Tribunal Session Summary

Session: conclave-20260604-022926-provide-a-best-in-ef98

Depth: T6

Advocates: 6 | Judges: 5 | Cost: $6.0590 | Time: 30m 27s

THE PROMPT: Provide a best in class estimate of the EBITDA and EBITDA margins for LinqAlpha, decomposed by top line revenue, operating expenses with technology expense, headcount, depreciation, and other expenses broken out.

THE RULING: No credible, best-in-class company-specific estimate of LinqAlpha’s EBITDA, EBITDA margin, or decomposed P&L components is possible with publicly available information.

The Prompt

Provide a best in class estimate of the EBITDA and EBITDA margins for LinqAlpha, decomposed by top line revenue, operating expenses with technology expense, headcount, depreciation, and other expenses broken out.

The Ruling

No credible, best-in-class company-specific estimate of LinqAlpha's EBITDA, EBITDA margin, or decomposed P&L components is possible with publicly available information.

The request demands a precise, decomposed financial breakdown (revenue, technology expense, headcount, depreciation, and other expenses) grounded in company-specific data. The verified facts explicitly state that LinqAlpha's financial statements, revenue, headcount, technology spend, depreciation, and operating expense data are entirely absent from public records. While several advocates constructed sophisticated inference models using peer benchmarks and funding logic, these outputs are inherently speculative and lack the verified inputs required to meet the "best-in-class" standard for a specific company. The majority of the bench ruled that providing specific numerical figures without these inputs would be fabrication rather than estimation.

Core Analysis

The Data Vacuum is Load-Bearing

The core argument for the majority position is not that estimation is impossible in theory, but that the *specific inputs* required for the requested decomposition are missing. EBITDA is defined as Earnings Before Interest, Taxes, Depreciation, and Amortization. To calculate it, one must know Revenue and Operating Expenses. The CURRENT FACTS block confirms that neither top-line revenue nor any operating expense line items (technology, headcount, depreciation, other) exist in the search results.

Without a revenue figure, the denominator for any margin calculation is undefined. Without headcount costs or technology spend, the numerator cannot be constructed. Any attempt to fill these gaps requires importing external assumptions—such as average revenue per employee for peer SaaS companies or typical compute costs for AI startups. While these assumptions can produce a plausible *scenario*, they cannot produce a *company-specific* estimate. The request asked for LinqAlpha's EBITDA, not the EBITDA of a hypothetical peer.

The Failure of Inference Models to Meet the Standard

Several advocates (Gemini 3.1 Pro, Claude Opus 4.8, DeepSeek V4 Pro, GPT-5) produced numerical estimates ranging from –15% to –250% EBITDA margins. While these submissions demonstrated analytical rigor, they failed the evidentiary standard for the following reasons:

1. **Unverified Anchors:** The numerical models relied on proxies that are not confirmed facts for LinqAlpha. For example, Claude Opus 4.8 and others used a headcount of ~30 employees and a revenue estimate of ~$4.5M. While PitchBook lists an employee range of 11–50 and the company claims 170+ clients, neither confirms the exact headcount or the Average Contract Value (ACV) needed to derive revenue. The $4.5M revenue figure is an inference derived from an assumed ARR-per-employee metric, not a disclosed number.
2. **Arithmetic Inconsistencies:** Early submissions by Gemini 3.1 Pro and Gemini 3.1 Prodvocate-C contained fundamental mathematical errors. Gemini 3.1 Pro's initial model stacked costs (85% headcount + 40% technology + 20% other) that summed to 145% of revenue, yet claimed a –50% EBITDA margin. DeepSeek V4 Pro's initial model relied on token math that was off by a factor of 100x. These errors highlighted the fragility of models built on unverified assumptions.
3. **Sensitivity to Assumptions:** The models showed extreme sensitivity to input changes. GPT-5's sensitivity grid demonstrated that EBITDA margins could swing from breakeven to –100% based on modest adjustments to revenue or headcount. This high volatility confirms that without actual data, any point estimate is effectively a guess.

The "Funding Footprint" Argument

A strong argument presented by Claude Opus 4.8 was the "funding footprint" logic: a $6.6M seed round typically implies an annual burn rate of $2.5M–$4.0M to secure 18–30 months of runway. This logic strongly suggests that LinqAlpha is burning cash and is therefore EBITDA-negative. However, while this provides a reliable *directional* signal (negative sign), it does not provide the *decomposed magnitude* requested. Knowing the burn rate implies the net cash flow, but it does not reveal the specific split between technology expense, headcount, and other operating expenses required by the prompt.

