

# Project Management Metrics: Measurement and Interpretation



**Course promise:** separate diagnostic signals from noise, then turn metrics into decisions



**Core tension:** metrics compress messy reality into a number — and distort it



**Bias, gaming, misaligned definitions, and false precision** are the distorters



**Roadmap:** metric types, measurement design, EVM and flow interpretation



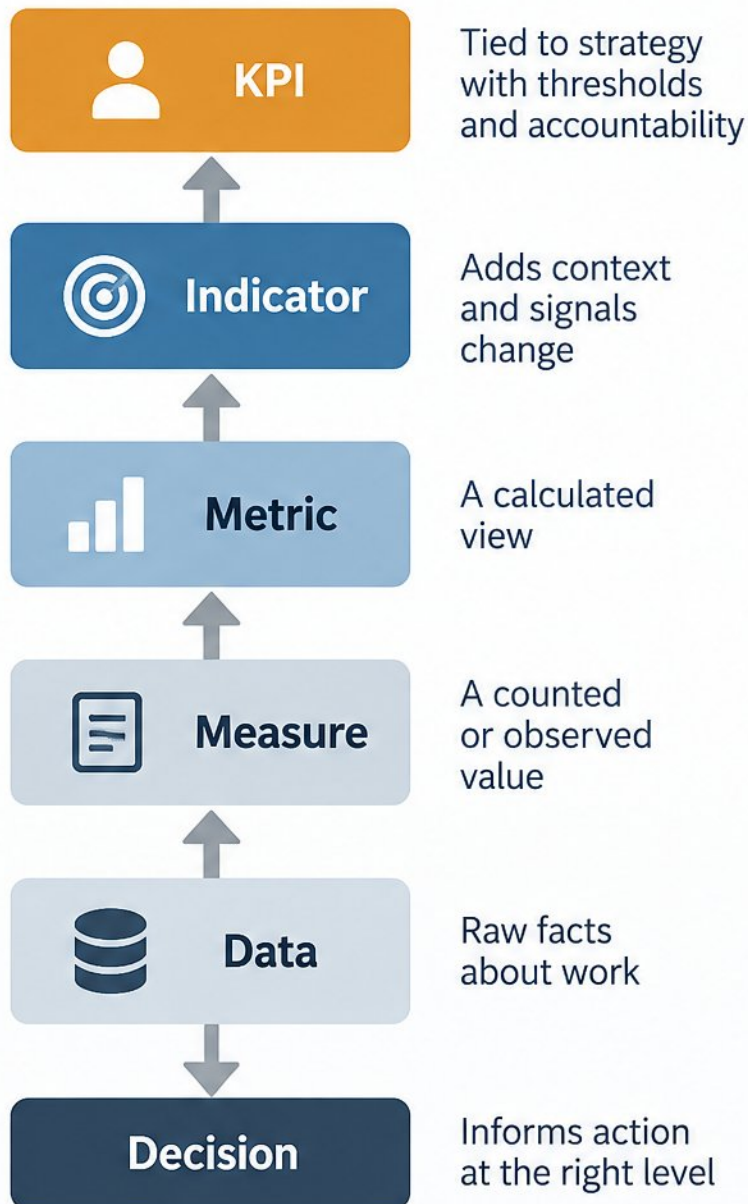
**Then:** probabilistic forecasting, uncertainty communication, storytelling, governance, AI



**Framing rule:** start from the decision, not the dashboard



# From Data to Indicator to Decision



- **Data → measure → metric → indicator → KPI:** each step adds meaning and accountability



- **KPI** = metric tied to strategic objectives, with thresholds and a review cadence



- **Leading indicators** (WIP, defect escape, float erosion) signal what comes next



- **Lagging indicators** (CPI, benefits, satisfaction) confirm what already happened



- **State the level:** process, project, portfolio, or business outcome — never mix silently

# Measurement Design: Validity, Reliability, and Feasibility



**Validity:** does the metric measure what we intend, or just a convenient proxy?



**Reliability:** consistent across time, teams, and reporters; subjective percent complete destroys comparability.



**Feasibility:** cost of collection versus decision value; if it changes no decision, value is zero.



**Operational definitions:** numerator, denominator, time window, scope, and clock start/stop — written, versioned, visible.



