Tribunal Session Summary

Session: conclave-20260604-011938-provide-a-best-in-b24a

Depth: T6

Advocates: 6 | Judges: 5 | Cost: $5.9349 | Time: 38m 12s

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

THE RULING: Rogo is operating at a deeply negative EBITDA of approximately **–$52M on an Adjusted (cash) basis** and **–$60M on a GAAP basis**, corresponding to margins of roughly **–120% and –139%** respectively.

The Prompt

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

The Ruling

Rogo is operating at a deeply negative EBITDA of approximately **–$52M on an Adjusted (cash) basis** and **–$60M on a GAAP basis**, corresponding to margins of roughly **–120% and –139%** respectively.

This conclusion is driven by a mechanical constraint: Rogo's verified revenue of $43.2M and revenue-per-employee of $148k imply a headcount of 292. Applying realistic NYC/AI-fintech compensation benchmarks (~$220k–$240k fully loaded) results in a personnel expense of ~$65M–$77M—which alone exceeds total revenue by 50–78%. When combined with AI-native cost of goods sold (compute and data licensing estimated at 30–37% of revenue) and other operating expenses, the company is burning roughly $5M–$6M in cash per month. The analysis relies entirely on industry benchmarks and the two verified operational metrics, as no public financial statements exist.

Core Analysis

The Core Thesis: Structural Losses from Headcount and Compute

Rogo is a classic hypergrowth, venture-subsidized AI startup. Its $2B valuation (46x revenue) and recent $160M Series D raise signal that investors are prioritizing growth over profitability. However, the unit economics at the current scale ($43.2M revenue) are structurally negative because the cost of acquiring and retaining the specialized talent required to build an AI-native financial product, combined with the high variable costs of AI inference and data licensing, significantly outpaces current revenue generation.

Decomposition and Unit Economics

The following breakdown synthesizes the consensus of the deliberation, anchored on the only verified facts: **Revenue ($43.2M)** and **Revenue Per Employee ($148k)**.

1. Revenue and Headcount Anchor

- **Revenue:** $43.2M (Verified).
- **Headcount:** ~292 employees. This is derived mathematically from the verified metrics ($43.2M / $148k = 291.9). This figure was treated as the analytical keystone after external trackers (Tracxn/PitchBook) showing 120–150 employees were rejected as lagging indicators of rapid hiring.

2. Operating Expenses

- **Headcount Cost:** $65.0M – $77.0M (150% – 178% of revenue).
  - *Mechanism:* A blended fully-loaded compensation rate of $220k–$240k per employee. This accounts for the mix of high-cost roles (ML engineers, ex-Wall Street domain experts) and lower-cost support functions in a NYC-based Series D company.
  - *Cash vs. SBC:* This total is split between cash payroll (~$195k–$230k) and Stock-Based Compensation (SBC) (~15% of cash payroll). Including SBC is necessary for GAAP EBITDA; excluding it yields Adjusted (cash) EBITDA.
  - *Why it matters:* Personnel is the dominant expense line. Even at the most conservative comp estimate ($220k), headcount cost ($64.2M) is nearly 50% higher than total revenue ($43.2M).
- **Technology / COGS:** $13.0M – $16.0M (30% – 37% of revenue).
  - *Mechanism:* Unlike traditional SaaS (often 80%+ gross margin), AI-native products face higher variable costs. This estimate includes:
  - **Compute & Inference:** ~25–30% of revenue. LLM API calls and vector database operations for retrieval-augmented generation (RAG) are expensive, though caching and model routing mitigate the upper end.
  - **Data Licensing:** ~5–10% of revenue. Serving bulge-bracket financial institutions requires premium, real-time data feeds (e.g., S&P, FactSet), which are a direct cost of revenue.
  - *Why it matters:* This compresses gross margin to 63–70%, leaving a smaller buffer to cover fixed personnel costs.
- **Other Operating Expenses:** $10.0M – $15.0M (23% – 35% of revenue).
  - *Mechanism:* Sales & Marketing (enterprise GTM, conferences), General & Administrative (legal, compliance, facilities), and software/tools.
  - *Why it matters:* While smaller than headcount, these costs are necessary for scaling a Series D company and add to the cash burn.

