

# The AI Document Review Playbook

---

Cut Document Review Time by 70%  
Without Sacrificing Accuracy

A practical field guide for law firms ready to justify, deploy, and defend AI-assisted review in 2026.

**70%**

Reduction in review time with TAR + GenAI

**\$13B**

Spent on document review by 2029

**92%**

Accuracy rate achieved with CAL workflows

## Why “More Reviewers” Is the Wrong Answer

The legal industry has spent decades solving a math problem the wrong way. When review sets grew from thousands to millions of documents, firms responded by hiring more contract attorneys. But the math has officially broken. At 500 documents per reviewer per day, a generous estimate, a 2-million-document set needs 4,000 reviewer-days. No deadline survives that.

The real problem isn't reviewer headcount. It's signal-to-noise ratio. Industry data consistently shows that only 1–3% of collected documents are actually relevant to a matter. AI review is the only mechanism that fixes the ratio problem before it becomes a cost problem.

## The Signal-to-Noise Crisis in Legal Review

### THE OLD WAY

Linear review. Every document gets equal time. Reviewers spend 97% of their hours on documents that will never matter to the case. The cost is baked in and unavoidable.

Result: \$18–\$22/GB reviewed, 6–8 weeks turnaround

### THE SMART WAY

AI-prioritized review. AI surfaces the most relevant documents first. Reviewers spend their hours where it matters. Irrelevant documents are culled before a human ever opens them.

Result: \$3–\$6/GB reviewed, 1–2 weeks turnaround

## Your AI Toolkit: TAR, CAL, and GenAI — What Each Does

Not all AI reviews are the same. Law firms that conflate TAR 1.0, CAL, and generative AI tools end up deploying the wrong technology for the wrong task. This table is the simplest way to separate the tools by purpose and defensibility.

TOOL	BEST FOR	COURT ACCEPTANCE
<b>TAR 1.0 (Simple Passive Learning)</b>	Fixed training sets, known-item search, matters where the issue scope is stable from day one	Well-established. Endorsed in <i>Da Silva Moore</i> (2012) and <i>Rio Tinto v. Vale</i> (2012)
<b>CAL / TAR 2.0 (Continuous Active Learning)</b>	Rolling training on reviewer decisions. Ideal for large, evolving matters with shifting issue codes.	Broadly accepted. Grossman & Cormack's research shows CAL outperforms random sampling at scale.
<b>GenAI (LLM-Assisted Review)</b>	Summarization, privilege log drafting, and issue spotting in complex documents (contracts, technical filings).	Emerging. Courts expect human-in-the-loop validation. Transparency in methodology is non-negotiable.

# The 70% Time Reduction and Where It Actually Comes From

PHASE	AI INTERVENTION	TIME SAVED
<b>Pre-review culling</b>	AI deduplication + email threading + language filtering	25–40% volume reduction
<b>First-pass relevance</b>	CAL prioritizes high-relevance documents; low-confidence docs batched separately	30–50% reviewer hours
<b>Privilege review</b>	GenAI drafts privilege log entries; attorney validates rather than drafts from scratch	60–70% privilege log time
<b>QC &amp; validation</b>	Statistical sampling + AI confidence scoring replaces full second-pass review	40–55% QC hours

## Justifying AI Review to Partners & Courts

The #1 reason law firms delay AI adoption isn't technical, it's political. Senior partners worry about malpractice exposure. Courts worry about reliability. Here's a framework to address both audiences.

### Audience 1: Your Managing Partner

Partners respond to risk-adjusted ROI. Frame AI review not as a cost-cutting exercise, but as a risk management upgrade. Use this three-point structure:

01

#### The Accuracy Argument

Courts have consistently found AI-assisted review to be at least as accurate as manual review, not as a concession, but as a documented finding. Cite Grossman & Cormack: In head-to-head tests, CAL achieved higher recall than manual review. Lower miss rate = lower malpractice risk.

02

#### The Cost Argument

Budget overruns on review are the #1 client complaint in litigation. AI review delivers predictable, bounded costs. Present a before/after model: linear review at \$0.08–\$0.12/page vs. AI-assisted at \$0.01–\$0.03/page. For a 5M-document matter, that delta runs into seven figures.

