AI Vendor Evaluation Worksheet for Small Law Firms

A documented gate for approving, limiting, piloting, or rejecting an AI tool

September 23, 2026

Important notice

This worksheet is educational and operational planning material. It is not legal advice, ethics counsel, cybersecurity certification, procurement advice, or a substitute for review by the firm's lawyer, security adviser, insurer, and other qualified professionals.

Complete one worksheet for one product, account type, configuration, and proposed use case. A vendor may receive different decisions for different features, integrations, data classes, or practice groups.

Product terms and technical controls change. Save the documents reviewed, record the date, and set a re-review trigger. Vendor statements in a sales call do not replace the contract, data protection terms, security documentation, or a controlled test.

Decision states

Decision	Meaning
Approve	The documented use case, users, data classes, settings, and review rules may proceed.
Approve with limits	Use is allowed only within the listed restrictions and approval gates.
Pilot	Use fictional, public, or fully de-identified data in a controlled test. No client-data use unless separately approved.
Reject	The product or use case does not meet the firm's requirements.
Needs follow up	A named person must resolve listed gaps before a decision.

How to use the worksheet

1. Name the exact use case before reviewing the brand or product category.
2. Map the data entering, leaving, and connecting to the system.
3. Review written terms and security evidence for the exact account and feature.
4. Run controlled output and abuse tests with non-client data.
5. Record restrictions, owners, approval gates, monitoring, exit steps, and the next review date.
6. Stop if a critical question is unanswered. A numeric average does not cure a failed confidentiality or data-flow gate.

Evaluation record

Field	Entry
Vendor legal name	
Product and feature	
Edition account type and contract	
Evaluator and date	
Business owner	

Tehnial or admin owner	
Lawyer approver	
Pratie area or team	
Proposed users	
Proposed go live date	
Next review date	

The proposed use case

Question	Aswer
What work problem will this solve?	
What exat task will the tool perform?	
What soure material will it reeive?	
What output will it produe?	
Who beks the output and against what soure?	
Where an the output go after approval?	
What ation is the tool prohibited from taking?	
What urrent proess or tool does it replae?	
What evidene will show that the use ase works?	

Stop rule

If the evaluator annot desribe the use ase in one sentene with a named input, output, reviewer, and destination, stop the purbase review. The firm is evaluating a demo, not a workflow.

Gate 1 Data classification and minimum access

Classify the proposed data before asking whether the tool is safe. The firm defines its own ategories and must aount for applible law, ethis duties, ontratual restritions, protetive orders, lient instrutions, insurane requirements, and internal poliy.

Example class	Examples	Default evaluation posture
Public	Published law, publiwebsite ontent, public forms	May be suitable for a ontrolled pilot after terms and output review
Internal	Firm proedures, nonpublitemplates, internal operations	Requires approved business aount, aess ontrols, and ontrat review
Client onfidential	Intake fats, lient ommuniations, matter douments, strategy	Requires legal, ontrat, seurity, data-flow, retention, and use-ase approval

Restricted	Credentials, full payment-card data, highly sensitive personal data, sealed material, protective-order data, regulated records	Prohibited unless the firm documents a specific lawful and secure path

Data decision

Question	Answer and evidence
Highest data class proposed	
Data fields necessary for the task	
Fields that must be removed or masked	
Client or third party restrictions	
Jurisdiction specific restrictions	
Minimum user permissions	
Minimum connector permissions	
Prohibited data	

Gate result

[ ] Pass for listed data classes and fields.

[ ] Pass only for public, fictional, or fully de-identified pilot data.

[ ] Fail pending legal or security review.

[ ] Fail for this use case.

Gate 2 Data flow and connected services

A vendor's main privacy page does not describe every path. Map the prompt, uploaded file, retrieval source, model provider, subprocessor, web search, plugin, connector, feedback channel, log, export, and downstream action used by the feature.

Data flow map

Step	System or party	Data received	Purpose	Location	Retention	Evidence
1						
2						
3						
4						
5						
6						

7						
8						

6nnected feature inventory

Feature	Enabled	Data reached	Permission	Separate terms	Decision
Web searb					
Email or alendar					
Doament repository					
Pratie management system					
Browser or omputer use					
Plugins or external tools					
Aents or autonomous ations					
Feedbak or produt improvement					

6itical questions

[ ] Every data destination is identified, intuding optional and feature-level paths.

[ ] The firm an disable unneeded onnetors and permissions.

[ ] Web searb, third party agents, and external tools have been reviewed separately.

[ ] The vendor identifies subproessors and provides a bange-notie proess.

[ ] The firm knows what remains after a user deletes a bat, file, aount, or workspae.

