

Report on AI Data Usage Oversight

Whereas: The immense and transformative potential of artificial intelligence comes with substantial risks.

The development and training of AI systems rely on vast amounts of data, and public information available via the Internet may not be enough to quench developers' insatiable thirst for high-quality training data.¹ Thus, stakeholders are concerned that developers will draw from unethical or illegal sources – such as personal information collected online,² copyrighted works,³ and proprietary commercial information provided by users.⁴

Supporting Statement: Microsoft Corporation (“Microsoft” or the “Company”) is an early leader in the AI arms race^{5 6} because of its extensive partnership with OpenAI,⁷ which has helped push the Company to the highest market capitalization in the world.⁸

But shareholders should be concerned with Microsoft’s record on data ethics:

- Microsoft employs generative AI models developed by OpenAI, which allegedly stole large amounts of personal information by scraping the web, including “private information and private conversations, medical data, information about children — essentially every piece of data exchanged on the internet it could take — without notice to the owners or users of such data, much less with anyone's permission.”⁹
- OpenAI recently appointed a former head of the National Security Agency – which has been criticized for spying on American citizens – to its board.¹⁰ The move increases fears that Microsoft and OpenAI are renegeing on their privacy assurances.
- Microsoft has received pushback against its proposed AI “Recall” feature, which screenshots everything a Windows user sees or does and turns it into searchable data. Users thought the feature was a gross violation of privacy and a cybersecurity risk.^{11 12}
- Microsoft or OpenAI, through their services, may inadvertently or deliberately access and utilize proprietary information provided by users, potentially leading to unauthorized use or exposure of sensitive business information.^{13 14}

¹ <https://www.wsj.com/tech/ai/ai-training-data-synthetic-openai-anthropic-9230f8d8>

² <https://www.scientificamerican.com/article/your-personal-information-is-probably-being-used-to-train-generative-ai-models/>

³ <https://www.axios.com/2024/01/02/copyright-law-violation-artificial-intelligence-courts>

⁴ <https://www.reuters.com/legal/legalindustry/risks-unleashing-generative-ai-company-data-2023-12-22/>

⁵ <https://www.americanbanker.com/news/aws-google-and-microsoft-are-in-an-ai-arms-race-banks-are-watching>

⁶ <https://www.cnbc.com/2023/04/26/google-had-a-kodak-moment-as-microsoft-takes-lead-in-ai-strategist.html>

⁷ <https://www.theverge.com/2024/5/1/24146302/microsoft-openai-investment-google-worries-internal-emails>

⁸ <https://companiesmarketcap.com/>

⁹ <https://www.businessinsider.com/openai-chatgpt-generative-ai-stole-personal-data-lawsuit-children-medical-2023-6>

¹⁰ <https://fortune.com/2024/06/14/edward-snowden-viscerates-openai-paul-nakasone-board-directors-decision/>

¹¹ <https://time.com/6980911/microsoft-copilot-recall-ai-features-privacy-concerns/>

¹² <https://www.theverge.com/2024/6/7/24173499/microsoft-windows-recall-response-security-concerns>

¹³ <https://www.legaldive.com/news/azure-openai-service-confidentiality-loophole-legaltech-news-reports/711006/>

¹⁴ <https://www.bloomberg.com/news/articles/2023-05-02/samsung-bans-chatgpt-and-other-generative-ai-use-by-staff-after-leak>

- Microsoft and OpenAI have been sued by the *New York Times*, among others, which alleged copyright infringement.¹⁵

Prioritizing data ethics in Microsoft's AI development may help avoid harmful fiduciary and regulatory^{16 17} consequences.¹⁸ Americans surveyed by Pew Research Center have expressed that data privacy and usage are among their main concerns with Big Tech's AI initiatives.¹⁹ Developers who prioritize ethical data usage will reap the benefits of consumer trust, while those that do not will suffer.

Microsoft's position in the AI arms race, and its associated historic valuation, hang in the balance.

Resolved: Shareholders request the Company prepare a report, at reasonable cost, omitting proprietary or legally privileged information, to be published within one year of the Annual Meeting and updated annually thereafter, which assesses the risks to the Company's operations and finances, and to public welfare, presented by the real or potential unethical or improper usage of external data in the development and training of its artificial intelligence offerings; what steps the Company takes to mitigate those risks; and how it measures the effectiveness of such efforts.

¹⁵ <https://www.reuters.com/legal/transactional/ny-times-sues-openai-microsoft-infringing-copyrighted-work-2023-12-27/>

¹⁶ <https://hbr.org/2023/05/who-is-going-to-regulate-ai>

¹⁷ <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/01/ai-companies-uphold-your-privacy-confidentiality-commitments>

¹⁸ <https://www.cnn.com/2024/01/05/microsoft-openai-sued-over-copyright-infringement-by-authors.html>

¹⁹ <https://www.pewresearch.org/short-reads/2023/10/18/key-findings-about-americans-and-data-privacy/>