Contracting with AI Vendors

A Practical Guide for Lawyers

Colin S. Levy

2026

This document is for informational purposes only and does not constitute legal advice.

About the Author

Colin S. Levy is a legal technology advocate, writer, and advisor who works at the intersection of law, technology, and business. With experience spanning in-house legal roles, legal technology companies, and legal operations, he brings a practical perspective to how legal teams can adopt and govern emerging technologies responsibly.

Colin writes and speaks extensively on legal innovation, artificial intelligence in legal practice, and the evolving role of legal professionals in a technology-driven landscape. His work focuses on helping legal teams move beyond the hype cycle to make sound, informed decisions about the tools they use and the workflows they build.

He is the author of *The Legal Tech Ecosystem* and editor of the *Handbook of Legal Tech*, and a regular contributor to publications covering legal technology and operations. He advises organizations on responsible AI adoption, legal operations strategy, and the practical governance frameworks that make innovation sustainable.

This guide is part of a series on AI and law that includes *AI for Lawyers*, *AI for Legal Teams*, *AI Agents Data Handling and Cybersecurity Guide*, *AI in the Courtroom*, *Human Judgment and AI in Legal Practice*, and the *AI Implementation Playbook for Legal Teams*.

Table of Contents

Part One

Why AI Vendor Contracts Are Different

AI vendor contracts are not just SaaS (Software-as-a-Service) agreements with a new label. They raise issues that traditional software contracts were never designed to address: who owns the output, whether your data trains someone else's model, what happens when the AI gets it wrong, and how you measure performance for a tool that produces different results every time it runs.

Industry analyses have found that the vast majority of AI vendors claim broad data usage rights in their standard terms. Only 17% commit to regulatory compliance warranties, and 88% impose liability caps that leave customers bearing most of the risk from AI-specific failures.[1] Most vendor contracts are written to protect the vendor. Your job is to close those gaps before you sign.

This guide is for lawyers negotiating contracts with AI solution providers, whether you are procuring AI tools for your own firm, advising clients on vendor selection, or reviewing AI agreements as part of due diligence. It provides specific clause language, illustrated redlining examples, and negotiation tactics grounded in what is actually happening in these deals right now.

A note on scope: this guide focuses on the contract itself, not on whether to adopt AI. It assumes you have already decided to evaluate or procure an AI tool and need to get the legal terms right.

How to use this guide: Each section identifies a specific contract risk, explains why it matters in practical terms, and provides concrete clause language you can use as a starting point. The redline examples show typical vendor language alongside revised language that better protects your interests. Adapt the specific thresholds and timeframes to your deal, but do not weaken the structural protections.

Part Two

The Clauses That Matter Most

Traditional software agreements focus on licensing, uptime, support, and data security. AI vendor agreements share those concerns but add an entirely new layer. The following clauses either do not exist in standard SaaS contracts or take on fundamentally different significance in the AI context.

Training Data Restrictions

This is the single most important clause in any AI vendor contract. The question is simple: can the vendor use your data, your clients' data, or the outputs generated from that data to train, retrain, or improve its AI models? If the answer is yes, you have a confidentiality problem, a privilege problem, and potentially a regulatory problem.

Most vendor standard terms permit broad data usage. Some frame it as "product improvement" or "service enhancement." Others bury it in definitions of "aggregated data" or "de-identified data" that may not actually be anonymous in any meaningful sense.

The Bonterms AI Standard Clauses (Version 1.0), an open-source industry template, offer three alternative positions on training data: no training permitted, training permitted only on de-identified and aggregated data, or training permitted with notice and opt-out.[2] The strictest alternative aligns with the revised language below. Whichever Bonterms alternative you select, the key is replacing the kind of broad vendor language shown here with an explicit, negotiated restriction.

VENDOR LANGUAGE (DELETE):

Customer acknowledges and agrees that Vendor may use Customer Data, including inputs, outputs, and usage data, in aggregated or de-identified form, to improve, develop, and enhance the Service and Vendor's other products, features, and machine learning models. Vendor may also use Customer Data to generate anonymous and aggregate statistics regarding use of the Service.

REVISED LANGUAGE (INSERT):

Vendor shall not use any Customer Data (including inputs, outputs, prompts, or derivatives) for model training, retraining, fine-tuning (adapting the model to specific data), or product improvement without Customer's prior written consent. Vendor may use only aggregated, anonymized usage statistics that cannot be reverse-engineered to identify Customer or any individual.

The Subprocessor Problem

Many AI vendors build their products on top of third-party AI platforms. A legal research tool might use OpenAI's GPT as the underlying model, with Microsoft Azure as the cloud infrastructure, and a third-party vector database to store embeddings (numerical representations) of your documents. Your data flows through all three entities. Your contract with the primary vendor must explicitly: identify all subprocessors and their roles; require that each subprocessor is bound by the same data restrictions (no training on your data); give you the right to object to new subprocessors with 30 days' notice; and require the vendor to terminate use of any subprocessor that violates data restrictions. Do not assume the vendor has negotiated these terms with its subprocessors. Ask.

Output Ownership and IP Rights

Who owns what AI generates? This question has no settled legal answer, and it affects every practice area. If an AI tool generates a memo analyzing a statute or a contract clause, and you rely on it in advice to your client, does the vendor retain any rights in that output? Can the vendor use your work product to improve a tool that serves your competitors? Most vendor contracts sidestep this with vague language. The U.S. Copyright Office has taken the position that purely AI-generated content is not copyrightable, but content created by humans using AI tools may be.[3] Your contract needs to address ownership regardless of how copyright law evolves.

VENDOR LANGUAGE (DELETE):

As between the parties, Customer retains ownership of Customer Data as submitted to the Service. Vendor retains all rights, title, and interest in and to the Service, including all improvements, modifications, derivative works, and any models, algorithms, or other technology developed or enhanced through operation of the Service, whether or not informed by Customer Data.

REVISED LANGUAGE (INSERT):

Customer owns all output generated by the Service using Customer Data. Vendor retains no rights in output and shall not use output for any purpose (including training) without Customer's written consent. Vendor retains ownership of the underlying Service, models, and algorithms, excluding any Customer Data or output.

Model Drift and Maintenance

AI models degrade over time in ways that are hard to detect. Industry surveys consistently find that a majority of businesses observe performance declines in deployed AI models that lack ongoing monitoring.[4] Example: a contract review tool that correctly identifies liability caps 95% of the time when deployed may degrade to 87% accuracy six months later without any warning, because the model's performance characteristics shift as it processes new data. Unlike traditional software, where a bug either exists or does not, AI model performance exists on a spectrum and can drift silently. Address this in the contract.

VENDOR LANGUAGE (DELETE):

Vendor shall use commercially reasonable efforts to maintain, update, and improve the Service. Vendor may, in its sole discretion, modify, retrain, or replace the model or models underlying the Service at any time without notice. Such modifications shall not constitute a material change to the Service.

REVISED LANGUAGE (INSERT):

Vendor shall monitor model performance continuously and maintain accuracy at or above the thresholds defined in Exhibit A. If performance degrades below threshold for two consecutive measurement periods, Vendor shall retrain the model at no additional cost within 30 days. Vendor shall provide 30 days' advance notice of any material changes to model versions and shall conduct regression testing before deployment.

