Full Text Searching

with Ruby on Rails

@claytonlz

1 Sphinx
1 ThinkingSphinx
1 ActiveRecord

Full Text Search in Rails

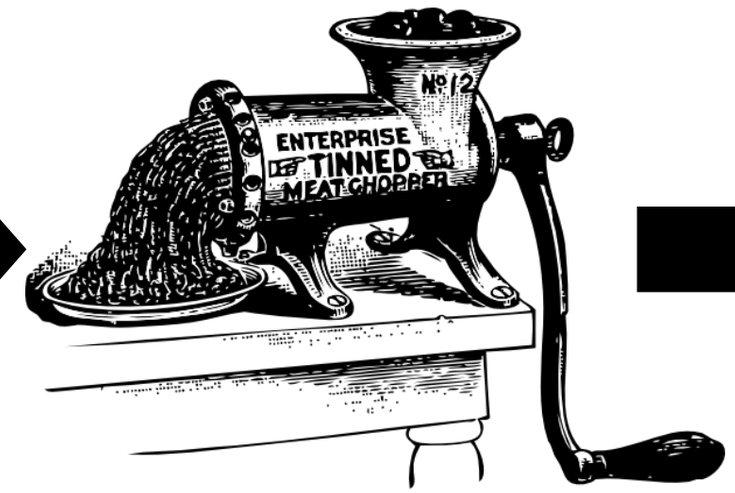
Sphinx

Sphinx

Documents

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem, porttitor et scelerisque eget, euismod eu lacus. Integer mollis leo sed quam euismod porta. Praesent convallis, nulla ac sodales lacinia, ante ipsum convallis turpis, in scelerisque eros tortor sit amet sapien. Ut sed sem risus. Praesent orci erat, consectetur non rhoncus eu, condimentum sed turpis.

Sphinx



Indexes

PostgreSQL

MySQL®

<? xml ?>

Fields

Attributes

ThinkingSphinx

ThinkingSphinx > Indexing

```
define_index do
  indexes language_id
  indexes address.state, :sortable => true
  indexes address.postal_code, :as => :zip_code
  indexes "LOWER(title)", :as => :title

  has is_private
  has trainers(:id), :as => :trainers

  where "active = 1"
  group_by "course_type_id"

  set_property :delta => true
end
```

ThinkingSphinx > Build Index

```
$> rake ts:configure
$> rake ts:start
$> rake ts:index
```

```
$> rake ts:rebuild
```

TS > Basic Searching

```
Course.search 'sphinx'

Course.search :conditions =>
  {:title => "Sphinx"}

Course.search 'sphinx', :with =>
  {:trainer_id => [5,19]}
```

TS > Fields vs. Attributes

```
Course.search 'rails',
  :conditions => {
    :title =>
      ["sphinx",
       "thinking sphinx"]
  }
```

DO NOT WANT

TS > Fields vs. Attributes

```
Course.search 'rails',
  :with => {
    :created_at =>
      Time.now..1.year.ago,
    :trainer_id => 1
  }
```

TS > Fields vs. Attributes

```
has "CRC32(category)",
   :as => :category,
   :type => :integer
```

```
Course.search 'sphinx',
   :with => {
     :category => 'Ruby'.to_crc32
   }
```

TS > Application Wide Searching

```
ThinkingSphinx.search 'clayton'

ThinkingSphinx.search 'clayton',
  :classes => [User, Article]
```

TS > Pagination

```
Course.search 'sphinx',
   :per_page => 10,
   :page => 3
```

```
Course.search 'code camp',
   :per_page => 20,
   :page => params[:page]
```

TS > Match Modes

```
Course.search 'specific',
  :match_mode => :all

Course.search 'this that',
  :match_mode => :any

Course.search 'nosql sucks',
  :match_mode => :phrase

Course.search 'just & kidding',
  :match_mode => :boolean
```

