Office 365 eDiscovery Background and Glossary

By
John P Collins, JD | DTI, Director of Information Governance & Office 365 Consulting

Background

What is Office 365?

It is a cloud-based service providing email, collaboration, real-time messaging (IM) and other productivity tools to individuals and organizations. Office 365 falls into the Software as a Service (SaaS) category of cloud computing: the user subscribes to the service, paying Microsoft a monthly fixed fee for a bundle of features and functions, including Exchange (email), SharePoint (collaboration and file storage) and Skype for Business (real-time communications-IM, voice.)

With Office 365, the underlying IT infrastructure (servers, storage, network plumbing) resides in Microsoft's data centers; the subscribing organization's users connect to Microsoft's data center to obtain access to the various tools.

Why is Office 365 a major force in eDiscovery?

First, it's where organizations of all types and sizes and across every vertical are storing their ESI: email, documents, files—consequently it's where a significant portion of documentary evidence is likely to reside. Those responsible for executing an organization's discovery obligations will need to learn how to identify, preserve, and collect from Office 365 since it will be (for many organizations) the primary source of ESI sought in discovery.

Second, Office 365 has built-in eDiscovery tools enabling organizations to identify, preserve, preview, collect, analyze, and cull ESI residing within the service. While, most organizations move to Office 365 for business, IT, and cost reasons, the presence of the built-in eDiscovery tools should trigger an evaluation of where and how the tools might be leveraged to meet the organization's eDiscovery obligations. Experience suggests organizations can fundamentally alter—for the better—their approach to eDiscovery by employing some or all the built-in tools.

Third, for many organizations the built-in eDiscovery tools are already included in their license plan—meaning there is no acquisition cost to obtain a set of new features to perform discovery tasks. Combined with the possibility that, by employing Office 365's eDiscovery tools, existing eDiscovery tools that carry licensing or maintenance costs can be eliminated—means an organization might derive significant cost savings through adoption of Office 365's built-in eDiscovery tools.

Office 365 is where a significant portion—often the majority—of the ESI sought in discovery resides; there are tools embedded in the system to perform discovery; and, there is often no cost to acquire the tools. Furthermore, the tools can dramatically reduce downstream eDiscovery spend. These dynamics reveal Office 365 for what it is: a major force in eDiscovery.

What are the key components of Office 365 from an eDiscovery perspective?

- Built-in eDiscovery: tools to identify, preserve, preview, collect, analyze, and cull ESI residing in Office 365. The tools are available in some—but not all—Office 365 plans. From a licensing perspective,[1] the tools break down into two categories: "Standard" and "Advanced:"
    - "Standard" (sometimes referred to as "E3") tools provide the ability to:
        - Search (identify)
        - Preserve-in-place (hold)
        - Preview (early case assessment "light")
        - Export (collect)
    - "Advanced" (sometimes referred to as "E5") tools (based on the Equivio technology Microsoft acquired in January 2015) provide the following analytic capabilities for ESI in Office 365:
        - Email threading
        - New-duplicate detection
        - Predictive coding
        - Clustering

- Exchange Online: Office 365 is powered, in part, by Microsoft Exchange—the email platform in existence for 20 years providing messaging, calendaring, tasks, contact management, etc. Exchange Online is subject to Office 365's built-in eDiscovery tools.

- Mashups: these are applications combining multiple elements of Office 365 to provide unique collaborative and communications functionality to users. These applications frequently store communications, files, and other ESI that is unique and separate from repositories of ESI from any given user's mailbox or OneDrive for Business. As noted below, mashups are not always subject to Office 365's eDiscovery tools when first released; in some instances, parts of a mashup are not subject to the tools while other parts are.

