

Information Governance

Deployment Acceleration Guide

Prepared by the Microsoft 365 Security + Compliance engineering team

Updated: January 2020

This document is provided “as-is”. Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. You bear the risk of using it.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

© 2019 Microsoft. All rights reserved.

Contents

[Advanced Data Governance Deployment Acceleration Guide 7](#_Toc25776463)

[Overview 7](#_Toc25776464)

[Getting Started 10](#_Toc25776465)

[General Requirements 11](#_Toc25776466)

[Import 13](#_Toc25776467)

[Govern 15](#_Toc25776468)

[Labels with retention and deletion policies 15](#_Toc25776469)

[Auto-apply labels to specific types of content 15](#_Toc25776470)

[Label Activity Explorer 15](#_Toc25776471)

[Content Explorer 15](#_Toc25776472)

[Comprehensive Audit 15](#_Toc25776473)

[Summary and best practices 16](#_Toc25776474)

[Import 16](#_Toc25776475)

[Govern 16](#_Toc25776476)

[Classify and apply policy 16](#_Toc25776477)

[Monitor 16](#_Toc25776478)

[Appendix 16](#_Toc25776479)

[Useful Links 16](#_Toc25776480)

[Additional Information 16](#_Toc25776481)

# Advanced Data Governance Deployment Acceleration Guide

## Overview

Information and its governance are an asset to every organization’s business performance. The quality of Business decisions and competencies are directly related to effective use and exchange of information. The same information is also subject to compliance laws and regulations and hence carries risks and obligations to manage information effectively. Therefore, Information technology tools used to manage information lifecycle require adequate security and compliance measures to Classify, Govern and Monitor the information they Create or Import.

Microsoft Information Governance is a cloud-based solution that enables organizations to manage compliance across enterprise information by ingesting or creating content into the Microsoft cloud, application of labels and associated policies to manage information lifecycle over time. Organizations of all sizes use Information Governance to keep items of information which must be retained, in their original form, for a defined period to meet the legal & regulatory or business requirements. Any information identified as a Record, cannot be altered or disposed of before the defined period has expired. This is done using organization defined schedule of policies and labels published to approved locations in Microsoft 365, as well as the auto-application of labels using sensitive information types, keywords or custom information types defined by the organization.

# Getting Started

With the general availability of the new Microsoft 365 Compliance Center[[1]](#footnote-1), customers now have a clear recipe to deploy and manage Information Governance suite of features across their Microsoft services.

A typical deployment consists of a ‘Compliance Program’ with a business focused roadmap for adoption of Microsoft 365 information governance (IG) driven practices and procedures.

Compliance Standards & Guidelines

Business Priorities

Compliance Priorities

IG deployment strategy

IG deployment roadmap

Business change Design

IG deployment Design

Business Change & adoption

IG deployment

Govern + Improve

Import

Classify

Compliance Implementation Program

IG Guide

IG solution blueprint

As seen above, the IG deployment is part of a wider compliance adoption program instead of being the program itself, and this highlights that this is a business change led initiative.

The key deliverables as per the above template are typically as follows:

* Information governance policies and guidelines
* Program business case and charter
* Program business and compliance requirement specifications
* Information governance solution strategy and blueprint
* Deployment plan, timeline and milestones
* End user adoption and knowledge management
* Technical support arrangements to monitor and report compliance driven KPI

Whilst the above template is meant to help guide your Information Governance adoption journey, the value and relevance of the template will vary subject to the size, complexity and specific context to the compliance program within your organization.

## General Requirements

For IG deployment to be successful, it must be aligned with both business and compliance priorities. Information may be created, imported (received), managed and disposed in many ways and integrated with business processes.

The Information Governance deployment guide will focus on the following key information lifecycle stages and provide a compliant toolkit for your staff to achieve effective compliance and value out of information.

The full list of Information Governance requirements can be found as part of our official documentation[[2]](#footnote-2). We have summarized some of the more important requirements in the Appendix that are essential for all deployments of Information Governance. If you have complex governance rules, and existing retention hierarchies we recommend reading the full documentation prior to deploying.

# Advanced Data Governance Deployment Guide

### Create or Import

One of the key drivers of a move to the cloud and digital transformation is an opportunity to take account of what digital assets and liabilities exist in your modern workplace.

