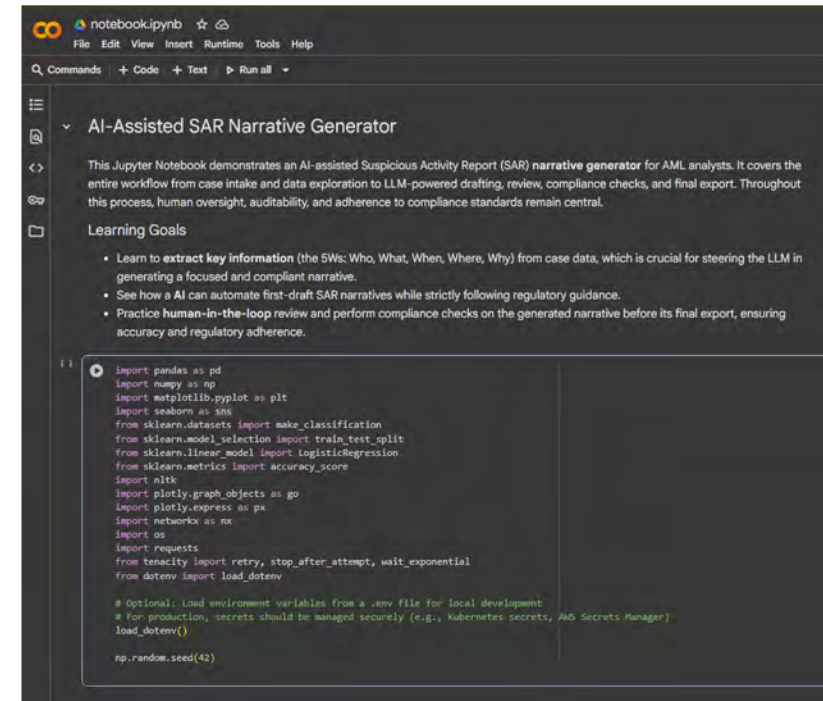
Learning Goals

- See how a **direct LLM call** can automate first-draft SAR narratives while following regulatory guidance (no RAG).
- Learn to **extract key information** (5Ws) from case data to steer the narrative.
- Practice **human-in-the-loop** review and compliance checks before export.

Code Requirements

Libraries

- pandas
- numpy
- matplotlib.pyplot
- seaborn
- sklearn (datasets, model selection, linear models, metrics)
- nltk
- plotly (e.g., plotly.express)
- **NEW (replace RAG/transformers deps):**
 - requests (HTTP call to your LLM endpoint)
 - tenacity (optional, for retry logic)



QCreate: Regulated Content Creation Sandbox

The product is not generation — the product is defensible generation.

GOAL

Create regulated artifacts with full traceability

CONTROLS

Tool allow-lists, redaction, citation requirements, approvals

OUTPUTS

Versioned drafts + validation reports + reviewer sign-offs

PRINCIPLE

Governance-by-construction (not governance-by-policy)



SAR Case Study

Suspicious Activity Report Generation

SARs are regulatory filings that document potentially suspicious financial activity for law enforcement review.

The hardest test case for agentic AI in compliance.

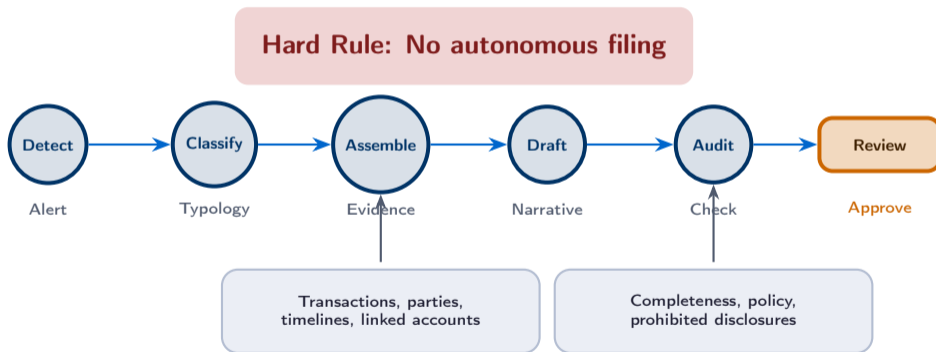
Why SAR Generation Is a Hard AI Problem

SARs are where LLMs meet law, deadlines, and accountability.

