

# Enterprise Vault Discovery Collector Getting Started Guide

Logan Sutterfield, E-Discovery Specialist  
Technical Field Enablement

# Contents

- Purpose of the document ..... 1
- Versions..... 1
- Use Case..... 1
- Preparing Enterprise Vault and Discovery Accelerator..... 2
- Preparing Discovery Collector..... 3
- Configuring Data Sources & Destinations ..... 7
- Setting Up the Retention Volume..... 7
- Setting Up Collection from a File Share..... 9
- Setting Up Collection from Exchange..... 11
- Setting Up Collection from Domino..... 15
- Setting Up Collection from SharePoint..... 18
- Setting Up Collection from Discovery Accelerator..... 20
- Server Collection..... 23**
- Creating a Workspace ..... 23
- Building a Volume Index..... 25
- Classification the Volume Data..... 31
- Import the Volume Data to Enterprise Vault ..... 34
- Summary ..... 40

## Purpose of the document

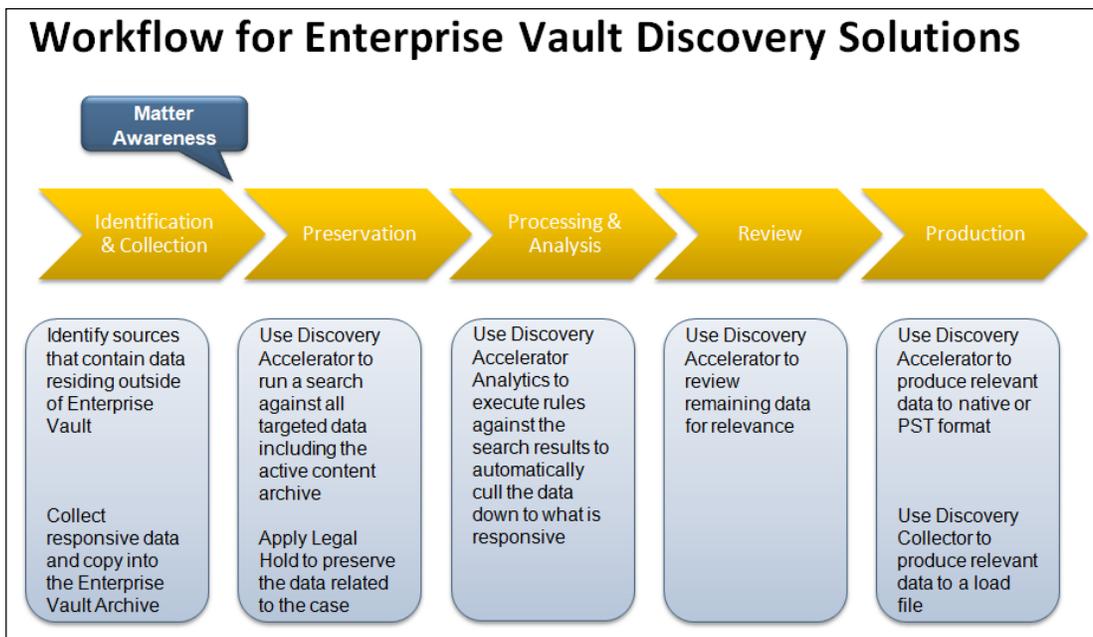
The purpose of this document is to provide the reader with the basic steps needed to initially configure the Enterprise Vault Discovery Collector. The various sections in this document will assist the reader in setting up source locations for collecting active content from servers and desktops, classifying the data, and copying the data into Enterprise Vault. As an added benefit, you should have the Discovery Collector User Guide handy to read further on details not covered in this document.

## Versions

9.0 All Service Packs

## Use Case

The basic use case for Discovery Collector would be to collect the 10-20% of data residing on the Exchange servers and file servers, that has not become eligible for archiving, and copy it into Enterprise Vault. Once this data is copied into Enterprise Vault, the legal team will be able to access the data along with previously archived data from within the same case, allowing a “single pane of glass” to search all user data in the enterprise.



## Preparing Enterprise Vault and Discovery Accelerator

The first step in preparing to collect active data into the archive is to configure Enterprise Vault and Discovery Accelerator for storing and searching of the active data. When using Discovery Collector for data in the wild, it is recommended that Enterprise Vault be utilized as the retention platform thereby giving the legal team one interface in Discovery Accelerator for managing the case work product.

### **Configuring Enterprise Vault as the Destination for Collected Content**

The destination archive in Enterprise Vault will be created using Discovery Collector once a collection request has been submitted to litigation support; this will require that the appropriate Vault Store and Vault Store Group be configured prior to the request.

The Vault Store Group is the container for Vault Stores and is the logical boundary for single instance storage. Where you do have the option to store the active data in an existing Vault Store; it is recommended that you separate the data out into a dedicated store.

### **Create a Retention Period to associate with the case data**

When data is copied into Enterprise Vault using the Discovery Collector, it must be assigned a retention category; the value of the retention category can be changed at the conclusion of the case. The retention category is selected in Discovery Collector when configuring the destination archive in Enterprise Vault. It is recommended that a new retention category is created for each active matter so that at the conclusion of the case, the data can be expired immediately.

Log in to the Enterprise Vault admin console and follow the instructions in Appendix B for creating a new retention category.

### **Discovery Accelerator Case Creation**

Typically a case will be created in Discovery Accelerator before a request for active content has been submitted. Having a case in place provides a point of reference for creating a retention volume in Discovery Collector. Once data has been collected, it will be imported to Enterprise Vault and will be available for search, Legal Hold, and review. This provides the legal team with one interface for processing of data from the wild or normally archived data in preparation for litigation.

## Preparing Discovery Collector

The second step in preparing to collect active data into the archive is to configure Discovery Collector to communicate with Enterprise Vault and Discovery Accelerator. When using Discovery Collector for data in the wild, it is recommended that Enterprise Vault be utilized as the retention platform thereby giving the legal team one interface in Discovery Accelerator for managing the case work product.

### Configure the Enterprise Vault Site on Discovery Collector

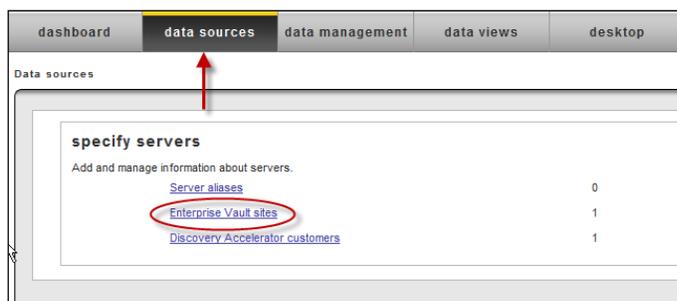
When data is collected from sources in the wild, it must be copied to a retention volume so that Legal Hold may be applied. The recommended retention volume location is Enterprise Vault therefore communication between Discovery Collector and Enterprise Vault must be established.

Follow the steps below to configure the Enterprise Vault site within the Discovery Collector interface to establish communication.

1. Login to Discovery collector as the administrator

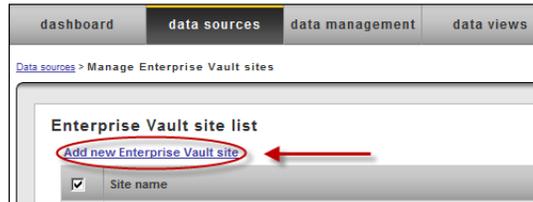


2. From the "data sources" menu, select "Enterprise Vault sites"



# Discovery Collector Getting Started Guide

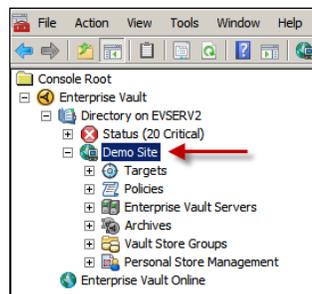
Then add a new site as seen below:



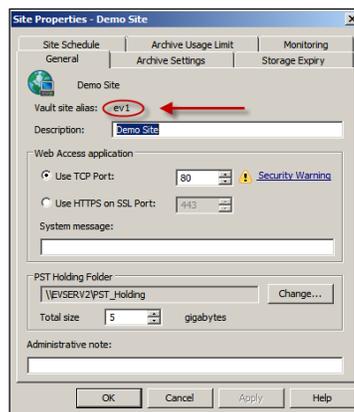
Configure the settings for the Enterprise Vault Site



- Site Name: This can be found in the Enterprise Vault Admin Console



- Enterprise Vault Site Alias: This can be found in the "Properties" of the site and will be a fully qualified domain name, e.g. ev1.evexample.local



- User name: domain name\username – This will be the Enterprise Vault admin account

3. Click “Ok” and confirm that the site name appears in the list

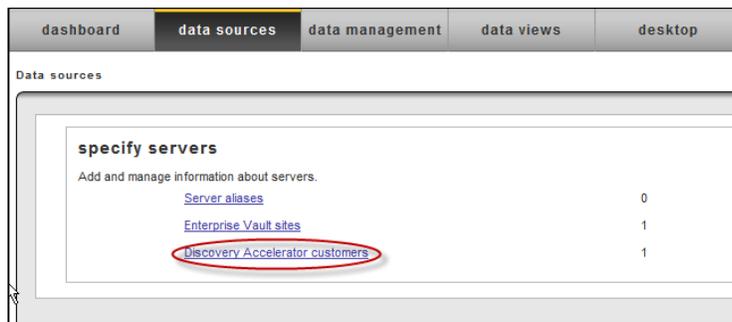


### Configure the Discovery Accelerator Customer on Discovery Collector

When active data has been collected and copied into Enterprise Vault; legal hold, search, and processing will be performed by the Discovery Accelerator. Once the search results have been culled down to what will be produced, the Discovery Collector can then collect the relevant results directly out of Discovery Accelerator and placed into a “Load File”. In order for this to work, communication between the Discovery Collector and Discovery Accelerator must be established.

Follow the steps below to configure a Discovery Accelerator customer in Discovery Collector.

1. From the “data sources” menu, select “Discovery Accelerator Customers”



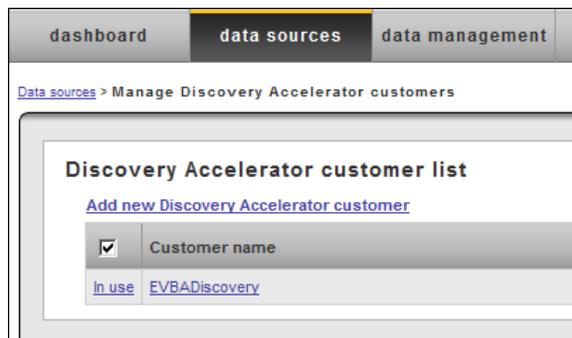
2. Then add a new customer and populate the fields respective to your environment.