A key consideration / challenge when mobilizing a compliance governance model in Office 365 is management of legacy data. There are a number of different directions an organization could pursue, Microsoft generally see three scenarios that customers want to pursue as they work to establish their information management strategy in the cloud:

1. Migrate (Lift and Shift) data for management in Microsoft Cloud
2. Maintain legacy data in place on-prem, 3rd party or other content repositories
3. Hybrid approach with some critical data Keeping in place, and majority moving to the cloud

As regards to this document, we will focus on scenarios (considerations and best practices) related to migrating and managing information governance in Microsoft 365.

In this case many organizations would be like to take the opportunity of a transition to the cloud, to identify critical data and bring only that data into the cloud, and to responsibility delete other data that has exceeded its useful life and is now considered redundant, trivial or obsolete (ROT) for the business. A transition to the cloud provides a great excuse for taking action on this data that may have been accruing over a long period of time without a particular purpose or governance strategy. There are capabilities available from Microsoft, such as the Azure Information Protection scanner, that can help you scan your non-Office 365 locations for sensitive content. Then you can use various import options to bring only the most critical data into the cloud.

Organizations can import data through a variety of methods into the Microsoft cloud for ongoing management.

1. Network upload provides the capability to upload PST files over the network to a temporary Azure storage location, and then use the Office 365 import service to import the PST data to mailboxes in your Office 365 organization.
2. Drive shipping – Copy the PST files to a BitLocker encrypted hard drive and then physically ship the drive to Microsoft.

Once you have gotten your data to Microsoft you can filter upon import to make sure that you only bring in the most relevant and valuable content. Filter based on data age, type and users to help make sure you are importing the right data. Read more about how to accomplish this in our technical documentation[[3]](#footnote-3)

Whatever the strategy for information governance and consolidation of data sources, understanding the broad landscape of enterprise data and intended management of that data is a critical first step to implementing a comprehensive data governance strategy.

### Classify

Once you have established a strategy for importing data into Office 365, you are ready to assess the information landscape and develop an information governance strategy. As most organizations start their information management journey, they want to ensure that important information of legal, regulatory or business value is kept and maintained as per a defined information governance schedule. Also equally important is to manage risk (e.g. data privacy) by deleting data that has exceeded its useful business life.

With the ever-increasing volumes of information ingested and created on a regular basis, the challenge organization face are regards to:

* Identify information (legacy and new) of value, and apply repeatable information lifecycle behaviors to manage the corporate knowledge and unlock business value
* Delete redundant data (e.g. personal data when no longer required) and reduce risks towards regulatory risks,
* Identify information that must be retained, in its original form, for a mandated period because of legal, regulatory purposes; or because they are evidence of a business activity or decision

These requirements ensure that records are retained (per defined policy), immutable, and destroyed in line with the legal requirements after the associated records retention period has expired.

The following section provides high level steps to develop and deploy a policy-based framework to identify and classify information of business value, legal and regulatory Records, and information which should be deleted.

* Step1: Create a schedule of Labels (File Plan)

Labels define compliance actions and duration (disposition date), and when used for classifying content managed in M365 applies a consistent set of governance and disposal rules.

Labels (or Retention schedules) are defined centrally in the Compliance Center for the entire organization, but often (multi geo organizations) includes exceptions / overrides to meet the compliance needs of country specific regulations.

Once you have determined the Labels applicable to your organization / departments, Microsoft recommends using the ‘File Plan[[4]](#footnote-4)’ capability to import Labels and associated configurations in Security and Compliance Centre.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Lightbulb and gear | Leverage ‘Description for users’ field to provide ‘just in time’ guidance to users in identifying the correct Label to select for specific information.  This description appears as a pop-up message when user ‘hovers’ mouse over the Label.   |  |  | | --- | --- | | Label Name | Compliance Auditing | | Description for Users | Use for internal and external compliance records including assurance, audit, assessment and training records. | |

Once your Labels are imported, you can make further changes using the ‘File Plan’ manager and prepare for deployment.

