



OpenAI's Fault Lines: Cracks in a Groundbreaking “Capped-Profit” Organization

Drew Collins-Burke*

Abstract

Influential artificial intelligence (AI) company OpenAI has a unique corporate structure featuring a non-profit/capped-profit (NP/CP) model. In the NP/CP model, a non-profit organization has control over a for-profit arm that offers financiers a fixed return based on their initial investment, as opposed to offering unlimited potential return. OpenAI's NP/CP structure is intended to reduce the negative impacts of shareholder capitalism on high-stakes artificial general intelligence (AGI) development projects. This paper evaluates OpenAI's organizational successes and failures, comparing its approach to the pitfalls many shareholder corporations fall into: excessive profit motives, lack of transparency, and negligence towards societal impacts. It also explores how OpenAI's structural features, such as investor profit caps and non-profit authority over the for-profit arm, have aided the company in avoiding some common issues with shareholder corporations. However, CEO Sam Altman's high-profile ousting and reinstatement, OpenAI's lack of open-source practices, and Microsoft's influence raise concerns about the overall efficacy of this structure. Through an analysis of OpenAI's structure, actions, and public statements, this paper investigates the hybrid NP/CP model's potential for mitigating the negative impacts of shareholder capitalism on responsible AGI development, highlighting its successes and limitations. The paper concludes that OpenAI's ability to develop AGI safely within this organization model is possible but uncertain.

Investing in OpenAI Global, LLC is a high-risk investment

Investors could lose their capital contribution and not see any return

It would be wise to view any investment in OpenAI Global, LLC in the spirit of a donation, with the understanding that it may be difficult to know what role money will play in a post-AGI world

—OpenAI's Operating Agreement

1. Introduction

OpenAI was founded as a non-profit in 2015 by current CEO Sam Altman, Elon Musk, Ilya Sutskever, Greg Brockman, and several elite research engineers. It received funding from

notable Silicon Valley personalities, like LinkedIn co-founder Reid Hoffman, influential startup incubator YCombinator co-founder Jessica Livingston, and PayPal co-founder Peter Thiel. It also received donations from Amazon Web Services; IT consulting firm Infosys; and YCombinator's charitable arm, YC Research. Despite being backed by billionaires and massively powerful companies, OpenAI asserts that it tried from the start to avoid being controlled by financial obligations: “Our goal is to advance digital intelligence in the way that is most likely to benefit humanity as a whole, unconstrained by a need to generate financial return” (Brockman et al.).

In 2018, OpenAI established a for-profit arm

*Drew Collins-Burke (drewcb2020@gmail.com) is a recent summa cum laude Honors College graduate with previous work for a Brookings-affiliated institution and multi-year contributions as a research assistant to a Carnegie Fellow. He is interested in AI policy, political polarization, men's issues, and loves to hike, play tennis, and listen to classic albums.

of the company. The non-profit legally controls the for-profit arm, subjecting the for-profit arm to the non-profit's obligation to benefit humanity. Ninety-nine percent of the company's personnel is currently employed by the for-profit arm of the organization, with OpenAI valued at 80 billion dollars and Microsoft owning 49 percent of the for-profit's shares (Andersen, "Does Sam Altman Know What He's Creating?"; Metz and Mickle). OpenAI asserts that the for-profit arm can only offer limited financial returns for investors, called a capped profit. Reportedly, the cap on returns is 100 times an initial investment (Wiggers). The for-profit arm has day-to-day control over all commercial endeavors OpenAI has promoted, such as ChatGPT. However, the non-profit can intervene should voting members perceive a violation of the organization's ethical obligations. The non-profit/capped-profit (NP/CP) model presents a unique and novel organizational structure with distinct dynamics from the more common shareholder corporation.

Some have criticized OpenAI's 2018 structural changes and subsequent organizational behavior. Elon Musk is suing OpenAI for breaking from its original mission (Nidumolu et al.). Former OpenAI researcher Dario Amodei has criticized the company's approach to safety and left to form rival AI development organization Anthropic (Fortune Editors). Former OpenAI board member Helen Toner praised the more cautious approach of Anthropic but condemned OpenAI's release of ChatGPT, citing the follow-on effect wherein other competing firms launched products to the market without proper safety testing (Toner et al.).

OpenAI has maintained that its non-profit's mission still governs the capped-profit arm and that the company has not fallen prey to capitalistic incentives. The warning OpenAI gives investors in its Operating Agreement demonstrates an earnest desire to avoid appearing as a traditional shareholder corporation. OpenAI's organizational structure demonstrates a recognition—whether performative or genuine—of the importance of preventing financial incentives from controlling

immensely powerful AI systems. Still, some developments have been concerning, such as OpenAI's lack of transparency or open-source code. Altman's brief ouster in late 2023 served as a stress test, demonstrating issues with the NP/CP structure's ability to commercialize its products without, in the eyes of some safety-focused board members, violating the non-profit's mission. These issues put the efficacy of the NP/CP structure for creating safe AGI into doubt; nonetheless, the NP/CP structure has some advantages that should be considered for other organizations.

Through a case study of OpenAI, which includes comparison between OpenAI and typical shareholder corporations, this paper aims to critically analyze the effectiveness of the NP/CP structure in developing safe AGI.

