

WHITE PAPER

The efficiency imperative:

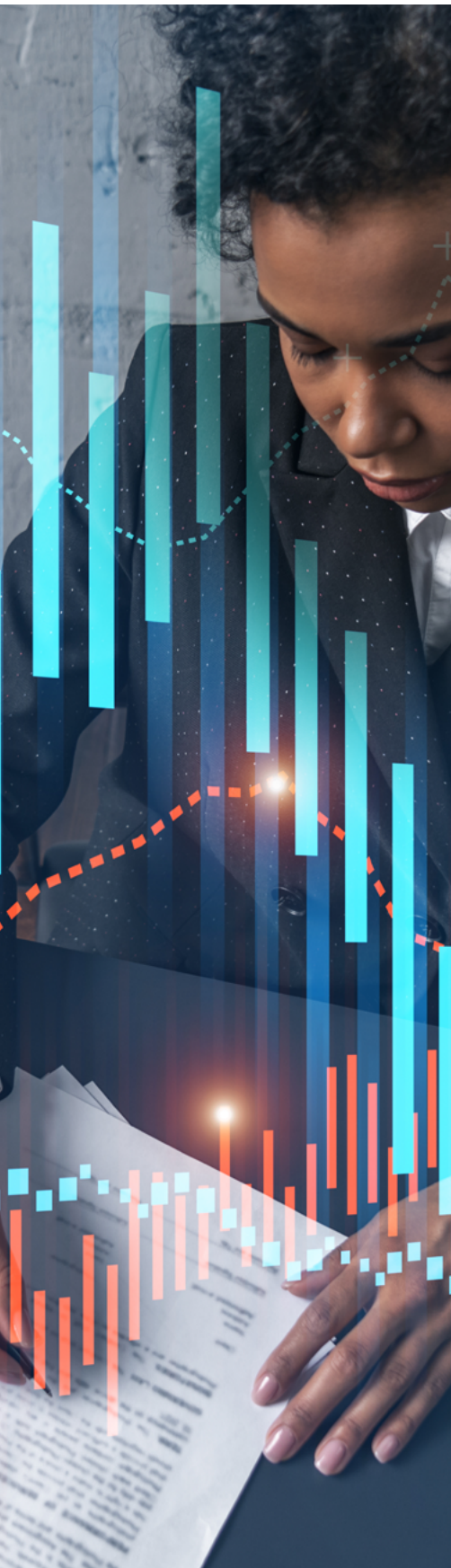
How corporate legal departments and law firms are bolstering their eDiscovery postures



CONNECTION
ANALYSIS
DATA
SEARCHING
VERIFICATION
CODING
SENDING

Contents

Executive summary	3
The efficiency imperative—drivers and implications	4
Current and emerging needs of legal departments and law firms: Key considerations in eDiscovery approaches	6
Architectural foundations of efficient eDiscovery	6
Cloud native versus on-premises deployments for eDiscovery technology	6
Integrated eDiscovery platforms versus point solutions	7
eDiscovery system scale and performance	7
New and emerging eDiscovery requirements	8
Visualization and visual analytics	8
Personal data identification and protection	9
Support for chat data review	10
Analytics designed for ECA and investigations	11
Enabling efficient eDiscovery: The EDRM perspective	12
Legal Hold	12
Collection and ingestion	12
Processing and narrowing the review set prior review	13
Analysis	13
Review	14
Production	15
Optimizing resource efficiency	16
Summary	16



More so than ever, legal departments are under intense resource and budget pressure. These challenges are exacerbated by escalating data volumes; new and emerging forms of business communication, such as chat; heightened regulatory and data privacy mandates; an increase in remote work; and rising numbers of investigations due to a whistleblower culture. This position paper explores the pressures affecting legal departments and law firms and discusses how eDiscovery solutions have evolved to help legal departments fulfill the efficiency imperative.

Executive summary

Gone are the days when litigation was viewed as a force that simply happened and could not be controlled, along with runaway spend on outside counsel. Today, corporate legal departments are under intense pressure to deliver more strategic value to their organizations and drive efficiency across all programs and processes within defined budgets. In fact, a key priority for 73 percent of in-house counsel is controlling spend on outside counsel.¹

While “do more with less” pressures aren’t new, the need for efficiency has never been more acute. Workload has increased by 51 percent while budgets have shrunk by 23 percent in the past two years. Seventy percent of legal department leaders cite a renewed focus on adopting new or better technologies to simplify workflows and reduce manual processes to cut costs.²

In response, law firms are introducing alternative fee agreements and other creative pricing strategies. Firms must try to balance the delivery of better service at a lower cost while improving profit per partner and the firm’s bottom line.

Meanwhile, concurrent trends are further increasing the costs, risks and overall complexity of eDiscovery—for example, the rise in the volume and variety of data, such as chat, the proliferation of personal data protection regulations, the ongoing threat of data breaches and the adjustment to remote work as an enduring reality.

As a result, corporate legal departments and law firms need eDiscovery solutions that help improve efficiency and control costs.

This position paper highlights strategies to drive efficiency and cost savings throughout the Electronic Discovery Reference Model (EDRM) and describes what legal teams should demand of their eDiscovery software to help them rise above the challenges.

