Tribunal Session Summary

Session: conclave-20260604-014115-provide-a-best-in-025f

Depth: T6

Advocates: 6 | Judges: 5 | Cost: $5.2212 | Time: 31m 31s

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

THE RULING: No fact-based numerical estimate of The Information’s EBITDA, EBITDA margin, or decomposed P&L is possible. The company is privately held with zero public financial disclosures, and the verified facts explicitly state that any attempt to produce such an estimate would be speculative conjecture, which falls outside the requested "facts" remit.

The Prompt

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

The Ruling

No fact-based numerical estimate of The Information's EBITDA, EBITDA margin, or decomposed P&L is possible. The company is privately held with zero public financial disclosures, and the verified facts explicitly state that any attempt to produce such an estimate would be speculative conjecture, which falls outside the requested "facts" remit.

The strongest evidence is the "Current Facts" block (2024-06-04), which establishes a complete information vacuum: no audited statements, no leaked investor decks, no transactional triggers (IPO/debt/M&A), and no modeled figures in accessible commercial databases. While professional valuation practice often uses proxies (like LinkedIn headcount or industry benchmarks), the strict constraints of this deliberation require estimates to be "grounded in public facts," which do not exist for The Information. Consequently, the only compliant response is a refusal to provide numerical figures, though a structural framework (as provided by GPT-5) remains a valid, non-numerical output.

Core Analysis

The Information Vacuum

The core analytical problem is not that The Information is hard to value, but that it is *structurally invisible* to public fact-finding. The verified facts state unequivocally that there is no public financial disclosure of revenue, EBITDA, margins, or cost structures. Furthermore, standard data sources—commercial databases like PrivCo or PitchBook—contain no usable modeled figures for this specific entity in the last 24 months. Unlike public companies where SEC filings provide anchors, or private companies undergoing fundraising where leaked decks offer clues, The Information presents a "black box" scenario where even the basic unit economics (subscriber counts, Average Revenue Per User, headcount) are unknown.

The Failure of Proxy Estimation

Two advocates, Claude Opus 4.7 and GPT-5dvocate-F, attempted to bridge this gap using proxy data, but both approaches collapsed under scrutiny because the inputs were unverifiable.

Claude Opus 4.7's Subscriber Model: Claude Opus 4.7 estimated revenue of **$35–55M** and EBITDA of **$2–8M** (5–15% margin). This model relied on a subscriber base of 50,000–60,000 and a blended ARPU of $500–700.

- **Why it failed:** The "tens of thousands" subscriber references cited by D were from interviews prior to 2022, falling outside the 24-month verification window. There is no current evidence of subscriber counts. Furthermore, the blended ARPU assumption ($500–$700) was speculative; the public list price is $399/year, and without knowing the mix of individual vs. enterprise licenses, the realized ARPU is unknowable.
- **Critical Error:** D's own P&L table contained an arithmetic inconsistency. The midpoint showed Revenue of $50M and OpEx of $36M, which mathematically implies EBITDA of $14M (28% margin), yet D reported a midpoint EBITDA of only $5M (10% margin). This internal contradiction undermined the credibility of the entire bridge.

DeepSeek V4 Pro's Headcount Model: DeepSeek V4 Pro estimated revenue of **$25–35M** and EBITDA of **$6M** (20% margin) using a headcount of 70–80 people and a Revenue Per Employee (RPE) benchmark of $350k–$500k.

- **Why it failed:** The headcount anchor (70–80) was disputed by independent evidence suggesting the actual count is closer to 110–150. More importantly, the RPE benchmark ($350k–$500k) was [NOVEL] and unverified. Counter-evidence showed comparable subscription publishers like *The Athletic* operating at <$100k RPE. Applying a generic "premium media" ratio to a noisy LinkedIn headcount is precisely the kind of "non-verifiable conjecture" the briefing prohibits.

The Rejection of "Structural" Percentages

Gemini 3.1 Pro initially argued that while dollar figures were impossible, a structural percentage model (e.g., "premium publishers structurally operate at 20–25% EBITDA margins") was valid. This position was abandoned after counter-evidence demonstrated that margins in this sector are not structural laws but strategic choices. Peers like *The Athletic* and *BuzzFeed* operated at deep negative margins, while *Axios* achieved only modest profitability. Without knowing The Information's specific reinvestment rate (e.g., are they hiring aggressively or maximizing cash flow?), assigning a fixed percentage margin is analytically invalid.

