

Viviana Márquez

Big Data - Prof. Camargo

Taller 6 - Septiembre 16, 2017

1. Descomprimir Solr después de descargarlo

```
Vivianas-MacBook-Pro:~ vivianamarquez$ tar -zxvf solr-6.6.1-src.tar

Vivianas-MacBook-Pro:~ vivianamarquez$ ls
Applications      Dropbox           Pictures          mongodb-osx-x86_64-3.0.4  solr-6.6.1-src.tar
Desktop           Library          Public           mongodb-osx-x86_64-3.0.4.tgz
Documents         Movies           anaconda         seaborn-data
Downloads         Music            mongodb          solr-6.6.1

Vivianas-MacBook-Pro:~ vivianamarquez$ cd solr-6.6.1
Vivianas-MacBook-Pro:solr-6.6.1 vivianamarquez$ cd solr/
Vivianas-MacBook-Pro:solr vivianamarquez$ ls
CHANGES.txt      NOTICE.txt      build.xml        contrib          example          site             test-framework
LICENSE.txt       README.txt      cloud-dev        core             licenses         solr-ref-guide  webapp
LUCENE_CHANGES.txt  bin             common-build.xml docs             server           solrj

Vivianas-MacBook-Pro:solr vivianamarquez$ bin/solr start
```

2. Intentar iniciar Solr

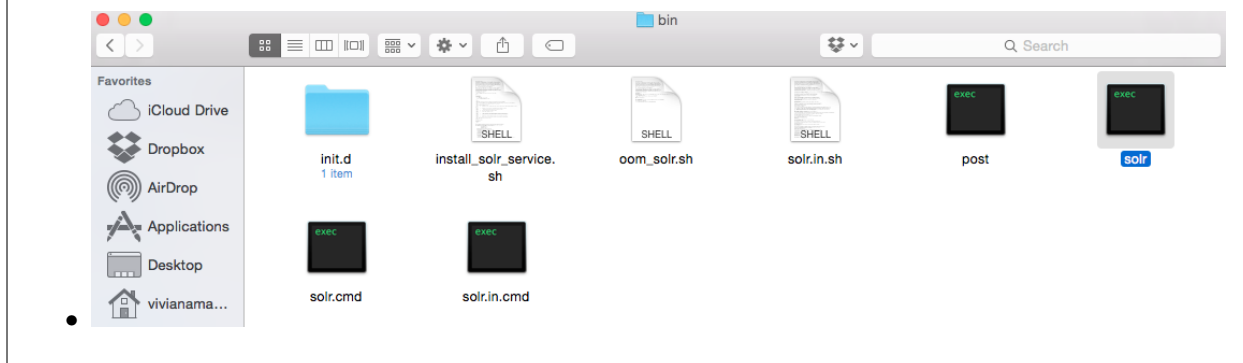
```
Vivianas-MacBook-Pro:solr-6.6.1 vivianamarquez$ bin/solr start
Java not found, or an error was encountered when running java.
A working Java 1.8 JRE is required to run Solr!
Please install latest version of Java 1.8 or set JAVA_HOME properly.
Command that we tried: 'java -version', with response:
No Java runtime present, requesting install.

Debug information:
JAVA_HOME: N/A
Active Path:
/Users/vivianamarquez/anaconda/bin:/Users/vivianamarquez/anaconda/bin:/usr/local/bin:/usr/bin:/bin:/usr/sbin:/sbin:/Library/TeX/texbin:/usr/texbin
Vivianas-MacBook-Pro:solr-6.6.1 vivianamarquez$ bin/solr start
Java not found, or an error was encountered when running java.
A working Java 1.8 JRE is required to run Solr!
Please install latest version of Java 1.8 or set JAVA_HOME properly.
Command that we tried: 'java -version', with response:
No Java runtime present, requesting install.

Debug information:
JAVA_HOME: N/A
Active Path:
/Users/vivianamarquez/anaconda/bin:/Users/vivianamarquez/anaconda/bin:/usr/local/bin:/usr/bin:/bin:/usr/sbin:/sbin:/Library/TeX/texbin:/usr/texbin
Vivianas-MacBook-Pro:solr-6.6.1 vivianamarquez$
```