TS > More Match Modes

```
Course.search 'advanced',
  :match_mode => :extended

Course.search 'expert',
  :match_mode => :extended2

Course.search 'sorting',
  :match_mode => :fullscan
```

TS > Sorting

```
Course.search 'sphinx',
  :with => {
    :created_at =>
      Time.now..1.year.from_now
  },
  :order => "created_at DESC,
    @relevance DESC"
```

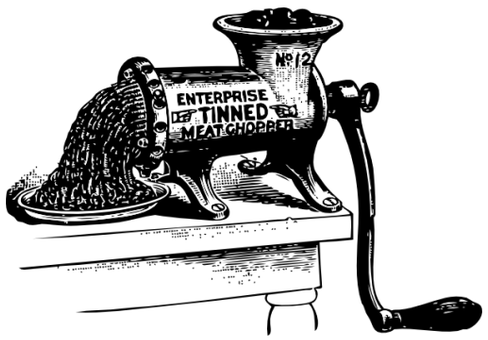
Deltas

TS > Delta Indexing

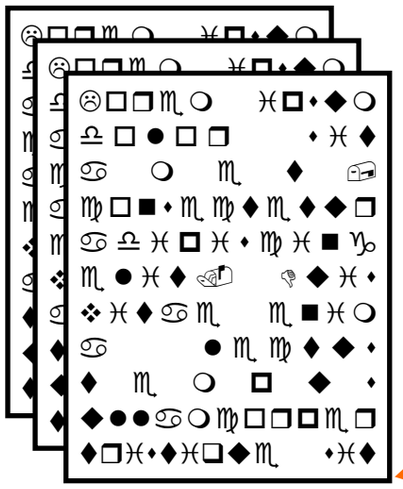
Documents

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

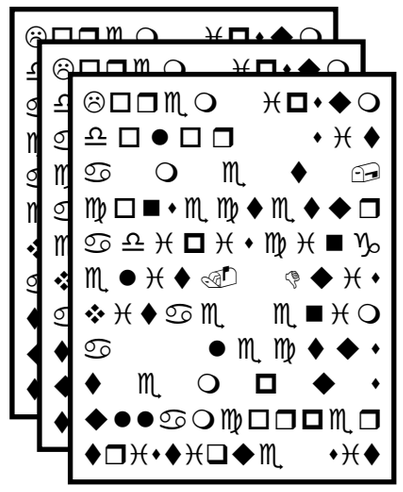
Sphinx



Indexes



Merged Index



V 1.0

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

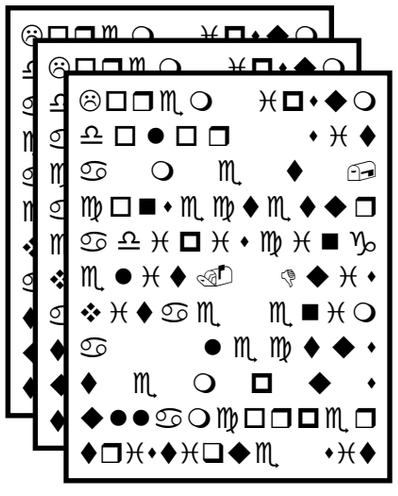
V 2.0

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

V 3.0

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

Delta Index



Beta 1.10

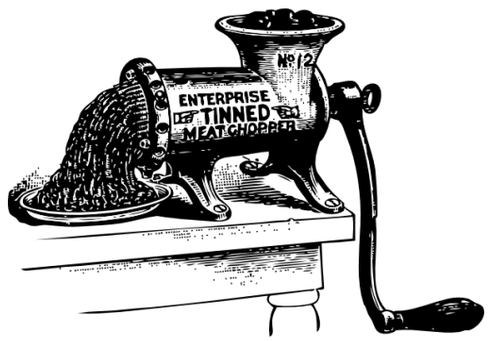
Testing

TS > Testing

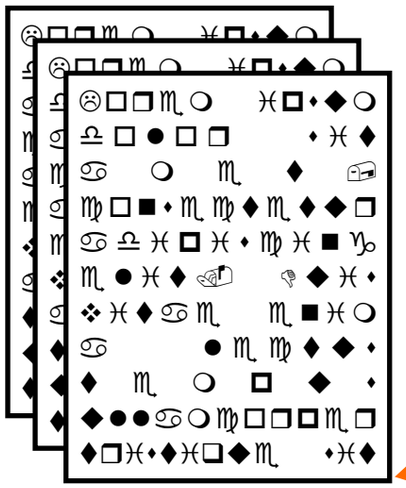
Documents

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

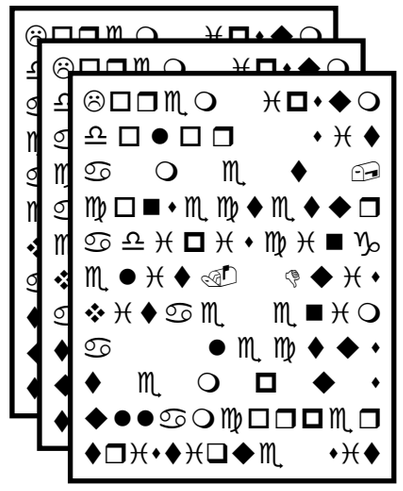
Sphinx



Indexes



Merged Index



V 1.0

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

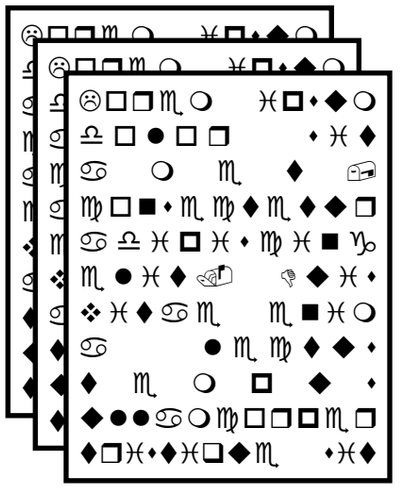
V 2.0

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

V 3.0