3. EBITDA Calculation

- **Adjusted EBITDA (Cash Basis):** –$50.0M to –$53.0M (–116% to –123%).
  - *Calculation:* Revenue ($43.2M) – Headcount Cash ($57M–$67M) – Tech/COGS ($13M–$16M) – Other OpEx ($10M–$15M).
  - *Significance:* This represents the actual cash burn rate, relevant for runway analysis. A burn of ~$50M/year implies a ~3-year runway on the recent $160M Series D, consistent with venture capital deployment norms.
- **GAAP EBITDA (Including SBC):** –$58.0M to –$63.0M (–134% to –146%).
  - *Calculation:* Adjusted EBITDA – SBC ($8M–$10M).
  - *Significance:* This represents the accounting operating loss, reflecting the full economic cost of the talent strategy.

Key Uncertainties and Sensitivities

- **Compensation Mix:** If Rogo has a higher proportion of offshore or lower-cost roles, the blended comp could drop to $200k, improving EBITDA by ~$12M. Conversely, a pure NYC onshore mix at $260k would worsen it by ~$12M.
- **Gross Margin:** If Rogo has achieved superior inference efficiency (e.g., 75% GM), EBITDA improves by ~$6M. If data licensing costs are higher (55% GM), EBITDA worsens by ~$6M.

Conclusion

Rogo is a high-growth entity where the cost of building and selling a complex AI product currently exceeds the revenue it generates. The "best in class" estimate is not a single precise number but a tight band derived from mechanical constraints (headcount) and industry benchmarks (comp and compute), converging on a deeply negative EBITDA of roughly **–$52M Adjusted** and **–$60M GAAP**.

How We Got Here

Scorecard

Judge	Grok 4	Claude Opus 4.7	DeepSeek V4 Pro	Magistral Medium	GPT-5	Gemini 3.1 Pro	Winner
Zhipu GLM-5.1	ADOPT	ADOPT	ADOPT	REJECT	ADOPT	ADOPT	SYNTHESIS
DeepCogito v2.1 671B	REJECT	REJECT	REJECT	REJECT	ACCEPT	REJECT	GPT-5
Zhipu GLM-4.7	ADOPT	ADOPT	ADOPT	REJECT	ADOPT	ADOPT	SYNTHESIS
MiniMax M2.7	ADOPT	ADOPT	ADOPT	REJECT	ADOPT	ADOPT	SYNTHESIS
GPT-OSS 120B	ADOPT	ADOPT	ADOPT	REJECT	ADOPT	ADOPT	SYNTHESIS

Convergence Assessment

The convergence was **epistemic (evidence-driven)**. Position shifts were triggered by specific, falsifiable challenges: arithmetic impossibilities (Magistral Medium), lack of sourcing for fabricated data (DeepSeek V4 Pro), and the introduction of superior benchmarks (SBC treatment, data licensing costs). The final consensus band (–$50M to –$63M) emerged from independent models applying the same mechanical constraints (292 heads) and similar industry priors, not from social pressure to agree.

Closing Statements

Grok 4: Maintained a deeply negative estimate of roughly –$50M (–116% margin). Acknowledged initial arithmetic inconsistencies and revised technology costs upward to 25–30% to reflect AI-native realities. Defended the $240k blended compensation as realistic for a mixed-role NYC organization. The final position is a conservative, bottom-up build anchored to the verified headcount.

Claude Opus 4.7: Shifted from an aggressive initial hypothesis (–$95M to –$115M) to a more grounded estimate of –$53M Adjusted / –$63M GAAP. Conceded that the initial headline was internally inconsistent and that the "burn-from-fundraise" logic was circular. Adopted a rigorous GAAP vs. Adjusted EBITDA distinction, separating cash payroll from SBC, which became a key component of the final synthesis.