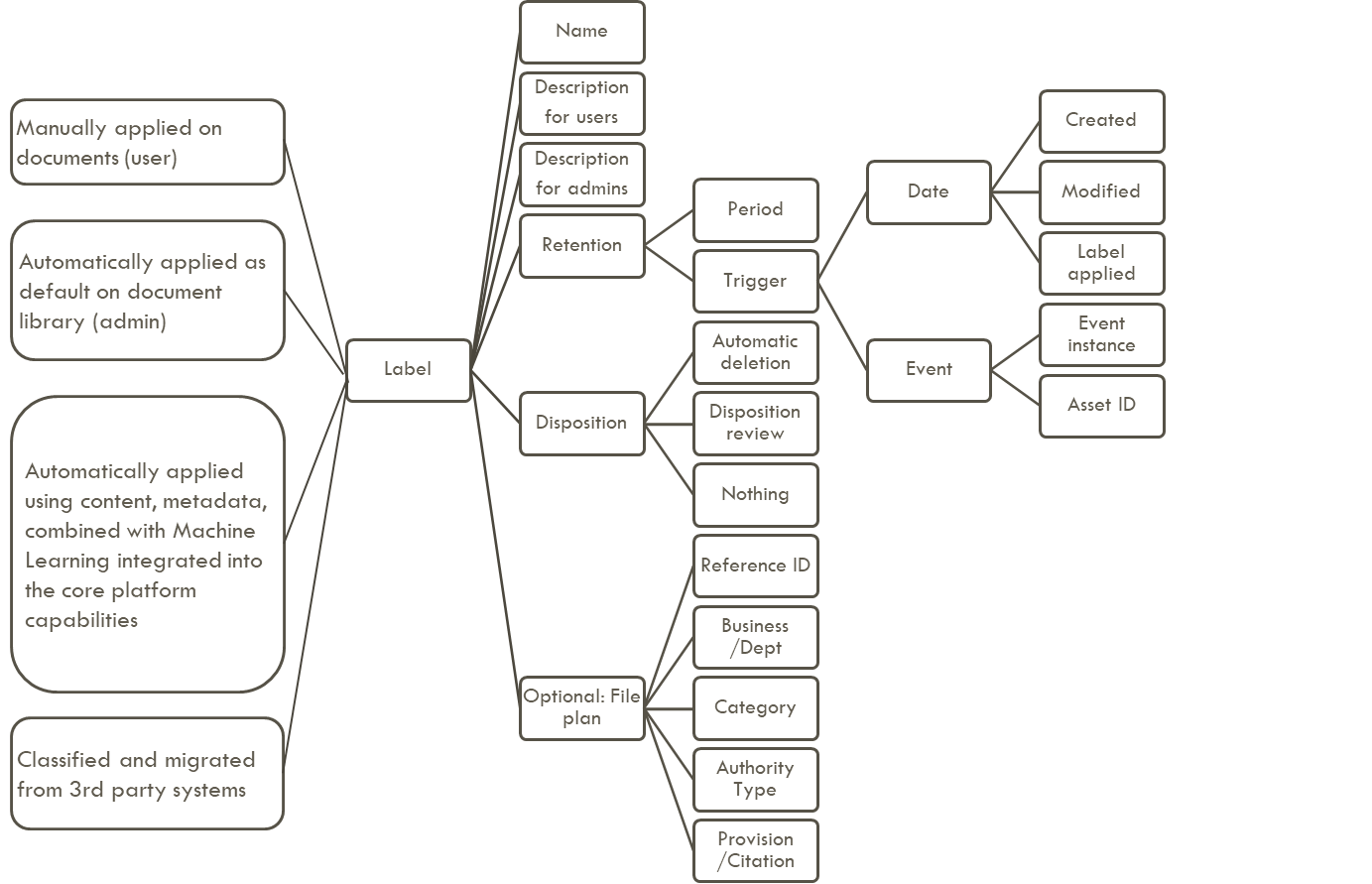
* Step2: Deploy Labels (via Label Policies)

Once retention labels are created, they can be deployed [[5]](#footnote-5)using the following options:

* + Label policies configured for specific workloads (SharePoint, Exchange Online etc.) / collections (SharePoint Site Collection, Exchange Mailbox)
  + Apply Labels automatically to content matching specific ‘keywords’ or ‘Keyword Query Language (KQL) based conditions
  + Other customized options using Power Shell scripts[[6]](#footnote-6) to automate classification of legacy / imported information

|  |  |
| --- | --- |
| Lightbulb and gear | It’s important to identify which Labels are published to each M365 collection, this requires collaboration with business teams to identify purpose of each collection and deploy relevant Label Policies. This should be a clear business engagement activity in every adoption program. |
| Lightbulb and gear | Increasingly organizations have chosen to publish an organization wide default deletion policy (e.g. Delete if not modified for 3 years), and complement it with publishing a set of retention / record labels that provides options for longer retention for important content.  This is a powerful tool to promote good information management practices and enable users to keep important information and let automated policies periodically prune content that has exceeded its useful life.  Microsoft roadmap includes empowering the automation capability with machine learning and cater for advanced scenarios involved in identification and management of information retention and disposition. |

Label is the core component to drive records management behaviors when applied to files. The diagram below provides an overview of the Label design



Most organizations are leveraging compliance programs to drive the big opportunity to reduce end user effort in managing information. This is likely to bring initial changes in ‘ways of working’ for many users, hence an important aspect of preparing End users for this change is to develop easily relatable set of guidance and training. The End user guidance should drive the benefits beyond just compliance to include the efficiency gains led by ability to identify and classify useful information.

* Classify (Label) information

Remember the labels you published to end users earlier in the deployment? If the label is configured as ‘record’ following behaviors will apply to the labelled file:

* + Files cannot be edited
  + Files cannot be deleted
  + File metadata (applied through content type) can be edited
  + Record label cannot be changed (Site collection admin permissions required)

Users can use following options to apply label and associate retention schedule to each information.

Manual application

Manually select single / multiple files / folder / document set and apply one of the published labels.

apply default labels to library

Users with ‘Site owner’ permissions in SharePoint online sites can configure libraries to apply a default label. The option is available in library settings ‘Apply label to items in this list or library’

The configuration includes following options:

* + Ability to select a label from the list published via ‘Label Policy’ to the site
  + Apply label to existing items in the library
    - If ‘not selected’ the selected label will only apply to newly uploaded / created files in the library

Auto-apply labels to specific types of content

Now we can empower the end user, reduce their burden of manually disposing information, and ease the mind of the compliance and risk teams about relying on end-user classification to meet compliance requirements with labels that are applied via auto labeling.

You can then create an auto-application policy to go ahead and apply the correct labels to the correct content based on sensitive information type or specific keyword or query.

Govern

Now, not all data is created equal, and organization admins are not always confident relying on their end-user’s ability to accurately classify content with the correct retention and deletion policies. We know that certain types of data require much more careful handling and management than other types. For example, personally identifiable information (PII) has specific requirements for management and handling, and with the advent of GDPR many organizations are moving very quickly to get a handle on managing this sensitive data. Another common high-risk data type is financial data. Organizations are focused on very quickly getting a handle on where this data is located and acting on this data.

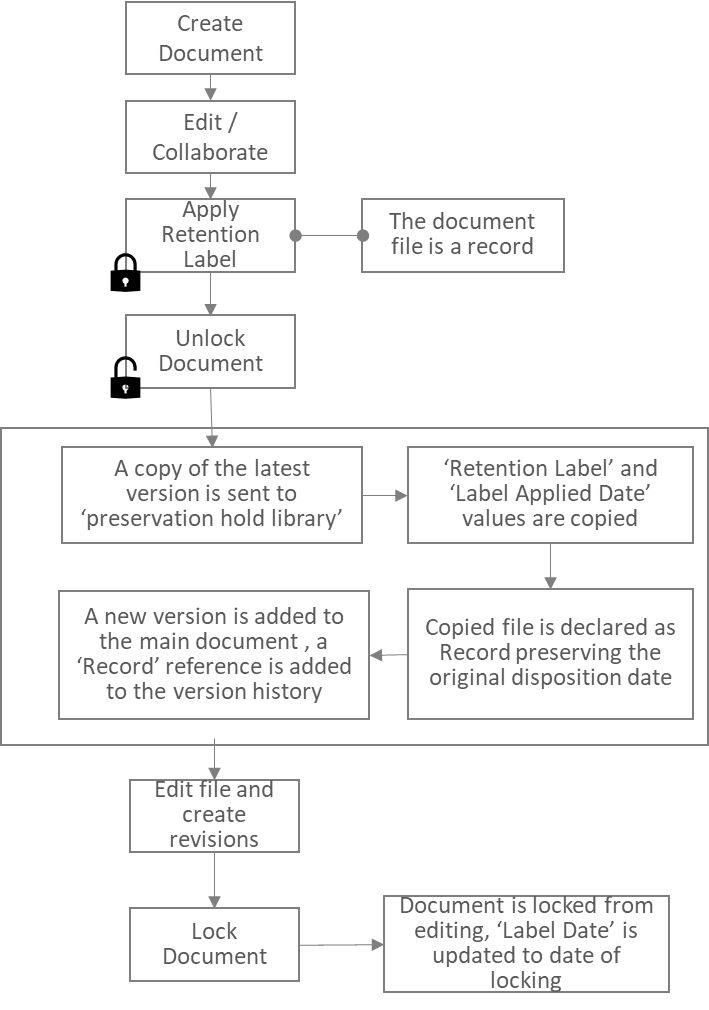
This section will cover available capabilities which can be readily used or extended to provide an effective governance of classified information (especially managing Records).

