



Smarter Approaches to Document Review

Cut time, control costs, and stay defensible in 2025

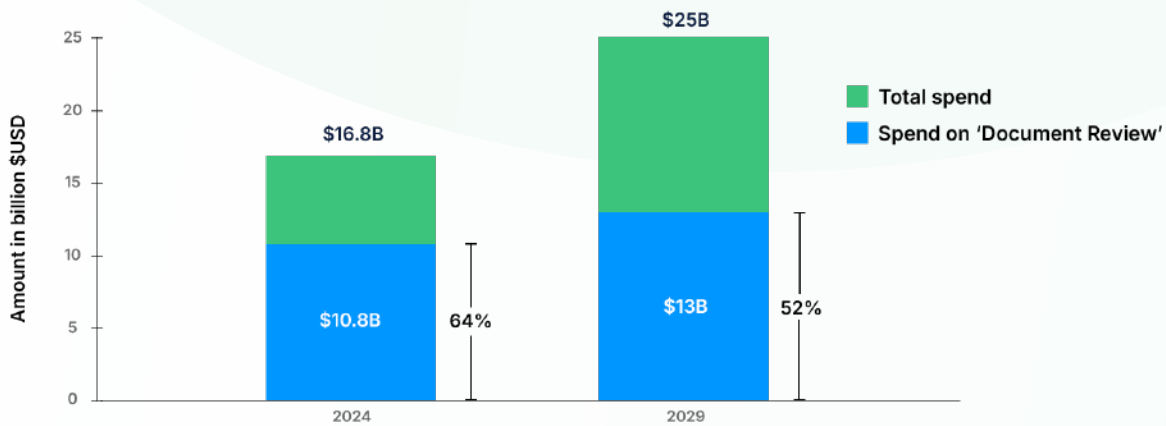
Inside the Guide

	Page
1. The Cost of Review	03
2. Challenge 1: Volume Overload	05
3. Challenge 2: Time Pressure	07
4. Challenge 3: Cost vs Accuracy	09
5. Smart and Practical Ways to Improve Review	11
6. Smart Document Review Starts Here	14

The Cost of Review

Did You Know 64% of eDiscovery Spend Goes Here?

As of 2024, document review accounts for 64% of total eDiscovery spending¹, while collection has risen to 16% and processing has stabilized at 20%. This makes review the single largest cost driver in litigation. In other words, most eDiscovery dollars aren't spent finding or processing data, they're spent reading it.



2024 vs. 2029 comparison of total spend on 'Document Review'

The Data Explosion

- By 2025, **global data creation will exceed 463 exabytes every single day** (World Economic Forum)².
- The **average organization now uses 112 SaaS applications**, each producing its own stream of files, chat logs, and metadata³.
- For legal teams, this translates into exponentially larger review sets — millions of files, often duplicative or irrelevant.

The Shift Toward Precision

By 2025, 70% of organizations will pivot from chasing “big data” to adopting “small and wide data” strategies — focusing on context-rich, diverse information sources for faster, more relevant insights. (Gartner 2024)⁴.

For legal teams, this means the challenge isn't just collecting more data. It's about finding the right data that drives defensible, cost-effective decisions.

The Pressures on Legal Teams

- Scale → Review sets now routinely include millions of files across email, chat, cloud, and mobile sources.
- Speed → Courts, regulators, and clients expect fast, accurate turnaround with little room for inefficiency or error.

Why It Matters

Document review is the single largest cost driver in eDiscovery. Even incremental gains in efficiency here can deliver disproportionate savings, making smarter review approaches essential for legal teams to stay defensible, efficient, and cost-conscious.



Challenge 1: Volume Overload

The single greatest obstacle in modern document review is sheer scale. What once meant reviewing boxes of paper or a few gigabytes of email has now evolved into terabytes of data across dozens of digital sources, all of which must be identified, processed, and reviewed defensibly.

The Data Explosion

- By **2025**, global data creation will exceed **463 exabytes every single day** (World Economic Forum)².
- Although **20.29% of organizations overall cite data growth as a top eDiscovery challenge, concern fell to 15% during Summer 2024**⁵.
- For legal teams, this translates into exponentially larger review sets — millions of files, often duplicative or irrelevant.

Legal Implications

Rising data volumes are not just an IT challenge, they create real legal risk:

- **Proportionality pressures** – Under FRCP 26(b)(1), it is increasingly difficult to prove proportionality when review sets include millions of files⁶.
- **Overproduction risk** – Disclosing irrelevant or privileged documents exposes sensitive data and may weaken legal strategy.
- **Underproduction risk** – Missing or withholding key evidence can trigger sanctions.
- **Regulatory urgency** – Agencies like the SEC, DOJ, and FTC now expect productions in weeks, not months.

