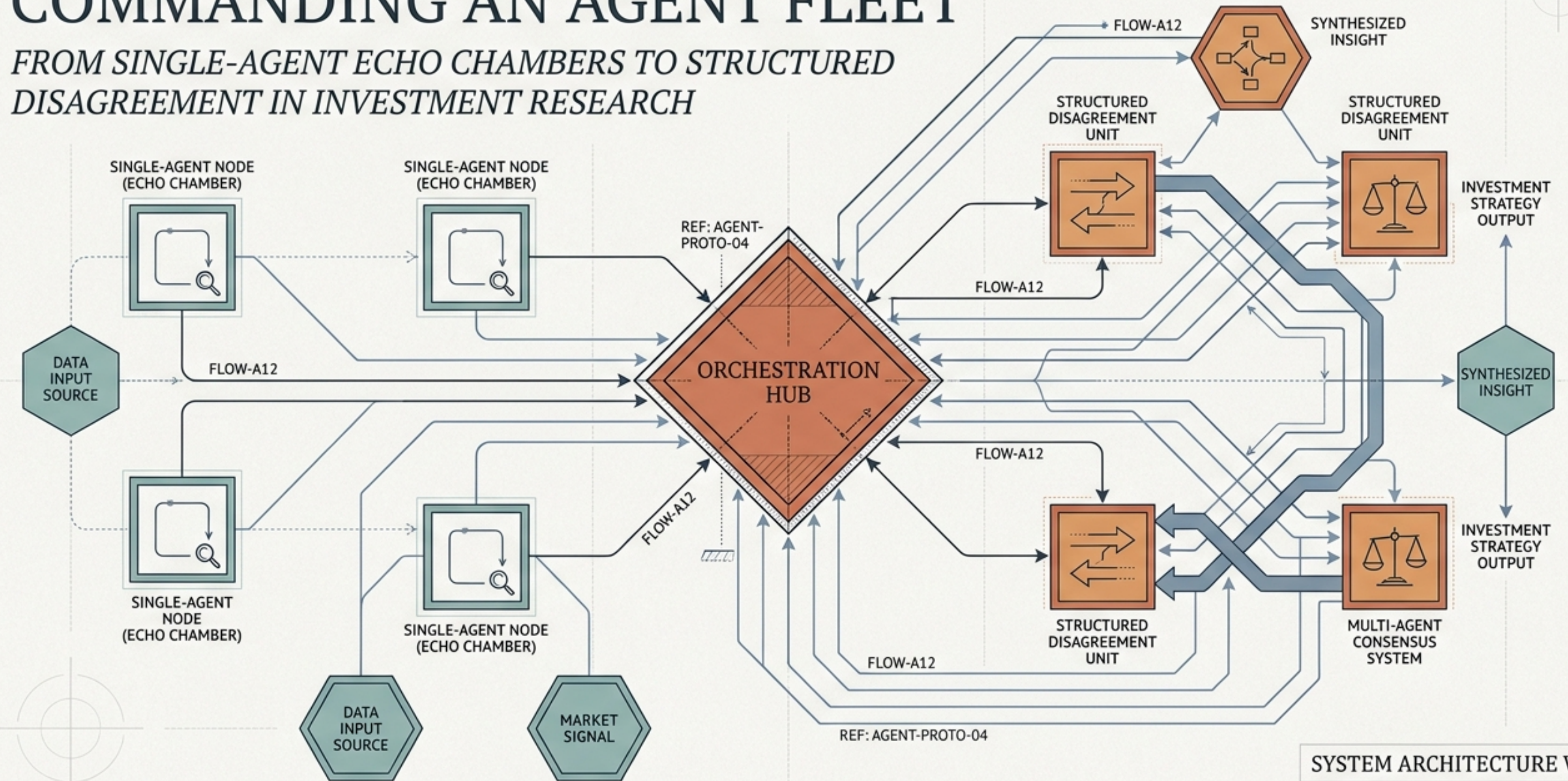


# COMMANDING AN AGENT FLEET

FROM SINGLE-AGENT ECHO CHAMBERS TO STRUCTURED DISAGREEMENT IN INVESTMENT RESEARCH



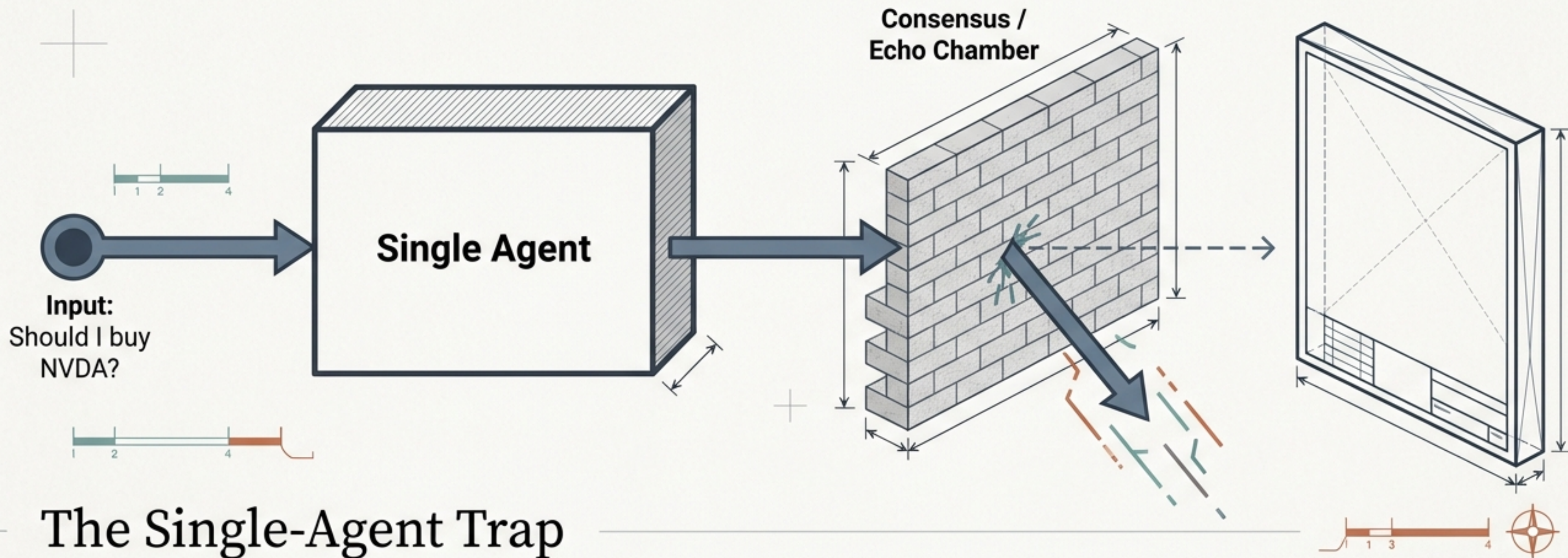
SYSTEM ARCHITECTURE V.2.1

SCALE: 1:100

TYPE: 1:10

DATE: OCT 26, 2024

STATUS: DEPLOYED



## The Single-Agent Trap

### The Status Quo

Most people use agents like they manage their first team: handing the whole problem to one entity and hoping for a good answer.

