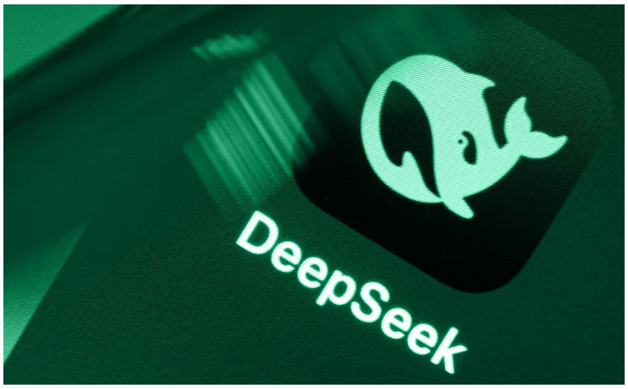


Expert Policy Research Article

DOC ILA - NEW TECH/AI/06

Deepseek AI:

Navigating Intellectual Property, Data Privacy, And Global Compliance Challenges

Authors: Favour Ernest, Paul Obala, Akinola Oluwatobi Ogo-oluwa.

Summary

The DeepSeek AI controversy brings to light critical issues surrounding data privacy, intellectual property violations, and the broader implications of artificial intelligence governance in an increasingly complex technological and regulatory environment. Allegations against DeepSeek center on the unauthorized use of proprietary content through model distillation, a technique that raises significant concerns about intellectual property rights and trade secrets. This practice has sparked debates over the boundaries of copyright and trade secret protections in the AI industry, particularly regarding the use of OpenAI's outputs to train DeepSeek's model.

In addition to these intellectual property concerns, DeepSeek's data handling practices have attracted scrutiny, with potential violations of privacy regulations such as the GDPR in Europe and the CCPA in California, and of particular concern are the company's data storage practices in China, where local laws mandate data sharing with authorities, as well as the potential risks posed by cross-border data transfers.These issues are compounded by geopolitical tensions, especially the ongoing U.S.-China rivalry in AI technology, raising national security concerns about the potential misuse of AI for military or surveillance purposes. This article delves into the legal, ethical, and operational risks that arise from DeepSeek's practices, shedding light on the broader challenges AI companies face when navigating the areas of data protection, intellectual property, and international regulations. The brief concludes with several key policy recommendations for DeepSeek to mitigate these risks. These include strengthening compliance with international data privacy laws, establishing clear and transparent data governance policies, enhancing cybersecurity measures, and fostering ethical AI development.

1.0 Introduction

The protection of information provided by users is a fundamental responsibility of data controllers, and when this data is not adequately protected through secure management practices and compliance with general data protection regulations, users and data subjects are left with concerns about

unauthorized access to their information. This became a point of contention in the instance of DeepSeek, when a cybersecurity firm identified an alleged data leak involving user data, including questions submitted through prompts.

The legality of the DeepSeek model has been a matter of concern, especially since there are very few statutes that govern the operations of AI, especially on the international scene. On this segment, we discuss the core legal issues that the DeepSeek allegation presents, in line with data privacy, IP theft, amidst the lack of concrete AI law on the international scene. This article explores the issue of data privacy in light of the alleged 'cybersecurity flaws' in the protection mechanisms for sensitive data.[1]

2.0 BRIEF OVERVIEW OF FACTS

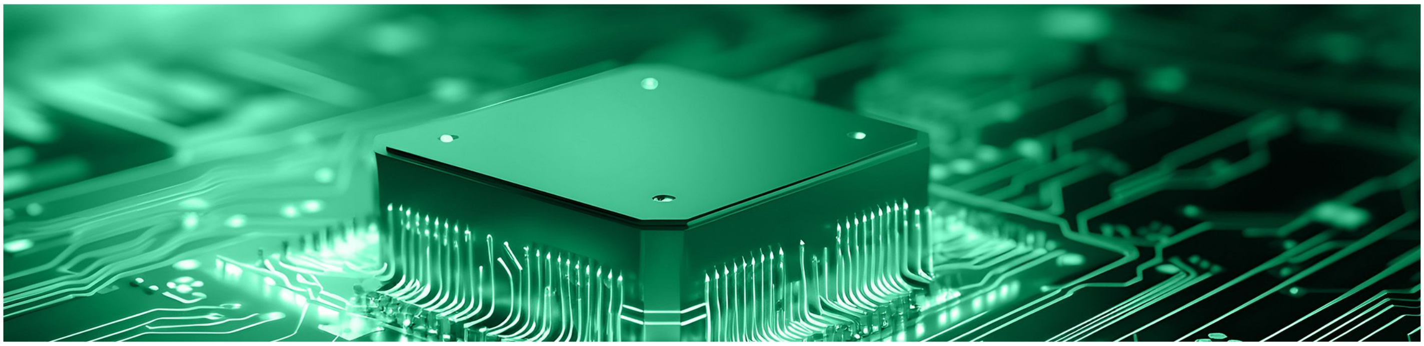
2.1 Intellectual Property Violations and Model Distillation