Creating revisions to files labelled as records in SharePoint online

It is possible for Microsoft E5 customers to leverage ‘Advanced versioning’ feature to create new revisions of files declared as records. This simplifies user interaction with records, with following benefits:

* + Multiple records can be declared from a single file
  + Records declared remain immutable, and retention integrity is maintained as new revisions are created
  + Original hyperlinks to file remain unchanged
  + A single file providing access to all document versions and record revisions

The system workflow for managing revisions of file with a record label is as follows:



When the file is locked, the latest version of the document is a record. When the file is unlocked, a reference to this version is added to the version history. You can check which version is a record by looking for ‘Record’ in the ‘Comments’. Subsequent edits to the file, when it is unlocked, creates versions which are not declared as record. To declare the latest version, you will need to lock the file, then it will become a record.

Any file which has multiple record versions will only be disposed after all the records versions are disposed. Such files are not subject to non-record deletion.

An example of a file through its records revisions lifecycle is described below:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Version** | **Action** | **Date** | **1 June ‘19** | **3 June ‘19** | **3 June ‘22** | **3 June ‘25** | **4 June ‘29** | **2 June ‘32** | **3 June ‘35** | **3 June ‘39** |
| **8** | **Lock to create new version** | 4 June ‘29 |  |  |  |  | 10 | Keep | Keep | Garbage |
| **7** | **Edit** | 1 June ‘29 |  |  |  |  | Keep | Keep | Keep | Garbage |
| **6** | **Unlock** | 11 June ‘22 |  |  |  | Keep | Keep | Garbage |  |  |
| **5** | **Lock to create new version** | 3 June ‘22 |  |  | 10 | Keep | Keep | Garbage |  |  |
| **4** | **Edit** | 1 June ‘22 |  |  | Keep | Keep | Keep | Keep | Keep | Garbage |
| **3** | **Unlock** | 5 June ‘19 |  |  | Keep | Keep | Garbage |  |  |  |
| **2** | **Label as a Record  (10 year retention label)** | 3 June ‘19 |  | 10 | Keep | Keep | Garbage |  |  |  |
| **1** | **Create File** | 1 June ‘19 | Keep | Keep | Keep | Keep | Keep | Keep | Keep | Garbage |

|  |  |
| --- | --- |
| Lightbulb and gear | It is important to note that the document will only be deleted once all record versions in its history have reached the end of their retention period. |
| Lightbulb and gear | A label can be selected only for the labelling of the first version to be declared. Subsequent versions, will use the label applied to the previous version.​  When locking and unlocking a document, the same record type will be applied to the record versions. ​  When a record version reaches its retention period, it will be permanently deleted. A Parent document will only be deleted once all record versions in its history have reached the end of their retention period and any more recent non-record versions are older than 3 years. ​ Each record version has its own Record retention period and will be disposed of accordingly.  ​  The parent file once labelled will be managed by the label policy throughout its lifecycle and will supersede any other policies published directly to the file or its container. |

Changing labels on files

A label can be changed or removed only by a System administrator (Site collection administrator), and it is recommended that system administrator role is assigned to centrally managed and authorized supporting roles. All label related activities are audited, this includes label change activities on each files which can be viewed either in ‘Label activity explorer’ or extracted for custom reporting using Office 365 management APIs.