Algorithmic Transparency and Audit Rights

You cannot evaluate what you cannot see. Audit rights in AI contracts go beyond traditional SOC 2 compliance. You need the right to understand how the model works, what data trained it, how it is tested for bias, and how it performs over time. The California AI Transparency Act (SB 942), with an operative date of August 2026, mandates certain transparency measures for providers with over one million monthly users.[5]

VENDOR LANGUAGE (DELETE):

Vendor shall provide Customer with access to standard reporting dashboards reflecting Service usage metrics, including volume of queries processed and system availability. Additional reporting, documentation regarding model architecture, training methodology, or internal testing is proprietary and not included in the Service.

REVISED LANGUAGE (INSERT):

Customer may audit Vendor's AI systems on reasonable notice (minimum 15 business days). Vendor shall provide access to model cards describing architecture, training data provenance, accuracy benchmarks, bias testing methodology, and ongoing monitoring results. Vendor shall maintain and update this documentation at least quarterly.

Bias and Fairness Obligations

If the AI tool is used for any decision that affects individuals, whether hiring, credit, housing, legal services eligibility, or benefits determination, bias is not a theoretical concern. It is a litigation risk. In *Mobley v. Workday* (N.D. Cal. 2024), a federal court allowed disparate impact discrimination claims to proceed directly against the AI vendor, not just the employer.[6] The potential class covered over one billion rejected applications processed by Workday's screening tool.

VENDOR LANGUAGE (DELETE):

Customer is solely responsible for determining the suitability of the Service for Customer's intended use case and for ensuring that Customer's use of the Service, including any decisions based on Service outputs, complies with all applicable laws, including non-discrimination, equal opportunity, and fair lending statutes. Vendor makes no representations regarding the suitability of outputs for use in legally regulated decision-making processes.

REVISED LANGUAGE (INSERT):

Vendor shall conduct bias testing and fairness audits at least annually, covering protected characteristics under applicable federal, state, and local law. Vendor shall provide audit reports to Customer upon request. For use cases involving employment, credit, housing, or legal services decisions, Vendor shall demonstrate no statistically significant disparate impact across protected groups and shall indemnify Customer for discrimination claims arising from the Service's outputs.

Key Principle:

The vendor built the model, trained it, and chose the data. The vendor is in the best position to test for bias and should bear contractual responsibility for it. "Customer is solely responsible" language in this context is a risk-shifting tactic, not a reasonable allocation.

Quick Reference: What to Check First

When you first receive an AI vendor agreement, scan for these five items before doing a full review. If any of these are missing or one-sided, flag them immediately:

- Training data restriction. Search the agreement for "train," "improve," "enhance," "aggregate," and "de-identify." If any of these terms appear in the vendor's rights section without an explicit carve-out for customer data, you have a problem.
- Output ownership. Search for "output," "generated," and "derivative." If the vendor claims any rights in output, or if the contract is silent on output ownership, this needs immediate revision.
- Liability cap structure. Look at the limitation of liability section. If there are no carve-outs for data breaches, IP infringement, or discrimination claims, the cap needs restructuring.
- Unilateral modification rights. Search for "modify," "change," and "update." If the vendor can change the model, terms, or data practices without notice, this is a non-starter for any enterprise use.
- Termination and data return. Go to the termination section. If it does not address embeddings, vector representations, and fine-tuned models, the deletion obligation is incomplete.

Part Three

Redlining in Practice

This section walks through the vendor provisions that most frequently need revision, with specific before-and-after examples. These are drawn from patterns found across major AI vendor agreements.

A note on approach: the redlines below are not aspirational wish lists. They reflect terms that enterprise customers are actually negotiating and, in many cases, obtaining. Where a provision is aggressive, we note fallback positions. Where a vendor is likely to push back, we explain why the provision matters and how to frame it.

The Liability Cap

Nearly every AI vendor limits its liability to fees paid in the prior 12 months. For a tool costing $5,000 per month, that is a $60,000 cap. If the tool produces a hallucinated case citation that leads to sanctions, or misses a material risk in a contract review, your exposure to your client could be many multiples of that cap.

The critical issue is not the cap itself but the scope of what falls under it. In traditional SaaS, a general liability cap is reasonable because the failure modes are predictable: downtime, data loss, service interruption. AI introduces failure modes that are qualitatively different: fabricated legal citations, discriminatory screening decisions, confidentiality breaches through model training. These risks require carve-outs.

VENDOR LANGUAGE (DELETE):

IN NO EVENT SHALL VENDOR'S AGGREGATE LIABILITY ARISING OUT OF OR RELATED TO THIS AGREEMENT EXCEED THE TOTAL FEES ACTUALLY PAID BY CUSTOMER TO VENDOR DURING THE TWELVE (12) MONTH PERIOD IMMEDIATELY PRECEDING THE EVENT GIVING RISE TO THE CLAIM. IN NO EVENT SHALL EITHER PARTY BE LIABLE TO THE OTHER FOR ANY INDIRECT, INCIDENTAL, CONSEQUENTIAL, SPECIAL, EXEMPLARY, OR PUNITIVE DAMAGES, INCLUDING WITHOUT LIMITATION DAMAGES FOR LOST PROFITS, LOST DATA, BUSINESS INTERRUPTION, OR LOSS OF GOODWILL, REGARDLESS OF THE CAUSE OF ACTION OR THE THEORY OF LIABILITY, EVEN IF SUCH PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

REVISED LANGUAGE (INSERT):

Vendor's total liability for general performance claims shall not exceed fees paid in the prior 12 months. However, the following are not subject to the liability cap: (a) Vendor's indemnification obligations; (b) breaches of data protection obligations; (c) Vendor's gross negligence or willful misconduct; (d) claims arising from AI hallucinations or discriminatory outputs where Customer used the Service in accordance with documentation. For uncapped claims, liability shall not exceed the greater of 12 months' fees or Vendor's applicable insurance coverage limits.

Fallback position: If the vendor will not accept uncapped carve-outs, propose a "super cap" for AI-specific claims at two to three times the general cap. This is a common middle ground in enterprise negotiations. The key is ensuring that the most dangerous failure modes (data breaches, discrimination, hallucination-caused harm) are not subject to the same cap as routine service issues.

The Indemnification Clause

Standard vendor indemnification typically covers only IP infringement in the vendor's underlying technology, not in the outputs. Given the wave of copyright litigation against AI companies (*Getty Images v. Stability AI*, *The New York Times v. OpenAI*, and others), output-level IP indemnification matters.[7] Microsoft, Adobe, and Google have each introduced some form of output-level indemnification for enterprise customers, which gives you leverage when negotiating with vendors that have not.

Watch for three common weaknesses in vendor indemnification clauses. First, knowledge qualifiers ("to the best of Vendor's knowledge") that effectively eliminate the indemnity because the vendor can always claim it did not know. Second, scope limitations that cover only the platform itself but not the outputs it generates. Third, conditions that require the customer to have used the service in a specific way, defined so narrowly that any real-world use falls outside the indemnity.