2. Methodology

This research was largely inspired by Gerald Davis's *Managed by the Markets: How Finance Reshaped America*, which presents theories and critiques of shareholder capitalism.

The present paper conducts a case study of OpenAI's organizational structure, evaluating theories describing disadvantages to shareholder corporations and analyzing OpenAI's attempts to mitigate those negative externalities. The case study involves an in-depth examination of the company's non-profit/capped-profit structure, actions, public statements, and recent events, such as the CEO's ousting and reinstatement. Qualitative data for the case study were collected from various sources, including news articles, company blogs, and interviews with OpenAI representatives.

3. Problems with Shareholder Capitalism OpenAI Aims to Avoid

Gerald Davis's *Managed by the Markets* demonstrates how financial incentives have led

modern corporations to become almost entirely influenced by profit, often disregarding the societal harms their businesses create. According to Davis, corporations used to be more amicable to societal goals during the era of managerial capitalism in the 1920s through 1980s (63). However, during the current era of shareholder capitalism, corporations tend to act according to financial incentives alone, operating less as social institutions and more as contractual nexuses associated with emotionally cold and economically rational behavior (63). These shareholder corporations are often heavily influenced by banks. Many have strongly disliked these corporations since their inception: “faceless monopolies were bad enough, but faceless monopolies controlled by a small handful of bankers in New York were worse still” (Davis 68).

Examples of the harm these dynamics create are numerous. Consulting firm McKinsey & Company advised Purdue Pharma to aggressively sell addictive opioids, heavily contributing to America’s current opioid crisis (Forsythe and Bogdanich). Enron and Shell knowingly engaged in behavior that led to major climate harm (Franta). Amazon subjected warehouse workers to conditions that led to injuries in more than half of their laborers over a three-year period (Day and Bloomberg). Social media companies like Meta, the owner of Facebook and Instagram, knowingly promote addictive apps with harmful mental health outcomes, to the extent that the U.S. Surgeon General and American Psychological Association have issued advisories for teenagers that caution against social media use (Katella). Despite this plethora of antisocial actions, these corporations remain dominant.

Shareholder corporations have been successful despite receiving heavy criticism for causing societal harm. Their success can largely be attributed to their economically effective behaviors and strategies. Market valuation has become the sole factor driving strategic decisions for the firm (Davis 93), which often prioritize increased profits for shareholders. Stock options

also reward CEOs for increasing the company’s value in a given quarter (Davis 87). With these motivations, company decision-makers often make choices that disregard morality for the sake of profit. Overall, the exclusive focus on financial incentives for corporate actions represents one of the most prominent and harmful traits of shareholder capitalism that OpenAI has tried to avoid.

Perhaps the most culturally prominent manifestation of the principal role of financial incentives in organizational behavior can be seen on Wall Street. Anthropologist Karen Ho’s *Liquidated: An Ethnography of Wall Street* details the culture, tendencies, and influence of Wall Street bankers, arguing that Wall Street’s intense fixation on maximizing shareholder value harms society. For instance, when banks acquire public companies, they institute organizational changes—layoffs, cutting benefits, and creating programs—to increase short-term profits and raise stock value. The threat of leveraged buyouts has even affected publicly traded firms once considered too big or stable for takeovers (Ho 144); these firms must then react and adjust their actions to account for Wall Street’s desires, even if no explicit declaration of a leveraged buyout has been made (Ho 145). Thus, Wall Street’s orientation towards shareholder value above all else leads companies to try to increase the company’s stock value at the expense of workers.

CEO Sam Altman’s perspective presents a marked aversion to Wall Street financing. In an interview for the Atlantic, Altman states, “you should never hand over control of your company to cokeheads on Wall Street,’ ...but he will otherwise raise ‘whatever it takes’ for [OpenAI] to succeed at its mission” (Andersen, “Does Sam Altman Know What He’s Creating?”). Altman’s aversion to collaborating with Wall Street bankers displays an ideological divergence from the typical shareholder corporation CEO. His commitment to the cause is perhaps strengthened by the fact that he has little financial stake in OpenAI, an abnormal position compared to most other CEOs

(Massa and Galpotthawela; Davis 86). Altman's attitude displays a determination to raise capital without Wall Street involvement.

In the context of AI, prioritizing shareholder value above societal welfare and long-term research and development goals could prove very harmful. Firms could use exploitative, profit-oriented AI systems to increase shareholder value by automating the complex pattern-recognition of consumer behavior—a practice often referred to as surveillance capitalism—and creating even more effective and addictive social media algorithms (Jones). This threat raises an important question: *Has OpenAI's hybrid structure been effective at ensuring the organization avoids recklessly adhering to financial incentives?*

4. OpenAI's Alternative Form: A Non-Profit and "Capped Profit" Hybrid

OpenAI has a unique company philosophy. It maintains that its work may lead to the creation of highly powerful AI systems known as artificial general intelligence (AGI), described by OpenAI as "a highly autonomous system that outperforms humans at most economically valuable work" (Our Structure). CEO Sam Altman seems to believe that AGI is such a significant technological development that society needs to be slowly introduced to less powerful AI technologies first to avoid massive social upheaval. He also believes that AGI will fundamentally alter the nature of our world, societies, and day-to-day lives, although he expresses doubt about what that theoretical future will look like (Andersen, "Does Sam Altman Know What He's Creating?").