Technical Complexity

Beyond sheer size, modern review sets introduce structural challenges:

- **Duplication** – A single email thread may appear in 20+ versions across multiple custodians
- **Unstructured data** – Platforms like Slack, Teams, and Zoom generate logs, chats, and media with limited metadata, making review cumbersome.
- **Dynamic files** – Cloud-based apps constantly update documents, complicating version tracking and defensibility.



Why It Matters

Volume overload isn't just a matter of "too much data." It creates cost pressure, technical strain, and legal exposure - all of which threaten defensibility.

Key Insight:

Without proactive strategies such as: AI-assisted culling, deduplication, and early case assessment, the rising tide of data threatens to overwhelm even the best-prepared legal teams. The organizations that succeed will be those that control scale early in the review process.

Challenge 2:

Time Pressure

If volume is the biggest obstacle, time is the sharpest constraint. Modern matters involve millions of files, yet the rules and regulators set tight, non-negotiable deadlines that teams must meet.

The Clocks That Matter

- **30 days to respond to document requests (RFPs)** – Under FRCP 34, parties must respond within 30 days (or 30 days after the Rule 26(f) conference if served early)⁷.
- **14 days for initial disclosures** – FRCP 26(a)(1)(C) requires disclosures within 14 days of the Rule 26(f) conference, unless otherwise agreed⁶.
- **20 days to petition CIDs** – Recipients of FTC Civil Investigative Demands (CIDs) generally have 20 days to petition to limit or quash; return dates are near-term and extensions are disfavored⁸.

Legal Implications

- **Proportionality + diligence (FRCP 26(b)(1))** – Courts expect reasonable, proportional efforts despite short timelines; incomplete or late productions risk sanctions⁶.
- **Regulatory leverage** – Agencies control CID deadlines; failure to show good-faith progress can trigger escalated demands.
- **Underproduction risk** – Missing or withholding key evidence can trigger sanctions.
- **Regulatory urgency** – Agencies like the SEC, DOJ, and FTC now expect productions in weeks, not months.

Technical Complexity

- **Linear review can't match the clock** – Traditional “eyes-on-every-document” review is impossible with millions of files and a 30-day response window.
- **Concurrency and speed matter** – Slow-loading infrastructure or limited reviewer concurrency wastes precious time under fixed FRCP/CID deadlines.
- **Dynamic files** – Cloud-based apps constantly update documents, complicating version tracking and defensibility.



Why It Matters

Time pressure in discovery isn't anecdotal; it's codified in the rules. FRCP deadlines (14 and 30 days) and CID regimes with 20-day petition windows force teams to execute at speed. Meeting these clocks requires front-loaded culling, analytics, and AI-assisted prioritization, and not just more reviewer hours.

Key Insight:

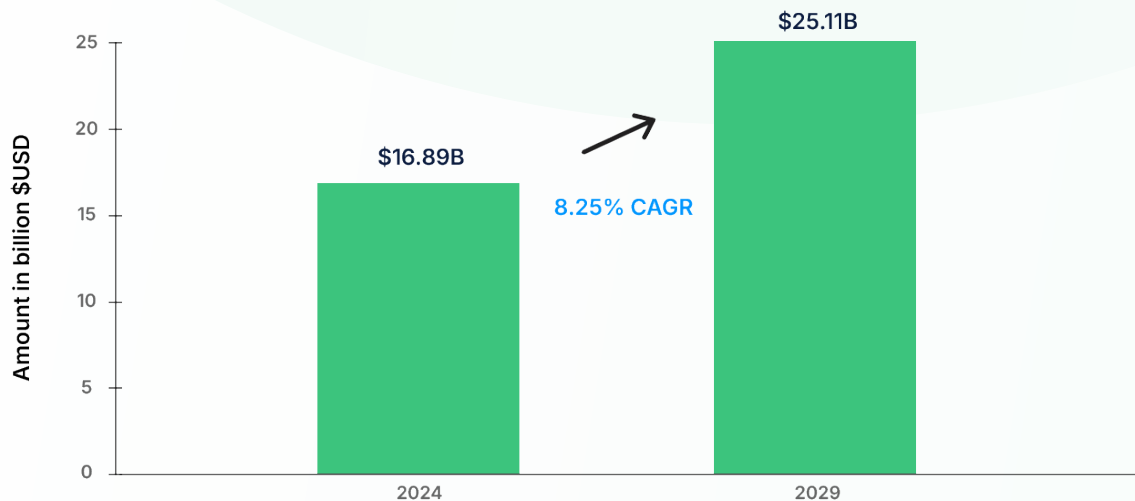
Deadlines in discovery are fixed and unforgiving. Success depends on building defensible, technology-driven workflows that can move at regulatory speed.

