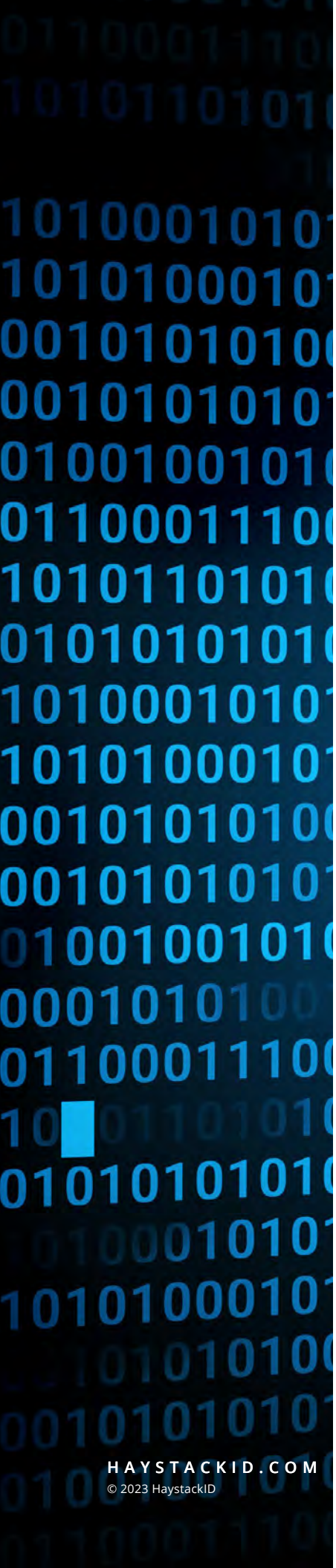
INFORMATION PAPER

Advanced ECA Techniques in Microsoft Purview eDiscovery (Premium)

A Process Design Framework from HaystackID®

By Jason L. Covey, M365 eDiscovery Consultant, HaystackID

HAYSTACK®



Introduction

Early Case Assessment (ECA)[1] is a set of techniques used in the field of eDiscovery to help rapidly compile and examine data related to audits, investigations, and litigation. While not a specific stage in the EDRM™ (Electronic Discovery Reference Model) process, ECA is regularly associated with the processing and analysis of electronically stored information (ESI) and can provide striking volume reduction and evidence identification benefits when properly applied to eDiscovery-centric projects. Despite its potential power as a discovery tool, ECA is sometimes underused and neglected due to intricacies associated with the challenge of employing multiple technologies and tools to accomplish specific ECA objectives. In this paper from M365 eDiscovery Expert Jason Covey, Microsoft Purview eDiscovery (Premium) (eDiscovery (Premium))[2] is highlighted as an advanced ECA tool that may, based on its integrated and unified capabilities, be able to remove the obstacle of real and perceived complexities of ECA for eDiscovery professionals and help them confidently leverage the power of ECA in their cases and matters.

Early Case Assessment

Despite its longstanding presence in the eDiscovery lexicon, ECA is not an official EDRM stage. Technically spanning multiple stages, ECA is most commonly associated with Processing and Analysis. ECA is also a technique that can be leveraged to achieve some of the most transformative results in an eDiscovery project – particularly in terms of volume reduction and expedited identification of high-value content. Unfortunately, ECA is generally not as well-understood as other eDiscovery processes, and with widely varying ECA capabilities among different technologies, it is sometimes avoided altogether. As this decision sometimes proves highly detrimental in terms of increased project time, effort and cost, it is advisable for case teams to fully consider the opportunities available via a well-executed ECA workflow and avoid missing out on its many potential benefits.

Although eluding any universally-recognized definition, ECA[3] refers to "...the process by which organizations rapidly gather and analyze data about potential matters in eDiscovery, compliance, or internal investigations to reach informed decisions about how to proceed..." Another definition describes "the process of attempting to quickly surface key electronically stored information (ESI), paper documents, and other potential evidence early on in a legal matter." ECA combines legal and technological analysis to quickly gather important case information, reduce the amount of data (and cost), and find relevant evidence as soon as possible in new matters to assist stakeholders in future decision-making.

Advanced ECA in Microsoft Purview eDiscovery (Premium)

With data culling (i.e., the rapid reduction of data volume via the identification and suppression of demonstrably irrelevant content, performed in a legally defensible manner) serving as a primary objective of an effective ECA workflow, the capabilities of eDiscovery (Premium) are of central importance.

The current state form of eDiscovery (Premium) provides organizations with a sophisticated toolset that eclipses those historically available outside of specialized, downstream eDiscovery platforms.

In a manner more on par with industry-leading applications, eDiscovery (Premium)'s native feature set includes: advanced indexing, advanced search capabilities, powerful filtering options, data reporting and visualization features, email threading, email domain extraction, textual near-duplicate detection, document theme extraction, data analytics, and predictive coding.

Advanced ECA Workflow Components

Analytics Processing

- Analytics processing in eDiscovery (Premium) is performed as a secondary (and optional) post-processing operation, manually executed on a "review set," to which a selection of documents has been committed for review. This process is straightforward to perform, with a few additional settings to be configured beforehand.

