

CASE STUDY

HaystackID® Core Intelligence AI Case Insight™  
in Practice: Documented GenAI Performance  
Across High-Stakes Legal Engagements  
Eight Matters. Data-Driven Volume Reductions.  
Defensible Results.

HAYSTACK®



# Introduction

Legal teams are contending with document populations that grow faster than review capacity, and the consequences of moving too slowly or spending too much is rarely abstract.

**Missed deadlines, runaway costs, and inconsistent productions carry real risk** in litigation, regulatory response, and government investigations.

Case Insight™, part of HaystackID's Core Intelligence AI™ suite, applies generative AI (GenAI) to classify, prioritize, and analyze document collections at scale. **The platform is not a replacement for legal judgment — it is built to extend it**, combining AI classification with structured workflows, expert oversight, and disciplined validation to surface what matters earlier and reduce the volume that requires human review.

Across these engagements, **Case Insight's classification was paired with Case Elements capabilities** that surfaced key documents, key people, and key events from within the data. These outputs provided structured fact development, relationship mapping, and event timelines early in each engagement.

The following eight engagements span litigation, regulatory response, government investigations, and internal matters. Across all the matters, the results follow a consistent pattern: **substantially reduced review populations, faster fact development, and classification accuracy that holds up to scrutiny.**

The methodology is designed to integrate with existing review strategies and analytics workflows — not displace them. **What changes is the speed and confidence** with which teams can act on what the data contains.

# Table of Contents

---

<b>Early Case Insight</b> .....	4
Hybrid TAR 2.0 Workflow	
<b>Case Insight Fact Finding</b> .....	6
Strategic Document Identification in Securities Fraud Litigation	
<b>Case Insight Received Production Analysis</b> .....	7
DOJ Criminal Investigation	
<b>Case Insight Quality Control</b> .....	8
Pre- and Post-Review Validation Across Regulatory Matters	
<b>Case Insight Internal Investigation</b> .....	9
48-Hour Fact Development for Regulatory Response	
<b>Case Insight Document Population Reduction</b> .....	10
92% Culling with Validated AI Analysis	
<b>Case Insight Benchmarking</b> .....	11
Validated Performance Against Keyword Search and Manual Review	
<b>Case Insight Enabled DSAR Response</b> .....	13
Employee Data Subject Access Request	
<b>Case Insight Witness Preparation</b> .....	15
Grand Jury Investigation Fact Development	

# Early Case Insight

## Hybrid TAR 2.0 Workflow

### Strategic Discovery Optimization Across 32.9M Documents

#### CHALLENGE

A major corporation engaged HaystackID to support a **significant commercial dispute**. The litigation centered on the interpretation and performance of complex agreements governing licensing, support, and payment obligations. The provider alleged contractual breaches tied to licensing scope, support, and fees, while the client disputed those claims and challenged the asserted entitlement and damages theories. The matter required a discovery approach capable of handling **exceptionally large data volumes** while quickly surfacing high-value communications and materials relevant to contract negotiation, performance, and dispute escalation.

#### SOLUTION

HaystackID led discovery analytics and review optimization across a very large and mixed data population, implementing a strategic **multi-phase workflow**:



- **Phase 1 - Data Processing and Preparation:** The team processed and loaded 5,863,406 client emails and Slack documents for potential review, drawn from a broader corpus of 32,946,517 documents. After processing, 5,344,037 documents were eligible for Case Insight analysis.
- **Phase 2 - Case Insight GenAI Classification:** HaystackID's analytics team consulted with counsel and designed a defensible workflow combining Case Insight and TAR 2.0 to maximize early reduction and prioritize high-value content. Documents classified as Key, Likely Responsive, or Potentially Responsive, were promoted to the review environment.
- **Phase 3 - TAR 2.0 Active Learning:** An active learning project containing 733,545 documents was launched. Reviewers coded 45,762 documents while the model trained and stabilized, after which an elusion test confirmed that only 6.2% of the active learning population required human review.
- **Phase 4 - Validation:** The team developed and executed validation procedures measuring efficacy on both the promoted population and a null set drawn from the remaining corpus, confirming accuracy and defensibility.

