

# Coveo Enterprise Search 6.1

## Microsoft Exchange Connector

Coveo's *Microsoft Exchange* connector allows users to index *Microsoft Exchange* content. By using a unique special administrative account, all the mailboxes and public folders can be indexed.

### Features

The *Microsoft Exchange* connector can be used to index mailboxes. It allows *Coveo Enterprise Search* (CES) to crawl and index the content of a *Microsoft Exchange Server 2003* or *2007* installation.

Furthermore, it offers a more efficient search than *Microsoft Exchange*'s built-in search components.

The following details the features available in the *Microsoft Exchange* connector:

- Mailbox indexing;
- Public Folder indexing;
- **Security:** CES indexes security by combining *Microsoft Exchange* permissions set in *Microsoft Exchange* by users and Active Directory by administrators;
- **Live indexing:** Allows the connector to periodically query *Microsoft Exchange* for the latest edits, keeping the index content up to date;
- The *Microsoft Exchange* connector collects a maximum of metadata for emails, contacts, tasks and calendars;
- The *Microsoft Exchange* connector supports form-based authentication, which should be used when the OWA server is configured to use it. Two types of form-based authentications are supported:
  - Standard Exchange Form Based Authentication;
  - ISA Server Form Based Authentication.

### Requirements

The *Microsoft Exchange* connector supports the standard *Microsoft Exchange* server configuration, as described in the following article: [http://technet.microsoft.com/en-us/library/aa997482\(EXCHG.65\).aspx](http://technet.microsoft.com/en-us/library/aa997482(EXCHG.65).aspx).

### Installation

#### ▶ How to Setup an Administrative Account

In order to crawl multiple mailboxes using a single source, it is important to setup a special account in the *Microsoft Exchange Server*. This account must be given read access on all the mailboxes to index, as well as the public folders if necessary.

For more information on how to setup such an account, click the following link for the *Microsoft Exchange Server 2003* setup <http://support.microsoft.com/default.aspx?scid=kb;EN-US;821897>.

As for the *Microsoft Exchange Server 2007* setup, the following procedure gives sufficient rights to a user to crawl using WebDave or WebServices.



To setup the user correctly, open the *Exchange Management Shell* and use the following command :

```
get-mailbox -server ServerName -ResultSize Unlimited | Add-MailboxPermission -User 'domain\user' -AccessRights 'FullAccess'
```

The *Microsoft Exchange* connector supports the common implementations of front-end and back-end server suggested by *Microsoft*. For more information, refer to [http://technet.microsoft.com/en-us/library/aa997482\(EXCHG.65\).aspx](http://technet.microsoft.com/en-us/library/aa997482(EXCHG.65).aspx) and [http://technet.microsoft.com/en-us/library/bb124804\(EXCHG.65\).aspx](http://technet.microsoft.com/en-us/library/bb124804(EXCHG.65).aspx).

▶ **How to Configure a User Identity based on an Administrative Account**

Once the connector’s administrative account has been setup, it is necessary to configure a CES user identity:

1. In the Administration Tool, access the **Administrators** page (Configuration > Security).
2. In the left navigation pane, click **User Identities**. The **User Identities** page is displayed.
3. Click . The **Modify User Identity** page is displayed.
4. In the **Name**, **User** and **Password** fields, enter the credentials of the administrative account.
5. Click  **Save**.

<b>Name</b>	<input type="text" value="companyname/jsmith"/>
<b>User</b>	<input type="text" value="companyname/jsmith"/>
<b>Password</b>	<input type="password" value="•••••"/>
<b>Option</b>	<input type="checkbox"/> Support basic authentication
<b>Client certificate</b>	<input type="checkbox"/> Use client certificate
<b>Store</b>	<input type="text" value="Personal"/>
<b>Certificate</b>	<input type="text"/>


**Configuration**

The *Microsoft Exchange* connector has two entries in the **Source Type** drop-down list of the **Add Source** page: *Exchange (WebDav)* and *Exchange (WebServices)*. Both are roughly equivalent; however, it is important to define which source type is used at the creation of the source. Once created, it is not possible to change to the other method—to do so, another source has to be created.