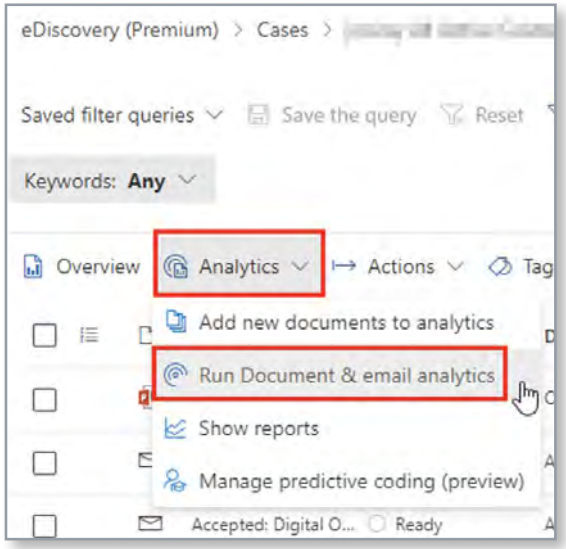


Figure 1

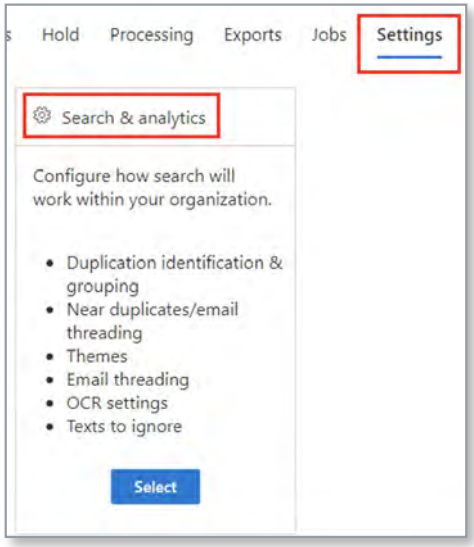


Figure 2

- When analytics processing is complete, a new "Saved filter query" is automatically created, and will appear as "[AutoGen] For Review." Note the pre-filtering item count of 5166.

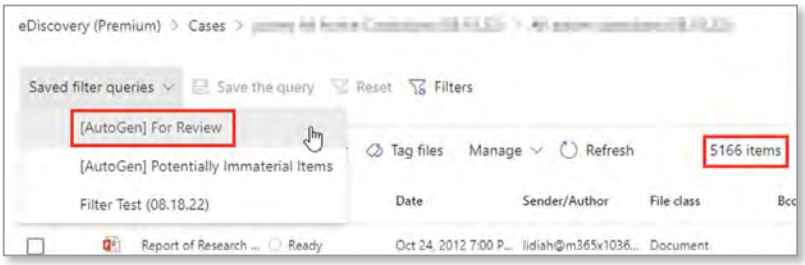


Figure 3

- When this filter is selected, eDiscovery (Premium) uses the new analytics results to suppress content[4] like non-inclusive email items, near-duplicate email items, duplicate documents, etc. **This technology reduced the very small collection of only 5166 documents by 63%, down to only 1880.**

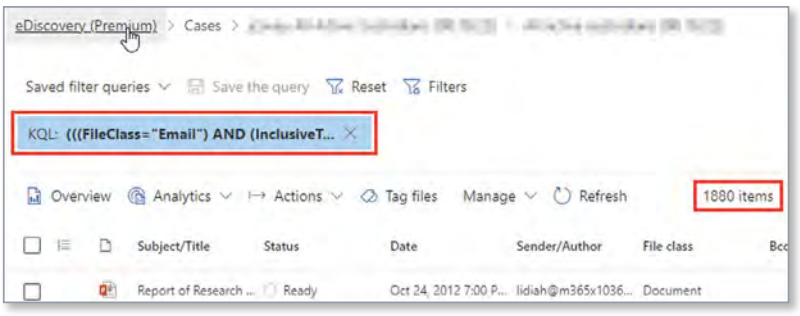


Figure 4

Filtering and Analysis

- Like the analytics results filter, eDiscovery (Premium) includes another automatically-generated filter to help suppress the appearance of "Potentially Immaterial Items" – a common problem in downstream document review.

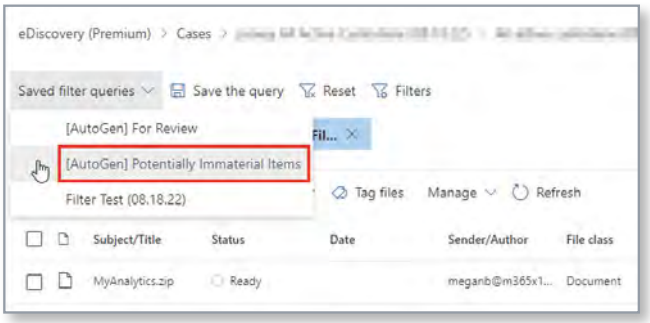


Figure 5

- Another of eDiscovery (Premium)'s important but underutilized analysis features is its easily-customizable document grid or table view. The table view supports a high-level evaluation of the review set content and a much more focused analysis.

- This view is easily configured from its default state via the "Customize columns" button to display whatever available metadata best supports the user's current objective. An example of this might involve a hypothetical effort to determine whether the subject of an ethics investigation communicated with a third party in a manner that violated the organization's policy on such interactions.

