newgenONE

OmniDocs EasySearch

User Guide

Version: 12.1

newgensoft.com

Disclaimer

This document contains information proprietary to Newgen Software Technologies Ltd. User may not disclose or use any proprietary information or use any part of this document without written permission from Newgen Software Technologies Ltd.

Newgen Software Technologies Ltd. makes no representations or warranties regarding any software or to the contents or use of this guide. It also specifically disclaims any express or implied warranties of merchantability, title, or fitness for any particular purpose. Even though Newgen Software Technologies Ltd. has tested the hardware and software and reviewed the documentation, it does not guarantee or imply that this document is error free or accurate regarding any particular specification. As a result, this product is sold as it is and user, the purchaser, is assuming the entire risk as to its quality and performance. Further, Newgen Software Technologies Ltd. reserves the right to revise this publication and make changes in its content without any obligation to notify any person, of such revisions or changes. Newgen Software Technologies Ltd. authorizes no Newgen agent, dealer or employee to make any modification, extension, or addition to the above statements.

Newgen Software Technologies Ltd. has attempted to supply trademark information about company names, products, and services mentioned in this document. Trademarks indicated below were derived from various sources.

Copyright © 2025 **Newgen Software Technologies Ltd.** All Rights Reserved.
No part of this publication may be reproduced and distributed without the prior permission of Newgen Software Technologies Ltd.

Newgen Software, Registered Office, New Delhi
E-44/13
Okhla Phase - II
New Delhi 110020
India
Phone: +91 1146 533 200
info@newgensoft.com

Contents

Preface **4**
Revision history 4
Intended audience 4
About this document 4
Documentation feedback 5
Introduction **6**
Searching folders and files using the Elasticsearch server **7**
Search result 8
Miscellaneous options 8
Document 9
Folder 9
Media 10
Advanced search 10
Fuzzy search 11
Proximity search 12
Boolean search 12
Phrase search 13
Searching folders and files without using Elasticsearch server **14**
Generating Q/A With NewgenONE Marvin **15**
Frequently asked questions **16**
General 16
Searching 19

Preface

This chapter provides information about the purpose of this guide, intended audience details, and revision history for NewgenONE OmniDocs EasySearch.

Revision history

Revision history	Description
December 2025	Initial publication

Intended audience

This user guide is intended for users who want to perform global searches and retrieve content by searching directly on specific keywords for quick reference or knowledge discovery. Readers must have a basic understanding of document operations, folder operations, and rights management. Readers must also belong to an administrative group with sufficient privileges to access and configure the required components for end users.

About this document

This guide enables you to easily search folders, documents, and media files. With a single point of contact, you can easily search any document and easily access all case-related documents. The platform supports multi-format content, including images, audio, video files, MS Office files, emails, and other large-size documents.

This guide is for business users and administrators assuming that you have a working knowledge of the following:

- The fundamentals and standard practices of your business area
- Basic understanding of Newgen OmniDocs
- Conceptual idea of ElasticSearch

Documentation feedback

To provide feedback or any improvement suggestions on technical documentation, write an email to docs.feedback@newgensoft.com.

To help capture your feedback effectively, share the following information in your email.

- Document name
- Version
- Chapter, topic, or section
- Feedback or suggestions

Introduction

EasySearch allows you to search documents, folders, and media files. It also simplifies searches with filters and sorting options.

Searching folders and files using the Elasticsearch server

Elasticsearch is a powerful open-source search and analytics engine commonly used to quickly index, search, and analyze large volumes of data. It is employed in applications where fast and efficient searching is essential.

This section describes the search operations that are supported when Elasticsearch is enabled.

The following filter options are available for searching folders and files:

Option	Description
DataClass Field Value	Enables you to search on a particular data field value of a dataclass.
Owner	Enables you to search for a particular owner.
Folder Name	Enables you to search on a particular folder name.
Document Name	Enables you to search for a particular document.
Global Indexes	Enables you to search on a particular global index.
Keywords	Enables you to search for a particular keyword.
Notes & Annotations	Enables you to search for a particular note or annotation.

