METHODOLOGY

WeWork

Retrodiction methodology

Paul Campillo

Version 1.0 · April 2026 · Complete

What a retrodiction is

A retrodiction applies a diagnostic framework to a historical case where the outcome is already known—but runs the analysis as if standing at a past date, using only evidence available at that date. It then scores the diagnosis against what actually happened.

The purpose isn't to explain the past (any narrative can do that after the fact). The purpose is to test whether the framework could have **predicted** the outcome from the evidence available **before** the outcome was known. If it can, the framework captures real structure. If it can't, the framework is just storytelling.

Why WeWork

WeWork is the second retrodiction case for three reasons:

It tests a different failure type. The first retrodiction (Kodak) was an information-channel failure: broken reality-sensing, 21 years of suppressed correction, three CEOs who each produced the same wrong model. WeWork is a resources-channel failure: the business model generates no margin, masked by external capital. If Genesis only worked on one failure type, it would be a pattern-matching trick, not a structural framework.

The conventional narrative is misleading. The popular story is "Adam Neumann was a bad CEO" or "WeWork was a fraud." These explanations are partially true but structurally incomplete—they can't explain why SoftBank invested $18.5B, why Goldman and JPMorgan underwrote an IPO at $47B, or why every standard metric (revenue growth, member growth, market expansion) pointed upward until six weeks before collapse. Genesis must produce something the conventional narrative cannot: a structural diagnosis that was **available at peak**, not only visible in hindsight.

It tests early detection under uncertainty. The Kodak retrodiction had rich public data (SEC filings, 20+ years of analyst coverage, academic case studies). WeWork at T0 was a private company with thin disclosure: one year of financial data (2016), no occupancy metrics, no per-unit economics, no precise lease obligations. This tests whether Genesis can produce useful signal from sparse evidence—the condition most real-world predictions face.

The protocol sequence

Genesis ran four protocols in sequence, each consuming the prior one's output:

STEP	PROTOCOL	WHAT IT DOES	OUTPUT
1	Research (-1)	Collects evidence at each checkpoint, admitting only information available at that date	Evidence packages: T0 (29 claims, scrubbed), T1 (29 + 10 cascaded), T2 (original)
2	Diagnostic (0)	Scores the system's capacity, boundaries, margin, navigation, and trajectory	SIRF scores, margin, posture, boundary inventory, bet register at each cutoff
3	Prediction (+0.5)	Generates counterfactual forecasts—what Genesis would have predicted standing at each date	Scenario bundles, trajectory projections, 15 bet register entries
4	Retrodiction (+1)	Scores all diagnoses and predictions against the known outcome	Brier scores, directional accuracy, error analysis, pattern library entry

The complete scored record is hash-locked in the Seer forecast registry as SEER-2026-0004.

The three checkpoints

CHECKPOINT	DATE	WHY THIS DATE	EVIDENCE CLAIMS
T0	August 2017	Peak private valuation. SoftBank's $4.4B just invested. Maximum distance from outcome. The hardest test—and the most informative if the framework can produce signal here.	29 (scrubbed)
T1	January 2019	The $16B SoftBank deal just collapsed. The first disconfirming signal. Tests whether Genesis tracks the transition from masked-Crisis to visible-Crisis.	29 + 10 (cascaded from T0 scrub)
T2	August 2019	S-1 filed. IPO imminent. The last moment before the outcome becomes visible to everyone. Tests whether Genesis identifies the dual trap before the market does.	Original package

The evidence scrub

The WeWork retrodiction introduced a protocol discipline that the Kodak case didn't require: a mandatory evidence scrub with hard temporal cutoff.

The original T0 evidence package contained 11 items that were not available until after August 2017: 2017 full-year financials (disclosed via April 2018 bond offering), per-desk loss data, precise lease obligations ($18.2B end-2017), occupancy rates, enterprise revenue share, the SpaceMob acquisition, WeGrow announcement, Rise by We launch, and Powered by We announcement.

These items were **useful** for the diagnostic—they sharpened the picture considerably. But they weren't **available** at the prediction date. Including them would contaminate the retrodiction with hindsight.

THE SCRUB PROTOCOL

Every claim in a T0 evidence package must cite a publication date on or before the cutoff. Items that fail the test are displaced to the next temporal window where they become available. The diagnostic and prediction are rebuilt from the scrubbed evidence only. Displaced items enter the later window as explicit delta claims with their actual availability dates.

The scrub displaced 11 items. Ten entered T1 (available between August 2017 and January 2019). One cascaded further to T2 (the 119% net member retention figure, disclosed only in the S-1 filing).

The impact was significant: Ψ_{norm} at T0 rose from 0.14 (Terminal) to 0.21 (Critical). The confidence ceiling dropped from 0.72 to 0.58. Several dramatic findings—the delusion cascade, the navigator field-emission audit, the precise lease obligations—were correctly attributed to T1 instead of T0.

This cost claim strength. But it gained scientific credibility—and revealed that the T0→T1 decline was **steeper** with honest baselines (74% vs. 61%), because the contaminated T0 was already reflecting mid-decline data.

The two-layer evidence structure

Each evidence claim in the WeWork retrodiction is split into two layers:

LAYER	CONTAINS	EXAMPLE
Research Layer	Observed facts and narrow inferences with explicit falsifiers. No construct classifications.	"2016 full-year: revenue $436M, net loss $429M. Loss-to-revenue ratio: 0.98:1." **Falsifier:** "H1 2017 data could show improvement."
Analyst Annotation	Genesis construct labels. Applied to Research Layer claims. Clearly separated.	"F-channel signal: loss ~ revenue indicates no self-sustaining economics." Applied to E001.