03

## The Speed Argument

Regulatory agencies expect productions in weeks. Clients demand responsiveness. AI review compresses timelines without burning out review teams. Speed is now a competitive differentiator for firms, not just a nice-to-have.

## Audience 2: The Court

Judicial acceptance of AI review hinges on one word: transparency. Courts don't reject AI, they reject opacity. Every defensible AI review protocol includes these five elements:

#	ELEMENT	WHAT TO DOCUMENT
1	<b>Methodology disclosure</b>	Describe the AI tool used, the training approach (TAR1.0 vs. CAL), and the seed set composition. Courts expect this upfront in your ESI protocol.
2	<b>Training set transparency</b>	Document how the training set was built, who coded it, and how inter-reviewer consistency was measured. Golden sets and judgment calls should be logged.
3	<b>Statistical validation</b>	Run elusion testing. Pull random samples from the "not relevant" set and verify AI decisions. Report recall and precision rates. Target recall $\geq 75\%$ , per EDRM guidance.
4	<b>Human-in-the-loop confirmation</b>	Every AI-assisted workflow needs a senior attorney sign-off layer. Courts require that a qualified professional is accountable for final decisions, not the algorithm.
5	<b>Iterative logging</b>	Maintain a review log that tracks model versions, training rounds, and decision changes over time. This is your audit trail if opposing counsel challenges the process.

# The 5-Step AI Review Deployment Playbook

Firms that fail with AI review usually skip steps 1 and 2. Firms that succeed treat AI deployment as a workflow redesign, not a software installation.

## STEP 1

### Scope Before You Train

Define your issue codes, custodian list, and date ranges before touching the AI tool. Garbage in = garbage model. The best AI review fails when the issue taxonomy is ambiguous. Run a 1-day scoping session with case team leads before training begins.

→ Deliverable: Issue taxonomy document + custodian priority matrix

## STEP 2

### Run Keyword Culling First (Yes, Still)

AI and keyword search are not competitors, they're sequentially complementary. Use targeted keywords to cull obvious noise (newsletter unsubscribes, automated system alerts, out-of-scope custodians) before training your model. This shrinks the training universe, improving model accuracy.

→ Deliverable: Pre-AI culling report showing volume reduction percentage

## STEP 3

### Build a Defensible Training Protocol

Seed your model with 500–1,000 attorney-coded documents that represent the full spectrum of relevance (not just obvious responsive docs). Include near-miss examples. Validate inter-coder agreement  $\geq 80\%$  before the first training round. Document everything.

→ Deliverable: Training protocol memo (shareable with opposing counsel if challenged)

#### STEP 4

### Review Smart, Not Linear

Use CAL's ranked output to prioritize your review queue. Review the top-ranked documents first, continue feeding decisions back to the model, and stop when elusion testing confirms the remaining set is below your recall threshold. Do not review everything.

→ Deliverable: Elusion test results + stopping criteria memo

#### STEP 5

### Automate the Privilege Log

Deploy GenAI to draft privilege log entries from coded privilege documents. Attorney prompts the model with: document type, date, parties, and identified privilege basis. Attorney reviews and approves every entry. This alone can cut privilege log time by 60%.

→ Deliverable: AI-assisted privilege log with attorney certification

## The Accuracy Equation: What “Good Enough” Actually Means

One of the most persistent myths about AI review is that 95% recall is the mandatory standard. It isn't — and treating it as such keeps firms trapped in expensive hybrid workflows that defeat the efficiency gains.

#### The FRCP Standard

FRCP 26(b)(1) requires proportional efforts, not perfect recall. Courts evaluate whether the review was reasonable relative to the matter's stakes, not whether every relevant document was found. This is the legal foundation for stopping review before 100% coverage.

#### The Industry Benchmark

EDRM guidance suggests targeting recall  $\geq 75\%$  as a reasonable threshold for most matters. High-stakes SEC/DOJ investigations may warrant higher targets. Routine civil discovery may accept lower, if documented and agreed. The key is elusion testing, not an arbitrary recall floor.

