Evaluating the Opportunities Presented by Analytics and Visualization in Legal Review

What You Need To Know

As volumes of electronically stored information (ESI) continue to expand, they contribute to the difficulty and cost of the legal review required to properly analyze and resolve matters. As this expansion has created the greatest pains in the collection, production, and document review stages of discovery, electronic discovery (eDiscovery) solution vendors have traditionally focused on solutions that automate and accelerate these aspects of review. Data analytics and visualization capabilities stand out in this landscape for the support they provide in facilitating legal analysis and comprehension beyond the acceleration of “heavy lifting” aspects of review.

Drawing from Blue Hill analysis twenty-one research interviews with corporate legal departments, law firms, and litigation support providers (LSPs), this report reviews these opportunities and provides key points to consider in evaluating these solutions.

Opportunities Presented by Visual Analytics

Visualization and analytics refers to a combination of data processing and representation technologies. Applied to legal review, visual analytics solutions assist with the organization and human comprehension of the document and data sets required by matters. Common capabilities offered by visual analytics solutions include:

- Multifactor analysis of document and custodian relationships
- Multifactor analysis of key terms
- Document clustering by similar concepts
- Trend and relationship visualization
- Predictive coding

The capabilities described above have direct applications to collection, culling, and document review. However, the deeper value of these solutions lies in the review and analysis of matters. Traditional,
linear legal review proceeds “blind,” with each document reviewed in isolation from its larger context until the completion of document review. By contrast, the use of analytics permits organizations to organize legal review with an understanding of document relationships and high-priority document clusters. Combining these capabilities with trend analysis and graphic representations of data facilitates rapid comprehension of a matter assessment and legal strategy, beyond traditional contributions to execution of document review.

To understand how, it is helpful to begin with an appreciation of human cognition. Our minds are extremely well equipped for pattern recognition, but comparatively poor at computational assessments. Visual representation of data supports the review of large sets of data by removing quantification aspects of information analysis, providing a shortcut to pattern recognition and identification of trends.

In the context of legal review, the combination of automated data analysis and trends visualization provides its deepest value by organizing documents for review and accelerating comprehension of matter information. In particular, by optimizing data for human analysis, these tools can help attorneys to identify key trends and bridge connections that would not readily appear in linear review. As such, the value proposition for visual analytics expands beyond core eDiscovery investments, which tend to relate to the acceleration of linear information processing in traditional review processes.

Because traditional eDiscovery capabilities primarily relate to the most mechanical and repetitive aspects of review, eDiscovery investments have typically faced volume or frequency thresholds to justify the expense.

By contrast, analytics and visualization capabilities support complex reasoning and strategic planning aspects of legal review.

This is because the eDiscovery solution market largely evolved in response to the pain points created as electronically stored information (ESI) expanded beyond what could be feasibly performed through manual processes. While these capabilities provide considerable value in accelerating collection, processing, review, and production of large data sets, their value primarily relates to the most mechanical and repetitive aspects of review. Because of this, eDiscovery investments have typically faced volume or frequency thresholds to justify the expense.

By contrast, analytics and visualization capabilities support the complex reasoning aspects of legal review as well as the effectiveness of review. Through the organization and facilitation of analysis, these tools have the potential to speed factual review, identify key case elements and challenges, as well as support decision-making and strategic planning. If traditional eDiscovery investments automate the most basic aspects of review, visual analytics investments further contribute to speed and

Research Participants

This report draws from insights derived from Blue Hill research interviews with twenty-one organizations. Participants were drawn from a cross-section of the eDiscovery market segments and include:

- 5 internal legal departments
- 10 law firms
- 6 litigation support providers
efficiency by capturing and clustering documents for review in context. This limits the haphazard and isolated nature of information discovery in linear review, removing wasted and redundant effort and redundancy. The resulting value derives from both more efficient handling of large volumes of information as well as improved attorney insight and, ultimately, the quality of legal advice and advocacy.

In addition to strict legal review use cases, the ability to rapidly sort, organize, and understand documents has applications in numerous fact-finding applications. This includes factual research involving large amounts of documents or internal investigation use cases as well. For example, in compliance or internal investigations involving large volumes of documents or communications, analytics tools can assist stakeholders to quickly group similar concepts and identify relevant facts. Ultimately, these applications help to reduce the time required for investigations and contribute to earlier resolution of issues. Figure 1 summarizes the essential value proposition of both analytics and tools supporting linear, binary review as each relates to the core legal review needs.

