d.velop enterprise search for OpenSearchServer: Administrator

Table of Contents

1. d.velop enterprise search for OpenSearchServer

1.1. Basic information on the application

This chapter contains general product information and conventions regarding documentation.

1.1.1. About d.velop enterprise search for OpenSearchServer

An ECM system enables structured data storage. As a rule, information objects are not only present in structured in your ECM system, but are also created in other systems and platforms. For example, the correspondence in your Exchange server also contains a lot of information in e-mails that could be relevant for a business transaction. With enterprise search, you can overcome data silos and find information not just in your ECM system, but in your heterogeneous business landscape using providers.

The d.velop enterprise search for OpenSearchServer search provider enables searches of any websites and file systems.

1.2. Installing and Uninstalling

This chapter provides information on the installation of d.velop enterprise search for OpenSearchServer.

1.2.1. System requirements

Pay attention to the general system requirements for d.velop products.

1.2.2. Installing the OpenSearchServer search provider

You install the software exclusively using d.velop software manager. If an application is required for different products, the corresponding software packages are also installed automatically.

For further information on installing the software, see the d.velop software manager manual.

After the installation, call the **Search provider** dialog in the configuration to refresh the list of available search providers. Call this dialog again every time the configuration is changed in the search provider to update the configuration.

1.2.3. Uninstalling the OpenSearchServer search provider

The software you installed using d.velop software manager can only be uninstalled with d.velop software manager. If the software to be uninstalled has dependencies with other software packages, you must resolve these conflicts accordingly.

For further information on uninstallation, see the d.velop software manager manual.

1.2.4. Installing updates of the OpenSearchServer search provider

You can only update the software using d.velop software manager.

For further information on updates, see the d.velop software manager manual.

1.3. Configuring the OpenSearchServer search provider

This chapter contains further information on the settings and configuration options.

1.3.1. Creating and configuring a searchable index

This section describes how to create an index in OpenSearchServer, that can then be searched.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.
2. In the **Create a new Index** section, enter the desired name of the new index in the **Index name** field.
3. Choose one of the following templates:
   - **file crawler:** Choose this template if you want to add files from a file server to the index.
   - **web crawler:** Choose this template if you want to add websites to the index.
4. Confirm creation by clicking **Create**.

The index is now created and is listed in a table.

When you have created the index, you must customize the index schema to enable the search.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.
2. Choose an index from the table.
3. Switch to the **Schema** tab.
4. Enable the **Fields** tab.
5. If you use an index for files from a file system, set the **TermVector** value for the following fields to **yes**:
   - **userAllow**
   - **userDeny**
   - **groupAllow**
   - **groupDeny**

> **Note**
> Create the **lastEditDate field**. The lastEditDate field is used when searching as a primary date field. This means that the content of this field is displayed for a result as the **Last update**. Do not use a default.ini. Use the following properties:
>
> - **Indexed: yes**
> - **Stored: no**
> - **TermVector: no**
>
> If you do not create a **lastEditDate** field, the standard fields **fileSystemDate** and **contentUpdateDate** are used to display the **last edit date**.

1.3.2. Considering User Permissions During a Search

This section explains how you can configure the system so that user rights are taken into account for SMB/CIFS connections.

> **Warning**
> User rights can only be taken into account for SMB/CIFS connections.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.
2. Choose an index from the table.
3. Switch to the **Schema** tab.
4. Switch to the **Authentication** tab.
5. Enable the selection box **Enable authentication**.

6. Chose the created index.
7. Under **User allow field,** enter the field **userAllow**.
8. Under **User deny field**, enter the field **userDeny**.
9. Under **Group allow field**, enter the field **groupAllow**.
10. Under**Group deny field**, enter the field **groupDeny**.
11. Under **Default group**, enter **"any"**.
12. Leave the field **Default user** blank.

> **Note**
> The Documentation of OpenSearch contains further information on compliance with permissions. When you have followed all steps, the permissions have already been checked.