As discussed earlier, OpenAI is a non-profit/capped-profit (NP/CP) organization. A capped-profit organization is a modified for-profit company that limits the maximum financial return investors can receive in order to create a balance between commercial viability and careful discretion (Our Structure). OpenAI is located in San Francisco and had 770 employees as of

November 2023 (Metz et al.). Its primary product, ChatGPT, is a widely used service, with 180.5 million unique visitors in August 2023 (Tong).

OpenAI's unique structure, extremely influential product, and stated mission of benefiting humanity make it a relevant organization to discuss when considering alternatives to the shareholder corporation. OpenAI has succeeded in creating AI safeguards and refused to monetize its technology as fully as possible, with measures against corporations owning AGI systems standing as a fundamental part of its structure. However, the company's movement away from transparent and open practices and recent tensions culminating in the ousting and subsequent reinstatement of CEO Sam Altman in late 2023 appear to show cracks within the organizational structure (Duhigg).

Nonetheless, OpenAI's NP/CP structure may qualify as an alternative to the shareholder corporation because it mediates the need for profit and capital by limiting shareholder returns. Initially, OpenAI was strictly a 501(c)(3) organization—a tax-exempt non-profit group. However, its board began to realize that a non-profit structure could not generate the capital required to fund the costs of creating powerful, novel AI systems that need enormous computational power and elite talent (Our Structure). Thus, the NP/CP structure emerged. Under this structure, investors have no reason to push OpenAI to monetize their products further once their maximum potential profits have been attained. Additionally, the non-profit's authority over the capped-profit arm—legally subjecting the capped-profit arm to the non-profit's obligation to benefit humanity—displays that OpenAI's organizational approach provides legal obligations distinct from the shareholder corporation.

OpenAI's unique organizational philosophy—that AGI could upend the world economy and potentially render many current societal structures obsolete—has also impacted its organizational structure and behavior. The

OpenAI website notes that the non-profit board will determine when the company has achieved AGI and that AGI is not included in intellectual property licenses and commercial agreements with Microsoft. Under the current NP/CP structure, OpenAI's overall goal—to create enormously powerful AI systems—cannot be controlled by any for-profit organization. Additionally, OpenAI's operating statement warns potential investors against expecting financial returns from their investments, claiming that the role of money after the completion of AGI is uncertain (Our Structure). This abnormal warning shows how OpenAI's capped-profit arm has a unique attitude towards investors that explicitly denies for-profit interests control of AGI. This behavior is far different from that of a typical shareholder corporation, which would be unlikely to discourage interested investors from expecting financial returns. OpenAI's organizational actions and attitudes towards financial backers show that the NP/CP structure can lead to organizational behavior that mitigates the harmful desire for "profit above all else" seen in shareholder corporations.

Cynical perspectives may question the genuineness of Altman and OpenAI's dedication to creating safe and beneficial AGI. After all, OpenAI has faced numerous lawsuits from parties like *The New York Times*, which claimed that the company's training of its AI on their work without permission violated copyright law. OpenAI received regulatory scrutiny from the US and EU and is under investigation by the Securities and Exchange Commission (Satariano et al.). Numerous industry leaders and academics, like Anthropic CEO Dario Amodei and former OpenAI board member Helen Toner, have also raised concerns about the safety and ethics of creating AGI. Further, the company's goal of achieving AGI is likely to displace workers. OpenAI has also strayed from its initial mission of remaining open-source and is currently being sued by co-founder and former board member Elon Musk for its closed-source products and ties with Microsoft

(Satariano et al.).

In particular, the current lack of open-source AI from OpenAI presents an important example of how its organizational behavior has changed after the addition of the for-profit arm. According to Musk, the organization initially intended to provide open-source code for its products. OpenAI has countered that it never intended its products to be open-source. OpenAI asserts that the "Open" in its name refers to its transparency practices in research and distributing AI's benefits (Metz). OpenAI's alleged shift away from its original open-source ambitions sparked controversy, with critics like Musk claiming it indicates the organization's adherence to financial incentives. Some have suggested that restricting access to AI's source code puts the potential power of AGI in the hands of corporations who will abuse it.

Defenders of OpenAI have claimed that keeping the source code private is justified because it limits the misuse of AI and increases the accountability of organizations for harmful AI usage. For instance, a technology ethics advocate at the Center for Humane Technology, Aza Raskin, demonstrated that Snapchat's AI feature could be manipulated for harmful ends. He found that the AI would tell a user (whom the AI believed to be 13) how to set the scene for a romantic getaway with a 31-year-old, recommending that the 13-year-old user set the mood for their first time having sex by lighting candles (Harris and Raskin). Snapchat quickly added safeguards to vary the AI's responses more appropriately based on age (Hutchinson).

Raskin's fellow Center for Humane Technology advocate, Tristan Harris, has pointed out that these public companies are afraid of financial consequences for negative AI behaviors and will curtail harmful behaviors that impact their companies' reputations. However, with the leakage of open-source models like Meta's LLaMA, Harris pointed out that governance and accountability for AI responses become more difficult, since anyone could modify AI for harmful ends (Lloyd). If ChatGPT was open-

source, users could create highly flexible and powerful AI systems from OpenAI's advanced code without any guardrails and limited accountability from OpenAI. Child abusers could utilize AI to help them craft persuasive messages to victims, and terrorists could use AI for advice on optimal attacks. These scenarios demonstrate a plausible way in which open-source AI could be exploited. Having publicly accountable organizations attached to AI responses prevents some harmful uses, curbing anonymous individuals from using AI for any purpose they desire.

