

Navigating generative AI in M&A transactions

July 5, 2023 | Client Update | 16-minute read

The recent rise of consumer and market interest in generative artificial intelligence (GAI) tools has spurred growing interest in GAI assets from strategic acquirers and private equity investors. This client update provides a brief introduction to GAI tools and their current uses, as well as an overview of the due diligence, transactional and other commercial considerations for investors and acquirers engaging in related investment and M&A activity.

The release of OpenAI's Chat Generative Pre-Trained Transformer (ChatGPT) has led to a surge of interest in GAI as companies, consumers, and governments search for ways to leverage GAI's rapid and increasingly sophisticated content generating capabilities. Although several prominent players are at the forefront of GAI technology, there are an increasing number of competitors entering this space and the GAI market has been projected to expand at a compound annual growth rate of 39%, reaching \$422.37 billion by 2028.¹ Additionally, increasing global interest in GAI has led to an uptick in investment and M&A activity involving targets in the GAI space across varied applications and in a wide range of industries – annual global venture capital investment in GAI companies increased from \$1.59 billion in 2020 to \$4.58 billion in 2022,² with significantly greater investment expected in 2023 and beyond. As activity involving targets in the GAI space has increased, potential investors and acquirers, and their counsel, are facing novel due diligence and commercial concerns.

What is generative AI?

GAI tools utilize combinations of supervised and unsupervised machine learning algorithms and large volumes of training data to develop models that are capable of quickly producing customizable output – such as audio, visual, and textual material – in response to relatively simple prompts from users. Where more familiar non-generative AI tools process or analyze existing data, GAI tools generate new material based on their training data and the relevant user prompt. The capabilities of GAI tools continue to expand – as of this writing, GAI tools are commonly categorized by the mediums of their respective inputs and outputs, including: (i) text-to-image; (ii) text-to-3D; (iii) image-to-text; (iv) text-to-video; (v) text-to-audio; (vi) text-to-text; and (vii) text-to-code. A few prevalent examples of current GAI tools include:

- ChatGPT: an AI chatbot, meaning a large language model that uses natural language processing and machine learning to understand and generate text-based responses to user prompts;
- DALL-E: an image generator that produces images of varying resolution in response to user prompts; and
- Copilot: a GAI tool that anticipates and produces code snippets – ranging from individual lines of code to entire functions – in response to user's natural language or code prompts.

These and other GAI tools may be applied in numerous contexts and across a variety of industries, each of which may raise unique concerns for potential investors and acquirers. Examples of such use cases include: generating new molecular compounds, improving medical scanning, providing customer support services, optimizing manufacturing processes and even developing and administering personalized educational experiences.

Due diligence concerns

Transactions involving companies in the GAI space present unique due diligence challenges, including with respect to intellectual property (IP) infringement and ownership, product design, and an evolving regulatory landscape.

Infringement: Similar to traditional companies in the technology sector, IP infringement issues may present material risks to targets in the GAI space. While many of the IP infringement considerations for targets in the GAI space are similar to those for traditional technology companies, the unique characteristics of GAI tools present a number of novel issues.

- **Training data:** The creation of any GAI tool requires the use of large data sets which the applicable machine learning model uses to “learn” and improve its output. The data sets used to train machine learning models may include images, text or other data or information subject to copyright and other IP protections, as well as contractual terms (e.g., terms of use; rights of attribution). Multiple GAI tool developers are facing allegations that their respective GAI tools were developed by ingesting certain protected content without a license to use such content.³ Potential investors and acquirers of targets developing or heavily reliant on the use of GAI tools should review and understand (i) the manner in which such GAI tools have been developed and trained, and (ii) any licensing arrangements related to the acquisition and use of such training data in order to evaluate the risk of copyright infringement (and other IP claims), or claims based on breach of contract, being brought against such targets. This is of particular importance given that destruction of a GAI tool’s underlying model may be a potential remedy where a court has determined that the model was trained in an infringing manner.
- **Generated content:** Due to the nature of GAI tools, there is an inherent risk that the output of GAI tools may reflect certain characteristics of the data used in training and developing such GAI tools. Accordingly, any perceived similarity between the output of such tools and any third-party IP may entail a risk of IP infringement claims being brought against the developer or user of the applicable GAI tool. Potential investors and acquirers should consider the target’s terms of service or other platform agreements, IP laws in the relevant jurisdictions, and any applicable license agreements to evaluate the risk that any content created by the GAI tools of a prospective target may lead to claims that such content infringes upon the IP rights of third parties or violates applicable contractual terms.