Example:

Customer name:	<input type="text" value="EVBADiscovery"/>
Discovery Accelerator server:	<input type="text" value="DAServer.domain.com"/>
Customer virtual directory:	<input type="text" value="EVBADiscovery"/>

3. Click "Ok" and confirm that the customer name in the list



Once the Discovery Collector has been implemented and configured to communicate with your existing Enterprise Vault and Discovery Accelerator servers, you can then begin setting up your target data sources as well as the target data destinations in the Enterprise Vault archive.

## Configuring Data Sources & Destinations

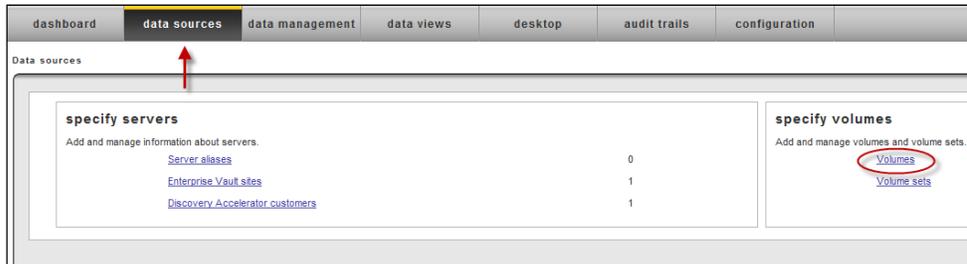
The second step in preparing to collect active data into the archive is to configure Discovery Collector to communicate with Enterprise Vault and Discovery Accelerator. When using Discovery Collector for data in the wild, it is recommended that Enterprise Vault be utilized as the retention platform thereby giving the legal team one interface in Discovery Accelerator for managing the case work product.

Prior to configuring the Discovery Collector System volume, you must configure the server platforms you will use for the different volume types. Each server type has requisite permissions and settings; these are described below.

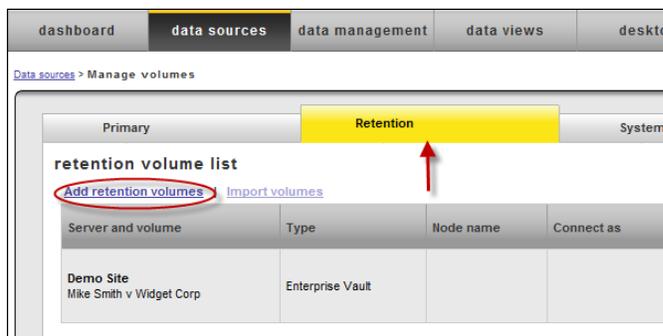
### Setting Up the Retention Volume

A retention volume is an archive in Enterprise Vault to store and preserve the responsive data throughout the matter. This retention volume should be associated with an existing case in Discovery Accelerator and should also have a dedicated retention category pre-configured in Enterprise Vault. With a dedicated retention category, you will be able to expire all active data collected for the case once it is closed.

1. While logged into Discovery Collector, select the “data sources” menu, and click the “Volumes” link.



2. Then select the “Retention” tab and click “Add Retention Volumes”.



3. Next you will configure the settings for the retention volume.

**Add volume**

Volume details \* Required

**Server definition**

Server type: \* Enterprise Vault

Site: \* Demo Site

Vault store: \* Active Content

**Volume details**

Matter archive name: \* Mike Smith v Widget Corp

Description: Intellectual Property Matter

Default retention category: \* Mike Smith v Widget Corp

[Show advanced options](#)

Retention category override:  Allow retention category to be overridden on policy

Constraints:  Only use one connection process(es)

OK Cancel

- **Server Type:** Enterprise Vault
  - **Site:** The name of your Enterprise Vault site that was configured in previous steps
  - **Vault Store:** Active Content or the name that you assigned the Vault Store for active data in previous steps
  - **Matter Archive Name:** A name for the archive to be created. It should be associated with the case name in Discovery Accelerator.
  - **Retention:** Select the retention category that you created in previous steps to associate the collected data to.
  - **Description:** Optional
  - **Advanced Options:** Leave default or refer to the product documentation for more details.
4. Click "Ok" and there will be a "Volume Successfully added" dialogue box. Click "Finished adding volumes" link to continue. You have just added an archive in Enterprise Vault.

Once the retention volume has been created from the Discovery Collector, it will reside in the Enterprise Vault "shared" archives folder in the admin console.

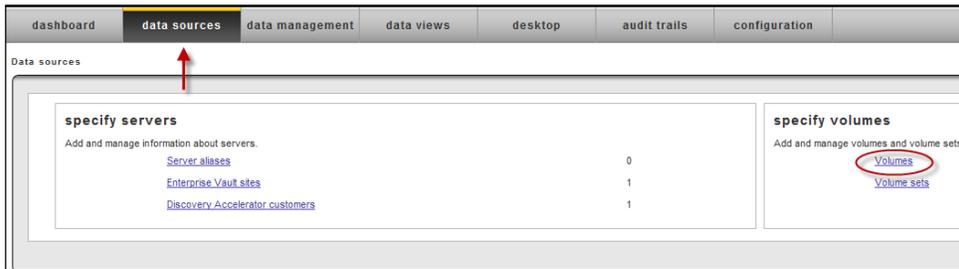
## Setting Up Collection from a File Share

Collecting data from a file server will only be needed for servers that have not been configured for archiving or there is responsive data that hasn't been archived yet by the scheduled archiving run. In most Enterprise Vault environments, you can expect anywhere from 10-20% of file data to be outside of the archive depending on how the archiving policies are configured.

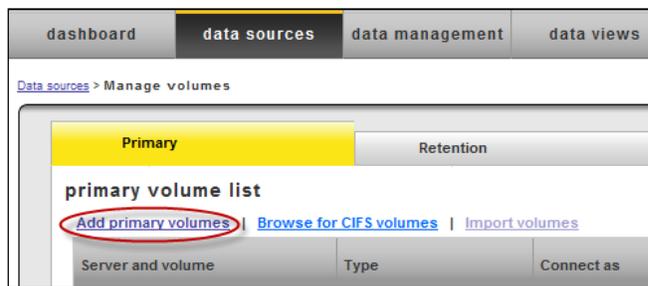
To collect data from volumes on File Servers or CIFS (Common Internet File System), the domain account used while creating this volume, must be in the backup operator group on the server. Typically you will target a specific folder on a user share server and a top level folder on a community or project file share.

## Configuring the File Share Primary Volume

1. While logged into Discovery Collector, select the "data sources" menu, and click the "Volumes" link.



2. Then from the "Primary" tab, click the "Add Primary Volumes" link



Next you will configure the settings for a File Share primary volume.

## Discovery Collector Getting Started Guide

The screenshot shows the 'Add volume' dialog box with the following fields and annotations:

- Server definition:**
  - Server type: CIFS (Windows platforms)
  - Server: FileServer.domain.com
  - Connect as: domain\user\_account (Annotated: Member of the Backup Operators security group)
  - Password: [Redacted]
- Volume details:**
  - Volume: Users (Annotated: Name of shared folder)
  - Initial directory: [Empty]
  - Index options:  include system metadata for data objects within containers,  include content tagging and full-text index

- **Server Type:** CIFS (Windows Platforms)
- **Server:** Name of the server where the share resides
- **Connect as:** domain\user account with adequate permissions
- **Volume:** This is the name of the folder as shared out through Windows File Sharing
- **Index options:** Leaving these in the default of "Selected" will result in indexing metadata of objects inside of zip files. If required, please refer to the product documentation for more details and considerations.

3. Click "Ok to continue and you should see a "Volume Successfully added" dialogue box.
4. Click "Finished adding volumes" link to continue and verify the volume has been added

evserv2 Users	CIFS	evexample\admin
------------------	------	-----------------

### Setting Up Collection from Exchange

Collecting data from an Exchange server will only be needed for employees that are not currently being “journalized” by Enterprise Vault or if there is residual data that existed in the mailbox before journaling was implemented. In most cases, the employee will be a member of a department set up for mailbox archiving and all residual data will already be available in the archive for searching.

First you should examine the permissions, authentication, and connection settings below that are specific to configuring and Exchange volume.

#### Permissions

The following permissions must be set on the Exchange 2003 server to the mailbox store or the mailboxes from which you will harvest.

- Read
- Execute
- Read permissions
- List contents
- Read properties
- List object
- Receive as

#### Permissions for Exchange 2007 and 2010

The Full Access permissions must be set on the Exchange server to the mailbox store or the mailboxes you plan on harvesting.

#### Recoverable items in Exchange 2010

To harvest the recoverable items folders in Exchange 2010, you must be a member of the Discovery Management role.

#### Deleted items

To harvest items that have been deleted from the Exchange server, enable Exchange’s transport dumpster settings. For more information, refer to *Microsoft® Exchange Server 2010 Administrator’s Pocket Consultant*. Configuration information is also available online at [www.microsoft.com](http://www.microsoft.com)

#### Windows Authentication

For all supported versions, enable Integrated Windows Authentication on each Exchange server, as follows:

1. From Microsoft Windows, log in to the Exchange Server.
2. Go to *Administrative Tools > Internet Information Services (IIS) Manager*.
3. Go to *Internet Information Services > Name of Exchange Server > Web Sites > Default Web Site*.
4. Right-click *Default Web Site*.
5. Click the *Directory Security* tab.
6. In the Authentication and access control pane, click *Edit*.
7. Select *Properties*. The Authentication Methods window appears.
8. In the Authentication access pane, click the *Integrated Windows authentication* checkbox.
9. Click *ok*.
10. Restart IIS Services.

## Improve performance for IIS 6.0 and Exchange 2003.

1. From Microsoft Windows, log onto the Exchange Server.
2. Go to *Administrative Tools > Internet Information Services (IIS) Manager*.
3. Select *Internet Information Services > <Name of Exchange Server> <Web Sites> Application Pools*.
4. Right-click *Application Pools* and select *Properties*.
5. On the *Performance* tab, locate the Web Garden section.
6. Enter a number of processes equal to or greater than the harvester processes you have running for the maximum number of worker processes.
7. Click *ok*.
8. Restart IIS Services

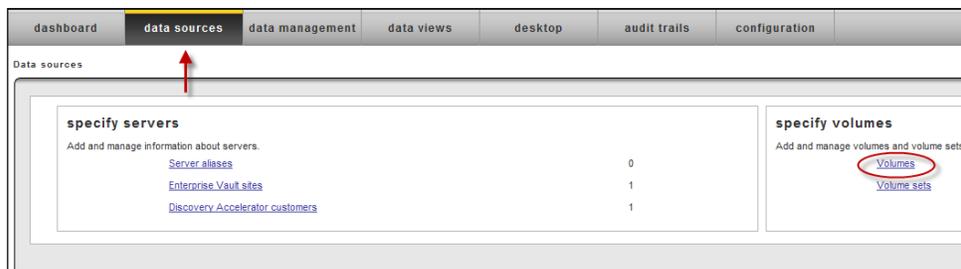
## Secure connection

If you want to connect to Exchange volumes over HTTPS, you must add port number 443 after the server name—for example, **exch2003.domain.local:443**. In some cases, this secure connection can result in some performance degradation due to SSL overhead. If you enter the volume information without the 443 suffix, the default connection will be over HTTP.