Investigative reporter Karen Hao spent three days at the OpenAI office and conducted over thirty interviews with relevant employees and experts. Hao concluded that "there is a misalignment between what the company publicly espouses and how it operates behind closed doors. Over time, it has allowed a fierce competitiveness and mounting pressure for ever more funding to erode its founding ideals of transparency, openness, and collaboration." Hao's assertion provides a worrying perspective on OpenAI's organizational structure. Hao's description sounds more like one of a shareholder corporation that follows financial incentives than holding with OpenAI's stated mission and obligation as a non-profit to benefit society (OpenAI, "About"). Hao's claims that the company has become competitive and opaque support the idea that OpenAI has strayed from its initial mission.

Still, some of OpenAI's organizational behavior has demonstrated continued concerns for AI safety. OpenAI has employed industry experts to conduct safety stress-testing on every AI model it has released, automatically registering known child abuse imagery to the National Center for Missing and Exploited Children and furnishing broader systems that aim to monitor harmful uses of its AI (Our Approach to AI Safety). One benefit of OpenAI's capped-profit arm is that it can distribute AI systems for people to interact, earning revenue while also enforcing guardrails to prevent harmful uses.

Significant revenue is necessary to pay for OpenAI's intensive expenses, including its team of elite talent and the large amounts of cutting-edge chips that power its AI services. These costs can be exorbitant. The most common salary range for engineering roles, as listed on OpenAI's website, is \$200,000 to \$370,000—a wage that bonuses can reportedly increase nearly threefold. Microsoft and OpenAI are reportedly planning to construct a \$100 billion data center, and ChatGPT reportedly costs \$700,000 a day to run (Constantz; Chervek; Elimian).

According to market research firm Sacra, 25 percent of OpenAI's revenue will be given to employees and early investors until they reach their profit cap, and 75 percent is expected to go to Microsoft until their principal investment of \$13 billion is recouped. Afterward, 49 percent of their revenue will go to early investors and employees. Microsoft will receive 50 percent of OpenAI profits until they receive an additional \$92 billion, and the remaining revenue will be given to the non-profit arm. Should that \$92 billion cap be reached, OpenAI will receive all further equity and 100 percent of profits under their current agreements ("OpenAI Revenue, Valuation & Growth Rate"). A non-profit alone could not legally have investors expecting returns; OpenAI's massive costs reinforce the idea that the capped-profit arm will be necessary to attain the capital and resources needed to create AGI. Still, Hao's claim that the company has strayed from its initial goal of transparency and become competitive, money-focused, and secretive creates worrying parallels to the pathologies of the shareholder corporation.

Concerns about the secrecy of OpenAI increased on November 17, 2023, when Altman was suddenly removed as CEO of OpenAI. In a vaguely worded press release, the board claimed that he "was not consistently candid in his communications" (Kerr). On November 21, 2023, Altman was reinstated as CEO after 95 percent of OpenAI employees signed a letter threatening to quit if Altman was not reinstated (Carter). As a non-profit, the board is legally obligated to hold

the organization to its mission of benefiting humanity and retains the ultimate say over hiring and firing employees (“Board Roles and Responsibilities”). Exactly what happened and why generally remains mysterious. Altman’s ouster had little to no warning—even close partner Microsoft was given only a few minutes of notice (Weise). Some details have been made public, however, and Altman’s ouster and subsequent reinstatement can be viewed as a stress test of OpenAI’s organizational structure.

Some possible malpractices can be ruled out as the reason for Altman’s ouster. OpenAI COO Brad Lightcap stated in an internal message: “We can say definitively that the board’s decision was not made in response to malfeasance or anything related to our financial, business, safety or security/privacy practices...This was a breakdown in communication between Sam [Altman] and the board” (Klein). In other words, Altman’s ouster was due to relational distress rather than the violation of law or company rules.

Altman’s removal only occurred because the non-profit was capable of ousting him for reasons beyond malpractice, incompetence, or negligence of his fiduciary obligations. In shareholder corporations, CEOs are generally only fired for cause—incompetence, insubordination, poor attendance, criminal behavior, harassment, or physical violence—if there is significant, legally admissible evidence of misconduct (“Acceptable Reasons for Termination”; Album). The board did not specify whether Altman’s removal was a for-cause termination. Their reasoning for his firing was a vague “breakdown in communication,” implying that their justification was either incompetence or insubordination. The board presented no evidence at the time of firing, and a third-party review of Altman’s behavior later cleared him of any wrongdoing (Metz and Ghaffary). Shareholder corporation CEOs are typically only fired without cause if they disappoint shareholders. The board’s statement clears Altman of disappointing shareholders by saying that his firing was not due to his financial

or business practices (Wiersema). Because of the non-profit board’s broad influence over the for-profit arm, the corporate norm of only firing CEOs for misconduct or disappointing shareholders was not sufficient to prevent the ousting of Altman. Arguably, the lack of rationale provided demonstrates a flaw in the NP/CP structure, since tense personal relations led to a high-profile event that ultimately was ineffectual at increasing OpenAI’s attention towards safety.