[ ] The firm an export required reords and end the data flow.

Gate 3 6ntract privacy and records

Florida Bar Ethis Opinion 24-1 instruts lawyers using third-party generative Awith onfidential information to investigate data retention, data sharing, self-learning, provider onfidentiality and seurity obligations, breab notie, reputation, seurity measures, liability limits, retention after termination, and laims to submitted information. Reord the atual doament and setion for eab answer.

Requirement	Aswer	Document section	Gap owner and date
Vendor onfidentiality obligation is enforeable			
Customer inputs and outputs are extuded from model training by default or by enforeable setting			

Optional feedback and improvement paths are identified			
Data processing terms and roles are clear			
Retention periods are stated for prompts, files, logs, backups, and abuse monitoring			
Deletion rights, timing, and confirmation are stated			
Ownership and permitted use of inputs and outputs are stated			
Subprocessor list and change notice are available			
Breach and legal-process notice obligations are stated			
Data location and cross-border terms are understood			
Audit, export, suspension, termination, and return rights are stated			
Liability, indemnity, warranty, and service limits are reviewed			
Regulated-data terms apply to the exact features proposed			
Changes to terms and product behavior trigger notice or re-review			

Contract gate result

[ ] Pass for the exact product, account, features, data classes, and use case.

[ ] Pass with the listed contractual or configuration limits.

[ ] Pilot only with public, fictional, or fully de-identified data.

[ ] Fail pending written terms.

[ ] Fail for this use case.

Gate 4 Security access and operations

A audit report or certification is evidence, not a universal approval. Match controls to the firm's data, use case, users, and integrations. Ask for the exact scope, date, exceptions, and products covered.

Control question	Evidence	Result and limits
Can the firm require multifactor authentication and centralized identity?		
Can admins add remove and review users and roles?		
Are least privilege and separate admin roles supported?		

Are data encrypted in transit and at rest?		
Are prompts files outputs admin actions and exports logged?		
Can the firm set retention legal hold and deletion where needed?		
Can the firm restrict sharing publidinks downloads and connectors?		
Are tenant or workspace boundaries documented and tested?		
Are vulnerability management penetration testing and incident processes documented?		
Does independent assurance cover the exact service and period?		
Can the firm receive security and subprocessor change notices?		
Are backup recovery continuity and service exit documented?		
Can the firm rapidly suspend users connectors and automated actions?		

Access review

Role	Users	Data access	Actions allowed	Reviewer	Review date

Gate 5 Aspecific risk screen

Use current Asecurity guidance as a question set, not as a certification. The OWASP GenALLM Top 10 2026 is the current release as of this worksheet's research date. NIST recommends risk management across governance, mapping, measurement, and management, including pre-deployment testing and incident disclosure.

Risk area	Test question	Control or limit	Result
Prompt injection	Can untrusted text instructions alter the task or trigger data access?	Separate trusted instructions, limit tools, test hostile documents	

Sensitive information disclosure	Can the system reveal data to the wrong user, output, log, or connector?	Least privilege, data filters, sharing limits, output review	
Improper output handling	Can generated content execute code, alter records, or reach another system without validation?	Treat output as untrusted, validate, require approval	
Excessive agency	Can the system send, delete, buy, file, schedule, or change records beyond the task?	Minimal tools and permissions, confirmation, logs, shutdown	
Supply chain	What models, plugins, libraries, agents, and subprocessors are involved?	Inventory dependencies, monitor changes, restrict additions	
Data and model poisoning	Can untrusted sources corrupt retrieval, examples, or instructions?	Source approval, provenance, versioning, test sets	
Misinformation and overreliance	Does fluent output hide unsupported facts, citations, or legal analysis?	Primary-source verification and named reviewer	
Model or system extraction	Could prompts or outputs expose system instructions, proprietary content, or secrets?	Do not store secrets in prompts, restrict access, monitor abuse	
Unbounded consumption	Can loops, large files, or repeated actions create unexpected cost or service impact?	Budgets, quotas, timeouts, alerts, manual stop	
Unexpected model change	Can vendor updates change behavior without the firm's test?	Version record, release monitoring, regression tests, re-review	

Agent and action controls

[ ] Every connected tool and permission is listed.

[ ] Read and write permissions are separated where possible.

[ ] The system cannot accept a matter, clear a conflict, give legal advice, send a filing, move money, or delete a record without the firm's named human gate.

[ ] High impact actions require a preview that shows recipient, destination, content, and changed fields.

[ ] Logs identify the user, tool, inputs, outputs, action, approval, and result.

[ ] The admin can disable the agent and revoke credentials quickly.

