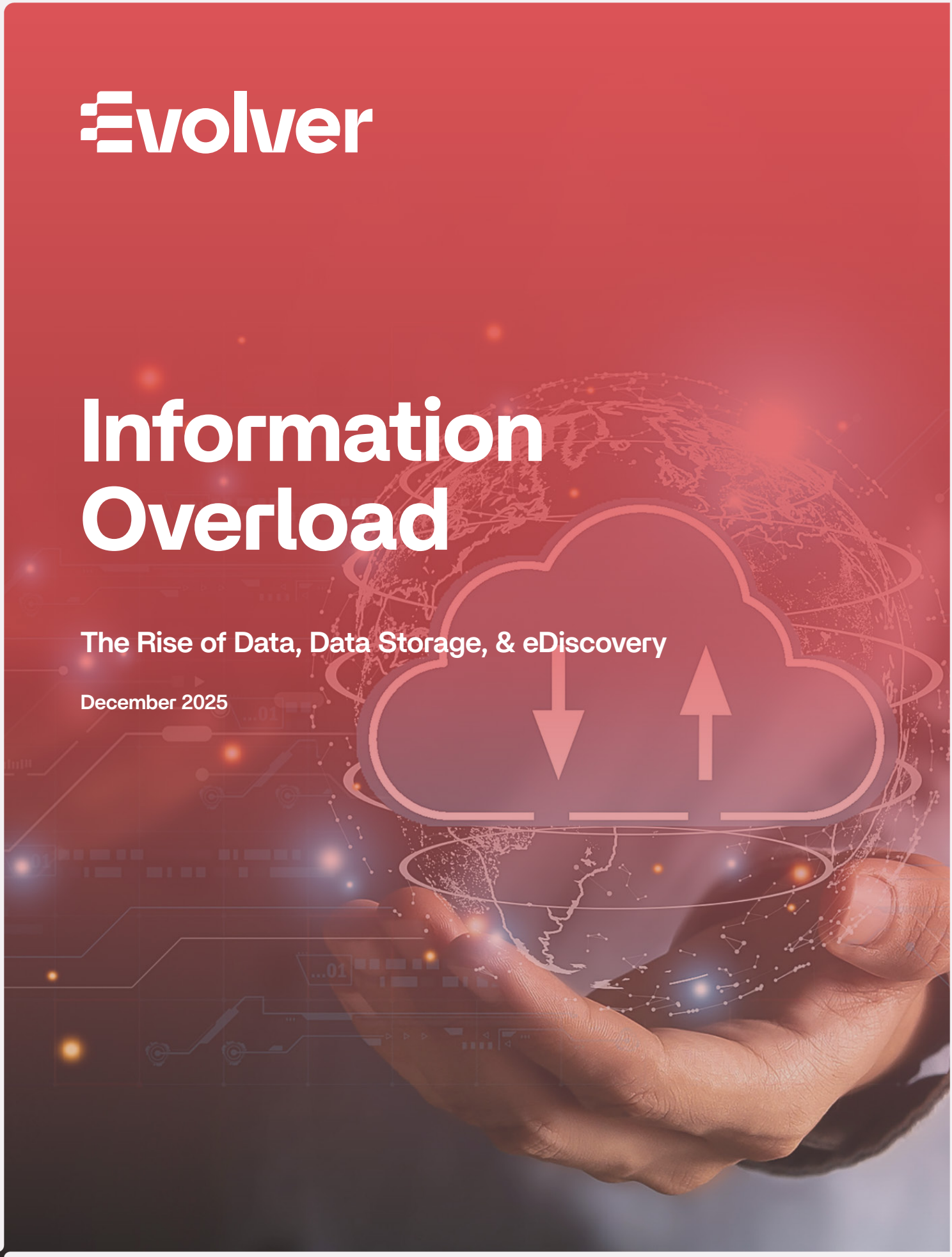
Evolver

Information Overload

The Rise of Data, Data Storage, & eDiscovery

December 2025



Executive Summary

Public and private organizations are contending with an unprecedented surge of digital information. U.S. Federal government mandates are accelerating the transition to electronic records and electronic storage of information (ESI), while the artificial intelligence (AI) revolution and the growth of quantum computing are pushing compute, storage, and networking to new limits.

At the same time, Freedom of Information Act (FOIA) requests and digital evidence requirements are rising, driving demand for AI-powered eDiscovery capabilities that can sift vast data estates quickly, defensibly, and affordably.