VENDOR LANGUAGE (DELETE):

Vendor shall defend, indemnify, and hold harmless Customer against third-party claims alleging that the Service, as provided by Vendor and used in accordance with the Agreement and applicable documentation, infringes any third-party intellectual property right, to the best of Vendor's knowledge. This indemnity shall not apply to claims arising from: (a) Customer's combination of the Service with third-party products or services; (b) any modification of the Service not made by Vendor; (c) Customer Data or Customer's inputs; or (d) use of the Service other than as documented.

REVISED LANGUAGE (INSERT):

Vendor shall indemnify Customer against third-party claims that: (a) the Service infringes third-party IP rights; (b) outputs generated by the Service infringe third-party copyright, trademark, or trade secret rights, provided Customer used the Service as documented and did not materially modify the output; (c) the Service produces discriminatory results in violation of applicable law. "To the best of Vendor's knowledge" is deleted because Vendor controls the training data and is in the best position to assess infringement risk.

Fallback position: If the vendor will not indemnify for all output-level IP claims, negotiate for indemnification limited to outputs generated when the customer follows the vendor's documented use guidelines. This narrows the vendor's exposure while still providing meaningful protection. Also consider requiring the vendor to maintain a content filter or watermarking system as a condition of limiting the indemnity scope.

The Data Use Clause

This clause often appears innocuous but carries enormous implications. A vendor that retains broad rights to "use" customer data for "product improvement" may be feeding your confidential client information into a model that serves your competitors. In February 2026, a federal court found that use of consumer AI tools with broad data collection terms could waive attorney-client privilege.[8]

VENDOR LANGUAGE (DELETE):

Customer hereby grants Vendor a non-exclusive, worldwide, royalty-free, sublicensable license to access, use, copy, transmit, store, and process Customer Data (including inputs, outputs, feedback, and usage data) as necessary to (a) provide and maintain the Service, (b) improve, develop, and enhance Vendor's products, services, and technology, including machine learning models, (c) generate aggregated and anonymized benchmarks, and (d) comply with applicable law. This license survives termination or expiration of this Agreement with respect to data processed prior to termination.

REVISED LANGUAGE (INSERT):

Vendor shall process Customer Data solely for the purpose of providing the Service to Customer. Vendor shall not use Customer Data (including inputs, outputs, feedback, or any derivatives) to train, improve, or develop any product, service, or model, whether for Vendor's own use or for any third party. No license to Customer Data is granted except as strictly necessary to deliver the Service during the term.

The Unilateral Modification Clause

Many AI vendors reserve the right to change models, features, or terms at any time without notice. For traditional SaaS, this might mean a UI redesign. For AI, it could mean the underlying model changes in ways that affect accuracy, introduce new biases, or alter how your data is processed.

VENDOR LANGUAGE (DELETE):

Vendor reserves the right to modify, update, or discontinue any features, functionality, or components of the Service at any time. Vendor will use reasonable efforts to notify Customer of material changes through the Service interface or by email to Customer's designated administrator. Continued use of the Service following notice of any modification constitutes Customer's acceptance of the modified Service.

REVISED LANGUAGE (INSERT):

Vendor shall provide 30 days' written notice before any material modification to the Service, including changes to the underlying model, data processing practices, or accuracy characteristics. "Material modification" includes any change to the model version, training data, or privacy practices. Customer may terminate without penalty within 30 days of receiving notice of any material modification it does not accept.

Negotiation Note:

If the vendor resists advance notice of model changes, propose a compromise: vendor provides notice within 48 hours after a model change, with a 30-day evaluation window during which Customer may terminate without penalty if the change materially affects performance or compliance.

The Confidentiality and Non-Disclosure Clause

Standard mutual NDAs in SaaS contracts typically protect both parties' confidential information from disclosure to third parties. In AI contracts, the risk is different: your confidential information may not be "disclosed" in the traditional sense but absorbed into a model that then serves other customers. The NDA needs to address this.

VENDOR LANGUAGE (DELETE):

Each party agrees to maintain the confidentiality of the other party's Confidential Information using at least the same degree of care it uses to protect its own confidential information (but no less than reasonable care), and not to disclose it to any third party without prior written consent. Confidential Information does not include information that: (a) becomes publicly available through no fault of the receiving party; (b) was known to the receiving party prior to disclosure; (c) is independently developed without reference to the disclosing party's Confidential Information; or (d) is required to be disclosed by law.

REVISED LANGUAGE (INSERT):

Each party agrees to maintain the confidentiality of the other party's Confidential Information and not to disclose it to any third party without prior written consent. For the avoidance of doubt, "disclosure" includes any use of Confidential Information to train, fine-tune, or improve any machine learning model, or any incorporation of Confidential Information (including patterns, structures, or derivatives) into any model, dataset, or product accessible by third parties. Vendor's confidentiality obligations survive termination for a period of five years.

The Force Majeure and Service Suspension Clause

AI vendors sometimes invoke force majeure or service suspension provisions when model performance degrades, when they need to retrain due to a safety issue, or when upstream providers (such as cloud infrastructure or foundational model providers) experience disruptions. Your contract should define what qualifies as force majeure and what does not.

VENDOR LANGUAGE (DELETE):

Neither party shall be liable for any failure or delay in performance caused by circumstances beyond its reasonable control, including but not limited to acts of God, natural disasters, pandemic or epidemic, government actions or orders, war or terrorism, labor disputes, power or internet outages, cyberattacks, failure or disruption of third-party services or infrastructure, or any other event beyond the party's reasonable control (each, a "Force Majeure Event").

REVISED LANGUAGE (INSERT):

Vendor shall not be liable for any failure or delay caused by circumstances beyond its reasonable control, including natural disasters, government actions, or widespread internet outages. The following do not constitute force majeure events: model performance degradation, upstream AI provider changes, training data issues, regulatory compliance requirements that were reasonably foreseeable, or failure of Vendor's subprocessors. If a force majeure event persists for more than 30 consecutive days, Customer may terminate without penalty and receive a pro-rata refund of prepaid fees.

Summary: The Redlining Process

When you receive a vendor agreement, work through it systematically. Start with the five items in the Quick Reference checklist in Part Two. Then review each clause in this section. For each provision: (1) identify the vendor's position, (2) compare it to the revised language above, (3) adapt the specific

thresholds and timeframes to your deal, (4) if the vendor pushes back, use the fallback positions noted above. Do not accept "this is our standard" as a reason to stop negotiating. Every term in this section is negotiable, and enterprise customers are obtaining these protections.

Part Four

Data, Privacy, and Privilege

For lawyers, data privacy in AI contracts is not just a GDPR compliance exercise. It is a professional responsibility issue. ABA Formal Opinion 512, issued in July 2024, makes clear that lawyers must understand whether AI tools are "self-learning" and whether confidential client information could be exposed through use of the tool.[9]

The Enterprise vs. Consumer Distinction

Most major AI providers offer both consumer and enterprise tiers. The difference is not just price. Consumer versions of tools like ChatGPT, Claude, and Gemini typically retain broad rights to use your data for model training. Enterprise versions generally do not, but the specific protections vary significantly by vendor and must be verified in the contract, not assumed from marketing materials.

If your firm or client is using the consumer version of any AI tool on matters involving confidential information, that is a privilege and confidentiality problem right now, regardless of what the vendor's privacy policy says. The contract you need to negotiate is the enterprise agreement.

