A Thinking Instrument for Attribution

A first-principles system for defensible, interpretable, and mathematically rigorous decision-making.

[CORE MANIFESTO]

This platform is built on a single vision: to move beyond heuristic guesswork and black-box opacity. It is designed for those who require attribution to be derivable from declared assumptions, with a complete audit trail from raw events to final conclusions.

Probabilistic Path Modeling (Markov) | Axiomatic Game Theory (Shapley) | Comprehensive Uncertainty Quantification (UQ)

The Crisis of Measurement: The Epistemic Gap

Heuristic Guesswork

- Relies on arbitrary rules (last-touch, linear).
- Ignores sequence and interaction effects.
- Cannot answer counterfactuals ('what if?').
- Violates fairness axioms.
- Provides no measure of confidence.

First-Principles Attribution

- Derives credit from probabilistic models.
- Explicitly models path structure.
- Built on counterfactual reasoning (removal effects).
- Guarantees fairness through mathematical axioms.
- Quantifies uncertainty as a core tenet.

The Gap Defined: Modern measurement creates a disconnect between correlation (what we see) and causation (what drives revenue). This isn't just an error; it is a structural risk to capital allocation.

The Three Pillars of a Defensible Model

Absorbing Markov Chains

Probabilistic Causality

Models customer journeys as a stochastic process. Captures the critical role of sequence and path dependency, allowing for true counterfactual analysis by measuring ‘removal effects’.

Shapley Value Theory

Axiomatic Fairness

Treats attribution as a cooperative game. It provides the only unique, axiomatically fair method to distribute credit based on a channel’s marginal contribution to every possible coalition.

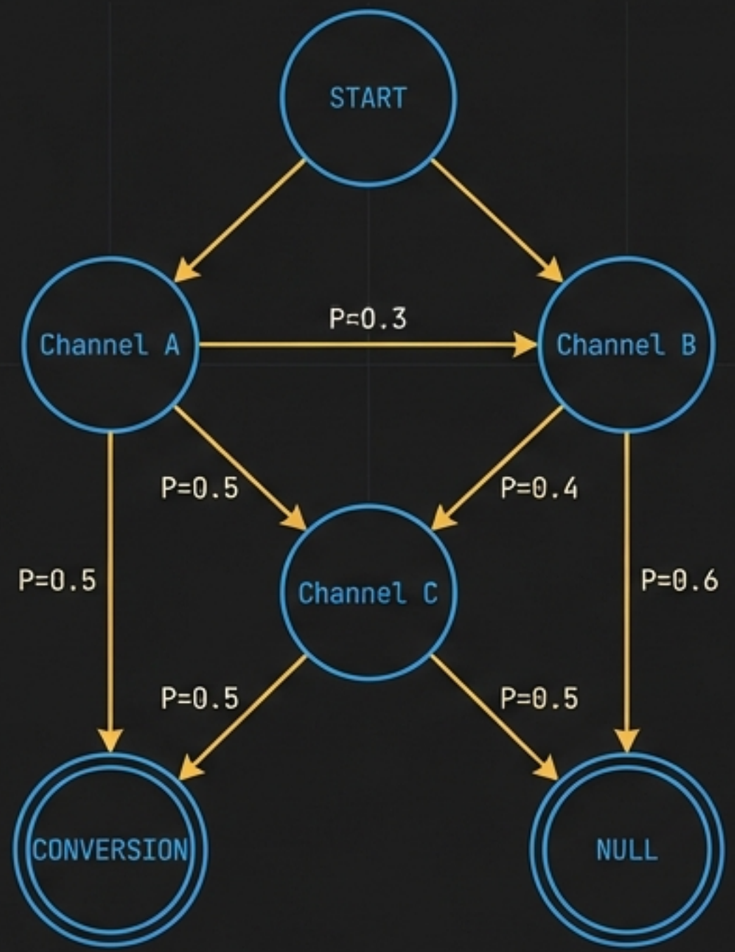
Psychographic Priors

Behavioral Context

Moves beyond raw frequencies by allowing domain knowledge to inform the model. It modulates transition probabilities based on behavioral context, such as user intent or device type.

