

The bottleneck isn't new ideas.

*It's knowing which beliefs are **wrong**.*

A falsification engine for human knowledge. We model science as a graph of claims, find the load-bearing assumptions resting on thin evidence, and compute the single experiment that would remove the most uncertainty.

THE MISUNDERSTOOD PROBLEM

We optimize the wrong end of the pipeline.

The folk model of progress:



Reality: progress is gated not by the supply of hypotheses but by *selection under uncertainty* — knowing which beliefs are false, and which experiment is worth a decade of a life.

We have industrial infrastructure for producing claims — and almost none for retiring them. Science has no “git blame” for belief.

Nothing represents knowledge as a dependency structure.

Citation graphs

Count attention, not logical dependence. A cited result can be wrong; a quiet one load-bearing.

Peer review

Judges one paper, once. No memory, no global view of contradictions elsewhere.

Large language models

Trained to reproduce consensus — they amplify fragile beliefs and call it truth.

Meta-analysis

Powerful but narrow, slow, manual. Cannot reason across questions.

The common blind spot: no signed, weighted, dependency-aware model. Without it, fragility and experimental value are undefined.

A belief is not true or false.

It has a probability and a set of dependencies.

01

Probability

Evidence is a Bayesian operator — each observation a signed log-likelihood ratio.

02

Dependency

Claims form a graph. A doubtful parent erodes its children's grounding.

03

Fragility

Retract one study, re-infer. The swing reveals single points of failure.

04

EVE

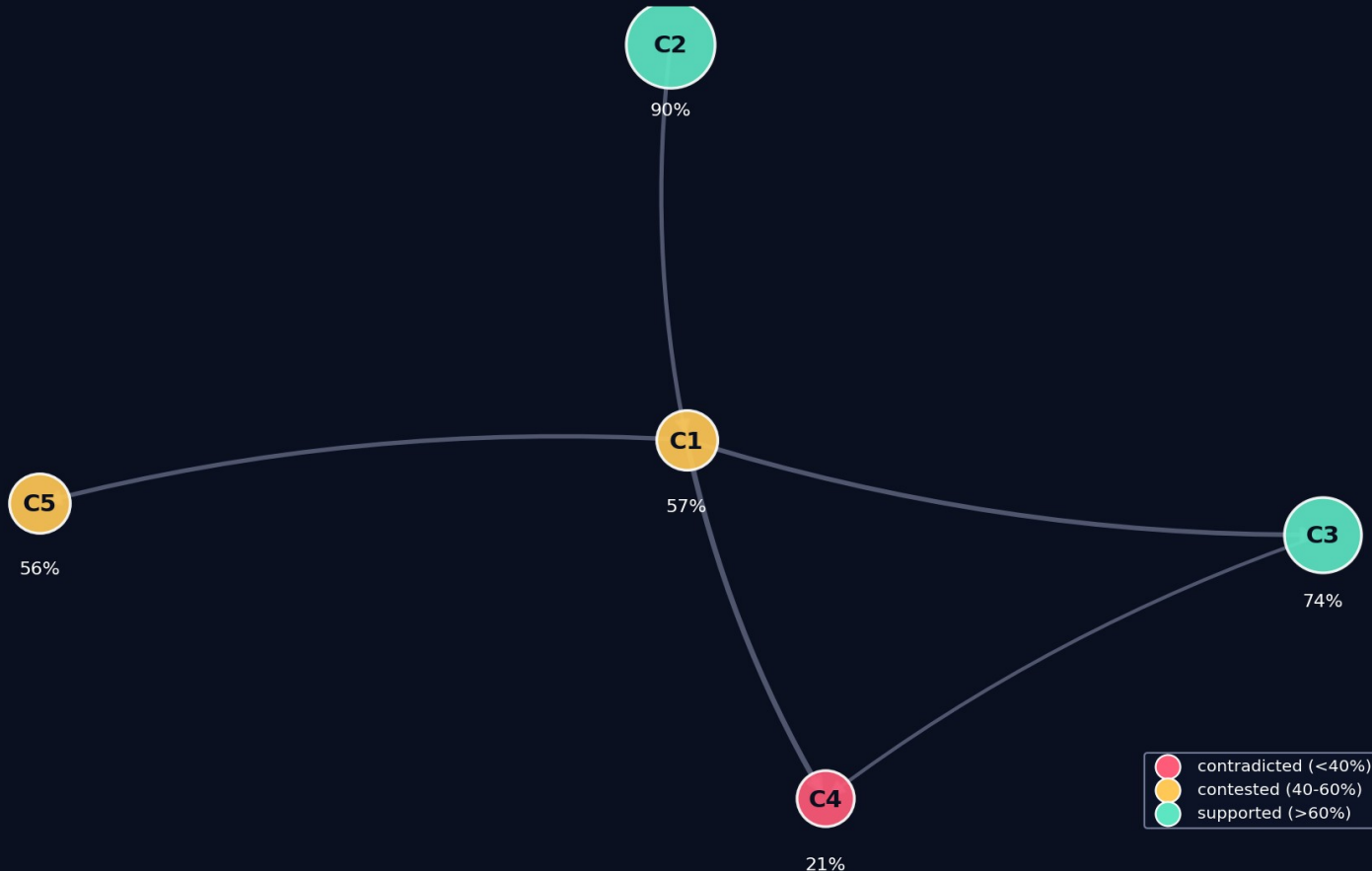
Price the unknown: rank experiments by expected uncertainty removed.