73% Counsel
in-house vs outside

51%
Workload

23%
Budgets

¹ LegalBillReview.com, Reducing Legal Spend: A survey of in-house attorneys. (2020)

² Thomson Reuters, State of corporate law departments report. (2021)

Petabytes per year



Investigations volume

predicted to increase in the next 1-2 years



65% of world's population

covered by law to protect personal data by 2023



The efficiency imperative—drivers and implications

More than ever, legal departments are under intense resource and budget pressures all driving the need to improve internal efficiencies. These include:

- **Continued growth in data generation:** In 2020 the total store of data worldwide was estimated at 44 million petabytes. It is estimated that 168 million petabytes will be generated in 2025 alone.³ The volume of data involved in litigation and investigations is rising in step, with legal teams reporting that data volumes have increased by 61 percent in the past two years.⁴
- **Continued expansion in data variety:** Unstructured data will continue to account for more than 90 percent of all data, partially due to the rise of newer forms of data such as chat. In 2021, more than 600 billion business chat messages were sent, with personal chat being at least an order of magnitude higher.⁵ New forms of data bring new challenges in collection, processing, review and production.
- **The rise of investigations:** Organizations across industries and sectors are witnessing a growth in the volume and types of investigations that legal and compliance teams must learn to conduct effectively and efficiently. In a recent survey by Compliance Week and OpenText, 43 percent of legal and compliance leaders expected the volume of investigations to increase in the next one to two years.⁶
- **Proliferation of personal data protection regulations:** By 2023, 65 percent of the world's population will be covered under modern laws to protect personal data.⁷ Data privacy regulations force organizations to honor all of the data that contributes to defining an individual as the exclusive property of the individual. Legal departments and law firms are under increased pressure, scrutiny and risk to protect personal data and efficiently respond to a Subject Rights Request (SRR), including GDPR Data Subject Access Requests (DSARs).



³ SeedScientific "How Much Data is Created Every Day", October 2021

⁴ OpenText internal client research

⁵ OpenText internal analysis – numerous sources

⁶ Compliance Week and OpenText, What factors are driving change in your corporate investigations process?. (November 2021)

⁷ Gartner Newsroom (September 2020)



\$9.6 billion

Cost of data breaches
in the U.S. in 2020



25% to 30%

Employees working from home



75%+ applications

migrated to the cloud

**3/4+ law firms
and corporate
legal departments**

use alternative legal service
providers on a regular basis



- **Data breaches remain a clear and present danger:** In the first nine months of 2021, the number of reported data breaches in the U.S. (1,291) had already exceeded the total for 2020 (1,108) by 17 percent.⁸ The total cost of breaches in 2020 reached a staggering \$9.6 billion in the U.S. alone (\$8.64 million per breach).⁹ In addition to protecting personal data within eDiscovery and investigations projects, legal teams now need to minimize the exposure of personal and sensitive data throughout its lifecycle.
- **Remote work will remain prevalent:** The percentage of employees working from home peaked at 69 percent in April 2021, but forecasts peg the number of people working from home long-term at 25 to 30 percent, approximately 10 times higher than pre-pandemic.¹⁰ Remote work elevates the demand for cloud-based solutions to aid collaboration on eDiscovery projects wherever reviewers reside. Remote work also increases the need for centralized data stores to avoid wasted effort and data porting security risks.
- **CIO evolution to faster service delivery via the cloud:** CIOs will continue to face a variety of challenges, including a focus on reducing the use of point solutions to speed service delivery by 50 percent.¹¹ Moreover, more than half of CIOs will migrate more than 75 percent of applications to the cloud to better accommodate remote work, improve ability to rapidly deploy and scale solutions and reduce CapEx in favor of OpEx.¹² Part of their objective is to offer choice of deployment models to fulfill their mandate as enablers of efficiency instead of being gatekeepers to the myriad concurrent requirements of the organization.¹³
- **The ongoing need to fill gaps in resources and expertise and speed service delivery:** Legal departments and law firms continue to rationalize in-house capabilities and resource levels with support from service providers for specialized expertise and added capacity for large projects and peak periods. In fact, 79 percent of law firms and 71 percent of corporate legal departments use alternative legal service providers (ALSPs) on a regular basis because of the breadth of their offerings and prolific availability.¹⁴

⁸ Identity Theft Resource Center, Number of Data Breaches in 2021 surpasses all of 2020. (October 2021)

⁹ Ponemon and IBM, How Much Does a Data Breach Cost? (2021)

¹⁰ Global Workplace Analytics, Work-at-Home After COVID-19 – Our Forecast. (2021)

¹¹ Gartner, CIO Agenda – The leadership, organizational and technology priorities CIOs must address in 2022. (January 2022)

¹² McKinsey, Cloud's trillion-dollar Prize is up for grabs. (February 2021)

¹³ Unily, Top 5 CIO Trends to Know for 2022. (December 2021)

¹⁴ Thomson Reuters, ALSP 2021 Report.