Pillar 1: Modeling Journeys with Probabilistic Causality

> "The order of touchpoints matters."



The Mechanism

We model user journeys as transitions between states. The probability of the next step depends only on the current state (first-order Markov assumption).

The Fundamental Matrix (N)

$N = (I - Q)^{-1}$. This calculates the expected number of visits to a state before conversion.

The "Removal Effect" (The Killer Feature)

$$M_i = v(N) - v(N \setminus \{i\})$$

Translation: This calculates the "collapse risk" of the funnel. It answers the question: "How much would the conversion probability drop if this channel never existed?"

Pillar 2: Distributing Credit with Axiomatic Fairness

> "Fairness is not a heuristic; it is a mathematical guarantee."

The Logic: Shapley calculates the expected marginal contribution of a channel, averaged over all possible sequences (coalitions) in which it could have appeared.

Efficiency

The sum of all channel credits equals the total value of the conversion.

Symmetry

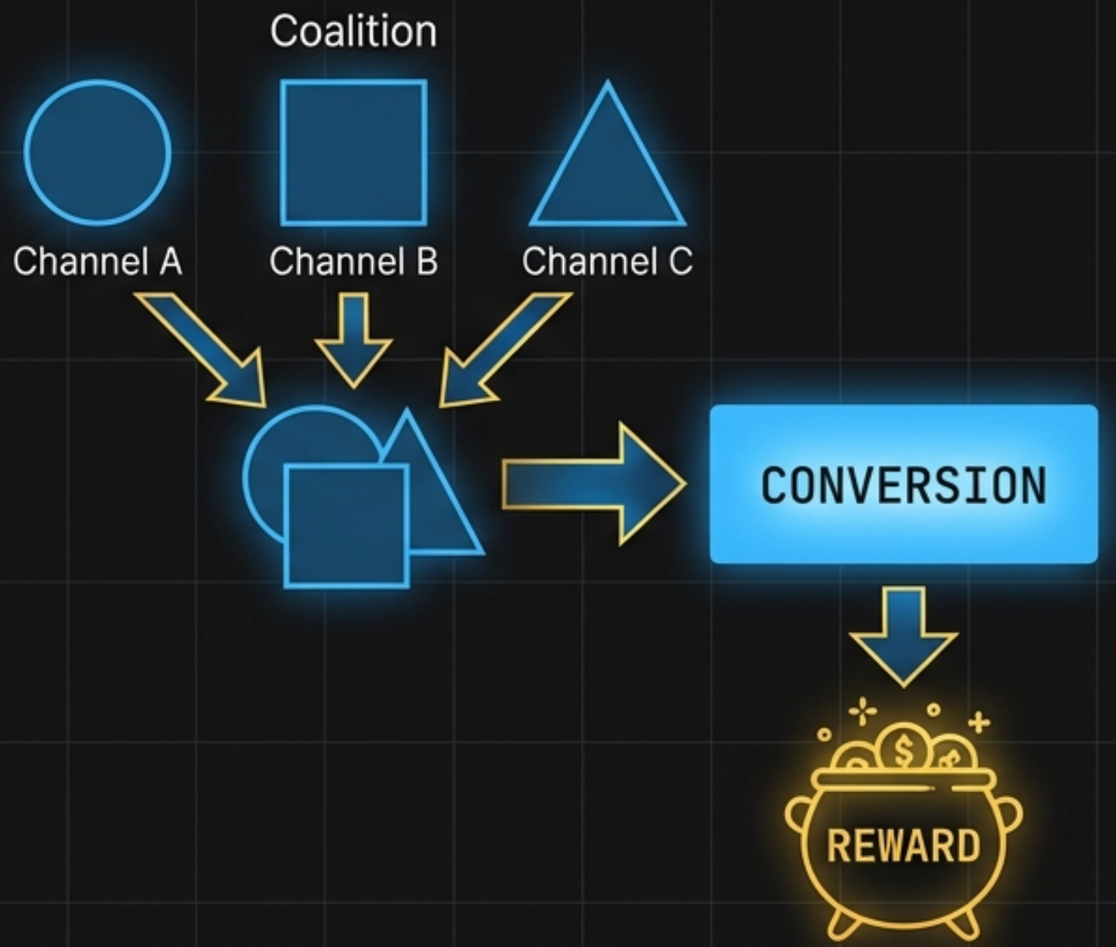
Channels that contribute equally receive equal credit.

