TOPIC ONE – GENERATIVE AI, DATA, AND HOW TO PROTECT WHAT'S YOURS

Large language models and other artificial intelligence programs like ChatGPT are unprecedented in their ability to absorb information and generate output that discloses, mimics or is closely tied to that information. Although AI programs have yet to gain widespread adoption, some companies have already connected their applications and datasets to AI systems, unaware that that these systems may collect far more information than they are aware of, and that sensitive information fed into the models as training data could resurface when prompted by the right queries. Some companies, like JPMorgan, Amazon, Microsoft and Wal-Mart, have taken steps to protect their information by restricting employee use of ChatGPT and issuing warnings to take care when using AI.

On the flip side, not only is there risk that a company's proprietary information may be misappropriated by others using AI programs, but companies relying on AI programs to generate output may themselves end up receiving and unwittingly using information that is copyrighted, trademarked, or the intellectual property of another person or entity.

At present, we have only a murky picture of how new and existing law might apply to AI systems. Issues yet to be decided include whether and to what extent the input and output of AI programs is protected under the fair use doctrine (see, e.g., the Getty Images case against Stability and the U.S. Supreme Court case involving Andy Warhol and Prince, whether the use of proprietary information in training AI systems can lead to unfair competition, and whether AI systems violate license agreements restricting the use of data (see class action filed by Github users against Microsoft and OpenAI).

As the law around AI develops, companies can take the following steps to protect their proprietary information and intellectual property:

- Implement technology (such as encryption or firewalls) to prevent AI systems from accessing company networks and data, and block employee access to AI programs.
- Block employee access to AI programs and amend employee confidentiality agreements and policies to prohibit employees from referring to or entering confidential, proprietary or trade secret information into AI programs.
- Add website and metadata tags warning AI companies that they may not scrape or otherwise use your data.
- If you contract with a service provider using AI, insist on robust contractual protections for all data you provide to them, as well as information derived from data you provide; consider if and to what extent the AI service provider may use your data to improve its AI systems.
- If you are using an AI system to generate output, insist that the AI provider bears the risk associated with your use of the AI system by indemnifying you against any third-party claims of intellectual property or privacy violations.