Jetspeed-2 Search Component v.2.1.2

Project Documentation

Table of Contents

....

1.1 **Search Overview**

....
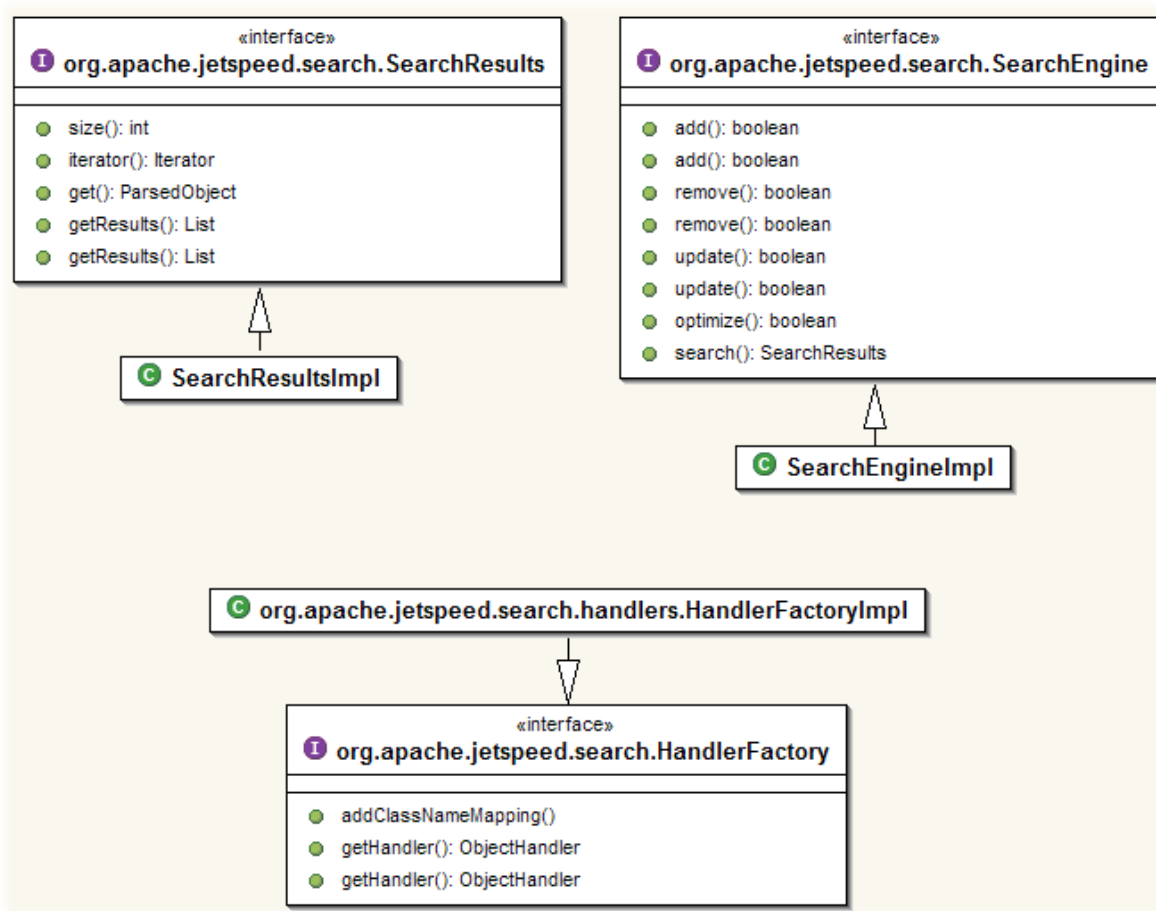
Search Overview

Jetspeed-2 provides an integration with the popular Apache Lucene , a high-performance, full-featured text search engine library written entirely in Java; a technology suitable for nearly any application that requires full-text search, especially cross-platform..

