

Generative AI

Due Diligence Framework

VENDOR ASSESSMENT PROTOCOL 2026

AUDIT ITEM	STATUS
01 Do you use our input data (prompts and documents) to retrain your base models or shared models? DATA PRIVACY	☐
02 Does your standard indemnity clause cover Intellectual Property infringement claims arising specifically from AI-generated outputs? RISK & IP	☐
03 Can you provide a documented 'Hallucination Rate' or accuracy benchmark for the specific legal tasks we are automating? MODEL INTEGRITY	☐
04 Is your solution 'Zero Retention' by default, or must we configure specific settings to prevent data persistence? DATA PRIVACY	☐

05 If the underlying LLM provider updates their model, how do you validate that our prompts/outputs remain consistent?

● MODEL INTEGRITY

☐

06 Do you offer a 'Private Cloud' or 'On-Premise' deployment option where the model weights sit entirely within our perimeter?

● SECURITY

☐

07 Are you SOC 2 Type II or ISO 27001 certified specifically for your AI infrastructure, not just your general corporate operations?

● SECURITY

☐

08 What is the distinction between 'Billable' vs. 'Administrative' AI usage in your pricing model (e.g., tokens, seats, or outcomes)?

● COMMERCIALS

☐

09 Does the platform provide an audit trail that tags specific document clauses or text segments as 'AI-Generated' versus 'Human-Drafted'?

● RISK & IP

☐

10 How do you handle 'Prompt Injection' attacks, and have you third-party pentested your prompts against jailbreaking?

● SECURITY

☐

11 Can we opt out of 'Quality Assurance' reviews where your human employees might view our anonymised data logs?

☐

● DATA PRIVACY

12 Do you maintain a 'Liability Cap' for AI errors that result in direct financial loss, and is it separate from the general software liability cap?

☐

● COMMERCIALS

13 Does the tool support 'Reference Anchoring'—providing direct citations/links to the source document for every claim made?

☐

● MODEL INTEGRITY

EVALUATED BY

DATE