- SARs blend structured facts + narrative (“modus operandi”)
- High cost of false positives (noise) and false negatives (misses)
- Regulators and law enforcement depend on narrative quality
- Process is audited: who decided, based on what, when

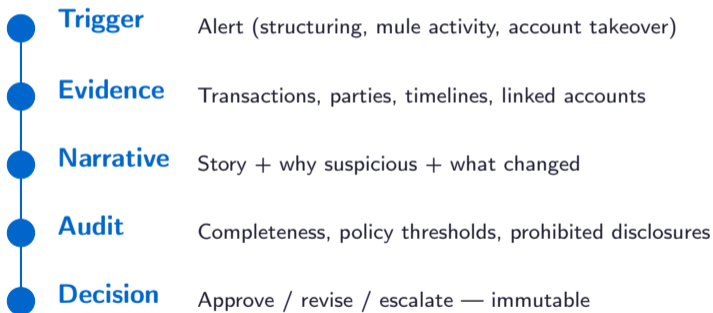
Agentic SAR Generation: Workflow

No autonomous filing — human sign-off required.



SAR Generator: Execution Trace

If you can't replay the SAR decision, you can't defend it.



Governance by Construction (SAR Case)

In compliance, the audit trail is the deliverable.

Lineage

prompt → sources → evidence
→ narrative → decision

Controls

Redaction, confidentiality rules,
approval gates

Metrics

Precision/recall, reviewer effort,
time-to-file

Outcome

Examiner-ready artifacts;
faster review with less noise

PE AI-Readiness Score

How can quantifying AI Readiness assist in PE, from Deal Screening through Exit!

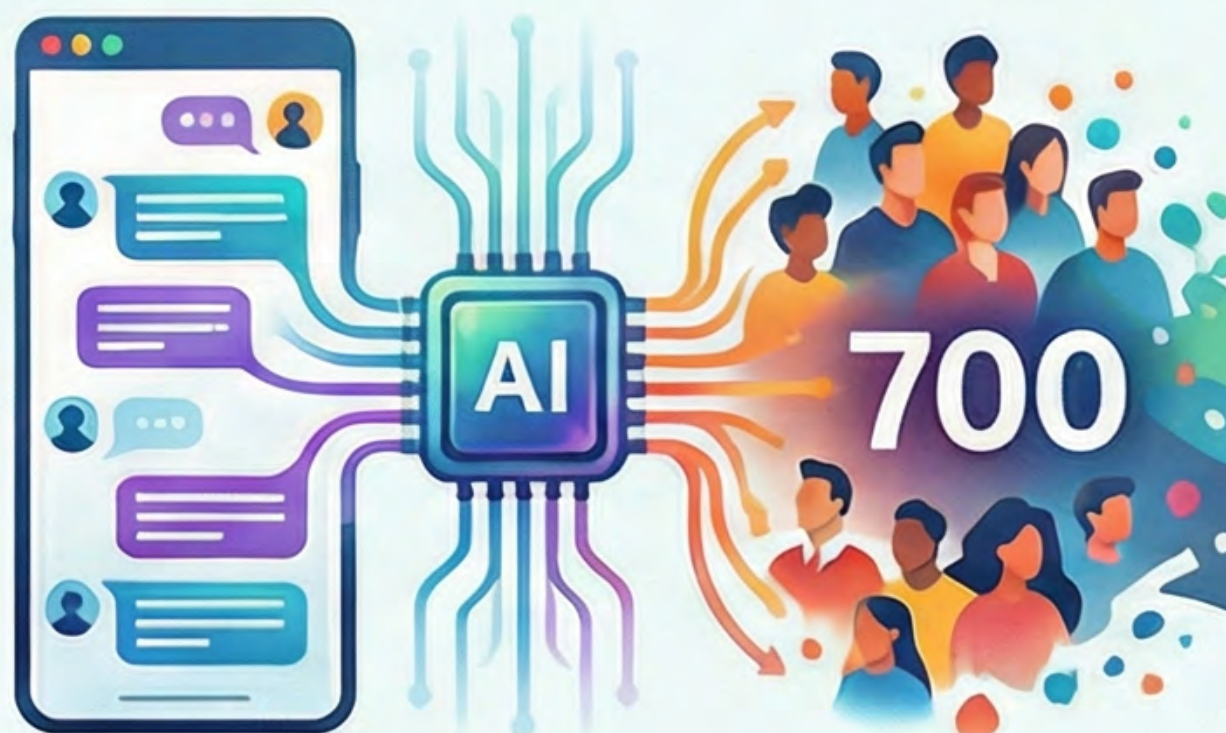
"If it can't be quantified, it can't be valued."

