

ReleaseNote47

February 2014, Apache Lucene™ 4.7 available

The Lucene PMC is pleased to announce the release of Apache Lucene 4.7

Apache Lucene is a high-performance, full-featured text search engine library written entirely in Java. It is a technology suitable for nearly any application that requires full-text search, especially cross-platform.

This release contains numerous bug fixes, optimizations, and improvements, some of which are highlighted below. The release is available for immediate download at:

<http://lucene.apache.org/core/mirrors-core-latest-redir.html>

See the CHANGES.txt file included with the release for a full list of details.

Lucene 4.7 Release Highlights:

- * When sorting by String (SortField.STRING), you can now specify whether missing values should be sorted first (the default), or last.
- * Add two memory resident dictionaries (FST terms dictionary and FSTOrd terms dictionary) to improve primary key lookups. The PostingsBaseFormat API is also changed so that term dictionaries get the ability to block encode term metadata, and all dictionary implementations can now plug in any PostingsBaseFormat. See <http://blog.mikemccandless.com/2013/09/lucene-now-has-in-memory-terms.html>
- * NRT support for file systems that do not have delete on last close or cannot delete while referenced semantics.
- * Add LongBitSet for managing more than 2.1B bits (otherwise use FixedBitSet).
- * Speed up Lucene range faceting from O(N) per hit to O(log(N)) per hit using segment trees. See <http://blog.mikemccandless.com/2013/12/fast-range-faceting-using-segment-trees.html>
- * Add SearcherTaxonomyManager over search and taxonomy index directories (i.e. not only NRT).
- * Drilling down or sideways on a Lucene facet range (using Range.getFilter()) is now faster for costly filters (uses random access, not iteration); range facet counts now accept a fast-match filter to avoid computing the value for documents that are out of bounds, e.g. using a bounding box filter with distance range faceting.
- * Add Analyzer for Kurdish.
- * Add Payload support to FileDictionary (Suggest) and make it more configurable.
- * Add a new BlendedInfixSuggester, which is like AnalyzingInfixSuggester but boosts suggestions that matched tokens with lower positions.
- * Add SimpleQueryParser: parser for human-entered queries.
- * Add multitermquery (wildcards,prefix,etc) to PostingsHighlighter.
- * Upgrade to Spatial4j 0.4.1: Parses WKT (including ENVELOPE) with extension "BUFFER"; buffering a point results in a Circle. JTS isn't needed for WKT any more but remains required for Polygons. New Shapes: ShapeCollection and BufferedLineString. More info: <https://github.com/spatial4j/spatial4j/blob/master/CHANGES.md>
- * Add spatial SerializedDVStrategy that serializes a binary representation of a shape into BinaryDocValues. It supports exact geometry relationship calculations.

* Various bugfixes and optimizations since the 4.6.1 release.

Please read CHANGES.txt for a full list of new features.

Please report any feedback to the mailing lists
(<http://lucene.apache.org/core/discussion.html>)

Note: The Apache Software Foundation uses an extensive mirroring network for distributing releases. It is possible that the mirror you are using may not have replicated the release yet. If that is the case, please try another mirror. This also goes for Maven access.