Current and emerging needs of legal departments and law firms: Key considerations in eDiscovery approaches

As organizations mature in their eDiscovery and investigations tasks, many are looking across the diverse criteria affecting their ability to be more efficient and reduce costs. These criteria can be grouped in three main categories—the architectural foundations of eDiscovery and investigations solutions, new requirements that span the EDRM and criteria that align to a particular stage of the EDRM. Finally, legal departments and law firms are assessing their requirements for efficiency within the context of their resources and skillsets to optimize what functions they perform in house and when to rely on service providers.

This position paper will look at each of these categories in order, along with the specific criteria within and how they affect the imperative for efficient eDiscovery and investigations.

Architectural foundations of efficient eDiscovery

Cloud native versus on-premises deployments for eDiscovery technology

Cloud-native deployments lower OpEx and CapEx (whether OnDemand or private cloud), improve ease and flexibility and address data privacy and data sensitivity considerations. With respect to the latter, the desire to retain sensitive data (confidential, personal, IP and trade secrets, etc.) in house varies between organizations, depending on their perceptions of cloud security, but often includes heightened sensitivity to keeping and being able to review sensitive data, such as the CEO's emails, on premises.

It is well established that cloud solutions are at least as secure as on-premises deployments. But on-premises remains a preference for specific types of sensitive data because of the perceived ability to retain closer control. As such, hybrid deployments, to get the best of both models, are likely to remain a long-term need for many corporate legal departments.

Cloud deployments in eDiscovery are on the rise due to 24/7 availability; ease of collaboration between legal departments, law firms and service providers; less reliance on IT; and scalability for burst, or unexpected data volumes that can add significant performance pressure to an on-premises system. Another significant efficiency gain is the immediate rollout of new features and functionality—versus waiting on IT for the hardware, software and time to install and test new versions.

Cloud subscription models are particularly effective at promoting efficient eDiscovery because cases and data volumes can be moved around if priorities change.

IT and legal departments often select on-premises deployments to retain end-to-end control over eDiscovery projects and the associated data, but many organizations recognize the burden that on-premises systems puts on CapEx.

To balance the benefits of cloud with the desire to retain sensitive data solely on-premises, hybrid models have emerged, where sensitive data can be reviewed (and maintained) on premises, with the rest of the project data hosted securely in the cloud.



Integrated eDiscovery platforms versus point solutions

Integrated eDiscovery addresses two of the predominant trends affecting legal departments and law firms—the need for data security and the efficiency imperative. The key attribute of integrated eDiscovery is that project data is held in a common data store to avoid having to secure multiple locations in point solutions and port data between those systems, which can take significant IT time and resources, and cause significant project delays. For example, one tool performs collection and ingestion, while another platform performs analytics (perhaps bolted onto a review platform) and another handles review and production.

The movement of data between point solutions is prone to human error or being compromised during transport between applications. The inefficiencies, risks and costs are magnified when data is moving between internal eDiscovery tools and external services providers, whether law firms, ALSPs or both.

With an integrated comprehensive platform, as projects progress across the EDM, instead of the data being moved from one application or module to another, the vantage point of the application changes—one tightly controlled central data store, multiple lenses.

Single and centralized eDiscovery systems avoid time wasted porting data between applications, which can take days, because multiple stakeholders need to sign off that the data is ready to be moved to the next stage of a project. Integrated eDiscovery systems also handle rolling loads better because new data is simply added to the central data store, where it can be processed and prepared for review within the project, instead of having to progress through different applications to catch up to the rest of the project.

Templatized workflows also allow for repeatable processes across matters, rather than reinventing the wheel with each new case. Integrated eDiscovery also eliminates the time and effort required to source and assemble multiple point solutions.

eDiscovery system scale and performance

Legal teams must handle large volumes of project data and ingest it at impressive speeds. Every minute spent watching an hourglass spin eats away at deadlines. Most leading solutions can ingest data quickly, at approximately three gigabytes of data per hour, which is about three million emails.

A key attraction of cloud solutions is their ability to support burst capacity through custom configuration to balance the amount of time available with the volume of data that needs to be ingested. Legal departments and law firms can get ahead of the need for burst capacity through cloud subscriptions—paying for active data across all projects while storing data not in current review projects (without incurring additional costs). They can diversify their requirements across projects within a given subscription size (number of terabytes) and forecast the need for periodic exceptional ingestion requirements and have machines pre-configured to avoid the time lag of setting up custom configurations.

New and emerging eDiscovery requirements

The second set of criteria that legal departments and law firms are assessing to fulfill the imperative for efficiency are related to new and emerging trends that span the EDRM and increase the complexity and cost of eDiscovery.