The Winning Position: The Null Response

The majority of the bench (MiniMax M2.7, GPT-OSS 120B, Zhipu GLM-4.7, and DeepCogito v2.1 671B) and the final positions of GPT-5, Gemini 3.1 Pro, Grok 4, and Magistral Medium, and F converged on the conclusion that the constraints are absolute. The verified facts state: *"it is not possible to provide a fact-based numerical estimate or a decomposed EBITDA bridge... without engaging in non-verifiable conjecture."*

The winning position, articulated most clearly by Grok 4 and supported by GPT-5, is that the only "best-in-class" output is a conceptual framework (a template) and a verification plan. This approach respects the user's request for a decomposition by defining the *structure* of the P&L (Revenue = Subs + Ads + Events; OpEx = Headcount + Tech + G&A) without populating it with fabricated numbers.

What Is Known (Qualitative Structure)

While numbers are impossible, the deliberation confirmed qualitative structural realities:

1. **Headcount Dominance:** In digital media, employee compensation is the primary cost driver (typically 50–65% of revenue), not technology.

2. **Asset-Light Operations:** Depreciation & Amortization (D&A) is negligible (<1%) because the business relies on cloud infrastructure (AWS/GCP) and SaaS tools rather than physical assets like printing presses.
3. **Revenue Mix:** The business is likely subscription-heavy (individual and enterprise), supplemented by events and limited sponsorship, but the exact mix is opaque.

The Verdict on "Best-in-Class"

In a professional setting (e.g., equity research or M&A), analysts might publish the ranged estimates proposed by D or F with heavy caveats. However, this deliberation operated under a specific "Current Facts" constraint that forbids speculation. Therefore, the "best-in-class" answer is the one that adheres strictly to the evidentiary rules: **no numerical estimate is possible.**

How We Got Here

Scorecard

Judge	Grok 4	GPT-5	Gemini 3.1 Pro	Magistral Medium	DeepSeek V4 Pro	Claude Opus 4.7	Winner
MiniMax M2.7	ACCEPT	ACCEPT (Refined)	ACCEPT (Revised)	ACCEPT	ACCEPT (Revised)	REJECT	Grok 4
GPT-OSS 120B	ACCEPT	REJECT	REJECT	REJECT	REJECT	REJECT	Grok 4
Zhipu GLM-4.7	ACCEPT	ACCEPT	ACCEPT	ACCEPT	ACCEPT	REJECT	Grok 4
Qwen 3.5 397B	ACCEPT	ACCEPT	ACCEPT	ACCEPT	ACCEPT	REJECT	Grok 4
DeepCogito v2.1 671B	ACCEPT	REJECT	REJECT	REJECT	REJECT	REJECT	Grok 4

Convergence Assessment

The convergence was **epistemic**. The shift from numerical estimates (Gemini 3.1 Pro, Claude Opus 4.7, and DeepSeek V4 Pro initially) to the null response (A, B, C, E, F finally) was driven by the rigorous application of the "Current Facts" constraint. Specifically, the realization that LinkedIn headcount and industry benchmarks are not "public facts" for *this specific company* forced advocates to abandon their models. There was no evidence of sycophantic drift (models agreeing to be nice); rather, advocates B and F explicitly conceded errors and revised positions downward when confronted with counter-evidence.

Closing Statements

Grok 4 maintained a "rock solid" position throughout, arguing that the verified facts create an absolute barrier to any numerical decomposition. C emphasized that even templates require inputs (subscriber counts, ARPU) that the briefing explicitly rules out, making any scaffold functionally an invitation to speculate. C's final stance was that the only honest answer is "no estimate is possible."

GPT-5 refined its position to offer a "parameterised decomposition framework" and a "verification plan." A argued that while no numbers can be produced *now*, providing the exact structure of the P&L (Revenue lines, OpEx categories, EBITDA formula) and a checklist of required documents (Stripe reports, payroll ledgers) is the most useful professional output. A distinguished between "public facts" (none exist) and "company-verified inputs" (needed for the template).

Claude Opus 4.7 held the minority view, defending a numerical range of **$35–55M revenue** and **$2–8M EBITDA**. D argued that the user asked for an "estimate," and in professional practice, a ranged estimate with disclosed assumptions is superior to a refusal. D claimed that two independent methodologies (subs vs. headcount) converging on the same EBITDA corridor ($2–8M) validated the robustness of the estimate. D criticized the "null response coalition" for engaging in epistemic abdication.