    - Groups: a core collaboration tool in Office 365 and one in which Microsoft is making significant investments. Each Group can have one or more participants (internal or external parties) and includes a standard set of shared resources: calendar, mailbox, document library, and OneNote notebook (recently added to each Group is a full SharePoint Team Site.) Members of a Group ostensibly communicate via group conversations—threaded email conversations captured and residing within the Group's shared mailbox; participants in the Group can receive copies of Group conversations in their standard Exchange mailbox but this is optional. Groups ARE subject to Office 365's built-in eDiscovery tools.
    - Planner: a Group but with an additional feature for project management. Using this feature, project tasks are created, assigned, and organized into "buckets"—groups of tasks tracked throughout the lifecycle of a project. Buckets are not (as of August 2017) subject to Office 365's built-in eDiscovery tools, but the features Planner shares with Groups ARE.
    - Teams: Microsoft describes Teams as "a chat-centered workspace...it brings conversations, files, and tools into one place so everyone has instant access to everything they need."[2] Teams has also been described as tool for "high-velocity" collaboration—situations where collaborators need and want to share information instantly. One of the central features of

[1] https://support.office.com/en-US/article/eDiscovery-FAQ-9d1a29ae-b7b4-4a27-9c8c-84289023dcae#q5
[2] https://support.office.com/en-US/article/Microsoft-Teams-Quick-Start-422bf3aa-9ae8-46f1-83a2-e65720e1a34d

Teams is the chat function—the ability to initiate and carry on conversations over an extended period. Teams also includes a Group—so each Team by default has a document library, OneNote, calendar, etc. It is useful to note Teams is a "millennial friendly" application—users can insert emoticons, Giphy, and other non-text artifacts into their communications.

- Office ProPlus: a subscription based model for acquiring the Office applications (Word, Excel, PowerPoint, Outlook); the applications are downloaded, via the internet, to the user's device. Included in some Office 365 plans (E3 and E5 for example), ProPlus allows users to download Office on to a maximum of 15 devices (5 tablets, 5 phones, 5 computers.) Implication: users storing their email in Outlook on 15 different devices. What happens when a user stops syncing one of their devices to Office 365? You have a unique island of email representing a snapshot in time. ProPlus can be managed to require the user to use only an approved device to download the software. Legal and eDiscovery teams should discuss ProPlus with their IT counterparts.

- OneDrive for Business (ODB): an online file and document storage repository—similar to Box and DropBox—assigned to each user in Office 365 (each user gets his or her own.) The concept is to have OneDrive for Business replace local PC ("My Documents" for example) and home directory/share (file server) as a user's primary file storage locations. ODB enables users to share files or folders with internal and (if permitted) external users. (NOTE: ODB comes with 1 terabyte of storage, and is a technically a SharePoint site.) The ESI stored in ODB IS subject to Office 365's built-in eDiscovery tools.

- SharePoint Online: SharePoint Online (like Exchange Online) is a fundamental building block upon with many of the Office 365 features and tools are built. Standing alone, SharePoint provides collaboration, intranet, extranet, application development, file sharing, document and records management capabilities. As a fundamental component of Office 365, elements of SharePoint power other features and tools—most notably Groups, Teams, and Planner. The ESI stored and created via SharePoint Online IS subject to Office 365's built-in eDiscovery tools.

- Skype for Business: a platform providing instant messaging, online meetings, VoIP, video chat, etc. In addition, Skype for Business can replace traditional on-premise PBX (phone systems.) IM and call logs generated by Skype for Business ARE subject to Office 365's built-in eDiscovery tools.

- Sway: a web-based tool that allows users to create presentations called "Sways" and share them (via links) with internal or external parties. There is no Sway file per se; rather, Sway's exist solely as a cloud artifact. Sway has significant adoption in the educational markets, but less so in corporations. As of 8/2017, Sway is NOT subject to Office 365's built-in eDiscovery tools, so preservation and collection requires use of forensic tools and techniques similar to those used for social media and other web-based repositories.

- Yammer: social media platform for the enterprise, providing a "Facebook-like" collaboration experience. Users can create groups, carry on conversations, "like" posts, upload files and pictures. As of 8/2017, Yammer ESI is NOT subject to Office 365's built-in eDiscovery tools. Also, new options are being introduced to create Yammer Groups, which combine elements of Office 365 Groups with Yammer functionality.