## RESULTS

86.7%  
reduction

6.2%  
reviewed

45,762  
documents

Reduced  
review spend

- **86.7% reduction:** 5.8M documents narrowed to 779,265 for review prior to TAR 2.0.
- **Only 6.2%** of the active learning population required human review.
- **45,762 documents** coded to train and stabilize the model.
- **Substantial attorney review hours avoided**, reducing outside counsel spend and accelerating time to insight.

By converting millions of potentially reviewable items into a tightly controlled review set, HaystackID enabled counsel to reduce costs, accelerate insight, and maintain a clear, auditable basis for discovery decisions.

# Case Insight Fact Finding

## Strategic Document Identification in Securities Fraud Litigation

### AI-Driven Document Prioritization in Complex Multi-Entity Financial Fraud

#### CHALLENGE

A private equity fund engaged HaystackID to support securities fraud litigation involving breach of fiduciary duty, fraudulent transactions, and misappropriation of fund assets. The matter involved **complex financial transactions across multiple entities**, including unwinding agreements, asset contributions, and secured loans that allegedly concealed conflicts of interest. With approximately **156,000 documents** in the workspace, the legal team needed to quickly identify communications and records related to transaction negotiations, undisclosed ownership interests, and asset transfers across multiple entities over a multi-year period.

#### SOLUTION

HaystackID deployed Case Insight to analyze **93,000 eligible documents** using the **three-tier classification system**: Key Responsive, Likely Responsive, and Potentially Responsive. The AI classification was combined with targeted search methodologies to ensure comprehensive coverage of fraud-related documents across all relevant entities and transaction types.

#### RESULTS

**95%**  
reduction

**5,600**  
documents

**Three-tier**  
classification

**Defensible**  
approach

- **95% reduction:** 156,000 documents narrowed to 5,600 for review.
- **Rapid identification** of fraud-related communications and transaction records.
- **Three-tier classification** enabled strategic review prioritization.
- **Comprehensive, defensible approach** combining AI classification with targeted search.

Case Insight enabled efficient document identification in a complex multi-entity securities fraud matter, dramatically reducing review volume while maintaining defensibility.

# Case Insight Received Production Analysis DOJ Criminal Investigation

## Rapid Key Document Identification from 1.29M Document Production

### CHALLENGE

A corporation faced a Department of Justice (DOJ) criminal investigation involving allegations of fraud and bribery related to government contracts. The client needed **rapid identification of key documents** from a received production in order to support defense strategy development.

### SOLUTION

Case Insight was applied to **1,290,290 documents**, using the **criminal indictment as the Case Background input**. This approach enabled the AI to classify documents based on relevance to the specific allegations, prioritizing communications and records most likely to inform defense preparation.

### RESULTS

**1.29M**  
documents

**<9%**  
reviewed

**Reduced**  
review burden

**Targeted**  
identification

- **1.29M documents analyzed** using criminal indictment as classification input.
- **Focused review on <9%** of total document population.
- **Dramatic reduction** in time, cost, and effort for the received production review.
- **Rapid defense strategy support** through targeted key document identification.

Case Insight enabled rapid, targeted analysis of a massive amount of received production, allowing the defense team to focus resources on the most critical documents in a high-stakes criminal matter.

# Case Insight Quality Control

## Pre- and Post-Review Validation Across Regulatory Matters

### AI-Driven QC for Consistent, Defensible Review Outcomes

#### CHALLENGE

Multiple regulatory matters involving outside counsel and internal review teams required **rigorous quality control** to ensure consistency, accuracy, and defensibility across large document populations.

#### SOLUTION

Case Insight was deployed for both pre- and post-review quality control workflows across multiple matters:

- **Validated Human Review:** Counsel used Case Insight to QC previously reviewed data, surfacing key documents and streamlining workflows for more consistent, defensible outcomes.
- **Improved Accuracy:** Pre-review deployment helped QC teams quickly flag missed responsive documents and relevance conflicts, raising review quality and lowering cleanup costs.

#### RESULTS

**Key documents**

**Conflicts resolved**

**Improved review quality**

**Consistent outcomes**

- **Key documents surfaced** that traditional review had missed.
- **Relevance conflicts identified** and resolved before production.
- **Review quality raised** while reducing cleanup costs.
- **Consistent, defensible outcomes** across multiple regulatory matters.

