4 Emerging Approaches To AI Protective Order Language

By **Joel Bush** (May 12, 2026)

Joel Bush

Federal courts are rapidly developing protective order language to address a new and significant risk: that confidential discovery materials will be exposed to generative artificial intelligence platforms that retain user inputs, use them for training or share them with third parties.

Between 2025 and early 2026, at least five federal district courts have issued or analyzed specific protective order provisions restricting the use of generative AI platforms with protected materials. These decisions establish that proactive AI-specific provisions are no longer optional — they must be standard practice.

AI Protective Order Language: Four Key Models

Federal courts have developed four principal models for addressing generative AI disclosure risks in protective orders.

1. The Blanket Prohibition

In Fiorito v. Metropolitan Council, decided in July 2025, the U.S. District Court for the District of Minnesota entered a protective order providing: "No party may upload or otherwise disclose Confidential Data to any generative artificial intelligence platform, such as (but not limited to) ChatGPT."[1]

This language is brief, clear and easy to apply, but it is also the bluntest instrument in the judicial tool kit. By treating all generative AI platforms identically, the blanket prohibition fails to account for the meaningful difference between open consumer platforms — where user inputs are retained, used for training and potentially disclosed to third parties — and enterprise or closed AI tools that contractually prohibit all of those practices.

The practical consequence is that parties subject to this order cannot use any AI tool with confidential materials, even a secure, enterprise-grade platform that offers data protections far exceeding what most human vendors provide. For a pro se plaintiff like Michael Fiorito, operating without a law firm's document review infrastructure, this restriction is particularly burdensome.

That said, the blanket prohibition has a distinct advantage: It is virtually impossible to violate unknowingly. There is no ambiguity about what is permitted, no contractual analysis required and no risk that a party will use an AI tool under the mistaken belief that its data protections are sufficient.

For courts managing cases with unsophisticated parties or where the volume of confidential material is small enough that AI-assisted review is unnecessary, this model remains a sensible default.

2. The Data Rights Approach

In both U.S. v. Jackson[2] and U.S. v. Hernandez,[3] — which were decided in June 2025

and November 2025, respectively — the U.S. District Court for the Western District of Washington entered protective orders providing: "The Protected Material shall not be loaded into an artificial intelligence (AI) tool or model, generative or not, where rights to the data are retained by any party that has not explicitly agreed to be covered by the protective order."

This formulation is notably more sophisticated than the blanket prohibition because it applies to AI tools of any kind — generative or otherwise — and pivots the inquiry from the nature of the platform to the nature of the data rights. The key question is whether the AI provider has agreed to be bound by the protective order's requirements, permitting use of tools that commit contractually to data protections aligned with the order.

In practice, this approach places the burden on the party using AI to confirm that the provider has explicitly agreed to be covered by the protective order. This is a meaningful but potentially difficult-to-satisfy requirement, because most commercial AI providers will not agree to be bound by a case-specific court order.

The data rights approach is particularly well suited to criminal cases (where it originated) because the stakes of data exposure are highest — e.g., informant recordings, personally identifiable information and grand jury material — and the universe of parties is small enough that tracking provider compliance is manageable.

It is also the most forward-looking of the four models because it does not limit itself to generative AI, capturing analytical and classification tools that increasingly process litigation materials.

Its chief weakness is enforceability: Determining whether an AI provider has retained rights to the data requires parsing the provider's terms of service, which may be ambiguous, subject to unilateral amendment or governed by foreign law.

3. The Contractual Safeguards Approach

The most detailed framework was crafted by the U.S. District Court for the District of Colorado in Morgan v. V2X Inc., which was decided on March 30.[4] The court rejected both parties' proposals and entered its own language:

> No party or authorized recipient may input, upload, or submit CONFIDENTIAL Information into any modern artificial intelligence platform, including any generative, analytical, or large language [model-based] tool ('AI'), unless the AI provider is contractually prohibited from: (1) storing or using inputs to train or improve its model; and (2) disclosing inputs to any third party except where such disclosure is essential to facilitating delivery of the service. Where disclosure to a third party is essential to service delivery, any such third party shall be bound by obligations no less protective than those required by this Order. In addition, the AI provider must contractually afford the party or authorized recipient the ability to remove or delete all CONFIDENTIAL information upon request. A party intending to use AI that it contends meets these requirements must retain written documentation of these contractual protections.