Two important, safety-focused former board members seem key figures in this event: Toner—an academic at the Georgetown Centre for Security and Emerging Technology—and Ilya Sutskever—OpenAI’s Chief Scientist. It appears likely that Altman’s relationship with these two deteriorated.

Toner co-authored an academic paper that praised rival AI company Anthropic—founded by former OpenAI researcher Dario Amodei—for its highly cautious approach to AI. Toner criticized OpenAI’s release of ChatGPT for creating race-to-the-bottom dynamics, where corporations strive to achieve a goal first by using progressively more harmful tactics to outcompete one another. In the case of ChatGPT, companies like Google rushed competitive AI products to market before proper safety tests could be completed (Toner et al. 30). Altman reportedly attempted to push Toner out due to this criticism (Metz et al.; Duhigg), and Toner has since been removed from the board. According to the limited information available, Toner’s publication of this critique likely led to her removal from the board, painting a worrying picture of OpenAI’s transparency and responses to criticism.

Sutskever, meanwhile, initially supported Altman’s ouster publicly, even delivering the news to Altman on a Zoom call November 17, 2023—a Friday. Sutskever’s initial support of the ouster was likely due to his concerns about AI alignment and the accelerated pace of AI products that Altman was encouraging. Ross Andersen, writing for *The Atlantic*, reported that Sutskever was concerned by Altman’s desire to ship out products at a rapid rate and fundraise with concerning

parties, such as new computer-chip production firms or even oppressive Middle Eastern governments. The same article also reported that Sutskever deeply fears AGI harming society because of the potential for corporations to misalign AGI's behavior for their own gain. Andersen based his assertion on his personal interactions with Sutskever and insider reports; due to the limited information available, his hypotheses should be taken as speculation ("The Man Who Tried to Overthrow Sam Altman").

Abruptly, Sutskever called for Altman's reinstatement the following Monday—November 20, 2023—stating, "I deeply regret my participation in the board's actions" (McMillan). Sutskever even signed the letter calling for Altman's reinstatement (Mann). Sutskever likely changed his mind due to pressure from OpenAI employees and Microsoft's initiative to briefly hire Altman and offer all OpenAI employees positions. Microsoft's move turned Sutskever's ousting of Altman against his intent to reduce the influence of large corporations on AGI.

Microsoft CEO Satya Nadella was reportedly furious at Altman's unexpected ouster (Knight). Microsoft flexed its power by briefly hiring Altman and offering OpenAI employees jobs, eventually leveraging a nonvoting board seat with OpenAI. Watchdog groups have raised concerns that Altman's reinstatement after Microsoft internally and externally pressured OpenAI sends a message that Microsoft holds the ultimate control over OpenAI. These groups claim that the board was merely doing its job as intended and assert that Altman's desire to monetize AI products at an increasingly aggressive pace has been enabled by his reinstatement (Duhigg). New, high-profile board members have now been installed, including Larry Summers, former Secretary of the Treasury, and Bret Taylor, Chairman of Twitter. With Toner and Sutskever's removal, the board lost two safety-focused members. These events diminished the influence of the sect of AI researchers who want to prioritize safety above monetization, raising significant concerns about

the efficacy of OpenAI's organizational structure for upholding its founding principles.

The ouster was a total failure for safety-concerned board members like Toner and Sutskever; instead, Altman, Microsoft, and AI commercialization came out stronger. This event serves as key evidence in evaluating whether OpenAI's structure has successfully maintained its goal of generating safe and beneficial AGI. The failure to oust Altman shows that the non-profit board could not overcome the incentives of commercialization and failed to increase AI safety. Although other aspects of OpenAI's behavior, such as its warning to investors and limitations to returns, remain demonstrable benefits of the structure, the idea that the board can declare the organization has achieved AGI and release it from commercial obligations seems to be a less effective safeguard after the removal of safety-focused board members. Utilizing this ouster as a case study and stress test for the NP/CP structure shows that the capped-profit arm's goals of creating value for investors may have greater influence over OpenAI than the non-profit. Thus, the NP/CP structure's efficacy is in doubt, although its advantages over the shareholder corporation should still be considered.

5. OpenAI's Organizational Structure's Successes and Failures

OpenAI's organizational structure has enabled it to be successful from an economic and innovation perspective, but whether it is currently having a positive societal effect and will succeed in making an AGI that benefits all of humanity remains questionable. OpenAI's success in avoiding the typical pitfalls of the shareholder corporation is mixed. Groups interested in using OpenAI's non-profit/capped-profit (NP/CP) model should approach it with caution. Some of OpenAI's actions, such as the warning to shareholders, capped profits for investors, board control of AGI, and Altman's statements on avoiding control from

Wall Street, demonstrate desires—likely genuine—to avoid the typical pathologies of shareholder corporations. However, the copyright lawsuits, fear of AGI's consequences, concerns of Microsoft's influence, diversion from the original mission, and ouster of Altman display that the NP/CP structure has cracks.

OpenAI's NP/CP structure should neither be seen as a perfect substitution for the shareholder corporation nor ruled out as a potential option. Its structure has enabled it to act differently from shareholder corporations; the warning to investors, the ability of the board to declare AGI, and capped profits for investors all represent legitimate, significant differences in OpenAI's organizational structure. It demonstrates that OpenAI is culturally unwilling and structurally unable to profit off its work at the expense of its morals. OpenAI could have structured its secondary arm to make as much money as possible at the expense of its mission, but instead, it purposely limited its financial incentives—an important deviation from the shareholder corporations that often act only in advancement of their goal of creating as much shareholder value as possible.