Case Insight’s QC capabilities demonstrate that GenAI adds value not only in initial classification but also in validating and improving the quality of human review at scale.

# Case Insight Internal Investigation

## 48-Hour Fact Development for Regulatory Response

### Enterprise-Scale Investigation Across 2.2M+ Documents

#### CHALLENGE

A financial services company received a regulatory inquiry requiring **immediate investigation of potential compliance violations**. In-house counsel needed to quickly identify all relevant facts across **2.2+ million documents**. The full scope of issues was unknown upfront, requiring a flexible investigative approach to surface unknown issues while responding to specific regulatory questions. Traditional document-by-document review would take months, but regulators expected rapid responses.

#### SOLUTION

Case Insight analyzed the full 2.2+ million document population **without predetermined search terms or extensive manual review**. The AI system rapidly surfaced key communications, decision points, compliance discussions, and trading activities relevant to the regulatory inquiry. Case Insight identified patterns and timelines that might not have been discovered through traditional keyword searching, providing a comprehensive fact picture even as the investigation scope evolved.

#### RESULTS

<b>48 hrs</b> fact discovery	<b>2.2M+</b> documents	<b>Swift</b> response	<b>Early</b> remediation
---------------------------------	---------------------------	--------------------------	-----------------------------

- **All critical investigation facts identified** within 48 hours.
- **Comprehensive fact development** without full-scale manual review of 2.2+ million documents.
- **Swift legal and business response** enabling regulatory cooperation.
- **Early identification** of remediation needs and compliance gaps.

Case Insight enabled rapid, comprehensive internal investigation at enterprise scale, supporting timely regulatory response and informed decision-making in a high-pressure compliance matter.

# Case Insight Document Population Reduction

92% Culling with Validated AI Analysis

AI-Driven Early Case Assessment for Targeted Review

## CHALLENGE

A legal team faced a large-scale discovery effort involving **277,000 documents** across a complex matter. The volume made traditional linear review impractical within the required timeframe and budget. The team needed to rapidly and defensibly reduce the document population to focus attorney resources on the most relevant materials, without risking the exclusion of key documents.

## SOLUTION

HaystackID deployed Case Insight to classify the full 277,000-document population, applying **validated GenAI analysis** to assess relevance at scale. The AI system evaluated each document against the case parameters and categorized documents by relevance tier, enabling the team to confidently exclude non-relevant materials and direct review efforts toward the highest-value subset.

## RESULTS

**92%**  
not relevant

**277K**  
documents

**Reduced**  
review set

**Defensible**  
methodology

- **Over 92% identified as not relevant** through validated AI analysis.
- **Review population reduced** from 277,000 to a focused, manageable subset.
- **Substantial time and cost savings** by eliminating non-relevant documents before review.
- **Defensible methodology** with validated classification supporting discovery obligations.

Case Insight’s early case assessment capabilities demonstrated that GenAI-driven document population reduction can deliver dramatic efficiency gains while maintaining the rigor and defensibility required for complex litigation.

# Case Insight Benchmarking

Validated Performance Against Keyword Search and Manual Review  
Independent Evaluation Across 131,000 Documents

## CHALLENGE

A law firm evaluated whether GenAI could **match or exceed traditional keyword-based approaches** for a complex legal matter involving 131,000 documents. The original review faced several challenges: a broad collection, a shortened timeline, technical subject matter requiring SME oversight, complex privilege relationships, and multiple regulatory filings. The firm needed to determine whether AI-powered classification could match accuracy while reducing iterative keyword refinement.

## SOLUTION

HaystackID deployed Case Insight against a **10,000-document subset** that had been fully coded by attorneys through rigorous managed review, creating a reliable ground truth baseline. **The evaluation proceeded in two phases:**

### Phase 1

#### Case Insight vs. Keyword Search

Case Insight classifications were compared against three iterations of keyword searches (initial broad draft, refined second draft, and final optimized version). Case Insight achieved results with minimal configuration, unlike the extended attorney effort required to refine keyword lists across multiple rounds.