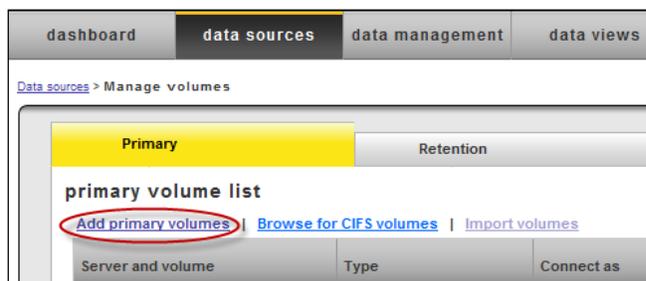
## Configuring the Exchange Primary Volume

Follow the steps below to configure Exchange for collection with Discovery Collector.

1. While logged into Discovery Collector, select the “data sources” menu, and click the “Volumes” link.

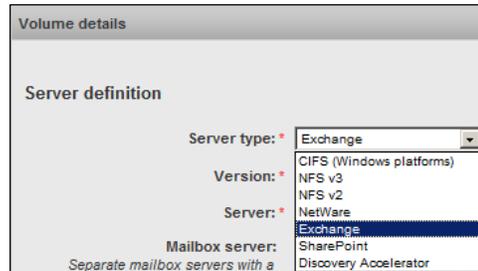


2. Then from the “Primary” tab, click the “Add Primary Volumes” link



## Discovery Collector Getting Started Guide

Next you will select "Exchange" from the server type drop down menu.



Volume details

Server definition

Server type: \* Exchange

Version: \* CIFS (Windows platforms)

Server: \* NFS v3

Mailbox server: \* NFS v2

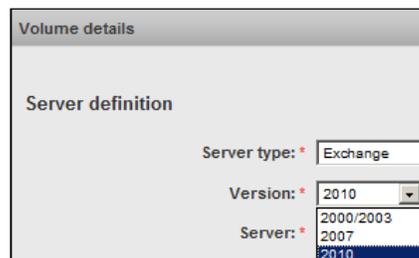
Separate mailbox servers with a \* NetWare

Exchange

SharePoint

Discovery Accelerator

You will then have the option to select which version of Exchange you are running in your environment



Volume details

Server definition

Server type: \* Exchange

Version: \* 2010

Server: \* 2000/2003

2007

2010

3. Now populate the fields according to your environment

## Discovery Collector Getting Started Guide

The screenshot shows the 'Add volume' dialog box with the following fields and values:

- Server type: Exchange
- Version: 2000/2003
- Server: exch2003.domain.com
- Mailbox server: (empty)
- Active directory server: (empty)
- Connect as: domain\exchangeadmin
- Password: (masked)
- Volume: Mailboxes
- Folder: Mailboxes (selected)
- Initial directory: mike.smith@domain.com (circled in red)
- Virtual root: (empty)
- Index options: (unchecked)

- **Server Type:** The type of server you will be targeting for collection
- **Version:** The version of the Exchange server you are targeting
- **Server:** The name of the Exchange server
- **Mailbox Server:** Only needed if using a Client Access Server (CAS)
- **Active Directory Server:** Only used if needing to specify a Global Catalog Server (GC)
- **Connect as:** Refer to the permission section above to ensure the adequate account is used
- **Volume:** The name that you would like this volume to appear as in the interface
- **Initial Directory:** If targeting a specific mailbox, enter the SMTP address
- **Index options:** Leaving these in the default of "Selected" will result in indexing metadata of objects inside of zip files. If required, please refer to the product documentation for more details and considerations.

4. Click "Ok to continue and you should see a "Volume Successfully added" dialogue box.
5. Click "Finished adding volumes" link to continue and verify the volume has been added

evserv1.evexample.local Mailboxes	Exchange (2010)	evexamplelevsvc
--------------------------------------	-----------------	-----------------

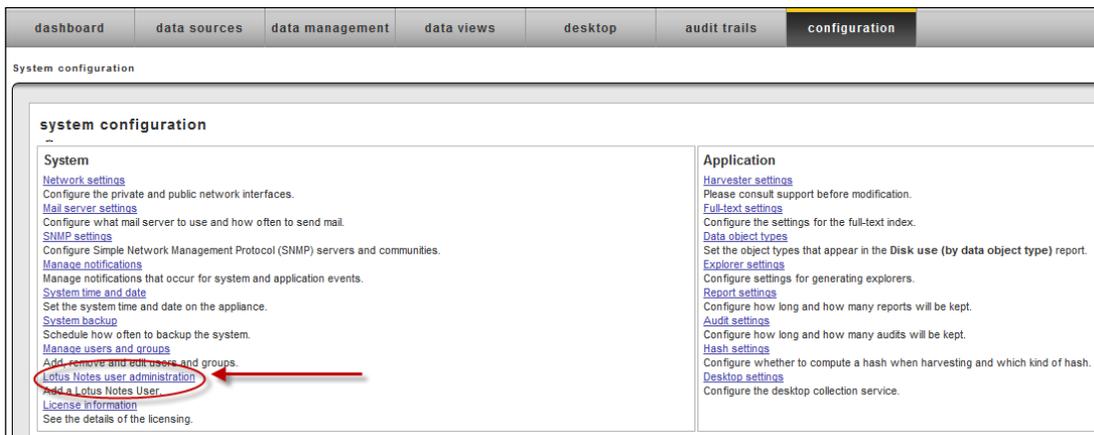
## Setting Up Collection from Domino

Collecting data from a Domino server will only be needed for employees that are not currently being “journalized” by Enterprise Vault or if there is residual data that existed in the mailbox before journaling was implemented. In most cases, the employee will be a member of a department set up for mailbox archiving and all residual data will already be available in the archive for searching.

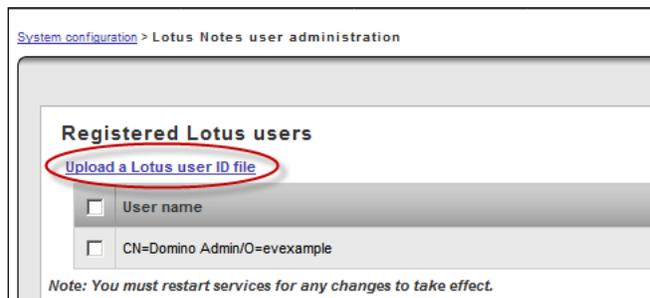
## Configuring the Domino Primary Volume

Follow the steps below for configuring Domino for collection with Discovery Collector

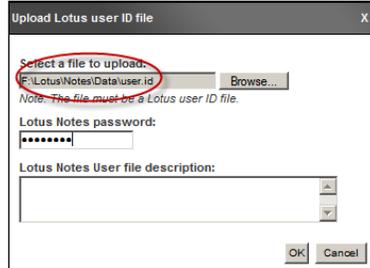
1. First you must add a Lotus Notes user by uploading its user ID file. While logged into Discovery Collector, select the “Configuration” menu and click on the “Lotus Notes User Administration” link.



2. Next click on the “Upload a Lotus user ID file” link



And browse to the location of the user ID file

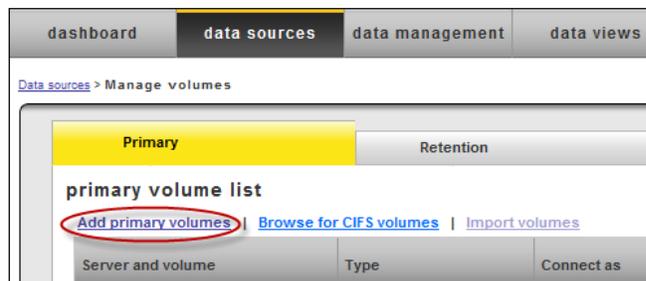


- If you would like to harvest a user's mailbox, add the user ID file for that user.
- If you would like to harvest multiple mailboxes within one volume definition, add the administrator's ID file.
- If the mailboxes have encrypted emails or NSFs, then you will need each user's user-id file in order to decrypt a given user's data.

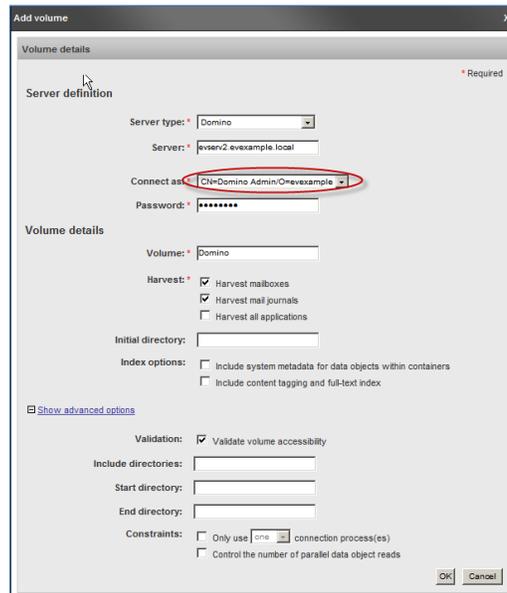
3. Once the ID file is uploaded, you can now select the "data sources" menu, and click the "Volumes" link.



Then from the "Primary" tab, click the "Add Primary Volumes" link



### 4. Now populate the fields according to your environment



- **Server Type:** Select Domino from the drop down menu
- **Server:** This will be the fully qualified name of your Domino Server
- **Connect As:** For every user ID that has been uploaded, a user will appear in the drop down menu.
- **Volume:** The name of the volume as you would like to see it in the interface. In this example, we simply used 'Domino'.
- **Index options:** Leaving these in the default of "Selected" will result in indexing metadata of objects inside of zip files. If required, please refer to the product documentation for more details and considerations.
- **Initial Directory:** If a single mailbox must be harvested, set the initial directory to be the path to the mailbox on the Domino server, such as *mail\USERNAME*.
  - To harvest mailboxes, select the *Harvest mailboxes* option, which obtains the list of all known Domino users and their NSF's. It then harvests those mailboxes unless it was pointed to a single mailbox using the initial directory.
  - To harvest all mail journals, select the *Harvest mail journals* option.
  - To harvest all mail applications, select the *Harvest all applications* option, which looks at all NSF's, including mail journals, on the Domino server.
- **Index options:** Leaving these in the default of "Selected" will result in indexing metadata of objects inside of zip files. If required, please refer to the product documentation for more details and considerations.

