

# Leveraging AI Analysis in Discovery: Why Structured Data Is Your Friend

Attorneys increasingly use specialized AI tools such as Westlaw Co-Counsel and RelAIr to securely analyze large bodies of confidential documents. These tools excel at summarizing and searching large documents or document collections, providing answers about complex data. In discovery, where practitioners routinely grapple with massive volumes of unstructured data—emails, contracts, depositions, expert reports, and financial records—this capability proves invaluable. Yet without the right workflows, AI analysis can feel like a glorified word search or summary generator: useful to a point, but not quite game-changing.

The key to unlocking AI's full potential lies in understanding the power of structured data. When we leverage AI to not only search and summarize, but also organize data—turning unstructured into structured—we unlock opportunities for extensive and efficient analysis.

## STRUCTURED VS. UNSTRUCTURED DATA

Structured data is information organized into a predefined, consistent format—typically rows and columns in spreadsheets or database tables—where each field has a specific data type and location. Unstructured data is freeform information without standardized organization—Word documents, emails, PDFs, scanned images—where relevant information can appear anywhere within the content.

Consider a large contract: The 50-page contract itself is unstructured data. A spreadsheet extracting that contract's key terms into labeled columns is structured data. The fundamental difference is accessibility. Structured data is instantly searchable, filterable, and analyzable because every data point has a predictable location and format.

## TO UNDERSTAND THE VALUE OF STRUCTURED DATA, THINK OF RELATIVITY

Practitioners who use Relativity or similar document-review platforms understand its fundamental architecture. The platform doesn't simply display raw documents for review, and Boolean searches aren't the only way to find what you need. Instead, Relativity extracts and standardizes metadata: sender, recipient, date, file type, file name, custodian. This metadata framework is structured data—the document list view is essentially a spreadsheet—enabling powerful filtering, searching, and analysis.

# Leveraging AI Analysis in Discovery: Why Structured Data Is Your Friend

By leveraging metadata filters and stacking them, you can quickly identify all emails between two specific individuals during a specific time period. Within that narrowed set, you can apply Boolean text searches to find exactly what you need. This metadata-driven approach transforms chaotic document collections into manageable, query-able datasets.

The same principle should guide your AI analysis strategy for large unstructured datasets in discovery. Analysis—including AI-assisted analysis—is generally more efficient when fed or guided by structured data rather than raw, unstructured documents.

But here's where it gets powerful: Unlike Relativity's predetermined metadata fields, you can create custom spreadsheets populated with whatever fields best suit your needs. Consider thousands of pages of product tests or medical records. Prompt your language model to create a table with fields such as: Test\_Number | Test\_Type | Test\_Date | Source\_File | Page\_Number | Test\_Results | Product\_or\_Patient | Test\_Technician. After iterating prompts and manual formatting until the table reaches your desired state, you've unlocked countless ways to engage in organized, precise analysis:

- Upload the spreadsheet back into the AI tool and reference it in prompts for precise guidance: "What is the total cost of Test Type A from January to March 2024?"
- Create similar spreadsheets from different document sets you want to compare—such as medical or financial records produced by different, similarly situated parties.

Leveraging structured data offers several critical advantages. Referring the language model back to your created table, rather than the original document collection, is far more efficient and produces more consistent results. Structured data dramatically reduces computational overhead, eliminates formatting inconsistencies and OCR errors, and enables repeatable, iterative analysis—essential when discovery strategies inevitably evolve.

## **DISCOVERY RESPONSES: ONE USE-CASE FOR OUR UNSTRUCTURED-TO-STRUCTURED STRATEGY**

Consider managing discovery responses in complex multi-party litigation where you receive hundreds or thousands of responses from dozens of defendants, each using different formats, objection language, and response styles. A two-stage approach transforms this chaos into actionable intelligence.

# Leveraging AI Analysis in Discovery: Why Structured Data Is Your Friend

Stage One converts chaos into clarity. Extract each party's responses into structured format: Party\_Name | Request\_Number | Request\_Text | Response\_Text | Objections\_Asserted | Documents\_Produced (Y/N) | Production\_Bates\_Range | Responding\_Attorney | Response\_Date. What were previously hundreds of separate PDF documents in varying formats becomes a searchable, sortable master database.

*Note: Language model tools such as Westlaw Co-Counsel alone may not suffice to create this spreadsheet, depending on document format or volume. Practitioners should consider specialized tools, including creating custom extraction tools using AI-assisted coding. We will address this workflow in a separate article.*

Stage Two unleashes AI's analytical power on this structured foundation. Now you can easily identify which parties failed to respond to—or supplement—specific requests, compare objection patterns across defendants, flag suspiciously similar response language suggesting coordination, and identify gaps where certain parties claim no responsive documents exist while others produced volumes.

Ask the AI to analyze your structured data: "Which parties objected to Request 15 on privacy grounds?" or "Show me all responses where parties claimed no responsive documents exist, sorted by request topic." The AI processes your structured table in seconds—analysis that would be more difficult, less consistent, and more time-consuming if constantly returning to raw response documents.

The structured approach also enables iterative refinement. Suppose opposing counsel supplements responses. In that case, simply have your AI tool add new rows to the existing table and re-run your analysis. When you identify a new pattern worth investigating, filter the structured data and ask AI targeted questions about that subset.

## **CONCLUSION: OPTIMIZING AI WORKFLOWS WITH STRUCTURED DATA**

The future of AI in discovery lies not in replacing legal staff but in amplifying their productivity through intelligent data organization. Just as Relativity would be far less useful without leveraging metadata, AI tools perform optimally when provided structured inputs.

# Leveraging AI Analysis in Discovery: Why Structured Data Is Your Friend

Crucially, a related advantage of AI tools is that we can use them to create these structured inputs. Employ the two-step approach: Use AI to create structured data first, then unleash AI's analytical power on that organized foundation.

In future articles we will dive deeper into examples and details of these workflows.

## AUTHOR

Andrew M. Mast is a BSP Law Partner focusing his practice on product liability and complex discovery matters. He has a proven record of innovation in e-discovery, as well as in litigating the biggest automotive product liability and class action matters.