Gate 6 Output testing and verification

Test the exact task with a controlled set before client use. Use fictional, public, or fully de-identified material until the data gates pass. Include ordinary examples, hard examples, missing information, conflicting sources, and an adversarial input.

Test plan

Test	Source set	Expected behavior	Reviewer	Pass rule
Typical task				
Missing fact				

Conflicting soures				
Outdated authority				
False itation request				
Prompt injetion in soure				
Wrong user or permission				
Unapproved ation request				
Large or ostly request				
Servie unavailable				

Output review record

Criterion	Evidence	Result	Defect and severity
Fats matb the soure			
Every itation resolves to the ited authority			
Athority is orrent for the task and jurisdition			
Unertainty and missing information are visible			
Instrutions and sope are followed			
Sensitive information is not exposed			
Output annot trigger an unapproved downstream ation			
Aqualified reviewer an reprodue the result from the soure			
Failure is detetable before lient ourt or system impat			

Verification protocol

Output type	Named checker	Source checked	Required checks	Destination after approval

Gate 7 Adoption cost and support

A secure tool that nobody uses is waste. A popular tool with no owner is risk. Record the whole operating cost, not only the advertised seat price.

Cost or effort	Assumption	Annual or one time amount	Evidence
Licenses and minimum seats			
Usage or consumption			
Implementation and configuration			
Security legal and contract review			
Integration or connector fees			
Training and supervised practice			
Quality review and monitoring			
Data migration export and exit			
Expected savings or value	Measured baseline required		

Adoption questions

[ ] The target users helped define the task and test set.

[ ] The tool fits the actual workflow and authoritative systems.

[ ] A primary owner and backup can administer it.

[ ] Training covers approved uses, prohibited data, output checks, incidents, and shutdown.

[ ] The firm can support the review burden at expected volume.

[ ] Success is measured against a baseline, not a demo estimate.

[ ] The firm has a fallback process when the service is unavailable.

Value record

Signal	Baseline	Pilot result	Source	Decision

Pilot plan

Field	Entry
Pilot objective	
Exact product feature and configuration	
Allowed users	
Allowed data	
Prohibited data and actions	
Test set and volume	
Start and stop dates	
Output reviewer and sample rule	
Security and activity logs reviewed	
Incident contact and shutdown method	
Success and failure signals	
Go limit or stop meeting date	

Pilot rules

[ ] Use fictional, public, or fully de-identified data unless all live-data gates have passed.

[ ] Test only the named use case and features.

[ ] Keep human approval before every external or high-impact action.

[ ] Preserve the source, output, reviewer decision, defects, cost, and time record.

[ ] Stop on unexpected data access, unapproved transfer, false authority, dangerous action, or loss of control.

[ ] Do not expand the pilot because the first demonstration worked.

Pilot issue log

Date	Test or event	Issue	Impact	Owner	Resolution	Retest

Decision record

Critical gates

Gate	Pass	Pass with limits	Fail or unknown	Evidence reference
Use case and owner				
Data classification				
Data flow and subprocessors				
Contract privacy and retention				
Security and access				
Specific risks				
Output testing and verification				
Adoption cost and support				
Incident and exit plan				

Decision

[ ] Approve.

[ ] Approve with limits.

[ ] Pilot only.

[ ] Reject.

[ ] Needs follow up.

Field	Entry
Approved use cases	

Approved users and roles	
Approved data lasses and fields	
Required settings	
Required human review	
Prohibited uses data and ations	
Monitoring and reporting	
Conditions to resolve	
Deision reason	
Deision date	
Business owner sign off	
Amin or seurity sign off	
Lawyer approver sign off	

Monitoring, review, and exit

Review triggers

[ ] Material ontrat, privay, retention, or subproessor hange.

[ ] New model, feature, onnetor, agent, data soure, or ation permission.

[ ] Change in ownership, hosting, seurity status, or servie sope.

[ ] New data lass, pratie area, jurisdition, or lient restrition.

[ ] Material output defet, inident, omplaint, or unauthorized use.

[ ] Change in applile law, rules, ethis guidane, ourt requirements, or insurer diretion.

[ ] Sheduled review date arrives.

Monitoring plan

Signal	Source	Owner	Frequency	Escalation rule
Users and permissions				
Conneted servies and agents				
Terms and subproessor hanges				
Output defets				
Inidents and near misses				
Usage and ost				

Adoption and measured value				

Exit plan

[ ] Export records the firm must retain in a usable format.

[ ] Confirm the authoritative destination and record owner.

[ ] Disable users, tokens, connectors, agents, plugins, and sharing links.