The AI Workforce Takeover: 4 Corporate Case Studies

Major corporations are publicly acknowledging their use of AI to automate tasks traditionally performed by humans. This has led to significant reductions in headcount, hiring freezes for specific roles, and shifts in workforce strategy, as illustrated by these prominent examples.

Klarna: AI Performs Workload of 700 Full-Time Agents

The AI assistant handled two-thirds of all customer service chats in its first month.



IBM: Pauses Hiring for Roles AI Can Replace

Leadership anticipates that approximately 7,800 back-office jobs could be replaced by AI over time.



Salesforce: AI Agents Enable Cutting ~4,000 Support Roles

The company reduced its support staff from approximately 9,000 to 5,000 employees.



Amazon: CEO Foresees Fewer Corporate Jobs Due to AI



The company expects its corporate workforce to shrink as “agentic” AI drives productivity.

The PE Decision Function

PE doesn't underwrite "potential" — it underwrites execution probability.

1

Can AI move EBITDA?

2

Can it happen within holding period?

3

Is execution institutional or individual-dependent?

4

Will buyers believe it is durable at exit?



The Org-AI-R Score

A Framework for Quantifying Enterprise AI Value

Org-AI-R is a parametric framework that measures an enterprise's readiness to deploy AI for competitive advantage and value creation.



The AI Value Gap
Only 26% of companies have developed the capabilities to **generate tangible value from AI**, highlighting the need for a rigorous assessment tool.

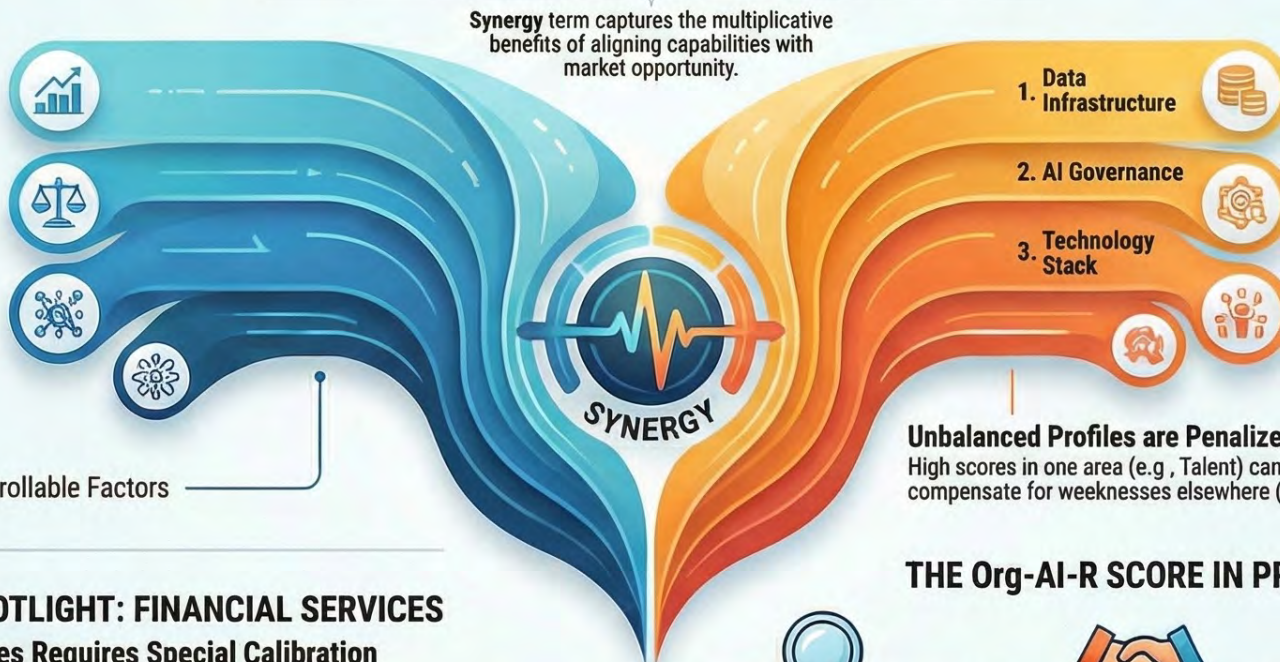
$$\text{Org-AI-R} = \alpha \cdot \text{Idiosyncratic Readiness} + (1-\alpha) \cdot \text{Systematic Opportunity} + \beta \cdot \text{Synergy}$$

