

Blockchain, AI Training Data, and Protecting Intellectual Property in the Next Deal

06 May 2026

Client Updates

The intersection of blockchain technology and AI training data disputes continue to generate significant market attention. The following offers practical perspective for companies seeking to protect and monetize their intellectual property in AI-related transactions, licensing, and investment. Copyright disputes have been prominent in the Courts concerning whether a company's data has been used to train institutional and foundational AI models.

Application of Blockchain for IP Protection

Blockchain has been applied to tracking AI training data given that blockchain is an immutable ledger that can help establish what data was used to train a model and when. In these instances, blockchain's value is as an IP audit and compliance layer embedded within commercial AI deals. Intellectual property diligence in AI transactions centers on the key inquiry of what data went into the model, and does the model have rights to that data for use in training. Blockchain-based logs provide tamper-resistant documentation of dataset provenance, the specific IP licenses in effect at the time of ingestion, and model version histories tied to particular copyright-cleared datasets. This kind of defensibility shortens IP diligence cycles and reduces the risk of post-closing indemnification claims related to training data disputes, particularly where the content is sourced from multiple data holders or updated on a rolling basis.

IP Representations and Warranties

In certain transactions, AI model developers have been increasingly expected to make representations and the intellectual property embedded in their models. This is stronger than the prior required representations limited to training sets as free of third party IP. Increasingly, as part of these transactions and representations, inquiries are made into auditable controls around such copyright and licensing compliance. Blockchain infrastructure can meaningfully support those IP representations, giving counterparties greater confidence that the company's core technology does not carry hidden infringement exposure. This is especially important in M&A contexts that have IP as a significant asset driving valuation.

For content owners, creators, and other IP holders licensing their works for AI training, the value of blockchain is not in detecting unauthorized use after the fact. Rather, it lies in embedding audit and logging requirements directly into IP license agreements and require licensees to record on-chain what content was ingested, under which license grant, and for what permitted purpose. This shifts the compliance burden to the front end of the relationship, creates a contemporaneous evidentiary record that can support license claims, reducing reliance on costly post-hoc forensic analysis of model outputs.

Impact on IP Strategy

Blockchain is valuable when treated as IP governance infrastructure. It can help evidence that your organization identified, tracked, and managed your own training data and third-party IP risk properly. Doing so will also reduce friction as markets are already beginning to demand clean-chain IP provenance in AI assets in term sheets, license negotiations, and diligence requests across the AI ecosystem.

ABOUT BAKER BOTTS L.L.P.

Baker Botts is an international law firm whose lawyers practice throughout a network of offices around the globe. Based on our experience and knowledge of our clients' industries, we are recognized as a leading firm in the energy, technology and life sciences sectors. Since 1840, we have provided creative and effective legal solutions for our clients while demonstrating an unrelenting commitment to excellence. For more information, please visit [bakerbotts.com](https://www.bakerbotts.com).

Related Professionals



Ali Dhanani

Partner