**Figure 1: Primary eDiscovery Investment Drivers and Contributions**

<table>
<thead>
<tr>
<th>Core eDiscovery</th>
<th>Visual Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Governance</td>
<td>Early case assessment</td>
</tr>
<tr>
<td>Collection and Forensics</td>
<td>Test and metadata analysis</td>
</tr>
<tr>
<td>Processing and Culling</td>
<td>Trend analysis</td>
</tr>
<tr>
<td>Review</td>
<td>Document clustering</td>
</tr>
<tr>
<td>Predictive Coding</td>
<td>Dashboards and visualization</td>
</tr>
</tbody>
</table>

Use Cases to Consider
- Compliance Review
- Internal Investigation
- Early case assessment
- Discovery preparation and production
- Custodian management
- Prioritization of review
- Factual review of matter
- Deposition preparation
- Litigation strategy

Source: Blue Hill Research, November 2015

**Case-in-Point: Dorsey & Whitney LegalMine™ and Ringtail Software**

In response to growing client concerns over both the cost of legal review and the quality of third-party review services, several years ago Dorsey & Whitney, an AmLaw 100 firm, launched a managed services practice devoted to legal review support. Through the practice, the firm currently manages over 600 client databases on Ringtail supporting a range of legal review needs, including: litigation discovery, transactional due diligence, corporate investigations, and compliance review.
In addition to scaling to handle over 600 databases on behalf of clients, Ringtail’s visual analytics capabilities enable the firm to improve both the efficiency and effectiveness of review. The firm identified three core use cases across the lifecycle of review projects:

- **Key Term Identification** – In initial stages, the firm uses Ringtail’s analytics to investigate the data set and validate search terms to use in culling documents for review. This process typically begins with attorney review of a random sample of documents to identify key terms within relevant documents. The firm uses document clustering and text analytics to understand trends and frequency in how terms occur across the document set, to identify additional words that tend to appear in connection with these terms, and to evaluate the effectiveness of specific terms. In this way, the firm reports that it uses “the documents to tell us what the key words should be.” In addition, relationships and connections between documents are used to identify potentially relevant documents that do not contain key terms. As a result, the firm obtains more effective sets of terms and flagged documents and is able to eliminate non-relevant materials from its document sets. These sample review sets can also be used in support of a proportionality argument regarding particular proposed search terms or for seeding in a predictive coding scenario. Ringtail’s analytics tools at this stage are also used to conduct a gap analysis on data received.

- **Review Prioritization and Acceleration** – Ringtail’s document mapping capabilities permit the firm to cluster documents according to timelines, narrative relationships, conceptual similarity or occurrence of key terms. In this way, the firm replaces linear review, organized by custodian, with topic-based review, organized according to potential relevance. The firm prioritizes clusters that appear to have the highest likelihood of offering information that is meaningful to the underlying matter. The firm performs this analysis both at the start of the review process as well as continuously throughout, based on ongoing feedback of reviewers and attorneys, to ensure that it always keeps the highest priority documents at the head of the review queue. This process helps the firm uncover relevant information early and accelerates review as a whole. In the case of a review of a received production, once the relevancy of clusters reaches a point of diminishing return, the legal team may opt to discontinue further review, again saving the client on review costs.

- **Quality Control** – Ringtail’s analytics capabilities assist in quality control checks of reviewed documents prior to production. The firm reports that Ringtail’s visualization capabilities play a crucial role in helping QC reviewers to quickly “look for, focus on, and target” potential issues in document sets. For example, visual clustering documents with responsive coding with those documents confirmed as privilege permits the firm to quickly identify clusters with co-mingled responsive and privilege documents, highlighting groups of documents which warrant quality control review to ensure the responsive documents are, indeed, eligible for production. These uses of visualization reduce the time required by quality control checks as well as the risks of mistaken or inadvertent disclosure of confidential or protected information.
Factors to Consideration in Solution Evaluation

Organizations considering investments in analytics and visualization capabilities within legal review will wish to begin by considering the potential business case. This requires an understanding of the potential upside and scope of value available. In part, this requires an assessment of the impact of usual pain points considered in eDiscovery solution investments, such as the cost and speed of reviewing large volumes of documents. However, if traditional eDiscovery investments depend on these factors, analytics and visualization capabilities also help to improve the effectiveness of attorney review. This expansion in value proposition and potential use cases also expands the number scenarios in which visual analytics investments will be appropriate.