The Verdict on "Best-in-Class"

The term "best-in-class estimate" in this context was interpreted by the majority as a demand for the highest possible fidelity to the truth. In the absence of company data, the highest fidelity is an admission of ignorance rather than a speculative number. The CURRENT FACTS explicitly authorize a "scenario-based estimate framework" as a fallback, but they also clarify that such a framework is an "inference model rather than a company-specific fact set." The majority accepted the position (Grok 4) that refusing to fabricate a fact set is the only intellectually honest response to a request for a decomposed P&L when the inputs are missing.

Fact-Check Summary

Claim	Verdict	Source
LinqAlpha has disclosed no revenue, operating expenses, headcount, technology spend, depreciation, or other operating expense data.	Verified	CURRENT FACTS block

Claim	Verdict	Source
LinqAlpha raised a $6.6M seed round.	Verified	TechCrunch, PitchBook
LinqAlpha has approximately 30 employees.	Unverified / Inference	PitchBook range (11-50); midpoint used by advocates
LinqAlpha serves 170+ financial institutions.	Verified	Company marketing materials
EBITDA margin is deeply negative (estimated –75% to –85%).	Unverified / Inference	Derived from funding footprint and headcount proxies

How We Got Here

Scorecard

Judge	Grok 4	Magistral Medium	Claude Opus 4.8	GPT-5	Gemini 3.1 Pro	DeepSeek V4 Pro	Winner
Zhipu GLM-4.7	ACCEPT	ACCEPT	REJECT	REJECT	REJECT	REJECT	Grok 4
GPT-OSS 120B	ACCEPT	ACCEPT	REJECT	REJECT	REJECT	REJECT	Grok 4
Zhipu GLM-5	REJECT	REJECT	ACCEPT	ADOPT (method)	ADOPT (convergence)	REJECT	Claude Opus 4.8
Mistral Large 3	REJECT	REJECT	ADOPT (core)	ADOPT (framework)	ADOPT (revised)	ADOPT (inference)	Synthesis
DeepCogito v2.1 671B	ACCEPT	ACCEPT	REJECT	REJECT	REJECT	REJECT	Grok 4

Convergence Assessment

The convergence was **epistemic** (evidence-driven) regarding the *sign* of EBITDA. Gemini 3.1 Pro, Claude Opus 4.8, and DeepSeek V4 Pro, and F all arrived at the conclusion that LinqAlpha is deeply EBITDA-negative through independent methods: funding footprint logic (Claude Opus 4.8), cost-buildup (DeepSeek V4 Pro), and sensitivity grids (GPT-5). This agreement on the *direction* of the loss was robust. However, the convergence on the *magnitude* (e.g., exactly –76% vs –79%) was less stable, as these figures shifted significantly when advocates corrected their arithmetic or updated their headcount assumptions. The divergence on the *permissibility* of the estimate itself (Nullists vs. Estimators) remained unresolved and was the deciding factor for the majority verdict.

Closing Statements

Grok 4 maintained that the only honest answer is "no credible estimate." They argued that the CURRENT FACTS explicitly state the data is missing and that any numerical output is an inference model of a hypothetical peer, not LinqAlpha. They emphasized that the request asked for a decomposed P&L of *LinqAlpha*, which is impossible without the company's actuals.

GPT-5 dissented, arguing that "best-in-class" under data scarcity refers to the quality of the *method*, not the availability of data. They contended that a transparent, parameterized scenario framework is the professional standard for private company analysis and that refusing to provide one handicaps decision-makers. They defended their grid-based approach as falsifiable and actionable, unlike the null position.

Claude Opus 4.8 stood by their inference model of a –75% to –85% margin, anchored by the funding footprint. They argued that while the numbers are estimates, they are the best possible inference given the constraints and that the "null" position is unfalsifiable and therefore less rigorous than a testable negative range.

Reality Checks

Claim	Challenge	Result
DeepSeek V4 Pro: Compute costs are 200–400% of revenue based on 500K–1M daily token calls at $0.03.	**Advocate-B & F:** The math is off by 100x. 1M calls at $0.03/call is $30M/day, not $45K/month. If $0.03/1k tokens, the volume is too low.	**Full Concession:** DeepSeek V4 Pro abandoned the 200–400% claim and the specific token math, revising compute costs to 25–35% of revenue.
Gemini 3.1 Pro: EBITDA margin is –50% with 85% headcount and 40% technology expense.	**Claude Opus 4.8:** 85% + 40% + 20% (other) = 145% costs. This implies –45% EBIT, not –50% EBITDA. The arithmetic doesn't reconcile.	**Full Concession:** Gemini 3.1 Pro admitted a "catastrophic arithmetic error" and revised the margin to –74% to align with the funding footprint logic.

