

Solr

sie 10, 2025

Polish language to Solr

<https://community.sap.com/t5/crm-and-cx-q-a/polish-language-to-solr/qaq-p/12021384>

HunspellStemFilterFactory which the current version supports

https://support.sitecore.com/kb?id=kb_article_view&sysparm_article=KB1001661

To add a new language to solr, you need put additional fields in schema.xml file under yoursolrserver/solr/server/multicore/conf folder. This is achieved by copy-pasting what's already there for existing languages, for example the text fieldType :

```
<fieldType name="text_pl" class="solr.TextField" positionIncrementGap="100">
  <analyzer>
    <tokenizer class="solr.StandardTokenizerFactory" />
    <filter class="solr.StandardFilterFactory" />
    <filter class="solr.LowerCaseFilterFactory" />
    <filter class="de.hybris.platform.solrfacetsearch.ysolr.synonyms.HybrisSynonymFilterFactory"
ignoreCase="true" synonyms="pl" coreName="${solr.core.name}"/>
    <filter class="solr.WordDelimiterFilterFactory"
      generateWordParts="1" generateNumberParts="1" catenateWords="1"
      catenateNumbers="1" catenateAll="0" splitOnCaseChange="0" />
    <filter class="de.hybris.platform.solrfacetsearch.ysolr.stopwords.HybrisStopWordsFilterFactory"
ignoreCase="true" lang="pl" coreName="${solr.core.name}"/>
    <filter class="solr.StopFilterFactory" words="stopwords_pl.txt" ignoreCase="true" />
    <filter class="solr.ASCIIFoldingFilterFactory" />
    <filter class="solr.SnowballPorterFilterFactory" language="Polish" />
  </analyzer>
</fieldType>
```

Implementacja

/opt/tomcat/webapps/opensource/WEB-INF/solr/configsets/default/conf

/opt/tomcat/webapps/opensource/WEB-INF/solr/configsets/default/conf/schema.xml

```
<!-- Polish -->
<fieldType name="text_pl" class="solr.TextField" positionIncrementGap="100">
  <analyzer>
    <tokenizer class="solr.StandardTokenizerFactory"/>
    <filter class="solr.LowerCaseFilterFactory"/>
    <!-- <filter class="solr.StempelPolishStemFilterFactory"/> -->
```

```

</analyzer>
</fieldType>

<!-- Russian -->
<fieldType name="text_ru" class="solr.TextField" positionIncrementGap="100">
  <analyzer>
    <tokenizer class="solr.StandardTokenizerFactory"/>
    <filter class="solr.LowerCaseFilterFactory"/>
    <!-- <filter class="solr.StempelRussianStemFilterFactory"/> -->
  </analyzer>
</fieldType>

```

```

<field name="text_pl"      type="text_pl"    indexed="true" stored="false" multiValued="true"/><!--
Catchall for PL text fields -->
<field name="text_ru"      type="text_ru"    indexed="true" stored="false" multiValued="true"/><!--
Catchall for RU text fields -->

```

```

<dynamicField name="*_pl"   type="text_pl"   indexed="true" stored="true" />
<dynamicField name="*_ru"   type="text_ru"   indexed="true" stored="true" />

<copyField source="*_pl"    dest="text_pl"/>
<copyField source="*_ru"    dest="text_ru"/>

```

[Read more](#)