This separation prevents construct smuggling—the tendency for a framework to read its own conclusions into the evidence rather than deriving them from it. The Diagnostic protocol cannot introduce new factual claims; it can only annotate the Research Layer claims with construct classifications.

The steel-man bull case

The WeWork retrodiction includes a mandatory section that the Kodak case didn't have: a steel-man bull case stated as an advocate would argue it, not as a strawman.

At T0, the bull case for WeWork was genuinely defensible:

- WeWork was 7 years old and doubling annually. IWG took 20+ years to reach profitability.
- Enterprise adoption was the fastest-growing segment, with stickier and higher-value contracts.
- The flex-work macro trend was real and accelerating.
- SoftBank's $4.4B provided 24+ months of runway regardless of operating losses.
- Loss-to-revenue ratios are common in high-growth startups; they don't necessarily indicate structural failure.

None of these arguments were falsified by T0 evidence. The diagnostic acknowledged this honestly, which is why the Upside scenario received 0.20 probability (higher than a weaker bull case would warrant) and the confidence ceiling was set at 0.58 rather than higher.

The bull case turned out to be wrong—enterprise economics didn't change the trajectory, and growth-phase losses proved to be structural. But the honest treatment of the bull case at T0 is what gives the retrodiction its scientific credibility. The framework didn't pretend to know more than it did.

Scoring methodology

Two levels of scoring, measuring different things:

Bet-level Brier score: Scores each individual prediction against its binary outcome. For each prediction with assigned probability p and outcome o (1 = correct, 0 = incorrect):

$$\text{BS} = (p - o)^2$$

The T0 and T1/T2 groups are scored separately because they reflect different evidence quality. T0 mean Brier: 0.217 (8 predictions at mean confidence 0.54). T1/T2 mean Brier: 0.042 (7 predictions at mean confidence 0.81). Combined: 0.136.

Directional accuracy: Did the prediction correctly identify the structural trajectory? Result: 15/15 (100%). No directional errors at any cutoff.

Differential accuracy: Each Genesis prediction compared against what SoftBank, Goldman Sachs, Fidelity, and short sellers would have predicted at the same checkpoint with the same evidence. Differential: +0.85. Genesis added substantial predictive value beyond all baseline approaches.

What the retrodiction does not claim

This retrodiction demonstrates that Genesis captures structural patterns in venture-backed hypergrowth that conventional analysis misses. It does **not** claim:

Certainty at T0. The T0 predictions carried 45–58% confidence—appropriately uncertain given sparse evidence and a defensible bull case. Genesis doesn't claim to have **known** WeWork would fail. It claims to have identified **why** it

was structurally vulnerable, at modest confidence, when every market signal said triumph.

Uniqueness. Short sellers identified the lease model as risky. The WeWork skeptic community existed. Genesis's advantage isn't the risk identification—it's the mechanism specificity: which channel was the persistent bottleneck, why the navigator wouldn't self-correct, why enterprise was the surviving boundary, and the structural comparison with IWG that predicted asymmetric outcomes from the same market conditions.

Hindsight elimination. Despite the evidence scrub, two-layer packaging, and bull-case discipline, the analyst running the retrodiction in 2026 knew the outcome. These safeguards mitigate but cannot fully eliminate unconscious bias toward the known result. The scrub protocol—developed during this case—represents a significant improvement in retrodiction hygiene over the Kodak case.

The honest claim: Genesis identified structural Crisis, specific bottleneck mechanisms, and behavioral predictions from peak-era evidence with a combined Brier of 0.136 and 100% directional accuracy. The early signal (T0) was correctly identified at modest confidence. The signal sharpened dramatically at T1 when evidence quality improved. This is what early structural detection looks like under real-world evidence constraints.

Protocol improvements from this case

The WeWork retrodiction produced five protocol improvements now mandatory for all future cases:

IMPROVEMENT	WHAT IT DOES	WHY IT MATTERS
Mandatory cutoff scrub	Every T0 claim must cite a publication date before the cutoff	Prevents hindsight contamination—the single biggest threat to credibility
Two-layer evidence	Research Layer (facts) separated from Analyst Annotation Layer (constructs)	Prevents construct smuggling—the framework can't read its own conclusions into the evidence
Steel-man bull case	State the strongest counter-argument as an advocate would	Forces honest uncertainty assessment—if the bull case is defensible, confidence must reflect that
Shared-evidence flagging	Flag when the same evidence loads on multiple constructs	Prevents false convergence—four constructs citing the same fact is one signal, not four
Binding confidence ceilings	Evidence ceiling propagates to all downstream artifacts	Prevents confidence inflation—you can't assign 0.85 to a bet from a 0.58 ceiling

Protocol: Protocol v0.5 | Research (-1) · Diagnostic (0) · Prediction (+0.5) · Retrodiction (+1) | **Evidence:** T0: 29 claims (scrubbed to Aug 2017). T1: 29 + 10 cascaded. T2: original. 15 scored bets.
Forecast registered: SEER-2026-0004 · github.com/wakeupsmiling/seer · Combined Brier: 0.136. Full analysis: genesistheory.org