Statement of Stephen M. Kohn
Attorney for Confidential AI Whistleblower

“Our firm was contacted by one or more whistleblowers concerning the legality of OpenAI’s non-disclosure agreements. Upon review, KKC determined that the agreements constitute clear violations of the Dodd Frank Act, and potentially violate federal obstruction of justice laws. On behalf of one or more employees or former employees, we filed a formal complaint with the U.S. Securities Exchange Commission (SEC) seeking a full federal investigation both into the illegal NDAs and the underlying issues or misconduct that employees may have been intimidated from disclosing.

“On July 1, 2024, we filed a follow up letter to the chairman of the SEC seeking to ensure that there be a vigorous and immediate investigation into the NDA issues. The ability of employees working in artificial intelligence to freely report concerns regarding AI must be quickly vindicated. Considering the serious safety issues posed by advanced AI, it is urgent that federal regulatory and law enforcement authorities such as the SEC take immediate and aggressive steps to ensure that NDAs are not interfering with the right to report.

“We cannot risk delaying effective oversight, the foundation of which is the ability of insiders to disclose potential violations safely, freely, and appropriately to law enforcement and regulatory authorities. Without whistleblower information, the ability of the U.S. government to police and regulate this newly developing technology will be handicapped, risking catastrophic disaster.

“The threats that AI poses to the public are well documented. Serious risks of ‘drastic accidents and societal disruption’ and risks that are ‘existential’ in their potential adverse effects have been acknowledged by OpenAI. Google DeepMind has noted that ‘it is plausible that future AI systems could conduct offensive cyber operations, deceive people through dialogue, manipulate people into carrying out harmful actions, develop weapons (e.g. biological, chemical), … due to failures of alignment, these AI models might take harmful actions even without anyone intending so.’ A presidential Executive Order noted that the ‘irresponsible’ use of AI ‘could exacerbate societal harms such as fraud, discrimination, bias, and disinformation; displace and disempower workers; stifle competition; and pose risks to national security.’”

“This is not about halting or destroying the development of AI — it’s about making sure the technology is safe. We must have a culture of transparency, openness and honest debate. You cannot build safe AGI in the dark.”

---

1 OpenAI: “Planning for AGI and beyond” February 2023
2 Google DeepMind: “AI Safety Summit: An update on our approach to safety and responsibility” October 2023
3 White House: “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence” October 2023
Kohn, Kohn & Colapinto LLP has successfully opposed illegal NDAs since 1988, filing the first successful SEC case against restrictive NDAs in 2015 as well as the first successful CFTC case against restrictive NDAs in 2024.

For more information contact:
   Sophie Luskin – sophie.luskin@kkc.com – (510) 507-2839 (Signal)