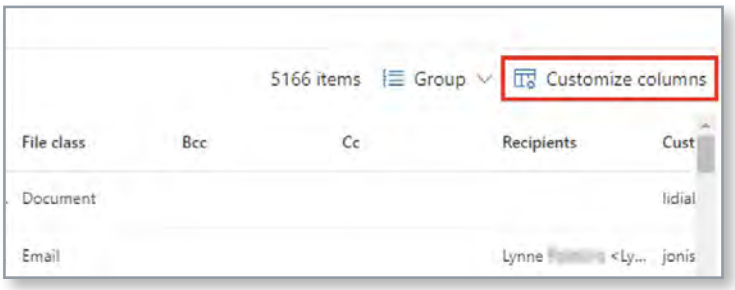


Figure 6

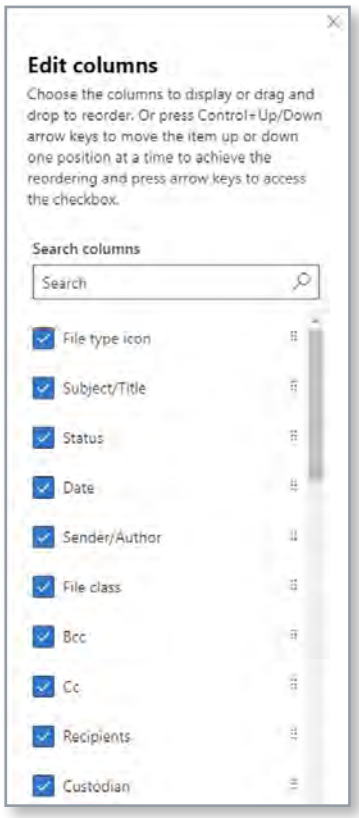


Figure 7

- Although revised in December 2022 to support a paginated format similar to the longstanding user experience in platforms like Relativity®, the document grid also remains available as a virtual table. This means that, for result sets that exceed the available screen real estate, the view will automatically update with additional documents when scrolling, without the need to navigate to subsequent pages manually.

Figure 8

This is a much-appreciated feature when applying the table view to perform iterative ECA tasks and can be enabled as follows:

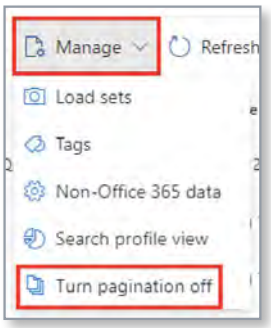


Figure 9

The opportunity for legal teams to directly leverage highly accessible features like the table view as early as possible to rapidly develop new insight represents a profound shift for corporate law departments.

The efficiency in simultaneously leveraging institutional knowledge and legal analysis across a broad range of real-world project scenarios while maintaining all data within the organization's Microsoft 365 (M365) tenant security environment represents a win for both legal and information security camps.

- The following collection of screenshots reflects the complete list of filters currently available in eDiscovery (Premium):

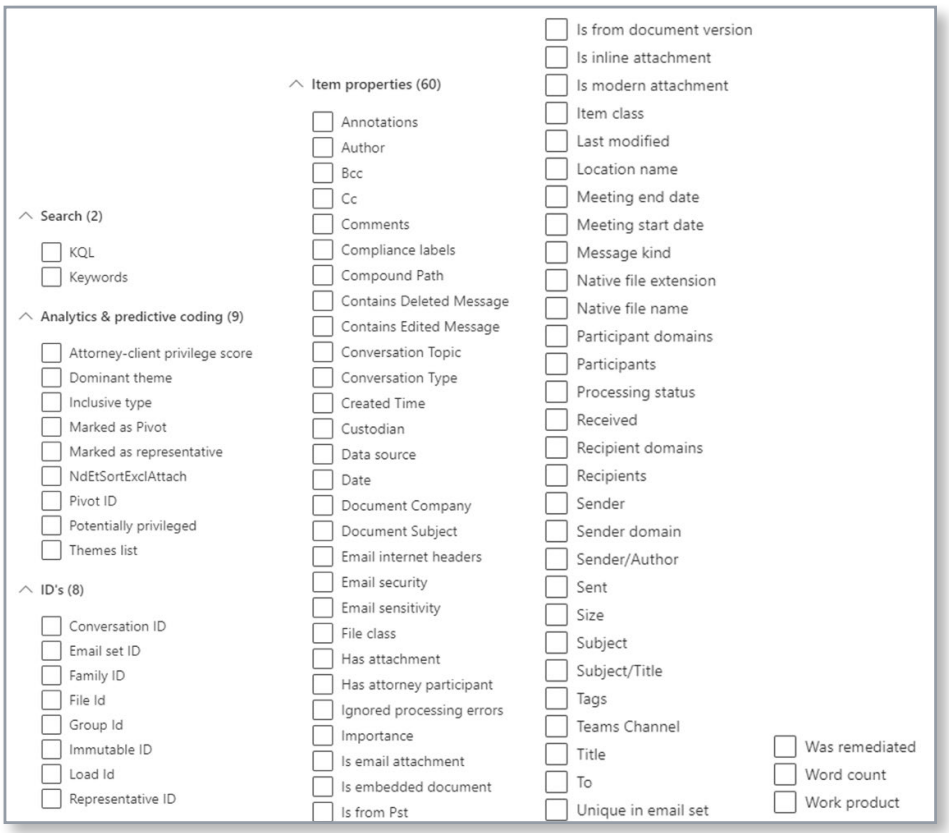


Figure 10

- Zooming into a specific use case for these features, with the "Message kind" filter selected within a review set, the following dialogue appears, with available selections for any combination of email, meetings, tasks, contacts, Microsoft Teams, and Microsoft Yammer:

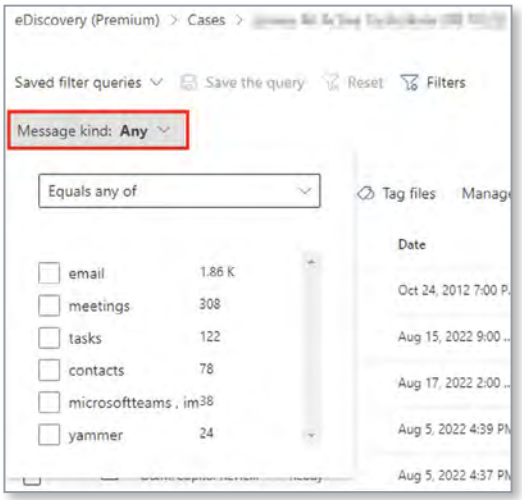


Figure 11

- Additional logic options are also provided to more precisely filter the specific content of interest:

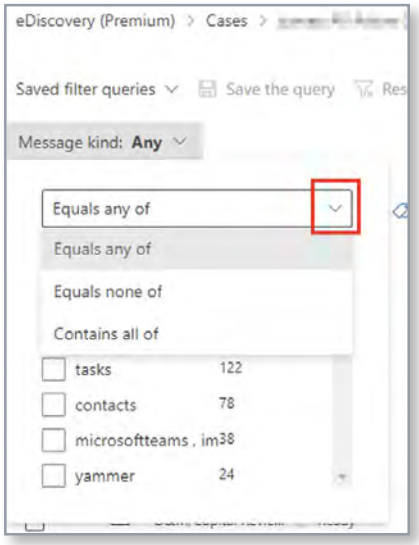


Figure 12

- Building on this concept, additional filters can be simultaneously added in order to apply multiple criteria.[5] This feature allows users to search through a complex dataset for specific information, similar to what has typically only been possible after data has been processed and indexed in a downstream eDiscovery platform. The benefit of this feature is that you can perform these targeted searches **before any data has left the organization's M365 tenant.**

- The screenshot below depicts a more advanced, multi-criteria scenario, with a keyword search for meet* OR review* OR schedule, within Exchange data only, and a date range of 8/1/2022 to 8/5/2022.

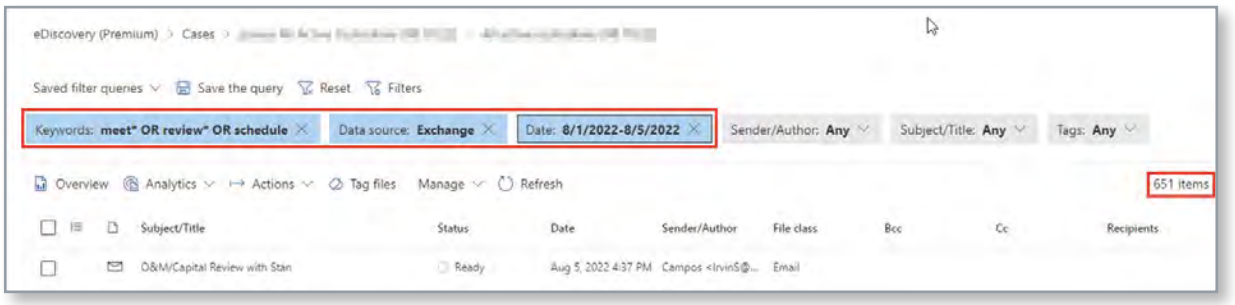


Figure 13

Email Domain Filtering

- Another capability introduced via eDiscovery (Premium)'s analytics processing is the extraction of email domain information and the resulting ability to isolate, then exclude, or include content by domain.

- Returning to our prior hypothetical of an ethics investigation, consider a set of email search results with a much higher-than-expected volume based on what is already known to be a minimal number of possible results. In such a situation, an assessment of email domains[6] could provide needed insight – perhaps in the form of a problematic, false-positive search hit associated with a specific domain.

  In such a case, let's imagine that filtering by sender domain immediately reveals obvious, false-positive participant domain hits for: amazon.com, linkedin.com, and tradejournal.com – all of which can't possibly be within the scope of the investigation.

  In this situation (in addition to the broader insight provided by easily viewing the complete list of sender domains), applying a simple "Equals none of" filter on those three domains would reduce the document count by over 4500 items in just a few clicks.

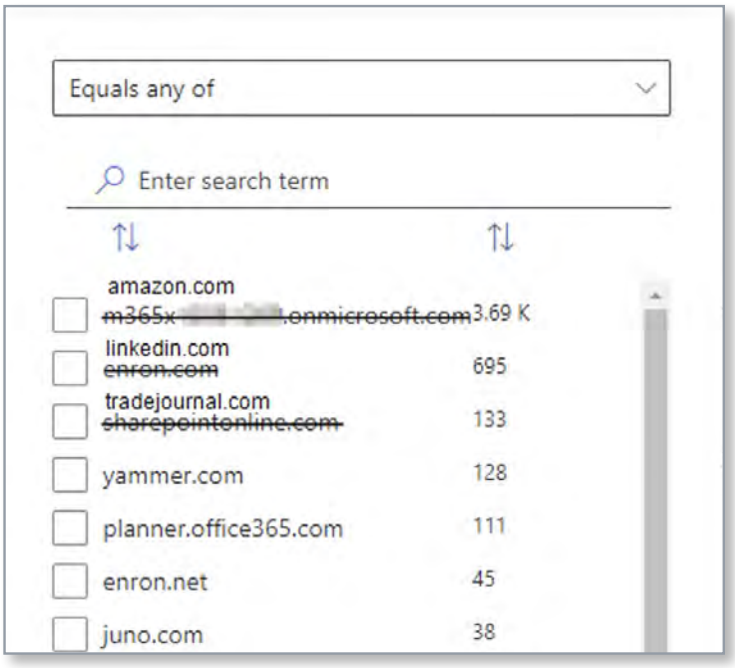


Figure 14

Keyword Search

- eDiscovery (Premium)'s keyword search capabilities[7] provide flexibility to be applied, either separately or in combination with its array of filters, to achieve precise control over the scope of a review set. It supports efficient, iterative fine-tuning of search parameters to help legal teams better identify content worthy of scrutiny by human reviewers.