**Avoid vanity metrics and composite indices;** use balanced sets so gaming one number distorts another.

## Measurement Definition



Numerator



Denominator



Time window



Scope



Clock start/stop

# Earned Value: What SPI, CPI, and TCPI Hide



**SPI = EV/PV:** cost-weighted efficiency, not time



Blends critical and non-critical work; forced to 1.00 at completion



**CPI = EV/AC:** distorted by late invoices, early earned material



**TCPI > CPI + 0.10** signals promised performance never demonstrated



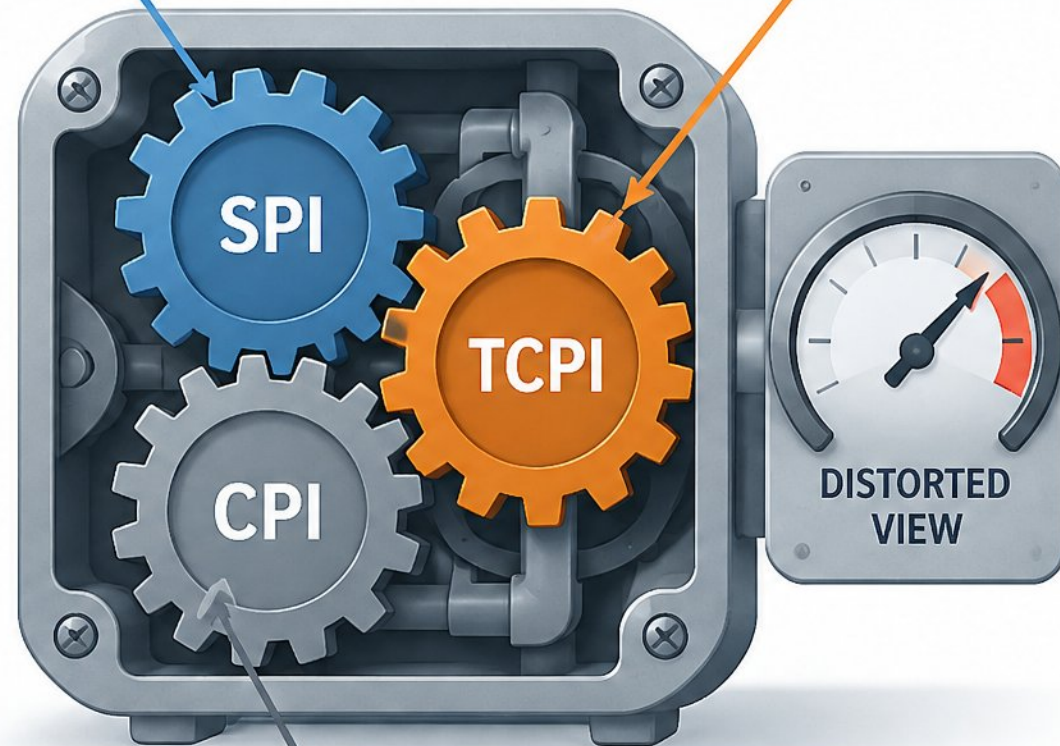
Cumulative CPI rarely improves after ~20% complete — forecast from reality



Read EVM with critical path float trend and earned schedule SPI(t)

Blends critical and non-critical work

Subjective progress distorts output



Late invoices, early earned material

# Forecasting Under Uncertainty: Ranges, Confidence, and Reference Classes



Never commit to a P50 date — that is a coin flip; commit at P80 internally, P95 for hard deadlines.



Reference class forecasting (RCF) corrects optimism bias with the outside view of comparable projects.



RCF accuracy depends critically on reference class formation and data stability, not just method choice.