Visualization and visual analytics

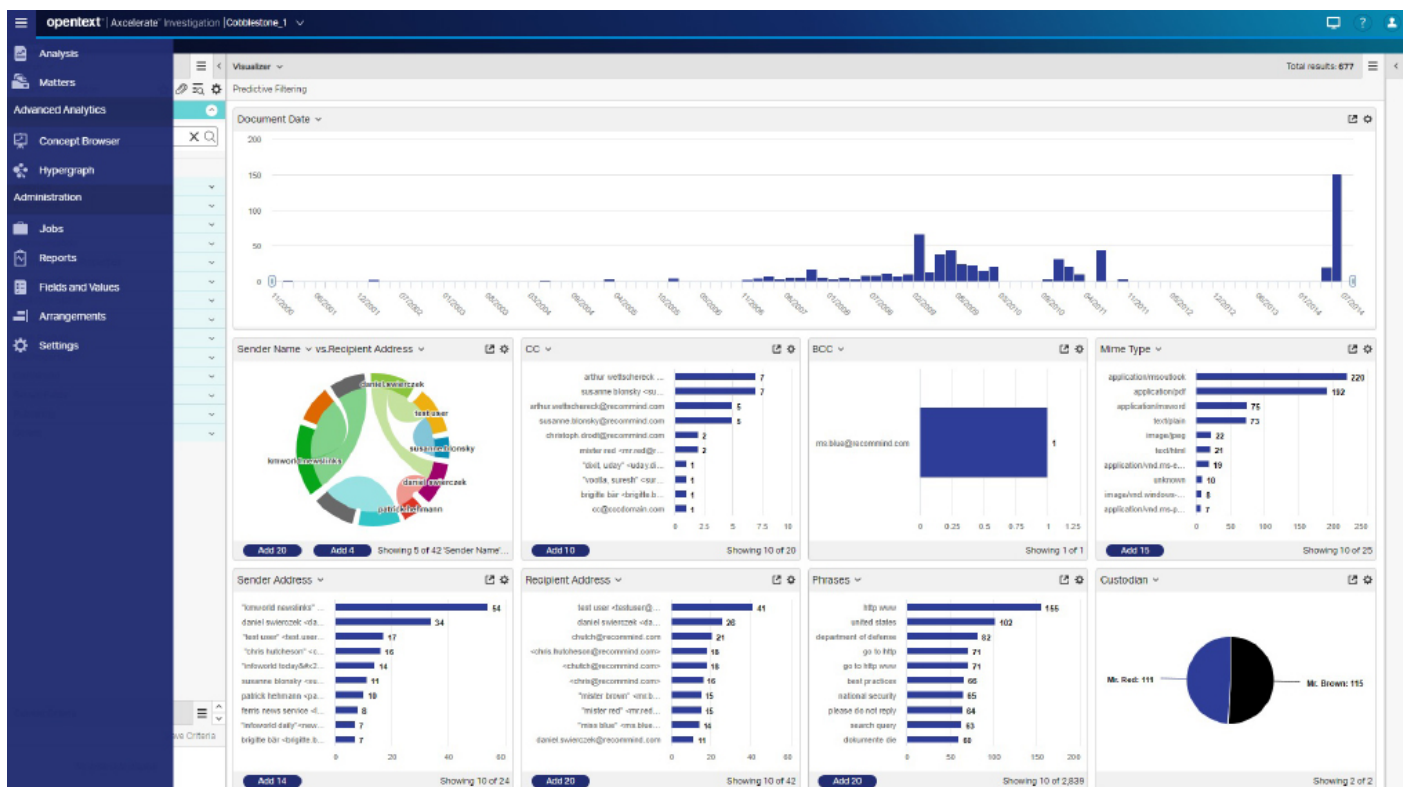
Visualization enables legal teams to navigate data more quickly to substantially expedite review.

Visualizations display specific data such as date histograms, communication interaction networks, concept groups and phrase analysis and support the ability to pivot within the data from one vantage point to another. This is essential for following communications chains between custodians, tracing interactions between domains and drilling down on the associations between concept groups, etc.

First, customizable visual dashboards can speed the eDiscovery process by providing the ability to select the most useful viewpoints into the data for a particular matter, showing specific data with the granularity appropriately customized for each.

Second, visual analytics introduce two-factor analysis to eDiscovery through the ability to visually associate any two data or analytics attributes. For example, instead of seeing file size and file type displayed independently, file size can be overlaid on file type for deeper understanding of the composition of data sets. When analytics are included, new opportunities for efficient review and deeper insights become possible. For example, sentiment analysis can be overlaid on fact vs. opinion analysis to quickly test whether there are patterns between custodians that are the most negative and those that are also most prone to unsubstantiated assertions. In addition to helping reviewers home in on key custodians faster, visual analytics offer new vantage points for the interpretation of evidence.

Visual dashboard – OpenText™
Accelerate™ Visualizer





Personal data identification and protection

For many years, legal teams have required eDiscovery solutions with tools for protecting certain types of data to safeguard against disclosing attorney-client privilege and other sensitive data in outbound productions. The rise in personal data protection regulations has increased requirements for enhanced personal data protection tools within eDiscovery solutions. These tools also allow legal departments and law firms to pursue related personal data protection use cases, including data breach response and the application of eDiscovery workflows to subject rights requests (SRRs), including data subject access requests (DSARs).

eDiscovery solutions with robust data protection tools are ideal for data breach response because they can efficiently extract specific relevant data from within large volumes of irrelevant data and produce detailed reports of the findings. Also, 23 percent of data breaches include confidential business information (CBI), such as IP, contracts and client lists, etc.¹⁶ CBI takes many forms and requires the same robust filters and text analytics tools as eDiscovery to assess both the content and context of data to determine relevance.