### The Reality

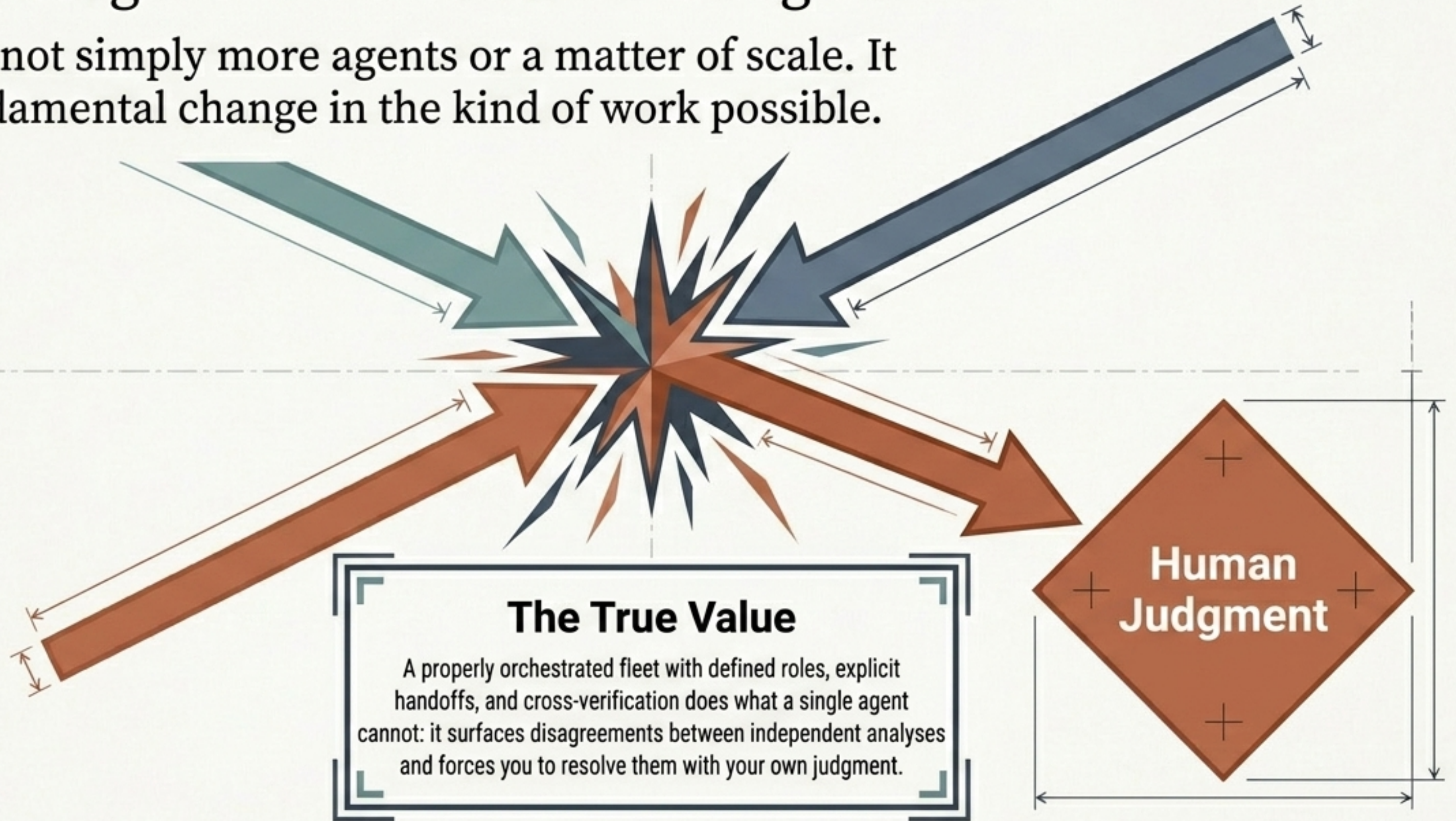
The output reads like original investment research, but it is actually a sophisticated summary of what everyone already thinks.

### The Flaw

A single agent defaults to the center of its training distribution. It cannot adopt two genuinely opposing viewpoints. It restates the median sell-side analyst opinion in clearer prose.

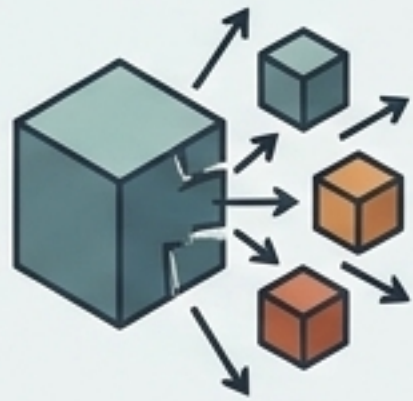
# The Paradigm Shift: Structured Disagreement

A fleet is not simply more agents or a matter of scale. It is a fundamental change in the kind of work possible.



# The 5 Core Orchestration Patterns

These patterns cover 90% of investment research workflows.



## 1. Delegation

Break the job before you hand it out.



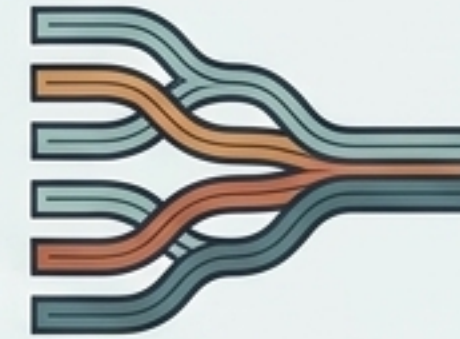
## 2. Verification

Independent checks, not proofreading.



## 3. Iteration

Loops that converge, not loops that spin.



## 4. Composition

Assembling partial outputs into a coherent whole.



