

[OpenAI has 'full confidence' in CEO Sam Altman after investigation. reinstates him to board](#)

AI Evaluation - Claude 3 Opus

My input:

Hello. I would like to discuss a prediction market with you regarding the question related to the rationale and thought processes of the board of an organization. First, I'll give you the data about the market and its description. When I give you data, it will always be between leading and closing "" (triple quote) starting now. Here is the claim: ""Title:Info that the board pushed out Altman from OpenAI due to personal, non-AI issues, by mid 2024 "" Description w/details: ""This is about the board's reasoning, not the reality. If the board kicked him out for some kind of personal issue, it is YES, even if the board was wrong

i.e. they thought there was some kind of personal issue for conduct outside of work, or personal-relationship violations within work. Personal issue = anything, usually selfish in some way, which benefits himself.

something like "lying about business issues" doesn't count.

Things that would count as YES

- behaving improperly towards someone in/out of work would count.
- stealing candy bars from walgreens
- stealing money from work
- using his influence to try to get a friend/partner hired or someone fired

Things that would not count

- being mean at work in order to change the priorities of the company
- being very stubborn about things, or too opinionated, or keeping secrets at work to manipulate the board into making decisions he wants

The hard cases are where he is alleged to have used his authority to change company direction, but also in a way which is alleged that he would profit from personally.

If no such info has come out by deadline, resolves NO."

Here is a recent news article regarding some new information which just was released, today, March 8 2024.

""<See attached article>""

Finally, here are some comments from users. I'm particularly interested in the logic and validity of the interaction between Shump and Ernie. ""<See 2nd attached comment thread>""

Your job:

1. please, based on your knowledge, the full claim, and the article, please make your own judgment about the state of the claim, whether any resolution would be appropriate now, probabilities, reasons, etc.
2. THEN evaluate Ernie's response for logic, grammar, spelling. Is he making sense? Are the points good? Is it a consistent and coherent argument? Is it respectful? Is it missing anything obvious? And most importantly, how correct or wrong might it be?
3. Finally, and to the best of your abilities, please make other arguments for all other possible points of view and actions for this claim, with the constraint that each one may be honest, justified, and clear. Each one should try to understand its opponents possible disagreements, and address each idea, honestly, fairly, etc.

Finally, you have my permission to answer with a VERY long form response, fully detailing everything you may think is relevant, in rich detail.

Claude 3 Opus's Response

Based on the provided claim details, news article, and user comments, here is my analysis:

1. Judgment on the state of the claim: The recent news article provides significant new information relevant to the claim. Key points:
 - The outside investigation by law firm WilmerHale found that Altman's ouster in November was due to a "breakdown in the relationship and loss of trust" between him and the prior board.
 - However, the investigation also concluded that Altman's "conduct did not mandate removal" and that both he and Brockman remained the right leaders for the company. And CC
 - The current board chair stated that Altman's firing "did not arise out of concerns regarding product safety or security" or the company's finances, investor/customer statements, etc.

Based on this, it appears the claim would still resolve NO at this time. The article portrays the conflict as centering around a breakdown of trust and working relationship between Altman and the prior board, but does not provide any substantive evidence that the board pushed him out specifically due to personal, non-AI related issues or misconduct as defined in the claim.

The description is quite specific that things like being stubborn, manipulative, or keeping secrets to influence company direction would NOT qualify, even if done partially for personal benefit, as long as no clear personal malfeasance occurred. And the board chair's statements seem to directly counter any notion that improper personal conduct was the driving factor.

That said, the full details remain murky and the possibility can't be completely ruled out based on this one article. If substantive new information emerges in the coming months indicating the board's primary motivation was personal misconduct by Altman, a YES resolution could still be justified by the deadline. But as it stands now, I believe the claim would resolve NO, with a rough probability of 80% NO / 20% YES based on the current publicly available information. More disclosures from OpenAI or leaks from those involved would likely be needed to shift those odds significantly.

2. Evaluation of Ernie's response: Ernie's response is generally well-reasoned, logically structured, and argues its points coherently. The writing is clear with no glaring grammatical or spelling errors. Ernie methodically goes through each sentence of the claim description, evaluating how the new information relates to the criteria for a YES or NO resolution.

