Apache Lucene

Copyright 2001-2022 The Apache Software Foundation

This product includes software developed at

The Apache Software Foundation (http://www.apache.org/).

Includes software from other Apache Software Foundation projects,

including, but not limited to:

- Apache Ant

- Apache Jakarta Regexp

- Apache Commons

- Apache Xerces

ICU4J, (under analysis/icu) is licensed under an MIT styles license

and Copyright (c) 1995-2008 International Business Machines Corporation and others

Some data files (under analysis/icu/src/data) are derived from Unicode data such

as the Unicode Character Database. See http://unicode.org/copyright.html for more

details.

Brics Automaton (under core/src/java/org/apache/lucene/util/automaton) is

BSD-licensed, created by Anders Møller. See http://www.brics.dk/automaton/

The levenshtein automata tables (under core/src/java/org/apache/lucene/util/automaton) were

automatically generated with the moman/finenight FSA library, created by

Jean-Philippe Barrette-LaPierre. This library is available under an MIT license,

see http://sites.google.com/site/rrettesite/moman and

http://bitbucket.org/jpbarrette/moman/overview/

The class org.apache.lucene.util.WeakIdentityMap was derived from

the Apache CXF project and is Apache License 2.0.

The class org.apache.lucene.util.compress.LZ4 is a Java rewrite of the LZ4

compression library (https://github.com/lz4/lz4/tree/dev/lib) that is licensed

under the 2-clause BSD license.

(https://opensource.org/licenses/bsd-license.php)

The Google Code Prettify is Apache License 2.0.

See http://code.google.com/p/google-code-prettify/

JUnit (junit-4.10) is licensed under the Common Public License v. 1.0

See http://junit.sourceforge.net/cpl-v10.html

This product includes code (JaspellTernarySearchTrie) from Java Spelling Checkin

g Package (jaspell): http://jaspell.sourceforge.net/

License: The BSD License (http://www.opensource.org/licenses/bsd-license.php)

The snowball stemmers in

analysis/common/src/java/net/sf/snowball

were developed by Martin Porter and Richard Boulton.

The snowball stopword lists in

analysis/common/src/resources/org/apache/lucene/analysis/snowball

were developed by Martin Porter and Richard Boulton.

The full snowball package is available from

https://snowballstem.org/

The KStem stemmer in

analysis/common/src/org/apache/lucene/analysis/en

was developed by Bob Krovetz and Sergio Guzman-Lara (CIIR-UMass Amherst)

under the BSD-license.

The Arabic,Persian,Romanian,Bulgarian, Hindi and Bengali analyzers (common) come with a default

stopword list that is BSD-licensed created by Jacques Savoy. These files reside in:

analysis/common/src/resources/org/apache/lucene/analysis/ar/stopwords.txt,

analysis/common/src/resources/org/apache/lucene/analysis/fa/stopwords.txt,

analysis/common/src/resources/org/apache/lucene/analysis/ro/stopwords.txt,

analysis/common/src/resources/org/apache/lucene/analysis/bg/stopwords.txt,

analysis/common/src/resources/org/apache/lucene/analysis/hi/stopwords.txt,

analysis/common/src/resources/org/apache/lucene/analysis/bn/stopwords.txt

See http://members.unine.ch/jacques.savoy/clef/index.html.

The German,Spanish,Finnish,French,Hungarian,Italian,Portuguese,Russian and Swedish light stemmers

(common) are based on BSD-licensed reference implementations created by Jacques Savoy and

Ljiljana Dolamic. These files reside in:

analysis/common/src/java/org/apache/lucene/analysis/de/GermanLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/de/GermanMinimalStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/es/SpanishLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/fi/FinnishLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/fr/FrenchLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/fr/FrenchMinimalStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/hu/HungarianLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/it/ItalianLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/pt/PortugueseLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/ru/RussianLightStemmer.java

analysis/common/src/java/org/apache/lucene/analysis/sv/SwedishLightStemmer.java

The Stempel analyzer (stempel) includes BSD-licensed software developed

by the Egothor project http://egothor.sf.net/, created by Leo Galambos, Martin Kvapil,

and Edmond Nolan.

The Polish analyzer (stempel) comes with a default

stopword list that is BSD-licensed created by the Carrot2 project. The file resides

in stempel/src/resources/org/apache/lucene/analysis/pl/stopwords.txt.

See https://github.com/carrot2/carrot2.