Dummy Player

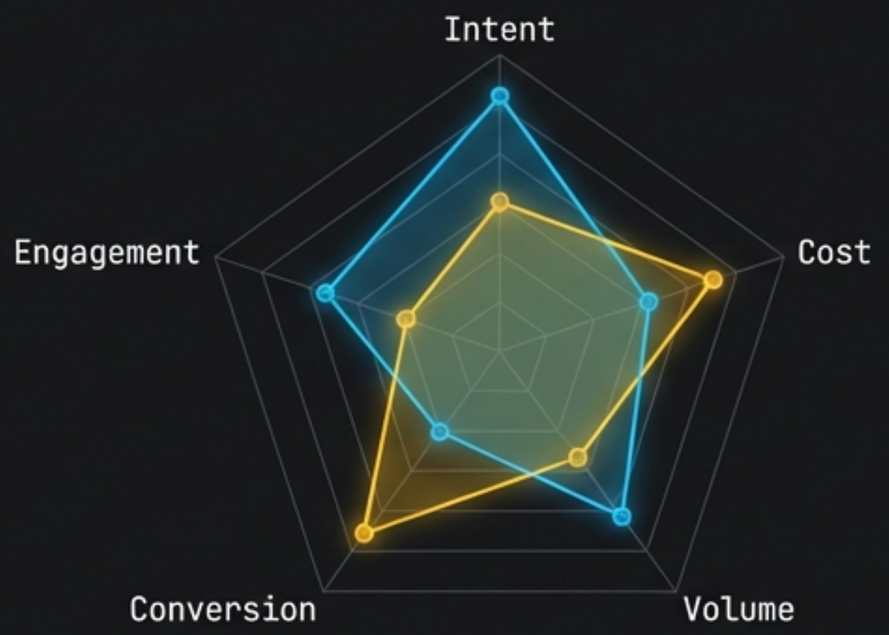
A channel that adds no value to any coalition receives zero credit.

Additivity

The model handles multiple games or value functions consistently.



Pillar 3: The Contextual Layer (Psychographic Priors)



Context Profiling Tab

Profile the behavioral context driving performance. The radar chart compares channels across Markov, Shapley, and Hybrid scores, while weight cards show the impact of psychographic priors like "high-intent search".

The Insight

A click is not just a click. Human decisions are context-dependent.

The Mechanism

We apply weights (w) to transitions based on the nature of the interaction.

- **Amplifiers (w > 1.0):**
  - high-intent search: 1.5x
  - desktop checkout: 1.3x
  - desktop checkout: 1.3x

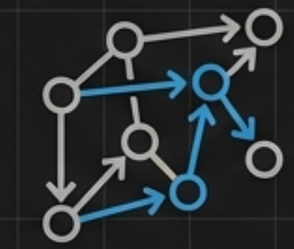
  Signals active problem solving or high friction.
- **Dampeners (w < 1.0):**
  - social feed passive: 0.8x
  - display ad view: 0.7x

  Signals accidental discovery or low engagement.

This bridges the gap between raw data logs and actual human intent.

The Synthesis: Stacking Causality and Fairness

This is not model averaging. It is model stacking.