Challenge 3: Cost vs. Accuracy

The third, and perhaps most punishing, challenge in document review is balancing rising costs with the uncompromising need for accuracy.

The Cost Pressure

- The global eDiscovery market will grow from \$16.89B (2024) to \$25.11B (2029) at an 8.25% CAGR⁹.



2024 vs. 2029 comparison of Global eDiscovery Market

- **31.67% of eDiscovery professionals** cite budget constraints as their top challenge, **up from 25.9%** earlier¹⁰.
- Review costs remain steep in 2025:
 - \$250–\$350/hour for forensic examiners.
 - \$100–\$200/hour for project management¹¹.

Legal Implications

- **Missed Privilege = Waiver Risk** – Producing privileged material without clawback protections risks permanent waiver.
- **Overproduction = Data Exposure** – Oversharing can expose trade secrets, PII, PHI, or regulatory-protected data.
- **Errors = Sanctions** – In SEC v. Collins & Aikman (2007), inadequate review practices led to sanctions.

Technical Complexity

- **Linear review inefficiency** – Traditional “eyes-on” review is unsustainable at scale.
- **Duplicative content** – Without deduplication/near-duplicate detection, teams waste effort.
- **Complex data types** – Chat logs, Zoom transcripts, CAD files, and other non-standard formats demand specialized handling.
- **Redaction risks** – Files like Excel pose unique redaction challenges; manual workflows raise leak risk.

Why It Matters

Costs might be tolerable if accuracy could be sacrificed, but in discovery, accuracy is non-negotiable. The cost-accuracy tradeoff is the heart of eDiscovery risk: cutting corners risks sanctions, while overspending can swallow the budget.

Key Insight:

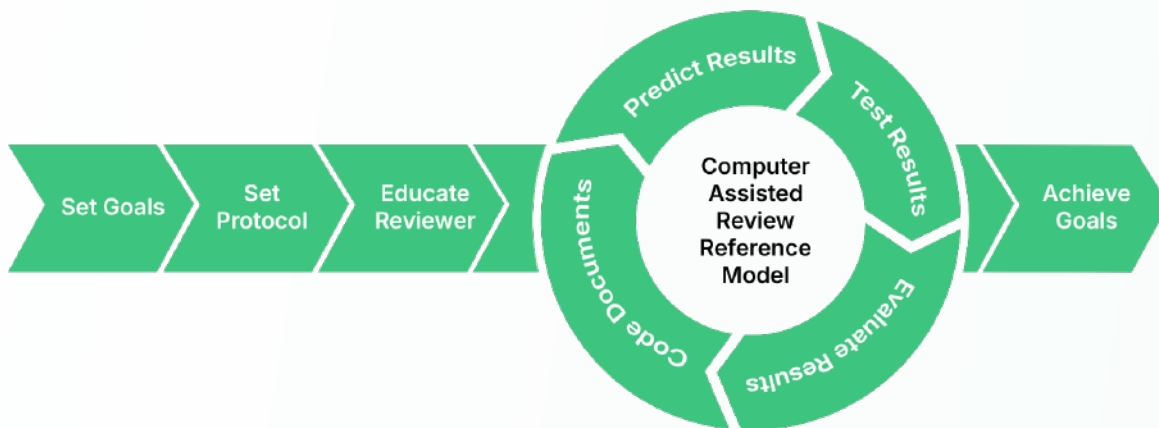
Document review is both the biggest cost driver and the highest risk vector in eDiscovery. The only sustainable path is leveraging AI, analytics, and defensible automation to contain costs without sacrificing accuracy.

Smart and Practical Ways to Improve Review

While the challenges of volume, time, and cost are daunting, proven strategies and technologies exist to make document review more efficient, accurate, and defensible.

Adopt TAR / Continuous Active Learning (CAL)

- Technology-Assisted Review (TAR) and Continuous Active Learning (CAL) consistently deliver higher accuracy and lower review volumes compared to linear, manual review.
 - **Evidence** – Grossman & Cormack’s seminal research and follow-up evaluations show TAR can equal or exceed human precision/recall while cutting human review hours dramatically.
 - **Impact** – Smaller review sets, faster results, and better defensibility.