Considerations When Your Organization or Client Moves to Office 365

1. What licenses does the organization own?
    a. Business Essentials
    b. Business
    c. Business Premium
    d. K1
    e. Education or Enterprise E1
    f. Government
    g. E3
    h. E4
    i. E5
2. Which of the following O365 services is being deployed by the organization:
    a. Exchange Online
    b. SharePoint Online
    c. Skype for Business
        i. Have you discussed "Conversation History" status (whether to retain/archive IM chats)?
    d. O365 Groups
    e. Teams
    f. Planner
    g. Yammer (NOTE: as of 8/2017 Yammer ESI is not subject to O365's built-in eDiscovery tools)
3. Is the organization migrating ESI from legacy repositories as follows:
    a. Current on-premise email to Exchange Online
    b. Current on-premise SharePoint to SharePoint Online
    c. Individual user's PC ESI to OneDrive for Business
    d. Individual user's "home-share" ESI to OneDrive for Business
    e. Individual user's PST files to Exchange Online
4. Who will be granted eDiscovery permissions in your tenant?
    a. eDiscovery Manager=
    b. eDiscovery Administrator=
    c. Custom=
5. Will you leverage the following eDiscovery tools:
    a. eDiscovery search
    b. Preservation (preserve ESI in place)
    c. Preview/Early Case Assessment
    d. Collection/Export
6. Can you eliminate licensing costs for other tools by leveraging O365's built-in eDiscovery tools?
7. What is your organization's annual spend for:
    a. eDiscovery analytics (email threading, near-duping)?
    b. Technology Assisted Review?
    c. eDiscovery processing?
8. Do you have the desire to move away from 3rd party email archiving and journaling to meet preservation needs?
9. Do you want to move away from custodian self-preservation?
10. Is there a need to set up clearly identified and enforced barriers between USA and EU based eDiscovery permissions (for example, prevent USA based staff from accessing EU staff ESI.)
11. Have you identified a PowerShell resource?

12. Do you have an audit and reporting strategy for using eDiscovery tools in O365?
13. Is the legal department connected with/working with IT and the business on O365 adoption and roll-out to ensure a proper legal (eDiscovery and RIM oriented) framework is in place?
14. Is your organization planning on deploying O365's built-in Information Governance (referred to by Microsoft as "Data Governance" tools)?
15. Have you received adequate training on O365's built-in eDiscovery and Information Governance tools, especially regarding limitations and caveats:
    a. Indexing
    b. Exports
    c. Change management
16. Is Office ProPlus deployed?
    a. If yes, are users permitted to download Office to non-domain joined devices?

Glossary

Content Search (in Security & Compliance Center): one of the several eDiscovery tools available in Office 365, Content Search allows a user with proper permissions to search one or more mailboxes, SharePoint sites, ODB, Groups, and Exchange Public Folders. After executing a search, the user may 1) preview the search results; 2) export the search results; 3) analyze the search results using O365 Equivio

Conversation History: a folder which is automatically created in user's Outlook mailbox that archives IM chats and call logs (for calls made via the Skype client client.) The ability to retain IM and call logs in the conversation history folder can be turned on and off—at either the user or organization level. Some organizations leave it up to the user. If conversation history is enabled and IM and call logs are captured, they are subject to Office 365's built-in eDiscovery tools.

Data Governance: a set of tools to retain and dispose of ESI across the various workloads and repositories in O365. There are two primary components to Data Governance: retention policies and labels. Retention policies can target certain repositories and users—for example, all mailboxes and OneDrive for Business sites—and apply specific retention and disposition requirements. Retention policies can be triggered based on when an item was created, sent/received, or by the presence of certain keywords, phrases, or sensitive information. Labels can be applied to specific items—such as an email or a document, and like retention policies, can be based on when an item was created, sent/received, or by the presence of certain keywords, phrases, or sensitive information. Labels add an additional trigger, which is when the item is labeled. Retention policies and labels can be combined to form a complex retention and disposition framework.

DLP (Data Loss Prevention): provides the ability to set up rules that trigger actions when certain types of content (sensitive, confidential, personal) is found in certain locations in Office 365 (Exchange and SharePoint primarily.) For example, if a user stores a spreadsheet with 25 credit card numbers in OneDrive for Business, an alert can be sent to a legal or a compliance officer; alternatively, if the user tries to email the same spreadsheet, the email itself can be blocked, encrypted, forwarded to a compliance officer, sent back to the user, or just deleted (with or without notifying the sender of these actions.)