3. Solr no iniciaba debido a problemas con JAVA

Aunque instalé JAVA, seguía mostrando el mismo error, así que intenté solucionar el problema de las siguientes maneras:



```

96 . "$SOLR_INCLUDE"
97 fi
98
99 if [ -z "$SOLR_PID_DIR" ]; then
100     SOLR_PID_DIR="$SOLR_TIP/bin"
101 fi
102
103 if [ -n "$SOLR_JAVA_HOME" ]; then
104     JAVA="$SOLR_JAVA_HOME/bin/java"
105 elif [ -n "$JAVA_HOME" ]; then
106     for java in "$JAVA_HOME"/bin/amd64/java "$JAVA_HOME"/bin/java; do
107         if [ -x "$java" ]; then
108             JAVA="$java"
109             break
110         fi
111     done
112     if [ -z "$JAVA" ]; then
113         echo >&2 "The currently defined JAVA_HOME ($JAVA_HOME) refers"
114         echo >&2 "to a location where Java could not be found. Aborting."
115         echo >&2 "Either fix the JAVA_HOME variable or remove it from the"
116         echo >&2 "environment so that the system PATH will be searched."
117         exit 1
118     fi
119 else
120     JAVA=/opt/java/bin/java
121 fi
122
123 if [ -z "$SOLR_STOP_WAIT" ]; then
124     SOLR_STOP_WAIT=180
125 fi
126 # test that Java exists, is executable and correct version
127 JAVA_VER=$(("$JAVA" -version 2>&1))
128 if [[ $? -ne 0 ]]; then
129     echo >&2 "Java not found, or an error was encountered when running java."
130     echo >&2 "A working Java $JAVA_VER_REQ JRE is required to run Solr!"
131     echo >&2 "Please install latest version of Java $JAVA_VER_REQ or set JAVA_HOME properly."
132     echo >&2 "Command that we tried: '$JAVA' -version, with response:"
133     echo >&2 "$JAVA_VER"
134     echo >&2
135     echo >&2 "Debug information:"
136     echo >&2 "JAVA_HOME: ${JAVA_HOME:-N/A}"
137     echo >&2 "Active Path:"

```

```

Vivianas-MacBook-Pro:solr-6.6.1 vivianamarquez$ mkdir /usr/local/jdk
Vivianas-MacBook-Pro:solr-6.6.1 vivianamarquez$ cp -R /usr/lib/jvm/java-1.8.0-openjdk-1.8.0.65-0.b17.el6_7.x86_64/jre/* /usr/local/jdk/
DLR URL SCHEME=https

```

```

vivianas-mbp:solr-6.6.1 vivianamarquez$ java -version
No Java runtime present, requesting install.
vivianas-mbp:solr-6.6.1 vivianamarquez$ $ which java
-bash: $: command not found
vivianas-mbp:solr-6.6.1 vivianamarquez$ which java
/usr/bin/java
vivianas-mbp:solr-6.6.1 vivianamarquez$ export JAVA_HOME=/usr/bin/java
vivianas-mbp:solr-6.6.1 vivianamarquez$ export PATH=$PATH:/usr/bin/java
vivianas-mbp:solr-6.6.1 vivianamarquez$ java -version
Unable to locate an executable at "/usr/bin/java/bin/java" (-1)
vivianas-mbp:solr-6.6.1 vivianamarquez$ ls
CHANGES.txt          LUCENE_CHANGES.txt  README.txt            contrib              docs                  licenses
LICENSE.txt           NOTICE.txt          bin                   dist                  example                server
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr start
The currently defined JAVA_HOME (/usr/bin/java) refers
to a location where Java could not be found. Aborting.
Either fix the JAVA_HOME variable or remove it from the
environment so that the system PATH will be searched.