DeepSeek V4 Pro: Underwent a major revision, abandoning an optimistic –$25M estimate based on fabricated unit economics. Conceded that compensation assumptions were too low and that the initial model contained arithmetic errors. Rebuilt the estimate to –$50M (–116% margin) using only peer benchmarks and the verified headcount anchor, effectively converging with the majority.

Magistral Medium: Abandoned the initial positive EBITDA thesis (+$13M) after the model was proven mathematically impossible (implying ~$52k/employee). Revised the position to a negative EBITDA of roughly –$45M to –$48M, acknowledging the high cost structure of AI startups. The final position aligned with the consensus but arrived late and with less methodological rigor than peers.

GPT-5: Presented a methodologically disciplined model centered on –$53M GAAP (–123% margin). Successfully argued for the inclusion of SBC in GAAP EBITDA and for the decomposition of COGS to include institutional data licensing. Retired an "offshore/high-GM" sensitivity case as infeasible for a regulated financial services vendor. The model's transparency and sensitivity framework were highly rated by judges.

Gemini 3.1 Pro: Defended the 292-headcount anchor against external data challenges and integrated critiques regarding SBC and data licensing to arrive at –$56M GAAP (–129% margin). Provided a clear distinction between GAAP and Adjusted EBITDA. The position underwent several revisions but ultimately provided a robust, conservative estimate that accounted for the high cost of NYC talent and compliance.

Key Moments

- **The Destruction of the Profitability Thesis (Magistral Medium):** Magistral Medium initially proposed a positive EBITDA of +$13M. This model was immediately dismantled by cross-examination because it implied a fully-loaded cost of ~$52,000 per employee for a NYC-based AI company—a mathematical impossibility given market rates. This forced the immediate abandonment of any positive-EBITDA scenario and established the floor for the debate as "deeply negative."
- **The Headcount Anchor vs. External Data (Gemini 3.1 Pro):** Early in the debate, some models relied on external headcount trackers (Tracxn/PitchBook) suggesting ~150 employees. Gemini 3.1 Pro rigorously defended the use of the verified Revenue Per Employee ($148k) metric to derive a headcount of 292. The argument that RPE is a company-reported metric (and thus more accurate than lagging trackers) prevailed, locking the 292-employee figure as the immutable denominator for all cost models.
- **The Collapse of Fabricated Precision (DeepSeek V4 Pro):** DeepSeek V4 Pro initially presented a highly specific "bottom-up" model claiming 400 clients and 20,000 inference requests per client to justify a -$25M loss. This was exposed as completely fabricated (no source existed for client count or usage).

Advocate-C conceded the fabrication and rebuilt the model using only benchmarks, shifting their estimate from -$25M to roughly -$50M, aligning with the consensus.

- **The SBC and GAAP Clarification (GPT-5 & B):** A significant methodological debate emerged regarding whether EBITDA should include Stock-Based Compensation (SBC). GPT-5 and Grok 4dvocate-B successfully argued that while "Adjusted EBITDA" (cash basis) is useful for runway, "GAAP EBITDA" must include SBC to reflect the true economic cost. This distinction clarified the ~$10M spread between some models and became a standard part of the final synthesis.
- **The Convergence on Compensation (Grok 4 & F):** Models initially ranged wildly on compensation ($170k to $285k). Through debate, the group converged on a blended range of $220k–$240k. Grok 4 successfully argued that a blended average must account for junior/support roles (preventing the $285k extreme), while Advocate-F successfully argued that NYC bulge-bracket requirements prevent the $170k extreme. This convergence tightened the final EBITDA band significantly.