The *WebDav* is faster than the *WebServices* connector, and is also compatible with *Microsoft Exchange Server 2003*. However, it can be harder to configure it correctly at the beginning because of form authentications and similar issues. Also, 2007 servers must have the 2003 Compatibility mode installed. The *WebServices* can only be used with *Microsoft Exchange 2007 Servers*, but usually does not need any configuration to work.

▶ **How to Index a Microsoft Exchange Source**

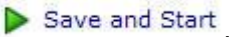
To use a single source to index all *Microsoft Exchange* mailboxes, enter a single URL in the **Addresses** box that stops at the mailboxes root (i.e. <https://owa.acme.com/exchange>). To index all public folders, enter a URL that stops at public folders root (i.e. <https://owa.acme.com/public>).


1. In the Administration Tool, access the **Sources and Collections** page (Index> Sources and Collections).
2. In the **Sources** section, click . The **Add Source** page is displayed.
3. Enter the appropriate values of the selected *Microsoft Exchange* server.

Field	Description	Example
Name	Any descriptive name.	All Microsoft Exchange mailboxes
Source Type	The connector used by this source.	Exchange
Addresses	List of starting points for the connector, one address per line.	Addresses are in the form of: <a href="https://servername/exchange">https://servername/exchange</a> <a href="https://servername/exchange/mailboxname">https://servername/exchange/mailboxname</a> <a href="https://servername/public">https://servername/public</a> The connector supports both http and https Ex.: <a href="https://owa.acme.com/exchange/johnsmith">https://owa.acme.com/exchange/johnsmith</a>
Max Number of Retries*	Number of retries to perform when a request fails due to a Microsoft Exchange Server <i>downtime or temporary unavailability</i> .	The default value is 3.
Retry Delay*	Delay between each retry attempt.	The default value is 30 seconds.
Exchange Folder	Name of the folder hosting all the mailboxes.	The default value is <i>exchange</i> . It corresponds to the segment of the URI that is after the server name and points to the mailboxes: <a href="https://servername/exchange/">https://servername/exchange/</a> .
Use Form Authentication	Whether to use form based authentication or not.	The default value is <i>False</i> . Necessary when <b>Form Based Authentication</b> is enabled on the <i>Microsoft Exchange</i> server. Two form authentications are supported: <ul style="list-style-type: none"> <li>• Standard <i>Microsoft Exchange</i> form authentication.</li> <li>• ISA server form authentication: When an ISA filter is put on front of the <i>Microsoft Exchange</i> OWA server.</li> </ul>
Index Exchange Security	Whether to index the securities specified in <i>Microsoft Outlook</i> or not.	The default value is <i>True</i> . Useful to prevent <i>Coveo</i> from indexing user defined securities. Only Active Directory Securities are used.
Index Active Directory Contacts	Whether to index contacts from Active Directory or not.	The default value is <i>False</i> . Useful to let everyone see the contact information available in Active Directory.
Public Folder	Name of the public folder.	The default value is <i>public</i> . It corresponds to the segment of the URI that is after the server name and that points to public folders: <a href="https://servername/public/">https://servername/public/</a> .
Index Junk E-mails	Whether to index junk emails or not.	Junk emails are indexed if the checkbox is selected.
Skip Archived Items	Whether to skip items archived by <i>Symantec Enterprise Vault</i> or not.	Useful to make the <i>Microsoft Exchange</i> connector ignore email stubs left behind by <i>Symantec Enterprise Vault</i> . The default value is <i>true</i> .

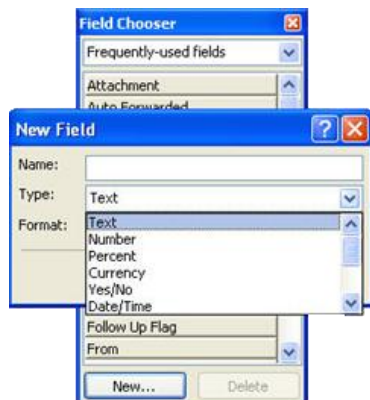
<p>Callback Service Url*</p>	<p>When installed, the Exchange callback service enhances the live indexing performance on <i>Microsoft Exchange 2007</i>. See the Live indexing: Using the Exchange Callback Service section.</p>	<p>Url of the Coveo Exchange Callback service: <a href="http://servername:8181">http://servername:8181</a>.</p>
------------------------------	--	---

\* These parameters are optional.