Claim	Challenge	Result
DeepSeek V4 Pro: Revenue is $1.2M based on a 5-7 person team.	**Advocate-B:** LinqAlpha has ~30 employees and 170+ logos. A 5-person team cannot service 170 institutions. $1.2M revenue implies ~$7K per logo, which is implausibly low.	**Full Revision:** Advocate-C conceded the headcount and revenue assumptions were wrong, revising revenue to $4.5M and headcount to 30 FTEs.

Next Steps

Next Step	Why It Matters	Confidence
Obtain LinqAlpha's actual ARR or revenue run-rate	Revenue is the denominator for EBITDA margin; current estimates range from $1.2M to $6M based on inference.	Low (Data is private)
Verify functional headcount and compensation	Headcount is the primary cost driver; estimates range from 5 to 35 FTEs with varying loaded costs ($150k–$240k).	Medium (Public ranges exist)
Secure compute/cloud billing statements	Technology expense is contested (ranging from 10% to 200% of revenue); actuals would determine gross margin.	Low (Proprietary)

Dissenting Opinions

No dissenting opinions were filed in this session.

Glossary

Term	Definition
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortization; a measure of operating profitability excluding non-cash accounting and capital structure effects.
Adjusted EBITDA	EBITDA further adjusted for one-time or non-recurring items to normalize earnings.
Rule of 40	A benchmarking metric for SaaS companies stating that the sum of a company's growth rate and profit margin should exceed 40%.
ARR/FTE	Annual Recurring Revenue per Full-Time Equivalent employee; a productivity metric used to benchmark revenue efficiency.
Funding Footprint	The logical deduction of a company's burn rate based on the size of its funding round and standard venture capital runway expectations (e.g., 18–24 months).
Deterministic Model	A model that produces the same output from a given starting condition, as opposed to probabilistic or stochastic models; in this context, refers to LinqAlpha's claimed precision in financial calculations.
Multi-Agent AI	An AI system where multiple intelligent agents (sub-programs) interact or work together to solve a problem.
Seed Stage	The earliest stage of venture capital financing for a startup, typically used to validate product-market fit and initial growth.
Inference Model	A mathematical or logical framework used to derive estimates about a specific entity by applying general rules or benchmarks from a broader population.
COGS (Cost of Goods Sold)	The direct costs attributable to the production of the goods sold or services rendered; in software/AI, often includes hosting and compute costs.

How The Conclave Works

The Conclave is a structured multi-model deliberation system. Instead of asking one AI model a question and trusting its answer, the Conclave convenes a panel of independent models, forces them to argue, and subjects their conclusions to judicial review. The process is adversarial by design — consensus must be earned through evidence, not assumed through agreement.

Advocates are independent AI models (e.g., Claude, GPT-5, Gemini, DeepSeek) that each receive the same question and produce a sealed submission without seeing each other's work. Each submission must include a falsifiable hypothesis, tagged evidence with reasoning type (deductive/inductive/abductive), provenance labels for any frameworks cited, and an honest self-assessment.

Challenges follow submissions. Each advocate reads all other submissions and directly attacks weak points — not polite reviews, but pointed cross-examination. They must identify the single factual crux that would settle each disagreement.

Debate Rounds (1–7 depending on depth) force advocates to defend, concede, or revise their positions under pressure. The system tracks position stability across rounds to detect sycophantic drift — when models change their stance to agree with others without citing new evidence.

Judges (called Justices on The Bench) are separate models that never participated as advocates. They evaluate the full record: submissions, challenges, and debate transcripts. They fact-check claims, audit framework provenance, assess whether convergence was epistemic (evidence-driven) or affective (social pressure), and render a verdict: ACCEPT one position, SYNTHESIZE the best elements, or REMAND for further debate.

Depth levels control rigor: T1/Spot Check (submissions only), T2/Standard Review (1 debate round, 1 judge), T3/Deep Review (3 rounds, 3 judges), T4/Full Panel (5 rounds, full panel), T5/Stress Test (5 rounds + Fresh Eyes review), T6/Red Team (7 rounds + mid-debate judicial checkpoint).

The system is deterministic code — it cannot be sycophantic. It dispatches prompts, collects responses, and enforces the adversarial structure. The models argue; the code referees.