Raw Path Data

Markov Engine

Computes Causal Value

v(S)

Shapley Engine

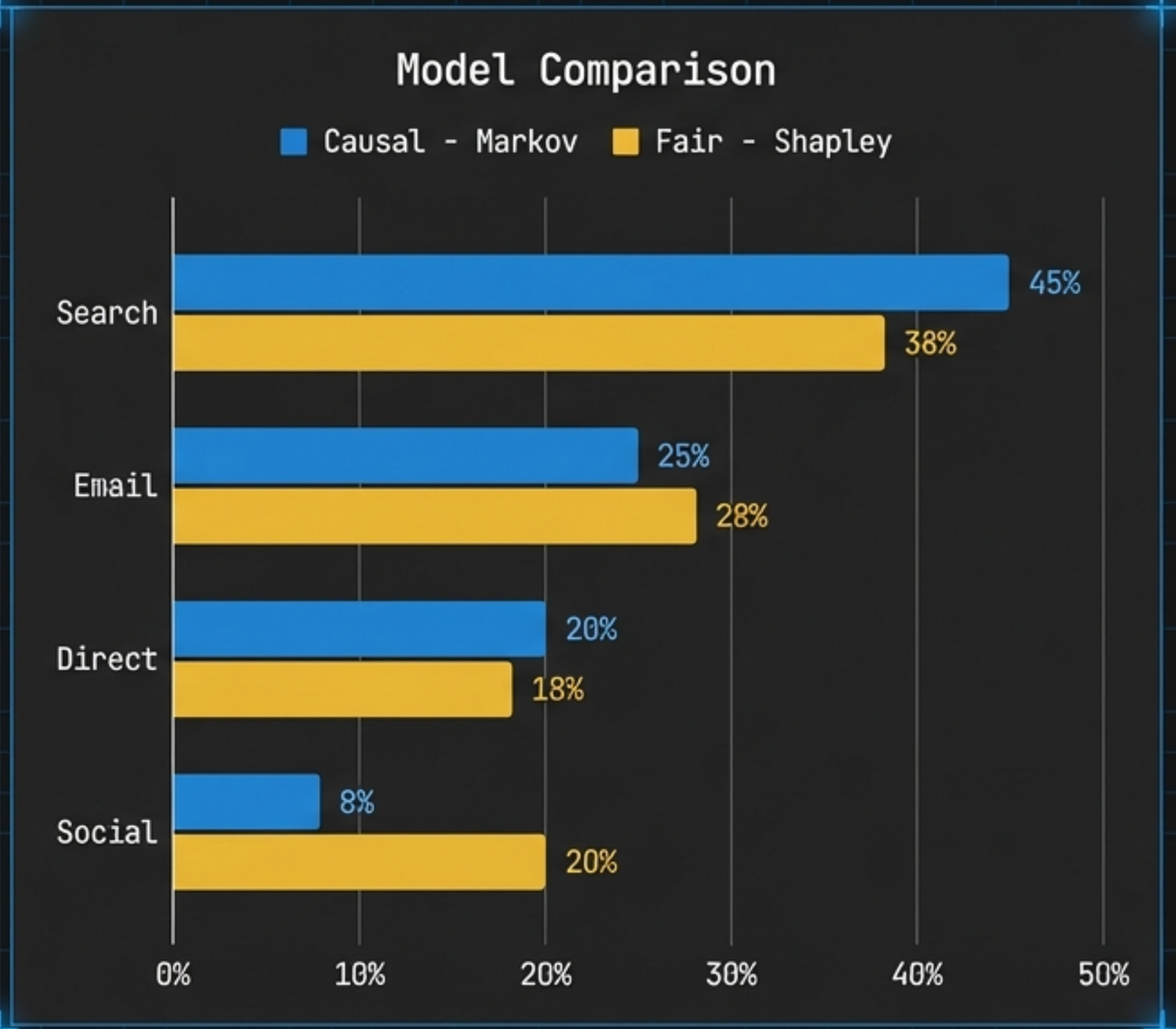
Distributes Value Fairly

Hybrid Attribution Shares

$$\text{Hybrid Share} = \alpha \times \text{Markov Share} + (1-\alpha) \times \text{Shapley Share}$$

α (Alpha) is the blend parameter. It controls the emphasis on pure causality (α=1) versus pure fairness (α=0). The default is a balanced α=0.5.

