ABDUCTIO (Free): A Framework for Deciding What to Do Next

How to avoid the hidden pitfalls that sink promising projects.

David Joseph

1   1. The Story of Project Chimera: The High-Credence Trap

Imagine you're a VP of R&D at a materials science company. A team of your brightest chemists presents "Project Chimera," a new catalyst that shows a 200% efficiency gain in lab-scale tests. The data is clean, the chemistry is sound, and the team is ecstatic. This looks like the breakthrough you've been waiting for.

The immediate impulse is to celebrate and fast-track a multi-million dollar pilot plant. After all, the evidence is right there. **Your belief that the catalyst *works* is incredibly high.**

This is the **High-Credence Trap**.

The project is a chimera not because the chemistry is wrong, but because an unexamined assumption makes it commercially non-viable. The synthesis requires a rare platinum-group metal as a precursor. At the milligram scale of the lab, its cost is trivial. But a back-of-the-envelope calculation—one that nobody has been incentivized to do yet—shows that producing the catalyst at an industrial ton-scale would cost more than the entire market for the end product is worth.

The project is doomed. But it's not doomed because of the science everyone is focused on; it's doomed because of a simple economic reality hiding in plain sight. An inexperienced team, chasing the excitement of the initial discovery, could waste a year and millions of dollars before hitting this wall.

ABDUCTIO is a system designed to spot these hidden-in-plain-sight pitfalls. It provides a lightweight but rigorous framework to force the right questions at the right time, ensuring that you discover the fatal flaw in two days of analysis, not two years of development.

2   2. The Two Numbers That Matter: Credence and Confidence

The High-Credence Trap arises because we typically only track one metric: our belief in a claim. ABDUCTIO's core insight is that you must track two.

2.1   Credence (p): *What do we believe?*

This is our best guess at the probability of a claim being true, from 0 to 1. For Project Chimera, credence in the claim "the catalyst has 200% efficiency" was high ($p \approx 0.9$).

2.2   Confidence (k): *How stable is our belief?*

This is the crucial second metric. Confidence measures the *robustness* of our belief. It asks: **"How much would our credence (p) move under reasonable re-examination?"**

For Project Chimera, confidence in the *commercial viability* was critically low ($k \approx 0.2$). Why? Because the belief was brittle—it rested on the unexamined assumption of economic scalability.

Crucially, these two numbers can be attached to *any* claim, from the top-level "Is this project a good bet?" to a granular prerequisite like "Can our supply chain source this material affordably?" This allows us to map out uncertainty with precision.

Table 1: The ABDUCTIO 2x2 matrix for triaging uncertainty.

	Low Credence (Probably False)	**High Credence** (Probably True)
High Confidence	*"The Obvious No"* Action: Reject & Stop.	*"The Solid Bet"* Action: Accept & Proceed.
Low Confidence	*"The Plausible Longshot"* Action: Investigate Bottleneck.	*"The High-Credence Trap"* (Project Chimera) Action: Investigate Bottleneck.

3 3. Reasoner-Agnostic Orchestration

ABDUCTIO standardizes the signal, not the method. It's a reasoner-agnostic framework: any assessor—a subject-matter expert, a cross-functional reviewer, an LLM, or an automated check—can produce a pair (p, k) for any proposition.

- **Uniform Signal:** Every assessor returns credence p, confidence k, and three short factors that most influenced their scores.

- **Method Neutrality:** Literature review, simulation, heuristics, or model inference are all valid ways to arrive at (p, k).

- **Composable Panels:** Mix human experts for depth, LLMs for speed and breadth, and automated "sanity checks" (e.g., unit tests, order-of-magnitude checks) for guardrails.

This approach gives you the power of collective intelligence without a monolithic method—a bigger brain, not a bigger bureaucracy.

4 4. The Engine: How ABDUCTIO Works

ABDUCTIO runs a simple loop guided by one idea: **only go deeper when it's likely to change the decision.** Before explaining the loop, let's look at its components.

4.1 Collective Intelligence, Without Chaos

Ensembles are powerful if disciplined. ABDUCTIO uses four guardrails:

1. **Independence First:** We collect the first round of (p, k) independently to avoid groupthink.

2. **Robust Pooling:** The free tier uses medians for both p and k across assessors, which resists outliers and basic gaming.

3. **Structured Disagreement:** We display a disagreement map showing the spread of p and k. Disagreement isn't a failure; it's a compass pointing to the most important sub-problem.

