PASSAGEWAYS

OnSemble

More Than An Intranet

dtSearch – Admin Guide and End User Guide

Table of Contents

# FAQ – Frequently Asked Questions

Q. How does dtSearch determine the order of my results?

A. This is a complicated answer. Be sure to take a look at the full explanation of how OnSemble determines the results order in the ***Search Result Order and Relevance*** sections below.

Q. How can I refine my search to find something more specific?

A. Use Document Tags. When viewing the properties of a document, you can add special keywords that are included in the searchable index. When selecting keywords, try not to use something that would be commonly found in other documents. As an example, our document called "2018 Account Policies" probably shouldn't include a keyword of "Policy". This would be a very common term likely to be found in many different documents in your portal. Instead, use a unique nickname for the document, like "GreenElephant". Sure, this seems silly, but can you name a single other document in your repository that has a special hidden keyword like this? Unique keywords will guarantee a much shorter list of results. Other unique examples might include random three-character codes.

Use the post-search filters. Once the search results are displayed, switch the category to Documents, and choose any of the filter dropdown menus to further refine your original search.
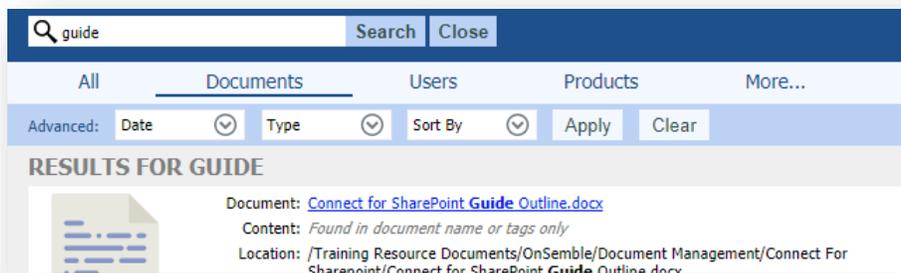
Q. I JUST added a new document/tag. Why isn't the search finding it?

A. The document or tag has not been added to the Search Index yet. The Search Indexer is an executable that runs on the web/application server. It is typically set up to run only a single time per day, so the document you've recently added will not be included in any search results until the indexer completes its next cycle.

Q. I just REMOVED a document/tag from OnSemble. Why is it still showing up in my searches?

A. This is the opposite of the previous FAQ question. A document or tag that has been recently removed is likely still in the Search Index. It will not be removed until the Indexer on the web/application server runs again.

Introduction

Welcome to the dtSearch Administration and Configuration Guide. This guide will show you how to set up and use the dtSearch application for document searching in OnSemble. Most of the content covered in this guide is geared towards users who have access to both the Web/Application Server and the dtSearch Module Tools section of OnSemble. However, OnSemble End Users will benefit from certain sections.

Core dtSearch Application Concepts

There are some core concepts and terms that you should be familiar with when dealing with the dtSearch application. These terms will give you a fundamental understanding of how the application operates and how each one can impact the portal overall.

Search Indexer

The Indexer is an executable program that runs on the Web/Application server as a frequent scheduled task. This program takes all of the documents that have been added to OnSemble and creates a searchable Index. Documents will not be able to be searched until the Indexer program has ran and added them to the index.

Search Index

Not to be confused with the Indexer, the Index is a series of files on the Web/Application server that contain OnSemble's searchable materials. As you add and remove documents from OnSemble, those documents will also be added and removed from the Search Index, but only as frequently as the Indexer's scheduled task is set to run.

Document Tags

To aid in searching for the correct document, Document Tags can be added to any individual document. For example, if you search the word "POLICY" and 30 results appear, you might still have a hunt on your hands. Adding special keywords to your documents can cut that list of results down significantly. For more information on how to add tags to documents, please refer to the Portal Document Management Guide.

Important Module Properties

Module Properties for dtSearch are found by navigating to Module Tools > dtSearch > Module Properties. These are special settings used to configure the application.

EnableFuzzySearch

This Module Property helps you find additional search results that are similar to your original search term. When this is enabled (set to TRUE), your search query will not need to be 100% exact. Searching for "paln" will still bring up results for "plan".

EnableStemming

Another setting that assists with grammatical variations. With this property enabled (set to TRUE), searching "apply" would also find results for "applying", "applied", and "applies".

ExcludeFilters

This setting helps keep your Search Index smaller, meaning it takes less time to run and takes up less disk space on the server. In addition, it keeps your search results from being cluttered up with document file types you don't care about. Adding filetypes to the exclusion list, such as *.mp3 means that your search results will not include any MP3 (audio) files.

Fuzziness

When the EnableFuzzySearch module property is enabled, the Fuzziness setting can determine how far off a search term can be and still return possible results. Fuzziness can be set anywhere between 1 and 10, but anything above 3 will likely return many unwanted results. With a Fuzziness of 3, searching "alpahket" would return results for "alphabet" (3 misplaced letters).

MaxFilesToRetrieve

Setting a number here will limit the amount of results returned in a single search. If your search would typically return with results containing 5000 different items, and your MaxFilesToRetrieve setting is configured to only return 20 of them, you'll see the top-most 20 results.