```

- Finalmente funcionó usando *brew*:

```

vivianas-mbp:solr-6.6.1 vivianamarquez$ brew cask install java
Error: Unknown command: cask
vivianas-mbp:solr-6.6.1 vivianamarquez$ brew update
^Cvivianas-mbp:solr-6.6.1 vivianamarquez$ cd $(brew --repo); git fetch; git reset --hard origin/master; brew update

remote: Counting objects: 90209, done.
remote: Compressing objects: 100% (22701/22701), done.
remote: Total 90209 (delta 65658), reused 89769 (delta 65263), pack-reused 0
Receiving objects: 100% (90209/90209), 20.21 MiB | 131.00 KiB/s, done.
Resolving deltas: 100% (65658/65658), completed with 300 local objects.
From https://github.com/Homebrew/brew
 0c6e307..c0063bb master -> origin/master
* [new tag]      0.1 -> 0.1
* [new tag]      0.2 -> 0.2
* [new tag]      0.3 -> 0.3
* [new tag]      0.4 -> 0.4
* [new tag]      0.5 -> 0.5
* [new tag]      0.6 -> 0.6
* [new tag]      0.7 -> 0.7
* [new tag]      0.7.1 -> 0.7.1
* [new tag]      0.8 -> 0.8
* [new tag]      0.8.1 -> 0.8.1
* [new tag]      0.9 -> 0.9
* [new tag]      0.9.1 -> 0.9.1
* [new tag]      0.9.2 -> 0.9.2
* [new tag]      0.9.3 -> 0.9.3
* [new tag]      0.9.4 -> 0.9.4
* [new tag]      0.9.5 -> 0.9.5
* [new tag]      0.9.8 -> 0.9.8
* [new tag]      0.9.9 -> 0.9.9
* [new tag]      1.0.0 -> 1.0.0
* [new tag]      1.0.1 -> 1.0.1
* [new tag]      1.0.2 -> 1.0.2
* [new tag]      1.0.3 -> 1.0.3

```

```

1.1.3 -> 1.1.3
* [new tag]      1.1.4 -> 1.1.4
* [new tag]      1.1.5 -> 1.1.5
* [new tag]      1.1.6 -> 1.1.6
* [new tag]      1.1.7 -> 1.1.7
* [new tag]      1.1.8 -> 1.1.8
* [new tag]      1.1.9 -> 1.1.9
* [new tag]      1.2.0 -> 1.2.0
* [new tag]      1.2.1 -> 1.2.1
* [new tag]      1.2.2 -> 1.2.2
* [new tag]      1.2.3 -> 1.2.3
* [new tag]      1.2.4 -> 1.2.4
* [new tag]      1.2.5 -> 1.2.5
* [new tag]      1.2.6 -> 1.2.6
* [new tag]      1.3.0 -> 1.3.0
* [new tag]      1.3.1 -> 1.3.1
* [new tag]      1.3.2 -> 1.3.2
HEAD is now at c0063bb Merge pull request #3156 from MikeMcQuaid/allow-developer-custom-ruby
^Cvivianas-mbp:local vivianamarquezbrew cask install java
==> Creating Caskroom at /usr/local/Caskroom
==> Caveats
This Cask makes minor modifications to the JRE to prevent issues with
packaged applications, as discussed here:

https://bugs.eclipse.org/bugs/show_bug.cgi?id=411361

If your Java application still asks for JRE installation, you might need
to reboot or logout/login.

Installing this Cask means you have AGREED to the Oracle Binary Code
License Agreement for Java SE at

https://www.oracle.com/technetwork/java/javase/terms/license/index.html

==> Satisfying dependencies
==> Downloading http://download.oracle.com/otn-pub/java/jdk/8u144-b01/090f390dda5b47b9b721c7dffa008135/jdk-8u144-macosx-x64.dmg
# 1.7%

```

```
Default
* [new tag] 1.3.0 -> 1.3.0
* [new tag] 1.3.1 -> 1.3.1
* [new tag] 1.3.2 -> 1.3.2
HEAD is now at c0063bb Merge pull request #3156 from MikeMcQuaid/allow-developer-custom-ruby
^Cvivianas-mbp:local vivianamarquezbrew cask install java
=> Creating Caskroom at /usr/local/Caskroom
=> Caveats
This Cask makes minor modifications to the JRE to prevent issues with
packaged applications, as discussed here:

https://bugs.eclipse.org/bugs/show_bug.cgi?id=411361

If your Java application still asks for JRE installation, you might need
to reboot or logout/login.

Installing this Cask means you have AGREED to the Oracle Binary Code
License Agreement for Java SE at

https://www.oracle.com/technetwork/java/javase/terms/license/index.html

=> Satisfying dependencies
=> Downloading http://download.oracle.com/otn-pub/java/jdk/8u144-b01/090f390dda5b47b9b721c7dffa008135/jdk-8u144-macosx-x64.d
##### 100.0%
=> Verifying checksum for Cask java
=> Installing Cask java
=> Running installer for java; your password may be necessary.
=> Package installers may write to any location; options such as --appdir are ignored.
Password:
=> installer: Package name is JDK 8 Update 144
=> installer: Upgrading at base path /
=> installer: The upgrade was successful.
[+] java was successfully installed!
```

