LAW.COM

Expert Opinion **Artificial Intelligence**

AI and the Law, Part 3: Whose Model Is It Anyway? The Intellectual Property Questions No One Thought to Ask

In Part 3 of his "AI and the Law" series, David J. Partida discusses the unresolved legal questions surrounding ownership, use and protection of AI-generated work and training data.

September 08, 2026 at 09:31 AM By **David J. Partida**

Organizations assuming their AI-assisted work automatically receives full copyright protection may be relying on legal assumptions that have yet to be tested. (Credit: Artram/Shutterstock.com)

Somewhere in most legal departments sits an AI vendor agreement negotiated by a technology procurement team, reviewed for commercial terms, and signed without anyone asking a handful of questions that are quickly becoming difficult to ignore.

Who owns the model a vendor fine-tuned using a company's proprietary data? Who owns the research, analysis, documentation, or software code the AI produces? And if a model's outputs

reflect confidential information embedded in its training data and that information later appears in a competitor's system, what legal remedies are available?

These are no longer hypothetical questions. They are already the subject of active litigation, with several cases moving beyond the pleading stage.

Training Data Liability: The Cases Are Moving

Whether training AI models on copyrighted works constitutes infringement is now being litigated in federal courts across the country. Publishers, authors, news organizations, and other rights holders have filed suits alleging that the use of their works in AI training datasets without permission or compensation violates copyright law. Several of those claims have already survived motions to dismiss.

Because the Second Circuit—which governs New York federal courts—has one of the country's most developed bodies of fair use jurisprudence, its eventual application of the four-factor fair use test to AI training will have significant implications for organizations that have fine-tuned models using content they did not create.

One of the central questions remains unresolved: does a license permitting the use of underlying content also permit its use to train an AI model? Most existing licensing agreements were negotiated before that question existed. Organizations that have built AI capabilities using licensed content without analyzing whether their agreements extend to model training may be carrying legal risk they have not yet measured.

The Ownership Gap

The U.S. Copyright Office has taken a position that is becoming increasingly important: works generated entirely by AI, without sufficient human creative authorship, are generally not eligible for copyright protection. At the same time, the Office has approved registrations for works involving substantial AI assistance where meaningful human authorship was present. Where that line ultimately falls remains unsettled.

For organizations relying on AI to generate commercially valuable material—research, legal analysis, marketing content, software code, technical documentation, and similar work—the consequences are significant. If an AI-generated work cannot be protected by copyright, competitors may be free to copy, adapt, or republish it. The commercial value of those outputs may therefore be considerably less secure than many organizations assume.

Meaningful human involvement makes copyright protection more likely, but exactly how much human contribution is required remains an evolving question. Organizations assuming their AI-

assisted work automatically receives full copyright protection may be relying on legal assumptions that have yet to be tested.

The Trade Secret Problem

Trade secret law presents a different set of challenges. If a model trained on proprietary information can be prompted in ways that reveal embedded knowledge—or if its outputs subtly reproduce confidential information that was never intended to be disclosed—existing trade secret doctrines provide only uncertain answers.

The Defend Trade Secrets Act and New York's common law developed around familiar forms of misappropriation: an employee taking a customer list, an engineer disclosing confidential designs, or a competitor acquiring proprietary information through improper means. They were not designed to address situations where an AI model absorbs information during training and later reproduces aspects of that knowledge without any deliberate disclosure by a human actor.

Some courts have begun considering whether trade secret protection extends to AI training data, model weights, or other aspects of machine learning systems. The doctrine, however, is still developing, and the outcomes remain difficult to predict.

Organizations providing proprietary information to AI vendors for training, fine-tuning, or evaluation should therefore treat those disclosures no differently than any other exchange of confidential information. The governing agreements should clearly define permissible uses of the data, ownership rights, retention obligations, and what happens to both the data and any derivative models when the relationship ends.

The Vendor Contract Gap

Many AI vendor agreements say remarkably little about ownership of fine-tuned model weights, AI-generated outputs, derivative works, or information learned from a customer's proprietary data. Most were negotiated as technology services agreements rather than intellectual property agreements. When disputes arise, courts will often be left applying general contract and IP principles to factual situations the parties never anticipated.

The consequences can cut in several directions. A company that gave a vendor access to years of proprietary business data may discover it effectively subsidized improvements that benefit other customers. A business built around AI-generated reports or software may find that ownership of those outputs is less certain than expected. And creators whose works were incorporated into training datasets without meaningful consent may ultimately benefit from legal standards emerging through the current wave of copyright litigation.

The Document That Matters

In many cases, the most important document will be the vendor agreement itself—specifically, what it says, and what it fails to say, about training data, model weights, outputs, derivative works, and ownership rights.

Data breach litigation followed a similar pattern. The vendor security agreement often became the critical document because it allocated responsibility among multiple parties, many of whom had signed standard terms without carefully considering how risk was actually being distributed. AI intellectual property disputes are likely to follow a similar path. The organizations that negotiated these provisions thoughtfully may find themselves in a very different position from those that simply accepted market-standard language.

The moment of recognition usually comes later: when an AI-generated work becomes commercially valuable, when a licensing dispute emerges over training data, or when a competitor's product begins producing outputs that look uncomfortably familiar. By then, the legal questions themselves are not especially complicated. The difficulty is that they should have been asked—and answered—before the dispute arose.

David J. Partida *is a litigator at Wilk Auslander, where his practice includes complex commercial litigation and emerging technology disputes. This article is for informational purposes and does not constitute legal advice.*

Page printed from: https://www.law.com/newyorklawjournal/2026/09/08/ai-and-the-law-part-3-whose-model-is-it-anyway-the-intellectual-property-questions-no-one-thought-to-ask/print/

NOT FOR REPRINT

© 2026 ALM Global, LLC, All Rights Reserved. Request academic re-use from www.copyright.com. All other uses, submit a request to asset-and-logo-licensing@alm.com. For more information visit Asset & Logo Licensing.