Label Activity Explorer

The label activity explorer helps you investigate and validate labeling activity. One of the most important components of any governance strategy is measurement, insights and adjustment. Within Microsoft 365 there are several controls and features provided to help you investigate, validate and take action on insights derived from various monitoring capabilities. These insights help ensure that your content is governed according to policy and help triage any suspicious or questionable activity taking place in your environment.

## Content Explorer

From within the file plan or label experience quickly get a sense for what content has the specific label of interest.

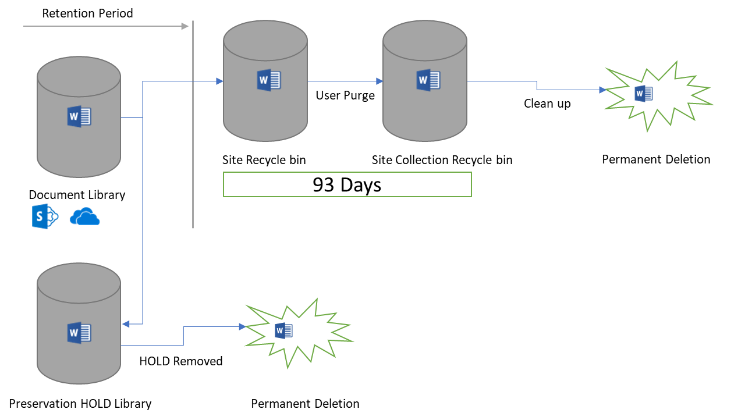
## Comprehensive Audit

Once you enable audit capability you can then drill into myriad events and activities within your environment or establish alerts for activities of interest.

# Preserve and Dispose

eDiscovery hold is a process implemented through Microsoft 365 Compliance Center to provide a single place to perform preservation activities in Office 365. As regards to SharePoint online, and One Drive for Business, the eDiscovery and hold process is described as follows:

* A preservation hold can be applied centrally (Compliance Center, Global litigation team) to a SharePoint site collection
* A special area (Preservation hold library) within each Site collection (or sub site) is created automatically to preserve items for the duration of hold
* This area is only accessible to Site collection administrators and authorized roles for eDiscovery
* Whilst the hold is active, existing and new files continue to remain editable and available for collaboration scenarios (e.g. Edit, search, Share link etc.)
* If a file has a Record retention label, the file is immutable as a record
* Any files without a retention Label modified or deleted (not by the retention / deletion policy), a copy of the original content as it existed when the HOLD was applied is copied to the hold
* The file is preserved in hold until all holds on the site are removed
* If a file is not modified or deleted, and is now due for disposition as per record retention policy, the file is permanently deleted
  + A copy of the file is maintained in the preservation hold
* Files in recycle bin are permanently deleted as standard
* Items restored from recycle bin do not change the status of file in preservation library



For a step guide to setup and use ediscovery with advanced data governance, use Microsoft documentation <https://docs.microsoft.com/en-us/microsoft-365/compliance/office-365-advanced-ediscovery>

As described in the Classify section, a Label configuration also includes the configuration to manage information status after the end of retention period. Most organizations prefer to configure an automated disposition at the end of lifecycle. This significantly minimizes the workload on record managers to review increasing volumes of records, and invest in implementing good information management practices. Where it is not possible to decide in advance whether records can be automatically disposed, Labels can also be configured to trigger a disposition review workflow at the end of retention cycle. For more information read manual disposition process here <https://docs.microsoft.com/en-us/microsoft-365/compliance/disposition-reviews>

Appendix

See here for helpful documentation and useful links to consult as you proceed on your deployment plan.

## Useful Links

1. M365 Compliance Center: <https://docs.microsoft.com/en-us/office365/securitycompliance/microsoft-365-compliance-center>
2. Content Explorer- <https://docs.microsoft.com/en-us/microsoft-365/compliance/data-classification-content-explorer>
3. Overview -<https://go.microsoft.com/fwlink/?linkid=2108832>
4. Activity Explorer - <https://go.microsoft.com/fwlink/?linkid=2109048>

## Microsoft Information Management ecosystem and partners

The Advanced Compliance Customer Experience Engineering (CXE) team in part of the Microsoft 365 Security + Compliance engineering division at Microsoft. Our team are subject matter experts who work with customers directly to help them realize the value of their investment with guidance for successful and rapid deployments.