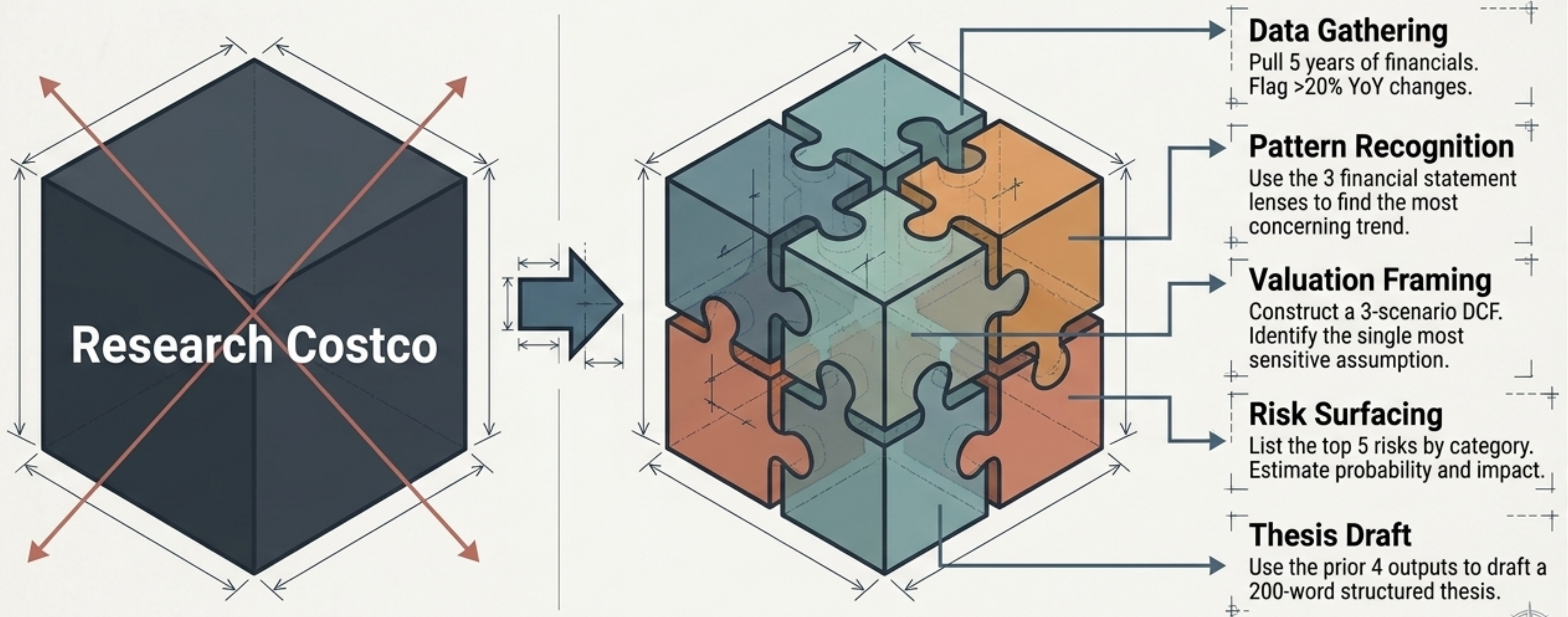
## 5. Failure Detection

Knowing exactly when the fleet is confidently wrong.



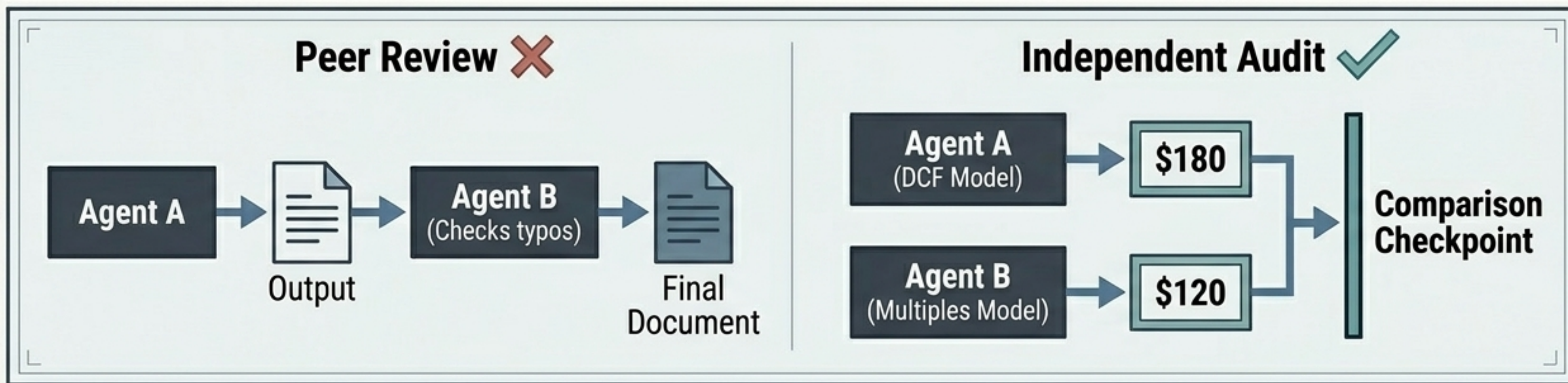
# Pattern 1: Delegation (Decomposition)

*Effective delegation requires breaking the work into pieces specific enough to verify independently.*



# Pattern 2: Verification (Independent Checks)

*Asking a second agent to check the first agent's work only creates an expensive echo chamber. The second agent **must never see** the first agent's output.*



## Parallel Valuation

Compare independent DCF models vs. market comparable models.



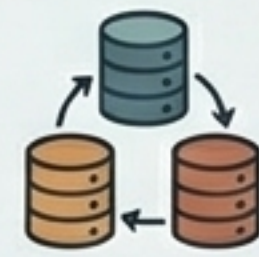
## Adversarial Fact-Check

Give Agent B only the claims, ask it to find contradicting evidence.



## Source Triangulation

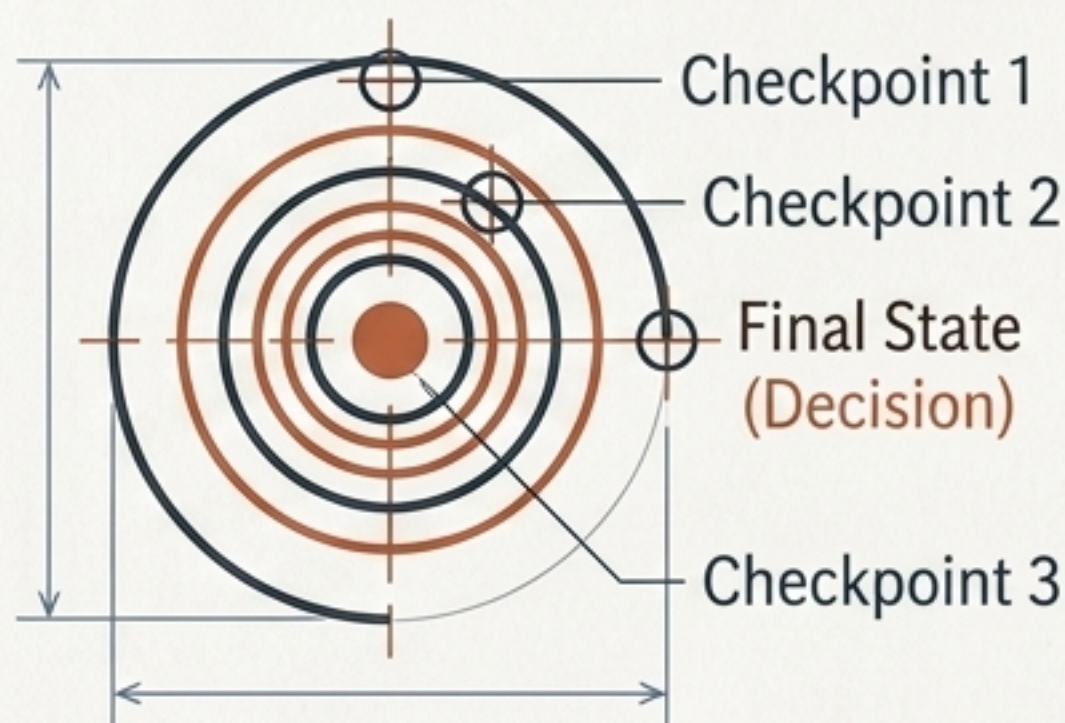
Verify foundational revenue figures across multiple data sources.