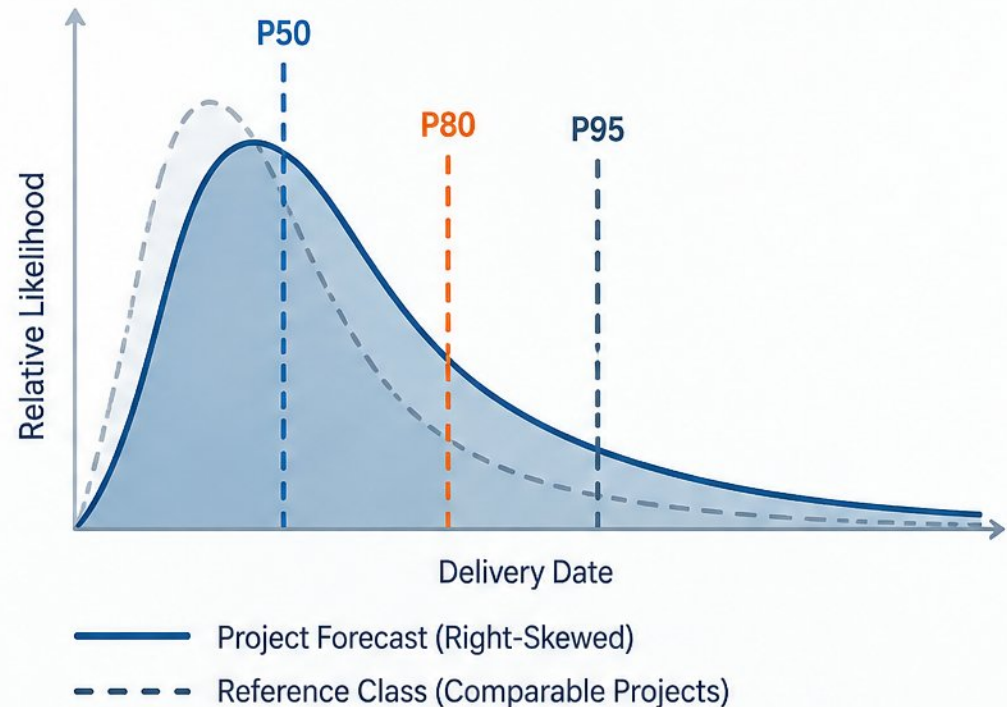


Read the distribution, not just percentiles: bimodal or right-skewed curves signal one or two tail-driving tasks.



Separate aleatory risk (probabilistic methods) from epistemic gaps (smart heuristics) and residual portfolio-level uncertainty.

## Delivery Date Distribution (Illustrative)



Wider tails often point to specific risks — investigate and mitigate the drivers, not just the numbers.

# Communicating Uncertainty Without Losing Sponsor Trust



Confidence intervals move when data moves, not on a quarterly cycle



Report distributions as decisions: ask "what is the P80?"



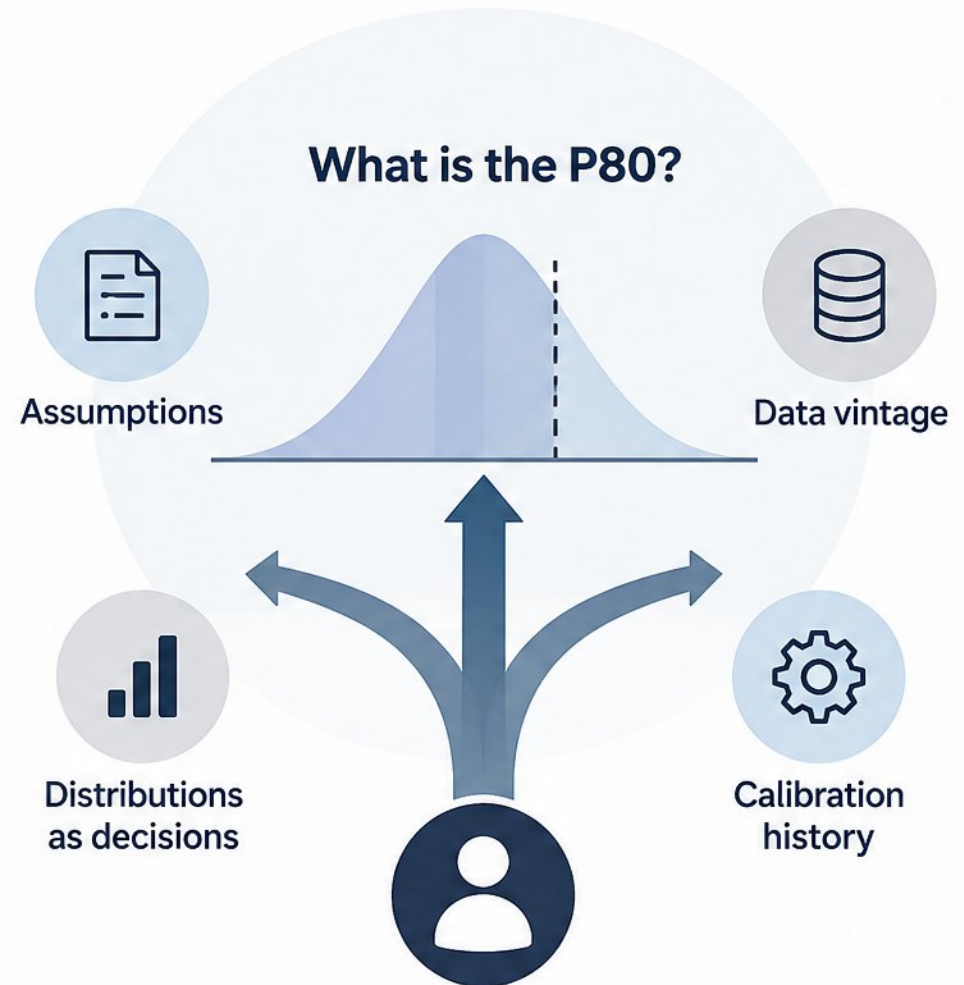
False precision to the day borrows credibility it has not earned