Microsoft 365 provides a number of ways for partners to customize, extend and compliment Microsoft 365 capabilities. The Microsoft Graph APIs[[7]](#footnote-7), Management Activity API and Audit capability all provide rich access to your organizational insights and activity. There are also APIs available to establish event-based retention and consume triggers from outside systems. Over time, the Security + Compliance engineering team plans to add more Graph like APIs for common compliance workflow.

Many customers find that a partner can provide helpful expertise, insight and compliment many of the deployment processes and activities required for a successful deployment. For a list of experience information management partners to help with your deployment, see our documentation[[8]](#footnote-8)

## Microsoft 365 Information Governance Roadmap

The Microsoft 365 Roadmap website[[9]](#footnote-9) has public details about when new features are in development and when they have been launched. To see details around new Information Management features in development, rolling out, and launched search for ‘Data Governance’ and check the boxes for the details you are interested in.

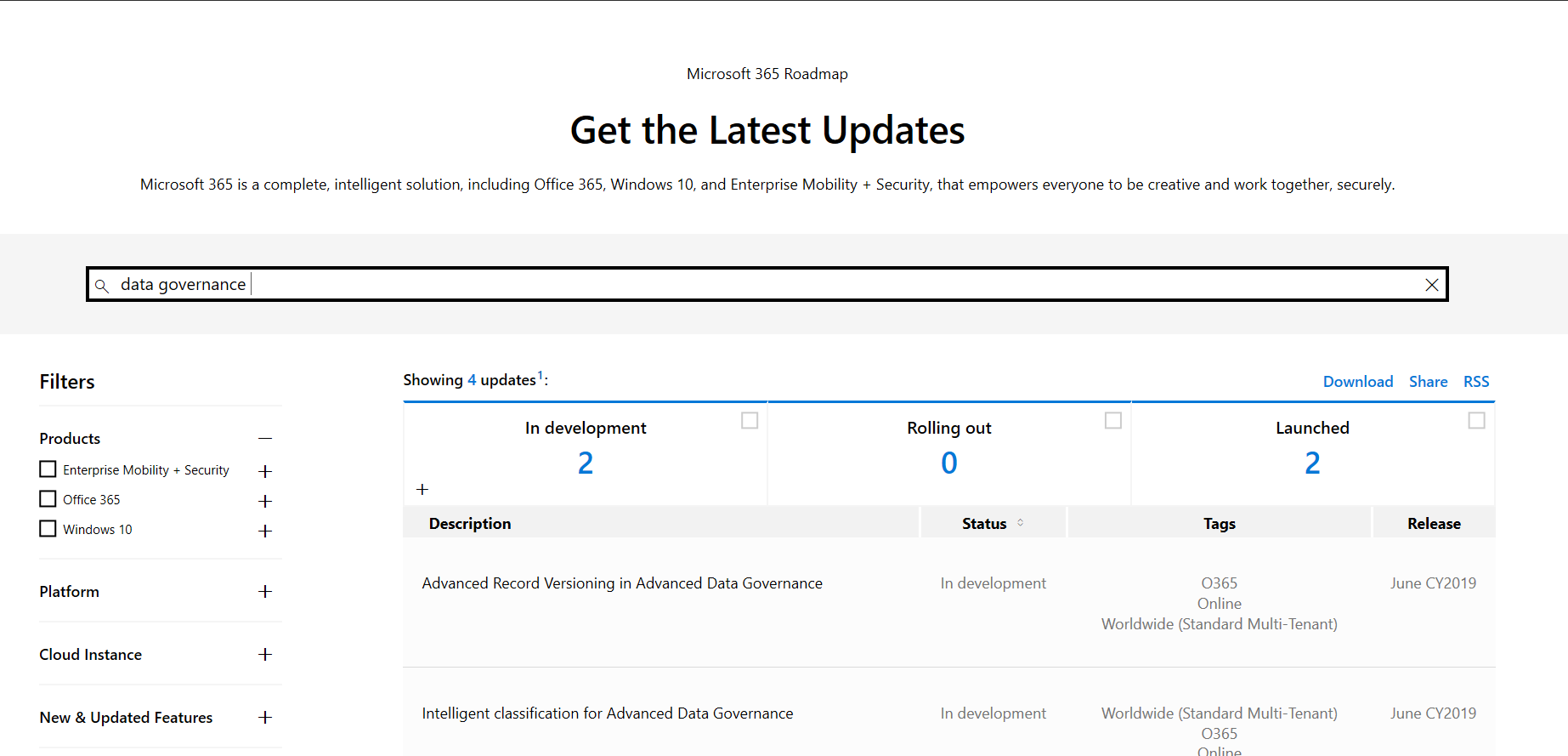


Figure 3. Microsoft 365 Data Governance Roadmap

## General Requirements

### Licensing Requirements

In this document, features of Information Governance and Advanced Information Governance are described, these features and capabilities are available via a number of subscriptions or add-ons. The Advanced Information Governance capabilities are included in Office 365 and Microsoft 365 E5, available as part of the Office 365 Advanced Compliance add-on or as part of the Microsoft 365 E5 Compliance add-on. Each user who will benefit from the capability should be licensed properly. Read more about this in our technical documentation.[[10]](#footnote-10)

### Azure Active Directory

Your organization must have an Azure Active Directory (AAD) to support user authentication and authorization for Advanced Data Governance. Configuring single sign-on (SSO) for Azure AD is also highly recommended for a seamless user experience. Additional details on requirements for Azure AD can be found in our documentation.[[11]](#footnote-11)

### Network Connectivity

A common stumbling block for many Advanced Data Governance deployments are the Firewall and Network requirements[[12]](#footnote-12). Successful deployments require a subset of the Office 365 URLs and IP address ranges. We recommend following the guidance in the Office 365 documentation and opening all Required URLs and IP addresses listed under the Microsoft 365 common and office online section[[13]](#footnote-13).

