Trainable Classifiers

examlabpractice.com

What are Trainable Classifiers?

- A Microsoft 365 trainable classifier is a tool you can train to recognize various types of content by giving it samples to look at.
- Once trained, you can use it to identify item for application of Office sensitivity labels, Communications compliance policies, and retention label policies.
- Creating a custom trainable classifier first involves giving it samples that are human picked and positively match the category.
- After processing the positively matched samples, you test the classifiers ability to predict by giving it a mix of positive and negative samples.
- Trainable Classifiers can be used for matching items for Retention Labels, Sensitivity Labels, and Communication Compliance Policies

Important Considerations

Licensing requirements

- Classifiers are a Microsoft 365 E5, or E5 Compliance feature. You must have one of these subscriptions to make use of them.

Permissions

To access classifiers in the UI:

- The Global admin needs to opt in for the tenant to create custom classifiers.
- Compliance Administrator role is required to train a classifier.

You'll need accounts with these permissions to use classifiers in these scenarios:

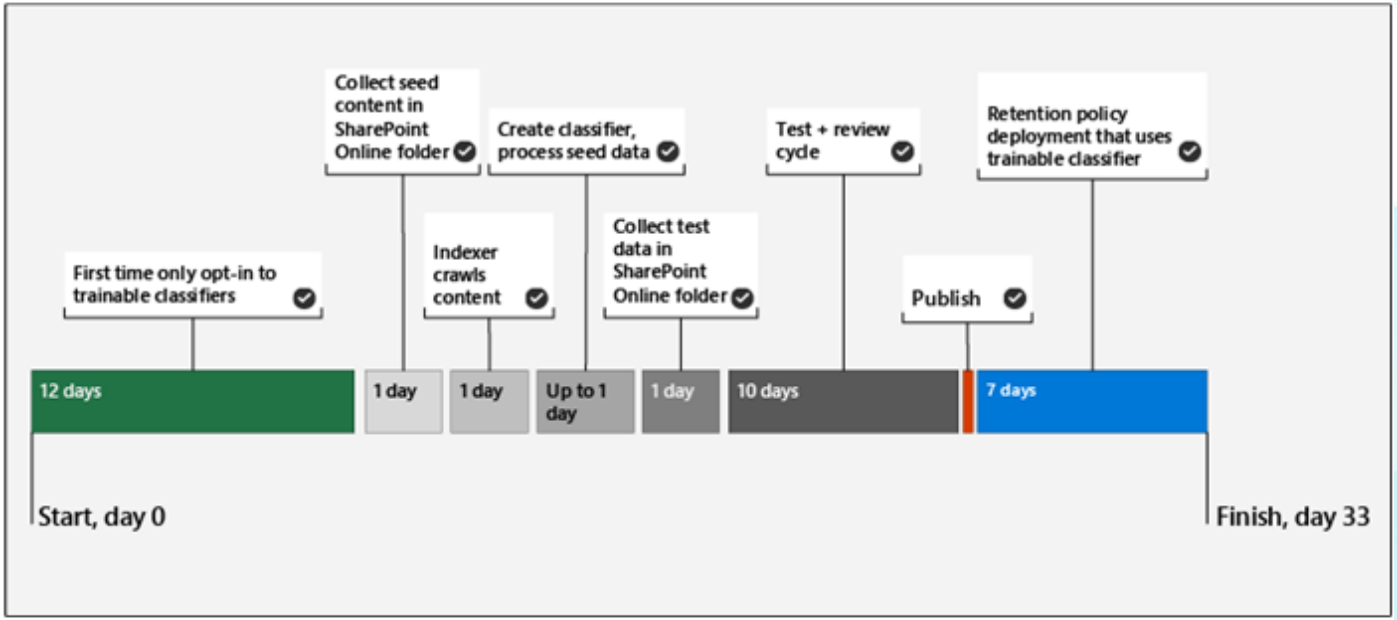
- Retention label policy scenario: Record Management and Retention Management roles
- Sensitivity label policy scenario: Security Administrator, Compliance Administrator, Compliance Data Administrator
- Communication compliance policy scenario: Insider Risk Management Admin, Supervisory Review Administrator

NOTE

- By default, only the user who creates a custom classifier can train and review predictions made by that classifier.

Timeline

This timeline reflects a sample deployment of trainable classifiers.



Seed content

- Seed content is selected by a human and is judged to represent the category of content.
- You need to have at least 50 positive samples and as many as 500.
- The trainable classifier will process up to the 500 most recent created samples (by file created date/time stamp). The more samples you provide, the more accurate the predictions the classifier will make.