Ship assumptions, data vintage, and calibration history with every forecast



If 90% ranges capture only 50% of outcomes, widen them



# Flow Metrics and Velocity: Interpretation Without Misuse



Core flow metrics: cycle time, throughput, WIP, work item age



Read cycle time as a percentile, never an average — work is right-skewed



Little's Law (cycle time = WIP / throughput) is a mental model

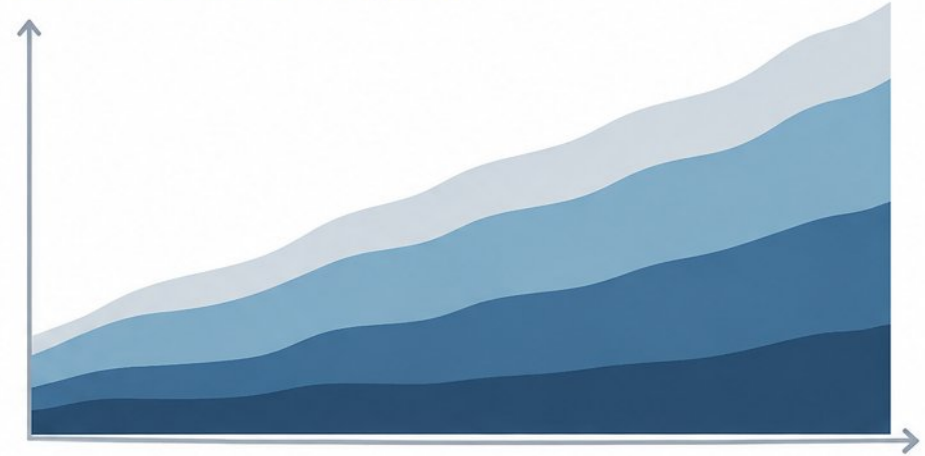


Cumulative flow diagrams: widening bands signal growing queues

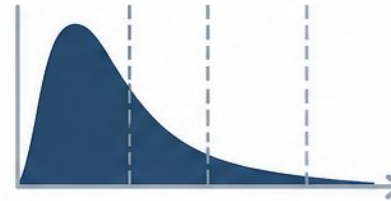


Velocity is a team planning tool, not a performance metric

Cumulative Flow Diagram



Cycle Time (Percentiles)



WIP Limit (Inflow Constraint)



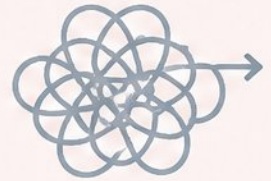
## Signal

Use metrics to understand flow and make better decisions.