SearchIndexPaths

After the Search Indexer has completed its process and generated the searchable index, that index is placed in a particular file path location on the server defined by the configuration files of the Indexer. The SearchIndexPath Module Property is used to let OnSemble know where to find that index file on your network.

SearchType

For this setting, we have the choice of two different values: ANYWORDS or ALLWORDS.

Using ANYWORDS will look for results that include any single one word in your search query. Searching "2018 Expense Reports" will pull any result with the word "reports" in it.

Using ALLWORDS will look for results that match every word in the query at least once. Searching "2018 Expense Reports" will pull results that have a match with all three separate words ("2018", "expense", and "report") in any way and in any order.

> ***Pro Tip:*** Regardless of the SearchType setting, you may perform a search for an entire exact phrase by using quotation marks around your search query.

Setting up the dtSearch Indexer

Configuring the dtSearch Indexer requires basic knowledge of Microsoft's Task Scheduler, and access to the OnSemble's application server. Generally, this step is completed by someone in the IT Department.

1. Locate the appropriate files on the Application Server.
The files we will be using are generally found in the following location on the application server: C:\Program Files\Passageways\DTSearch\.
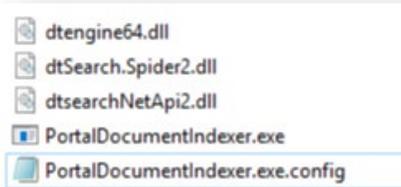
2. Open and Edit the Configuration File.
Use Notepad to open the PortalDocumentIndexer.exe.config file. Very specific information will need to be keyed into this file. Change the following locations and then save and close the file:

add UniqueNameNoSpaces="OurPortalName" – Change "OurPortalName" to be the unique name of your instance of OnSemble.

PortalUrl="http://OurPortal/" – Change "OurPortal" to be the correct URL used to access your instance of OnSemble.

IndexPath="C:\Inetpub\wwwroot\Passageways\Modules\dtSearch\Index\" – This file path can likely remain the same, as long as the file path exists. This location is where the Index will ultimately live and is also the location that needs to be entered into the SearchIndexPath Module Property described earlier in this guide.

Domain="OurDomain" – Change "OurDomain" to be the name of the domain where your installation files exist.

Username="IdentityAccount" – Change "IdentityAccount" to the username who is set as the Portal's Identity Account. If you're unsure who this is, it can be found by navigating to Reports > Portal Framework > System Summary and looking at the field: Portal Account the Portal is Running Under. (Username Only. Omit the Domain Name)
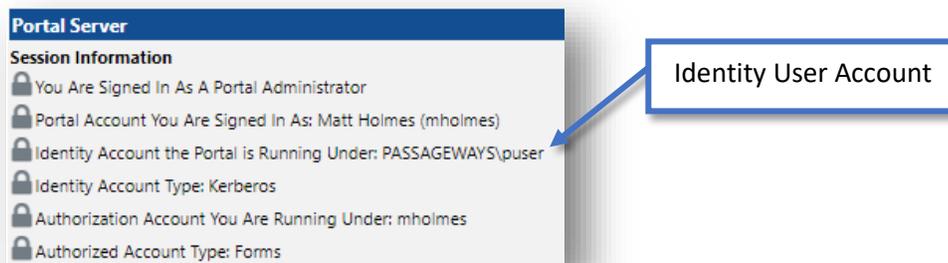
Password="Password" – Change the password to be the Windows Credentials password of the Identity User Account listed above.

3. Create a new Scheduled Task.

On the OnSemble application server, use the Windows Task Scheduler to create a new automatically recurring task. More information on how to create this task can be found in Microsoft documentation found HERE.

When configuring the Triggers, Actions, Conditions and Settings of this new task, we recommend the following:

- Run the task daily. Set it to begin after normal business hours.
- Have the task run the PortalDocumentIndexer.exe file shown in the screenshot on page 6.
- Run the task as the Portal Identity Account User, described on page 6.
- Run the task with the highest available privileges.
- Run the task whether the user is logged in or not.
- Wake the computer to run the task.
- Allow the task to be ran on-demand.
- Stop the task if it runs longer than 3 days.

After these three steps are finished, you can manually run the task for the very first time or wait until the next scheduled time for it to run automatically. Keep in mind that you will not be able to return any search results until the Indexer has finished running.

> **_Pro Tip:_** The Search Indexer task run time will vary drastically, depending on how many documents are in your Portal Document Repository. The more documents, and the larger they are, the longer (and larger) the Index will be.