# Pattern 3: Iteration (Converging Loops)

*Iteration without boundaries is procrastination disguised as productivity.*

*Productive loops require three explicit constraints:*



## 1. Specific Target

For example: "The kill criteria are price-based. Rewrite them as falsifiable business conditions."

## 2. Clear Convergence Criteria

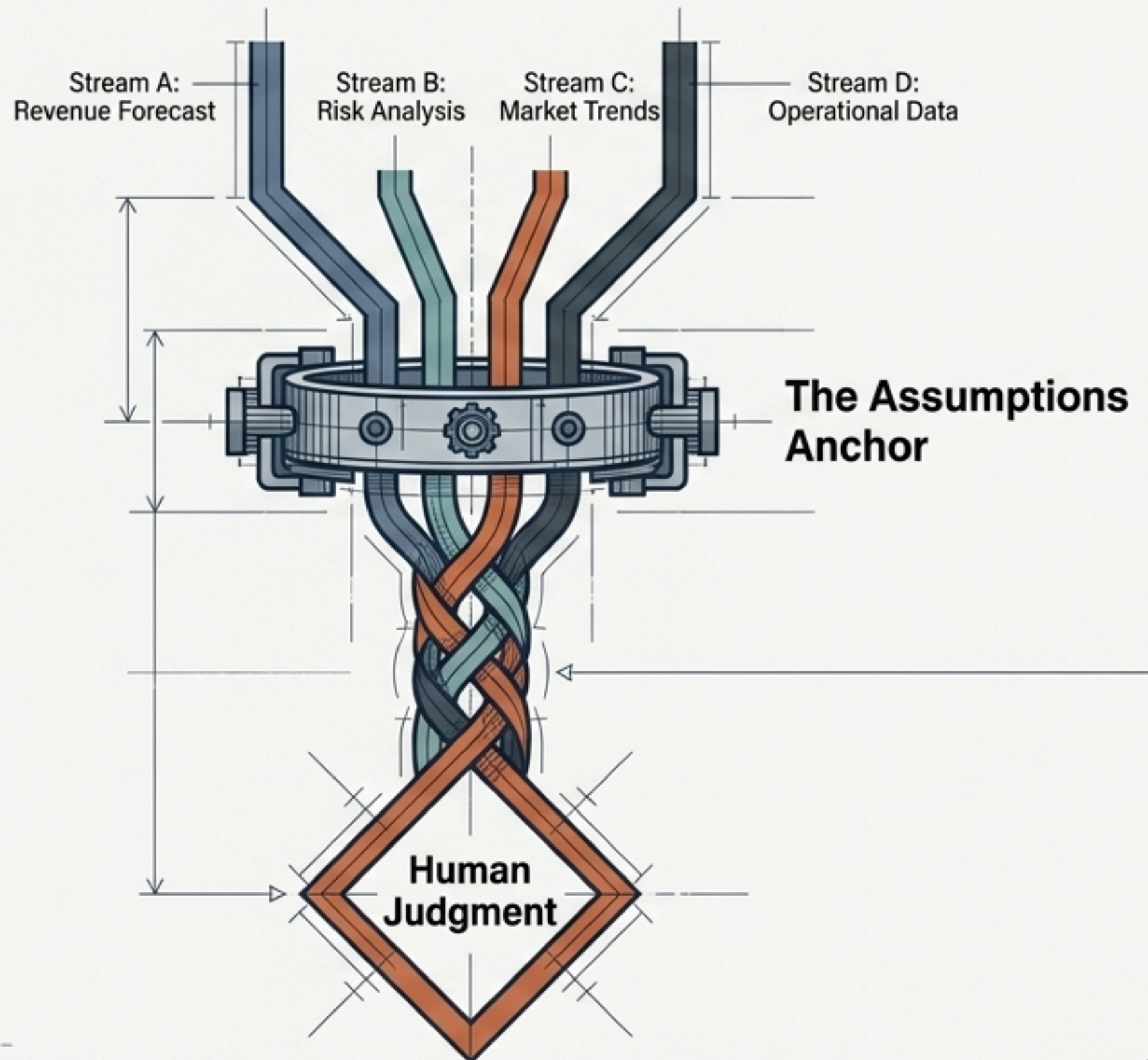
Stop when the "red team" agent's attacks become generic and no longer reveal new weaknesses.

## 3. Maximum Cycles

Restrict to 3 to 5 rounds for a thesis document. If you reach round 8, you are avoiding making a decision.

# Pattern 4: Composition (Assembling the Whole)

*Combining specialized outputs naively creates a Frankenstein memo: bullish on revenue, pessimistic on risk, with contradictory assumptions.*



## The Assumptions Anchor

State key assumptions (growth rate, macro conditions) before agents start. Every output must adhere to them or explicitly justify a deviation.



## The Integration Step

Do not smooth over contradictions. If a revenue forecast conflicts with a risk analysis, that tension is the most valuable part of the process.



## The Narrative Thread

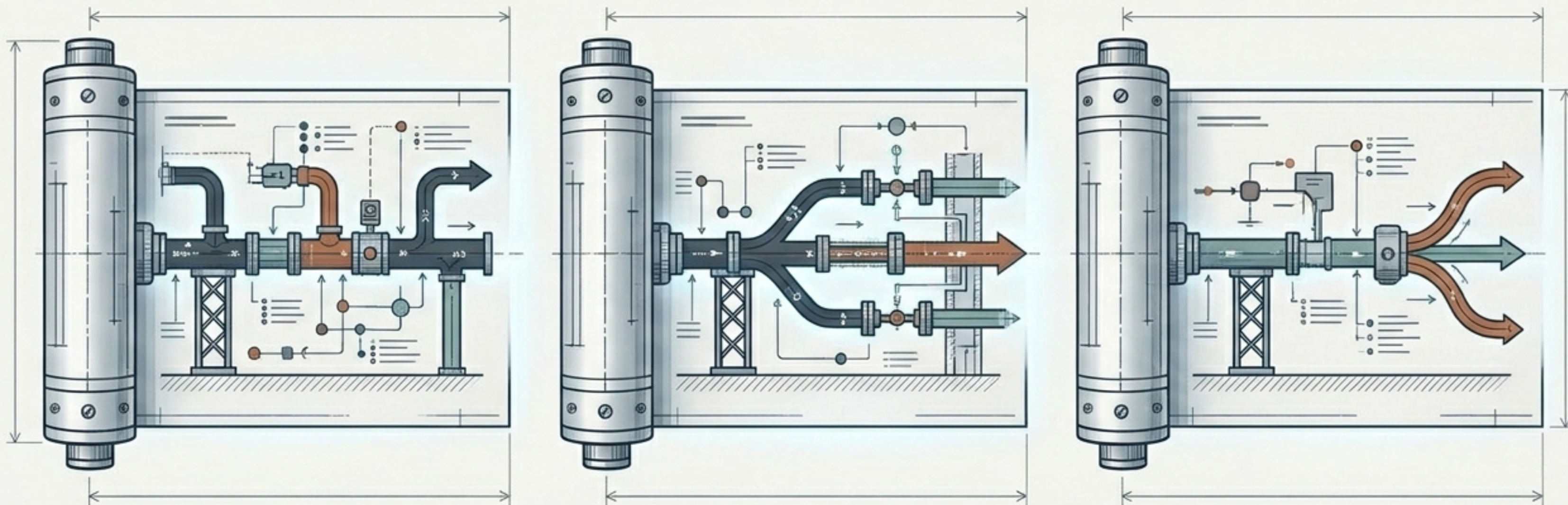
The final story connecting evidence to conclusion must be written by you.