When you initiate the search action, the search results appear. The results include all documents, folders, and media files that match the specified keyword.

Search result

Search Results lists the results based on the entered search criteria. You can apply filters to refine the result set.

The search results consist of the following filter tabs:

- Document
- Folder
- Media

Miscellaneous options

This section describes how to sort the search results.

You can sort search results based on different criteria. The following options are available:

Option	Description
Relevance	It sorts the search results based on the entered search query, with the most relevant result appearing at the top.
Recently Added	It sorts the search results based on the dates when files and folders were added. The most recently added appears at the top.
Modified on	It sorts the files and folders based on the modification date.
Owner: A to Z	It sorts the search results based on the owner's name in ascending order.
Owner: Z to A	It sorts the search results based on the owner's name in descending order.
Filename: A to Z	It sorts the files and folders based on their file name in ascending order.
Filename: Z to A	It sorts the files and folders based on their file name in descending order.

Document

The Document tab displays the searched document list. You can perform various operations such as viewing properties, setting alarms or reminders, and other operations on the document.

You can access additional operations for a document by using the available document actions. The list of document operations appears when you select the required option.

For details about document operations, refer to the documentation for your specific product version.

When you select the Document tab, additional filters become available. These filters help refine the search results. The following table describes the filters:

Filter	Description
Data Class	It filters search results based on the selected data class.
Keyword	It filters search results based on the selected keyword.
Doc Type	It filters search results based on the available document format.
Document Size	It filters search results based on the document size.

If any metadata is associated with respective search results, they appear as facets in the left pane under the Filters option.

Folder

The Folder tab displays the searched folder list. You can perform various operations such as viewing properties, setting alarms or reminders, and other operations on the folders.

You can access additional folder operations by using the available folder actions. The list of operations appears when you select the required option.

For more information about folder operations, refer to the documentation for your specific product version.

Media

The Media tab displays the searched media list. You can perform various operations such as viewing properties, setting alarms or reminders and other operations.

You can access additional media operations by using the available media actions. The list of media operations appears when you select the required option.

For more information about media operations, refer to the documentation for your specific product version.

When you select the Media tab, additional filters become available based on the search results. The following table describes the filters:

Filter	Description
Keyword	It filters media based on the selected keywords.
Media Size	It filters media based on the specified media size.
Duration	It filters media based on the specified duration.
Media Type	Filter out the media file based on the available media file format.

Advanced search

You can use Elasticsearch's advanced search techniques to retrieve relevant documents from the repository. Elasticsearch provides flexible and powerful search capabilities that allow you to perform a wide range of searches, including fuzzy, proximity, boolean, and phrase searches.

This section includes the following search options:

- Fuzzy search
- Proximity search
- Boolean search
- Phrase search

Fuzzy search

A fuzzy search is an approximate string-matching search. It matches the string with one edit distance for a given string. It allows you to obtain the right results with one character correction, insertion, deletion, or one character transposition.

A fuzzy search is useful when there are potential spelling mistakes, typos, or variations in the data you're searching for.

You can configure fuzzy search in the *easysearch.xml* file. Setting the tag value to 0 disables fuzzy search, while values 1, 2, and 3 enable fuzzy matching for one-character, two-character, and full-word variations respectively.

The cases with the above-stated string are as follows:

Case 1 – One character correction

For example, if you type a search string "Anelysis", the fuzzy search automatically returns the result matching the correct string "Analysis" with one edit by changing the 'e' at the third position to 'a'.

Case 2 – One extra character insertion

For example, if you type a search string "Analysisz", the fuzzy search automatically returns the result matching the correct string "Analysis" with one edit by removing the 'z' from the last position.

Case 3 – One character missing

For example, if you type a search string "Analysi", the fuzzy search automatically returns the result matching the correct string "Analysis" with one edit by adding the 's' at the last position.

