

Apache Lucene - Overview

Jon S. Stevens Ted Husted Doug Cutting Peter Carlson

Table of contents

1 Apache Lucene.....	2
2 Lucene News.....	2
2.1 19 June 2007 - Release 2.2 available.....	2
2.2 18 February 2007 - Lucene at ApacheCon Europe.....	2
2.3 17 February 2007 - Release 2.1 available.....	3
2.4 3 January 2007 - Nightly Source builds available.....	3
2.5 22 December 2006.....	3
2.6 26 November 2006.....	3
2.7 26 May 2006 - Release 2.0.0 available.....	3
2.8 2 March 2006 - Release 1.9.1 available.....	3
2.9 27 February 2006 - 1.9 final available.....	4
2.10 26 January 2006 - Nightly builds available.....	4
2.11 28 October 2005 - Lucene at ApacheCon.....	4
2.12 14 February 2005 - Lucene moves to Apache top-level.....	5
2.13 December 2004 - Lucene in Action is published	5
2.14 29 November 2004 - Lucene 1.4.3 Released.....	5

1. Apache Lucene

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.

Apache Lucene is an open source project available for free download. Please use the links on the left to access Lucene.

2. Lucene News

2.1. 19 June 2007 - Release 2.2 available

This release has many improvements since release 2.1. New major features:

- "Point-in-time" searching
- Payloads
- Function queries
- New API for preanalyzed fields

In addition Lucene 2.2 has many performance improvements, bug fixes, etc. See CHANGES.txt for details.

Lucene 2.2 includes index format changes that are not readable by older versions of Lucene. Lucene 2.2 can both read and update older Lucene indexes. Adding to an index with an older format will cause it to be converted to the newer format.

Binary and source distributions are available here.

2.2. 18 February 2007 - Lucene at ApacheCon Europe

ApacheCon Europe logo

Lucene Java and related Lucene projects will have extensive representation at ApacheCon Europe in Amsterdam this year. For the 2007 session, Yonik Seeley will be giving the Full-Text Search with Lucene talk at 10:30 am on May 2nd. Immediately following, Grant Ingersoll will be presenting Advanced Lucene at 11:30. Grant will also be leading a full day tutorial session on May 1st titled Lucene Boot Camp.

Lucene related talks include Solr committer Bertrand Delacrétaiz's talk titled Beyond full-text searches with Solr and Lucene and Hadoop committer Owen O'Malley's Introduction to Hadoop.

Registration is now open on the ApacheCon website.

2.3. 17 February 2007 - Release 2.1 available

This release has many improvements since release 2.0, including new features, performance improvements, bug fixes, etc. See CHANGES.txt for details.

Lucene 2.1 includes index format changes that are not readable by older versions of Lucene. Lucene 2.1 can both read and update older Lucene indexes. Adding to an index with an older format will cause it to be converted to the newer format.

Binary and source distributions are available here.

2.4. 3 January 2007 - Nightly Source builds available

Nightly source builds of the current development version of Lucene are now available at <http://people.apache.org/builds/lucene/java/nightly/>. Files are named lucene-DATE-src.tar.gz where DATE is the date of the build.

2.5. 22 December 2006

The Lucene Java main site is now based on the nightly build of the documentation contained in Subversion. The API documentation is also based on the nightly build of the source. Previous released versions of this site (including API references) are available under the "Site Versions" menu item on the left.

2.6. 26 November 2006

New Forrest based site released. The Lucene Java website now has a consistent look and feel with its Lucene siblings.

2.7. 26 May 2006 - Release 2.0.0 available

This is mostly a bugfix release from release 1.9.1. Note however that deprecated 1.x features have now been removed. Any code that compiles against Lucene 1.9.1 without deprecation warnings should work without further changes with any 2.x release. For more information about this release, please read CHANGES.txt .

Binary and source distributions are available here .

2.8. 2 March 2006 - Release 1.9.1 available

This fixes a serious bug in release 1.9-final. CHANGES.txt for details.

Binary and source distributions are available here .

2.9. 27 February 2006 - 1.9 final available

This release has many improvements since release 1.4.3, including new features, performance improvements, bug fixes, etc. See CHANGES.txt for details.

1.9 will be the last 1.x release. It is both back-compatible with 1.4.3 and forward-compatible with the upcoming 2.0 release. Many methods and classes in 1.4.3 have been deprecated in 1.9 and will be removed in 2.0. Applications must compile against 1.9 without deprecation warnings before they are compatible with 2.0.

Binary and source distributions are available here .

2.10. 26 January 2006 - Nightly builds available

Nightly builds of the current development version of Lucene, to be released as Lucene 1.9, are now available at <http://cvs.apache.org/dist/lucene/java/nightly/> .

2.11. 28 October 2005 - Lucene at ApacheCon



Monday, December 12, 2005 at 3pm by Grant Ingersoll:

Abstract:

Lucene is a high performance, scalable, cross-platform search engine that contains many advanced features that often go untapped by the majority of users. In this session, designed for those familiar with Lucene, we will examine some of Lucene's more advanced topics and their application, including:

1. Term Vectors: Manual and Pseudo relevance feedback; Advanced document collection analysis for domain specialization
2. Span Queries: Better phrase matching; Candidate Identification for Question Answering
3. Tying it all Together: Building a search framework for experimentation and rapid deployment
4. Case Studies from CNLP : Crosslingual/multilingual retrieval in Arabic, English and Dutch; Sublanguage specialization for commercial trouble ticket analysis; Passage retrieval and analysis for Question Answering application

Topics 1 through 3 will provide technical details on implementing the advanced Lucene features, while the fourth topic will provide a broader context for understanding when and where to use these features.

2.12. 14 February 2005 - Lucene moves to Apache top-level

Lucene has migrated from Apache's Jakarta project to the top-level. Along with this migration, the source code repository has been converted to Subversion. The migration is in progress with some loose ends. Please stay tuned!

2.13. December 2004 - Lucene in Action is published

The first book dedicated solely to Lucene is published. The "search inside the book" feature implemented with Lucene can be seen at lucenebook.com .

2.14. 29 November 2004 - Lucene 1.4.3 Released

This fixes a few bugs in 1.4.2. See [CHANGES.txt](#) for details. Binary and source distributions are available here . After choosing your mirror, navigate to the archive section via the java link.