What "Training on Your Data" Actually Means

When a vendor "trains" on your data, the practical effect is that your information becomes part of the model itself. The model does not store your documents in a database it can retrieve later. Instead, training adjusts the model's internal parameters (weights) so that patterns from your data influence every future output the model produces, for every customer. Once your data has been used for training, extracting it from the model is not possible in any practical sense. There is no "undo" button.

This creates three distinct risks. First, confidentiality loss: information from your client matters can surface in outputs the vendor generates for other users. The model will not reproduce your documents verbatim, but it may reflect patterns, legal strategies, deal structures, or factual details that originated in your data. Second, competitive exposure: if the vendor serves your opposing counsel or a competitor to your client, your work product may indirectly inform the other side's analysis. Third, regulatory violation: if your data includes personal information subject to GDPR, CCPA, or sector-specific privacy laws, incorporating it into model weights may constitute unauthorized processing, because the data is now being used for a purpose (improving a commercial product) that the data subject never consented to.

The third-party dimension compounds each of these risks (and is why the subprocessor provisions discussed in Part Two matter so much). Most AI vendors do not operate in isolation. A vendor that builds on top of a foundation model provider (OpenAI, Anthropic, Google, or others) may pass your data through to the underlying provider as part of its normal operations. If the foundation model provider retains data for its own training purposes, your information can end up in a model used by thousands of other applications you have no visibility into and no contractual relationship with. The same risk applies to cloud infrastructure providers, vector database services, and any other subprocessor in the vendor's data pipeline.

Your contract must address this chain directly. Requiring the primary vendor not to train on your data is necessary but insufficient if the vendor's subprocessors retain that right. Every entity that touches your data must be bound by the same restriction, and the primary vendor must be contractually responsible for enforcing it downstream. Ask the vendor to identify every third party that will process your data and confirm, in the contract, that none of them will use it for training.

Data Processing Agreements

Every AI vendor contract involving personal data should include a Data Processing Agreement (DPA). But standard DPAs were designed for traditional data processors, not AI systems that learn from the data they process. Your DPA needs to address AI-specific risks.

At minimum, the DPA should explicitly state that: the vendor processes data solely on your instructions; the vendor does not use data to train or improve models; subprocessors are disclosed with objection rights (industry standard is 15 days' notice); data is deleted within 30 days of termination, including embeddings and vector representations; and the vendor cooperates with data subject access requests.

Privilege Protection

ABA Formal Opinion 512 states that generic consent in engagement letters is insufficient for AI tool use. Lawyers must provide specific disclosure about which AI tools they use, how data is processed, and what risks are involved.[9]

Your vendor contract should include a representation that use of the enterprise version does not waive attorney-client privilege, that the vendor will execute a confidentiality agreement or DPA with protections at least as stringent as your ethical obligations require, and that the vendor maintains SOC 2 Type II certification for security and confidentiality.

Practical Step:

Before signing any AI vendor contract, confirm in writing (not just from the sales representative, but in the contract itself) that the enterprise version does not train on customer data. If the vendor's standard terms include any exception for "safety," "abuse prevention," or "feedback," negotiate those exceptions down to the narrowest possible scope and require notice before any data is used.

Practice-Area-Specific Privilege Concerns

Different practice areas carry different privilege and confidentiality risks when using AI tools. Your vendor contract should address the specific sensitivities of your practice:

- **Criminal defense.** Attorney work product protection is at its strongest in criminal cases. Client strategy, witness evaluations, and plea negotiation positions are highly sensitive. Require explicit representation that the enterprise version does not waive work product privilege. Limit use to research pre-screening with mandatory independent verification, given that errors in sentencing guidelines or case law can directly affect client outcomes.
- **Family law.** Family law matters involve spousal privilege, settlement positions, child custody information, and financial disclosures. Require the vendor to confirm that no family law matter details, settlement positions, or information about spouses or children will be used for model training or improvement. If the vendor is breached and family law data is exposed, you need immediate termination rights and indemnification for privilege waiver claims.
- **Immigration law.** Client data includes national origin, visa status, family structure, and travel history. If disclosed, this information could expose clients to enforcement risk. Require explicit confirmation that no immigration-related prompts or case details will be used for training. Additionally, immigration clients' data may be subject to GDPR if family members reside in the EU, so confirm the vendor's data residency capabilities.

- **Healthcare law.** Any AI tool processing protected health information requires a HIPAA Business Associate Agreement executed separately from the main contract. Do not rely on the vendor's general data protection clauses to satisfy HIPAA. Require explicit prohibition on using PHI for model training, HIPAA-compliant encryption, and breach notification within 60 days.
- **Real estate.** Hallucinations in property-specific research, such as fabricated recorded documents, misidentified jurisdictional authority, or incorrect zoning citations, can affect title insurance and escrow compliance. Require audit trails showing which underlying data sources were used for each output.

Client Disclosure Requirements

ABA Formal Opinion 512 requires specific disclosure to clients about AI tool use. For new clients, include this in the engagement letter. For existing matters, confirm consent before deploying a new AI tool on their work. The disclosure should address: which specific AI tools you use and which versions (consumer vs. enterprise), what categories of client data they process, whether the vendor trains on that data, and the professional responsibility risks you have assessed. This disclosure is not a waiver. It is informed consent. If a client objects, respect that and use non-AI tools for their matters.

DPA Provisions Specific to AI

Your standard DPA template almost certainly does not cover AI-specific risks. Add these provisions explicitly:

- **No model training.** "Processor shall not use Personal Data or any derivatives thereof (including embeddings, vector representations, or aggregated patterns) for the purpose of training, fine-tuning, or improving any machine learning model."
- **Subprocessor AI restrictions.** "Processor shall ensure that no subprocessor uses Personal Data for model training purposes, and shall include equivalent restrictions in all subprocessor agreements."
- **Deletion scope.** "Upon termination, Processor shall delete all Personal Data including any embeddings, cached representations, and data stored in vector databases, within 30 days. Processor shall certify that Personal Data does not persist in any model weights or training datasets."

Part Five

Liability and Risk Allocation

The fundamental problem with AI vendor liability provisions is a mismatch between who creates the risk and who bears the consequences. The vendor builds the model, trains it on data the vendor selects, and determines how it processes inputs. But when the model hallucinates, discriminates, or produces infringing output, the standard contract puts the liability squarely on the customer.

The Hallucination Problem

A 2024 Stanford HAI study found hallucination rates of 17% for Lexis+ AI and up to 33% for Westlaw AI in legal research tasks, with general-purpose models performing significantly worse.[10] These are not edge cases. They are baseline performance characteristics. In practical terms: if a tool returns 10 citations, one to three of them may not exist or may be inaccurate. If you rely on a fabricated citation in a brief without independent verification, you face sanctions risk under Rule 11 and potential malpractice liability to your client. Allocate this risk in the contract before you sign.

If the vendor cannot guarantee that outputs will be factually accurate, then the vendor should not be able to disclaim all liability when they are not. The practical solution is to tie liability to the vendor's documented performance standards: if the vendor promises accuracy above a certain threshold and fails to deliver, the liability cap should not protect that failure.

Discrimination and Bias Liability

In *Mobley v. Workday*, a federal court allowed disparate impact claims to proceed directly against the AI vendor for employment discrimination, even though Workday was not the employer.[6] The case involved over one billion rejected job applications processed by Workday's AI screening tool and achieved preliminary collective certification in May 2025.