Case 4 – One character transposition

For example, if you type a search string "Analyssi", the fuzzy search automatically returns the result matching the correct string "Analysis" with one edit by interchanging the characters 's' and 'i'.

Proximity search

Proximity search enables you to set the maximum edit distance of words in a phrase and how far apart the words can be from each other.

For example, if a document contains the phrase "It gives us vital information", the proximity search feature allows you to search for the phrase by specifying the distance between the words 'It' and 'information'. In this case, the distance is 3 as there are three words between them.

Proximity search has a specific syntax that requires keywords to be within double quotes and a number after the '~' symbol specifying the distance between the words. The syntax is "In Folder"~2.

To obtain an exact search result, you must enclose the keyword in double quotes " ". This is known as Phrase Search.

Boolean search

Boolean search allows you to search the required folders and files using the AND, OR, and NOT logical operators.

- The AND operator enables you to search for documents and folders that have all the search terms associated with them. For example, if you want only those documents and folders that contain the keywords “Project” AND “Product” associated with them.
- By using the OR operator, you can search for documents and folders that contain any of the search terms associated with them. For example, to search for all documents and folders that contain the keywords "Project" OR "Product".
- NOT operator excludes search on those keywords that are in the search criteria followed by NOT operator. For example, if you want to search only those documents and folders that contain the keyword "Project" NOT "Product".

Phrase search

Phrase search provides the flexibility to get the exact search result by inserting the keywords within double quotes " ". It is useful when you want to find documents that include a particular set of words in the exact order specified, ensuring more precise results.

Searching folders and files without using Elasticsearch server

This section describes how searching functions when Elasticsearch is not used. In this mode, searches are performed on document content or on metadata associated with documents and folders. However, certain advanced search functionalities, such as proximity search, phrase search, fuzzy search, boolean search, and synonym search, are unavailable to perform. Also, the facets do not appear to filter the search results.

The following filter options are available when searching without Elasticsearch:

Option	Description
Content	Allows you to search a document based on its content.
MetaData	Allows you to search a folder or document based on the metadata associated with the folder or document.
All	Allows you to search a folder or document based on both the content and metadata of the folder or document.

For more information about folder and document operations, refer to the documentation for your specific product version.

Generating Q/A With NewgenONE Marvin

NewgenONE Marvin is a Generative Artificial Intelligence (GenAI) tool used to generate questions and answers based on selected documents. It allows you to create a large set of questions and answer pairs from the available data, reducing the effort and time required compared to manually searching specific content or metadata.

When searching with EasySearch, the system also provides a direct option to perform Marvin operations on the selected file.

For details about using NewgenONE Marvin, refer to the documentation for your specific product version.

Frequently asked questions

This section contains the frequently asked questions on EasySearch.

General

1. What is an Elasticsearch server?
    - Elasticsearch is a search server based on Lucene. It provides a distributed, multitenant-capable full-text search engine with an HTTP web interface and schema-free JSON documents. Elasticsearch is developed in Java and released as open source under the Apache License.
2. What is an Ingestion Service?
    - An Ingestion Service retrieves content and metadata from various repositories and sends this data to search engines such as Elasticsearch. It also supports document-level security to ensure that only authorized content appears in search results.
3. How does the Elasticsearch server work?
    - Elasticsearch is a search engine based on Lucene. It takes data in JSON format and indexes it. Accessible through an extensive API, Elasticsearch powers fast searches that support data discovery applications.
4. What are the benefits of using Elasticsearch?
    - Elasticsearch is a highly scalable, open-source full-text search and analytics engine. It allows you to store, search, and analyze large volumes of data quickly.
        - Provides search results based on relevance.
        - Supports proximity search, wildcard search, Boolean search, and more.
        - Supports faceted search.