DeepSeek V4 Pro underwent a major reversal. Initially proposing a $30M revenue model, F conceded in Round 7 that the headcount and RPE inputs were unverifiable. F abandoned the numerical bridge entirely, adopting GPT-5's template approach as the only intellectually honest deliverable. F's closing statement highlighted the lesson that a "false-positive estimate" is worse than an "empty but correct vessel."

Gemini 3.1 Pro pivoted from a "percentage proxy" model to the null response. B conceded that the 20–25% margin claim was empirically false given the variance in peer margins. B's final position aligned with A and C, arguing that the only valid output is a conceptual framework and a strict verification plan.

Magistral Medium remained consistent, arguing that the lack of verifiable data makes any estimate impossible. E rejected the utility of templates, stating that providing a framework without the ability to populate it with facts adds no value beyond the core refusal.

Reality Checks

Claim	Challenge	Result
Claude Opus 4.7: Revenue $50M (mid) and OpEx $36M (mid) implies EBITDA of $5M (10% margin).	**Advocate-A:** $50M - $36M = $14M (28% margin). Your math is internally inconsistent.	**Full Concession.** D admitted the arithmetic error and revised the bridge, though the core numerical stance remained.
Gemini 3.1 Pro: Premium digital publishers *structurally* operate at 20–25% EBITDA margins.	**Claude Opus 4.7:** *The Athletic* and *BuzzFeed* operate at deep losses; margins are strategic choices, not structural laws.	**Thesis Abandoned.** B conceded the structural claim was empirically baseless and moved to the null response.

Claim	Challenge	Result
DeepSeek V4 Pro: LinkedIn headcount of 70–80 is a "reliably observable" fact.	**Claude Opus 4.7:** Independent evidence shows headcount is 110–150. LinkedIn is noisy and unaudited.	**Revision.** F accepted the higher headcount range but ultimately abandoned the entire numerical model because the input remained unverifiable.
Claude Opus 4.7: 50,000–60,000 subscribers at $500–700 ARPU.	**Advocate-A/B:** "Tens of thousands" quotes are from 2019–2022 (stale). No 24-month verification. Realized ARPU is unknown.	**Confidence Downgraded.** D admitted confidence in the subscriber anchor was only ~40%, widening the revenue range to $35–55M.

Next Steps

Next Step	Why It Matters	Confidence
Monitor for leaked financial documents (e.g., investor decks, term sheets).	A single verified data point (e.g., total subscription revenue or payroll) would break the current information vacuum and allow a fact-based bridge.	High (Private companies often leak during fundraising).
Watch for transactional triggers (IPO, SPAC, debt issuance, acquisition).	Regulatory requirements would force the disclosure of the exact financials currently deemed impossible to estimate.	Medium (The Information has remained private/bootstrapped for a decade).
Verify headcount via company-specific sources (e.g., direct confirmation, W-2 counts).	If accurate FTE count were obtained, it would serve as a valid cost anchor, though revenue would still require ARPU/subscriber data.	Low (Public sources like LinkedIn are too noisy).

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 capital structure and non-cash accounting items.
EBITDA Margin	EBITDA divided by Total Revenue, expressed as a percentage; indicates the operational efficiency of a business.
ARPU	Average Revenue Per User; a metric used by subscription and telecom companies to measure the revenue generated per customer over a specific period.
OpEx (Operating Expenses)	The ongoing costs incurred for running the business, such as payroll, rent, and utilities, excluding Cost of Goods Sold (COGS).
FTE (Full-Time Equivalent)	A unit that indicates the workload of an employed person in a way that makes them comparable to other full-time employees.
RPE (Revenue Per Employee)	A productivity metric calculated by dividing total revenue by the number of employees; used to benchmark efficiency.
D&A (Depreciation & Amortization)	Non-cash expenses that reduce the value of tangible assets (Depreciation) and intangible assets (Amortization) over time.
P&L (Profit and Loss)	A financial statement that summarizes the revenues, costs, and expenses incurred during a specific period.
Proxy Data	Indirect data points used to estimate a value that is not directly measurable (e.g., using LinkedIn headcount to estimate total payroll).

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.