SYSTEMATIC OPPORTUNITY (The Market You're In)

Macro-level AI potential. Organizations can position themselves to capture this, but cannot create it.

1. Industry AI Disruption Potential (30%)
2. Regulatory Environment (20%)
3. Competitive Dynamics (25%)
4. AI Talent Market (15%)
5. Technology Ecosystem (10%)

Key Finding: Separating Controllable vs. Uncontrollable Factors



IDIOSYNCRATIC READINESS (The Capabilities You Build)

Specific, internal capabilities you control to execute AI strategy.

1. Data Infrastructure
2. AI Governance
3. Technology Stack
4. Talent & Workforce (25%)
5. AI Governance (20%)
6. Use Case Portfolio (15%)
7. Organizational Culture (15%)

Unbalanced Profiles are Penalized:
High scores in one area (e.g., Talent) cannot fully compensate for weaknesses elsewhere (e.g., Data).

Critical Readiness Bottlenecks: Only 4% report data is AI-ready. 75% have responsible AI programs, but only 14% have operationalized them.



APPLICATION SPOTLIGHT: FINANCIAL SERVICES

Why Financial Services Requires Special Calibration

High regulatory intensity, documented AI links to financial metrics, and mature use cases.

Navigating a Complex Regulatory Environment

Measures compliance with SR 11-7, EU AI Act, and Fair Lending laws. Non-compliance with EU AI Act can result in fines up to **€35M** or 7% of global turnover.



Documented Performance Impact
15-25% reduction in loan losses;
2.1-4.2 percentage point improvement in insurance loss ratios.



Tailored Capability Weighting
AI Governance & Risk Management weight increased from 20% to 20% for this sector.

THE Org-AI-R SCORE IN PRIVATE EQUITY DUE DILIGENCE



The Ultimate Goal: EBITDA Improvement
PE firms use the framework to build value creation playbooks targeting a 5-25% EBITDA improvement from AI initiatives.

Real-World Valuation Impact
A distribution company with AI demand forecasting saw its valuation multiple expand from -7x to -9x EBITDA.



PE AI Readiness Equation

Readiness prices time-to-value and durability.

$$\text{PE-AIR} = \alpha \cdot H^{PE} + (1 - \alpha) \cdot V^{Org} + \beta \cdot \text{Synergy}$$

H^{PE} : Opportunity

EBITDA-relevant drivers

V^{Org} : Execution

Data, work-flow, leadership

Synergy: Upside

Scale wins across platform

What PE Actually Measures

If it can't be repeated, it can't be valued.

Data Usability

Availability, quality, lineage

Workflow Integration

AI embedded in decision loops

Leadership

Ownership, incentives, cadence

Scalability

Replicate wins across sites

The AI Value Lever: A Private Equity Framework for Quantifying AI Readiness

A systematic methodology to assess, improve, and monetize AI capabilities across the investment lifecycle.

THE PROBLEM: AN AI MEASUREMENT GAP IN PRIVATE EQUITY



AI-Enabled Companies Command Major Premiums
Companies with embedded AI capabilities achieve

5-25% EBITDA

improvement compared to their non-AI peers.