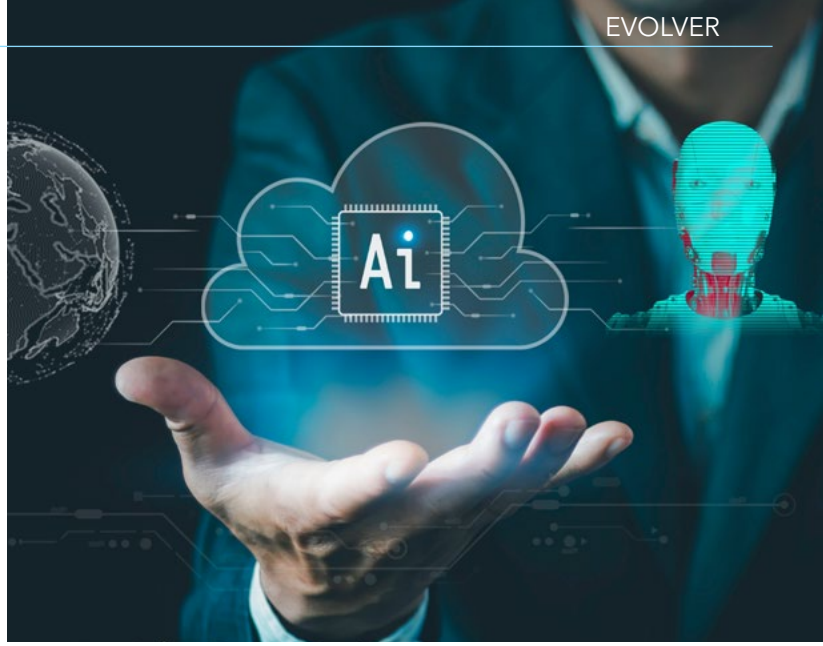
In this paper, we'll outline the forces behind information overload; the implications for data centers, governance, and legal/FOIA workflows; and a pragmatic approach for agencies to adopt modern eDiscovery AI-powered platforms augmented by dashboards, chat interfaces, and automated evidence packaging — all without disrupting mission operations. [1]

The Rapid Growth of Data

Estimates of the "Global Datasphere" continue to climb. While methodologies differ, major analysts converge on a range of roughly 175–181 zettabytes by the end of 2025. Drivers for this explosive growth in data include cloud adoption, mobile communications, Internet of Things (IoT) devices, and machine-generated telemetry. This is not just more data; it is more diverse (text, chat, audio, video, sensor), more distributed (on-premises, cloud, edge), and more regulated (privacy, retention, accessibility). [2]

In the U.S. Federal government sector, that growth meets policy: Office of the Management & Budget (OMB)'s M-23-07 reinforced the requirement that by June 30, 2024, agencies manage permanent records, electronically cementing the operational shift from paper to digital repositories and ESI-ready processes. [3]

In addition to federal mandates, commercial enterprises such as financial management firms, supply chain management companies, and law firms are all experiencing significant data growth, driving demand for scalable eDiscovery solutions that can handle both regulatory and business pressures. Exploding data volumes increase the scope, cost, and risk of litigation, investigations, and FOIA responses; without smarter approaches to collection, culling, review, and production, costs and timelines become untenable.