This is not a hypothetical risk. If your client uses an AI tool for hiring, credit decisions, insurance underwriting, or any process with a disparate impact, the vendor should share liability. The contract should require the vendor to conduct bias testing, disclose results, and indemnify against discrimination claims arising from model design.

IP Infringement in Outputs

Multiple lawsuits are pending against major AI vendors over training data that included copyrighted material without authorization: *Getty Images v. Stability AI*, *The New York Times v. OpenAI*, and class actions by authors against multiple AI companies.[7] If the model was trained on infringing data, the outputs may carry infringement risk. This risk belongs to the party that chose the training data: the vendor.

Microsoft offers indemnification for copyright infringement in Copilot outputs. Adobe indemnifies for Firefly-generated images. Use these as benchmarks when negotiating with vendors that disclaim output-level IP protection.

Negotiation Leverage:

"Your competitor provides output-level IP indemnification. We expect equivalent protection. If you cannot provide it, we need to understand why, and our risk assessment of this tool changes significantly."

Part Six

Performance Standards and SLAs

Traditional SLAs measure uptime and response time. AI SLAs need to measure something fundamentally different: the quality and reliability of outputs that change every time the model runs.

Accuracy and Hallucination Thresholds

Define what acceptable performance looks like in measurable terms. A hallucination rate threshold of 5% is a reasonable starting point for legal research tools, based on current benchmarks showing best-in-class tools achieving rates below 1% for straightforward tasks but significantly higher for complex reasoning.[11]

VENDOR LANGUAGE (DELETE):

The Service is provided on an "as is" and "as available" basis. Vendor shall use commercially reasonable efforts to provide accurate and reliable results, but does not guarantee the accuracy, completeness, or fitness for any particular purpose of any output generated by the Service. Customer acknowledges that AI-generated outputs may contain errors, omissions, or inaccuracies and that Customer is solely responsible for independently verifying all outputs before reliance.

REVISED LANGUAGE (INSERT):

Vendor shall maintain a hallucination rate not exceeding [X]% on the benchmark evaluation set defined in Exhibit A, measured quarterly. If quarterly performance falls below threshold: (a) Vendor provides service credit of 10% of monthly fees for each 1% above threshold; (b) at Customer's request, Vendor retrains the model at no charge; (c) if underperformance persists beyond 90 days, Customer may terminate without penalty.

Bias and Fairness Metrics

For tools used in decision-making, performance must include fairness metrics. For example, if an AI hiring screener accepts 20% of male applicants but only 15% of female applicants with comparable qualifications, that is a disparate impact and a legal liability. Require the vendor to measure acceptance, rejection, or recommendation rates broken down by protected characteristics (race, gender, age, disability status, depending on jurisdiction) at least quarterly, and commit that no characteristic shows more than a defined performance difference. Tie SLA credits to fairness failures.

Model Drift Monitoring

Require the vendor to detect performance degradation within 24 hours, notify you within 48 hours, and provide a remediation plan within five business days. Define "drift" concretely: an accuracy decline of more than two percentage points month-over-month, or failure to meet performance thresholds on current test data.

SLA Credits That Actually Matter

Standard SLA credits of 5-10% of monthly fees provide no meaningful incentive for the vendor to perform. For AI-specific failures (high hallucination rates, bias above threshold), credits should escalate

rapidly, and persistent failure should trigger termination rights.

A workable credit escalation structure:

- **Tier 1 (minor):** Performance 1-2 percentage points below threshold for one quarter: 10% credit on monthly fees for that quarter.
- **Tier 2 (moderate):** Performance 3-5 percentage points below threshold, or Tier 1 failure for two consecutive quarters: 25% credit and mandatory remediation plan within 15 business days.
- **Tier 3 (severe):** Performance more than 5 percentage points below threshold, or Tier 2 failure for two consecutive quarters: 50% credit and Customer right to terminate without penalty.
- **Bias SLA:** Any statistically significant disparate impact finding triggers immediate suspension of the affected use case until remediated, plus indemnification for any claims arising during the affected period.

Part Seven

Termination and Getting Out

AI vendor lock-in is more dangerous than traditional software lock-in. With standard SaaS, you lose access to a tool. With AI, you may lose access to custom models trained on your data, embeddings that represent years of document analysis, and workflows that depend on specific model behaviors. The EU Data Act, effective September 2025, now mandates specific data portability and switching rights that give customers leverage in these negotiations.[12]

Data Portability

Your contract should guarantee that all customer data can be exported in open, machine-readable formats within 30 days of termination. This includes not just the raw data you input but also any embeddings, vector representations, indexes, and derivatives the vendor created from your data.

Model and Artifact Unwinding

If you have fine-tuned models or custom configurations, the contract must address who owns those artifacts and what happens to them on termination. Can you download model weights? Can the vendor continue using a model trained on your data to serve other customers? These questions need answers in the contract, not after termination.

VENDOR LANGUAGE (DELETE):

Upon termination or expiration of this Agreement, Vendor shall delete Customer Data from its production systems within thirty (30) days, and from backup and archival systems within ninety (90) days, in accordance with Vendor's standard data retention policies. Vendor shall provide written certification of deletion upon Customer's request. Data that has been incorporated into aggregated or de-identified datasets, or that Vendor is required to retain by applicable law, is excluded from the deletion obligation.

REVISED LANGUAGE (INSERT):

Upon termination, Vendor shall within 30 days: (a) return or securely delete all Customer Data, including embeddings, vector representations, and cached outputs; (b) delete all fine-tuned models and derivatives created using Customer Data; (c) certify in writing that no Customer Data persists in any training dataset, model weights, or backup system; (d) provide transition assistance at no additional cost, including data export in open formats and up to [X] hours of technical support for migration to a replacement service.

Transition Assistance

Negotiate for specific transition obligations: data export in your choice of format, technical support for migration, cooperation with your replacement vendor, and continued access to the service during the transition period at existing rates. The EU Data Act requires vendors to complete switching activities within 30 days, with a maximum extension of seven months only when technically infeasible.[12]

Lock-in Red Flag:

If the vendor stores fine-tuned models exclusively on its own infrastructure and cannot export model weights in standard formats (safetensors, ONNX), you are locked in from day one. Address this before signing, not at termination.

Part Eight

Negotiation Tactics That Work

AI vendor negotiations have their own dynamics. Vendors are competing hard for market share with products that change quarterly, which gives buyers more leverage than many realize. Here are specific tactics that work.

Understand the Vendor's Incentives

Before you negotiate, understand what the vendor actually needs. Most AI vendors are under pressure to prove they have paying customers in regulated industries (law, healthcare, finance) because investors care about that. They also want multi-year contracts, not one-year pilots, because cancellations hurt their growth metrics. You can use both of these: "We will commit to a three-year deal if you give us the data protections we need, and we will be a reference customer in our industry if the tool performs." That is leverage every size firm can claim.

The vendor's sales team has authority to negotiate certain terms but not others. Data training restrictions, liability caps, and indemnification scope often require escalation to legal or product leadership. Ask early in the process what the vendor's approval chain looks like for contract modifications. This tells you who actually makes decisions and how long the process will take.