The core arguments Ernie makes are:

- The description implies that even if Altman acted improperly or manipulatively to change OpenAI's direction in ways that benefited him personally, this alone would NOT qualify as the "personal issues" needed for a YES resolution, as long as it was still ostensibly in service of his vision for the company. Truly personal misconduct outside of his leadership role would be required.
- The article portrays the conflict as a loss of trust and relationship breakdown between Altman and the board, but provides no evidence of the kind of personal malfeasance that would qualify under the claim criteria.
- The board chair's statements that the firing was not due to product safety, security, finances, communications, etc. counter the notion that clear personal misconduct was the driving factor.

These are fair points that align with my own read of the claim and article. Ernie's conclusion that the claim would likely resolve NO based on current evidence is reasonable and well-supported.

The one potential weak point is that Ernie is inferring quite a bit about Altman's underlying intentions, portraying his actions as a reasonable CEO trying to free OpenAI from restrictions to pursue his vision, even if his tactics were controversial. But the reality is we don't know Altman's true motivations, and it's conceivable that personal interests were a larger factor than Ernie assumes. The response could benefit from acknowledging more uncertainty on this point.

However, Ernie is correct that, regardless of Altman's intentions, the bar for a YES resolution is high and would require evidence of overt personal misconduct, not just controversial leadership. So while Ernie's portrayal of Altman is arguable, it doesn't undermine the overall conclusion that NO is likely the proper resolution unless new evidence emerges.

Overall, Ernie's response is respectful, argues its points honestly and persuasively, and shows strong reasoning on par with how I analyzed the claim. I believe Ernie's conclusion is

directionally correct based on the available facts, even if some of the details around Altman's motivations are more speculative.

3. Other possible arguments and perspectives:

Argument for YES: While the WilmerHale investigation and board chair's statements suggest Altman's ouster was not primarily due to clear personal misconduct, we still don't know the full story. The vague allusions to a "breakdown of trust" and Altman's conduct leave room for interpretation. It's possible there were concerns about personal improprieties that contributed to the board's loss of confidence in his leadership, even if they fell short of overt malfeasance. Altman's own apology for not handling a dispute with a former board member with enough "grace and care" hints at possible personal friction beyond just professional disagreement. If new leaks or disclosures over the coming months reveal that things like inappropriate relationships, harassment, self-dealing, or other personal issues were a significant factor in the board's decision, even if not the only factor, a YES resolution could still be justified by the deadline.

Argument for NO: The claim sets a high bar for a YES resolution, requiring fairly conclusive evidence that the board's decision was primarily motivated by non-AI personal issues, not just leadership controversies. The WilmerHale investigation appears to have been thorough, and nothing in its findings or the board's statements point to personal scandal as the driving concern. Even if Altman acted aggressively to change OpenAI's direction in ways that benefited his own vision, this alone would not qualify as disqualifying "personal issues" under the claim criteria. Altman's rapid reinstatement and the current board's "full confidence" in his leadership further suggest that any personal friction fell short of firing-level misconduct in their eyes. Barring surprise new revelations, the claim is likely to resolve NO based on what we know now.

Middle Ground: The reality is likely complex, with both professional and personal factors contributing to the breakdown between Altman and the prior board. While serious personal scandal doesn't appear to have been the primary cause based on current evidence, we can't rule out that concerns about Altman's personal conduct and character were part of the mix, even if they fell short of criminal activity or stark ethical breaches. The board's statements leave wiggle room, and we're still missing key details. It's quite possible there were elements of personal impropriety in the "loss of trust" that added to the tension, even if they weren't as egregious as the hypotheticals outlined in the claim description. New information could still shift the odds, and a YES resolution can't be definitively ruled out. But as it stands, the weight of available evidence points to NO as the most likely outcome.

In conclusion, based on the information provided, I believe the claim would currently resolve NO, as the article and WilmerHale investigation do not provide clear evidence that non-AI personal issues were the primary motivation for Altman's ouster, as required for a YES resolution. However, uncertainty remains and new disclosures in the coming months could conceivably still shift the balance to YES. Ernie's analysis arguing for an 80% NO / 20% YES outlook is logically sound and in line with my own assessment of the evidence. While room for debate remains, I believe Ernie's overall conclusion is directionally correct based on what we

know now. The situation appears more complex than strictly personal misconduct driving the board's decision. But only time and further revelations can clarify the full truth behind this leadership fracas at one of the world's most watched AI