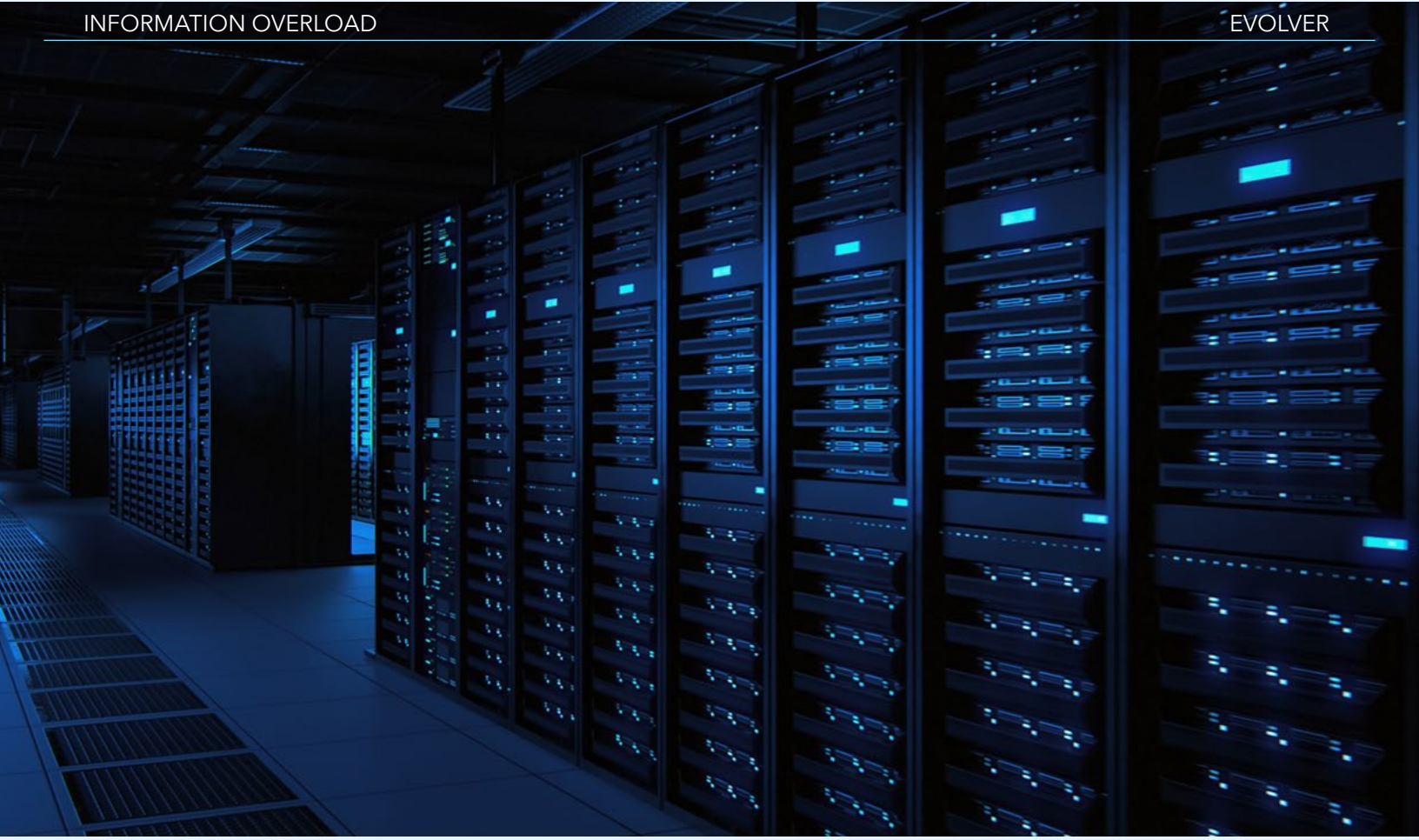


The AI Revolution: More Data and Fast Analysis

Generative AI and machine learning (ML) models intensify both data creation (logs, embeddings, prompts, outputs) and data gravity (centralized high-bandwidth storage with low-latency access). They also enable accelerated sense-making, from semantic search and clustering to technology-assisted review (TAR) and continuous active learning (CAL) that prioritize likely-relevant material for human reviewers.

Operationally, AI is already reshaping infrastructure: the International Energy Agency estimates data centers consumed ~1.5% of global electricity in 2024 and projects sustained growth as AI workloads scale. [4]

Government agencies can augment AI-powered eDiscovery with intuitive dashboards and chatbot interfaces, enabling analysts and attorneys to query authorized corporate documentation in natural language and accelerate decision-making. Government agencies need AI-assisted eDiscovery that is explainable, auditable, and aligned with U.S. Department of Justice (DOJ) expectations. [5]



The Rise of Data Centers

The physical backbone of information overload is expanding rapidly. Hyperscale data centers represent the core hosting platforms for cloud and AI. These centers have doubled in count over five years to more than 1,100 facilities by the end of 2024, with the U.S. accounting for the largest share of capacity.

Analysts expect capacity to nearly triple by 2030, with Generative AI driving most of the new demand. For public-sector missions, this means greater elasticity for bursty FOIA/eDiscovery workloads and near-real-time analytics for investigations, provided governance and cost controls keep pace. [6]

Data centers are not only the backbone of AI workloads but also the repositories for Electronic Storage of Information (ESI), enabling agencies to transition from paper-based archives (e.g., Iron Mountain) to digital repositories aligned with FOIA and investigative needs.

Key data center operational aspects include: planning and providing power, managing the sustainability constraints, ensuring data residency and cross-border transfer legal limitations, and meeting the need for enterprise metadata management so discovery platforms can find the right data, fast. [7]

The Need for AI-Powered eDiscovery Tools

Market data confirms the shift: the eDiscovery market is growing steadily (8%-10% CAGR in the U.S.; double-digit globally), reflecting both ESI expansion and adoption of AI-driven workflows that reduce unit costs of review.

