NEW YORK LAW JOURNAL

DeepSeek's Legal and Cross-Cultural Risks and Benefits

By William Tanenbaum and Dr. Matthias Orthwein | March 7, 2025

The cultural and business biases in the Chinese-built AI Large Language Model (LLM) for DeepSeek, and the datasets used to train the LLM, can adversely influence business-focused AI research. These cross-cultural factors particularly apply in cross border business transactions. This article focuses on the potential legal and business risks as well as the potential benefits of DeepSeek and addresses the application of EU privacy law to DeepSeek technology and business practices.

What is the Impact of the *Thomson Reuters* AI Training Data Case?

DeepSeek's launch was followed closely by the first U.S. case to substantively address whether the use of copyrighted material to train AI is an activity protected by the copyright doctrine of fair use, *Thomson Reuters Enterprise Center Center GMBH and West Publishing Corp. v. Ross Intelligence Inc.* decided by the U.S. District Court for the District of Delaware (Case No. 1:20- CV-SB on), decided on Feb. 11, 2025. While not a Generative AI case, its ruling focuses on a key issue that other courts will face in claims of assessing copyright infringement involving training data and Generative AI.

Thompson Reuters argued that the subcontractor retained by Ross (a legal research platform company that is no longer in business) copied Westlaw's proprietary case law headnotes to build a set of materials (called "Bulk Memos") to train Ross' AI powered legal research product.

While many issues were deferred to trial, the court granted summary judgment and rejected Ross' fair use defense for using copyrighted material by finding in relevant part that Ross "meant to compete with Westlaw by developing a market substitute."

Many Generative AI companies rely on the fair use defense in creating their training data; DeepSeek, like others, will face arguments based on this Court's ruling in litigation over the right to use copyrighted material in training data free of infringement.

What is the Cross-Cultural Impact of DeepSeek in Business?

Putting aside the Chinese government's political and historical censorship affecting DeepSeek AI outcomes, DeepSeek reflects Chinese cultural and business norms and the Chinese economic system. These factors are different from their equivalent in U.S. LLMs.

From this perspective, DeepSeek AI answers are not necessarily "wrong," but do complicate cross-border business transactions because of these differences. United States companies using DeepSeek in business transactions should take cognizance of Chinese norms that will be reflected in DeepSeek output.

As background, in AI, data turns into information, information turns into insights, and insights turn into actionable business plans. Moreover, the meaningfulness of data depends on its context. The fact that the temperature is 30 degrees is not in itself useful data.

It lacks context. That temperature means it is cold in Hawaii but it is time to put on shorts in Maine. In the DeepSeek context, business plans resulting from this paradigm means that the plans should be evaluated in terms of the norms and assumptions built into the DeepSeek model, and the facts provided need to be evaluated in context, especially when the use of data is repurposed from its original context.

However, while DeepSeek outputs may have limitations in certain business use cases, the Chinese

assumptions built into DeepSeek and the omissions in its results can provide useful positive insights, which are discussed further below.

What is the Data Supply Chain for DeepSeek?
In AI, the data itself is the supply chain. AI data is analogous to components acquired from different sources and assembled to manufacture a finished machine. In the manufacturing context, problems arise when a defective component is not discovered until it causes the machine to malfunction.

Turning to DeepSeek, factual biases and omissions (including historical events such as Tiananmen Square) underlie DeepSeek. Moreover, DeepSeek uses the Chinese Internet as a source of data, and the information generated in Internet search results is subject to distortion by the political control exerted over the content of information provided through the Chinese internet.

Not only is the DeepSeek LLM programmed to censor government-designated sensitive data, but the underlying data compromises DeepSeek LLM even before the LLM is "fine-tuned" to meet Chinese political and social goals.

Both the Chinese Internet and the DeepSeek LLM are intentionally designed to promote Chinese political policy and the Chinese Political Party's positioning of China on the world stage. Thus, the DeepSeek data supply chain is designed to generate AI output that includes meeting government goals. This differs in material ways from the inherent bias in U.S. private sector LLMs.

Can DeepSeek Be a Political Research Assistant?
Because DeepSeek outputs will reflect Chinese political positions, as a result of what is omitted or distorted when its results are compared to other AI solutions, DeepSeek can help companies get a better understanding of Chinese political goals, including as they apply to economic activity. DeepSeek may provide a leading-edge indicator of changes in political and economic goals and how China wants to position itself in global trade.

Can DeepSeek be a Cultural Negotiation Mentor?
U.S. and Chinese negotiation styles and tactics differ (as do Western and Asian practices in general). For reasons similar to those above for using DeepSeek to gain political insights, strategic use of DeepSeek can help negotiators understand what techniques work with Chinese companies and what negotiation styles are counter-productive.

This can help negotiators understand cross cultural differences. It can also help younger negotiators learn from AI as well as from human-to-human training.

What Risks to the Confidentiality of Business Data Can DeepSeek Create?
Data input into DeepSeek and its AI output are stored in servers in China, where the government can demand access. Chinese cyber security laws and data security laws require that technology companies provide data from their servers to the Chinese government.

Putting aside personally identifiable information, the DeepSeek "privacy" policy results in government access to what would be considered confidential business information under U.S. law and under the terms of U.S. private contracts.

DeepSeek and other AI solutions expose a company to liability for breach of its confidentiality obligations. This can occur if a company holds information about its customers in its database that was provided under different confidentiality and nondisclosure agreements. The scope of a company's authorized employees can vary from those who "have a need to know," to those in a business unit, or to those in a department with personnel in multiple offices in multiple jurisdictions.