Subject rights requests are entrenched within data privacy laws, such as GDPR and CCPA, and are available to customers, employees and former employees alike. Requests from employees are often the most complex and typically involve data across a wide variety of systems in a wide range of formats and spanning longer time periods. Seventy-one percent of EU organizations have received DSAR requests from employees.¹⁷

The tools that legal teams need for the efficient discovery and protection of personal data within eDiscovery solutions include a library of pre-configured patterns (strings) to quickly detect common patterns (email addresses, credit card patterns, etc.), custom regular expression (RegEx) tools for inputting any custom string, discrete and bulk redaction, automated quality control (QC) and the ability to automatically detect the names of people, organizations and places.

Efficient data protection processes require the provision of a broad set of easily accessible and easy-to-use tools. eDiscovery products with embedded data protection tools offer the most efficient protection of personal data. Including personal data detection and redaction tools in central actions menus and tying them into search filters improves efficiency because reviewers can execute data protection during processing and review instead of having to find and redact personal data as a standalone process.

Regex Pattern Search (Enrich)

30 documents will be processed - This will start batch jobs for the selected patterns listed below and the enrichment results will be available once the jobs have finished.

Select at least one pattern to start enrichment.

Social Security Number (close match)

☐ The fuzzy Social Security Number pattern finds US social security numbers similar to the SSN format. Separators are more flexible, but therefore a SSN related text is expected.

Social Security Number (SSN)

☐ The exact Social Security Number pattern finds US social security numbers exactly matching the SSN format.

Credit Card

☐ The "Credit Cards" pattern finds VISA, MasterCard, Amex and Discover credit cards in a combined expression.

Email

☐ The "Email" pattern finds email addresses.

Tax Identification Number (TIN)

☐ The "TIN" pattern finds tax payers identification numbers.

Birthdate

☐ The "Date of Birth" pattern finds birth dates in both DD/MM/YYYY and MM/DD/YYYY format.

Employee Identification Number (EIN)

☐ The "EIN" pattern finds employer identification numbers.

Medicare

☐ The "Medicare" pattern finds US Medicare numbers close after the word "medicare".

Passport

☐ The "Passport" pattern finds US passports numbers close after the word "passport" or "passports".

Cancel Enrich

Library of pre-configured personal data patterns

¹⁶ Ponemon and IBM, How Much Does a Data Breach Cost. (2021)

¹⁷ GRC World Forum, Employee DSARs: The Coming Deluge. (August 2021)

Support for chat data review

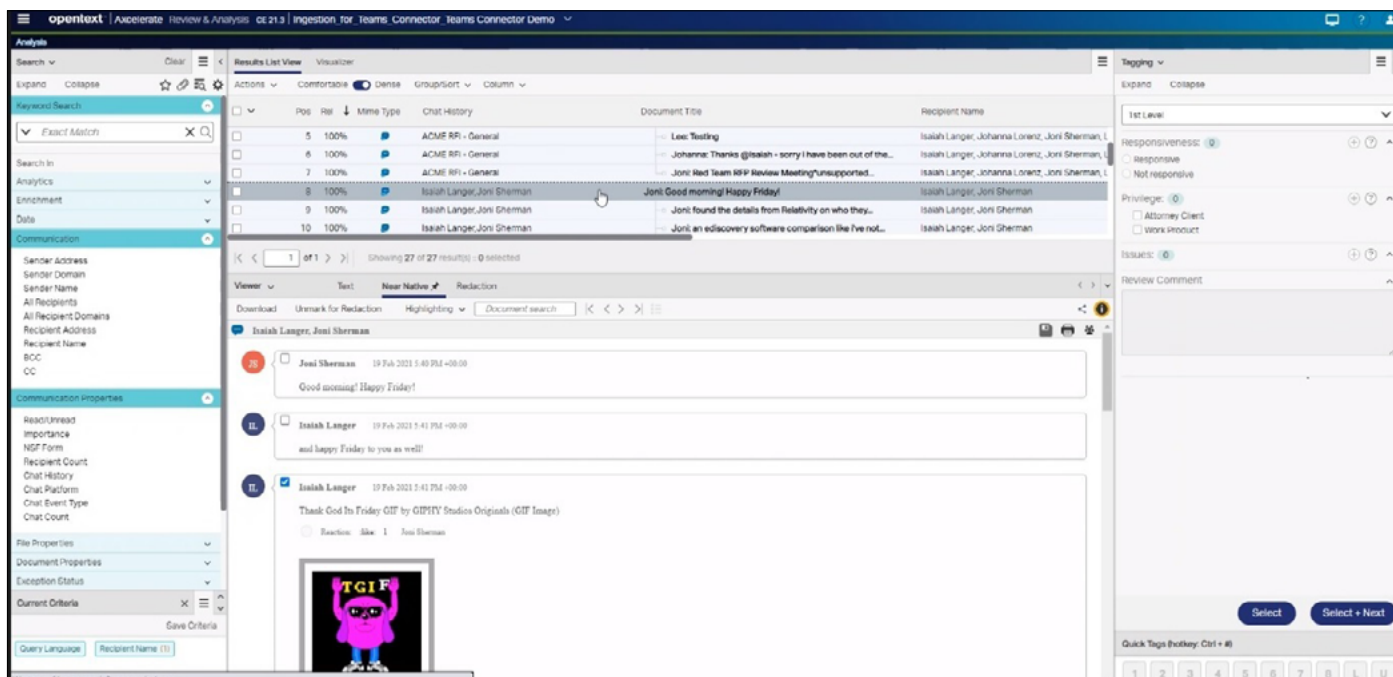
Business chat messages are expected to soon exceed one trillion per year, in addition to the trillions of personal chat messages sent each year. Thirty-seven percent of eDiscovery professionals say that chat data is involved in 50 percent or more of their litigation matters today.¹⁹