# Access

You can provide access to Microsoft 365 via role-based access control to the compliance center and the critical activities with specific roles groups for retention management, records management, disposition review and more. Review the specific role groupings with your teams and start to identify a logical, simple role and access model where you provide enough access for critical tasks and no more. Read more about granular permissions available in the technical documentation[[14]](#footnote-14)

## Additional Information

For further information contact your representative, check the blogs, websites and other public locations for new features, use cases and customer success stories

1. <https://docs.microsoft.com/en-us/office365/securitycompliance/microsoft-365-compliance-center> [↑](#footnote-ref-1)
2. <https://docs.microsoft.com/en-us/office365/securitycompliance/retention-policies> [↑](#footnote-ref-2)
3. <https://docs.microsoft.com/en-us/office365/securitycompliance/importing-pst-files-to-office-365> [↑](#footnote-ref-3)
4. <https://docs.microsoft.com/en-us/office365/securitycompliance/file-plan-manager> [↑](#footnote-ref-4)
5. <https://docs.microsoft.com/en-us/office365/securitycompliance/labels> [↑](#footnote-ref-5)
6. <https://docs.microsoft.com/en-us/office365/securitycompliance/bulk-create-publish-labels-using-powershell> [↑](#footnote-ref-6)
7. [www.graph.microsoft.com](http://www.graph.microsoft.com) [↑](#footnote-ref-7)
8. <https://docs.microsoft.com/en-us/office365/securitycompliance/work-with-partner-to-archive-third-party-data> [↑](#footnote-ref-8)
9. [www.roadmap.office.com](http://www.roadmap.office.com) [↑](#footnote-ref-9)
10. <https://docs.microsoft.com/en-us/office365/securitycompliance/retention-policies> [↑](#footnote-ref-10)
11. <https://docs.microsoft.com/en-us/azure/information-protection/requirements-azure-ad> [↑](#footnote-ref-11)
12. <https://docs.microsoft.com/en-us/azure/information-protection/requirements#firewalls-and-network-infrastructure> [↑](#footnote-ref-12)
13. <https://docs.microsoft.com/en-us/office365/enterprise/urls-and-ip-address-ranges-microsoft-365-common-and-office-online> [↑](#footnote-ref-13)
14. <https://docs.microsoft.com/en-us/office365/securitycompliance/permissions-in-the-security-and-compliance-center> [↑](#footnote-ref-14)