Wallace Tech Law LLC
Privacy | AI Governance | Technology
Portland, OR
lauren@wallacetechlaw.com

Sleeper Terms in Legacy SaaS Contracts

A quick diagnostic for AI risk in vendor renewals

Many SaaS contracts signed between 2018 (the GDPR/CCPA inflection point, when privacy and security clauses entered the contracting mainstream) and today's AI-feature wave were negotiated for a different technical reality.

Agreements were negotiated for a world in which vendors hosted software, stored customer data, and improved products through ordinary debugging and feature refinement. Today, those same platforms may include AI-enabled functionality that changes what familiar clauses mean in practice.

This guide examines common SaaS terms from the GDPR/CCPA era, and reframes them for the AI era.

Each section identifies sleeper clauses to spot, shows how the "before" and "after" meanings can diverge, and offers questions to help you assess whether legacy contract language still fits your organization's current use of AI-enabled products.

---

1. "Improve the Services" Rights

What to look for
Vendor may use Customer Data to "improve," "enhance," or "develop" the services.

Before
Usually meant debugging, performance tuning, and feature refinement.

After
May now support model training, tuning, dataset development, and cross-customer learning.

Questions to ask now
Does "improve" include training or tuning AI models?
Is use limited to support and service delivery, or does it extend to product development?

---

2. Definition of Customer Data

What to look for
An older or narrow definition of “Customer Data.”

Before
Often focused on uploaded files, records, and account information.

After
May fail to capture prompts, outputs, annotations, corrections, logs, or other AI-related interaction data.

Questions to ask now
Are prompts and outputs expressly included?
Are logs, metadata, and user corrections covered by the same protections?

---

3. Usage Data, Telemetry, and Analytics

What to look for
Vendor rights to collect usage data, analytics, telemetry, or statistical information.

Before
Usually meant service monitoring, uptime, and product analytics.

After
May include prompts, outputs, user behavior, and corrections that are highly valuable for AI optimization.

Questions to ask now
What exactly is included in telemetry?
Are prompts, outputs, and correction data excluded from broader reuse?

---

4. De-Identified or Aggregated Data

What to look for
Vendor may use de-identified, anonymized, or aggregated data without restriction.

Before
Often accepted as low-risk benchmarking or reporting language.

After
May support training datasets, commercialization, or uses that create re-identification

and inference risk (especially where inferences generated by the AI are considered "personal data" under various privacy regimes).

Questions to ask now
What standard of de-identification applies?
Can de-identified data be used for model training or product development?
Does the definition of de-identification align with your regulatory environment? Terms such as anonymization, pseudonymization, and de-identification have different impact across various regulatory regimes (GDPR, CCPA/CPRA, HIPAA).

---

5. Feedback Clauses

What to look for
Broad vendor rights to use "feedback," "suggestions," or "ideas."

Before
Usually understood as ordinary comments on product usability.

After
May be read to include prompt refinements, workflow explanations, annotations, or corrections to AI outputs.

Questions to ask now
How is "feedback" defined?
Does it exclude prompts, outputs, confidential information, and customer workflows?

---

6. Confidentiality Provisions

What to look for
Standard confidentiality language with no AI-specific use restrictions.

Before
Designed to prevent disclosure and misuse in conventional SaaS operations.

After
May not clearly prevent confidential inputs from being used to train, tune, or shape AI systems.

Questions to ask now
Does confidentiality prohibit training or tuning on confidential information?
Does it address outputs that may reflect customer inputs?

7. License Grants

What to look for
Broad licenses allowing vendor to use, copy, modify, or create works from Customer Data.

Before
Usually tied to hosting and operating the service.

After
May be cited as authority for reuse, transformation, or long-term retention in AI-enabled environments.

Questions to ask now
Is the license limited to providing and supporting the service?
Does it expressly exclude model training, derivative development, and commercialization?

---

8. Data Security Clauses

What to look for
Strong security commitments, but no limits on authorized internal AI use.

Before
Focus was on unauthorized access and breach risk.

After
Security language may coexist with broad rights to use data in ways the customer never expected.

Questions to ask now
Are data use limits separate from security promises?
Does the contract address authorized but expansive internal use?

---

9. Warranties and Disclaimers

What to look for
“As is” language and broad disclaimers around output quality.

Before
Often tolerated in lower-risk software contexts.

After
Can become much more consequential when customers are encouraged to rely on AI-generated outputs in real workflows.

Questions to ask now
What responsibility does the vendor take for AI-enabled functionality?
Are there any commitments around output reliability, transparency, or human review?

---

10. Indemnification

What to look for
Traditional IP and third-party claim allocation.

Before
Drafted for conventional software risk.

After
May not clearly cover output-related IP claims, training data provenance issues, or embedded third-party content.

Questions to ask now
Does indemnity cover AI-generated outputs?
Who bears the risk if training data or generated content creates third-party claims?

---

11. Termination and Deletion

What to look for
Vendor will return or delete Customer Data on termination.

Before
Focus was on active stored data.

After
May not address whether data has already influenced models, embeddings, derived datasets, or tuning artifacts.

Questions to ask now
What happens to data already used in training or tuning?
Does deletion extend to derived AI artifacts or only source data?

12. Audit Rights and Transparency

What to look for
Conventional audit rights with limited operational visibility.

Before
Often enough for privacy and security verification.

After
May not provide meaningful visibility into training practices, data lineage, or downstream AI providers.

Questions to ask now
Can the customer verify whether its data is used in training or testing?
What transparency rights exist around AI functionality and data flows?

---

13. Subprocessors and AI Supply Chain

What to look for
Routine subprocessor language.

Before
Usually referred to hosting, storage, or support vendors.

After
May now include model providers, API layers, and other downstream AI infrastructure.

Questions to ask now
Who are the relevant downstream AI providers?
Do the same data use restrictions flow through the vendor's AI stack?

---

14. Derivative Works / Derived Data

What to look for
Vendor rights to create derivative works, derived data, or derived analytics.

Before
Often read as ordinary internal reporting or analysis language.

After
May be invoked to justify retention of downstream value extracted from prompts, outputs, feedback, or usage data, including trained model assets.

Questions to ask now
Can the vendor create derivative models or datasets from customer interactions?
Does the contract clearly distinguish service delivery from value extraction?

What to do next

Start with the contracts that sit closest to sensitive data, important workflows, or newly activated AI features, not necessarily the highest-spend vendors.

A targeted AI Addendum can be an efficient way to begin the conversation with critical vendors about training, reuse, retention, and other AI-era issues that older SaaS terms may not clearly address.

Need a starting point? I help legal teams identify priority contracts for AI review and draft practical AI Addenda.