The court observed that the practical effect is to bar use of "most, if not all, mainstream low-to-no-cost AI to process Confidential Information," including "standard ChatGPT, Claude, Gemini, or similar platforms."

The court in Morgan rejected the plaintiff's proposed language permitting uploads to AI operating in a "secure, closed-circuit environment," finding that it addressed cybersecurity concerns about unauthorized access "rather than the distinct risks associated with today's widely available AI platforms."

It also rejected the defendant's proposed language as "over-engineered" and "crafted to fit the precise bounds of Defendant's contractual engagement with AI providers," resulting in provisions that were vague from the opposing side's perspective. The court further cautioned parties against overdesignating information as confidential in light of this restriction, and ordered the plaintiff to disclose the name of any AI platform previously used with confidential materials.

Of the four models, the contractual safeguards approach is the strongest candidate for a default standard in complex civil litigation. Unlike the blanket prohibition, it does not foreclose AI use entirely. And unlike the data rights approach, it does not require the AI provider to agree to be bound by the protective order itself — a condition few providers will accept.

Instead, it focuses on four discrete, verifiable contractual commitments: (1) no training on inputs, (2) no third-party disclosure except for essential service delivery, (3) downstream protections on any necessary third-party disclosures, and (4) deletion rights. These requirements map closely to the data-processing provisions found in enterprise AI agreements, making compliance achievable for parties willing to invest in enterprise-tier tools.

4. The Notice-and-Secure-Environment Approach

The protective order established by the U.S. District Court for the District of Kansas in Jeffries v. Harcros Chemicals Inc., which was decided on March 25, presents the most comprehensive AI framework.[5]

The original protective order required that a party intending to use an AI tool must:

- "[P]rovide other parties with notice and an opportunity to object";
- "[E]nsure that the AI Tool is used in a secure environment and that Confidential Information is not used to train or improve any AI Tool except one used exclusively in this action";
- Ensure that any confidential information used to train an AI tool is destroyed at the conclusion of litigation and not made accessible to unauthorized persons; and
- "[E]nsure all confidential information is deleted at the conclusion of [the] action."

The court subsequently granted a motion to expand these AI restrictions to all discovery materials — including nonconfidential documents — prohibiting their input into open-loop or public generative AI tools entirely.

The court found good cause based on the practical impossibility of clawing back data from open AI tools, security risks of a centralized repository at unprecedented scale, potential General Data Protection Regulation violations and the critical infrastructure designation of some defendants' operations.

Additionally, the court rejected First Amendment objections, noting the order was limited to the pretrial discovery context and did not prevent public dissemination of nonconfidential information from other sources.

The Jeffries framework is the most comprehensive of the four models and the only one extended beyond confidential materials to cover all discovery. Its notice-and-opportunity-to-object mechanism introduces a procedural safeguard no other model provides: Opposing parties can evaluate and challenge the specific AI tool before it touches any discovery material. This makes the model uniquely well suited to cases involving critical infrastructure, national security, regulated industries or GDPR-subject entities.

However, the Jeffries model is also the most resource-intensive to administer. The notice requirement creates a procedural checkpoint that can slow case progression, and the requirement to destroy AI-trained models at the conclusion of litigation may be impractical if a party has used a tool that commingles training data across matters.

For routine commercial disputes, this framework is likely overkill, but for high-stakes cases with sensitive data, it represents the gold standard.

Choosing the Right Model: A Practical Framework

No single model is optimal for every case. The choice depends on the sensitivity of the materials, the sophistication of the parties, the volume of discovery and the industry context.

For routine commercial litigation with moderate discovery volumes and represented parties on both sides, the contractual safeguards model should be the default starting point. It provides the clearest, most enforceable standard while preserving the ability to use enterprise AI tools, which is a practical necessity in an era when AI-assisted document review is becoming the norm rather than the exception.

For criminal cases, particularly those involving sensitive investigative materials, the data rights model offers a streamlined approach that integrates naturally into the standard criminal discovery protective order framework, as demonstrated by its adoption across multiple cases in the Western District of Washington.

For cases involving critical infrastructure, GDPR-regulated entities or national security concerns, the notice-and-secure-environment model provides the most robust protection, even at the cost of increased procedural complexity.

