Risks of unleashing generative AI on company data

By Ken D. Kumayama, Esq., and Pramode Chiruvolu, Esq., Skadden, Arps, Slate, Meagher & Flom LLP

DECEMBER 22, 2023

"We have all this data — let's do something with it!" You have doubtless heard something like this in a meeting — or perhaps on a call from a coworker eager to unleash an army of data scientists on the company's vast data lakes. The benefits could be game-changing; the consequences of doing nothing could be dire. Your competitors may have already deployed tools that dramatically increased productivity, slashed costs and unlocked insights and enhancements that will soon put you out of business!

What is legal counsel to do, faced with such an imperative? You already have analysts generating actionable business intelligence and AI isn't new. What are the real risks vs. hype?

Generative AI ("GenAI") generates risks

The crescendo of calls to leverage data comes from the hype surrounding GenAI models — massive math functions with coefficients found through feeding in training data and updating the coefficients to reduce the error rate. By some miracle of data science, carrying out this process at massive scale has built models capable of composing new sonnets, and developing complex software and almost anything else you can imagine.

Despite their promise, three categories of risks arise when companies train or prompt GenAI models with data: (1) infringement of intellectual property ("IP") rights, (2) disclosure of confidential information (along with related risks), and (3) compliance with laws and regulations. While, in theory, any use of data raises these concerns, the risks are amplified by the proliferation of GenAI tools, the volume of GenAI inputs and outputs and the often-heavy investments required to train or deploy GenAI models.

IP infringement

A number of GenAI companies have been sued for copyright infringement based on allegations that they trained models using copyrighted data scraped from the internet. Even if these GenAI companies ultimately prevail in asserting that this training is "fair use" (and not infringement), a definitive answer to this question is likely years away.

In the meantime, companies should assess whether the data in question is subject to third-party IP rights and whether training or prompting a GenAI model with that data will violate those rights. Companies should also consider their potential exposure, including in light of statutory damages for copyright infringement.

Confidentiality and related considerations

GenAI systems often require large amounts of computing power to operate, so, most GenAI systems are deployed on third-party servers in the cloud. For example, the state-of-the-art language model, GPT-4, is currently accessible only through OpenAI's or Microsoft's systems. Companies training their own models often have vendors carry out the training process using third-party cloud services, but it's important to keep in mind that disclosure of confidential information to any third party creates the potential for leaks of company confidential and sensitive information. For instance, a company could lose protection of trade secrets or waive attorney-client privilege as to communications used with a GenAI model, particularly where the terms of use give a GenAI vendor broad rights to use such data.

Despite their promise, three categories of risks arise when companies train or prompt GenAI models with data: (1) infringement of intellectual property rights, (2) disclosure of confidential information (along with related risks), and (3) compliance with laws and regulations.

In addition, companies might violate confidentiality obligations or other contractual obligations if they share third-party data with GenAI vendors. And training a GenAI system entirely in-house does not necessarily shield you from this risk — doing so may still breach confidentiality and other obligations restricting use of third parties' data.

Companies should therefore carefully consider their contractual obligations, as well as the potential for trade secret leakage or loss of privilege, when assessing any planned use of company data in a GenAI system.

Legal compliance

Training or prompting GenAI models with personally identifiable information ("PII") may violate privacy and other laws used to enforce privacy commitments, including the California Consumer Privacy Act ("CCPA") and Section 5 of the Federal Trade Commission Act. Notably, the FTC has required companies that trained AI models in violation of their privacy commitments to delete not only the data containing PII, but also the models themselves.

Companies contemplating using GenAI tools in recruiting, hiring or promoting employees should note that in addition to general anti-discrimination laws, there are now an increasing number of AI-specific employment laws, including New York City's Local Law 144 of 2021, which prohibit employers from using an automated employment decision tool in New York City unless the tool undergoes an independent audit for bias and proper notice to the potential hires is given.

Companies should assess whether the data in question is subject to third-party IP rights and whether training or prompting a GenAI model with that data will violate those rights.

Sector-specific regulations, such as the Securities and Exchange Commission's regulation of conflicts of interest between clients and broker-dealers or investment advisers, can also impact companies' GenAI plans. Companies that have data subject to export controls should also consider whether training or prompting GenAI with such data could violate those controls.