5. Click "Ok to continue and you should see a "Volume Successfully added" dialogue box.

6. Click "Finished adding volumes" link to continue and verify the volume has been added

evserv2.evexample.local Domino	Domino	CN=Domino Admin/O=evexample
-----------------------------------	--------	--------------------------------

### Setting Up Collection from SharePoint

Collecting data from a SharePoint server will only be needed for employees with data on sites that is not currently archived by Enterprise Vault. In most Enterprise Vault environments, you can expect anywhere from 10-20% of SharePoint data to be outside of the archive depending on how the archiving policies are configured.

First you should examine the permissions and connection settings below that are specific to configuring and SharePoint volume.

#### Roles

To run policies on SharePoint servers, you must configure the following roles on them.

#### Contributor

To harvest, copy from SharePoint as a source volume, and reset existing attributes.

#### Web Designer

To set custom Discovery Collector attributes, reset existing attributes, harvest, and run policies on SharePoint as a source or destination volume.

**Secure connection.** If you wish to connect to SharePoint volumes over HTTPS, you must add the port number 443 after the server name when setting up the volume on the Symantec Discovery Collector Platform. For example: **sps01.company.com:443**.

In some cases, this secure connection can result in some performance degradation, due to Secure Socket Layer (SSL) overhead. If you enter the volume information without the 443 suffix, the default connection is over HTTP.

### Configuring the SharePoint Primary Volume

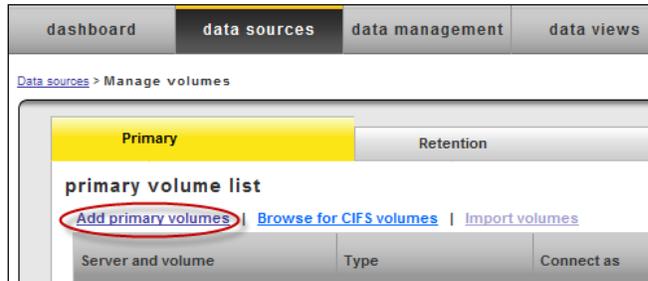
Follow the steps below for configuring SharePoint for collection with Discovery Collector

1. While logged into Discovery Collector, select the "data sources" menu, and click the "Volumes" link.



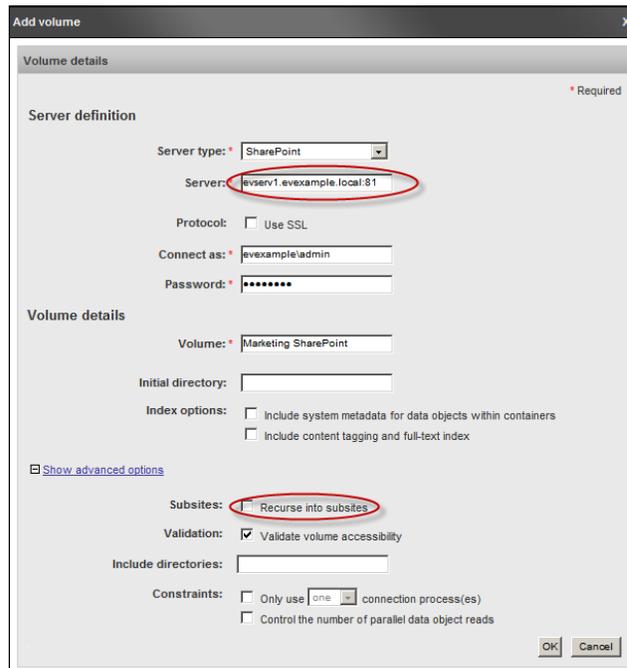
## Discovery Collector Getting Started Guide

2. Then from the "Primary" tab, click the "Add Primary Volumes" link



Next you will select "SharePoint" from the server type drop down menu.

3. Now populate the fields according to your environment

The screenshot shows a dialog box titled 'Add volume' with a close button (X) in the top right corner. The dialog is divided into two main sections: 'Server definition' and 'Volume details'. In the 'Server definition' section, there is a 'Server type' dropdown menu set to 'SharePoint', a 'Server' text field containing 'evserv1.evexample.local:81' (circled in red), a 'Protocol' checkbox for 'Use SSL', a 'Connect as' text field with 'evexample\admin', and a 'Password' text field with masked characters. The 'Volume details' section includes a 'Volume' text field with 'Marketing SharePoint', an 'Initial directory' text field, and 'Index options' with two unchecked checkboxes: 'Include system metadata for data objects within containers' and 'Include content tagging and full-text index'. There is a 'Show advanced options' link. Below this, the 'Subsites' dropdown is set to 'Recurse into subsites' (circled in red), and the 'Validation' checkbox for 'Validate volume accessibility' is checked. At the bottom, there are 'Include directories' and 'Constraints' sections with several unchecked checkboxes. 'OK' and 'Cancel' buttons are located at the bottom right.

- **Server Type:** SharePoint
- **Server:** The name of your SharePoint server followed by the port number if applicable
- **Connect As:** The user with an adequate SharePoint role
- **Volume:** The name as you would like to see it in the interface. In this example, we used "Marketing SharePoint"
- **Initial Directory:** A specific repository may be specified if appropriate

- **Index options:** Leaving these in the default of “Selected” will result in indexing metadata of objects inside of zip files. If required, please refer to the product documentation for more details and considerations.
  - **Sub-sites:** This should be selected if you need to recurse into sub-sites for the collection
7. Click “Ok to continue and you should see a “Volume Successfully added” dialogue box.
  8. Click “Finished adding volumes” link to continue and verify the volume has been added

evserv1.evexample.local:81 Marketing SharePoint	SharePoint	evexample\admin
--	------------	-----------------

### Setting Up Collection from Discovery Accelerator

Prior to configuring Discovery Accelerator primary volumes, you must configure Discovery Accelerator customer information and Enterprise Volume sites in Discovery Collector so that communication has been established. If properly configured, you will see Discovery Accelerator appear in the volume configuration drop-down menus.

All case data can be exported from either Discovery Accelerator or Discovery Collector. Data will only need to be collected from Discovery Accelerator if you plan to export to a load file for ingestion into another analytics tool. If your outside counsel is requesting the data in PST, NSF, or original format only; this volume is not required. In order to export Lotus Domino items archived by Enterprise Vault; Lotus Notes 7.0.3 or later is required on the Discovery Accelerator server.

Before configuring this volume, you should review the permissions, configuration, and performance information below.

#### Permissions

It is recommended that the credentials used for referencing the Enterprise Vault Site are those of the Vault User or any other administrator. Discovery Collector will validate that the credentials are strong enough for it to:

- Login remotely to the specified server
- Perform DCOM operations over RPC remotely.

To successfully harvest a Discovery Accelerator volume, a user must have the following privileges:

1. A Role defined in the Discovery Accelerator Web Application.
2. Review Messages permission for the Case used in the volume definition.
3. Folder Review permissions on a Case, if a Folder (sometimes also called a ‘Research Folder’) in the Case is going to be harvested.
4. Permission to set all of the Review Marks that will be selected for the Volume definition.

#### Discovery Accelerator Web-Services Interface:

In order for Discovery Collector to interface with Enterprise Vault via the Discovery Accelerator web-services, the following configurations must be made on the Discovery Accelerator service by logging in to the Discovery Accelerator client interface as a Vault User or Discovery Administrator.

1. Click the *Configuration* tab, and expand the API options on the Settings page.
2. In the API settings group, ensure that the API Enabled setting is enabled.
3. Configure a Temporary Storage Area, as needed. Ensure that it has sufficient free space and that any authenticated users that will define volumes against Discovery Accelerator have Full Control permissions on this storage area. Additionally, consider configuring the Temporary Storage Area Cleanup Interval, depending on the size of the cluster that will be deployed—typically, the default

value of 30 minutes should be sufficient. If greater than 4 nodes will be deployed in the cluster, this interval must be reduced accordingly for more frequent cleanups to free up storage space.

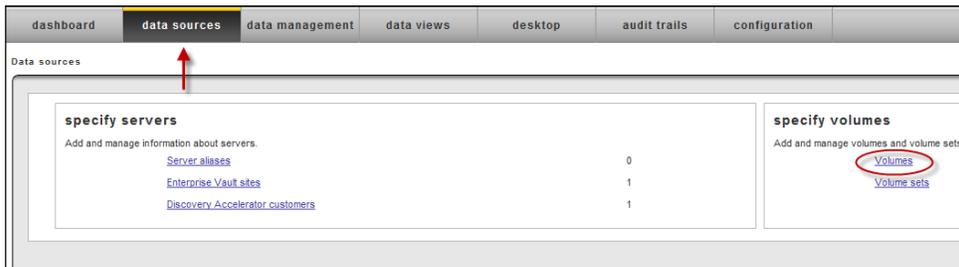
### Improve performance for IIS 6.0:

If the Discovery Accelerator server runs over IIS 6.0, an existing bug in IIS causes severe performance degradation when used along with Kerberos authentication. It is recommended that the hotfix described in Microsoft Knowledge Base article 917557 (<http://support.microsoft.com/kb/917557>) be applied to the server in this case.

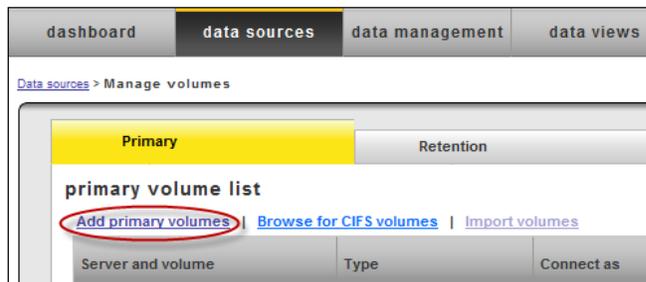
### Configuring the Discovery Accelerator Primary Volume

Follow the steps below to configure a primary volume for Discovery Accelerator.

1. While logged into Discovery Collector, select the "data sources" menu, and click the "Volumes" link.



2. Then from the "Primary" tab, click the "Add Primary Volumes" link



Next you will configure the settings for a Discovery Accelerator primary volume.

## Discovery Collector Getting Started Guide

Volume details

Server definition

Server type: Discovery Accelerator

Customer: EVBADiscovery

Active directory server:

Protocol:  Use SSL

Connect as: evexample\admin

Password: \*\*\*\*\*

Site: Demo Site

Volume details

Volume: Mike Smith v Widget Corp

Discovery Accelerator case: Loading cases...

Scope of data:  All data in this case  
 Only selected review marks

Index options:  Include system metadata for data objects within containers  
 Include content tagging and full-text index

Show advanced options

OK Cancel

- **Server Type:** Discovery Accelerator
- **Customer:** The name of the customer in Discovery Accelerator
- **Connect as:** domain\case admin account (Just being a case owner will not work). Once you have entered credentials that have Discovery Accelerator case admin rights, and click in the "Volume" field, you will see the cases and marks being populated.
- **Site:** The Enterprise Vault Site name will be populated
- **Volume:** Name this the same as the Discovery Accelerator Case
- **Discovery Accelerator Case:** The cases that your account has administrative rights to will appear in the drop down as seen below.