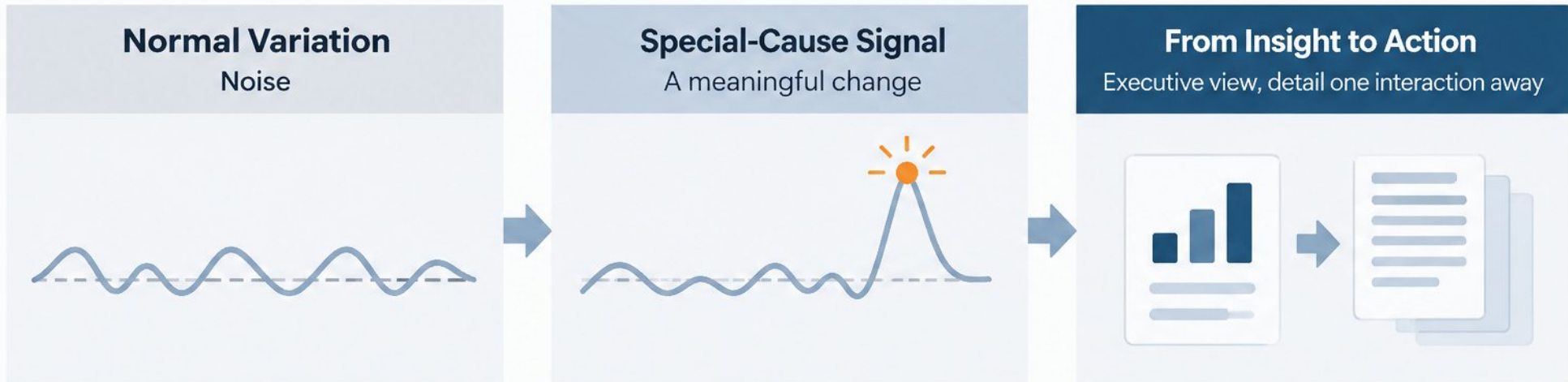


## Noise

Avoid misusing metrics or treating velocity as a performance metric.



# Analysis and Storytelling: Signal, Noise, and Narrative



SPC separates normal variation from special-cause signals; one reporting cycle is a data point, not a trend.



Variance analysis: distinguish timing variances (work shifted) from performance variances (work genuinely costing more).



Segment by workstream, vendor, or team to see if a change is broad-based or isolated.



Executive dashboards favor a few dominant signals; detail belongs one interaction away.



Sponsor KPIs answer three questions: where are we, where are we heading, what do we need from you?