# The Architectural Blueprints

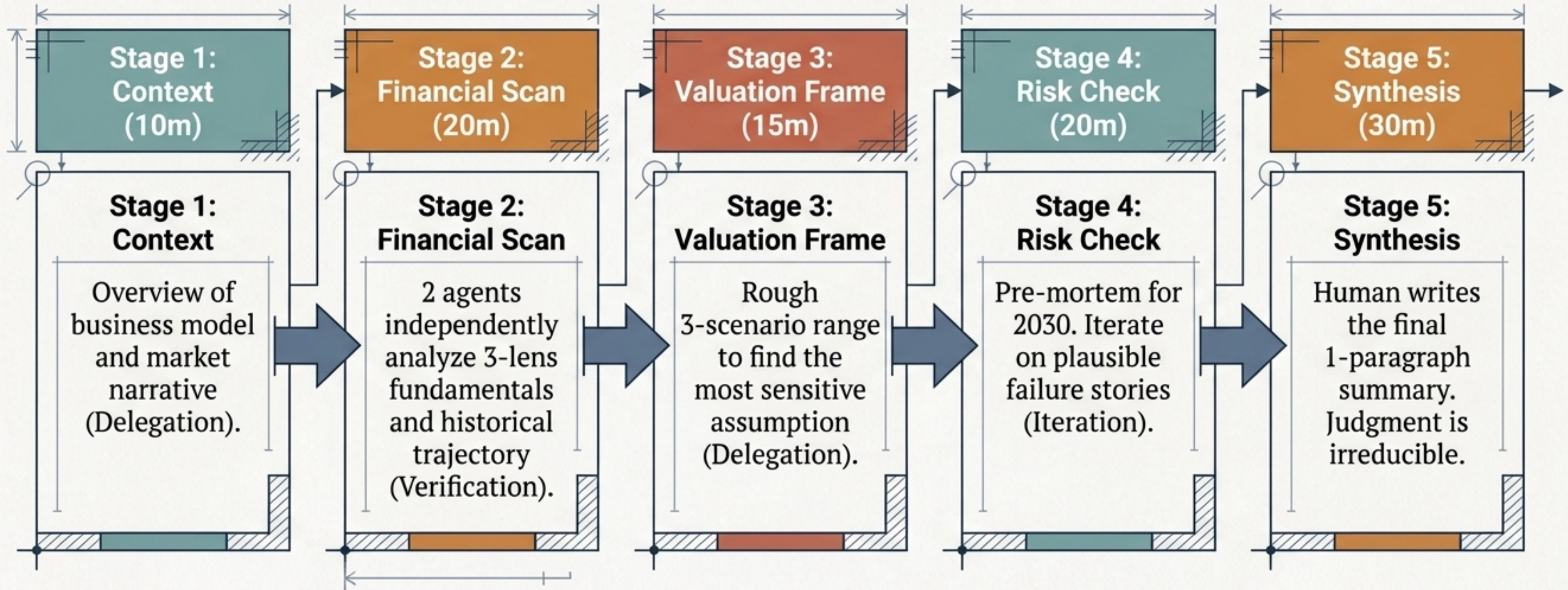
*How the 5 patterns combine in the real world.*



Three reference implementations covering the primary investment research tasks, engineered for individual investors.

# Workflow A: The Research Pipeline

Goal: From zero to an informed opinion in 2 hours.



# Workflow B: Thesis Refinement

Goal: Stress-test an existing thesis in 90 minutes.

## Stage 1: Adversarial Review

Agent acts as **short-seller**. Produces the 3 strongest, evidence-based attacks on key assumptions.

## Stage 2: Your Response

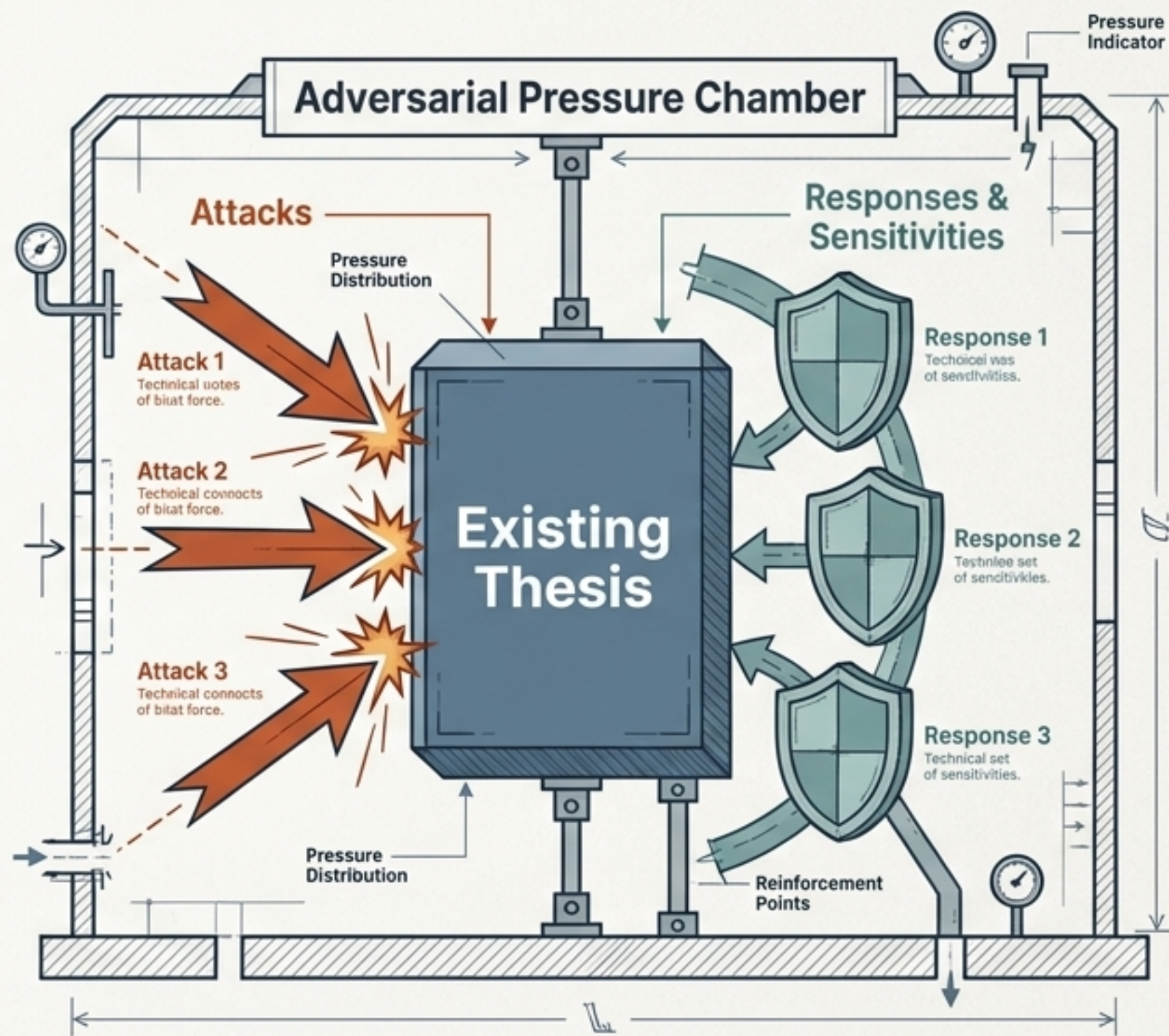
Human writes specific refutations or adjusts position sizing.