SearchEngine Overview

Jetspeed-2 provides a `SearchEngine` component configured as a spring component. The `SearchEngine` component is configured in `WEB-INF/assembly/search.xml`. The default implementation based on the embedded Lucene search engine must specify the location of the search index file, the name of the analyzer class (if null the default analyzer `StandardAnalyzer` is used), whether to optimize after update and the `HandlerFactory`:

```
    <bean id="org.apache.jetspeed.search.SearchEngine"
            class="org.apache.jetspeed.search.lucene.SearchEngineImpl">
            <constructor-arg
 index="0"><value>${applicationRoot}/WEB-INF/search_index</value></constructor-arg>
            <constructor-arg index="1"><null /></constructor-arg>
            <constructor-arg type="boolean"><valu>true</value></constructor-arg>
            <constructor-arg><ref
 bean="org.apache.jetspeed.search.HandlerFactory"/></constructor-arg>
    </bean>
```

The `HandlerFactory` provides the `SearchEngine` with a list of `ObjectHandler` that will handle the various document types supported by Jetspeed-2 for search. By default, Jetspeed-2 supports portlet instances and portlet definitions as searchable entities. When portlets are registered to the portal, `searchEngine.add(pa)` and `searchEngine.add(pa.getPortletDefinitions())` are invoked. This action updates the Jetspeed-2 search index. For more information on how portlets are registered to the search engine, see `org.apache.jetspeed.tools.pamanager.PortletApplicationManager`.
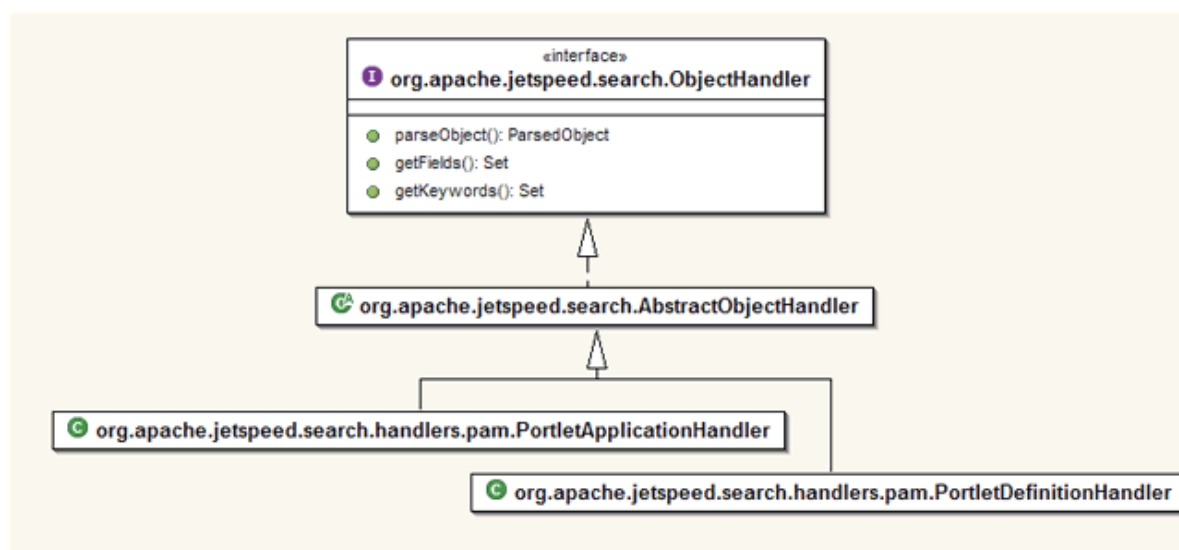
Document Handlers Overview

Document Handlers are responsible for the parsing of a specific document type in order to index the relevant document fields.

Jetspeed-2 provides 2 document handlers implementations responsible for parsing parsing the documents supported by Jetspeed-2 as `org.apache.jetspeed.search.ParsedObject`. The `ParsedObject` specify the list of fields and documents supported by Jetspeed-2 that can then be added to `org.apache.lucene.document.Document` and written to the index through `indexWriter.addDocument(doc)` operation of the `IndexWriter`.

By default, Jetspeed-2 can index portlet applications and portlet definitions respectively through the `PortletApplicationHandler` and `PortletDefinitionHandler`.

An Extensible Framework

As most components in Jetspeed-2, the search engine can easily be extended to support addional document types.

1.2 **Using the Search Engine**

....

Using Jetspeed Search Engine

Jetspeed-2 `SearchEngine` exposes a `search` operation returning `SearchResults`. The `SearchResult` can be iterated through to display the result of the search.

Portlet Search

Jetspeed-2 `SearchEngine` is used by the `PortletApplicationBrowser` portlet to search and retrieve portlet according to a given criteria.

Performing a search for "security" returns all the matching portlets and portlet definitions.