Chat data can be a cumbersome and error-prone process from collection through to review, but legal teams need to be able to handle chat data as efficiently as any other form of data in the eDiscovery dataset. Supporting chat data for investigations and litigation requires capabilities across the EDRM, including:

- Connectors to the most used business chat applications, such as Microsoft Teams® and Slack™.
- Dedicated parsers and/or generic parsers for all other sources of business and consumer chat applications.
- Proprietary chat formats within eDiscovery applications, so chat data can be displayed in a familiar chat layout.
- The ability to include chat data alongside all other forms of data, so that visualizations and analytics tools provide a comprehensive picture and analysis of all data, including chat.
- Broad inclusion of all chat activities, including emoticons and other reactions, and the ability include leave/join events.
- Full support for attachments to chat messages.
- The ability to narrow the volume of chat data for review by both blocks of time and by issue.
- The ability to batch out chat data in its own groups so reviewers don't have to jump between document or email layout structures and chat layout structures.
- The ability to apply redactions granularly within individual messages, across entire threads and in bulk across all chat data and the inclusion of chat data in automated redaction QC processes.

Chat data is among the most complex eDiscovery data formats that legal departments and law firms must accommodate. Handling chat data efficiently and effectively requires a broad set of tools as described above. Legal teams should look for these capabilities to be provided within an easy-to-use framework.

¹⁹ eDiscovery Today, 2022 State of the Industry Report. (2022)



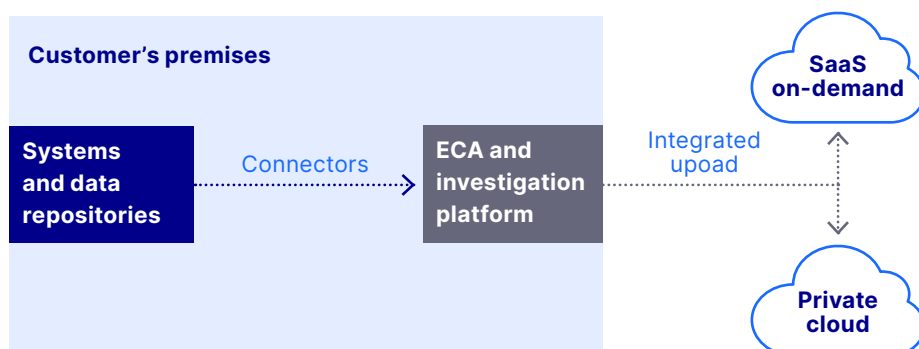
Familiar chat layout in
OpenText™ Axcelerate™

Analytics designed for ECA and investigations

Getting to the root of matters and discovering the facts to inform case strategy and make “fight or flight” decisions is becoming increasingly difficult due to the rise in the volume and variety of data. Legal teams require eDiscovery software that exposes many of the analytics tools associated with review so they are available earlier in the process to get to key evidence—without the time and associated costs of completing a full review workflow.

Specialized solutions purpose-built for early case assessment (ECA) and investigations with front-loaded analytics help find the facts quickly without a protracted eDiscovery-style review.²⁰ These solutions promote efficient investigative processes by providing search filters, analytics and visualizations but do not include tools more aligned to reviewing data for litigation, such as full technology-assisted review (TAR) and production workflows.

Specialized ECA and investigations solutions also drive efficiency if a full review is required by supporting the ability to easily port data to on-demand eDiscovery solutions for review and production.



²⁰ View the OpenText Axcelerate Investigation product page as an example of a specialized investigations product



Enabling efficient eDiscovery: The EDRM perspective

The third set of criteria that legal departments and law firms are assessing to optimize efficient eDiscovery are related to specific stages of the EDRM. While these criteria are considered discretely, their contribution to integrated eDiscovery is also key.

Legal hold

Legal holds arise from the duty to preserve information from pending or reasonably anticipated litigation. In practice, many legal departments still issue holds via email and track them using spreadsheets—a time-consuming, risky and expensive “one off” way to manage holds. Many legal departments still use a traditional workflow in which outside counsel conducts in-person interviews of custodians to identify potentially relevant data sources. Then, a forensic collection vendor may gather data and a specialized eDiscovery vendor may process the data and load it for review. Under this model, the process is inefficient and time-consuming, not to mention disruptive to employee productivity.