Volume details

Volume: Mike Smith v Widget Corp

Discovery Accelerator case: Mike Smith v Widget Corp

Scope of data:  All data in this case  
 Only selected review marks

Helps Us  
Helps Them  
No Mark  
Relevant  
Query

- **Scope of Data:** You have the ability to only export items with a specific mark or you can export the entire case data set regardless of the mark.
  - **Index options:** Uncheck all boxes. Full text indexing is not required during the export; items are exported based on marks.
3. Click "Ok to continue and you should see a "Volume Successfully added" dialogue box.
  4. Click "Finished adding volumes" link to continue and verify the volume has been added

EVBADiscovery Mike Smith v Widget Corp	Discovery Accelerator	evexample\admin
---	--------------------------	-----------------

## Server Collection

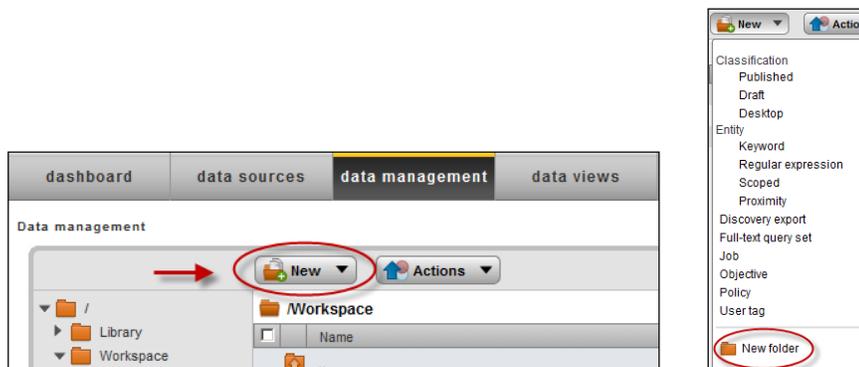
Once the Enterprise Vault Discovery Collector system has been installed and configured, most work is performed in folders. The *data management* menu displays two first-level folders: *workspace* and *library*. The workspace folder, as its name suggests, is your workspace and allows you to create a hierarchy of folders that reflect your use of the system. The library folder contains the attributes, classifications, and jobs that are standard in the system. These cannot be edited or removed. If your server is part of a federated system, a federated folder is also present.

For example, you could start by creating folders named *Active Matters* and *Closed Matters*. As new matters arise, you can then create new folders within the Active Matters folder and keep all the necessary classifications, entities, tags, policies, and jobs in one place. When the matter completes, you can simply move the folder in its entirety to the Closed Matters.

## Creating a Workspace

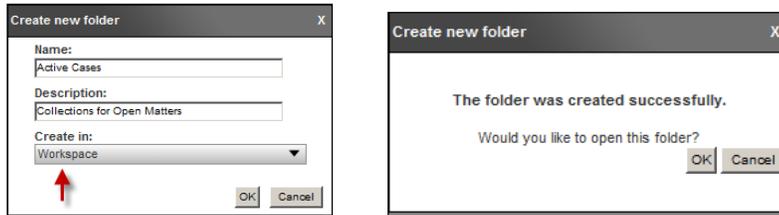
To create a workspace, you must decide how you would like to organize the user generated content. In this section we will use the most common method of organization. Use the following steps to create your workspace.

1. While logged into Discovery Collector, select the "Data Management" menu and click on the "New" button underneath and choose "New Folder"



2. The new folder dialogue appears; populate the fields with the name of your top level folder. In this example, we will use "Active Cases"

## Discovery Collector Getting Started Guide



Be sure to create the new folder in the "Workspace" folder. Click "OK" and then again to open the newly created folder. Follow the same process for creating a "Closed Cases" folder.

3. You may also follow the same process to create subfolders for Classifications, Jobs, Policies, etc. as seen below, or choose to store all work in one single folder and organize later by using the "Move" option.



4. You are now ready to begin collecting, proceed to the next page.

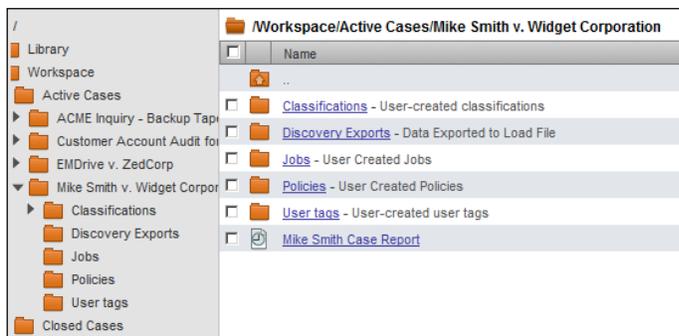
### Building a Volume Index

When collecting active data or “Data in the Wild”, the first step is to build an index of what is not in the archive so that you can then take actions such as classifying or tagging. We refer to the build of an index of non-archived data as “Harvesting”

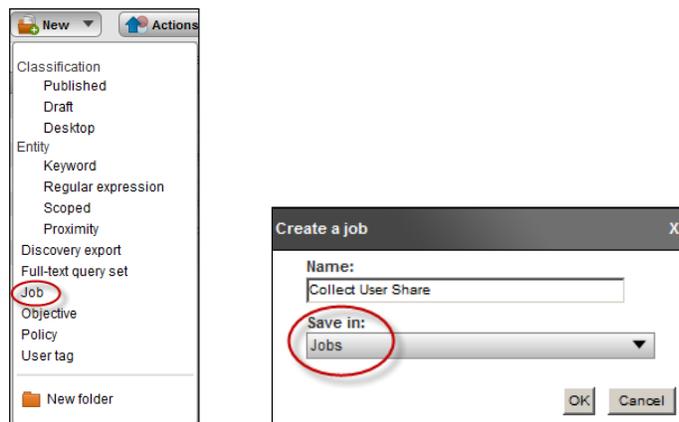
Harvesting is the process or task by which Discovery Collector examines data in your network and builds an index. You can run a full or incremental harvest; noting that the full harvest may be run against all volumes and incremental will only return changes on specifically requested volumes. These options are selected when you create a job for the harvest.

In this example, we will harvest from a user share located on a file server that is currently being archived by Enterprise Vault; we will only collect the data that is not already in the archive.

1. While logged into Discovery Collector, select the “Data Management” menu, and drill down to your case folder underneath “Active Cases”. In this example, we will use the “Mike Smith v. Widget Corporation” case folder as seen below.



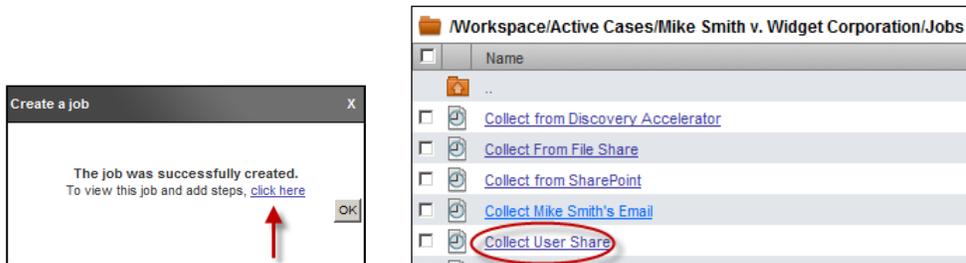
2. Next from the “New” dropdown menu, select “Job” and use the following settings:



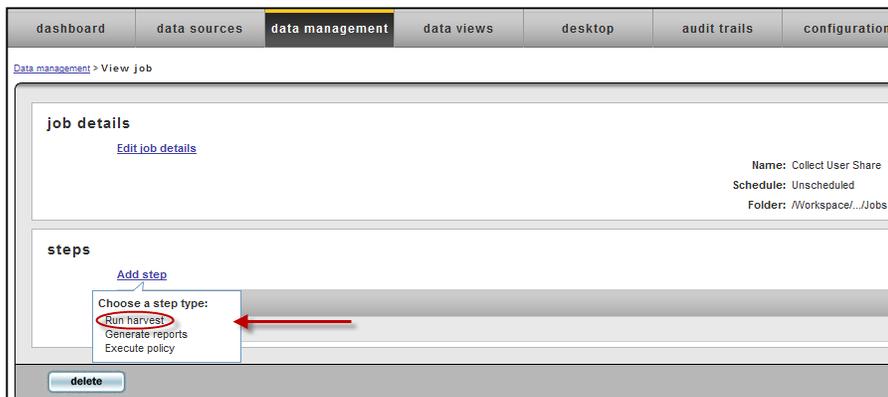
- **Name:** Collect User Share
- **Save in:** Jobs (under Mike Smith v. Widget Corporation)

## Discovery Collector Getting Started Guide

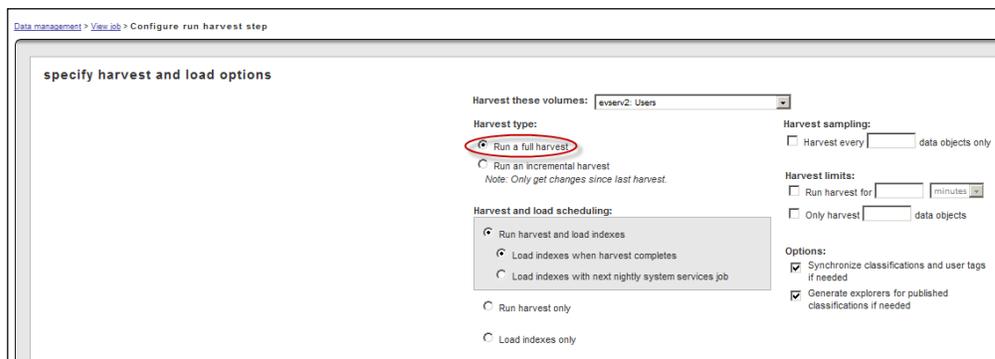
- Once the job has been created, use the link "click here" to add steps. If you clicked "OK" instead, you can still configure the steps by clicking on the job link in the "Jobs" folder "Collect User Share" to open



When you have the job open, click "Add Step" and select "Run Harvest"



Configure the job step as follows:

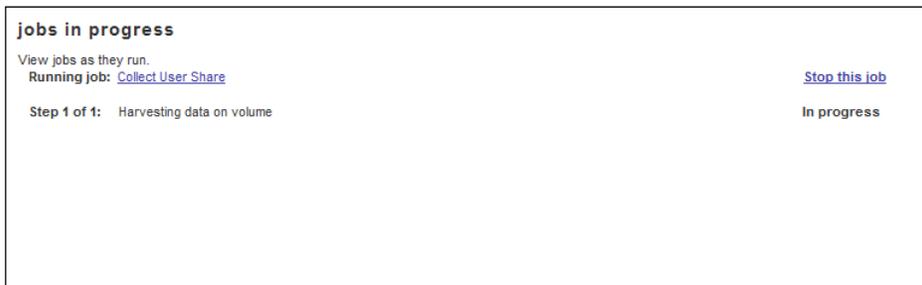


- Harvest these volumes:** EVSERV2:Users (This will be the File Share Primary Volume that you define in previous steps)
- Harvest type:** Run a full harvest (An incremental can be run to collect any delta information since the last collection)