DeepSeek is facing allegations of intellectual property violations due to its use of model distillation, a technique where a smaller AI model is trained using the outputs of a larger, pre-existing model. In this case, it is claimed that DeepSeek used outputs from OpenAI's models to train its own, raising concerns over copyright and trade secret violations.[2] Model distillation allows for the extraction of insights from a more sophisticated model to improve the performance of a smaller model, but the use of proprietary outputs in this manner could be seen as unauthorized use of protected materials.

The allegations came to light when a cybersecurity firm identified that DeepSeek's model was trained on large-scale data scraped from sources that could include proprietary content from OpenAI. While AI-generated text is generally not protected under copyright, the underlying data and methodologies used to train the model may still be subject to intellectual property protections. OpenAI, along with other potential stakeholders, may argue that the training data and algorithms involved are proprietary, potentially violating trade secret laws. DeepSeek may defend its actions by claiming that it used publicly available outputs or relied on fair use provisions, asserting that no proprietary data or internal methods were

[1] Lars Daniel, 'DeepSeek Data Leak Exposes 1 Million Sensitive Records' (1 February 2025) <https://www.forbes.com/sites/larsdaniel/2025/02/01/deepseek-data-leak-exposes--1000000-sensitive-records/> accessed 4 February 2025.

[2] Andres Guadamuz, 'Will DeepSeek Impact the AI Copyright Wars?' (30 January 2025) <https://www.technollama.co.uk/will-deepseek-impact-the-ai-copyright-wars> accessed 11 February 2025.



accessed. This conflict highlights the ongoing legal complexities surrounding AI training, especially in cases where data ownership and access rights are not clearly defined.[3]

3.0 DATA PRIVACY AND CONSUMER PROTECTION CONCERNS

Concerns about DeepSeek's data collection practices involve possible violations of privacy laws, cross-border data transfer risks, and content moderation issues. While these concerns have not been conclusively validated, they merit further examination given the regulatory complexities involved.

3.1 User Data Collection and Storage

AI models like DeepSeek rely on extensive datasets, raising questions about compliance with data protection regulations such as the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA) in the U.S. These laws impose strict requirements regarding data collection, processing, and consent.

One concern is DeepSeek's reported practice of storing user data in China, where laws mandate that companies must share data with authorities upon request. This has led to scrutiny, especially in regions like Italy[4] and Taiwan[5], where the company was banned after an investigation into its data handling. DeepSeek may argue that it adheres to international privacy standards through anonymization and consent-based collection methods, though questions about compliance with GDPR and CCPA remain.

3.2 Cross-Border Data Transfers

China's Personal Information Protection Law (PIPL) and Data Security Law (DSL) require companies to store specific data within the country and regulate the conditions under which data can be transferred internationally. In contrast, U.S. laws, such as the CLOUD Act, allow American authorities to request access to data stored overseas, raising concerns about the potential exposure of U.S. user data

[3] Lars Daniel, 'DeepSeek Data Leak Exposes 1 Million Sensitive Records' (1 February 2025) <https://www.forbes.com/sites/larsdaniel/2025/02/01/deepseek-data-leak-exposes--1000000-sensitive-records/> accessed 4 February 2025.

[4] Reuters, 'Italy's Regulator Blocks Chinese AI App DeepSeek on Data Protection' (30 January 2025) <https://www.reuters.com/technology/artificial-intelligence/italys-privacy-watchdog-blocks-chinese-ai-app-deepseek-2025-01-30/> accessed 10 February 2025

[5] Reuters, 'Taiwan Bans Government Departments from Using DeepSeek AI' (3 February 2025) <https://www.reuters.com/technology/taiwan-bans-government-departments-using-deepseek-ai-2025-02-03/> accessed 10 February 2025.

to foreign government surveillance.[6]

DeepSeek may argue that it complies with both Chinese and international data protection laws, maintaining strict data localization and transferring data in line with recognized legal frameworks. However, cross-border data transfer remains a complex issue for global companies and could present regulatory challenges.[7]

4.0 REGULATORY AND NATIONAL SECURITY CONCERNS SURROUNDING DEEPSEEK

DeepSeek's operations have attracted attention due to the ongoing geopolitical tensions between the United States and China, raising concerns about national security implications. Two primary issues are central to this:

4.1 U.S. Export Control Regulations: The U.S. government has placed restrictions on American companies selling advanced AI hardware to China, particularly high-performance GPUs from companies like Nvidia. If DeepSeek used such technology before the restrictions took effect, it could face investigations regarding potential violations of U.S. export controls. Additionally, DeepSeek could face legal scrutiny under U.S. export laws, which may result in penalties or sanctions if the company is found to have violated these regulations.