Ownership and protection of IP: IP rights are critical to companies in the GAI space. Due in part to the ways in which existing laws and regulations are being applied to GAI tools, the IP rights and obligations of targets using or developing GAI tools are currently determined largely by contract. Potential investors and acquirers of such targets should closely analyze the specific mechanisms at play in producing the output of their GAI tools as well as the terms and conditions applicable to their GAI tools to evaluate whether such targets may face any challenges relating to the ownership or protection of their IP rights.

- **Protection of output:** U.S. courts have held that only output authored or invented by a human being is protectable under current U.S. IP law.⁴ Further, the U.S. Copyright Office has indicated that not all GAI tool output produced in response to user prompts is protectable.⁵ Whether any GAI tool’s output is protectable turns on the manner in which such GAI tool produces such output, including whether such output was sufficiently predictable in advance, and the specific prompts provided by the user of such GAI tool with respect to such output, amongst other considerations.⁶ As of this writing, there are no clear-cut guidelines for assessing the ownership and protectability of the output of GAI tools under U.S. IP law, and potential investors and acquirers should evaluate this risk in any context where the ownership of GAI tool output is important to the proposed transaction.
- **Terms and conditions:** For practical purposes, assuming the output of GAI tools is protectable under U.S. law, whether the output of a GAI tool is owned by the applicable GAI tool provider or by the user of such GAI tool will depend on the terms and conditions associated with the use of the applicable GAI tool. For example, while the market generally seems to be shifting to a paradigm in which the GAI tool provider assigns ownership of output to the users of such GAI tool,⁷ certain GAI tool providers continue to retain ownership and license the output to users for certain limited purposes, or alternatively require users to make certain payments (often a subscription or service fee) in exchange for ownership of the output.⁸ Additionally, the terms applicable to GAI tools frequently provide GAI tool providers certain reserved rights to both the output (including the ability to refine or improve the GAI tool using such output) and the prompts provided by users (including for testing and maintenance purposes). Important rights and provisions to consider when reviewing the terms and conditions applicable to GAI tools include:
 - the scope of rights granted to users of GAI tools with respect to the output;
 - the scope of rights retained by the GAI tool provider in any output or user prompts;

- any use restrictions or obligations with respect to the output (including, for example, limitations on publishing or commercial use of output and requirements to disclose the role of GAI tools in generating such output);
- any royalties or other fees that may be owed in the event of any secondary sale of the output;
- indemnification obligations or representations and warranties provided by the GAI tool provider; and
- obligations to limit use of such GAI tools in accordance with acceptable use policies and applicable law.

Safeguards: Companies in the GAI space generally employ technical, administrative and contractual safeguards to mitigate against a broad array of potential pitfalls arising from the use of GAI tools, including with respect to misuse, bias or errors, and content regulations. Potential investors and acquirers should evaluate such risks in the context of the use of the applicable GAI tool, and consider whether the relevant company has adopted, maintained and implemented any policies or procedures to mitigate these risks.

- **Misuse:** Due in part to the novelty and sophistication of GAI tools, there is a risk that users may employ GAI tools in ways not anticipated by targets developing or commercializing GAI tools, potentially leading to unlawful or harmful consequences. Potential investors and acquirers should review whether the applicable GAI tools were developed in a manner designed to address these and other risks, including, for example, by incorporating built-in filters, response limitations, design guardrails or other safety features incorporated in such GAI tools. Beyond GAI tool development, potential investors and acquirers should also evaluate any applicable acceptable use policies or other contractual terms designed to limit the misuse of the applicable GAI tool or guard against any legal liability or reputational damage stemming from any misuse.
- **Bias and error:** As discussed above, the output of GAI tools often reflects the characteristics of the training data used to develop such tools. If the training data used to develop a GAI tool is inaccurate, biased, or otherwise flawed, then such GAI tool and its output may in turn be inaccurate, biased, or otherwise flawed. Potential investors and acquirers should review the accuracy and reliability of GAI tools, any biases in the training data used in developing such GAI tools, and any disclosure such targets may make to their users with respect to these risks.
- **Open source software:** GAI tools are frequently trained on large data sets that may include open source software (OSS). Users of OSS, including companies using or developing GAI tools, are obligated to comply with the terms of any licenses applicable to such OSS, which may include restrictions on the use of such software or additional obligations. For example, companies using or developing GAI tools that improve, modify or distribute any OSS may be obligated under such licenses to make any output produced by such GAI tools publicly available at no cost, or to provide required attribution in connection with such output. Importantly, any restrictions applicable to OSS used by GAI tools in their development or operation may also apply to the output produced by such GAI tools, undermining users' perceived ownership of or exclusive rights to such output. Potential investors and acquirers should diligence the manner in which any OSS is used by the applicable GAI tools to evaluate (i) whether the target is in violation of any such licenses and (ii) whether the output of such GAI tools may be subject to any OSS license restrictions or obligations. In each case, potential investors and acquirers should diligence the policies and procedures implemented and maintained by such target with respect to the use of OSS in developing or commercializing GAI tools.
- **Content regulations:** Targets in the GAI space also may be subject to content regulations in varying jurisdictions that may penalize such targets for any GAI tool use that produces illicit or harmful output, including, for example, false or misleading information. Potential investors and acquirers should evaluate any applicable content regulations and whether the applicable GAI tool's content policies and practices may tarnish or impair the brand or reputation of the applicable target.