4. Accept remaining defaults and click "ok"
5. Now you are ready to run the harvest job. Click "Start Job"



6. You may proceed to the "Dashboard" menu to monitor progress.



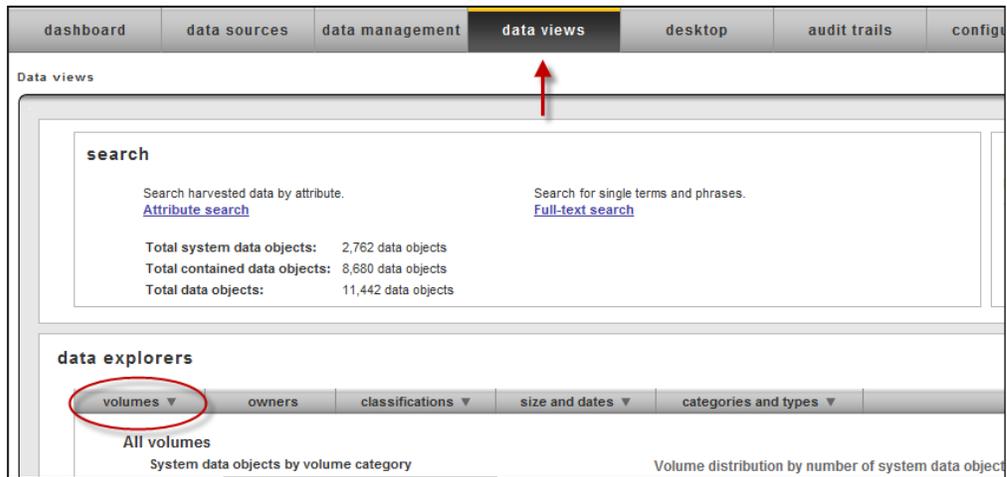
Once the harvest has completed, you can browse the index to get a feel for what is on the volume. Follow the steps below to browse the index.

### Browsing the Index

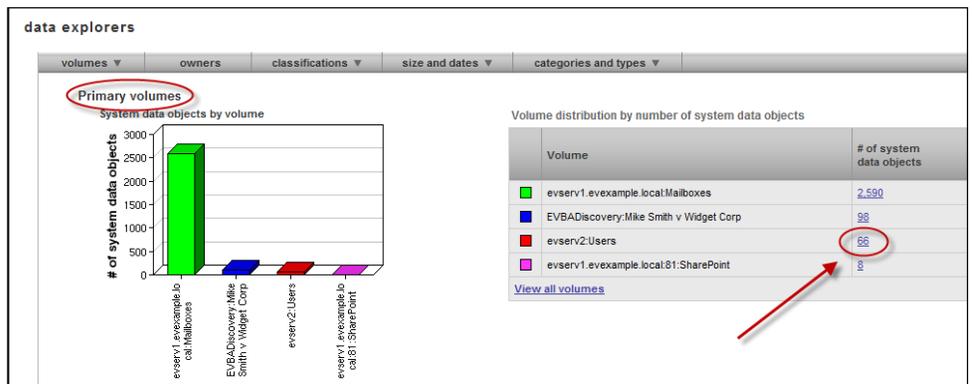
Once the harvest has completed, you can perform data analysis to ensure that you have collected the correct data set. This analysis will prepare you for classifying in preparation for importing the responsive data into Enterprise Vault for Legal Hold.

1. While logged into Discovery Collector, select the "Data Views" menu, and click on the "Volumes" link and select "Primary Volumes" in the "Data Explorers" section.

# Discovery Collector Getting Started Guide

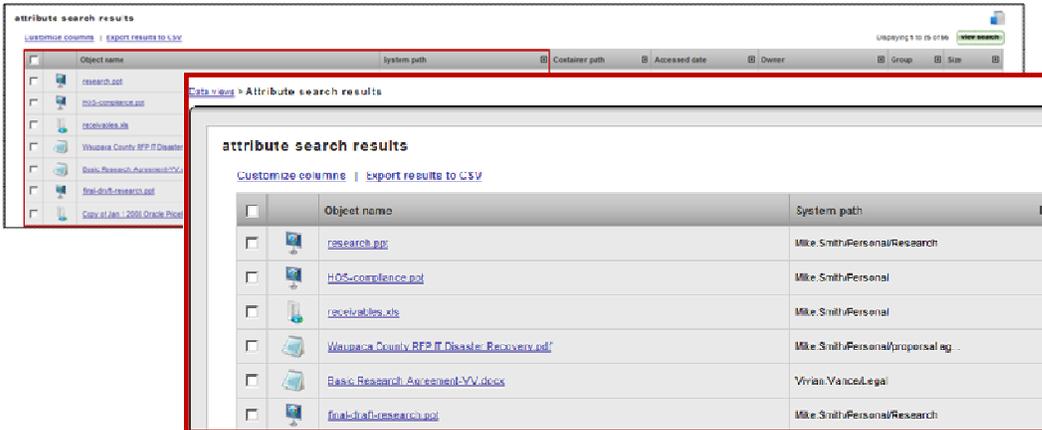


2. Once you are on the "Primary Volumes" page, there will be statistical information displayed about the data that has been indexed from each volume. Click on the link for the volume that you have indexed.

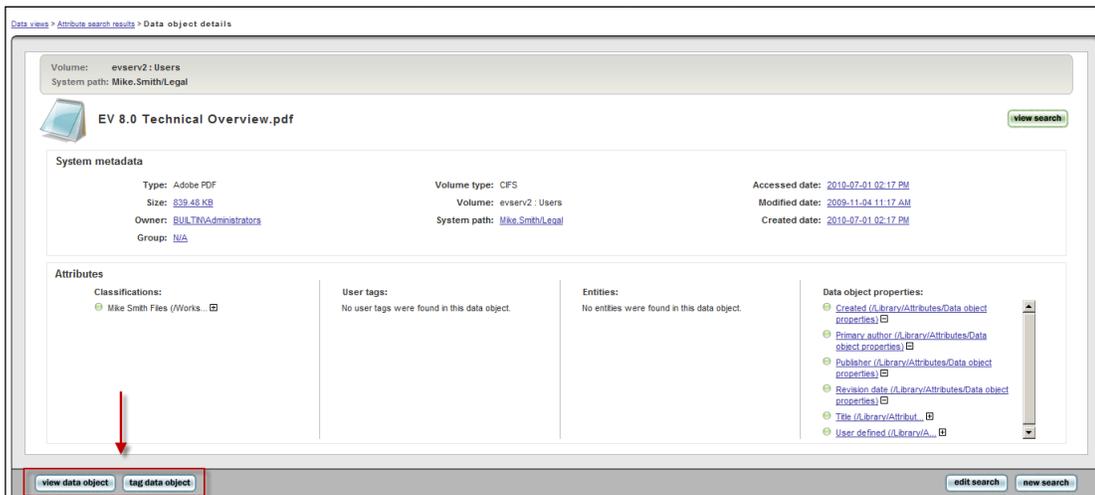


# Discovery Collector Getting Started Guide

- Now you will see a list of the objects that have been indexed identifying the owner, size, exact folder where they are located, and the date they were last accessed. This is helpful in determining where the bulk of the responsive data may be located.

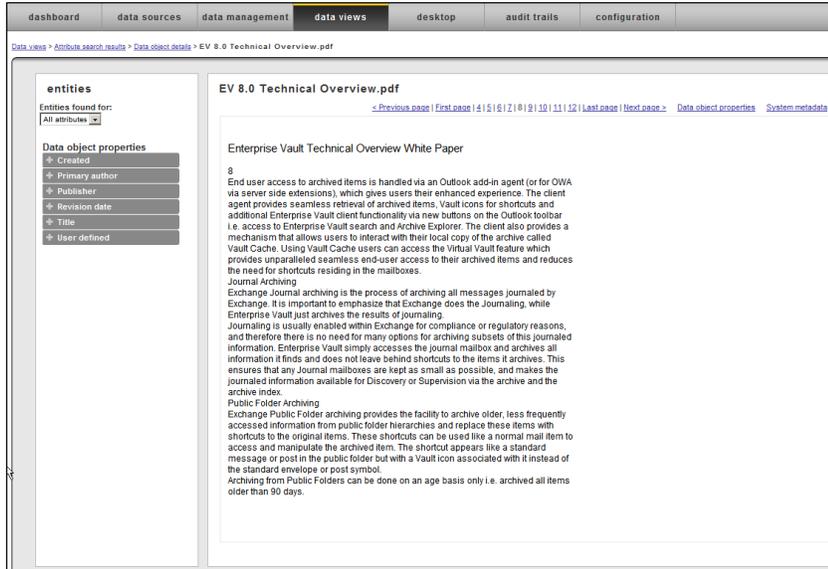


- If you click on the link for the object, more detailed information for that object will be displayed as seen below. Notice there are buttons to "View" or "Tag" the object; click on the "View Data Object" button.



## Discovery Collector Getting Started Guide

- From here, you can look at an HTML version of the object to get a feel for the context and even add a tag to be used when classifying the object.



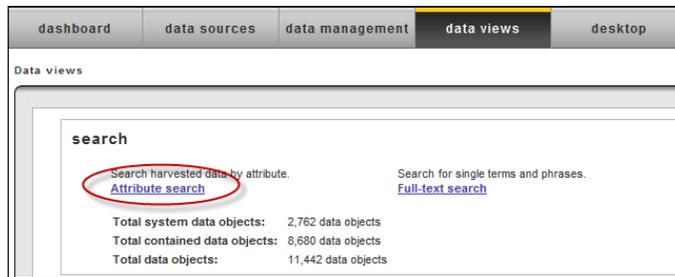
For the purpose of collecting this data for searching in Discovery Accelerator, this functionality will not be used. For more information about the data analysis features, please refer to the user guide in the product documentation.

## Classification the Volume Data

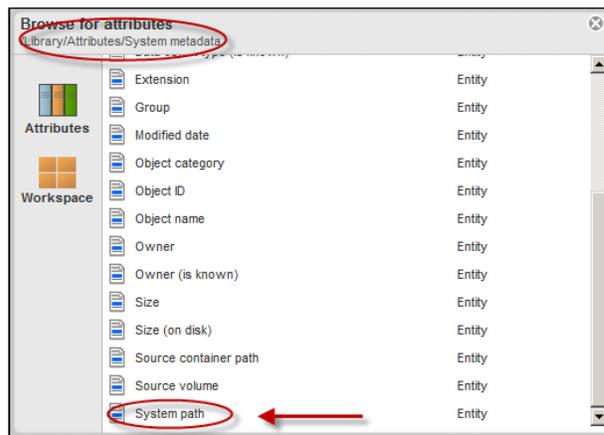
Now that you have built an index of the targeted volume, the next step is to classify the data. The classification of the data will allow you to import to Enterprise Vault based on certain attributed of the data object.