4. Click  .

The following custom source parameters are not displayed in the source. They must be created or added to the connector in the **Modify the parameters of the additional connector** page (Administration Tool > Configuration > Connectors > Additional Connector > Symantec Enterprise Vault >  ):

- **LDAPSearchRoot (string):** Root of the search in Active Directory for the current domain. For more information, refer to the appendix on LDAP Searches.
- **LDAPFilters (string):** Filter applied to the search of Active Directory in order to find mailboxes. For more information, refer to the appendix on LDAP Searches.
- **MaxNumberPrefetchChanges (integer):** Maximum number of items that can be stored in the prefetch cache, this cache is useful to reduce the delay between each query to *Microsoft Exchange*. The default value is 200. Increasing this value increases performance, but also the amount of memory used by the connector. If the connector is running on a server with limited memory resources, lowering this number to 100 or 50, can prevent out of memory issues.
- **Batchsize (integer):** Number of items to fetch per request made to the *Microsoft Exchange Server*. The default value is 50. The minimum value is 1 and the maximum 512. A small value forces the connector to make small but frequent queries to *Microsoft Exchange*. On the other hand, a bigger value leads to scarcer connections, but longer ones.
- **IndexJunkattachments (boolean):** When this parameter is set to *true*, the attachment of junk emails is indexed, otherwise it is ignored. The default value is *false*.
- **SkipEmbeddedImageAttachment (boolean):** When this parameter is set to *true*, the images embedded directly within the text of an email (like signature images) are skipped, otherwise they are indexed. The default value is *true*.
- **UserDefinedFields (string):** This parameter can be added in order to retrieve user defined fields from *Microsoft Exchange* Items. This string is in the following format: *Fieldtype:FieldName;FieldType:FieldName*, etc. The FieldType value is taken directly from the *Microsoft Exchange* Form Designer. The FieldName value is the name of the field as entered in the Form Designer.



- **ConnectionLimit (integer):** This parameter can be added in order to increase performance. If you plan on increasing the number of threads in the connector, you will also need to increase this value. The suggested value is the sum of all the threads you want to be able to connect to *Exchange*, in all your sources. Increasing this value will increase the load on the *Exchange* Server. The default value is 2.
- **IndexOutbox (boolean):** This parameter can be added in order to crawl the outbox folder. This folder usually only contains mails temporarily until they are sent by the *Exchange* server. The default value is *false*.
- **IgnoreUnresolvedDeniedSecurities (boolean):** This parameter can be added in order to allow the connector to keep crawling even after finding denied securities that could not be resolved. This behavior could create a security hole, so this parameter should be used with caution. The default value is *false*.
- **ResolveDisplayValuesWithContacts (boolean):** This parameter can be added to allow the connector to perform a resolution of the email address values found on emails with the mailbox' contacts. For example, if the connector finds the 'To' value "jsmith <jsmith@acme.com>", it will go look into the contacts folder of this mailbox, and try to form a new value using the first and last name of the contact: "John Smith <jsmith@acme.com>". The default value is *false*.

<b>Name</b>	<input type="text" value="All Microsoft Exchange mailboxes"/> ?
<b>Source Type</b>	<input type="text" value="Exchange (WebDav)"/>
<b>Addresses</b>	<input type="text" value="http://servername/exchange"/> Depends on the additional connector used. One entry per line.
<b>Rating</b>	<input type="text" value="Normal"/> ?
<b>Document Types</b>	<input type="text" value="Default"/> ?
<b>Fields</b>	<input type="text" value="Default Scheme"/> ?
<b>Refresh Schedule</b>	<input type="text" value="Every day"/> ?
<b>Max Number of Retries</b>	<input type="text" value="3"/> ?
<b>Retry Delay</b>	<input type="text" value="30"/> ?
<b>Exchange Folder</b>	<input type="text" value="exchange"/> ?
<b>Use Form Authentication</b>	<input type="checkbox"/> ?
<b>Index Exchange Security</b>	<input type="checkbox"/> ?
<b>Index Active Directory Contacts</b>	<input type="checkbox"/> ?
<b>Public Folder</b>	<input type="text" value="public"/> ?
<b>Index Junk E-mails</b>	<input checked="" type="checkbox"/> ?
<b>Skip Archived Items</b>	<input checked="" type="checkbox"/> ?
<b>Option</b>	<input checked="" type="checkbox"/> Index subfolders ? <input type="checkbox"/> Index the document's metadata ? <input checked="" type="checkbox"/> Generate a cached HTML version of indexed documents ? <input type="checkbox"/> Open results with cached version ?
<b>Authentication</b>	User Identity <input type="text" value="(none)"/>