## Stage 3: Assumption Sensitivity

Agent models valuation impact (e.g., margins compress 200bps).

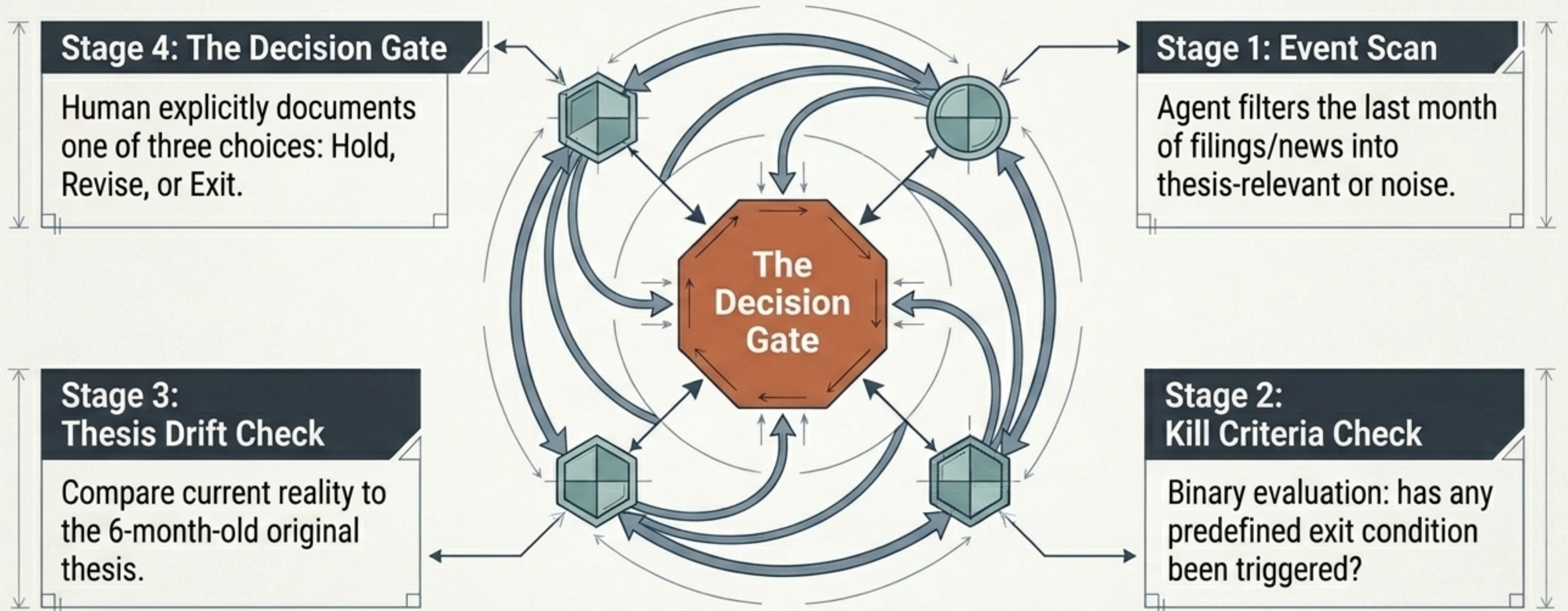
## Stage 4: Kill Criteria Check

Map every kill criterion to a specific, observable data source.



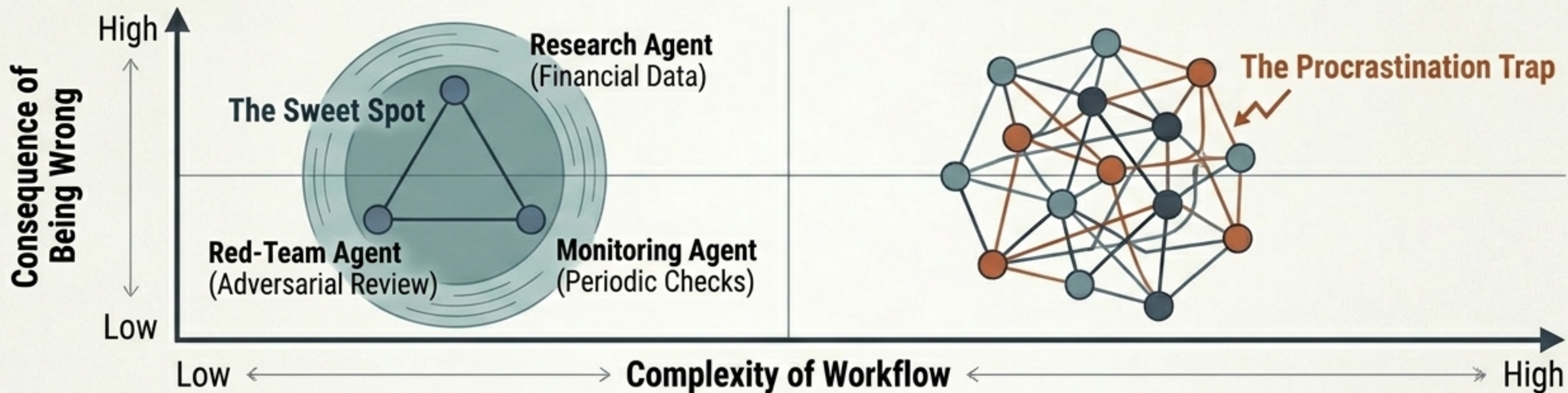
# Workflow C: The Monitoring Loop

**Goal:** Minimum viable tracking process (30-45 mins/month per position).



# The Fleet Complexity Trap

Building a 12-agent pipeline with branching logic is procrastination wearing productivity's clothes.



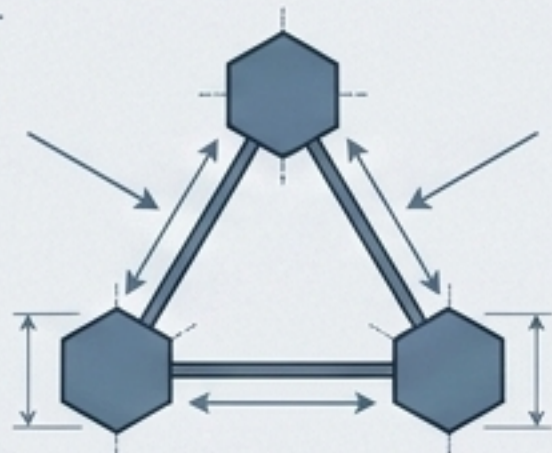
## The Decision Rule

Do not scale the fleet based on the complexity of the question. Scale it based on the consequence of being wrong.