1.3.3. Configuring a Crawler for a Local File Directory

This section describes how you can create a crawler for SMB/CFID sharing in your OpenSearchServer installation.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.
2. Choose an index from the table.
3. For the creation of a specific directory, switch to the **Crawler** tab.
4. Click **Files > Locations**.
5. Click **Create new location**.
6. Choose the types **SMB/CIFS.**
7. In the fields **User name** and **Domain**, enter the corresponding values of a user who has access to all documents in the share.
8. Choose the option **File Permission** in the **Security Permissions** field to index permissions.
9. In the **Password** field, enter the password of the user.
10. In the **Host** field, enter the computer name for the share.
11. In the **Path** field, enter the relative path for the shared directory.
12. Enable **With sub directories.**
13. Enable **Enabled**.
14. Switch to **Field mapping**.
15. In the first selection list, choose the indexed field **fileSystemDate**.
16. In the second selection list, choose **lastEditDate**.
17. Click **Add** to add the mapping.
18. Switch to **Crawl process** and start generating the index. There is the option **run once** for a one-time entry or **run forever** for continuous indexing.
19. Under **File browser**, you can view and check already indexed files.

1.3.4. Configuring a Crawler for SMB/CIFS File Sharing

This section explains how you can create a crawler for a file directory in your OpenSearchServer installation.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.
2. Choose an index from the table.
3. For the creation of a specific directory, switch to the **Crawler** tab.
4. Click **Files > Locations**.

5. Click **Create new location**.
6. Now choose the type **Local files.**
7. Enter the desired path in the **Path** field.
8. Enable **With sub directories.**
9. Enable **Enabled**.
10. Confirm the settings by clicking **Save**.
11. Switch to **Field mapping**.
12. In the first selection list, choose the indexed field **fileSystemDate**.
13. In the second selection list, choose **lastEditDate**.
14. Click **Add** to add the mapping.
15. Switch to **Crawl process** and start generating the index. There is the option **run once** for a one-time entry or **run forever** for continuous indexing.
16. Under **File browser**, you can view and check already indexed files.

1.3.5. Configuring a Crawler for a Website

This section describes how you can create a crawler for a website in your OpenSearchServer installation.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.
2. Choose an index from the table.
3. To create a specific website, switch to **Crawler**.
4. Choose **Web > Pattern List**.
5. Enter the websites to be indexed. If you also want to index the subdomains of a page, add an asterisk, e.g.: **(www.example.com/*)**.
6. Enable **Enabled**.
7. Switch to **Field mapping** and assign the corresponding indexed field**contentUpdateDate**to the **lastEditDate** field.
8. Switch to **Crawl process** and start generating the index. There is the option **run once** for a one-time entry or **run forever** for continuous indexing.

> **Note**
> If the subdomains are not indexed, try the following settings:
>
> 1. Go to **Schema > Parser list**.
> 2. Choose **HTML parser**.
> 3. Set **Ignore non-canonical pages** to **False**.
> 4. Restart the indexing.

1.3.6. Creating the Access Data for Your OpenSearchServer Installation

You can restrict access to the OpenSearchServer installation by setting up users. The following section describes how to set up the users and access the API key for the connection.

> **Warning**
> Make sure that you have already created an administrative user before you create a non-administrative user. Otherwise, you will lose administrative access to the OpenSearch-Server installation. You can find further information here.

This is how it works

1. Call the web interface of your OpenSearchServer installation. This interface can usually be accessed via port 9090.

2. Switch to **Privileges**
3. Create a new user. Call it dvelopenterprisesearch, for example.
4. The API key is generated automatically upon creation of the user. Copy the user name and API key for later use (e.g. in an editor).
5. Choose an existing index that this user is allowed to search.
6. For a search, choose the entry **Index: query the index**.
7. Click **Add** to add the selection of the index and the role to the user.
8. Click **Create** to create the user.