Companies should also watch the coming implementation of President Biden's Executive Order 14110 (the "AIEO") and related federal initiatives and assess their impact on using company data with GenAI models. For example, the AIEO includes an obligation that companies planning to train models above a certain size submit reports about their activities and perform certain assessments and "red-teaming" of their models.

Finally, companies operating internationally should monitor developments in relevant jurisdictions to ensure their plans comply with existing and future regulations.

Consider how data will be used

Companies can use GenAI to leverage their data in three ways.

First, companies can use their data to train or "fine-tune" a new GenAI model. "Fine-tuning" takes an existing general model to tailor it to a specific context. For example, a health care company might fine-tune an existing model on patient data or a corpus of medical knowledge to create a chatbot to answer medical questions.

Training GenAI models on a company's confidential information or third-party data can raise infringement and confidentiality risks. But if a company is training a GenAI model third parties don't access, the confidentiality risks and the risks of detection of infringement may be low.

Secondly, companies can use their data for retrieval augmented generation (RAG) — retrieving relevant data from a database to include in the prompt to ground the GenAI system's responses in the specific context of the retrieved data. For example, a health care company could deploy a RAG system that can retrieve patient records and allow a medical professional to ask natural language questions about a patient's medical history.

If a RAG system submits data to third parties' models, consider whether the terms of use for those third-party models give the third party broad rights to use prompt data. If so, trade secret protection can be lost or confidentiality or use restrictions may be breached. The RAG system can also undermine a company's data access controls if the system is not carefully designed to honor such controls in retrieving data for prompting the GenAI model.

Finally, company data may be licensed to others seeking to train or fine-tune their own AI systems or to use as a data source for RAG. Licensors should consider limiting their liability for IP infringement; ensuring the license does not result in trade secret loss or violate confidentiality or similar obligations to third parties; and shifting the risks of the data's disclosure and use to the licensee, including through appropriate indemnification obligations and disclaimers of liability.

Consider the data source

Company data can be internal data, external data or synthetic data (synthetic data can itself be generated internally or externally). Internal data originates from within the company, including product data and employee data. External data comes from third parties, including customer data and data scraped from the internet. Synthetic data is artificially generated — for example, a number of smaller language models have been trained using synthetic datasets consisting of responses generated by larger language models.

Internal datasets are more likely to raise confidentiality concerns with respect to the company's own proprietary information, but external data introduces greater infringement risk and the risk of breaching any related confidentiality or similar restrictions when that data is ingested into a GenAI system.

Both internal and external datasets may include PII that is connected to a specific individual and can be used to uncover their identity. Using synthetic data may limit the privacy risks by reducing the chance that the PII of a real person is included in the data. But synthetic data can be difficult to generate and overreliance on synthetic data might result in models that perform poorly because the synthetic data is not representative of real world data.

Key takeaways

- Legal counsel should carefully evaluate proposals to train or prompt GenAI with company data, as it raises risks related to IP infringement, confidentiality (and related considerations), and general and sector-specific laws and regulations.

- How the data will be used matters — generally, data is used in training, prompting or for licensing, and each case presents its own risks.
- Where data comes from also matters — data can be internal, external or synthetic, and the risk profile of each varies.
- While this article focuses on the risks on the input side, the outputs of GenAI systems also create risks that we will explore in a future article in this series.

About the authors

Ken D. Kumayama (L) is a partner in the intellectual property and technology group at **Skadden, Arps, Slate, Meagher & Flom LLP**. He concentrates his practice on transactional matters relating to intellectual property, technology, privacy and cybersecurity, as well as artificial intelligence and machine learning. He can be reached at ken.kumayama@skadden.com. **Pramode Chiruvolu** (R) is a counsel in the intellectual property and technology group at the firm. He advises clients on complex transactional matters involving emerging technologies, including artificial intelligence, digital health and biotechnology, the internet of things and 5G networks. He can be reached at pramode.chiruvolu@skadden.com.

This article was first published on Reuters Legal News and Westlaw Today on December 22, 2023.