Major Steps in the Technology Assisted Review (TAR) Framework Process

Front-load Early Case Assessment (ECA)

- Early Case Assessment (ECA) uses sampling, rapid search, custodian analysis, and analytics to shrink datasets before review.
 - **Best Practice** – Triage custodians, date ranges, and data types early.
 - **Why It Matters** – Many cases settle or narrow quickly; early culling preserves defensibility and delivers outsized cost savings.

Deduplication & Near-Duplicate Detection

- Automated deduplication and near-dedupe remove redundant documents before human review.
 - **Typical Result** – Data *reductions of 30–40%* (practitioner-reported averages)¹².
 - **Benefits** – Fewer files to host and review, reducing costs at the outset.

Clustering, Smart Batching & Auto-Tagging

- Modern tools accelerate review through intelligent grouping and tagging:
 - **Clustering** – Groups similar documents for bulk tagging.
 - **Smart batching** – Keeps reviewer focus by mixing prioritized files.
 - **Auto-tagging** – Applies issue codes across near-identical content.
 - **Outcome** – Faster, more consistent coding decisions with reduced cognitive load.

Native Redaction & Specialized Viewers

- Redacting directly within native files (Excel, PowerPoint, multimedia) avoids error-prone export/reimport loops.
 - **Value** – Saves time, prevents hidden data exposure, and improves defensibility.

Scalable Cloud or Hybrid Deployments

- Flexible cloud and hybrid infrastructure provides elastic capacity that scales with matter size.
 - **Advantage** – Avoids fixed infrastructure costs.
 - **When Combined With** – Aggressive culling + TAR, this model keeps hosted data (and costs) to a minimum.

Modern Pricing & Market Trends

- Vendor surveys show material reductions in processing and hosting costs:
 - **Processing** – \$25–\$75/GB (completion-based)¹¹
 - **Hosting** – \$10–\$20/GB/month (analytics off)¹¹
 - **Insight** – Selecting modern providers and culling aggressively before hosting can reduce run-rate costs substantially.

Key Takeaway:

Smarter approaches don't just cut costs — they reduce risk, accelerate timelines, and strengthen defensibility. Teams that adopt TAR, ECA, and scalable infrastructure will be best positioned to handle today's discovery demands.

Smart Document Review Starts Here

Document review doesn't have to drain time, budgets, or teams. With smarter approaches and the right platform, you can cut review costs, meet deadlines, and ensure defensible outcomes.

Why Venio Systems

Venio Systems delivers smarter, faster, and more defensible eDiscovery. Our end-to-end platform streamlines collection, processing, review, and production, so legal teams can:

- Reduce costs without sacrificing accuracy
- Handle rising data volumes with confidence
- Protect sensitive information with enterprise-grade security



What Sets Us Apart

- **Smarter Review** – Built-in AI, automation, and predictive coding reduce review time.
- **End-to-End Platform** – From collection to production in one streamlined solution.
- **Defensible Security** – PII/PHI redaction, compliance, and enterprise-grade controls.
- **Trusted Worldwide** – Law firms, corporations, and government agencies rely on Venio for their most critical matters.

Key Insight:

Smarter review isn't about choosing between speed, cost, or accuracy. With Venio, you can achieve all three — while staying defensible in every matter.

See Venio Review in action

Book a 30-Minute Live Demo



Data Sources

1. <https://www.jdsupra.com/legalnews/2025-ediscovery-review-update-tasks-7435306/>
2. <https://www.weforum.org/stories/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/>
3. <https://www.prnewswire.com/news-releases/the-number-of-saas-applications-at-companies-declined-for-the-first-time-in-over-a-decade-302199899.html>
4. <https://www.gartner.com/en/newsroom/press-releases/2021-05-19-gartner-says-70-percent-of-organizations-will-shift-their-focus-from-big-to-small-and-wide-data-by-2025>
5. <https://edrm.net/2024/10/ediscovery-challenges-impacting-business-performance-insights-from-the-summer-2024-and-aggregate-business-confidence-surveys/>
6. https://www.law.cornell.edu/rules/frcp/rule_34
7. https://www.law.cornell.edu/rules/frcp/rule_26
8. <https://www.ftc.gov/business-guidance/blog/2018/01/so-you-received-cid-faqs-small-businesses>
9. <https://www.linkedin.com/pulse/complete-look-complexdiscoverys-2024-2029-ediscovery-market-robinson-ex5ec/>
10. <https://complexdiscovery.com/ediscovery-challenges-impacting-business-performance-insights-from-the-summer-2024-and-aggregate-business-confidence-surveys/>
11. <https://complexdiscovery.com/whats-new-in-ediscovery-pricing-summer-2024-market-report/>
12. <https://www.allaboutediscovery.com/2017/01/what-is-de-duplication-and-how-do-i-do-it/>

VENIO