Other parties considering ways to avoid the pathologies of shareholder capitalism should consider the NP/CP structure while keeping its flaws and failures in mind, because OpenAI's behavior has some worrying parallels to the shareholder corporation. The culture of secrecy and competition that reporter Karen Hao highlights seems strikingly similar to a company trying to beat other firms to release a product. Although its step away from open-source code may be justifiable from a safety perspective, this decision still insulates the inner workings of OpenAI's products, as opposed to democratizing AI. Most concerningly, Altman's reinstatement can be read as a victory for the commercialization of AI and a loss for those advocating a slower, safety-oriented approach to AI's integration into society.

The successes of OpenAI's NP/CP structure

could be incorporated into other organizational frameworks through capped profits for investors, offering an incentive to invest without creating an incentive to squeeze the company for profits. Potential options for improving this structure might include adding additional measures to prevent something like Altman's ouster and reinstatement from recurring. Specifically, it may be beneficial for the non-profit board to hold an employee vote before making decisions on hiring or firing C-level executives.

In the end, evaluating OpenAI's organizational structure presents areas of hope and gloom for the safe handling of powerful AI systems, including AGI. The NP/CP structure should be considered by other corporations to curb an exclusive focus on financial incentives but should not be qualified as a total fix of the pathologies of the shareholder corporation. Although it has some demonstrable benefits, the ability of the NP/CP structure to safely handle AGI remains in doubt.

Acknowledgements

I would like to acknowledge Dr. Gerald Berk for teaching a fascinating class that helped produce this paper. I would also like to thank Dr. Neil O'Brian and Dr. Trond Jacobsen for helping me refine my research and argumentation skills.

Works Cited

- "Acceptable Reasons for Termination." *The Hartford*, <https://www.thehartford.com/business-insurance/strategy/employee-termination/valid-reasons>. Accessed 14 May 2024.
- Album, Michael. "Terminating a CEO for Cause." *Employee Benefits & Executive Compensation Blog*, 20 Aug. 2020, <https://www.erisapracticecenter.com/2020/08/terminating-a-ceo-for-cause/>. Accessed 14 May 2024.
- Andersen, Ross. "Does Sam Altman Know What He's Creating?" *The Atlantic*, 24 July 2023, <https://www.theatlantic.com/magazine/archive/2023/09/sam-altman-openai-chatgpt-gpt-4/674764/>. Accessed 1 Mar. 2024.
- Andersen, Ross. "The Man Who Tried to Overthrow Sam Altman." *The Atlantic*, 21 Nov. 2023, <https://www.theatlantic.com/technology/archive/2023/11/openai-ilya-sutskever-sam-altman-fired/676072/>. Accessed 22 Mar. 2024.

- B Corp Certification Demonstrates a Company's Entire Social and Environmental Impact.* <https://www.bcorporation.net/en-us/certification/>. Accessed 28 May 2024.
- "Board Roles and Responsibilities." *National Council of Nonprofits*, <https://www.councilofnonprofits.org/running-nonprofit/governance-leadership/board-roles-and-responsibilities>. Accessed 28 May 2024.
- Brockman, Greg, et al. *Introducing OpenAI*. 11 Dec. 2015, <https://openai.com/blog/introducing-openai>. Accessed 20 Mar. 2024.
- Carter, Tom. "95% of OpenAI Workers Have Threatened to Quit If Sam Altman Is Not Reinstated as CEO." *Insider*, 20 Nov. 2023, <https://www.businessinsider.com/openai-workers-threaten-to-quit-over-sam-altman-firing-2023-11>. Accessed 21 Mar. 2024.
- Chervek, Emma. "Microsoft, OpenAI Plan \$100B Stargate AI Data Center, Eases Reliance on Nvidia." *SDxCentral*, 1 Apr. 2024, <https://www.sdxcentral.com/articles/news/microsoft-openai-plan-100b-stargate-ai-data-center-eases-reliance-on-nvidia/2024/04/>. Accessed 14 May 2024.
- Constantz, Jo. "OpenAI Engineers Earning \$800,000 a Year Turn Rare Skillset Into Leverage." *Yahoo Finance*, 22 Nov. 2023, <https://finance.yahoo.com/news/openai-engineers-earning-800-000-183139353.html>. Accessed 14 May 2024.
- Davis, Gerald F. *Managed by the Markets: How Finance Re-Shaped America*. Oxford University Press, 2009.
- Day, Matt and Bloomberg. "Half of Amazon's Warehouse Workers Are Injured after Just 3 Years, According to Study That Revealed Far More 'Injury and Pain' than Previously Known." *Fortune*, 25 Oct. 2023, <https://fortune.com/2023/10/25/amazon-worker-injuries-warehouse-study/>. Accessed 12 May 2024.
- Duhigg, Charles. "The Inside Story of Microsoft's Partnership with OpenAI." *The New Yorker*, 1 Dec. 2023, <https://www.newyorker.com/magazine/2023/12/11/the-inside-story-of-microsofts-partnership-with-openai>. Accessed 22 Mar. 2024.
- Elimian, Godfrey. "ChatGPT Costs \$700,000 to Run Daily, OpenAI May Go Bankrupt in 2024." *TechNext.Ng*, 14 Aug. 2023, <https://technext24.com/2023/08/14/chatgpt-costs-700000-daily-openai/>. Accessed 13 May 2024.
- Forsythe, Michael, and Walt Bogdanich. "McKinsey Settles for Nearly \$600 Million Over Role in Opioid Crisis." *The New York Times*, 4 Feb. 2021, <https://www.nytimes.com/2021/02/03/business/mckinsey-opioids-settlement.html>. Accessed 12 May 2024.
- Fortune Editors. "Anthropic's CEO Says Why He Quit His Job at OpenAI to Start a Competitor That Just Received Billions from Amazon and Google." *Yahoo Finance*, 26 Sept. 2023, <https://finance.yahoo.com/news/anthropic-ceo-says-why-quit-194409797.html?guccounter=1>. Accessed 20 Mar. 2024.
- Franta, Benjamin. "Shell and Exxon's Secret 1980s Climate Change Warnings." *The Guardian*, 19 Sept. 2018, <https://www.theguardian.com/environment/climate-consensus-97-per-cent/2018/sep/19/shell-and-exxons-secret-1980s-climate-change-warnings>. Accessed 12 May 2024.
- Grynbaum, Michael M., and Ryan Mac. "New York Times Sues OpenAI and Microsoft Over Use of Copyrighted Work." *The New York Times*, 27 Dec. 2023, <https://www.nytimes.com/2023/12/27/business/media/ne>