- In addition to eDiscovery (Premium)'s primary keyword search capability, even more extensive capabilities are available via the KQL-based "Advanced Query Builder." KQL[8] stands for "keyword query language" and is designed to retrieve data based on the combination of keywords and conditions specified according to the KQL syntax.

- Microsoft updated eDiscovery (Premium)'s Advanced Query Builder in 2022 for an improved user experience in constructing complex queries. It allows users to "add conditions or add condition groups that are made up of multiple conditions that are logically connected by AND OR relationships."[9]

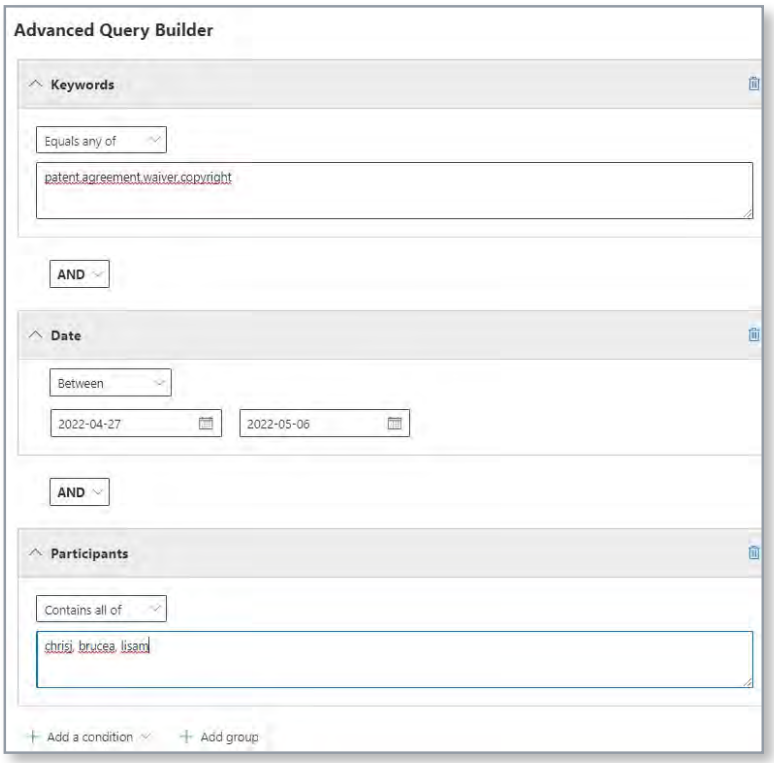


Figure 15

Reporting

- Multiple reporting options are available in eDiscovery (Premium), which facilitate the assessment of review set content and further support ECA efforts.

- The following are two different examples of the types of fixed reporting provided in the review sets tab:

Figure 16

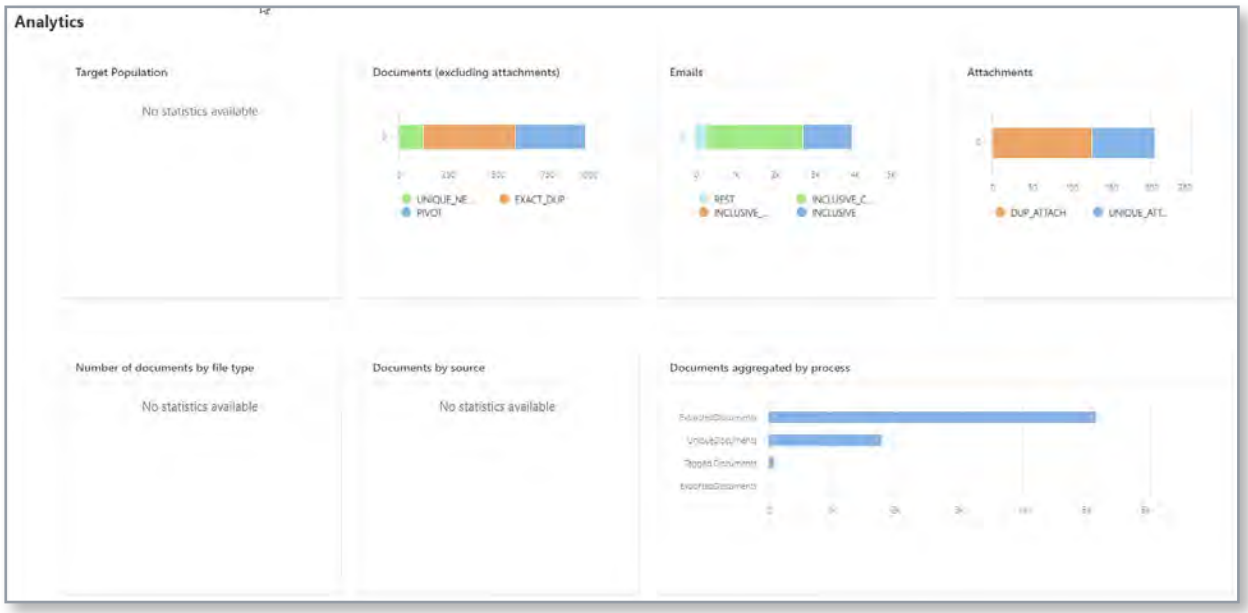


Figure 17

Search Profile View[10]

- A final aspect of eDiscovery (Premium)'s support for advanced ECA activities resides in a corner of the Review tab UI that is very easily overlooked. However, the visual analytics capabilities that lurk there are easily the most impactful of its feature set.