Use Competitive Intelligence

Know what the vendor's competitors offer. If Microsoft indemnifies for Copilot outputs, say so when negotiating with a vendor that does not. If Google offers data residency controls in Gemini Enterprise, cite that when your vendor claims it is "not technically feasible." Vendors track each other's contract terms closely. Showing that you do too changes the negotiation.

Specific examples of competitive terms you can cite: Microsoft's Copilot Copyright Commitment provides indemnification for copyright infringement in outputs. Adobe's Firefly IP indemnification covers images generated by its tools. Anthropic's Enterprise terms include zero data retention for enterprise API customers. Google's Gemini Enterprise provides data residency controls and no training on customer data. These are public commitments you can reference directly.

Counter Common Vendor Objections

AI vendors have a standard playbook for resisting customer requests. Knowing the playbook in advance lets you respond immediately rather than taking issues "back to the team." Here are six objections you will hear repeatedly, and how to respond:

"We cannot identify our training data sources."

Then the vendor cannot warrant that the model does not infringe third-party IP. Point this out directly. OpenAI and Google negotiated licensing agreements with Reddit and other content platforms after facing similar pressure. If those companies can do it, your vendor can start.

"We cannot agree to minimum accuracy thresholds."

Then you cannot rely on outputs for anything sensitive. Propose a compromise: restrict use to suggestions only with mandatory human review, and define a joint monitoring protocol to establish baselines over time. Frame this as a partnership: "We want to use your tool. Help us define when it is

safe to do so."

"Data deletion upon termination is technically impossible."

The EU Data Act requires data portability and deletion within 30 days.[12] If the vendor cannot comply with EU law, that is a significant infrastructure concern, not a negotiation position. If the vendor claims model weights cannot be separated from training data, ask them to explain exactly how customer data is stored, processed, and retained. The technical explanation often reveals more options than the initial objection suggests.

"Our liability cap is 12 months of fees and that is final."

The *Mobley v. Workday* case shows courts may hold vendors liable despite contract limitations for discrimination claims.[6] Propose carve-outs: keep the cap for general performance issues, but remove it for discrimination, data breaches, and IP infringement indemnification. Tie the uncapped amount to the vendor's insurance coverage.

"We do not negotiate our standard terms."

Every vendor negotiates with customers who are willing to walk. The question is whether your deal is large enough or strategic enough to justify the vendor's time. If you are a small buyer, consider aggregating your purchase with other firms or departments to increase leverage. If you are buying for a regulated industry (healthcare, financial services, legal), emphasize that regulatory compliance is not optional and that the vendor's standard terms create compliance risk for both parties.

"Our security certifications cover these concerns."

SOC 2 and ISO 27001 certifications do not address AI-specific risks like model training on customer data, output accuracy, or bias. These certifications verify information security controls, not AI governance. Point out that the certification scope likely does not cover the specific risk you are raising, and request documentation of the certification's scope.

Trade, Don't Concede

If the vendor will not reduce the liability cap, trade for something else: broader audit rights, faster deletion timelines, advance notice of model changes, or enhanced SLA credits. Identify your true priorities before the negotiation begins and know which points you are willing to trade.

Effective trades in AI vendor negotiations often follow this pattern: the vendor resists a structural change (such as removing the liability cap for AI-specific claims) but accepts operational concessions that reduce the underlying risk (such as mandatory bias testing, advance notice of model changes, or expanded audit rights). A contract that requires the vendor to test for bias quarterly is often more valuable than a higher liability cap, because it reduces the likelihood of the claim arising in the first place.

Start With the DPA

If the vendor is resistant to negotiating the main agreement, start with the Data Processing Agreement. Vendors are more accustomed to DPA negotiations, and a strong DPA establishes precedent for the data protections you want in the master agreement. Once the vendor agrees to no-training-on-customer-data in the DPA, it is difficult to argue against the same restriction in the main contract.

Use Pilots and Phased Rollouts as Leverage

If the vendor will not agree to your terms for a full enterprise deployment, propose a pilot with a limited scope and a shorter term (90 to 180 days). Use the pilot period to establish performance baselines that become the SLA metrics in the full agreement. The pilot also gives you data on the vendor's actual performance, which strengthens your negotiating position for the full contract.

Structure the pilot agreement to include: a defined scope of use, performance metrics you will measure, a clear path to the full agreement if metrics are met, and the right to terminate without penalty if they are not. The pilot is not a concession. It is a way to de-risk the deal for both sides while building the data you need to negotiate informed terms.

Before You Start: Preparation Checklist

Effective AI vendor negotiations require preparation. Before the first call or redline exchange, complete these steps:

- **Map your data flows.** Know exactly what data will enter the AI system, what categories of personal data are involved, and what regulatory regimes apply. You cannot negotiate data protections if you do not know what data needs protecting.
- **Identify your non-negotiables.** Rank your priorities. For most organizations: (1) no training on customer data, (2) output ownership, (3) adequate liability allocation, (4) termination and data portability. Know which points you will walk away over.
- **Research the vendor's competitors.** Get quotes and contract terms from at least two competing vendors. Specific knowledge of competitor terms is the most effective leverage you can bring to the table.
- **Check the vendor's standard terms against industry templates.** Compare the vendor's terms to the Bonterms AI Standard Clauses or the IAPP AI Governance Center resources. Deviations from industry standards are your negotiation targets.
- **Assemble your team.** AI vendor negotiations benefit from having both legal and technical voices at the table. Your IT or data science team can challenge vendor claims about technical impossibility. Your privacy officer can speak to regulatory requirements. Your procurement team knows the vendor's commercial pressure points.

For Solo Practitioners and Small Firms

If you are a solo practitioner or small firm, you likely lack the scale to demand custom terms from major vendors. Do not give up on these protections. Instead, pursue them strategically:

- **Reference industry templates.** Tell the vendor: "The Bonterms AI Standard Clauses are industry-standard. We expect you to meet them." This shifts the conversation from "our firm's unusual request" to "your failure to meet market norms."
- **Aggregate your leverage.** If you are part of a bar association, practice group, or peer review organization, see if other members also want this tool. A group negotiation has more leverage than individual negotiations.
- **Prioritize ruthlessly.** Your non-negotiables should be: no training on client data, confidentiality protections adequate for privileged information, and the right to terminate without penalty if the vendor materially changes its data practices. You may have to accept broader liability caps than enterprise customers, but these three items are worth fighting for.
- **Emphasize professional responsibility.** Every lawyer, regardless of firm size, can tell a vendor: "My ethical obligations under the Rules of Professional Conduct require these protections. If you

cannot provide them, I cannot use your tool." This is not a negotiation tactic. It is a statement of fact.

Strongest Position:

The strongest negotiation position is the willingness to walk. If the vendor will not agree to basic data training restrictions, adequate liability allocation, or reasonable exit terms, the risk of signing may outweigh the benefit of the tool. There are other vendors.

Part Nine

Regulatory Compliance

AI regulation is moving on multiple fronts simultaneously. Contracts signed today must account for obligations already in effect and those taking force over the next 12 to 24 months.

