eServices Administrator's Guide

Configuring Analyzers for UCS Search

12/15/2025

Contents

Configuring Analyzers for UCS Search

Basic options controlling indexing and searching are described in the eServices 8.1 Reference Manual. This page describes one further option that controls the use of analyzers.

> ### Warning
> Any change in configuration of indexing must be followed by a reindexing of the full content of all UCS indexed objects.

The analyzers that are supplied with UCS are

- WhitespaceAnalyzer—Splits the text into tokens separated by white space characters (specifically, SPACE_SEPARATOR, LINE_SEPARATOR, PARAGRAPH_SEPARATOR, HORIZONTAL TABULATION, LINE FEED, VERTICAL TABULATION, FORM FEED, CARRIAGE RETURN, FILE SEPARATOR, GROUP SEPARATOR, RECORD SEPARATOR, or UNIT SEPARATOR).

- StandardAnalyzer—Converts the text to lower case, splits the text into tokens separated by the white space character, and removes high-frequency English words (called *stop words*).

- LowerCaseAnalyzer—Converts the text to lower case and splits the text into tokens separated by the white space character.

- SimpleAnalyzer—Divides text at non-letters and converts to lower case. This works well for languages in which words are separated by spaces, such as most European languages, but is of little use for languages in which words are not separated by spaces, such as many Asian languages.

- KeywordAnalyzer—Treats the entire text as a single token. This is useful for data like zip codes, IDs, and some product names.

In the default case, UCS search uses the StandardAnalyzer for all fields in all tables in the database. To override the default analyzer, use the following option.

<table_name>-field-analyzer<any>

Optional: Yes

Default value: StandardAnalyzer

Valid values: See below

Changes take effect: After restart

Sets the analyzer used for any table or field. In the option name, `<table_name>` is one of the following tables in the UCS database:

- Callback
- Chat

- CoBrowse
- Contact
- ContactAttribute
- EmailIn
- EmailOut
- Interaction
- PhoneCall
- StandardResponse

<any> can be anything, including zero. Use it to differentiate among multiple `field-analyzer` options referring to the same table.

Values for this option have the general form

<field>=<analyzer>, <field>=<analyzer>, ...

where <field> is the name of a field in the table and <analyzer> is the name of a supported and installed analyzer. For example:

- Option name: `interaction-field-analyzer`
- Option value: `Text=GermanAnalyzer,StructuredText=StandardAnalyzer`

With this option name and value, when searching the `Interaction` table, the search operation applies GermanAnalyzer to the Text field and StandardAnalyzer to the StructuredText field.

You can achieve the same result by creating two options:

- `Name:interaction-field-analyzer-01,value:Text=GermanAnalyzer`
- `Name:interaction-field-analyzer-02,value:StructureText=StandardAnalyzer`

Supported Analyzers

General Analyzers

- WhitespaceAnalyzer
- LowerCaseAnalyzer
- SimpleAnalyzer
- KeywordAnalyzer

Language-specific Analyzers

These are the same as SimpleAnalyzer but also remove *stop words:* words that are so common that there is little to be gained in searching for them or listing their occurrences.

As an example, the stop words used by StandardAnalyzer, the language-specific analyzer for English, are *a an and are as at be but by for if in into is it no not of on or such that the their then there these they this to was will with.*

The language-specific analyzers installed with UCS are:

- BrazilianAnalyzer
- ChineseAnalyzer
- CJKAnalyzer (Chinese/Japanese/Korean; any language that uses Chinese characters/kanji/hanja)
- CzechAnalyzer
- DutchAnalyzer
- FrenchAnalyzer
- GermanAnalyzer
- GreekAnalyzer
- RussianAnalyzer
- StandardAnalyzer (English)
- ThaiAnalyzer
- SpanishAnalyzer
- ItalianAnalyzer

Note that `Interaction.AllAttributes` requires the usage of KVListAnalyzer, it is not possible to change it or customize it. The **AllAttribute** field is indexed as a text with following tokens: `.@-_<>'"[]|=`

- It is not possible to do accurate wildcard matches on **AllAttributes**.
- It is difficult to do exact searches in the form key=value.