Complex matters with large volumes of data create clear use cases for visualization and analytics, where these solutions can help to organize documents and supply context for attorney review. While this impacts the efficiency of attorney review processes directly, the net benefits are similar to traditional eDiscovery benefits: the time and costs imposed by large volumes of ESI. Because analytics contributes to the speed and effectiveness of review, rather than simply automating processes, these solutions have applications in a wide range of types and sizes of matters, rather than simply larger, complex cases.

In assessing business cases, organizations will also need to consider ways to measure how improvements in attorney effectiveness offer value to the organization. For law departments, the impact may relate to (1) improved understanding of risk and liability, (2) the cost and time required to resolve matters, or (3) costs related to collection, data processing, and external legal services resulting from more intelligent internal review and production. For law firms, the impact on the quality of legal service should be measured in terms of client satisfaction, effectiveness in advocacy, as well as firm differentiation, client retention, and increases in revenue. Not all factors can be measured as directly as contributions to the cost or time required to review materials. Nonetheless, they will justify the use of analytics and
visualization to support review in scenarios where data volume or complexity would be insufficient to justify the use of traditional eDiscovery tools.

Once an organization possesses an estimate of the potential value presented, it must balance that investment against the potential costs of the solution and net contribution available. This is rarely a comparison of cost alone. Rather, Blue Hill’s research frequently highlights how this evaluation involves a range of factors related to the functionality, ease of use, and cost of ownership related to the solution.

Blue Hill’s eDiscovery Solution Evaluation Framework report identified ten factors commonly used in eDiscovery solution evaluations reported by law department, law firm, and LSP buyers. Due to the prioritization placed on the speed and efficiency of collection, data processing, and review in traditional eDiscovery, not all of these factors have direct application to the assessment of analytics and visualization. Table 1 summarizes a select set of these capabilities that are relevant to these investments.

**Table 1: Analytics & Visualization Solution Evaluation Factors**

<table>
<thead>
<tr>
<th>Factors to Consider</th>
<th>Impact on Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fit to Needs</strong></td>
<td></td>
</tr>
<tr>
<td>● Number of capabilities matching investment needs out of the box</td>
<td>● Impact on underlying investment driver</td>
</tr>
<tr>
<td>● Fit of capabilities to use and business case</td>
<td>● Opportunity for consolidation of investment costs</td>
</tr>
<tr>
<td>● Impact of capabilities on underlying need</td>
<td>● Impact on quality of insight supporting matter resolution</td>
</tr>
<tr>
<td>● Amount of customization required</td>
<td></td>
</tr>
<tr>
<td><strong>Ease of Implementation</strong></td>
<td>● Time required to set up solution to particular project</td>
</tr>
<tr>
<td>● Difficulty of implementation of the solution</td>
<td>● Specialization required of stakeholders to set up solution</td>
</tr>
<tr>
<td>● Difficulty of setting up projects / processes</td>
<td>● Customization and costs required to set up solution</td>
</tr>
<tr>
<td>● Burdens placed on IT infrastructure and staff</td>
<td></td>
</tr>
<tr>
<td><strong>Ease of Adoption</strong></td>
<td>● Level of specialization and education required for use</td>
</tr>
<tr>
<td>● Ease of data import and management</td>
<td>● Cultural resistance to use of solution</td>
</tr>
<tr>
<td>● Ease of adoption and use by attorney users</td>
<td>● Efficiency of solution users in execution of processes</td>
</tr>
<tr>
<td>● Availability of integrations with other solutions</td>
<td></td>
</tr>
<tr>
<td><strong>Cost</strong></td>
<td>● Software license costs presented by solution</td>
</tr>
<tr>
<td>● License costs of solution</td>
<td>● Number of solutions required to achieve goals</td>
</tr>
<tr>
<td>● Hardware and IT infrastructure costs required</td>
<td>● Data storage and utility costs presented by solution</td>
</tr>
<tr>
<td>● Impact on data storage and processing costs</td>
<td>● Production and data sharing costs presented IT and support staff time required by solution</td>
</tr>
<tr>
<td>● Integration with existing IT ecosystem</td>
<td></td>
</tr>
<tr>
<td>● Difficulty of integration with related solutions</td>
<td></td>
</tr>
<tr>
<td>● Amount of customization required</td>
<td></td>
</tr>
</tbody>
</table>