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vitae enim a lectus tempus ullamcorper tristique sit amet dui. Maecenas bibendum, nisi adipiscing vehicula eleifend, mi leo aliquam dui, et rhoncus erat eros a orci. Sed nibh lorem,

Delta Index



TS > Testing

```
# Setup
require 'cucumber/thinking_sphinx/external_world'
Cucumber::ThinkingSphinx::ExternalWorld.new

# No transactional Fixtures (DB Cleaner)
Cucumber::Rails::World.use_transactional_fixtures
  = false

# Index the test data
ThinkingSphinx::Test.index
```

TS > Testing

```
# SPHINX
Before('@sphinx') do
  require 'cucumber/thinking_sphinx/external_world'
  Cucumber::ThinkingSphinx::ExternalWorld.new
  ThinkingSphinx::Test.start
  ThinkingSphinx.deltas_enabled = true
end

After('@sphinx') do
  ThinkingSphinx.deltas_enabled = false
  ThinkingSphinx::Test.stop
end
```

Extras

TS > Extras

```
[
  "Advanced Sphinx Conf.",
  "Complex Algorithms",
  "Excerpts",
  "Geo-Searching",
  "Sphinx Scopes"
]
```

Deployment

TS > Deployment

1. "Deploy Code"
2. "Symlink (maybe)"
3. "Stop Sphinx"
4. "Generate Configuration File"
5. "Start Sphinx"
6. "Rebuild (maybe)"
7. "Reindex periodically"

TS > Deployment Gotchas

1. "Symbolic Links :("
2. "Slow Rebuilding :("
3. "Permissions :("
4. "Monit :("
5. "Binary paths :("

http://bit.ly/9iVJaC

http://bit.ly/cZYPyW

@claytonlz

http://claytonlz.com