Automated legal hold technology delivered via the cloud simplifies these processes within a single, centralized system that automates legal hold notification and integrates with email systems for in-place preservation, Microsoft® Active Directory® for custodian management and HR systems for employee change management. This allows legal department stakeholders to run preservation protocols appropriately. Automation reduces errors and risks and when connected to review platforms, it can provide legal teams with quick, seamless access to critical documents—rather than taking days (or weeks!) when relying on law firms and/or internal IT.

Collection and ingestion

Legal departments and law firms primarily use two methods for collecting and ingesting potentially relevant data. Drag-and-drop ‘self-service’ is the main method typically offered by cloud-only vendors, while data source connectors are most commonly associated with on-premises solutions. (Some cloud products also provide connectors but often only for cloud sources of data.)

- **Drag-and-drop:** Because of limited IT involvement and skills needed for drag-and-drop, this method is often viewed as faster and easier. However, when search filters are not applied to isolate specific data within data sources, the result can be over-collection, with lower proportions of relevant data collected, more to process and higher review costs. Also, for certain types of data, such as chat, the ease and speed of drag-and-drop is compromised when exports from source applications, data staging and the use of custom scripts to parse data is required.
- **Connectors:** Connectors can collect more precisely and while they may require more effort up-front, that time is saved through reduced downstream effort compared to the time spent processing and preparing less refined collections for review. Expending the effort earlier improves the relevance ratios of data from the outset to promote overall project efficiency. Further, limited IT involvement reduces collaboration cycles and time to complete collections by putting collection processes in the hands of corporate legal teams. Finally, connectors are more adept at handling complex data forms, such as chat. The downside of connectors is that data source architecture changes, so software development must evolve connectors in step and add new ones for emerging sources.

An efficient approach is to employ drag-and-drop in concert with front-loaded analytics and easy-to-use search filters to build streamlined review sets quickly after the data is ingested.

Processing and narrowing the review set prior review

Legal departments and law firms have fueled a substantial evolution in how data is processed for investigations and eDiscovery. Processing has taken on a broader context than deduplicating, de-NISTing and handling exceptions. Processing now includes an enhanced focus of applying pre-review analytics to narrow review sets even more, to tackle the cost of eDiscovery where it is highest—70 percent of eDiscovery cost is spent on review.²¹ Additional tools and processes increasingly valued by legal teams and being applied to the processing and pre-review stage include:

- Concept grouping and phrase analytics.
- Front-loaded analytics, such as document summaries, sentiment analysis, fact vs. opinion analysis and the automated identification of people, places and organizations.
- Predictive search and/or predictive filters that use small sets of known relevant data to find similar data that is also likely to be relevant by finding data with related content and context.

Legal departments and law firms benefit from the enhanced efficiency by being able to more quickly narrow initial collections and effectively complete the pre-review processes. Refined review sets with higher proportions of relevant data also substantially reduce the volume of data that requires eyes-on review, driving further efficiency.

Analysis

Analytics help to quickly assess the content and context of data so that data likely to be relevant can be pulled forward for review and data unlikely to be relevant left behind.

How efficiency is enabled:

- **Automated concept grouping and phrase analytics:** Machine-driven algorithms can find data that reviewers cannot. Reviewers only know what they know and typically only surface a fraction of the relevant content when concepts and phrases are search-driven. Richer concept groups and phrase analytics provide more accurate details on the contextual depths of data.
- **Context-based document summaries:** Document summaries are sometimes an aggregate of headers or copies of executive summaries. Allowing reviewers to quickly see what data warrants closer inspection eliminates time spent on data with low or no relevance. Document summaries that assess and aggregate the lexical constructs and context of data are essential for focusing time where it is required.
- **Sentiment analysis** assesses the concepts and phrases within data alongside a library of terms that typically indicate sentiment (love, like, dislike, hate, etc.) and provides a score for the percent of content that is positive, neutral or negative. Sentiment analysis helps to efficiently narrow review sets because very strong positive or negative sentiment often points to the key custodians of interest.

²¹ EDRM.



- **Fact vs. opinion analysis** is similar to sentiment analysis but uses a library of terms that suggest whether assertions are substantiated or not ("just guessing" etc. "the data indicates" etc.). Fact vs. opinion analysis expedites the selection of data for deeper assessment of the assertions within.
- **Automated identification of people, places and organizations** fuels efficiency by helping to find all relevant custodians faster, by homing in on cases tied to specific locations or narrowing data to one entity or subsidiary involved in a matter.

There is an evolving appreciation of the interplay between long-standing forms of text analytics (e.g. concepts and phrases) with newer forms (e.g. fact vs. opinion analysis), with each supporting the other. For example, sentiment analysis can help identify the key concepts to focus on and as a package, these tools help to substantially expedite review.

Review

Review is the costliest component of eDiscovery and many corporate legal departments are seeking ways to cut costs by limiting human review. To optimize review, legal teams are looking for technology-assisted review, predictive search, granular control over batching and reporting to help drive review efficiency.