Regulatory: Regulatory activity related to GAI tools is rapidly increasing⁹ and, given the concerns related to automated decision making in the data privacy sphere¹⁰ and the cybersecurity issues raised by GAI tools, data privacy and cybersecurity regulations will be especially relevant to potential investors and acquirers of targets developing or heavily reliant on the use of GAI tools.

- **Data privacy:** Due to the scale of the data sets used to train machine learning models, and the costs associated with screening such large volumes of data for any personal information, there is a risk that the training data used to develop GAI tools may incorporate personal or sensitive information. Additionally, users of GAI tools may submit certain personal information (related to themselves or others) in their prompts to GAI tools. Any personal or other sensitive information used in the training of GAI tools or contained in such user prompts may be incorporated in the output of such tools. Any of the foregoing uses or perceived uses of personal information may constitute a violation of applicable privacy laws or regulations, particularly in jurisdictions that restrict automated decision making or other automated processing of personal information. Potential investors and acquirers should consider, amongst other concerns, (i) whether appropriate privacy protections were implemented in acquiring or refining the relevant training data for the development of the applicable GAI tools, (ii) whether the users of such GAI tools are party to any privacy

policies or any other terms or conditions governing the processing of personal information in connection with such GAI tools, and (iii) whether applicable targets have implemented other practical or contractual data privacy protections (e.g. data anonymization and encryption, data subject access request mechanisms and data processing agreements) and acted in accordance with such policies.

- **Cybersecurity:** GAI tools also provide threat actors with powerful new vectors of attack, including, for example, using GAI to execute sophisticated phishing attacks, producing false or misleading information (including “deepfakes”), rapidly developing malware or other malicious software, and exploiting vulnerabilities in source code produced by GAI tools. Potential investors and acquirers should consider whether the applicable GAI tool has been designed to prevent the use of such GAI tools for these purposes, and whether the target has any contractual or other remedies available in the event such GAI tool is used in any cyberattack. Additionally, GAI tools are trained on large volumes of publicly available software (including OSS) that may contain security weaknesses, vulnerabilities, errors and other flaws, increasing the risk that the source code produced by a GAI tool may also contain these defects. Accordingly, potential investors and acquirers should consider the cybersecurity risks associated with the use of such targets’ GAI tools, and whether any specific cybersecurity protections are warranted in the context of the applicable transaction.

Potential investors and acquirers of targets developing or heavily reliant on the use of GAI tools should be aware that additional due diligence may be required depending on the GAI tools at issue and the specific industries in which they may be applied, including related to antitrust, executive compensation, tax, investment management and regulatory or compliance concerns.

Transaction agreement considerations

Asset transactions: With asset sales (as opposed to stock or equity transactions), acquirers should ensure the asset purchase agreement fully covers all of the technology and data required to operate and continue the development of the applicable GAI tools owned or used by the target. Acquirers should also ensure the agreement is drafted such that purchased assets capture all of the target’s IP rights in the applicable GAI tool and all rights to commercialize the target’s use of such GAI tool and its output, including all related GAI tool technology, software, algorithms, models and datasets. In an asset sale, the purchase agreement should also specify any licenses being transferred in connection with the purchase, and prospective acquirers should evaluate whether such licenses include the full set of rights needed for underlying training data and should be mindful of any restrictions regarding the use of applicable GAI tools, including with respect to industry, territory, time period, or type of use (e.g. commercial vs. non-commercial).

Representations and warranties: The operation and continued development of GAI tools requires the use of many supporting technologies and large data sets. Accordingly, and in light of the diligence considerations outlined above, potential investors and acquirers may want to include representations and warranties in the relevant transaction agreement that are specifically designed to backstop their due diligence (including with respect to the nature and provenance of the training data and any safeguards employed in the development of such GAI tool), in addition to a typical, fulsome set of IP, IT, privacy and cybersecurity representations and warranties.

Interim operating covenants: In addition to traditional covenants to maintain IP assets and not to amend or terminate material agreements, given the pace of change in the GAI market, practitioners should consider whether to include covenants restricting the target’s ability to materially change, without the acquirer’s consent (except where the change is required by applicable law), (i) the nature of the training data used by the target, (ii) the terms of use or other agreements governing the target’s use or development of GAI tools or related assets, or (iii) the target’s data privacy or security policies or practices.