A Widespread Lack of Measurement

Despite a surge in AI adoption, **36%** of PE firms with AI strategies have no specific milestones or RPIs for measuring its impact on value creation.



High Failure Rate for AI Pilots

A staggering **95%** of generative AI pilots fail to deliver, often due to a lack of rigorous assessment and planning frameworks.

THE SOLUTION: THE PE ORG-AI-R FRAMEWORK

Quantifying AI Readiness
A parametric score that assesses a company's ability to capture AI opportunity by combining industry potential with its specific internal capabilities.



SYSTEMATIC OPPORTUNITY
(Industry Potential)

IDIOSYNCRATIC READINESS
(Organizational Capabilities)

Tailored for the PE Model

The framework is specifically designed for the PE lifecycle, focusing on EBITDA Improvement, exit-readiness, and value creation within a 3-7 year held period.

THE FRAMEWORK IN ACTION: THE 5 STAGES OF THE PE LIFECYCLE



STAGE 1: SCREENING
(1-2 Days)

Rapidly assess deal opportunities using external signals like AI-related job postings, patent filings, technology stack indicators, and leadership experience.



STAGE 2: DUE DILIGENCE
(2-4 Weeks)

Conduct a comprehensive assessment through management interviews, technical reviews of data infrastructure, and an organizational talent inventory.



STAGE 3: 100-DAY PLAN
(Post-Close)

Develop a value creation roadmap by identifying quick-wm AI initiatives, addressing data gaps, and setting a 3-year AI strategy with clear milestones.



STAGE 4: MONITORING
(Quarterly)

Track progress against KPIs such as Org-AI-R score improvement, incosiment efficiency (score points per \$M invested), and documented AI-driven EBITDA impact.



STAGE 5: EXIT PREP
(2-6 Months Pre-Exit)

Build a compelling, evidence-based AI narrative for buyers, focusing on documented RDI, sustainable capabilities, and future potential to maximize valuation.

SECTOR-SPECIFIC AI PLAYBOOKS



MANUFACTURING: FOCUS ON OPERATIONS

High-value use cases include predictive maintenance (reducing downtime 15-25%) and computer vision for quality control (improving yield 5-10%).



HEALTHCARE: FOCUS ON EFFICIENCY & COMPLIANCE

Key opportunities are in AI-driven claims processing (reducing denials 15-25%) and patient flow optimization. AI governance is critically important.



RETAIL: FOCUS ON THE CUSTOMER

AI drives value through personalization engines (increasing conversion 10-30%), demand forecasting, and dynamic pricing (improving margins 2-5%).



BUSINESS SERVICES: FOCUS ON PRODUCTIVITY

Generative AI tools can improve knowledge worker productivity by 20-30% and reduce manual effort in document processing by 50-70%.



TECHNOLOGY: FOCUS ON PRODUCT & MOAT

Embedding AI into software can loat to a 40-102% solution up/dt. Assesment focuses on competitive defensibility and laint depth.

THE FINANCIAL IMPACT: LINKING AI READINESS TO RETURNS

Three Channels of Value Creation



Direct Link to EBITDA Improvement

A 10-point improvement in the Org-AI-R score is modeled to visid a 1.6-4.0% EBITDA improvement in a high-opportunity industry.

Quantifiable Multiple Expansion

A strong, documented AI exit narrative can contribute an estimated 1.0 to 3.0+ turns of additional EBITDA multiple upon exit.

EXAMPLE: CASE STUDY: INDUSTRIAL MANUFACTURER

A 52.7M investment in AI initiatives led to an 8% cumulative EBITDA improvement and increased the exit multiple from 8.5x to 5.0x.

