

# Adding Full Text Search to Your Rails App



# Lenio





Why?





Lucene

# Lucene

From Wikipedia, the free encyclopedia

**Lucene** is a [free/open source information retrieval library](#), originally implemented in [Java](#). It is supported by the [Apache Software Foundation](#) and is released under the [Apache Software License](#). Lucene has been ported to programming languages including [Delphi](#), [Perl](#), [C#](#), [C++](#), [Python](#), [Ruby](#) and [PHP](#).

While suitable for any application which requires full text indexing and searching capability, Lucene has been widely recognized for its utility in the implementation of Internet search engines and local, single-site searching. Lucene itself is just an indexing and search library and does not contain [crawling](#) and [HTML parsing](#) functionality. The Apache project [Nutch](#) is based on Lucene and provides this functionality; the Apache project [Solr](#) is a fully-featured search server based on Lucene.

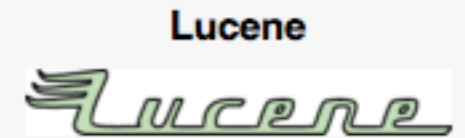
At the core of Lucene's logical architecture is the idea of a **document** containing **fields** of text. This flexibility allows Lucene's API to be independent of file format. Text from [PDFs](#), [HTML](#), [Microsoft Word](#) documents, as well as many others can all be indexed so long as their textual information can be extracted.

## Contents [hide]

- 1 [Software using Lucene](#)
- 2 [Ports](#)
- 3 [References](#)
- 4 [See also](#)
- 5 [External links](#)

## Software using Lucene

- [OpenGrok](#) A fast and usable source code search and cross reference engine
- [Puggle](#) A graphical desktop search engine that uses Lucene for full-text and metadata search [1] [↗](#)
- [isoHunt](#) Uses Lucene for the site search [2] [↗](#)
- [Gplex Database](#) ([homepage](#) [↗](#)) Uses a C# version of Lucene to index your database metadata.
- [EB-eye\\_EBI's\\_Search\\_Engine](#) EMBL-EBI's Biomedical databases search engine (contains than 200 million documents)
- [Joost](#) Internet TV uses Lucene to search for programs.
- [MediaWiki](#) can use Lucene for full-text search.
- [QuestAgent](#) [↗](#) offline search engine uses Lucene in applet that provides full text search.
- [Liferay](#) open source portal, uses Lucene for full-text search.
- [Beagle](#) uses a port of Lucene to [C#](#) called [Lucene.Net](#) [↗](#) as its indexer.



<b>Developed by</b>	<a href="#">Apache Software Foundation</a>
<b>Latest release</b>	2.3.1 / February 23, 2008
<b>Written in</b>	<a href="#">Java</a>
<b>OS</b>	<a href="#">Cross-platform</a>
<b>Genre</b>	<a href="#">Search and index</a>
<b>License</b>	<a href="#">Apache License 2.0</a>
<b>Website</b>	<a href="#">lucene.apache.org</a> <a href="#">↗</a>

[\[edit\]](#)



**WIKIPEDIA**  
The Free Encyclopedia

### navigation

- [Main Page](#)
- [Contents](#)
- [Featured content](#)
- [Current events](#)
- [Random article](#)

### interaction

- [About Wikipedia](#)
- [Community portal](#)
- [Recent changes](#)
- [Contact Wikipedia](#)
- [Donate to Wikipedia](#)
- [Help](#)

### search




### toolbox

- [What links here](#)
- [Related changes](#)
- [Upload file](#)
- [Special pages](#)
- [Printable version](#)
- [Permanent link](#)
- [Cite this page](#)

### languages

- [Deutsch](#)
- [Español](#)
- [Français](#)
- [Italiano](#)
- [Nederlands](#)
- [日本語](#)

# Lucene - key concepts

Tokens

Fields

Documents



# Solr



# acts\_as\_solr

```
script/plugin install svn://svn.railsfreaks.com/  
  projects/acts_as_solr/trunk
```

```
add  
act_as_solr to  
model
```

```
rake solr:start
```

# Adding it to a model

```
class Movie < ActiveRecord::Base  
  acts_as_solr  
  
end
```

```
class Movie < ActiveRecord::Base
  acts_as_solr :fields => [:title, :production_year]
end
```

```
Movie.find_by_solr('love').docs
Movie.find_by_solr('love').total_hits
```

# Sample queries

1984

love

title:1984

love AND hate

lov\*

```
class Movie < ActiveRecord::Base

  has_many :actors

  acts_as_solr :fields => [:title, :production_year,
                        :actor_names]

  def actor_names
    actors.map(&:name).join(" ")
  end

end
```

# Multimodel search

```
Movie.multi_solr_search "Paris", :models => [Actor]
```

```
results.docs =>
```

```
 [#<Actor id: 41, name: "Paris Hilton">,
  #<Movie id: 1195, title: "Return to Paris">]
```

# Facets

amazon.com

Hello, Jørgen. We have [recommendations](#) for you. (Not [Jørgen?](#))