You have successfully created a user that you can use in the configuration of the OpenSearchServer search provider of the enterprise search.

1.3.7. Adding an OpenSearchServer Connection

This section explains how you can configure the connection for an OpenSearchServer installation.

This is how it works

1. In the Start section, click **Configuration**.
2. Under **Enterprise search**, select the entry **Search provider**.
3. Click the cog icon in the group **OpenSearchServer** and select the entry **OpenSearchServer connection**.
4. Click the plus character to add your OpenSearch installation.
5. Enter a descriptive name.
6. Specify a URL for an icon with which this connection should appear. Use the file extension **.png** or **.jpg**. Of course, you can also use the default URL.
7. Enable the option **Take user rights into account for display** to check access to the document when displayed.
8. Specify the base address for your OpenSearchServer installation. This installation usually accesses port 9090.
9. If your OpenSearchServer installation is secured for user accesses, specify the access data (user name and API key) that has already been created.
10. Choose the index for the connection.

You have successfully created the OpenSearchServer search provider. Enable the search provider in search provider management so your users can use the search provider.

You can deploy static content and access it via the enterprise search in a web-based format. For example, you can deploy your own image files or logos to customize the appearance according to your own needs.

This is how it works

1. Open the installation directory of d.velop enterprise search. The path is usually:

```
C:\d3\d.velop enterprise search
```

2. Create a **files** subfolder. The path is usually as follows:

```
C:\d3\d.velop enterprise search\files
```

3. Store the desired files in the **files** subfolder.

You have now stored your files and made them accessible in a web-based format via d.velop enterprise search. The files can be called in the browser with the following path:

```
https://<Base address>/enterprisesearch/static/*
```

When you deploy a file **Logo.png** in the folder **C:\d3\d.velop enterprise search\files** , the logo can be accessed in the browser with the following call:

```
https://<Base address>/enterprisesearch/static/Logo.png
```

1.3.8. Blocking Content when Websites are Crawled

Websites often contain elements, such as their navigation, that appear on every subpage and that are thus not considered desirable information where searches are concerned. OpenSearchServer allows you to block elements during the crawling process.

This is how it works

1. Choose a web crawler index in the OpenSearchServer configuration dialog.
2. Click **Schema > Parser list > HTML parser**.
3. Edit the parser in the area **XPATH Exclusion** by blocking an area with a selector. To do this, specify the selector of a CSS class, for example, to ignore the entire surrounding block. The documentation of OpenSearchServer contains instructions on how to do this.

Example: //*[@class="mega-menu-container"]

> **Warning**
>
> The area you select in this manner will be completely ignored. Any URLs it contains, such as URLs in navigation menus, will then no longer be processed, and the crawler may not search all existing websites.
>
> If possible, enter a site map URL from which the crawler can dynamically request all URLs to be searched on the tab **Crawler > Site Map**. Many content management systems provide a site map URL.

1.4. Frequently Asked Questions

This chapter answers frequently asked questions.

1.4.1. Why Can't I Find Partial Terms During Searches?

The enterprise search for OpenSearchServer currently only supports searches with whole terms.

1.4.2. Where Can I Find the Official Website of OpenSearchServer?

The official page can be reached at http://www.opensearchserver.com/.

1.5. Additional information sources and imprint

If you want to deepen your knowledge of d.velop software, visit the d.velop academy digital learning platform at https://dvelopacademy.keelearning.de/.

Our E-learning modules let you develop a more in-depth knowledge and specialist expertise at your own speed. A huge number of E-learning modules are free for you to access without registering beforehand.

Visit our Knowledge Base on the d.velop service portal. In the Knowledge Base, you can find all our latest solutions, answers to frequently asked questions and how-to topics for specific tasks. You can find the Knowledge Base at the following address: https://kb.d-velop.de/

Find the central imprint at https://www.d-velop.com/imprint.