▶ **How to Index All Mailboxes using a Single Source**

To index all *Microsoft Exchange* mailboxes using a single source, enter a single URL in the **Addresses** box that stops at the mailboxes root (i.e. <https://owa.acme.com/exchange>). To index all public folders, enter a URL that stops at public folders root (i.e. <https://owa.acme.com/public>).

## Appendix

### ▶ LDAP Searches

LDAP Searches are an important part of the *Microsoft Exchange* connector. For example, when the starting address <https://servername> or <https://servername/exchange> is used, a query is performed against the Active Directory of the current domain in order to find out the mailboxes available.

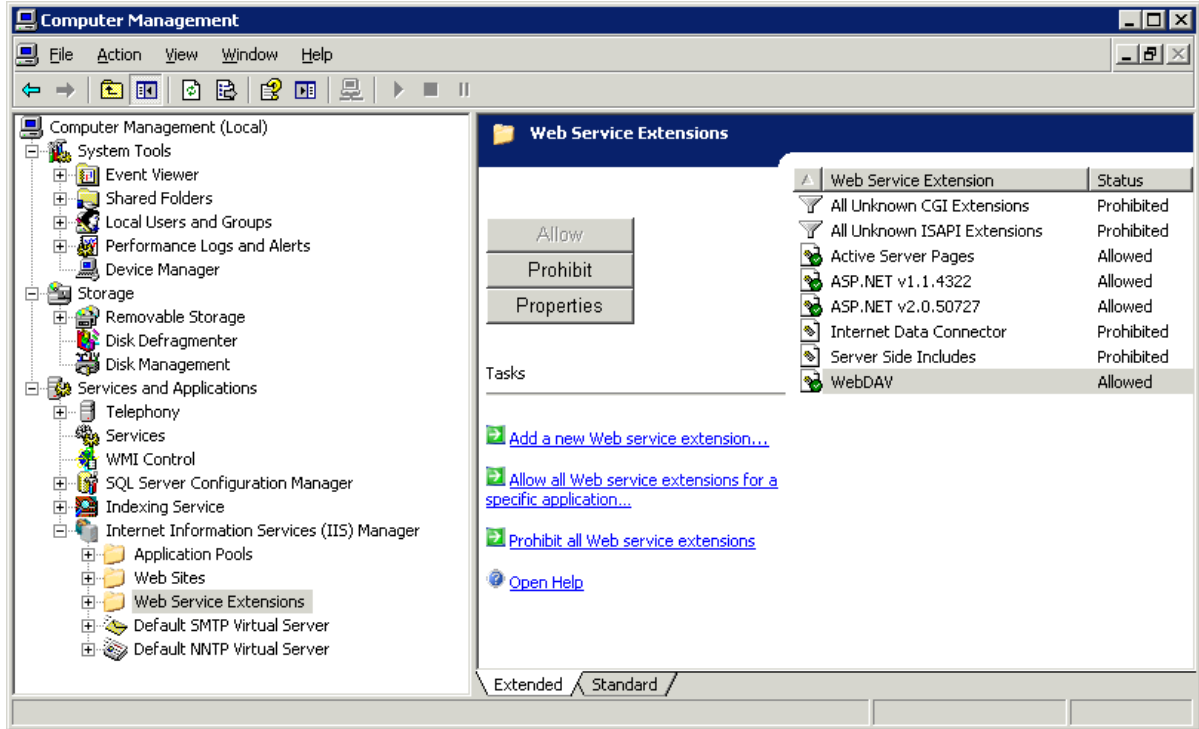
The results should reflect all the users in Active Directory that own a mailbox. Depending on the configuration of the Active Directory, it is possible the results will not match the results anticipated. Two parameters of the connector allow to fine tune the search.