4. Solr is running!

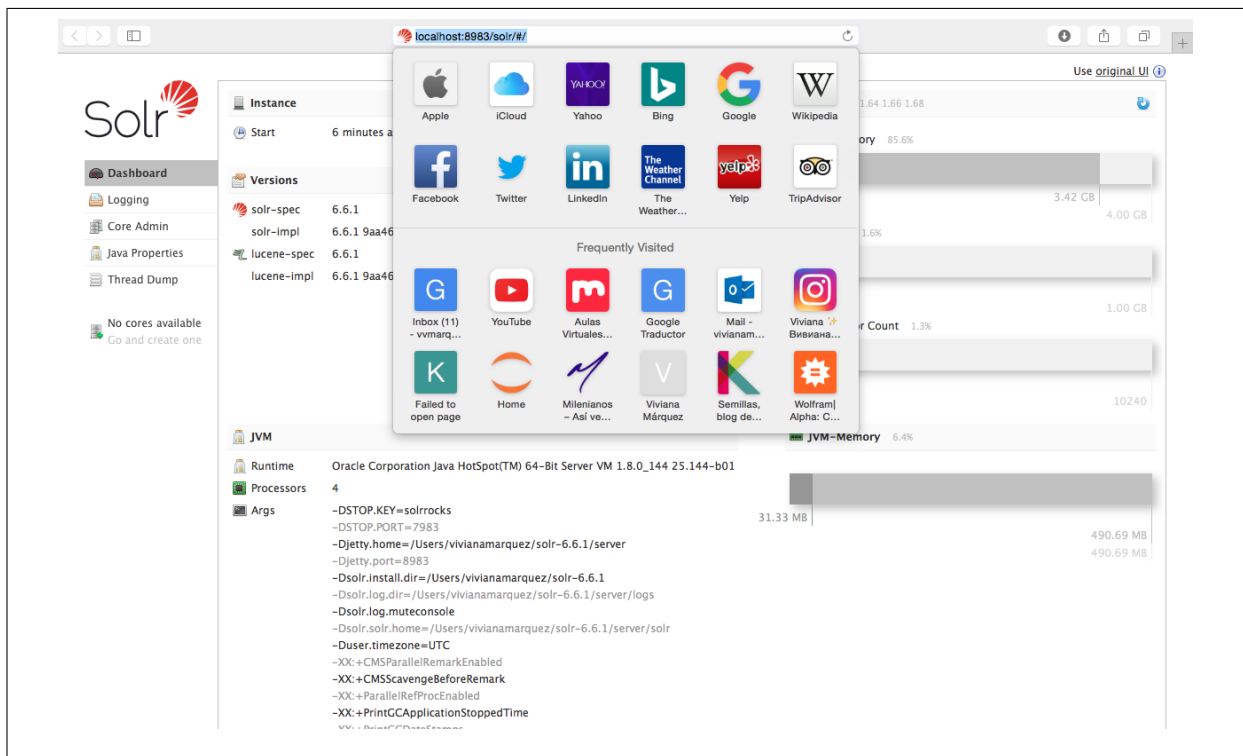
```
vivianas-mbp:~ vivianamarquez$ cd solr-6.6.1
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr start
Waiting up to 180 seconds to see Solr running on port 8983 [\]
Started Solr server on port 8983 (pid=3327). Happy searching!

vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr status

Found 1 Solr nodes:

Solr process 3327 running on port 8983
{
  "solr_home":"/Users/vivianamarquez/solr-6.6.1/server/solr",
  "version":"6.6.1 9aa465a89b64ff2dabe7b4d50c472de32c298683 - varunthacker - 2017-08-29 22:00:32",
  "startTime":"2017-09-16T12:21:22.625Z",
  "uptime":"0 days, 0 hours, 5 minutes, 45 seconds",
  "memory":"49.3 MB (%10) of 490.7 MB"}

vivianas-mbp:solr-6.6.1 vivianamarquez$
```