4. **Lightweight Iteration:** After pooling, assessors may optionally revise once. The process stops if the median k passes the confidence gate.

4.2 Decomposition Grammar (Free Edition)

We keep the reasoning graph simple and auditable:

- **Nodes:** Propositions with (p, k) and a 1-3 line rationale.

- **Edges: AND** ("all must hold") or **OR** ("any route suffices").

- **Aggregation:** Conservative defaults: AND uses $p = \min, k = \min$; OR uses $p = \max, k = \max$.

4.3 The Loop: Assess → Gate → Decompose → Repeat

1. **Assess the claim.** Elicit **Credence** (p) and **Confidence** (k) from your panel of reasoners.

2. **Gate on confidence.** If the pooled median confidence $k \geq \tau$ (your chosen confidence gate), act on what you have—proceed or reject. If $k < \tau$, don't debate: decompose.

3. **Decompose the claim.** Ask: "What must be true for this to be true?" Create sub-claims with AND/OR links.

4. **Find the bottleneck.** Pick the lowest-confidence prerequisite that gates the parent (an AND-link) or unlocks a major route (an OR-link). This becomes the single next step for investigation.

5. **Act and Update.** The free tool asks you to input a rough cost and turnaround time for this next step, showing you the decision leverage versus the effort. Once you run that one analysis or experiment, you update its (p, k) and re-run the gate.

5 5. ABDUCTIO in Action: Applications

This framework brings clarity to any high-stakes problem:

- **Technology:** An AI model shows state-of-the-art accuracy (high credence), but confidence is low because its robustness to adversarial attacks is unknown. ABDUCTIO directs effort to that specific robustness test, not more accuracy benchmarks.

- **Strategy:** A proposed merger looks great on paper (high credence), but confidence is low due to unexamined assumptions about cultural integration. ABDUCTIO prioritizes a cultural due diligence step over further financial modeling.

- **Research:** A new Alzheimer's drug shows promise in mice (high credence), but confidence in its human efficacy is low. ABDUCTIO helps decompose the "translatability" problem to pinpoint the most uncertain biological mechanism to test next.

6 6. The ABDUCTIO Difference: Radical Transparency

ABDUCTIO favors inspectable reasoning over black boxes. The free tier gives your team the tools to build confidence in your process.

6.1 Provenance & Audit (Free "Lite")

- Every node logs who, when, (p, k), and the three influencing factors.

- Each change creates a new revision; old states remain visible.

- Reasoning trees are shareable within your organization and exportable.

6.2 Calibration, Without the Lab Coat

Confidence is only useful if it's calibrated. The free tier gives teams a light, honest kit:

- **Anchors:** Use a handful of known-outcome "calibration cards" from your domain to see how your estimates perform.

- **Reliability Strip:** A quick plot of your stated p vs. actual outcomes on anchor cases provides a simple sanity check on your team's calibration.

- **Self-Awareness Prompts:** Nudge assessors to name what would change their minds, often revealing the next decomposition step.

6.3 Where Enterprise Takes Over

- **Confidence you can trust:** Compute k from stability checks (paraphrase/sensitivity tests, rater calibration) instead of self-report.

- **Prioritization that pays:** Automated, cost-aware prioritization that weighs decision leverage and declared costs.

- **Assurance at scale:** Sealed audit bundles, portfolio views, and governance hooks for compliance-heavy teams.

7 7. Conclusion

The challenge in modern decision-making isn't a lack of information; it's a lack of a rational framework for processing it. We fall into traps like Project Chimera not because we are unintelligent, but because we are undisciplined in how we examine our own beliefs.

ABDUCTIO provides that discipline. It's a simple, powerful system for surfacing the unexamined assumptions that put our most important projects at risk. It helps us understand not just what we don't know, but **what's worth knowing next.**

8 Appendix: Interchange Card Schema

To combine humans, models, and automated checks, ABDUCTIO uses a simple, uniform data structure for every assessment.

```
{
  "proposition": "SAE level-4 stack is safe for downtown deployment in 2026",
  "assessment": {
    "p": 0.58,
    "k": 0.42,
```

```
      "factors": [
        "rare-edge-case handling unclear",
        "simulation-to-road gap unresolved",
        "sensor redundancy promising"
      ]
    },
    "assessor": {
      "type": "human|llm|check",
      "role": "safety_engineer",
      "id": "anon-1372"
    },
    "context": {
      "evidence_refs": ["doc://safety-PRD#42", "sim://run-0198"],
      "timestamp": "2025-09-27T17:01:13Z"
    }
}
```