EU AI Act

If your firm or clients operate in the EU, or if the vendor serves EU customers, the EU AI Act creates new obligations effective through 2026.[13] Prohibited practices (social scoring, subliminal manipulation) took effect in February 2025. High-risk AI system obligations, including documentation, conformity assessment, and quality management requirements, take effect in August 2026. Penalties reach up to 7% of global annual revenue. Do not assume the vendor has assessed its compliance. Most have not.

Your contract should require the vendor to: assess whether the tool constitutes a high-risk AI system under Annex III; maintain technical documentation including training data provenance, testing results, and monitoring plans; conduct conformity assessments and register in the EU AI Register where required; and provide all documentation necessary for your own compliance as a deployer.

Colorado AI Act

Colorado's SB 24-205, effective June 30, 2026, requires developers and deployers of high-risk AI systems to use "reasonable care" to protect against algorithmic discrimination.[14] It applies when AI contributes to decisions about employment, education, housing, healthcare, or access to financial or legal services. Developers must provide use case statements, training data summaries, known limitations, evaluation methods, and impact assessment materials.

Sector-Specific Requirements

Your vendor contract must flow down whatever sector-specific requirements apply to your use case. The following are the most common:

- **Healthcare.** Any AI tool processing protected health information requires a HIPAA Business Associate Agreement (BAA) executed separately from the main contract. The BAA must explicitly prohibit using PHI for model training, require HIPAA-compliant encryption and access controls, mandate breach notification within 60 days, and extend to all subprocessors.
- **Employment and hiring.** If the AI tool is used for hiring, performance evaluation, or termination decisions, you must comply with: NYC Local Law 144 (annual bias audits and applicant notification); state hiring disclosure laws (New York, Illinois, and others require notice to applicants that AI is being used); and FCRA requirements if the tool affects hiring decisions. Require the vendor to provide documentation of bias testing methodology and to indemnify for discrimination claims arising from the tool's outputs.
- **Financial services.** If the tool handles financial data, GLBA privacy requirements and FCRA compliance obligations apply. If it contributes to credit or insurance decisions, the Equal Credit Opportunity Act and state insurance regulations may also apply.
- **Education.** FERPA applies to AI tools processing student education records. The vendor must restrict use of student data to the educational purpose specified in the contract and must not use it for marketing, profiling, or model training.

Compliance Assistance, Not Compliance Shifting

Vendors often include language making compliance the customer's sole responsibility. Push back. The vendor controls the model, the data, and the infrastructure. The contract should require the vendor to: provide documentation sufficient for your compliance obligations; conduct bias audits at the vendor's expense; deliver compliance certifications; and cooperate with regulatory audits.

Part Ten

Insurance and Risk Transfer

Your existing professional liability insurance likely does not cover AI-related errors. Search your current policy for "artificial intelligence," "algorithmic," "machine learning," and "hallucination." If you find nothing, or only exclusions, you have a coverage gap. Contact your broker and ask directly: does this policy cover malpractice claims arising from reliance on AI-generated output? Most professional liability and cyber policies do not explicitly address AI-specific failures: hallucinations causing financial loss, discriminatory outputs, or IP infringement in model-generated content. Many carriers are now introducing AI-specific exclusions, creating a "silent coverage" gap where AI risks are neither explicitly covered nor excluded.[15] Before signing any AI vendor contract, resolve this coverage question.

Require Vendor Insurance

Your contract should require the vendor to maintain: professional liability or E&O insurance with minimum coverage of $5 to $10 million that does not exclude AI-related claims; cyber liability insurance with minimum coverage of $10 to $20 million; and naming your organization as an additional insured on all policies. Require annual certificates of insurance and 30 days' notice of cancellation or non-renewal.

Tie Liability Caps to Insurance

A powerful negotiation tool: tie the vendor's liability cap for indemnification claims to its insurance coverage rather than to subscription fees. If the vendor maintains $5 million in E&O coverage, the liability cap for indemnifiable claims should be $5 million, not $60,000 in annual fees. This aligns the vendor's incentives: maintain adequate insurance or accept higher contractual exposure.

Evaluate Your Own Coverage

Review your firm's or client's existing policies for AI coverage gaps. Confirm whether your professional liability policy covers errors arising from reliance on AI-generated output. Confirm whether your cyber policy covers data breaches caused by AI vendors. If there are gaps, address them through vendor indemnification, supplemental insurance, or both.

Conclusion

AI vendor contracts are not standard procurement. They involve risks that traditional contract templates were never built for: data training that can compromise confidentiality, outputs that hallucinate or discriminate, performance that drifts without warning, and termination that can leave you locked into infrastructure you cannot replicate.

The market is moving in your favor. Vendors are competing for enterprise customers, and enterprise customers are demanding better terms. Microsoft, Adobe, and Google have introduced indemnification provisions that were unthinkable two years ago. The EU AI Act and Colorado AI Act are creating compliance obligations that require vendor cooperation. And case law like *Mobley v. Workday* is making clear that courts will hold vendors accountable.

Your leverage in these negotiations is real. Use it. Be specific about what you need, know what the competition offers, and be willing to walk if the terms do not protect your clients. The vendor wants your business. Make them earn it with a contract that allocates risk where it belongs.

Glossary of Key Terms

AI vendor contracts use terminology from both the legal and technical domains. This glossary defines terms as they appear in vendor agreements and as they should be understood for negotiation purposes. Where a term has a different meaning in an AI context than in traditional software contracting, the AI-specific meaning is noted.

AI and Technical Terms

Algorithmic Discrimination / Algorithmic Bias.

Systematic and repeatable errors in an AI system that produce unfair outcomes for particular groups, typically along lines of race, gender, age, or other protected characteristics. In a contractual context, this is the risk that vendor language attempts to shift entirely to the customer. The Colorado AI Act uses "algorithmic discrimination" as its operative term.

Embeddings.

Numerical representations of text, images, or other data that an AI model creates to understand relationships between inputs. Embeddings are stored in vector databases and can contain encoded representations of confidential information. When a vendor agreement references "deletion of customer data," embeddings and vector representations must be explicitly included or they will likely persist after termination.

Fine-Tuning.

The process of training a pre-existing AI model on a customer's specific data to improve performance for a particular use case. Fine-tuned models incorporate customer data into the model weights themselves, which raises ownership, portability, and deletion issues that standard data processing concepts do not address.

Foundation Model / Large Language Model (LLM).

A large AI model trained on broad data that can be adapted to many tasks. In vendor agreements, the distinction matters because many vendors build their products on top of foundation models provided by third parties (OpenAI, Anthropic, Google). Your data may flow through multiple parties, each with its own data practices, and the vendor's contract needs to address this chain.

Hallucination.

An AI output that is factually incorrect, fabricated, or unsupported by the input data but presented with apparent confidence. In legal tools, hallucinations have included fabricated case citations, invented statutes, and non-existent regulatory provisions. Contractually, hallucination is the AI-specific failure mode that most directly creates malpractice risk for lawyers.

Inference.

The process of an AI model generating an output from an input (as distinct from training, which is the process of building the model). Most AI vendor services involve inference: you send data in, the model processes it, and you receive output. The distinction matters because vendors may argue that processing data for inference is different from using data for training, even though both involve the vendor accessing your data.

Model Card.

A document describing an AI model's intended use, performance characteristics, training data, limitations, and evaluation results. Model cards are the AI equivalent of a product specification sheet. The EU AI Act requires documentation that functions similarly to a model card for high-risk systems.