Reality Checks

Claim	Challenge	Result
Magistral Medium: Rogo's EBITDA is approximately +$12.96M with a 30% margin.	**Advocate-F:** This implies a fully-loaded cost of $51,780 per employee for 292 heads, which is below minimum wage for NYC AI talent.	**Full Concession:** Magistral Medium abandoned the positive profitability thesis entirely.
DeepSeek V4 Pro: Rogo's technology expense is exactly $10M, based on 400 clients × 20k requests × $1.25/request.	**Gemini 3.1 Pro:** You have hallucinated a complete unit-economics model (client count, usage volume, cost per query) with absolutely no basis in the verified facts.	**Full Concession:** DeepSeek V4 Pro admitted the data was fabricated and withdrew the specific $10M claim.
Claude Opus 4.7: Rogo's EBITDA is –$95M to –$115M (headline), based on a $285k average loaded comp.	**Advocate-C:** Your own table shows a base case of –$75.5M; you literally do not produce any scenario that reaches your headline range. Also, $285k assumes homogeneity of compensation that ignores junior/support roles.	**Revision:** Advocate-B conceded the headline inconsistency and revised comp down to $220k–$240k, moving the estimate to ~–$62M.
Grok 4: Technology expense is only $5.5M (12.7% of revenue).	**Advocate-C:** This is invisible to AI-native economics; peers routinely see 25–50% compute costs. You offer no mechanism for a 4x cost advantage.	**Revision:** Advocate-A conceded the tech line was too low and revised it upward to 25–30% (~$11M–$13M).

Next Steps

Next Step	Why It Matters	Confidence
Verify actual headcount and geographic distribution	The 292 figure is derived from RPE; a significant offshore mix could lower the blended comp cost by $20k–$40k per head.	Medium (RPE is a verified metric, but geo mix is unknown).
Obtain actual gross margin data	The 30–37% COGS estimate is a benchmark; actual efficiency in inference or negotiated data rates could swing EBITDA by $5M–$10M.	Low (No public data exists; relies on peer analogues).
Confirm Stock-Based Compensation (SBC) policy	The 15% SBC load is a standard assumption; actual refresh rates and equity grants could vary materially.	Medium (Standard VC convention, but company-specific policies vary).

Dissenting Opinions

No dissenting opinions were filed in this session.

Glossary

Term	Definition
Adjusted EBITDA	A non-GAAP financial metric that adds back interest, taxes, depreciation, amortization, and often stock-based compensation (SBC) to net income. It is used to analyze cash operating performance.
AI-Native	A company whose products and services are fundamentally built around artificial intelligence technologies (e.g., LLMs) rather than integrating AI as an add-on feature.
Benchmark	A standard or point of reference against which things may be compared or assessed. In this context, industry averages for compensation and margins.
Blended Compensation	The average cost per employee across an entire organization, weighting high-cost roles (e.g., engineers) against lower-cost roles (e.g., support staff).
COGS (Cost of Goods Sold)	The direct costs attributable to the production of the goods sold by a company. For Rogo, this includes cloud compute, AI inference, and data licensing.
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortization. A measure of a company's operating performance.
EBITDA Margin	EBITDA expressed as a percentage of revenue.
Fully-Loaded Cost	The total cost of an employee including base salary, bonuses, benefits, payroll taxes, and often stock-based compensation.
GAAP (Generally Accepted Accounting Principles)	The standard framework of guidelines for financial accounting. GAAP EBITDA typically includes SBC as an expense.
Gross Margin	Revenue minus the cost of goods sold (COGS), expressed as a percentage of revenue.
Headcount	The total number of employees working for a company.
Inference	The process of a machine learning model making a prediction or generating output based on input data. A primary cost driver for AI products.
RAG (Retrieval-Augmented Generation)	an AI framework that retrieves data from external sources before generating a response, improving accuracy but adding infrastructure costs.
Revenue Per Employee (RPE)	A productivity metric calculated by dividing total revenue by total headcount.
SBC (Stock-Based Compensation)	The expense a company recognizes for granting stock options or restricted stock units to employees.
Series D	A fourth round of venture capital funding, typically for companies that have achieved significant scale and are preparing for an IPO.

Term	Definition
SaaS (Software as a Service)	A software licensing and delivery model in which software is licensed on a subscription basis and is centrally hosted.

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.