# Governance, Incentives, and Metric Integrity



**Goodhart's Law:** when a measure becomes a target, it stops being a good measure



**Gaming** is a rational response to incentive design — not a character flaw



**Failure patterns:** relaxed progress rules, re-baselining, renegotiated thresholds, anchored EACs



**Countermeasures:** balanced metric sets, published thresholds, written rules, independent audits



**One named owner per metric,** versioned registry, lineage, deprecation rules

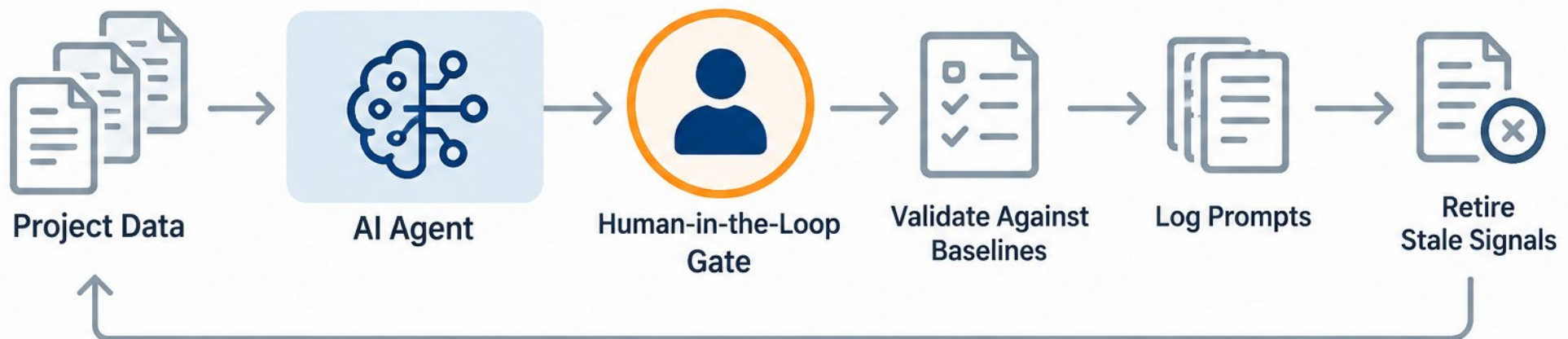


**Governance forums** pre-specify reviewer, cadence, threshold, and defined response



# AI-Augmented Metrics: Opportunities and Guardrails

- AI adds anomaly detection, forecast adjustment, risk sensing, and narrative reporting
- Gains depend on data hygiene and real-time visibility into project KPIs
- Failure modes: false precision, fragmented data, and automation bias eroding judgment
- PMI AI Standard (June 2026): eight principles, five domains, human-in-the-loop
- EU AI Act high-risk duties from August 2026; UK DUAA 2025 requires meaningful human involvement
- Keep the gate outside the agent; log prompts, validate against baselines, retire stale signals



# Case Clinic: Interpreting a Real Dashboard Under Pressure



**Scenario:** SPI near 1.00 and CPI 0.98, yet defects, rework, and float erosion worsen



Check the decision each metric supports before trusting the number



Distortions: low-invoicing period props up CPI; non-critical work dilutes SPI



Progress percent complete hides quality-driven forecast risk



Rewrite the status view around true delivery drivers, not convenient metrics



# Metric Interpretation Playbook: A Five-Step Discipline



## Step 1

Clarify the decision;  
no decision, zero  
information value



## Step 2

Test validity,  
reliability, and  
Goodhart gaming  
risk



## Step 3

Compare baseline,  
three-period trend,  
and reference class



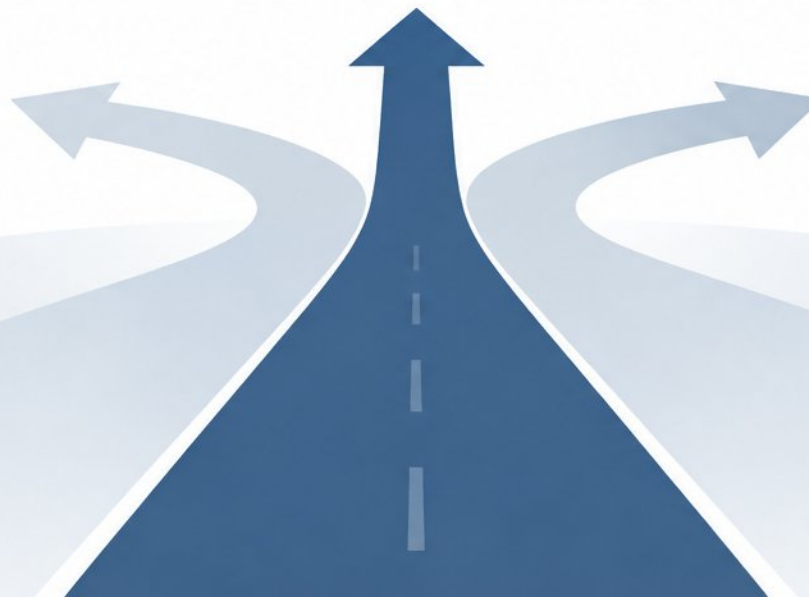
## Step 4

State an uncertainty  
range with explicit  
confidence level



## Step 5

Record assumptions,  
owners, thresholds,  
and review cadence



# Putting It Into Practice: Workshop and Takeaways



Run the five-step playbook end to end on a sample dashboard



Which metric is misleading, why, and what would falsify it?



Debrief pitfalls: one cycle  $\neq$  trend, familiar  $\neq$  useful, informal changes



Measure for understanding, not evaluation; balance linked metrics



Thresholds and definitions non-negotiable; humans own interpretation

# PersonWise.

PersonWise • AI digital-human course platform



Scan the code or click anywhere to watch this course live, with voice Q&A

[personwise.ai/t/8ad58948/public](https://personwise.ai/t/8ad58948/public)