w-york-times-open-ai-microsoft-lawsuit.html. Accessed 28 May 2024.

Hao, Karen. "The Messy, Secretive Reality behind OpenAI's Bid to Save the World." *MIT Technology Review*, 17 Feb. 2020, <https://www.technologyreview.com/2020/02/17/844721/ai-openai-moonshot-elon-musk-sam-altman-greg-brockman-messy-secretive-reality/>. Accessed 22 Mar. 2024.

Harris, Tristan, and Ava Raskin. "The AI Dilemma." *Center for Humane Technology*, 5 Apr. 2023, <https://www.humanetech.com/podcast/the-ai-dilemma>. Accessed 22 Mar. 2024.

Ho, Karen. *Liquidated: An Ethnography of Wall Street*. Duke University Press, 2009.

Hutchinson, Andrew. "Snap Outlines New Safeguards for Its 'My AI' Chatbot Tool." *Social Media Today*, 5 Apr. 2023, <https://www.socialmediatoday.com/news/snap-outlines-new-safeguards-for-its-my-ai-chatbot-tool/646972/>. Accessed 22 Mar. 2024.

Jones, Joseph. "Don't Fear Artificial Intelligence, Question the Business Model: How Surveillance Capitalists Use Media to Invade Privacy, Disrupt Moral Autonomy, and Harm Democracy." *Journal of Communication Inquiry*, Feb. 2024, <https://doi.org/10.1177/01968599241235209>.

Katella, Kathy. "How Social Media Affects Your Teen's Mental Health: A Parent's Guide." *Yale Medicine*, 8 Jan. 2024, <https://www.yalemedicine.org/news/social-media-teen-mental-health-a-parents-guide>. Accessed 12 May 2024.

Kerr, Dara. "OpenAI Reinstates Sam Altman as Its Chief Executive." *NPR*, 22 Nov. 2023, <https://www.npr.org/2023/11/22/1214621010/openai-reinstates-sam-altman-as-its-chief-executive>. Accessed 21 Mar. 2024.

Klein, Ezra. "The Unsettling Lesson of the OpenAI Mess." *The New York Times*, 22 Nov. 2023, <https://www.nytimes.com/2023/11/22/opinion/openai-sam-altman.html>. Accessed 23 Mar. 2024.

Knight, Will. "Sam Altman's Sudden Exit Sends Shockwaves Through OpenAI and Beyond." *WIRED*, 18 Nov. 2023, <https://www.wired.com/story/openai-sam-altman-ousted-what-happened/>. Accessed 22 Mar. 2024.

Lloyd, Jay. "Interview With Tristan Harris." *Issues in Science and Technology*, 16 May 2023, <https://issues.org/tristan-harris-humane-technology-misinformation-ai-democracy/>. Accessed 22 Mar. 2024.

Mann, Jyoti. "OpenAI Cofounder and Chief Scientist Says He Deeply Regrets Participating in Ousting Sam Altman." *Insider*, 20 Nov. 2023, <https://www.businessinsider.com/openai-cofounder-ilya-sutskever-deeply-regrets-participating-ousting-sam-altman-2023-11>. Accessed 22 Mar. 2024.

Massa, Annie, and Vernal Galpotthawela. "Bloomberg." *Sam Altman Is Worth \$2 Billion—That Doesn't Include OpenAI*, 1 Mar. 2024, <https://www.bloomberg.com/news/articles/2024-03-01/sam-altman-is-a-billionaire-thanks-to-vc-funds-startups>. Accessed 2 Mar. 2024.

McMillan, Robert. "Ilya Sutskever: The OpenAI Genius Who Told Sam Altman He Was Fired." *The Wall Street Journal*, Dow Jones & Company, Inc., 21 Nov. 2023, <https://www.wsj.com/tech/ai/ilya-sutskever-the-openai-genius-who-told-sam-altman-he-was-fired-26a3381c>. Accessed 21 Mar. 2024.