### Phase 2

#### Supporting Manual Review

Case Insight was applied across the full 131,000-document population and compared directly to attorney coding. Discrepancies typically involved benign inclusions such as cover emails, which attorneys easily triaged. Case Insight served as a quality control layer, highlighting inconsistencies and surfacing edge cases for re-review while providing comprehensive matter intelligence before review even began.

Beyond classification accuracy, Case Insight delivered contextual intelligence that transformed the review workflow. It clustered communications into themes, built timelines, and mapped key actors—giving attorneys earlier access to case narratives and **shifting focus from document identification to strategy.**

## RESULTS

89%  
F-score

2-4 pts  
variance

131K  
documents

Single-pass  
analysis

- **88% to 93% recall and 85% to 89% precision**, with an F-score near 89%.
- **Performed within 2 to 4 points** of fully optimized keyword searches, with minimal configuration.
- **Validated attorney coding** across 131,000 documents, flagging inconsistencies and improving defensibility.
- **Dramatically reduced time investment:** comprehensive results in a single pass vs. weeks of keyword iteration.
- **Strategic intelligence delivered upstream**, enabling proactive case strategy rather than reactive document discovery.

The independent evaluation confirmed that Case Insight consistently matched or exceeded keyword-based approaches and strengthened manual review. By combining high recall and precision with contextual narrative generation, it set a new benchmark for efficiency, accuracy, and strategic value in legal data analysis.

# Case Insight Enabled DSAR Response

## Employee Data Subject Access Request

### Precision AI Filtering and Automated Redaction for Contentious Employment Matter

#### CHALLENGE

A multinational organization received a Data Subject Access Request (DSAR) from a former employee in anticipation of an employment tribunal matter. The request encompassed **32,407 documents** spanning email, messaging platforms, and shared file repositories. The matter was steeped in workplace controversy, and the volume, review complexity, and **extensive redaction requirements** exceeded internal team capacity for traditional DSAR response protocols within the regulatory timeline.

#### SOLUTION

HaystackID deployed its AI-enabled DSAR Response Service, integrating Case Insight's precision filtering with automated redaction technology to streamline the end-to-end response workflow:

- **AI-Powered Scope Reduction:** Case Insight classified the full 32,407-document population, identifying documents relevant to the data subject's request and removing 30,480 documents (94%) from review scope through validated AI analysis.
- **Targeted Review:** The narrowed review set of 1,927 documents (6% of the original collection) was dominated by high-value content, with Key documents constituting 79% of the final review population, demonstrating precision in scope determination.
- **Automated Redaction:** AI-driven redaction protocols protected third-party personal information across disclosed documents, dramatically reducing the manual effort typically required for DSAR redaction compliance.

## RESULTS

94%  
reduction

1,927  
documents

1,107  
disclosed

79%  
key documents

- **94% volume reduction:** 30,480 documents removed from review scope through AI filtering.
- **1,927 documents reviewed** (6% of the original 32,407-document collection).
- **1,107 documents disclosed**, with automated redactions applied to 740 (67%).
- **79% of final review set classified as key documents**, demonstrating targeting precision.
- **Defensible compliance delivered** within regulatory timeframe at competitive pricing.

HaystackID transformed a contentious, high-volume employee DSAR into a manageable, defensible response, relieving internal teams of complex data challenges and mitigating compliance risk through AI-powered precision filtering and automated redaction.

# Case Insight Witness Preparation

## Grand Jury Investigation Fact Development

### AI-Driven Fact Development and Document Analysis for 13 Witnesses in Federal Grand Jury Investigation

#### CHALLENGE

A technology company faced a **federal Grand Jury investigation** into its content and transactions. Thirteen employees received Grand Jury subpoenas compelling testimony and personal records. Outside counsel needed to rapidly develop individualized witness preparation materials for **all 13 witnesses**, each of whom had distinct roles spanning multiple enforcement areas. The document universe relevant to these witnesses was vast, spanning internal investigations, enforcement actions, escalation records, policy development materials, and cross-functional program documentation. Traditional witness preparation would have required weeks of attorney-led document review across each witness's individual record set.