And for cases with pro se litigants, limited confidential materials or courts seeking maximum simplicity, the blanket prohibition remains a pragmatic choice: a bright-line rule that requires no further analysis.

Practitioners should also consider hybrid approaches. A protective order could adopt the contractual safeguards model as the baseline for confidential materials while incorporating the notice-and-secure-environment model's notice requirement for particularly sensitive categories of documents, such as trade secrets, medical records or materials subject to foreign data protection regimes.

What You Should Be Doing Now

Given the rapid development of this area, practitioners should take the following steps.

Proactively propose AI-specific provisions; do not wait for the other side to raise the issue.

Courts have demonstrated a willingness to impose AI restrictions even when the parties have not initially included them, and the Morgan court went further by rejecting both parties' proposals and crafting its own language from scratch.

The lesson is clear: If you do not propose protective order language that addresses AI, the court may impose terms that are either more restrictive than you need or less protective than your client requires.

The most effective approach is to propose the contractual safeguards framework as a starting point and negotiate from there. This positions your client as thoughtful and proactive while establishing a standard that most enterprise AI tools can satisfy.

Waiting to address AI until a dispute arises mid-discovery is a losing strategy — by that point, protected materials may already have been submitted to a noncompliant platform, creating exposure that no subsequent order can cure.

Audit existing protective orders for AI gaps, because silence may not protect you.

Submitting protected materials to AI platforms that retain data rights may independently violate a broad existing protective order, as the Morgan court observed. But most protective orders entered before mid-2025 contain no AI-specific language.

Practitioners should review every active protective order in their litigation portfolio and assess whether the existing confidentiality provisions cover AI platform disclosures. If the order prohibits disclosure to third parties without exception, an argument exists that AI platform submissions are already covered — but that argument has not been tested in most jurisdictions.

The safer course is to negotiate an amendment incorporating AI-specific provisions before the issue becomes contested.

Deploy the Jeffries framework strategically in sensitive cases — but understand its limits.

In cases involving critical infrastructure, GDPR-subject entities or heightened data sensitivity, the Jeffries court's expansion of AI restrictions to all discovery materials provides powerful precedent. But practitioners should be deliberate about when to invoke this framework.

The Jeffries expansion was grounded in case-specific facts: The defendants operated in a critical infrastructure sector, were subject to the GDPR and the court found the open/closed AI distinction more meaningful than the confidential/nonconfidential designation for data protection purposes. Attempting to invoke Jeffries-style restrictions in a routine breach of contract case risks being characterized as obstructionist.

The strongest use of this framework is defensive: If your client's documents are being produced and you want to ensure opposing counsel cannot submit them to open AI platforms for large-scale analysis, Jeffries gives you a well-reasoned basis for seeking that

protection. The court's rejection of First Amendment objections — holding that restrictions were limited to the pretrial discovery context — removes one of the strongest potential counterarguments.

Build and maintain an AI compliance record — because courts are starting to ask.

Under the framework established in the Morgan case, a party intending to use AI must retain written documentation of contractual protections, and, specifically in the Morgan case, the court ordered the plaintiff to disclose the name of any AI platform previously used with confidential materials.

Practitioners should maintain a running log of every AI platform used in connection with any matter involving protected materials, including the platform name, date of use, applicable contractual protections and confirmation that the platform satisfies the relevant protective order's requirements. This documentation should be treated with the same rigor as privilege logs.

When a court asks — and the trend suggests courts increasingly will — the ability to produce a comprehensive compliance record on short notice will distinguish the diligent practitioner from the one scrambling to reconstruct a history of AI use after the fact.

Joel D. Bush is a partner at Kilpatrick Townsend & Stockton LLP.

[1] Fiorito v. Metropolitan Council, 2025 WL 1806612 (D. Minn. July 1, 2025).

[2] United States v. Jackson, 2025 WL 1666249 (W.D. Wash. June 12, 2025).

[3] United States v. Navarro Hernandez, 2025 WL 3290753 (W.D. Wash. Nov. 26, 2025).

[4] Morgan v. V2X, 2026 WL 864223 (D. Colo. Mar. 30, 2026).

[5] Jeffries v. Harcros Chemicals Inc., 2026 WL 820218 (D. Kan. Mar. 25, 2026).