In the steps below, we will create a simple classification so that objects in a specific user folder can be imported to Enterprise Vault.

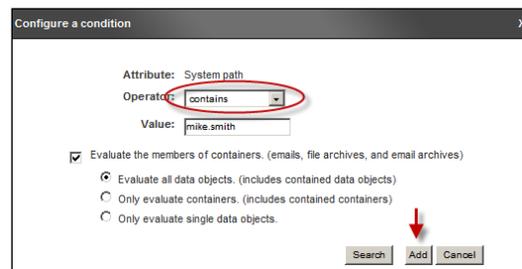
1. While logged into Discovery Collector, select the "Data Views" menu, and then select the "Attribute Search" link in the "Search" section of the page.



2. Next we will build a query against the index. In the "Define a Condition" box, click the "browse" button and select system metadata → system path



3. Once the condition configuration page opens, use the following settings:



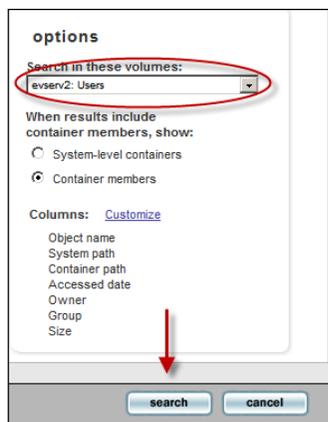
# Discovery Collector Getting Started Guide

- **Operator:** contains
- **Value:** mike.smith
- **Evaluate the members of containers:** selected, evaluate all data objects

Click "Add" and verify the condition is now part of the query.



4. Next, while on the same page, access the "Options" section on the far right. Then select your volume or in this example "EVSERV2: Users" from the drop down menu



Click "search" to see results

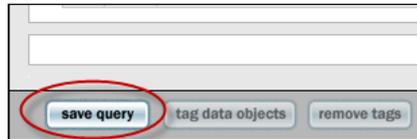
attribute search results

[Customize columns](#) | [Export results to CSV](#)

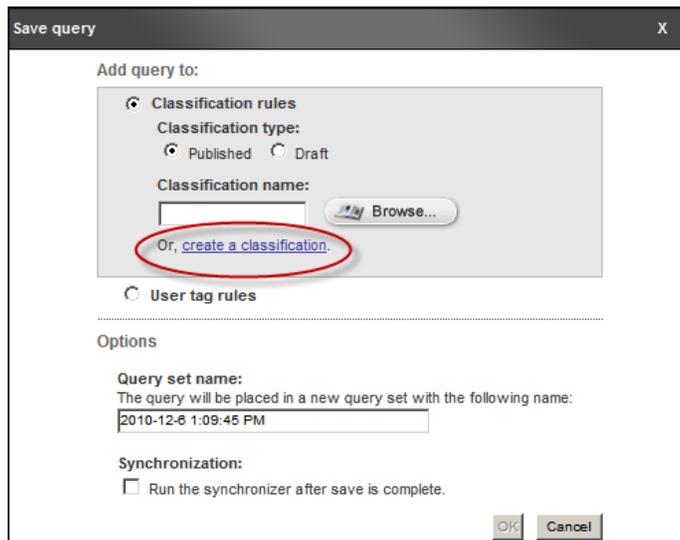
<input type="checkbox"/>	Object name	System path
<input type="checkbox"/>	<a href="#">save.zip</a>	Mike.Smith.Personal
<input type="checkbox"/>	<a href="#">RS-QOS.ppt</a>	Mike.Smith.Personal/Research
<input type="checkbox"/>	<a href="#">Copy of Jan 1 2008 Oracle Pricelist EU.xls</a>	Mike.Smith.3 Years
<input type="checkbox"/>	<a href="#">Best Practice for Upgrading Enterprise Vault-2009-11-04.docx</a>	Mike.Smith.Legal
<input type="checkbox"/>	<a href="#">EV 8.0 Technical Overview.pdf</a>	Mike.Smith.Legal
<input type="checkbox"/>	<a href="#">NBU-EV-Agent-Overview.pdf</a>	Mike.Smith.Legal

## Discovery Collector Getting Started Guide

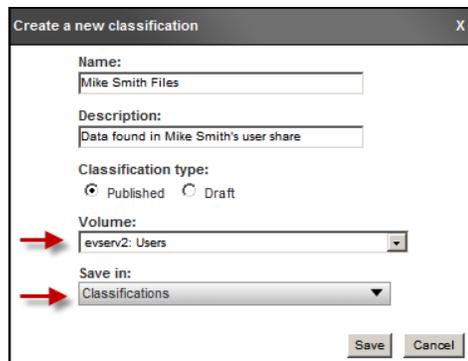
5. Notice that in this example, only the objects that were in the "Mike.Smith" user folder were returned in this query. Now scroll to the bottom left of the page and Click "save query" on the lower left to create the classification.



6. Next, select the radio button "Classification Rules" leaving the default of "Published", click "Create a classification" link



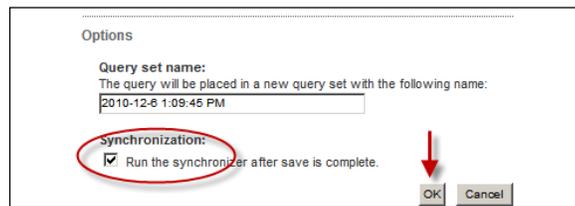
Then use the following settings:



- **Name:** Mike Smith Files
- **Description:** Data found in Mike Smith's user share
- **Classification Type:** Published

- **Volume:** EVSERV2:Users (This should be the volume you built the index for in previous steps)
- **Save In:** Mike Smith v. Widget Corporation (This will be the case folder that you save the work in)

Click "Save"



- Select "Run the synchronizer after save is complete"

Click "ok" twice, then scroll up and select the "dashboard" tab to see synchronizer progress.

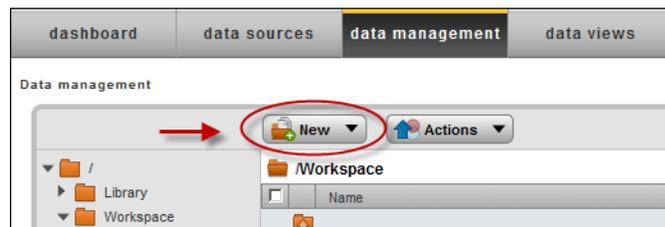
All of the objects found with that query will now have a classification attribute assigned so that policies may be run against the data; more specifically a policy to copy that specific data into Enterprise Vault.

## Import the Volume Data to Enterprise Vault

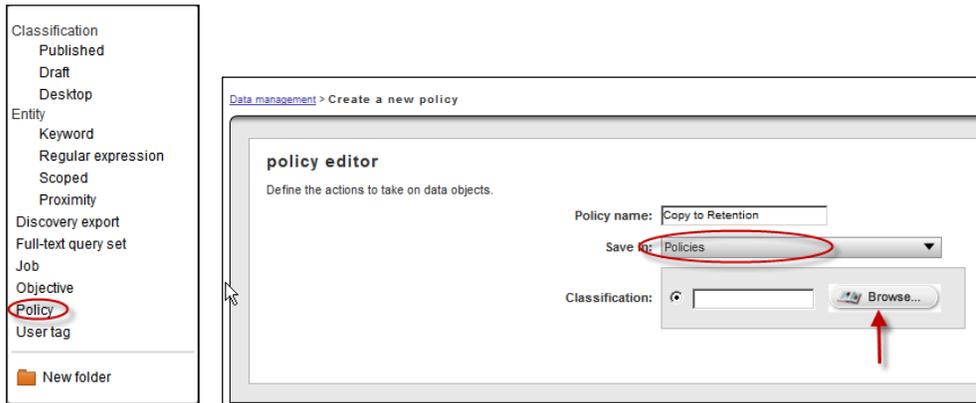
Now that you have created a classification and have synchronized with the volume index, the classification attribute can be leveraged to import the file server to Enterprise Vault. This section describes the steps need to copy the classified data from the file server directly to the archive in Enterprise Vault created during the configuration of the Retention Volume.

Follow the steps below for importing data:

1. While logged into Discovery Collector, select the "Data Management" menu and click on the "New" button underneath



Select "Policy" from the drop down menu

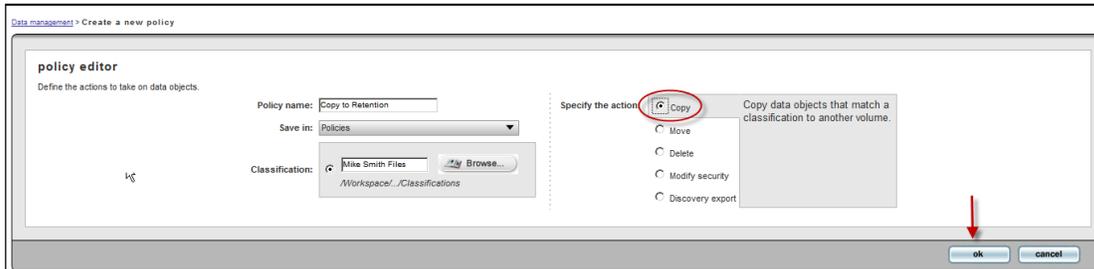


2. Once the Policy Editor is displayed, you will name the policy something relevant and recognizable; in this example we used "Copy to Retention" since we will be copying data to the retention volume. Save the policy in the "Policies" folder created underneath your case folder and click browse for the classification.

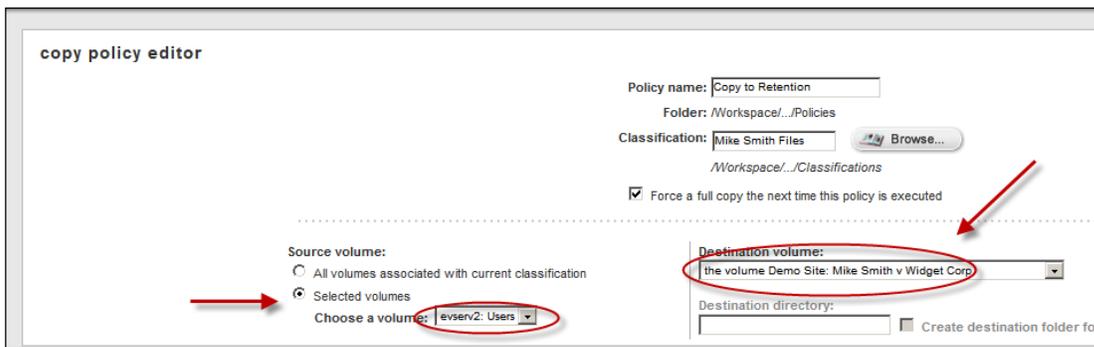


Drill down to the "Classifications" folder underneath your case folder and select the classification you created in the previous section. In this example, it is named "Mike Smith Files"

3. Once you have selected the classification, the Policy Editor will be displayed. Notice now that the "Specify the Action" section is activated as seen below.