Source: Blue Hill Research, November 2015

**Key Observation and Takeaways**

As data volumes and the complexity of matters continue to grow, these capabilities will play an increasingly important role in the efficient and the effective management and completion of legal review.
Where traditional eDiscovery investments primarily helped to prepare documents and to accelerate basic aspects of linear review, their ultimate contributions are limited, often requiring volume and complexity thresholds in order to justify the expense.

Analytics and visualization help to prioritize, organize, and maintain document context for intelligent review. This shift from “all in” data production and linear review to selective identification of responsive data and context-sensitive attorney review expand on the value offered by traditional eDiscovery investments. In particular, this shift can contribute significantly to reductions in the data management and data processing costs of litigation as well as the time and costs associated with review. Visual analytics solutions also offer value by enhancing attorney comprehension of matters as well as in a variety of fact-finding and investigatory scenarios. While the number of documents involved certainly plays a role here as well, primary drivers should expand from solely the volume and speed of document review to also include improvements to the quality and effectiveness of legal review and analysis.

Organizations evaluating analytics and visualization solutions will begin by building business cases considering the impact on both the speed and cost of attorney review. In addition, they will wish to consider whether the impact on attorney comprehension and effectiveness offers further business case value outside of these factors. Due to differences in the underlying business pressures, how the organization ways this impact will differ between internal law departments and among law firms:

- **For law departments**, the ultimate impact is found in improvements in its ability to assess exposure, early resolution of matters, as well as intelligent production of data. These benefits contribute to improved resolution of matters, and reduced dispute resolution and litigation cost.

- **Law firms** will find that the ultimate business benefits offered largely derive from enhanced client value and satisfaction that result from gains in review efficiency and attorney effectiveness. While an indirect source of value, these benefits can help provide competitive differentiation, client retention, and long-term revenue gains.

As organizations weigh the potential gains, they must also consider how the functionality offered by a vendor fits their needs, the complexity and difficulty of use and implementation, the ease of adoption and use by attorneys, and the total ownership costs imposed. A comparison of these factors against expected value contributions offers a basis from which an organization can understand both whether investment is justified and which solutions best fit their needs. As data volumes grow, an increasing number of corporations and law firms will need to begin to make these calculations to better understand the potential benefits of visual analytics.
ABOUT THE AUTHOR

David Houlihan
Principal Analyst

David Houlihan researches enterprise risk management, compliance and policy management, and legal technology. He is an experienced advisor in legal and technology fields with a unique understanding of complex information environments and business legal needs.

CONNECT ON SOCIAL MEDIA
@DWHoulihan
www.linkedin.com/in/houlihandavid
bluehillresearch.com/author/david-houlihan/

For further information or questions, please contact us:

Phone: +1 (617) 624-3600
Fax: +1 (617) 367-4210

Twitter: @BlueHillBoston
LinkedIn: www.linkedin.com/company/blue-hill-research

Contact Research: research@bluehillresearch.com

Blue Hill Research is the only industry analyst firm with a success-based methodology. Based on the Path to Success, Blue Hill Research provides unique and differentiated guidance to translate corporate technology investments into success for the three key stakeholders: the technologist, the financial buyer, and the line of business executive.

Unless otherwise noted, the contents of this publication are copyrighted by Blue Hill Research and may not be hosted, archived, transmitted or reproduced, in any form or by any means without prior permission from Blue Hill Research.
James Haight is a research analyst at Blue Hill Research focusing on analytics and emerging enterprise technologies. His primary research includes exploring the business case development and solution assessment for data warehousing, data integration, advanced analytics and business intelligence applications. He also hosts Blue Hill's Emerging Tech Roundup Podcast, which features interviews with industry leaders and CEOs on the forefront of a variety of emerging technologies. Prior to Blue Hill Research, James worked in Radford Consulting's Executive and Board of Director Compensation practice, specializing in the high tech and life sciences industries. Currently he serves on the strategic advisory board of the Bentley Microfinance Group, a 501(c)(3) non-profit organization dedicated to community development through funding and consulting entrepreneurs in the Greater Boston area.