The SmartChineseAnalyzer source code (smartcn) was

provided by Xiaoping Gao and copyright 2009 by www.imdict.net.

WordBreakTestUnicode\_\*.java (under modules/analysis/common/src/test/)

is derived from Unicode data such as the Unicode Character Database.

See http://unicode.org/copyright.html for more details.

The Morfologik analyzer (morfologik) includes BSD-licensed software

developed by Dawid Weiss and Marcin Miłkowski

(https://github.com/morfologik/morfologik-stemming) and uses

data from the BSD-licensed dictionary of Polish (SGJP, http://sgjp.pl/morfeusz/).

Servlet-api.jar and javax.servlet-\*.jar are under the CDDL license, the original

source code for this can be found at http://www.eclipse.org/jetty/downloads.php

===========================================================================

Kuromoji Japanese Morphological Analyzer - Apache Lucene Integration

===========================================================================

This software includes a binary and/or source version of data from

mecab-ipadic-2.7.0-20070801

which can be obtained from

http://atilika.com/releases/mecab-ipadic/mecab-ipadic-2.7.0-20070801.tar.gz

or

http://jaist.dl.sourceforge.net/project/mecab/mecab-ipadic/2.7.0-20070801/mecab-ipadic-2.7.0-20070801.tar.gz

===========================================================================

mecab-ipadic-2.7.0-20070801 Notice

===========================================================================

Nara Institute of Science and Technology (NAIST),

the copyright holders, disclaims all warranties with regard to this

software, including all implied warranties of merchantability and

fitness, in no event shall NAIST be liable for

any special, indirect or consequential damages or any damages

whatsoever resulting from loss of use, data or profits, whether in an

action of contract, negligence or other tortuous action, arising out

of or in connection with the use or performance of this software.

A large portion of the dictionary entries

originate from ICOT Free Software. The following conditions for ICOT

Free Software applies to the current dictionary as well.

Each User may also freely distribute the Program, whether in its

original form or modified, to any third party or parties, PROVIDED

that the provisions of Section 3 ("NO WARRANTY") will ALWAYS appear

on, or be attached to, the Program, which is distributed substantially

in the same form as set out herein and that such intended

distribution, if actually made, will neither violate or otherwise

contravene any of the laws and regulations of the countries having

jurisdiction over the User or the intended distribution itself.

NO WARRANTY

The program was produced on an experimental basis in the course of the

research and development conducted during the project and is provided

to users as so produced on an experimental basis. Accordingly, the

program is provided without any warranty whatsoever, whether express,

implied, statutory or otherwise. The term "warranty" used herein

includes, but is not limited to, any warranty of the quality,

performance, merchantability and fitness for a particular purpose of

the program and the nonexistence of any infringement or violation of

any right of any third party.

Each user of the program will agree and understand, and be deemed to

have agreed and understood, that there is no warranty whatsoever for

the program and, accordingly, the entire risk arising from or

otherwise connected with the program is assumed by the user.

Therefore, neither ICOT, the copyright holder, or any other

organization that participated in or was otherwise related to the

development of the program and their respective officials, directors,

officers and other employees shall be held liable for any and all

damages, including, without limitation, general, special, incidental

and consequential damages, arising out of or otherwise in connection

with the use or inability to use the program or any product, material

or result produced or otherwise obtained by using the program,

regardless of whether they have been advised of, or otherwise had

knowledge of, the possibility of such damages at any time during the

project or thereafter. Each user will be deemed to have agreed to the

foregoing by his or her commencement of use of the program. The term

"use" as used herein includes, but is not limited to, the use,

modification, copying and distribution of the program and the

production of secondary products from the program.

In the case where the program, whether in its original form or

modified, was distributed or delivered to or received by a user from

any person, organization or entity other than ICOT, unless it makes or

grants independently of ICOT any specific warranty to the user in

writing, such person, organization or entity, will also be exempted

from and not be held liable to the user for any such damages as noted

above as far as the program is concerned.

===========================================================================

Nori Korean Morphological Analyzer - Apache Lucene Integration

===========================================================================

This software includes a binary and/or source version of data from

mecab-ko-dic-2.1.1-20180720

which can be obtained from

https://bitbucket.org/eunjeon/mecab-ko-dic/downloads/mecab-ko-dic-2.1.1-20180720.tar.gz

The floating point precision conversion in NumericUtils.Float16Converter is derived from work by

Jeroen van der Zijp, granted for use under the Apache license.