FOIA volumes in the U.S. Federal & State governments reached a record of over 1.5 million requests in FY 2024 (a 25% year-over-year surge) putting additional pressure on U.S. Federal and State government agency tools and staffing. [8]

Legal and policy guardrails are stable and well-understood: the Federal Rules of Civil Procedure (Rules 26 & 34) define proportional discovery and ESI production, while the Sedona Principles provide best practices on scope, cooperation, preservation, and the defensible use of advanced analytics (including TAR/CAL). [9, 10]

AI-powered eDiscovery tools are increasingly vital for DOJ, FBI, Army JAG, Navy NCIS, and DHS/ICE, where rapid triage and defensible evidence packaging are mission critical.

Government agencies such as IRS, DOI, DOL, SSA, FDA, HUD, and NIH face rising FOIA and data management requirements, necessitating AI-enabled platforms that automate classification, enforce retention rules, and deliver timely, defensible productions.

Use Cases (Public & Private Sectors)

Financial, Legal, & Criminal Investigations

Faster triage of devices and cloud sources; entity/communication mapping; timeline reconstruction; link to evidentiary standards. [11]

FOIA

Intake, deduplication, auto-classification, PII/PHI redaction, and production tracking against statutory deadlines. [12]

Data Management & Analysis

Authoritative data catalogs, retention/disposition rules, and eDiscovery-ready storage policies synchronized with the electronic-records mandate. [13]

A Suggested eDiscovery Approach for U.S. Federal Government Agencies

We recommend an eDiscovery capability that leverages an AI-powered toolkit with a data standards-aligned approach, which U.S. government agencies can adopt incrementally, using best-in-class commercial eDiscovery platforms built to minimize operational disruption and maximize speed and cost-effectiveness.

Architecture

- Data estate integration: connectors to email, collaboration, endpoints, cloud object stores, databases, line-of-business systems, and archives.
- Ingestion & normalization: hash-based deduplication, container expansion, OCR/ transcription for audio/video, time zone normalization, and metadata harmonization.
- AI/ML services: semantic search, clustering, near-duplicate detection, CAL/TAR, entity extraction, and machine redaction, with quality gates (precision/recall sampling).
- Governance & security: role-based access, legal holds, audit trails, FedRAMP-authorized hosting options, encryption and key management, and defensible retention/disposition.
- Dashboards & chat: operational dashboards (throughput, backlog, SLAs) and chat interfaces for analysts/attorneys to ask natural-language questions over authorized corpora.
- Production & export: load files, Bates stamping, privilege logs, and FRCP-compliant packaging for litigation or FOIA release. [14]

Operating Model

- Policy-to-product mapping: translate FOIA, FRCP, and records directives (M-23-07) into workflows, SLAs, and metrics.
- Data inventory & risk register: catalog systems of record; align retention and legal-hold posture.
- Pilot on high-impact matters: choose a FOIA queue or investigative caseload to validate gains in cycle time and quality.
- Scale with guardrails: implement sampling and validation for AI outputs; codify privilege/PII rules; introduce multi-party governance.
- Continuous improvement: track cost-to-produce, backlog aging, model performance, and error rates; roll lessons learned into playbooks. [15]

This model establishes an evidence-ready, AI-assisted eDiscovery operating model that reduces reviewer hours, shortens FOIA and production timelines, and measurably improves defensibility.

Implementation Considerations

1

Records & Compliance

- Electronic records: ensure repositories and archives meet M-23-07 goals with metadata completeness and exportability to NARA.
- Discovery proportionality: operationalize Rule 26(b)(1) proportionality in scoping decisions and sampling thresholds.
- Production formats: align to Rule 34(b) expectations for ESI organization and labeling; maintain chain-of-custody. [16]

2

Security & Privacy

- Zero Trust controls: identity-aware access, device health, micro-segmented review environments; immutable audit logs.
- Sensitive data handling: automated detection and machine redaction of PII/PHI; differential access for law enforcement vs. civil discovery.
- Transparency of AI: document training data, model versions, and validation results; preserve explanations for privilege/redaction decisions. [18]

3

Infrastructure & Cost

- Right-sized capacity: use cloud/hyperscale options for burst workloads; leverage nearline storage for cold data; monitor cost per GB reviewed.
- Sustainability: anticipate rising data-center energy needs from AI workloads; factor location and power availability into platform choices. [17]

4

Measuring Success - Program KPIs

- Throughput & timeliness: median days from FOIA request or legal hold to production; backlog age; % on-time responses.
- Quality: precision/recall from sampling; privilege error rate; re-work after QC.
- Cost/Price: subscription-based services, cost per GB processed/reviewed; reviewer hours saved via TAR/CAL; TCO.
- Compliance: electronic records completeness; legal adherence; defensibility of processes (audit findings). [19]