- **LDAPSearchRoot:** Specifies to the connector where to start looking in Active Directory. When this parameter is not specified, the connector looks at the root of Active Directory, which can be extremely large. By specifying a value, it is possible to refine the search. For example, to search only within the OU “companynameOU” within the domain “corp.companyname.com”, enter: “LDAP://OU=companynameOU, DC=corp, DC=companyname, DC=com”.
- **LDAPFilters:** Allows to add properties to the results the connector is looking for. The default filter is: “(&(mail=\*)(objectclass=user)(objectclass=person))”. This AND operation is performed on a few properties in order to find the most mailboxes without receiving any unwanted ones. If this filter is not specific enough, it is possible to complete the filter by adding a value to it. The value entered is added to the AND operation on the filter. For example, if the value “cn=JohnSmith” is entered in the parameter, the resulting filter is: “(&(mail=\*)(objectclass=user)(objectclass=person)(cn=JohnSmith))”.

### ▶ Starting Addresses for Microsoft Exchange Server 2007

In *Microsoft Exchange Server 2007*, the URI used to connect to OWA through the browser is no longer the same as the one used by the connector. *Microsoft* has performed several modifications on OWA and the virtual sites for *Microsoft Exchange*. While the URI used to connect to OWA is different, the starting address for the connector does not change. The connector communicates with the *Microsoft Exchange* virtual site in IIS; therefore, the starting addresses are not different.

Furthermore, the *Microsoft Exchange* virtual site can no longer be accessed by default. It is important to install the *2003 Compatibility* module on the *Microsoft Exchange Server 2007*; hence recreating the 2003 virtual sites. By default, the WebDav protocol is prohibited in IIS. It is important to make sure it is allowed.



► **Live indexing: Using the Exchange Callback Service**

It is possible to enhance the live indexing performance by using the *Coveo Exchange Callback Service*. The service uses the *push subscription* mechanism of *Microsoft Exchange* to index changes on mailboxes and public folders.

Note that this service is only available with *Microsoft Exchange 2007*.

**Installing the service**

The service can be installed by running CES installation kit, and selecting the **Advanced** option.

**Configuring the service**

Although the service is configured to run correctly with the default parameters, some of them can be configured manually to override out-of-the-box behaviors. Parameters can be edited in the *ExchangeCallbackService.exe.config* file located in the bin folder of CES. The custom parameters are located in the **configuration/appsettings** section.

```

<?xml version="1.0" encoding="utf-8"?>
<configuration>
  <appSettings>
    <add key="MaxConnections" value="6"/>
  </appSettings>
  <system.serviceModel>
    <bindings>
      <wsHttpBinding>
        <binding name="WSHttpBinding_INotif.
    
```

The different parameters that can be configured for the service are listed in the table below.

Field	Description	Example
LogFileLifeTime	Duration before a log file will be automatically deleted by the service.	Default is 30 days.
LogMaxSize	Maximum size before a log file will be splitted.	Default is 10 mega bytes (10 485 760 bytes)
MaxConnections	Maximum number of concurrent connections that the service will create when calling the Exchange web services.	Default is 4 connections.
ShowAllCounters	Displays the usage statistics about the service in the CES console and logs before each live indexing run.	Default is false.
DataFilePath	Location of the data files used by the service.	By default, data files are located in the Index/Crawlers/ExchangeCallbackService folder, under the folder where is located the CES files (config, index, logs, etc.)
LogFilePath	Location of the log files created by the service.	By default, log files are located in the same folder than the CES log files.
SubscriptionTimeout	Duration before an Exchange subscription will be considered invalid. This value should not be changed unless there are issues using the default value.	Default is 30 minutes. Maximum allowed is 1440 minutes.

HostName	Host name that will be used internally to build the web service url. Changing this value affects directly the <b>Callback Service Url</b> parameter that is set on an Exchange source.	Default is the computer name.
PortNumber	Port number that will be used internally to build the web service url. Changing this value affects directly the <b>Callback Service Url</b> parameter that is set on an Exchange source.	Default is 8181.