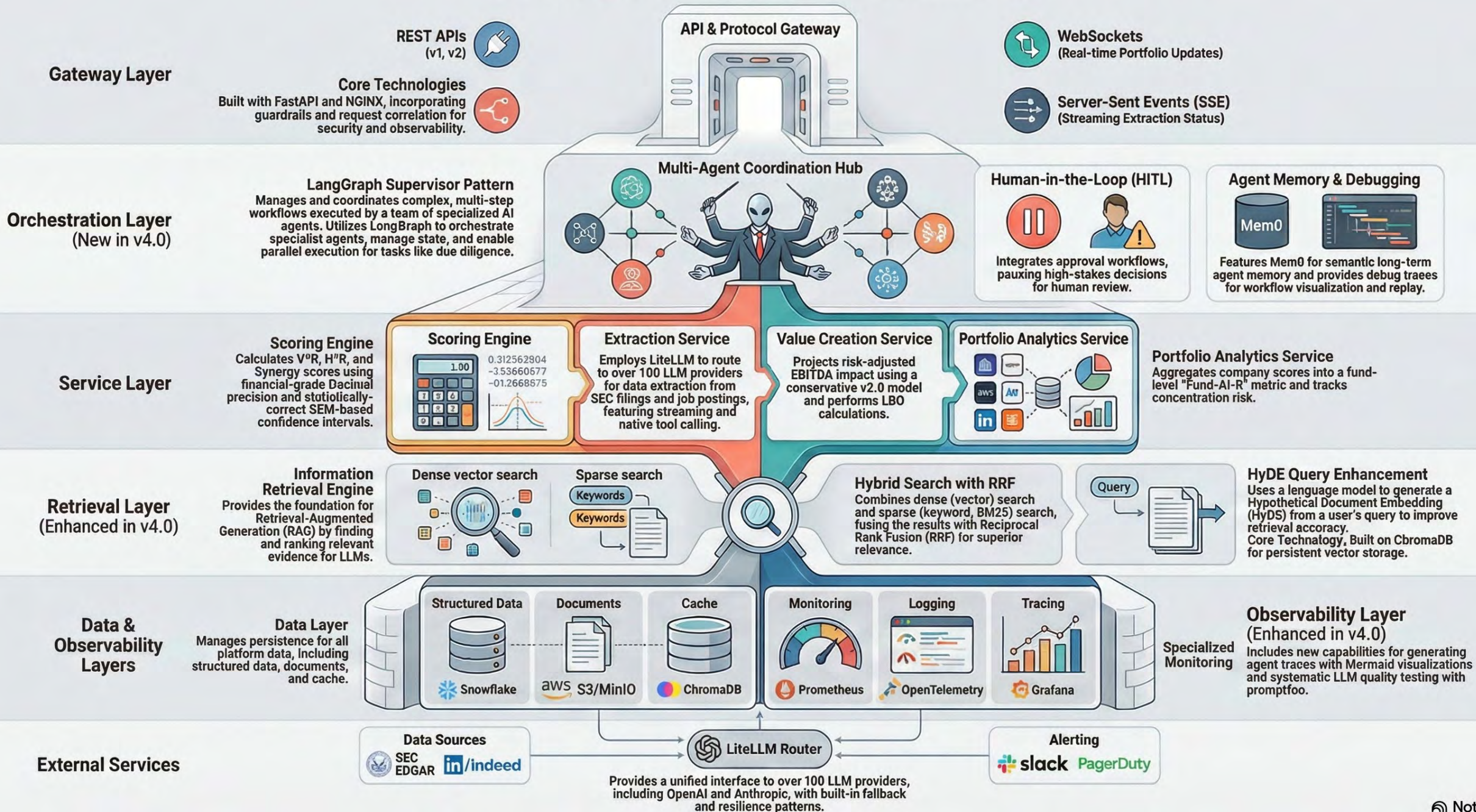
Close

The AI Quant of the Future

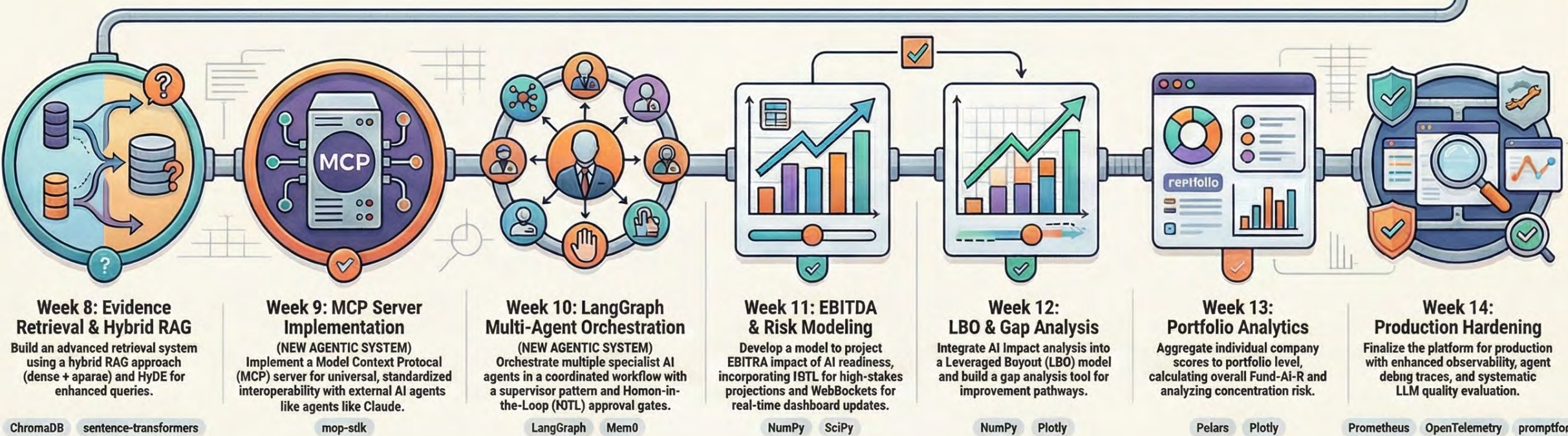
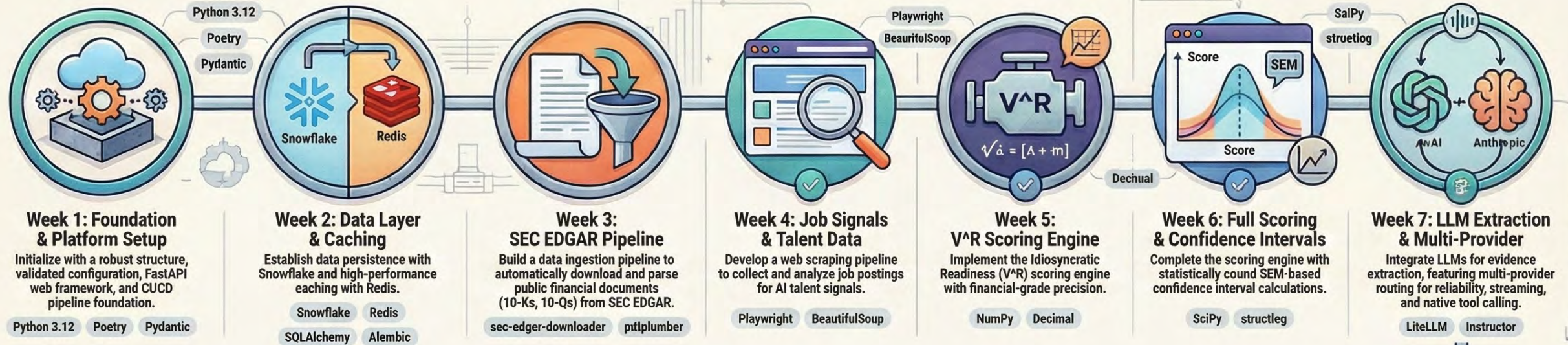
The future quant is a workflow architect with risk discipline.

- ▶ Less model tinkering, more **system orchestration**
- ▶ **Judgment** > Prediction
Especially under regime shifts
- ▶ **Accountability** > Accuracy
For regulated decisions
- ▶ **Artifacts + Controls = Trust**

Anatomy of the PE Org-AI-R Platform v4.0



The 14-Week Journey to Building a Portfolio Intelligence System



Thank You

Sri Krishnamurthy, CFA, CAP

sri@quantuniversity.com

www.quantuniversity.com

<https://tinyurl.com/PEAIR2026>



“The best time to design governance is before you need it.”