Further, confidentiality clauses often provide for the right to seek injunctive relief in a jurisdiction that is an exception to the governing law, jurisdiction, and venue designated in the agreement. While liability for breach of confidentiality is often conducted on a

contract-by-contract basis, the risk presented by AI is that it will disclose all confidential information all at once and without context.

In addition to liability for breach of confidentiality obligations, the company can then face liability for violating the injunction. In addition, there may be liability owed by the AI provider to the company, raising different contract, governing law, and jurisdictional issues.

Is DeepSeek a New Frontier in Corporate Espionage?
DeepSeek is potentially an enhanced tool for corporate espionage. If a corporate AI Acceptable Use Policy does not prohibit disclosure of sensitive corporate information, or the policy does but is not forced, then sensitive business information (including trade secrets) will be stored on Chinese servers and potentially be made available to Chinese companies seeking inside information about the technology, IP, and business plans of U.S. companies.

Significantly, this risk of disclosure may also apply to information about a company's customers. That, in turn, may subject the company to the risk of liability for breach of confidentiality obligations, as noted above.

Is DeepSeek an Engine to Reverse Engineer Corporate Strategies?
Business plans, new products, and research and development activities can be reverse engineered from data collected and stored by DeepSeek. To a person with the right skills, the identity of subcontractors, components used, RFPs for product development assistance, raw materials used, and even visits to specific facilities drawn from company travel information can be used to reverse engineer what a company is planning.

If sufficient information of this type is accessible from DeepSeek servers, then the interpretation of such information can provide a window into a company's internal business affairs and weaken that company's competitive market advantage.

Can DeepSeek Enable Cyber-Attacks on Steroids?
As discussed above, DeepSeek collects data provided to it, including about company personnel and technology. This can give malicious actors information that can provide an easier pathway to gain unauthorized access to a company's IT systems and data.

Can DeepSeek Silently Undermine Scientific Research?
In addition to historical omissions in DeepSeek, even such neutral subjects as scientific research can be affected by distorted or omitted data. While reports of scientific research may look authoritative, the research reported may be incomplete or misleading.

Even though this can be a limitation for internal Chinese research use of DeepSeek, it also can provide limitations when other parties plan to use Chinese research reported by DeepSeek to combine it with research conducted by academic and scientific institutions located in other countries.

What is the Effect of DeepSeek on Corporate AI Acceptable Use Policies?
Government agencies and U.S. companies are prohibiting the use of DeepSeek to prevent the Chinese government from gaining access to critical information and providing it to Chinese companies. Corporate internal AI Acceptable Use Policies should require adherence to regulatory requirements to protect the company against liability and regulatory sanctions.

In this connection, companies with cross-border operations must consider the requirements of the EU General Data Protection Regulation (GDPR) in formulating these policies, and face challenges because DeepSeek's practices raise significant issues that could render its use by companies as a failure to comply with GDPR requirements.

A primary risk created by DeepSeek is its extensive data collection and storage practices. It gathers a wide range of user data, including personal

information, chat histories, search queries, device and browser data, and even keystroke patterns. Such comprehensive data collection occurs without explicit user consent, thus violating many principles and rules under GDPR, including a failure or potential failure to meet the requirements for limiting the data collected and the purpose for which it is maintained. (See, Article 5, Paragraph 1).

The DeepSeek app can enable the unencrypted transmission of sensitive data as defined under the GDPR, which is covered by higher privacy requirements. This increases the risk of data breaches and fails to meet requirements for the security and confidentiality of personally identifiable information in GDPR Article 32.

DeepSeek fails to offer a Data Processing Agreement (DPA) to its corporate users, which is essential for legal data processing by third parties on behalf of a business. Without a DPA, companies cannot rely on the privilege of data processing under Article 28 of the GDPR when allowing DeepSeek access to personal information. This exposes a company to risk in connection with the use of personal information such as that belonging to its customers.

As a practical matter, companies should take note that because of privacy and security risks, European Data Protection Authorities have taken action or open investigations into DeepSeek's lack of privacy data protection. The Italian data protection authority has restricted access to DeepSeek for Italian users and has opened an investigation into DeepSeek.

The Irish Data Protection Commission has requested information about the processing of data of data subjects in Ireland. German data protection authorities are planning a joint investigation into DeepSeek, and the European Data Protection Board is coordinating a task force to ensure a standardized approach by EU countries.

For the reasons above, many corporate AI Acceptable Use Polices will prohibit or severely limit the use of DeepSeek in various ways depending upon the nature and scope of the company's business, the practical requirements not introducing risk into its business relationships with its customers and business partners.

What are the DeepSeek Wider AI Lessons?
Many of the risks and potential risks of DeepSeek identified above also apply to other AI solutions. For example, as a general principle, unreliable data can corrupt AI results in any AI tool, and such data limitations reduce the utility of AI results.

DeepSeek puts these risks in high profile because of the intentional distortions included in the training data, Chinese censorship of the Internet, and political engineering of LLM results. The wider lesson is that it is important to be aware of the assumptions and cultural and economic biases in any AI solution because they will reflect the cultural norms and economic practices inherent in the culture of the country in which the AI solution is built.

William Tanenbaum *is a partner at Moses Singer.*
Dr. Matthias Orthwein *is a partner at SKW Schwarz.*