Combining aspects of eDiscovery (Premium)'s analytics processing and search capabilities, the "Search profile view" or Review tab "dashboard" supports dynamic visualizations based on a selection of available pivot fields, as well as filter state.

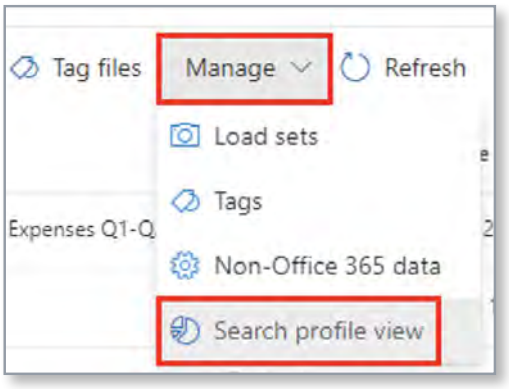


Figure 18

What this functionality provides is the ability to create visualization widgets, which represent the current result set via one of four available formats:

1. Pie/ring chart
2. Horizontal bar graph
3. Vertical bar graph/chart
4. Line graph

By default, three widgets are included:

1. Pie/ring chart on ItemClass
2. Bar chart on RecipientDomains
3. Pie/ring chart on the To field

- The following example illustrates several additional widgets, reporting results based on a limited number of extensions via the native file extension filter:

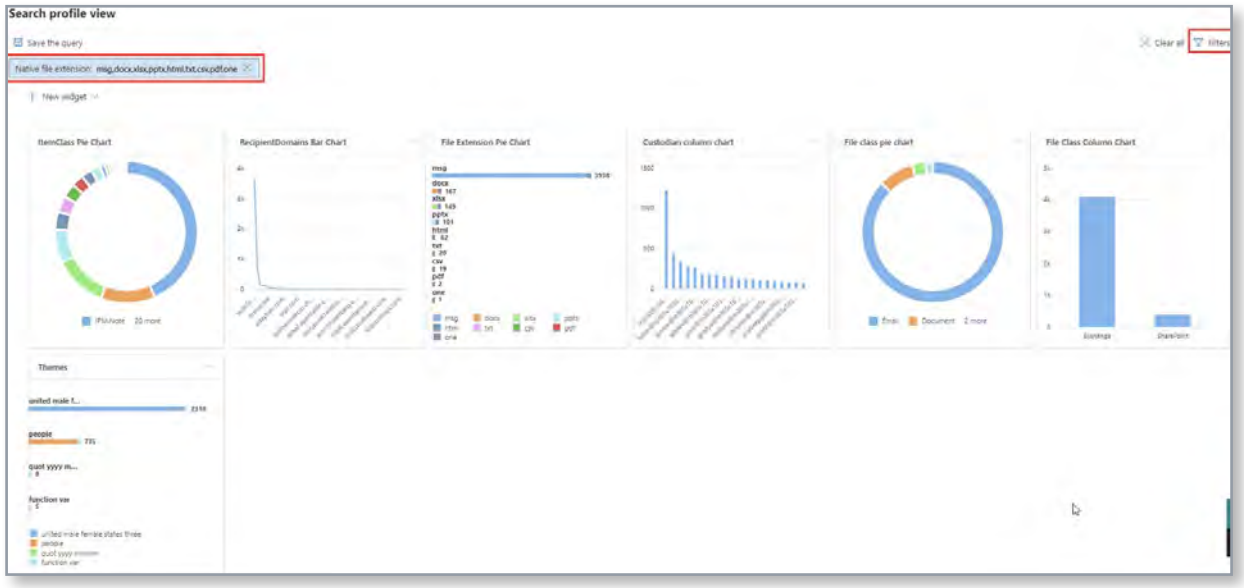


Figure 19

- In the following example, the visualizations are modified by applying two additional filters – a keyword query and a date range. The visualization widgets update to reflect the new filters and report on the new totals for each category:

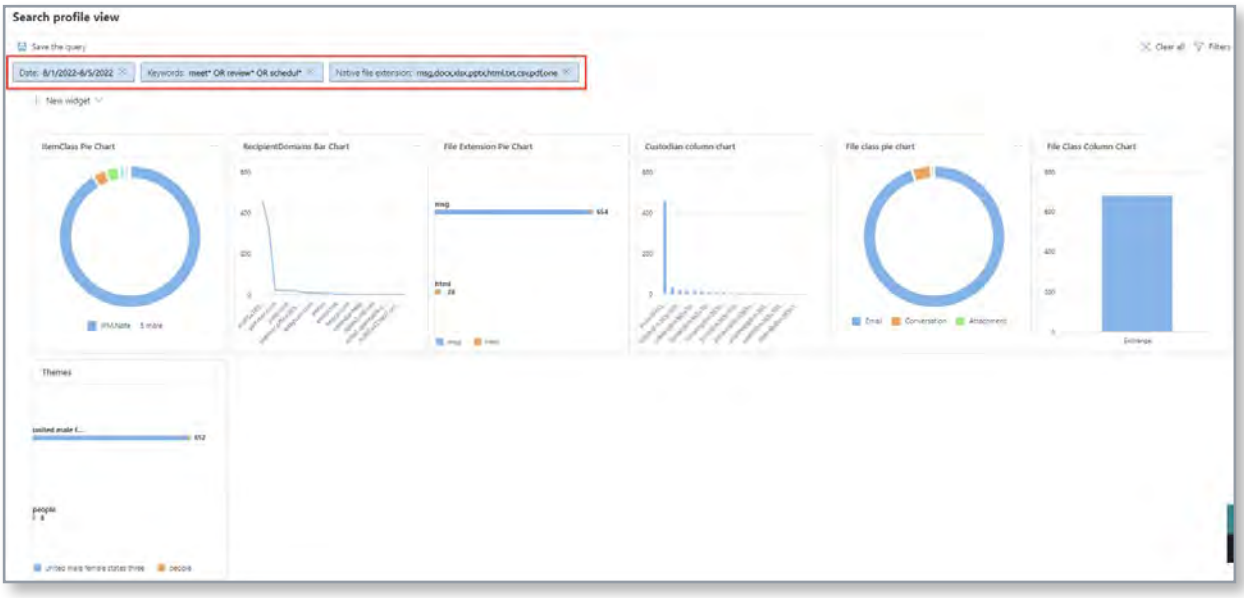


Figure 20

- The following are examples of the individual widgets, which are viewable in higher resolution form via:

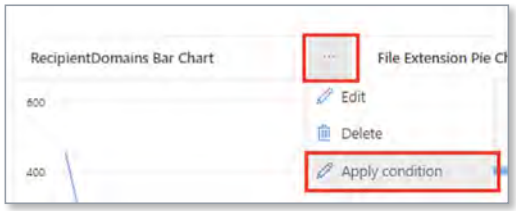


Figure 21

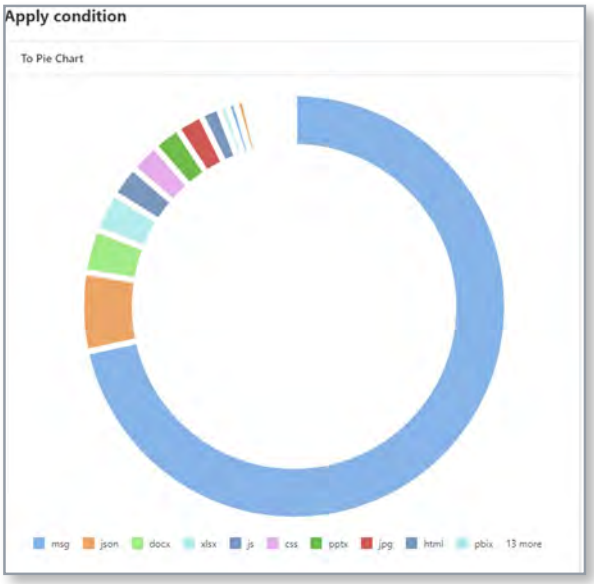


Figure 22

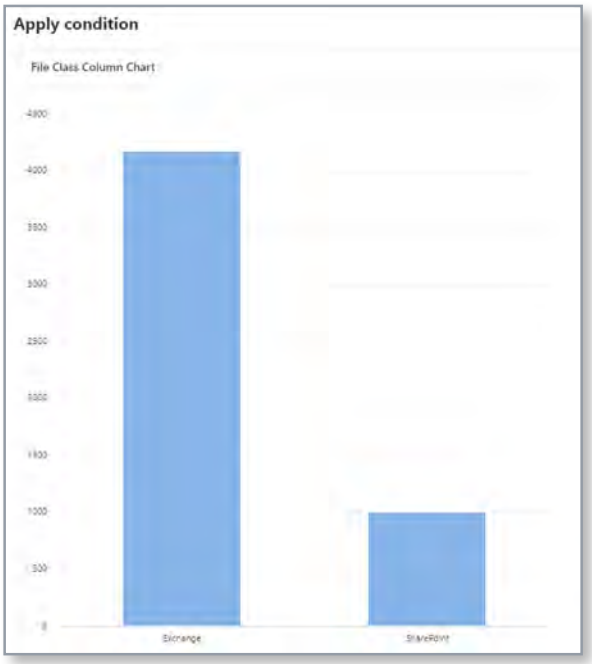


Figure 23

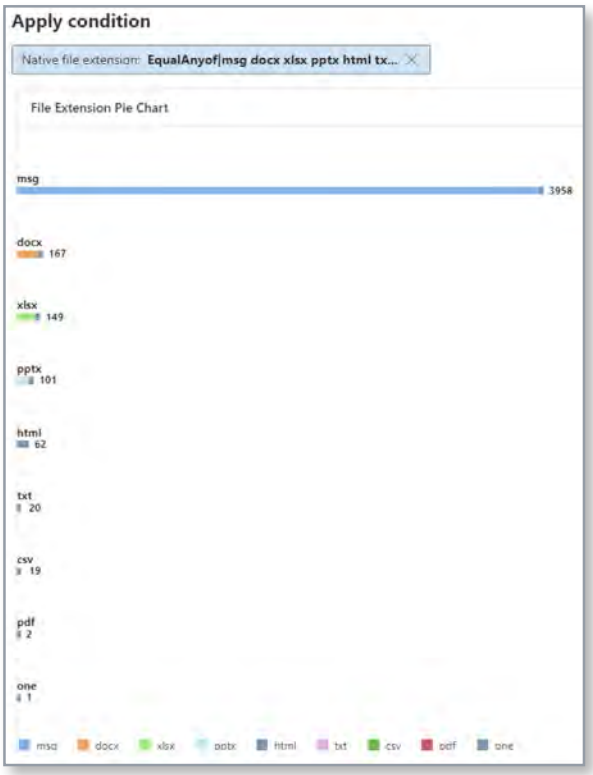


Figure 24

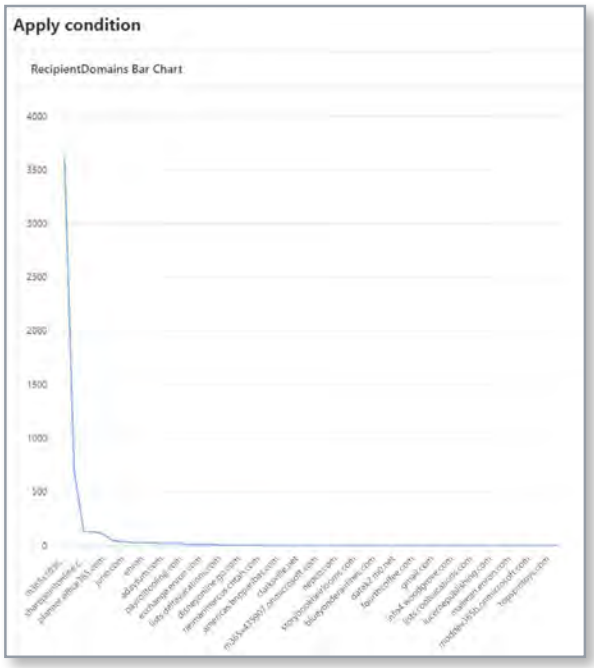


Figure 25

Conclusion

Although not a comprehensive survey of the capabilities available in eDiscovery (Premium) to support ECA workflows, the information presented should be construed as a broader framework for process design, and provide food for thought from which more specific, customized workflows can be developed.

Learn More. Today.

Contact us today to learn more about our Microsoft 365 capabilities and how we can help assess, augment, accelerate, and support your cyber, data, and legal discovery operations.

References:

[1] Also referred to Early Data Analysis or EDA.

[2] https://learn.microsoft.com/en-us/microsoft-365/compliance/overview-ediscovery-20?view=o365-worldwide. Microsoft Purview eDiscovery (Premium) is also commonly referred to as "Premium eDiscovery," "Premium" or "eDP."

[3] See the following links for a more in-depth discussion of ECA:
https://zapproved.com/blog/eca-early-case-assessment/.
https://www.exterro.com/basics-of-e-discovery/early-case-assessment.

[4] See the following Microsoft article for more information on this functionality: https://learn.microsoft.com/en-us/microsoft-365/compliance/analyzing-data-in-review-set?view=o365-worldwide#using-the-for-review-filter-query.

[5] It's important to note that, by default, filters are evaluated in eDiscovery (Premium) with Boolean AND logic applied between them when multiple filters are used, which best supports the objective of filtering content. In order to perform the most flexible and advanced searches possible in eDiscovery (Premium), the KQL (keyword query language filter) should be used, which supports Boolean OR logic, among many other parameters, and is addressed in the next section.

[6] Sender, recipient, and participant domains are all supported.

[7] https://learn.microsoft.com/en-us/microsoft-365/compliance/keyword-queries-and-search-conditions?view=o365-worldwide.

[8] https://learn.microsoft.com/en-us/microsoft-365/compliance/ediscovery-kql-editor?view=o365-worldwide.

[9] https://learn.microsoft.com/en-us/microsoft-365/compliance/review-set-search?view=o365-worldwide#advanced-query-builder.

[10] https://learn.microsoft.com/en-us/microsoft-365/compliance/advanced-ediscovery-dashboard?view=o365-worldwide.

About HaystackID®

HaystackID is a specialized eDiscovery services firm that supports law firms and corporate legal departments through its HaystackID Discovery Intelligence, HaystackID Core, and HaystackID Global Advisory offerings. In addition to increased offerings, HaystackID has expanded with five investments since 2018. Repeatedly recognized as a trusted service provider by prestigious publishers such as Chambers, Gartner, IDC MarketScape, and The National Law Journal, HaystackID implements innovative cyber discovery services, enterprise solutions, and legal discovery offerings to leading companies across North America and Europe, all while providing best-in-class customer service and prioritizing security, privacy, and integrity. For more information about its suite of services, including programs and solutions for unique legal enterprise needs, please visit HaystackID.com.

About the Author

As the lead M365 eDiscovery Consultant in HaystackID's Client Services group, Jason Covey provides a unique combination of consulting and technical expertise in operationalizing M365's toolsets for eDiscovery and investigation matters. Working closely with HaystackID's forensics group, Jason also develops and delivers training content and technical support for law firms, corporate legal departments, and IT professionals. Jason has extensive industry experience with more than 20 years of experience in law firm civil litigation, eDiscovery, and litigation support, primarily in Am Law 200 firms. Contact HaystackID's Global Advisory Practice at Info@HaystackID.com to learn more.