THE PROTOTYPE · IT WORKS

Seeded with a real contested field, the engine found the keystone itself.

ALETHEIA · The Belief Graph

node colour = collective confidence · node size = fragility (collapse risk if one study fails)



57%

Keystone causal claim — propped up by a confounded correlation.

21%

Human-trial claim is actively contradicted by a large null result.

3×

Top experiment removes 3× more uncertainty than any alternative.

Which experiment should humanity run next?



Result, in nats of information

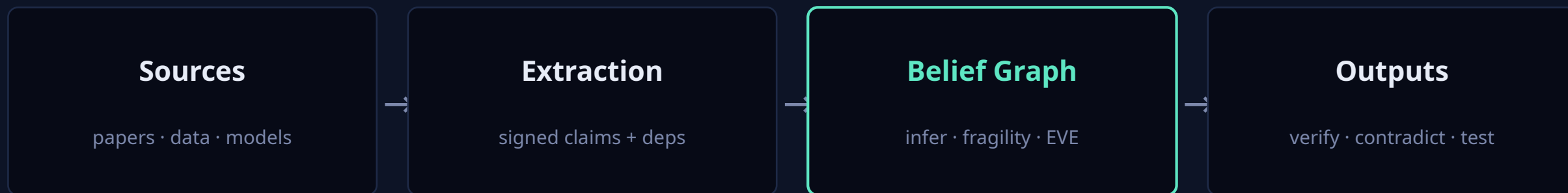
0.67

removed by the single
best experiment — the
keystone the whole
field rests on.

For the first time, the most important experiment we are NOT running is a computable number.

FROM PROTOTYPE TO SUBSTRATE

A self-auditing, forkable commons for knowledge.



↻ Every new result re-enters at the top. The loop closes — knowledge becomes self-auditing.

Forkable like version control. Competing communities keep different priors and see exactly where — and why — their conclusions diverge.

Human knowledge, with a memory and a conscience.

Funding becomes information-optimal

Fund the experiments that remove the most collective uncertainty — not the most fashionable.

Errors self-heal

Overtake a result, and every dependent conclusion is flagged and recomputed. No more silent rot.

AI becomes a scientist

Not a parrot of consensus — an agent that proposes the experiments that most advance us, and says why.

We did not get version control for prettier code — we got it because coordinating many minds is impossible without a shared, auditable history. Knowledge is the largest changing artifact we have, and it has none.