4.2 National Security Risks: DeepSeek's use of advanced AI technology has raised concerns about potential military applications, particularly given the company's ties to China. The U.S. government may view DeepSeek's operations as a national security risk, especially if it is perceived as having close connections with the Chinese government. These concerns could shape future regulatory decisions and government policies.

5.0 EFFECTS OF DATA PRIVACY ISSUES ON DEEPSEEK AI

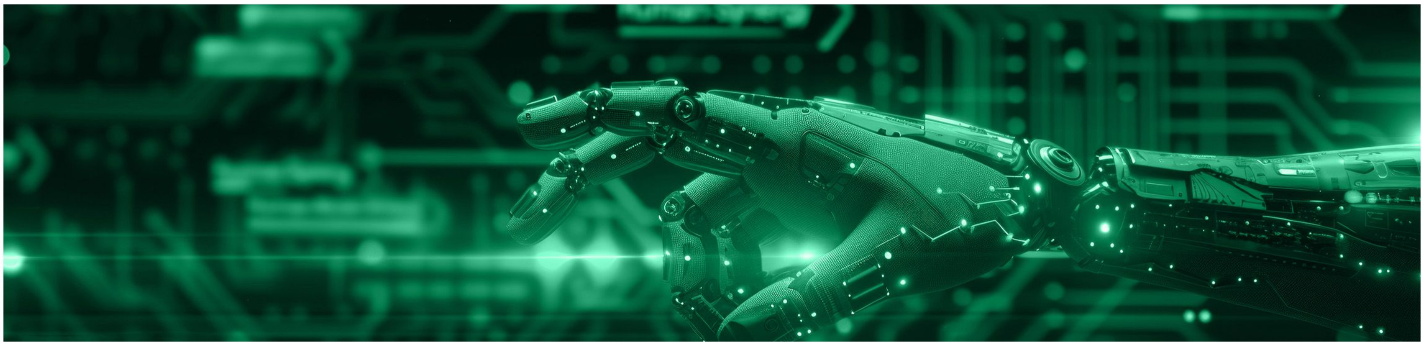
The ongoing discussions about DeepSeek's data privacy practices highlight the broader challenges faced by companies in managing user data responsibly and maintaining public trust.

As AI continues its relentless march forward, the importance of airtight data protection has never been more pronounced.[8]

[6] Feroot, 'Security Research Reveals DeepSeek AI's Hidden Data Pipeline to China' (5 February 2025) <https://www.feroot.com/news/feroot-security-research-reveals-deepseek-ais-hidden-data-pipeline-to-china/> accessed 10 February 2025.

[7] ibid

[8] Cate F. H., & Mayer-Schönberger, V. (2020). Data ownership: who owns our data? International Data Privacy Law, 10(1), 1-7.



5.1 Legal Consequences: If DeepSeek is found to be non-compliant with data protection laws such as GDPR or CCPA, it could face significant legal consequences, including fines and operational restrictions. The potential for cross-border data transfer issues could also expose DeepSeek to further legal challenges, particularly in Europe and the U.S.

5.2 Reputational Damage: Public concerns about the potential misuse of personal data, as well as the possibility of government surveillance, could damage DeepSeek's reputation. As surveys have shown, many consumers are wary of how AI companies handle their data, and any negative publicity could lead to a loss of user trust and credibility.

5.3 Business and Operational Risks: Tighter regulations and increased scrutiny of cross-border data transfers could present operational challenges for DeepSeek. These challenges could result in additional costs associated with ensuring compliance and adapting to local laws, which may slow down expansion or require significant infrastructure changes.[9]

5.4 International Trade and Geopolitical Tensions: DeepSeek has unwittingly become a pawn in the global chessboard of technological supremacy. The U.S.-China AI rivalry has placed an intense spotlight on the company, with fears that its operations could provide a strategic advantage to state-backed initiatives. Governments wary of AI's military applications have already imposed trade restrictions, curbing DeepSeek's access to critical hardware like high-performance AI chips (Lewis, 2022). If these restrictions intensify, DeepSeek could find itself technologically hamstrung, unable to keep pace with Western competitors.[10]

[9] European Data Protection Board. (2023). Guidelines on Data Protection Impact Assessment. EDPB Publications.