[ ] Request deletion under the contract and retain confirmation.

[ ] Address backups, logs, legal holds, and subprocessor copies.

[ ] Remove the tool from approved lists, training, templates, and workflows.

[ ] Notify affected users and provide the fallback process.

[ ] Record the final data disposition and closure date.

Vendor document request

Send this list before the demo becomes the decision. Ask the vendor to identify the exact document, section, account type, and feature for each answer.

[ ] Current master agreement, product terms, privacy terms, and data processing addendum.

[ ] Exact statement on use of customer inputs, outputs, files, feedback, and logs for training or product improvement.

[ ] Retention and deletion schedule for prompts, outputs, files, logs, backups, safety review, and terminated accounts.

[ ] Subprocessor list, model providers, hosting locations, and change-notice process.

[ ] Data-flow description for web search, connectors, plugins, agents, retrieval, and external actions.

[ ] Security overview, independent assurance reports, scope, exceptions, and current period.

[ ] Identity, multifactor authentication, role, audit, sharing, export, retention, and administrative controls for the proposed edition.

[ ] Incident response, breach notice, legal-process notice, continuity, recovery, and support commitments.

[ ] Regulated-data terms and a feature-level list of what is included or excluded.

[ ] Release notes and notice process for model, feature, privacy, contract, and security changes.

[ ] Pricing for seats, usage, storage, connectors, agents, support, implementation, export, and termination.

[ ] Test environment or pilot terms that allow the firm to use non-client data and preserve results.

Four questions before approval

1. What exact data will this feature receive, where will it go, and what will remain after deletion or termination?
2. What enforceable terms and settings control training, confidentiality, security, access, subprocessors, and incidents?
3. How will a qualified person verify the output or approve an action before it affects a client, court, counterparty, system, or payment?
4. Can the firm explain and document this use to a client, regulator, court, insurer, or bar authority if asked?

Source notes

Sources were checked September 7, 2026. They are starting points for the firm's own review.

Florida Bar Ethics Opinion 24-1. Advisory guidance on confidentiality, vendor investigation, supervision, output verification, fees, advertising, and intake. Official source

ABA Formal Opinion 512. Model Rules guidance on competence, confidentiality, communication, supervision, candor, and fees. Official source

NIST AI Risk Management Framework 1.0. Voluntary and non-prescriptive framework organized around Govern, Map, Measure, and Manage. Official source

NIST Generative AI Profile. Cross-sector profile addressing governance, provenance, pre-deployment testing, incident disclosure, and other generative-AI risks. Official source

NIST Cybersecurity Framework 2.0. Cybersecurity outcomes covering governance, identification, protection, detection, response, recovery, and supplier risk. Official source

OWASP GenAI LLM Top 10 2026. Current OWASP security-risk release for LLM and generative-AI applications. Official source

Dated vendor policy examples

The following examples show why the exact product context matters. They do not approve any service for a law firm or client data.

OpenAI business data commitments. OpenAI states that business-product inputs and outputs are not used for training by default. Official source

OpenAI individual data controls. Individual users can change model-improvement controls, and Temporary Chat has separate behavior. Official source

Microsoft Copilot privacy and protections. Microsoft describes enterprise data protection for organizational prompts and responses and separate considerations for web search. Official source

Google Workspace Gemini FAQ. Google describes protections for qualifying Workspace accounts and the importance of license and administrator controls. Official source

Anthropic commercial data use. Anthropic states that commercial inputs and outputs are not used for training by default, with separate treatment for feedback or opt-in. Official source

The next action

Choose one proposed tool and one narrow task. Complete the use-case, data, and data-flow gates before the next vendor meeting. Send the document request in advance. If the vendor cannot answer in writing, record the gap and limit the evaluation to non-client data.

Want help evaluating a product or designing the pilot? Schedule a complimentary 15 minute introductory call.

Schedule with Greg Yates. Current Actionable AI scheduling page, checked September 7, 2026. Official source

Actionable AI helps small law firms fix one workflow at a time using the systems they already pay for.

Actionable AI

Download the free Florida AI Starter Playbook.

Eight practical resources for managing partners and firm administrators: a readiness checklist, five pilot workflows, a tool-approval scorecard, a lawyer verification checklist, and a 90-day plan. It is free, and the button below downloads it directly.

Scan to download the Starter Playbook

Download the Starter Playbook (PDF)

Direct download: Florida AI Starter Playbook (PDF)

Want a second set of eyes on which workflow to pick?

I take 15-minute calls with managing partners and administrators. I will tell you whether there is a fit. No deck.

Book a 15-minute introductory call

Book a 15-minute introductory call