Due to the increasing focus of regulators on GAI tools,¹¹ potential investors and acquirers also may wish to pay particular attention to the target’s obligations to provide notice of any investigations or regulatory inquiries, and considering bargaining for rights to participate in any regulatory discussions related to the target’s use or development of such GAI tools (to the extent such participation would be permitted under applicable law).

Recourse: Depending on the risks identified in due diligence and the availability of representation and warranty insurance, potential investors and acquirers should consider incorporating specific indemnities in the relevant transaction agreement to address any such known risks, including with respect to the use and acquisition of training data, pending or threatened litigation proceedings, and any other salient concerns identified in due diligence.

Post transaction considerations

Applicable terms of use, terms of service, and other written agreements are often central to determining the IP rights held by GAI tool providers in the output of such tools. Acquirers should pay careful attention to a target’s scope of ownership

rights and licensing rights with respect to such tools and their output when implementing any plan to integrate such GAI tools or their output with the acquirer's existing operations. Given the current uncertainty regarding the protectability of the output of GAI tools, combining such output with otherwise protectable content may present unique risks.

Further, acquirers should evaluate the reputational considerations from integrating GAI tool output with their own proprietary tools or systems, especially given the risks related to potential inaccurate, biased or flawed GAI tool output as described above, and potential requirements to disclose the role of GAI in producing any such output.

If you have any questions regarding the matters covered in this publication, please reach out to any of the lawyers listed below or your usual Davis Polk contact.

Frank Azzopardi

+1 212 450 6277
frank.azzopardi@davispolk.com

Matthew J. Bacal

+1 212 450 4790
matthew.bacal@davispolk.com

David R. Bauer

+1 212 450 4995
david.bauer@davispolk.com

Pritesh P. Shah

+1 212 450 4147
pritesh.shah@davispolk.com

Samantha Lefland

+1 212 450 4807
samantha.lefland@davispolk.com

Christopher C. Woller

+1 212 450 3428
christopher.woller@davispolk.com

This communication, which we believe may be of interest to our clients and friends of the firm, is for general information only. It is not a full analysis of the matters presented and should not be relied upon as legal advice. This may be considered attorney advertising in some jurisdictions. Please refer to the firm's privacy notice for further details.

- ¹ \$422.37+ Billion Global Artificial Intelligence (AI) Market Size Likely to Grow at 39.4% CAGR During 2022-2028, Bloomberg, accessed July 3, 2023. Available from: <https://www.bloomberg.com/press-releases/2022-06-27/422-37-billion-global-artificial-intelligence-ai-market-size-likely-to-grow-at-39-4-cagr-during-2022-2028-industry>.
- ² *Generative AI startups jockey for VC dollars*, PitchBook, accessed July 3, 2023. Available from: <https://pitchbook.com/news/articles/Amazon-Bedrock-generative-ai-q1-2023-vc-deals>.
- ³ See e.g., *Getty Images (US), Inc. v. Stability AI, Inc.*, 1:23-cv-00135, (D. Del.); and *DOE 1 v. GitHub, Inc.*, 4:22-cv-06823, (N.D. Cal.).
- ⁴ See e.g. *Thaler v. Hirshfeld*, 558 F. Supp. 3d 238 (E.D. Va. 2021).
- ⁵ Letter from United States Copyright Office, RE: Zarya of the Dawn (Registration #VAu001480196), February 21, 2023. Available from: <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>.
- ⁶ *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence*, United States Copyright Office. Updated March 16, 2023, accessed July 3, 2023. Available from: https://copyright.gov/ai/ai_policy_guidance.pdf
- ⁷ *Terms of Use*, Open AI, Updated March 14, 2023, accessed July 3, 2023. Available from: <https://openai.com/policies/terms-of-use>.
- ⁸ *Terms of Service*, Midjourney, Updated June 8, 2023, accessed July 3, 2023. Available from: <https://docs.midjourney.com/docs/terms-of-service>.
- ⁹ *The AI Act, Development*, European Commission, accessed July 3, 2023. Available from: <https://artificialintelligenceact.eu/developments/>; *Artificial Intelligence Act*, European Commission, accessed July 3, 2023. Available from: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf
- ¹⁰ Automated individual decision-making, including profiling, Article 22, Regulation (EU) 2016/679 (General Data Protection Regulation) (GDPR). Available from: <https://gdpr-info.eu/art-22-gdpr/>.

¹¹ *FTC Chair Khan and Officials from DOJ, CFPB and EEOC Release Joint Statement on AI*, Federal Trade Commission (FTC), accessed July 3, 2023. Available from: <https://www.ftc.gov/news-events/news/press-releases/2023/04/ftc-chair-khan-officials-doj-cfpb-eoc-release-joint-statement-ai>.