## The Sweet Spot (5-15 positions): Just 3 roles.

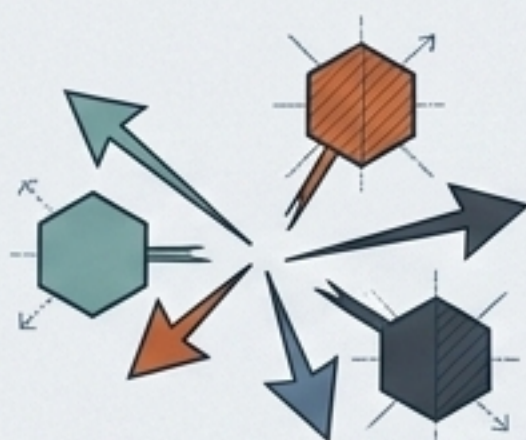
- 1 Research Agent (Financial Data)
- 1 Red-Team Agent (Adversarial Review)
- 1 Monitoring Agent (Periodic Checks)

# The Trust Meta-Question



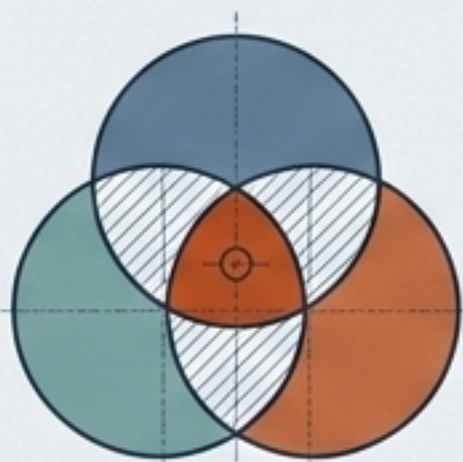
## All Agents Agree (Suspicious)

Unanimous agreement usually means the agents are just reproducing the same consensus view from their shared training data. Test it by asking one to steelman the opposite.



## All Agents Disagree (Ambiguous)

The underlying data is inherently unclear or the company is in transition. “I don’t have enough clarity” is a valid, money-saving output.



## Partial Agreement (Informative & Optimal)

Agent A flags revenue, Agent B flags balance sheet, Agent C flags regulatory risk. They identified different dimensions. Your job is integration.

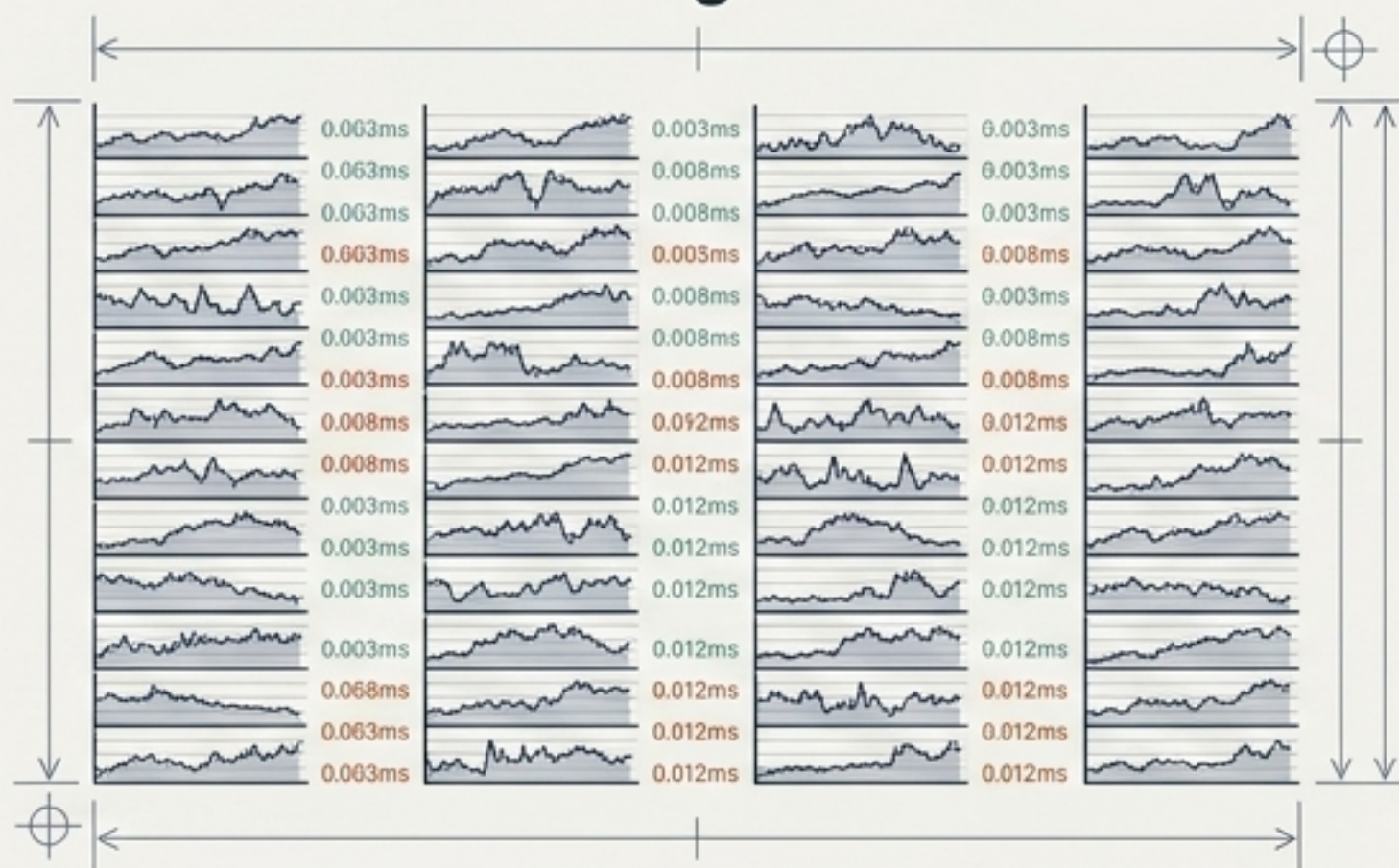
# Diagnostic Matrix: 4 Financial Agent Failure Modes

Failure Mode	Why it is dangerous	How to Detect	How to Mitigate
<b>1. Outdated Data</b>	Uses training data instead of latest 10-Q.	Verify dates.	Pull latest SEC filing yourself to use as an anchor.
<b>2. Consensus Blindness</b>	Accurately states strengths but ignores if the stock price already reflects them.	Ask what would surprise a pro.	Explicitly prompt for pricing/valuation justification.
<b>3. Fabricated Sources</b>	Invents non-existent analyst notes (e.g., fake Goldman report).	Search for the primary source.	Instruct agent to separate verifiable facts from estimates.
<b>4. Unit Errors</b>	Confuses millions with billions, inflating valuations by 1000x.	Cross-check revenue, market cap, and employee count.	Add a final sanity check verification step.

# Reality Check: Bloomberg vs. The Fleet

Agents cannot replace the proprietary data or millisecond speed of a \$24k Bloomberg terminal.