## GenAI Guardrails: What Attorneys Must Control

GenAI in document review is powerful and legally treacherous in equal measure. These are the non-negotiable guardrails for any firm using LLMs in review workflows:

### ✓ DO THIS

- ✓ Use GenAI for summarization and privilege log drafting, tasks with attorney review at the end
- ✓ Establish a system prompt that defines the legal matter context, confidentiality requirements, and output format
- ✓ Validate AI outputs with statistical sampling before relying on them in production
- ✓ Disclose GenAI tool usage in your ESI protocol, where opposing counsel may reasonably inquire
- ✓ Maintain prompt logs and output versions as part of your audit trail

### ✗ AVOID THIS

- ✗ Don't use GenAI to make final relevance calls without attorney review, courts require human accountability
- ✗ Don't feed privileged documents into public LLM APIs; use on-premise or contractually secured environments only
- ✗ Don't present AI-generated privilege log entries as attorney work product without reviewing and certifying each entry
- ✗ Don't skip validation because the model "seems accurate"; elusion testing is your legal protection, not optional
- ✗ Don't deploy GenAI on a matter without a written AI usage policy approved by firm leadership

# The ROI Case: Numbers Your CFO Will Believe

Abstract efficiency claims don't move law firm budgets. Concrete math does. Here's a model you can adapt to your last major matter and present as a business case for AI review investment.

SAMPLE MATTER ROI MODEL – 1 MILLION DOCUMENT REVIEW		
METRIC	LINEAR REVIEW	AI-ASSISTED REVIEW
Documents reviewed by humans	1,000,000	200,000–350,000
Average review rate (docs/hour)	40–60	40–60 (on prioritized set)
Reviewer hours required	~18,000 hrs	~5,000–7,000 hrs
Contract attorney cost (@\$50/hr billed)	~\$900,000	~\$250,000–\$350,000
Platform + AI tooling cost	\$0 (manual)	~\$40,000–\$80,000
Total estimated review cost	\$900,000+	\$290,000–\$430,000
Timeline	10–14 weeks	3–4 weeks
Estimated savings	—	\$470,000–\$610,000 (52–68%)

 **NEXT STEP**

Ready to cut review time by 70%? See how Venio's AI-powered platform puts this playbook into practice. → Book a 30-Minute Live Demo at [veniosystems.com](https://veniosystems.com)

## Verified Sources & Further Reading

- 1 Grossman & Cormack, “Technology-Assisted Review in E-Discovery” — Richmond J. of Law & Tech. (2011)
- 2 Da Silva Moore v. Publicis Groupe, 287 F.R.D. 182 (S.D.N.Y. 2012)
- 3 Rio Tinto PLC v. Vale S.A., 306 F.R.D. 125 (S.D.N.Y. 2015)
- 4 Hyles v. New York City, No. 10 Civ. 3119 (S.D.N.Y. Aug. 1, 2016)
- 5 ABA Formal Opinion 512 on Generative AI (2023) — [americanbar.org](https://www.americanbar.org)
- 6 EDRM TAR Framework & Guidelines — [edm.net/resources/frameworks-and-standards](https://edm.net/resources/frameworks-and-standards)
- 7 ComplexDiscovery 2024 Summer Pricing Report — [complexdiscovery.com/whats-new-in-ediscovery-pricing-summer-2024-market-report/](https://complexdiscovery.com/whats-new-in-ediscovery-pricing-summer-2024-market-report/)
- 8 Ari Kaplan Advisors, AI in Legal: 2024 Benchmark Study — [arikaplanadvisors.com](https://arikaplanadvisors.com)
- 9 Relativity 2023 State of E-Discovery Report — [relativity.com/resources/](https://relativity.com/resources/)
- 10 EDRM Business Confidence Survey Summer 2024 — [edm.net/2024/10/ediscovery-challenges-impacting-business-performance](https://edm.net/2024/10/ediscovery-challenges-impacting-business-performance)

VENIO

[veniosystems.com](http://veniosystems.com)