Strategic Archetypes: Reading the Hybrid Output



Archetype Definitions

1. Structural Winners (High Causality, High Fairness)

Definition: High removal effect. Removing this channel causes funnel collapse.

Action: Defend budget at all costs.

2. Coalition Riders (Low Causality, High Fairness)

Definition: Present in many successful paths but removing them barely impacts conversion probability. They are 'along for the ride'.

Action: Audit for efficiency or redundancy.

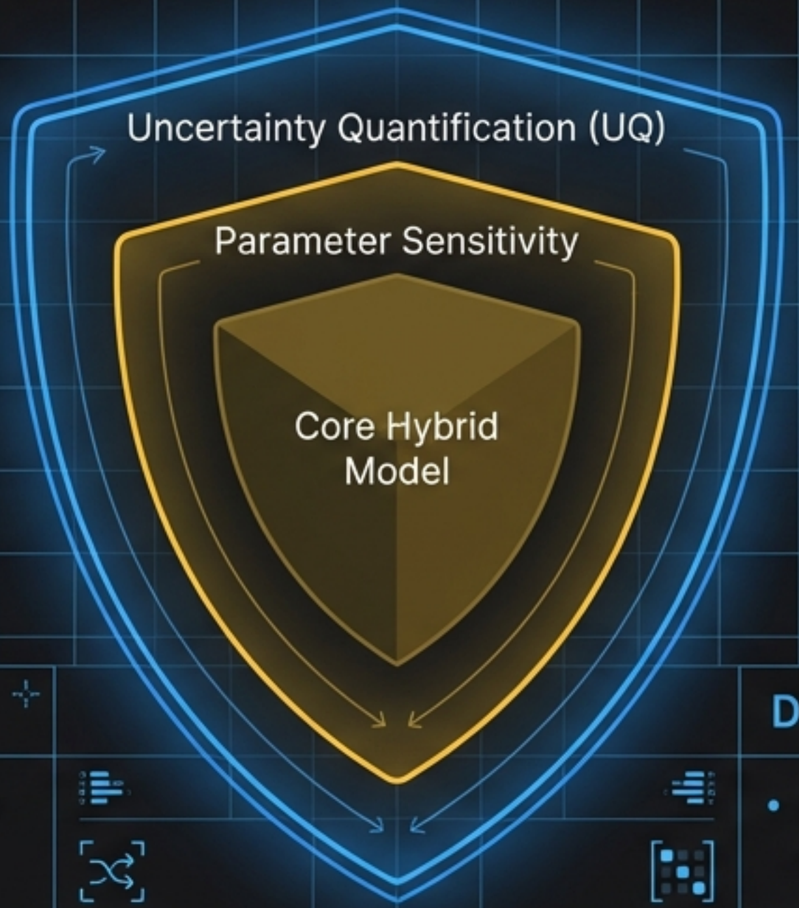
3. The 'Closer' Trap

Definition: High last-touch credit, but low Markov/Shapley scores.

Action: Reallocate budget to earlier influencers.

Making Truth Defensible: The Robustness Stack

"An answer without a measure of confidence is an opinion, not an analysis."



Bootstrap UQ (Path Sampling)

- The Question: "What if we had a different set of users?"
- The Method: Resamples user journeys (B=500) to measure data sufficiency.

Dirichlet UQ (Parameter Uncertainty)

- The Question: "How confident are we in the transition probabilities?"
- The Method: Samples transition matrices to measure model stability.