## Bloomberg Terminal

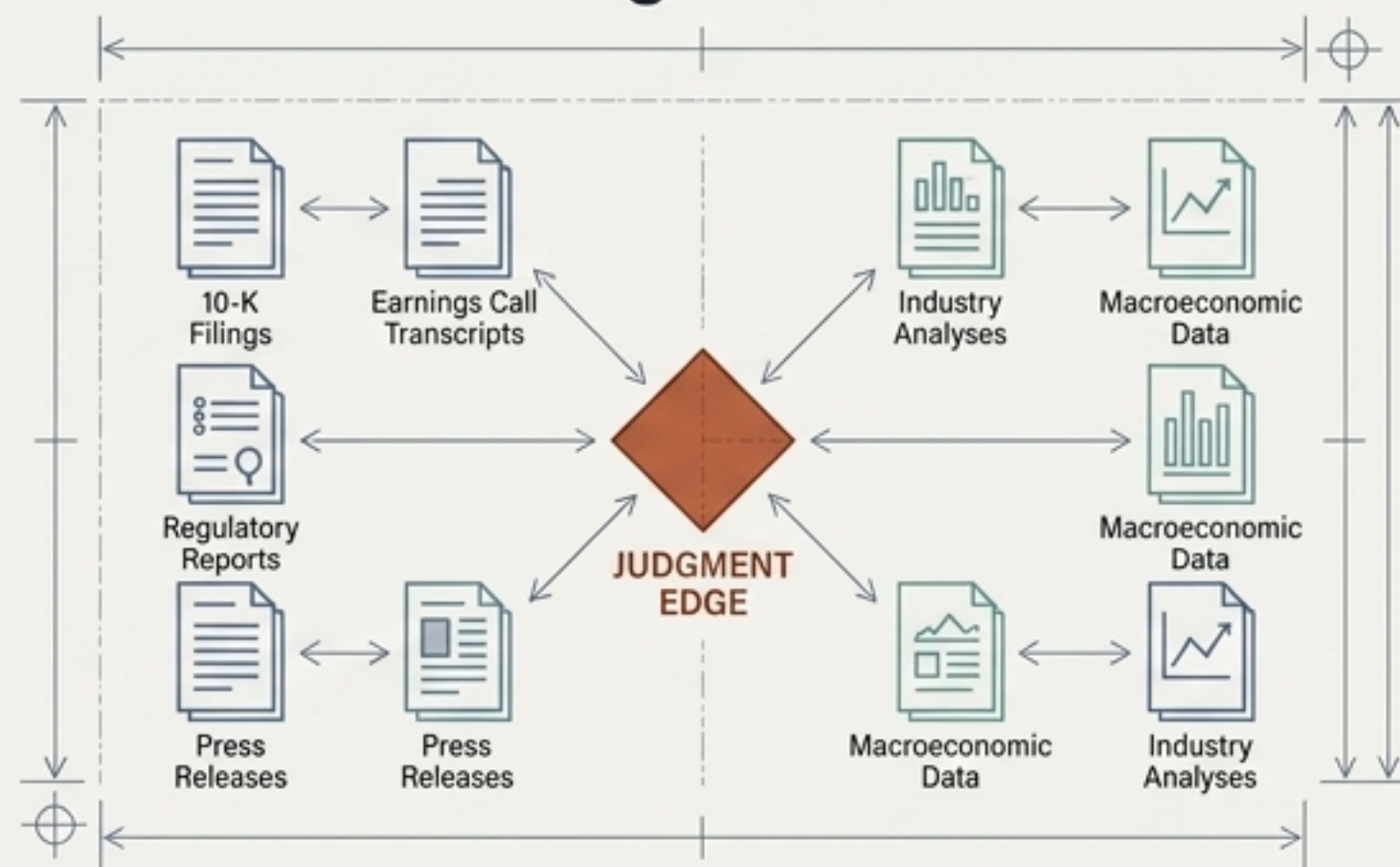


### Different Games

You are not competing on high-frequency order flow.

You are competing on long-term judgment.

## The Agent Fleet

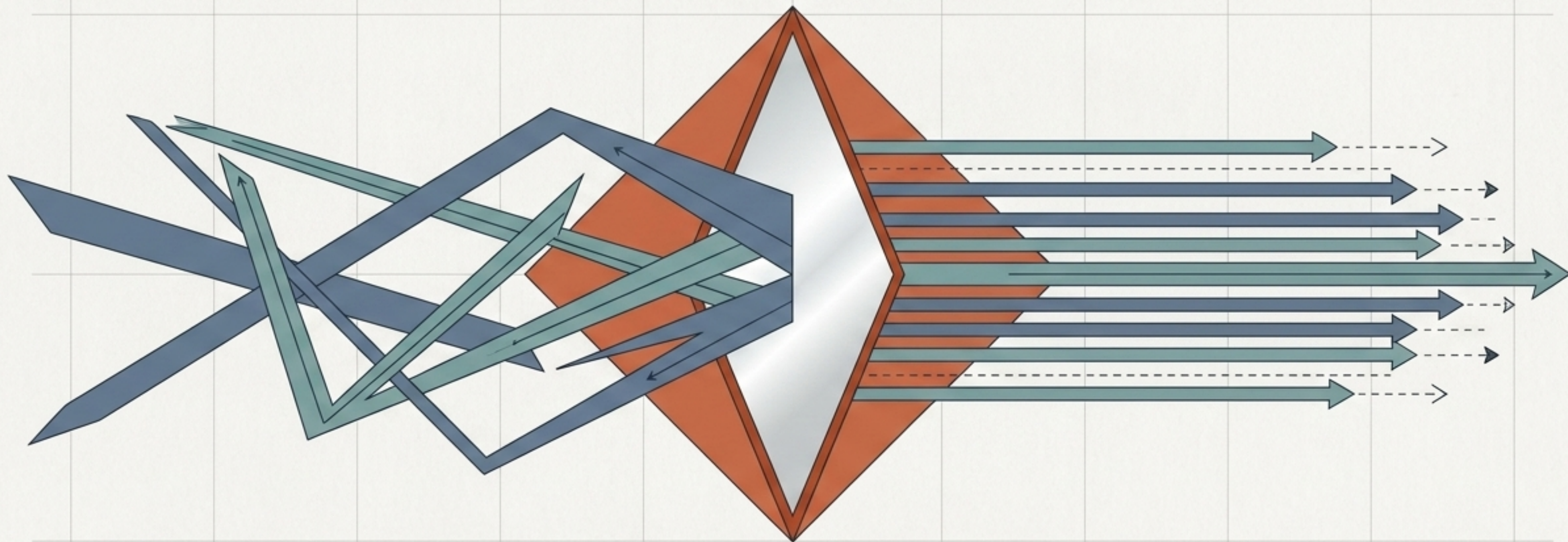


### The Structural Advantage

Public data (10-Ks, earnings calls) + individual patience + lack of institutional career risk.

Agents compress the time required to gather public data; your judgment provides the actual edge.

# What the Fleet is Actually For



The scarcest resource in investing is not data, analysis, or time. It is **clarity** about your own **reasoning**.  
The agents do not make the decision. They create the conditions under which a better decision becomes possible.  
A well-designed fleet is simply a **mirror**. It reflects the quality of your questions back at you.  
Every improvement in orchestration is an improvement in the precision of your own judgment.