Spending is shifting upstream (collection/processing) and into AI-enabled review, lowering the proportion of total budget consumed by manual review, evidence that automation is working where properly governed. [20]

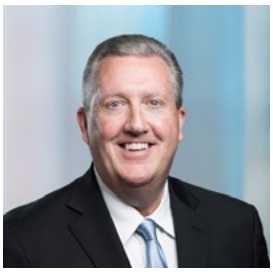
Summary & Conclusion

Information overload is very real and now quite structural in nature (more data, more modalities, more information requests, more Large Language Models (LLMs) required to support the growth of AI tools, and more urgency). U.S. Federal government directives have locked in an electronic-records future, while AI and quantum computing are changing both the scale of the problem and the speed of potential solutions.

We recommend both commercial organizations and U.S. government agencies respond to this data challenge with AI-powered, data standards-aligned eDiscovery capabilities that protects rights, improves timeliness, and demonstrate stewardship of resources, without forklifting existing information systems.

When eDiscovery is done well, it becomes a mission accelerator, turning data into informed and accurate decisions at the speed of business. [21]

About the Authors



Gregory A. Garrett is the Chief Operating Officer of Converged Security Solutions (CSS), leading expert teams within business units Evolver and eVigilant. These groups provide advanced IT, AI, cybersecurity, electronic security, and eDiscovery services to federal agencies and Fortune 500 companies.

With over 25 years of executive experience and a background as a CIO, CTO, CISO, and COO, he has managed more than $40 billion in global tech contracts and served in key roles across both large and mid-sized firms.

A retired U.S. Air Force Colonel and accomplished author, Garrett has also contributed extensively to industry thought leadership through 24 books, more than 150 articles, expert testimony, and over 250 speaking engagements.

Jamie Berry, J.D., is the President of Evolver Legal Services (ELS), where he leads the strategic growth, technical development, and delivery of high-quality legal technology and eDiscovery solutions for clients across various industries and sectors. With 20+ years of robust litigation services experience within internal law firms and Fortune 500 corporate legal departments, he has a wealth of knowledge and expertise in computer forensics, cyber privacy and compliance, and investigative incident response.

Berry has held key leadership roles at Integreon, Thomson Reuters, DLA Piper, and UnitedLex. He is an Adjunct Professor at Catholic University's Law School, where he earned his J.D., a graduate of Marquette University, and a recognized thought leader in the litigation support and eDiscovery communities.

Dr. Brian McElyea is an accomplished cybersecurity executive recognized for leading largescale digital transformation initiatives and building high-performing Cyber Centers of Excellence that drive mission-aligned security outcomes.

With over two decades of leadership across healthcare, defense, and federal sectors, he has guided global organizations through the complexities of risk management, compliance, and modernization with measurable results.

Works Cited

- **[1, 12, 19] Office of Information Policy, U.S. Department of Justice.** 2024 Annual FOIA Report Summary. **April 28, 2025.**
- **[2] Seagate.** Data Age 2025: "The Digitization of the World From Edge to Core." **May 2020.**
- **[3, 13, 15, 16, 21] OMB/NARA.** M-23-07, Update to Transition to Electronic Records. **December 23, 2022.**
- **[4, 17] International Energy Agency.** "Energy and AI—Executive Summary." **2024.**
- **[5, 9, 14] Federal Rules of Civil Procedure.** Rule 26: Duty to Disclose, General Provisions Governing Discovery. **December 2024.**
- **[6] Synergy Research Group.** "Hyperscale Data Center Count Hits 1,136; US Accounts for 54% of Total Capacity." **19 Mar. 2025.**
- **[7] McKinsey & Company.** "The Cost of Compute: A $7 Trillion Race to Scale Data Centers." **28 Apr. 2025.**
- **[8] Grand View Research.** U.S. eDiscovery Market Size & Share. **2024–2030.**
- **[10] Federal Rules of Civil Procedure.** Rule 34: Producing Documents, Electronically Stored Information, and Tangible Things, or Entering onto Land, for Inspection and Other Purposes. **December 2024.**
- **[11] Major Cities Chiefs Association.** Digital Evidence: Challenges for Large Police Agencies. **Oct. 2024.**
- **[18] The Sedona Conference.** The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production. **2018.**
- **[20] HaystackID.** 2025 eDiscovery Review Update: Tasks, Cost Data, and Spending Patterns. **June 25, 2025.**

www.evolverinc.com