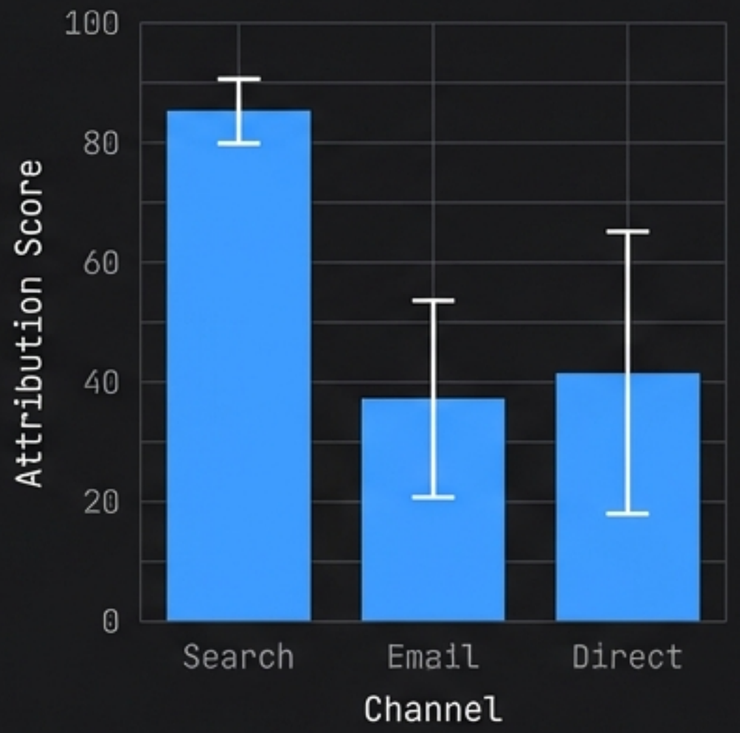
Reading the Robustness: "How Sure Are We?"

Rank Stability Table

	Rank #1	Top 2	Top 3
Search	85%	98%	100%
Email	...	...	...
Direct	...	...	...

Search is a dominant winner (85% Rank #1).

UQ Confidence Intervals



Interpretation

- Wide Bars = High Uncertainty. We need more data.
- Narrow Bars = High Confidence. Trust this decision.
- Rank Stability: Quantifies decision risk. An 85% dominant winner is a safe bet.

The Instrument: From Analysis to Action

The Tuner: Alpha-parameter slider to calibrate Causality vs. Fairness

The Readout: Hybrid Attribution Breakdown

Context Profiling

Transition Flow River

The Invariants of a Correct System

Mathematical guarantees enforced at runtime (Tolerance ≤ 1e-6).

Mathematical Invariants

- **Shares Sum to 1:** All attribution shares must sum to exactly 1.0.
- **Row-Stochastic Matrix**: All transition rows must sum to 1.0.
- **Quantile Ordering**: p05 ≤ p50 ≤ p95.
- **Non-Negative Probabilities**: All probabilities must be ≥ 0.

Computational Guardrails

- **Complexity**: Exact Shapley enumeration capped at n ≤ 12 channels.
- **Reproducibility**: Seeded RNG ensures deterministic outputs.
- **Privacy-First**: No PII stored. Fingerprint-based resolution.
- **Schema Validation**: Strictly validated against ir-schema.json.

The Permanent Record: Frozen & Auditable Artifacts

Analysis is reproducible only when its outputs are canonical. The platform exports five schema-validated Intermediate Representation (IR) artifacts, each with version stamping and generation timestamps.

Architecture Notes:

- **Privacy-First:** No PII stored. Fingerprint-based path resolution.
- **Reproducibility:** Seeded random number generators (RNG).
- **Schema Validation:** All outputs strictly validated against `ir-schema.json`.

Business Impact & Strategic ROI

15-30%

Higher ROI by reallocating budget from ‘Closers’ to ‘Structural Drivers’.

Case Study Insight: Digital News Publisher

- **Discovery:** Last-touch assigned 44% credit to a specific channel.
- **The Truth:** Hybrid Shapley revealed it was a ‘Coalition Rider’ with only 18% structural value.
- **Action:** Budget reallocated to undervalued mid-funnel drivers.

Note: Designed for strategic batch analysis, not real-time bidding.

The Guarantees of First-Principles Attribution

Axiomatically Fair:
Satisfies Shapley axioms: Efficiency, Symmetry, Dummy Player.

Sequence-Aware:
Path order is encoded in the transition matrix.

Counterfactual:
Built to answer “what if?” via removal effects.

Defensible:
Outputs wrapped in a comprehensive robustness stack (Dual UQ).

A complete, correct, and defensible attribution system.