TAR based on continuous active learning

Today's more advanced legal teams use technology-assisted review (TAR) to limit human review. TAR based on a continuous active learning (CAL) protocol is one of the easiest ways to reduce discovery costs. The ratio of responsive documents reviewed with TAR based on continuous active learning is far superior to that of keyword search, followed by linear review or other less sophisticated forms of TAR, achieving an average ratio of only two documents to review for every document that is responsive (or even lower). Relevance ranks are adjusted as additional data is coded and quickly achieves a hierarchy that aligns to the parameters of a matter where proportional review can be achieved with a fraction of the effort.

CAL systems also mix in contextually diverse documents, a process by which the algorithms actively find documents that may be related but are unlike other documents that have been reviewed. This helps reviewers overcome bias and find documents they might not otherwise see.

Predictive search

Predictive search enables efficiency for legal teams as an alternative to TAR and is often called "predictive coding on-the-fly." Instead of proceeding through review based on relevance rankings, a small set of data known to be highly relevant is held up for comparison against the entire review set to find similar data that matches the concepts and phrases of the relevant data. This quickly isolates the data that requires eyes-on review. Predictive search can also be used as a quick and effective QC tool towards the end of review. If the most relevant data does not surface documents that have not been included for review yet, case managers can reasonably conclude that proportional review has been achieved.

Batching

Granular control over batching enables efficiency by allowing case managers to assign documents out for review based on the criteria most related to the case at hand. Reviewers can better focus and achieve more consistent review decisions with less rework on overturns by covering all data related to the same issue, with the same concepts or in the same format, such as chat.

Redaction

In addition to the ability to redact data within complex formats, such as Microsoft® Excel® and chat, legal teams need easily accessible and easy-to-use redaction tools for all forms of data. Requirements include the support for highlighting, so reviewers can see and approve redactions, the ability to redact discretely to specific terms within individual documents and the ability to apply redactions in bulk across large sets of data. Streamlined, yet robust, redaction processes substantially reduce the time and effort required to de-risk data and are best used alongside automated QC processes, where machine-driven processes help assure accuracy while reducing manual QC effort.

Business intelligence and reporting

Legal departments and law firms rely on business intelligence (BI) in the form of reports and dashboards to improve the oversight of review processes, including review progress and reviewer productivity reports, along with overturn reports with clickable links to the specific documents.

Production

For legal teams, producing accurate and de-risked discovery sets is among the most critical aspects of eDiscovery. Features that promote efficiency include:

- Flexible output options, including native, TIFF and PDF.
- Templates for re-using common production settings to avoid starting from scratch each time.
- Automated QC for checking the accuracy of productions.
- Production reports that enable easy assessment that the data produced is inclusive of all the intended data and only the intended data.



Resource links

- [➔ OpenText end-to-end eDiscovery](#)
- [➔ OpenText Axcelerate](#)
- [➔ What factors are driving change in your corporate investigation processes](#)
- [➔ How to handle chat data in eDiscovery and investigations](#)
- [➔ Contact us](#)

Optimizing resource efficiency

Finally, as organizations look for optimally efficient eDiscovery models and technology to meet their needs, they are also assessing how to optimize and balance in-house resources and expertise with third-party expertise to augment gaps in resources and accelerate project timelines—from collections and forensics to managed document review, consulting services, investigation support and expert witness services. Third-party providers can come in when last-minute needs arise and can also implement repeatable and efficient processes for consistent support and process optimization across cases.

ALSPs and vendors generally fall into two categories: those that provide services (including hosting) using third-party software platforms and those that develop proprietary technology and provide services leveraging that technology.

As corporations seek to streamline the eDiscovery process, many are looking to direct-from-vendor professional services due to the simplified process of sourcing services, streamlined billing, one-stop accountability and enhanced confidence that vendors are experts in optimizing the technologies they develop and deliver.

Summary

Legal teams are looking for integrated and highly scalable solutions as a foundation for smooth and repeatable workflows without having to port data between modules. These systems include data connectors for efficient collection, front-loaded analytics in the processing and culling stage, rich analytics, data protection tools, effective management of new forms of data, rich technology-assisted review and more.

By taking a holistic and multi-faceted approach to reducing obstacles and enabling fast and effective processes across the EDRM, legal departments and law firms can deliver more efficient and effective eDiscovery programs.

About OpenText Axcelerate and OpenText eDiscovery Services

OpenText Axcelerate is a flexible and powerful end-to-end eDiscovery and investigations platform built around proprietary advanced analytics, machine learning and automation. Axcelerate delivers best-in-class investigative capabilities in a fully integrated, intuitive review interface that helps legal teams get to the facts that matter sooner and inform case strategy.

OpenText eDiscovery Services augment client resources and provide expertise to help corporations and law firms master the demands of complex litigation, investigations and regulatory matters. Forging best practices from thousands of matters and clients across the globe, eDiscovery Services fuel innovative approaches to lowering costs and improving efficiency.

About OpenText

OpenText, The Information Company, enables organizations to gain insight through market leading information management solutions, on-premises or in the cloud. For more information about OpenText (NASDAQ: OTEX, TSX: OTEX) visit: [opentext.com](https://www.opentext.com).

Connect with us:

- [OpenText CEO Mark Barrenechea's blog](#)
- [Twitter](#) | [LinkedIn](#)