Jørgen's Amazon.com



Today's Deals

Gifts & Wish Lists

Gift Cards

Your Account | Help

Shop All Departments

Search Amazon.com

love

GO



Your Lists

## Category

### Any Category

- Apparel (1,589,642)
- Books (590,195)
- MP3 Downloads (185,861)
- Home & Garden (168,357)
- Music (48,157)
- Sports & Outdoors (18,616)
- DVD (8,378)
- Jewelry (7,482)
- Health & Personal Care (6,541)
- Everything Else (4,906)
- VHS (3,088)
- Kindle Store (2,144)
- Automotive (2,086)
- Baby (1,692)
- Toys & Games (1,571)
- Unbox Video Downloads (1,238)
- Gourmet Food (731)
- Watches (672)
- Office Products (647)
- Beauty (385)
- Electronics (241)
- Home Improvement (240)
- Software (144)
- Grocery (144)

## "love"

Related Searches: [i love you](#), [romance](#), [valentine](#).

Showing Top Results

« Previous | Page: 1 2 3 | Next »



Explore over 4,700 products tagged **love** by customers

See what the community is talking about in the [love discussion forum](#).

1.



**Love** by The Beatles (**Audio CD** - Nov 21, 2006)

Buy new: ~~\$18.98~~ **\$8.99** [98 Used & new](#) from **\$7.15**

Get it by **Thursday, April 3** if you order in the next **14 hours** and choose one-day shipping. Eligible for **FREE** Super Saver Shipping.

★★★★☆ (536)

**Music:** [See all 48,157 items](#)

2.



**Forever Changes** by Love (**Audio CD** - Feb 20, 2001)

Buy new: ~~\$11.98~~ **\$10.99** [73 Used & new](#) from **\$5.32**

Get it by **Thursday, April 3** if you order in the next **12 hours** and choose one-day shipping. Eligible for **FREE** Super Saver Shipping.

★★★★☆ (161)

**Music:** [See all 48,157 items](#)

3.



**Other People's Love Letters: 150 Letters You Were Never Meant to See** by Bill Shapiro (**Hardcover** - Oct 30, 2007)

Buy new: ~~\$22.50~~ **\$15.30** [27 Used & new](#) from **\$11.30**



```
class Movie < ActiveRecord::Base
```

```
  acts_as_solr :fields => [:title, :production_year],  
              :facets => [:production_year]
```

```
end
```

```
Movie.create(:title => 'The Tenth Man', :production_year => 1988)
```

```
Movie.create(:title => 'Last Man Standing', :production_year => 1996)
```

```
Movie.create(:title => 'Rain Man', :production_year => 1988)
```

```
results = Movie.find_by_solr('man', :facets => {:fields => [:production_year]})
```

```
results.facets['facet_fields']['production_year_facet']  
=> {"1996" => 1, "1988" => 2}
```

```
class Movie < ActiveRecord::Base
```

```
  acts_as_solr :fields => [:title, :production_year],  
              :facets => [:production_year => :range_integer]
```

```
end
```

```
Movie.create(:title => 'The Tenth Man', :production_year => 1988)
```

```
Movie.create(:title => 'Last Man Standing', :production_year => 1996)
```

```
Movie.create(:title => 'Rain Man', :production_year => 1988)
```

```
results = Movie.find_by_solr('man',  
  :facets => {:fields => [:production_year],  
             :query => ['production_year:[1980 TO 1989]',  
                       'production_year:[1990 TO 1999]']})
```

```
results.facets['facet_queries']  
=> {"production_year_t:[1980 TO 1989]" => 2,  
    "production_year_t:[1990 TO 1999]" => 1}
```

# Highlighting

```
Movie.create(:title => 'Last Man Standing', :production_year => 1996)
results = Movie.find_by_solr('man', :highlight => {:fields => 'title'})
results.highlights
=> {1 => {:title => "Last <em>Man</em> Standing"}}
```

# Other languages

Danish Dutch English Finnish French  
German Italian Norwegian Portuguese  
Russian Spanish Swedish

```
<filter class="solr.EnglishPorterFilterFactory"  
  protected="protwords.txt"/>
```

```
<filter class="solr.SnowballPorterFilterFactory"  
  language="Danish" />
```

# Other options

Ferret / `acts_as_ferret`

Sphinx / `Ultrasphinx`

# Resources

[acts-as-solr.railsfreaks.com](http://acts-as-solr.railsfreaks.com)

[www.quarkruby.com/2007/8/12/  
acts\\_as\\_solr-for-search-and-faceting](http://www.quarkruby.com/2007/8/12/acts_as_solr-for-search-and-faceting)

[www.quarkruby.com/  
2007/9/14/advanced-  
acts\\_as\\_solr](http://www.quarkruby.com/2007/9/14/advanced-acts_as_solr)

# Questions?



# Credits

iStockphoto.com

[www.flickr.com/photos/astrowoosie/120750647](http://www.flickr.com/photos/astrowoosie/120750647)