Metz, Rachel, and Shirin Ghaffary. "Bloomberg." *OpenAI's Sam*

- Altman Returns to Board After Probe Clears Him*, 8 Mar. 2024, <https://www.bloomberg.com/news/articles/2024-03-08/openai-s-altman-returns-to-board-after-probe-clears-him>. Accessed 14 May 2024.
- Metz, Cade, et al. "OpenAI Staff Threatens Exodus, Jeopardizing Company's Future." *The New York Times*, 20 Nov. 2023, <https://www.nytimes.com/2023/11/20/business/openai-staff-exodus-turmoil.html>. Accessed 29 Feb. 2024.
- Metz, Cade, and Tripp Mickle. "OpenAI Completes Deal That Values the Company at \$80 Billion." *The New York Times*, 16 Feb. 2024, <https://www.nytimes.com/2024/02/16/technology/openai-artificial-intelligence-deal-valuation.html>. Accessed 20 Mar. 2024.
- Metz, Rachel. "OpenAI and the Fierce AI Industry Debate Over Open Source." *Bloomberg*, 15 Mar. 2024, <https://www.bloomberg.com/news/newsletters/2024-03-15/openai-tumult-raises-question-of-how-open-an-ai-company-should-be>. Accessed 13 May 2024.
- Nidumolu, Jahnvi, et al. "Elon Musk Sues OpenAI for Abandoning Original Mission for Profit." *Reuters*, 1 Mar. 2024, <https://www.reuters.com/legal/elon-musk-sues-openai-ceo-sam-altman-breach-contract-2024-03-01/>. Accessed 20 Mar. 2024.
- Novet, Jordan. "Microsoft's \$13 Billion Bet on OpenAI Carries Huge Potential along with Plenty of Uncertainty." *CNBC*, 8 Apr. 2023, <https://www.cnbc.com/2023/04/08/microsofts-complex-bet-on-openai-brings-potential-and-uncertainty.html>. Accessed 8 Feb. 2024.
- OpenAI. "About." *OpenAI*, <https://openai.com/about/>. Accessed 13 May 2024.
- . "Our Approach to AI Safety." *OpenAI*, <https://openai.com/blog/our-approach-to-ai-safety>. Accessed 22 Mar. 2024.
- . "Our Structure." <https://openai.com/our-structure>. Accessed 7 Feb. 2024.
- "OpenAI Revenue, Valuation & Growth Rate." *Sacra*, <https://sacra.com/c/openai/>. Accessed 14 May 2024.
- Reuters. "Elon Musk Takes Another Swing at OpenAI, Makes xAI's Grok Chatbot Open-Source." *Reuters*, 11 Mar. 2024, <https://www.reuters.com/technology/elon-musk-says-his-ai-startup-xai-will-open-source-grok-chatbot-2024-03-11>. Accessed 13 May 2024.
- Roose, Kevin, et al. "Dario Amodei, C.E.O. of Anthropic, on the Paradoxes of A.I. Safety and Netflix's 'Deep Fake Love.'" *The New York Times*, 21 July 2023, <https://www.nytimes.com/2023/07/21/podcasts/dario-amodei-ceo-of-anthropic-on-the-paradoxes-of-ai-safety-and-netflixs-deep-fake-love.html>. Accessed 8 Feb. 2024.
- Satariano, Adam, et al. "Elon Musk Sues OpenAI and Sam Altman for Violating the Company's Principles." *The New York Times*, 1 Mar. 2024, <https://www.nytimes.com/2024/03/01/technology/elon-musk-openai-sam-altman-lawsuit.html>. Accessed 1 Mar. 2024.
- Toner, Helen, et al. "Decoding Intentions." *Center for Security and Emerging Technology*, 23 Oct. 2023, <https://cset.georgetown.edu/publication/decoding-intentions/>. Accessed 8 Feb. 2024.
- Tong, Anna. "Exclusive: ChatGPT Traffic Slips Again for Third Month in a Row." *Reuters*, 7 Sept. 2023, <https://www.reuters.com/technology/chatgpt-traffic-slips-again-third-month-row-2023-09-07/>. Accessed 29 Feb. 2024.
- Verma, Pranshu, and Gerrit De Vynck. "ChatGPT Took Their Jobs. Now They Walk Dogs and Fix Air Conditioners." *The Washington Post*, 2 June 2023, <https://www.washingtonpost.com/technology/2023/06/02/ai-taking-jobs/>. Accessed 8 Feb. 2024.
- Weise, Karen. "How Microsoft's Satya Nadella Kept the OpenAI Partnership Alive." *The New York Times*, 20 Nov. 2023, <https://www.nytimes.com/2023/11/20/technology/openai-microsoft-altman-nadella.html>. Accessed 21 Mar. 2024.
- Wiersema, Margarethe. "Holes at the Top: Why CEO Firings Backfire." *Harvard Business Review*, 1 Dec. 2002, <https://hbr.org/2002/12/holes-at-the-top-why-ceo-firings-backfire>. Accessed 14 May 2024.
- Wiggers, Kyle. "Investors Are Souring on OpenAI's Nonprofit Governance Model." *TechCrunch*, 20 Nov. 2023, <https://techcrunch.com/2023/11/20/openai-governance-model-investors/>. Accessed 12 May 2024.