5. How is EasySearch different from the current search of OmniDocs?
    - EasySearch enhances the search capability of OmniDocs, making it more efficient. The major differences from the existing search are:
        - Supports faceted search. Facets narrow search results based on data class, keyword, file type, and global index.
        - Requires minimal input. The search runs throughout the document for specified keywords, eliminating the need for location-specific queries.
        - Orders search results based on relevance.
6. What is the relevance of documents?
    - The *_score* is a positive floating-point number that represents a document's relevance score. Documents with higher scores are more relevant. A query clause generates a *_score* for each document based on the following factors:
        - **Term frequency**: A term's relevance increases with its frequency in a field.
        - **Inverse document frequency**: Terms that appear in fewer documents are considered more relevant.
        - **Field-length norm**: Shorter fields, such as titles, assign more weight to terms compared to longer fields.
7. Is it possible to tweak relevance?
    - Yes. You can modify the relevance score by using different methods, such as score boosting and index-time field-level boosting.
8. Is it possible to change the facets classification?
    - Yes. You can manage the view of facets in the user interface by updating the name and location in the configuration file.
9. How do I boost metadata fields (such as document name, owner, or content) in Elasticsearch?
    - Add the metadata key in *EasySearch.xml* under the *BoostedFields tag*. You can add one or multiple fields.
      Example: *<BoostedFields>documentname,content</BoostedFields>*
10. How do I disable fuzziness in Elasticsearch?
    - Set the *Fuzziness* tag value to **0** in *EasySearch.xml*.
      Example: *<Fuzziness>0</Fuzziness>*

11. Where do I configure primary and replica shards?
    - You configure primary and replica shards at the time of Elasticsearch index creation. Update the *ESConfigurator/conf/Settings.json* file.
      Example:

```
"number_of_shards": 3, // Primary shards
"number_of_replicas": 1 // Replica shards
```

12. How do I enable rollover for Elasticsearch index creation?
    - Perform the following steps:
      1. Set *ESRolloverPolicy=Y* in *ESConfigurator/conf/ESConfig.ini*.
      2. Update *cabinet_name* with your CabinetName in *ESConfigurator/conf/RollOverTemplate.json*.
      3. Update primary and replica shard settings in *RollOverTemplate.json*.
13. Where do I define the rollover policy?
    - You define the rollover policy in *ESConfigurator/conf/RollOverPolicy.json*.

Searching

1. How to perform a search?
    - You can perform a search using the keyword in the search text box and press the 'Enter' key. Then, the documents related to that keyword appear as search results, and facets based on that keyword appear as search filters. For more information, see *Searching files and folders*.
2. What are the supported search operators types?
    - The EasySearch supports the following search operators:
        - Boolean (AND, OR)
        - Wildcard (*?)
        - Single and multiple words search
        - Phrase search
        - Fuzzy search
        - Proximity search
3. How to filter and refine search results?
    - Search results can be refined by selecting different fields or facets in the Advanced search section.
4. How to use Boolean operators with search text?
    - In case to search for a document having two or more keywords, then use AND or && operator while searching. For more information, see *Boolean search*.
5. What is multi-select faceting?
    - Selecting multiple values from the same facet results in an OR filter between each of the values: (facet1.value1 OR facet1.value2).
    - Faceting on more than one facet results in an AND filter between each facet: (facet1.value1 OR facet1.value2) AND (facet2.value1).
6. What are facets or faceted search?
    - A Faceted search helps you to explore large amounts of data by displaying summaries about various partitions of the data. Then, it further allows you to narrow down the navigation to a specific partition.

7. How to use wildcard queries?
    - It matches documents that have fields matching a wildcard expression. The supported wildcards are
        - * - It matches any character sequence, including the empty one.
        - ? - It matches any single character.
8. Is Phrase search supported? How can a user use this feature?
    - Yes, write the Keyword within double quotes like- “Account report”. For more information, see *Phrase search*.
9. What type of data or document formats can be searched using EasySearch?
    - You can search for documents, folders, images, audio, and videos.
10. Which ranking algorithms does EasySearch use?
    - BM25
    - Minimum field length (mfl)
    - TF or IDF (term frequency or inverse document frequency)