#### SOLUTION

HaystackID deployed Case Insight to support a multi-witness preparation workflow spanning all 13 subpoenaed employees. **The engagement proceeded in coordinated phases:**

- **Case Background Development:** The team worked with outside counsel to develop detailed Case Background inputs for each witness, incorporating the subpoena scope, known investigation focus areas, key document categories of interest, and each witness's specific role and enforcement history.
- **Witness-Specific Document Classification:** Case Insight analyzed the document populations associated with each witness, classifying materials by relevance to the specific allegations and the witness's area of responsibility. The system identified key communications, internal investigation records, enforcement decisions, escalation chains, and policy discussions relevant to each witness's anticipated testimony.
- **Fact Development and Narrative Generation:** Case Insight generated witness-specific fact chronologies, mapped relationships between witnesses and overlapping investigations, and surfaced cross-cutting themes across the 13 witnesses, including patterns in enforcement gaps, resourcing constraints, revenue-driven policy decisions, and coordination failures between integrity teams.
- **Preparation Material Delivery:** Individualized witness preparation packages were delivered to outside counsel, each containing classified and prioritized document sets, key document summaries, fact timelines, and identified areas of potential exposure, enabling focused and efficient preparation sessions.

## RESULTS

**13**  
witnesses

**Faster**  
preparation

**Shared**  
themes

**Exposure**  
areas

- **13 witnesses prepared** with individualized, AI-driven fact development packages.
- **Rapid, parallel witness preparation** replacing weeks of sequential attorney-led document review.
- **Cross-witness themes identified** across enforcement gaps, resourcing constraints, and coordination failures.
- **Key exposure areas surfaced**, including revenue-tied content, AML control gaps, and enforcement resourcing deficiencies.
- **Comprehensive fact chronologies delivered** enabling focused preparation sessions aligned to anticipated lines of questioning.

Case Insight transformed a high-stakes, multi-witness Grand Jury preparation effort from a months-long document review exercise into a structured, AI-driven fact development operation. By delivering individualized preparation materials for all 13 witnesses in parallel, HaystackID enabled outside counsel to focus preparation time on strategy and testimony, not document identification.

# Conclusion

By applying Case Insight within structured, validated workflows, teams consistently **reduced review volumes, developed facts faster, and produced results that held up to scrutiny** across matter types, data sizes, and legal contexts. In several engagements, volume reductions exceeded 90%. In others, fact development that would have taken weeks was completed in days. The performance was not incidental—it was the product of disciplined methodology, not just the technology.

In each engagement, Case Insight operated alongside established review strategies, legal expertise, and quality control processes. It was applied to received productions, internal investigations, regulatory responses, and witness preparation, each with different requirements, timelines, and defensibility standards. **Across all of them, the workflow held.** The difference was in what teams could see earlier, and how much less data they had to work through to get there.

The demands driving these engagements—volume, speed, cost pressure, defensibility—are not temporary conditions. They are the baseline against which every matter is now measured. Organizations that treat AI as an add-on, applied after traditional methods have already shaped the review, will capture only a fraction of the available benefit. The results here reflect **what becomes possible when AI is integrated from the outset, validated against ground truth, and guided by practitioners** who understand both the technology and the legal obligations it serves.



# Learn More Today.

Discover how HaystackID helps legal teams apply GenAI to reduce document volumes, accelerate fact development, and maintain defensible workflows. [Connect with us](#) to explore solutions that improve speed, clarity, and control across complex matters.

---

## **About HaystackID®**

[HaystackID](#) solves complex data challenges related to legal, compliance, regulatory, and cyber requirements. Core offerings include Global Advisory, Cybersecurity, Core Intelligence AI™, and ReviewRight® Global Managed Review, supported by its unified CoreFlex™ service interface and eDiscovery AI™ technology. Recognized globally by industry leaders, including Chambers, Gartner, IDC, and Legaltech News, HaystackID helps corporations and legal practices manage data gravity, where information demands action, and workflow gravity, where critical requirements demand coordinated expertise, delivering innovative solutions with a continual focus on security, privacy, and integrity. Learn more at [HaystackID.com](#).

Assisted by GAI and LLM technologies.