Model Drift.

The gradual degradation of an AI model's performance over time as real-world data diverges from the data on which the model was trained. Unlike a software bug, model drift is silent: the tool continues to operate but produces increasingly unreliable results. Contracts should define drift thresholds and monitoring obligations.

Prompt / Prompt Engineering.

The text input a user provides to an AI system to generate a response. Prompts may contain confidential information, legal analysis, or client data. In vendor agreements, prompts are part of "Customer Data" and must be subject to the same training restrictions and confidentiality protections as any other input.

RAG (Retrieval-Augmented Generation).

A technique where an AI system retrieves information from a specific knowledge base before generating a response, rather than relying solely on its training data. RAG systems are common in legal AI tools because they reduce hallucinations by grounding outputs in source material. Contractually, RAG raises questions about where the knowledge base is stored, who controls it, and whether the vendor can access it.

Training Data.

The data used to build or improve an AI model. This is the core contractual issue: whether the vendor can use your data, your clients' data, or the outputs generated from that data as training data. "Aggregated" or "de-identified" training data may still be identifiable, especially in specialized legal contexts.

Vector Database.

A specialized database that stores embeddings and enables similarity searches. If a vendor uses a vector database to store representations of your documents, those representations are customer data and must be addressed in deletion and portability provisions.

Legal and Contractual Terms in AI Context

Commercially Reasonable Efforts.

A standard of care that appears frequently in AI vendor agreements, typically in accuracy, maintenance, and security provisions. In the AI context, this standard is often too vague to be meaningful. A vendor exercising "commercially reasonable efforts" to maintain accuracy has no measurable obligation. Replace with specific thresholds and timelines wherever possible.

Data Processing Agreement (DPA).

A contract (or contractual annex) required under GDPR and similar privacy laws that governs how a processor handles personal data on behalf of a controller. For AI vendors, standard DPA templates are insufficient because they do not address model training, embeddings, or the AI-specific ways in which data may persist beyond the service relationship.

Deployer.

Under the EU AI Act, the entity that puts an AI system into use in a professional context. If your firm or client uses an AI vendor's tool, you are the deployer. Deployers have specific obligations under the EU AI Act, including conducting impact assessments, monitoring system performance, and maintaining records.

High-Risk AI System.

Under the EU AI Act, an AI system used in specified domains (including employment, credit, education, law enforcement, and access to essential services) that is subject to heightened requirements for documentation, testing, transparency, and human oversight. Under the Colorado AI Act, a similar concept applies to systems that make or substantially contribute to consequential decisions.

Indemnification (AI-Specific).

In traditional software contracts, indemnification typically covers IP infringement in the vendor's code. In AI contracts, the scope must extend to outputs: IP infringement in generated content, discriminatory results, and claims arising from hallucinations. "Output-level indemnification" is the term of art for this broader coverage.

Model Weights.

The numerical parameters that define an AI model's behavior. Model weights are the product of training and represent the vendor's core intellectual property. However, if your data was used to fine-tune the model, a portion of those weights encodes your data. Ownership and portability of fine-tuned model weights is a key termination issue.

Provider / Developer.

Under the EU AI Act, the entity that develops an AI system or has one developed and places it on the market. If the AI vendor developed the tool, it is the provider and bears the primary compliance burden, including conformity assessment and registration obligations.

Silent Coverage Gap.

In the insurance context, a risk that is neither explicitly covered nor explicitly excluded by an insurance policy. AI risks (hallucination liability, algorithmic discrimination claims, IP infringement in outputs) often fall into silent coverage gaps in existing professional liability and cyber policies. The gap creates uncertainty about whether a claim will be covered.

Subprocessor.

A third party that processes personal data on behalf of the primary vendor. In AI, subprocessors are especially important because many vendors rely on foundational model providers (OpenAI, Anthropic, Google) as subprocessors. Your data may flow through multiple entities, each with its own data practices, and each must be subject to the same restrictions as the primary vendor.

Endnotes

[1] Industry analysis of AI vendor contract terms, 2024. See, e.g., Jones Walker LLP, "The AI Vendor Liability Squeeze," 2025; Gouchev Law, "10 Critical Clauses for AI Vendor Contracts," 2025.

[2] Bonterms, "AI Standard Clauses (Version 1.0)," 2023. Open-source contract clauses for AI-specific provisions. Available at https://bonterms.com/forms/ai-standard-clauses-version-1-0

[3] U.S. Copyright Office, "Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence," 88 Fed. Reg. 16190 (March 16, 2023).

[4] Industry surveys on AI model performance degradation have consistently found that the majority of deployed AI models experience accuracy declines without ongoing monitoring. See, e.g., Gartner, "AI in Production: Managing Model Degradation," 2024; MLOps.community surveys on model monitoring practices, 2024.

[5] California AI Transparency Act (SB 942), originally effective January 1, 2026; operative date extended to August 2, 2026 by October 2025 amendments. Mandates detection tools and watermarking for providers with over one million monthly U.S. users.

[6] *Mobley v. Workday, Inc.*, No. 3:23-cv-00770 (N.D. Cal.). Court allowed disparate impact discrimination claims to proceed directly against AI vendor (agency theory claims dismissed). Preliminary collective certification granted May 2025.

[7] See *Getty Images v. Stability AI* (D. Del., filed Feb. 2023); *The New York Times v. OpenAI* (S.D.N.Y., filed Dec. 2023); and multiple author class actions against AI vendors for training on copyrighted material.

[8] K&L Gates, "Generative AI Data, Attorney-Client Privilege, and the Work-Product Doctrine," February 2026. Discusses federal court ruling on privilege waiver from consumer AI tool use.

[9] ABA Standing Committee on Ethics and Professional Responsibility, Formal Opinion 512, "Generative Artificial Intelligence Tools," July 29, 2024. Addresses competence, confidentiality, communication, candor, supervision, and fees.

[10] Magesh et al., "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools," 22 J. Empirical Legal Studies 358 (2025). Originally published as Stanford HAI working paper, 2024. Found hallucination rates of 17-33% across major legal research AI tools.

[11] Hallucination rate benchmarks, 2025: Google Gemini-2.0-Flash at 0.7%; OpenAI variants at 0.8-1.5%; reasoning models at 33-51% on open-ended tasks. See Visual Capitalist, "Ranked: AI Hallucination Rates by Model," 2025.

[12] EU Data Act, effective September 12, 2025. Mandates data portability and switching rights for data processing services, including 30-day transition period and maximum 7-month extension.

[13] EU Artificial Intelligence Act (Regulation (EU) 2024/1689), entered into force August 1, 2024. Prohibited practices effective February 2, 2025; high-risk obligations effective August 2, 2026. Penalties up to 7% of global annual revenue.

[14] Colorado Artificial Intelligence Act (SB 24-205), signed May 17, 2024. Originally effective February 1, 2026; enforcement delayed to June 30, 2026 by SB 25B-004. Requires reasonable care to protect

against algorithmic discrimination.

[15] See K&L Gates, "Navigating the New Frontier: Insurance for Artificial Intelligence Risks," 2024; ABA Journal, "Does Your Professional Liability Insurance Cover AI Mistakes?," 2025.