Using dtSearch – Searching for Documents

Document Searching happens in the OnSemble Toolbar. Simply key in your search terms into the search bar and hit the enter key or click the eyeglass icon. *(If you do not see the search bar, contact your Portal Administrator and ask them to enable a toolbar layout that includes it.)*

Enter Search Term

Execute Search

guide

Filter by Documents Only

RESULTS FOR GUIDE

Page: Guides and Rulers
Content: *Found in title or tags only*
Location: TRAINING - 6.2 Webinar\Guides and Rulers   Page Tags: *Page has no tags*
Created On: 12/3/2018 11:02:48 AM
Modified On: 12/3/2018 11:02:48 AM

Page: Guides and Rulers
Content: *Found in title or tags only*
Location: Matt Holmes\Guides and Rulers
Created On: 11/19/2018 1:18:07 PM
Modified On: 11/19/2018 1:18:07 PM

Search Results

Document: Connect for SharePoint Guide Outline.docx
Content: *Found in document name or tags only*
Location: /Training Resource Documents/OnSemble/Document Management/Connect For Sharepoint/Connect for SharePoint Guide Outline.docx
Modified On: 10/28/2015

Top Results

Once you've completed a search and clicked on one of the results, you'll begin to see an additional panel in the search results area. **Top 5 Results**, and **My Top 5 Results**. Sometimes, these two categories will contain the same set of documents. These areas indicate frequently accessed documents in your portal. For this list to change, it does require you to actually click on one of the results in the list. Simply performing the search alone is not enough to change the Top 5 Results list.

Top 5 Results

This category will show the top 5 most frequently accessed documents accessed across all users in the entire organization.

My Top 5 Results

This category will show the top 5 most frequently accessed documents accessed by you, and you alone.
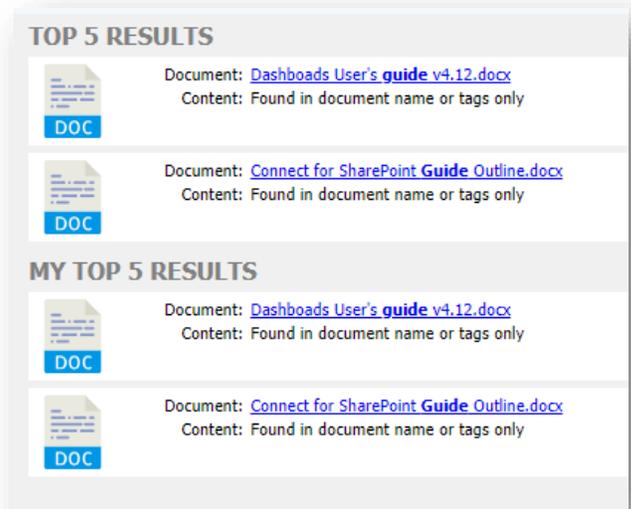
Search Result Order

OnSemble's dtSearch application will comb through many different areas of a document to try and find what you're looking for. Document title, document contents, and tags are all captured by the search indexer and are completely searchable. However, occasionally, the order in which the results display can be a little confusing. Below, you'll find information about how the order is calculated and how to better find what you're looking for on the first try.

Search Relevance

Each different searchable area of a document (title, body, tags, etc…) has been appropriately weighted relative to its importance. This value is called "Relevance" and it is capped at 100 points. For example, the document title is more important than the document's contents (in terms of search), so the title is given a much higher Relevance value than a single instance of the search term found in the body of the document. This means that when performing a search, you will generally see results that match the title before ones that match any of the content. The more times a search match is made in a single document, the higher it will appear in the search results list.

Here is a chart that describes what values are given to each searchable document object:

Search Match:	Relevancy Point Value
Document Title (Exact Match)	100pts
Document Title (Begins with search query)	+25pts
Document Title (Contains search query)	+25pts
Document Tag	+15pts
Document Content	+15pts

With the information shown in this chart, and the knowledge that the Relevancy value caps itself at 100 points, we can start to understand what to expect when performing searches. Here are some examples:

Example #1 – The General Expectation
Searching "Administrator's Guide", you might expect to see an order of the following:

- A document with the title of "Administrator's Guide" (exact match) will be assigned 100 points.
- A document with the title of "Administrator's Guide 2.0" (beginning match) will be assigned 25 points.
- A document with the title of "Portal Administrator's Guide" (contained match) will be assigned 25 points.
- A document containing a single instance of "Administrator's Guide" in the body will be assigned 15 points.
- A document with the document tag of "Administrator's Guide" will be assigned 15 points.

Example #2 – More Common than Uncommon
Searching "Administrator's Guide", the following can also occur:

- Regardless of the document's title, a document with 7 different instances of "Administrator's Guide" in the body will be assigned 105 points (capped at 100pts). Meaning this document will likely appear near the top of the results list.
- Matching a document's tag of "Administrator Guide" is only valued at 15 points. However, if the search matches many other documents with multiple instances of the search term in the body, the document tag matched result will be pushed further down the list. This is the reason it is important to choose UNIQUE document tags that exist nowhere else in a document.

Example #3 – Maxing Everything Out / No Order
Searching "Administrator's Guide", this situation is also possible:

- Regardless of the document's title, 3 different documents that all contain many instances of "Administrator's Guide" in the body will be assigned hundreds of points, which are all capped at maximum of 100 points each.
- A document with the title of "Administrator's Guide" (exact match) will be assigned 100 points. This means that several seemingly unrelated documents all have the same relevancy value, and all appear at the top of the search results in a random order. **Even if a document title matches your search query 100%, it is not guaranteed to be the top result. The Relevancy value is the key to this order.**

Take Note!

Last Updated March 5, 2019