[10] Ibid

6.0 CONCLUSION

The DeepSeek AI controversy is a stark reminder that cutting-edge technology must be tempered with responsibility. The company's alleged missteps in handling user data, coupled with its opaque privacy policies, have reignited debates on AI ethics and regulatory enforcement. As AI legislation struggles to keep up with rapid innovation, DeepSeek's case highlights the pressing need for well-defined global AI governance structures.

While DeepSeek maintains its compliance with Chinese data laws, international scrutiny suggests that a localized approach to privacy will not suffice in a globally interconnected world. Regulatory bodies, industry leaders, and consumers now demand a more transparent and accountable AI ecosystem.[11] If DeepSeek wishes to regain public confidence and secure its place in the industry, proactive reforms and meaningful commitments to data protection are non-negotiable.

7.0 RECOMMENDATIONS

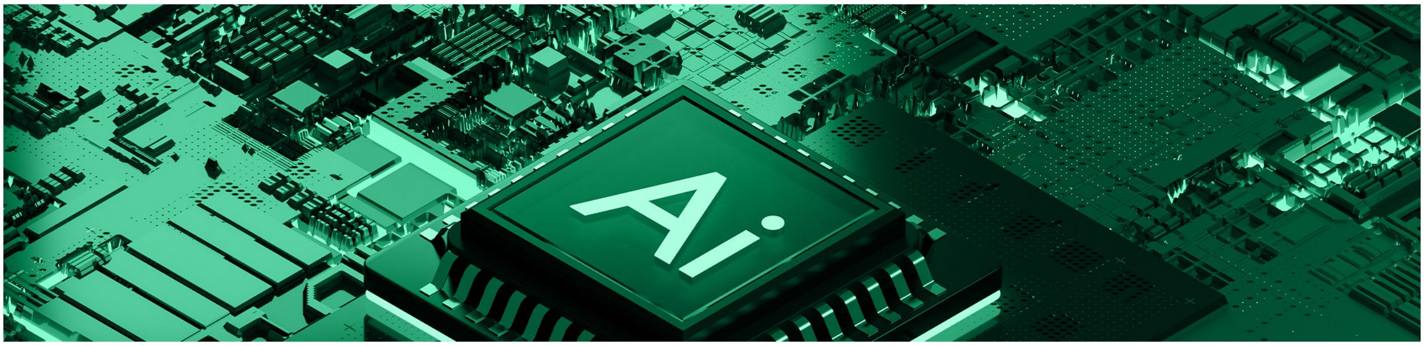
To navigate this treacherous landscape, DeepSeek AI must embrace strategic, proactive measures that reinforce compliance, transparency, and ethical responsibility.

7.1 Strengthen Compliance with International Data Privacy Laws: DeepSeek must go beyond surface-level compliance and adopt a rigorous, global-first approach to data protection. Aligning with GDPR, CCPA, and other regulations requires meticulous consent management, data minimization strategies, and a robust anonymization framework.

7.2 Establish Transparent Data Governance Policies: Vagueness breeds distrust. DeepSeek must draft and publicly disclose comprehensive data governance policies, updated continuously to reflect regulatory shifts. Independent third-party audits should validate compliance, assuring stakeholders of DeepSeek's commitment to ethical AI.

7.3 Implement Unbreakable Cybersecurity Measures: With high stakes, cybersecurity cannot be an afterthought. DeepSeek must adopt state-of-the-art encryption protocols, multi-factor authentication, and real-time intrusion detection systems. Frequent, rigorous security audits will help uncover and

[11] Burdon, M. (2021). Digital Data and the Law: Privacy, Security and Surveillance.



address vulnerabilities before they escalate into catastrophic breaches.[12]

7.4 Rethink Cross-Border Data Transfers: To ease concerns surrounding government oversight and data localization, DeepSeek must establish regional data centres in key markets, ensuring compliance with local storage laws. A legal task force specializing in international data transfer policies should guide these efforts, mitigating regulatory risks.

7.5 Commit to Ethical AI Development: DeepSeek should spearhead industry-wide discussions on ethical AI, engaging with regulators, researchers, and civil society groups to craft AI policies that balance innovation with privacy protection. Such engagement will not only enhance public perception but also contribute to shaping a responsible AI landscape.

7.6 Rebuild Consumer Trust with Radical Transparency Users must be empowered to control their data. DeepSeek should implement clear, user-friendly data management options, allowing individuals to delete, modify, or opt out of data collection. An ongoing dialogue with regulators and users can help reshape its reputation and restore faith in the platform.

[12] Nissenbaum, H. (2019). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.