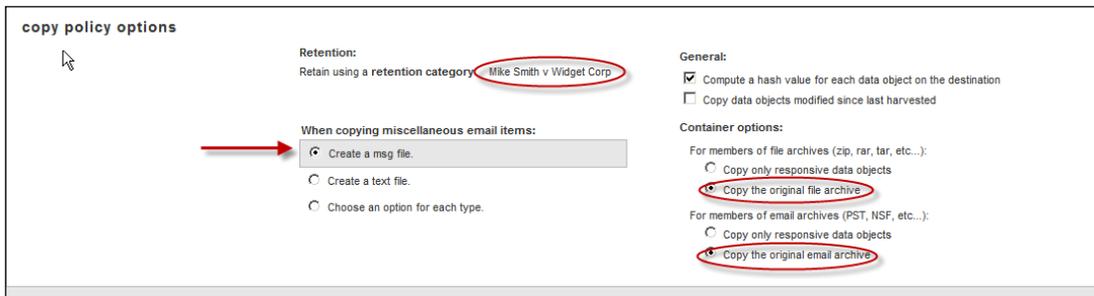


Choose "Copy" and click "OK". With the "Copy Policy Editor" displayed, you can configure the volume information



- **Source Volume:** Choose the volume that you created in previous steps
- **Destination Volume:** Choose the retention volume that you created for the case. This volume is mapped to the case archive in Enterprise Vault.

4. Once you choose the destination volume, the lower half of the page will be displayed to configure copy policy options.



- **Retention:** If not already selected, choose the retention category that will be associated with the case data. Typically this will be a dedicated retention category named for the case itself.
- **Miscellaneous Email:** Choose how to handle email items found on the volume you are copying data from. You can get granular and select different options for appointments, contacts, tacks, etc.

- **General:** You can compute a hash for each object at the destination before copying. Keep in mind; this will affect the copy speed.
- **Container options:** This setting determines whether the entire container file (ZIP, PST, NSF, etc) into Enterprise Vault or only the selected files that were classified. In our example, the container file would have been found in Mike Smith's folder, so you would want the entire container file. If the container file was found on a Project share, then you would only want the individual files found within the container file.

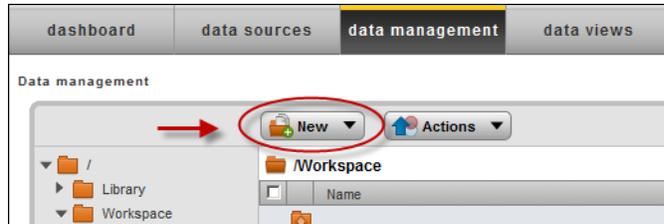
The policy has now been configured and requires that a "Job" be run to execute the policy.

## Create Job to Execute Copy Policy

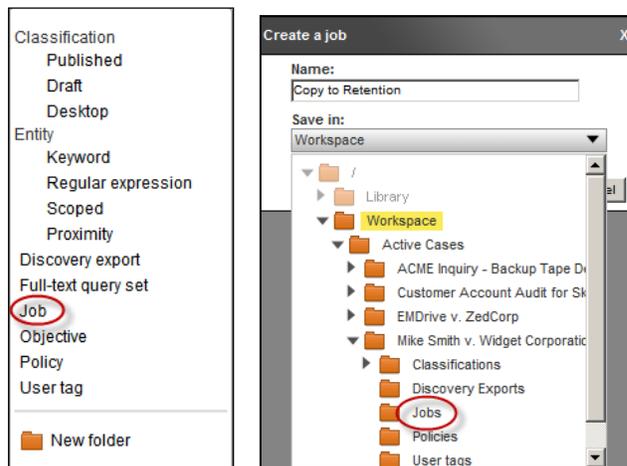
Now that you have the data classified and a copy policy configured; a job will be created to execute the policy. This will start the actual process of copying the data from the file server into Enterprise Vault.

Follow the steps below for creating a job:

1. While logged into Discovery Collector, select the "Data Management" menu and click on the "New" button underneath

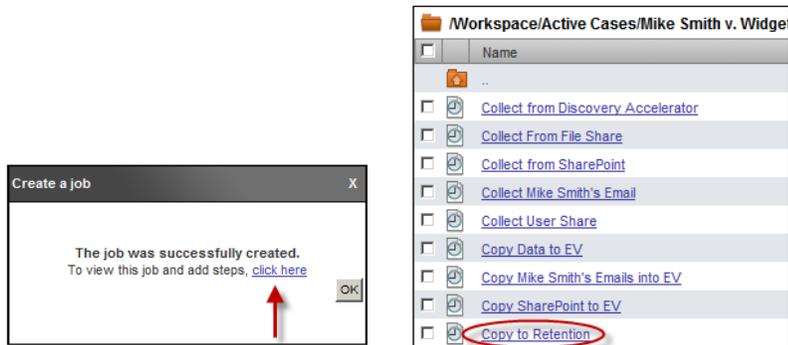


Select "Job" from the drop down menu; name the job something recognizable, and save it in the "Jobs" folder underneath your case folder.

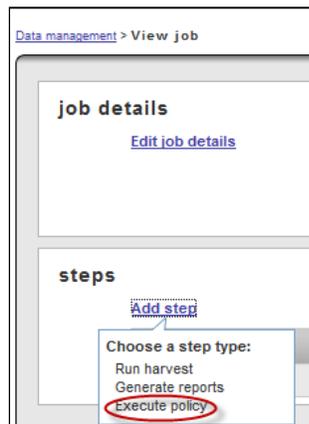


## Discovery Collector Getting Started Guide

- Once the job has been created, use the link “click here” to add steps. If you clicked “OK” instead, you can still configure the steps by clicking on the job link inside the “jobs” folder “Copy to Retention” to open



When you have the job open, click “Add Step” and select “Execute Policy”



The “Specify Policy Step Options” will be displayed; choose the suitable option for when to update the indexes

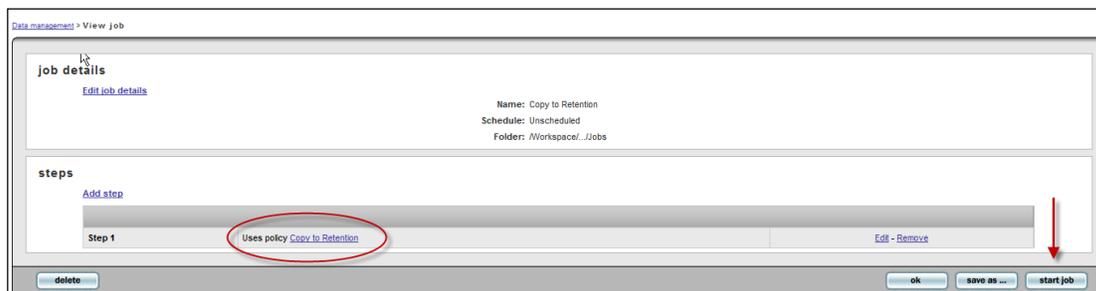


## Discovery Collector Getting Started Guide

Next click "Browse" to locate your policy, and then click "OK"



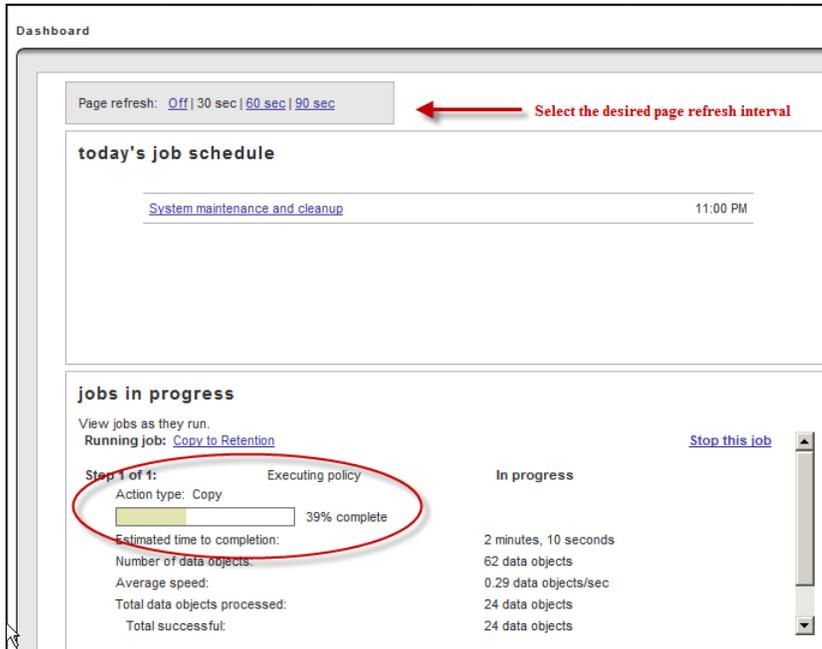
3. Confirm that all is configured and ready to go; the name of the policy should be displayed in the job steps. You can begin the import by clicking the "Start Job" button as seen below.



4. You can proceed to the dashboard to monitor progress



If you do not see progress, ensure that the "Page Refresh" is turned "On" as seen below



Once the job has successfully completed, depending on how much data was copied into Enterprise Vault, you should be able to search using Discovery Accelerator. It is best practice to wait before searching; for example, if 20 GB of data was copied to Enterprise Vault, then you should wait the amount of time it typically takes your environment to process that amount of data before searching in Discovery Accelerator.

## Summary

In conclusion, this document has focused on the basics of using Discovery Collector to collect data from server platforms. High level details needed to configure data sources and destinations should provide the reader with enough information to ready the system before training has occurred. We also covered the basics of indexing, classifying, and importing the data into Enterprise Vault for preservation, search, and review using Discovery Accelerator. While this document is not a replacement for formal training, it should enable you and your team to begin using Discovery Collector today.

For more detailed information, please refer to the User Guide which ships with the product documentation.

## About Symantec

Symantec is a global leader in providing security, storage and systems management solutions to help businesses and consumers secure and manage their information. Headquartered in Mountain View, Calif., Symantec has operations in 40 countries. More information is available at [www.symantec.com](http://www.symantec.com).

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation  
World Headquarters  
350 Ellis Street  
Mountain View, CA 94043 USA  
+1 (650) 527 8000  
1 (800) 721 3934  
[www.symantec.com](http://www.symantec.com)

Copyright © 2009 Symantec Corporation. All rights reserved. Symantec and the Symantec logo are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.