Document Deletion Policies: broad-brush retention policies that can be applied to SharePoint and OneDrive for Business document libraries. If a document deletion policy is in effect, files are deleted a certain number of days after EITHER the file's created or last modified date. These policies can be mandatory or optional; users may also be given the option to choose from among several policies. This feature is being deprecated in favor of the new data governance features.

eDiscovery (in Security & Compliance Center): one of several eDiscovery tools available in Office 365. Those with appropriate permissions can 1) search one or more mailboxes, SharePoint sites, ODB, Groups, Teams, and Exchange Public Folders; 2) can place ESI residing in the locations searched on in-place hold; 3) can "preview" emails, documents, and other files; 4) can export the ESI out for further review in down-stream review tools. These activities are all done without disrupting the end-users/custodians.

E3: refers to one of the O365 licensing plans. E3 licensing includes the "Standard" Office 365 built-in eDiscovery tools (identification/search, preservation/hold, preview, and collection/export.)

E5: refers to one of the O365 licensing plans. E5 includes all the features of E3 plus additional functionality for eDiscovery commonly referred to as "Advanced eDiscovery." Advanced eDiscovery includes the former Equivio suite of tools for structured analytics (near-duping, email threading, and clustering) and predictive coding.

Hybrid: denotes an Office 365 environment where a portion of the organization's Exchange, SharePoint, and Skype for Business environment remains "on-prem" (the infrastructure, i.e., servers, data storage, etc. remain in the organization's own data centers.) For organizations with more than several hundred email users they will likely be in a hybrid mode for a period (typically for the duration of mailbox migration.) Some organizations elect to stay in hybrid mode on a permanent or extended basis. eDiscovery in a hybrid environment requires integrations between cloud and on-prem components, and has a number of limitations[3].

In-Place Archive: Microsoft's terminology for what is in essence an expansion of a user's primary mailbox, it provides additional email storage capacity—in some plans, its unlimited storage. Items can be moved systematically from a user's primary mailbox to the archive. The In-Place Archive has to be turned on for a user (by default it is NOT enabled.) Office 365's built-in eDiscovery tools can search, preserve, and collect from the In-Place Archive.

Labels: labels are a data governance tool. An administrator sets up various labels based on the organization's retention and disposition requirements and makes them available for application with the tenant. Labels can be applied manually or automatically. Labels are typically applied to individual items, but may also be applied to all items uploaded or created in a SharePoint document library. Like retention policies, labels can be based on when an item was created, sent/received, or by the presence of certain keywords, phrases, or sensitive information, however, labels add an additional trigger, which is when the item is labeled.

Messaging Records Management (MRM): legacy (they are being deprecated in favor of the new data governance features) tools for email retention and disposition in Exchange and Outlook. MRM enables an administrator to set up policies which determine how long email (and other items) are retained in a user's mailbox. For example, a policy could be configured wherein all emails in a given employee's mailbox are automatically purged after 180 days EXCEPT items "tagged" by the employee (or via a rule.) In this example, users could be provided tags for 1 year, 3 year, and permanent retention; any items users apply one of these tags to will be retained for that period of time and thus not subject to being purged after 180 days. Multiple policies can be created to meet the needs of different departments and user groups.

PowerShell: a command line scripting language that underlies many Microsoft technologies, including Office 365. PowerShell allows for the automation of tasks via its command-line interface, and is

[3] https://technet.microsoft.com/en-us/library/dn497703(v=exchg.150).aspx

particularly helpful in automating execution and reporting of eDiscovery activities in Office 365. Organizations seeking to employ Office 365's eDiscovery tools on a large-scale basis typically employ some level of PowerShell scripting and automation.

Retention Policies: retention policies can target certain repositories and users—for example, all mailboxes and OneDrive for Business sites—and apply specific retention and disposition rules. For example: keep all items in OneDrive for Business, SharePoint, and Teams for 5 years after the last modified date. Retention policies can be triggered based on when an item was created, sent/received, or by the presence of certain keywords, phrases, or sensitive information. For example: delete any item with the phrase "for training purposes only" 3 years after creation date.

Security & Compliance Center: Office 365 administrative interface from which users with appropriate permissions can access tools for eDiscovery, information security, retention, and other related compliance features.