5. Intentar indexar archivos de ejemplo

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/post -c gettingstarted example/exampledocs/books.json
java -classpath /Users/vivianamarquez/solr-6.6.1/dist/solr-core-6.6.1.jar -Dauto=yes -Dc=gettingstarted -Ddata=files org.apache
solr.util.SimplePostTool example/exampledocs/books.json
SimplePostTool version 5.0.0
Posting files to [base] url http://localhost:8983/solr/gettingstarted/update...
Entering auto mode. File endings considered are xml,json,jsonl,csv,pdf,doc,docx,ppt,pptx,xls,xlsx,odt,odp,ods,ott,otp,ots,rtf,rt
m,html,txt,log
POSTing file books.json (application/json) to [base]/json/docs
SimplePostTool: WARNING: Solr returned an error #404 (Not Found) for url: http://localhost:8983/solr/gettingstarted/update/json
docs
SimplePostTool: WARNING: Response: <html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
<title>Error 404 Not Found</title>
</head>
<body><h2>HTTP ERROR 404</h2>
<p>Problem accessing /solr/gettingstarted/update/json/docs. Reason:
<pre>    Not Found</pre></p>
</body>
</html>
SimplePostTool: WARNING: IOException while reading response: java.io.FileNotFoundException: http://localhost:8983/solr/gettings
arted/update/json/docs
1 files indexed.
COMMITting Solr index changes to http://localhost:8983/solr/gettingstarted/update...
SimplePostTool: WARNING: Solr returned an error #404 (Not Found) for url: http://localhost:8983/solr/gettingstarted/update?comm
t=true
SimplePostTool: WARNING: Response: <html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
<title>Error 404 Not Found</title>
</head>
<body><h2>HTTP ERROR 404</h2>
<p>Problem accessing /solr/gettingstarted/update. Reason:
<pre>    Not Found</pre></p>
</body>
</html>
Time spent: 0:00:00.182
vivianas-mbp:solr-6.6.1 vivianamarquez$
```

6. Crear colección

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr create -c demo
```

Copying configuration to new core instance directory:
/Users/vivianamarquez/solr-6.6.1/server/solr/demo

Creating new core 'demo' using command:
<http://localhost:8983/solr/admin/cores?action=CREATE&name=demo&instanceDir=demo>

```
{
  "responseHeader":{
    "status":0,
    "QTime":3099},
  "core":"demo"}
```

```
vivianas-mbp:solr-6.6.1 vivianamarquez$
```

The screenshot shows the Solr Admin web interface in a browser window. The browser's address bar shows 'localhost'. The page has a sidebar on the left with the Solr logo and navigation links: Dashboard, Logging, Core Admin, Java Properties, and Thread Dump. Below these is the 'Core Selector' which shows a search bar and a list of cores, with 'demo' selected. The main content area is titled 'Request-Handler (qt)' and shows the configuration for the '/select' handler. It includes a 'common' section with a 'q' field set to '*.*'. Below this are fields for 'fq', 'sort', 'start, rows' (with 'start' set to 0 and 'rows' set to 10), 'fl', and 'df'. There is a 'Raw Query Parameters' section with a text area containing 'key1=val1&key2=val2'. At the bottom, there is a 'wt' dropdown set to 'json', and several checkboxes: 'indent' (checked), 'debugQuery', 'dismax', 'edismax', 'hl', 'facet', and 'spatial'.

Solr Admin

Apache Solr -

Solr Tutorial - Getting started with Solr 6

Use original UI

Solr

Dashboard

Logging

Core Admin

Java Properties

Thread Dump

demo

Overview

Analysis

Dataimport

Documents

Files

Ping

Plugins / Stats

Query

Replication

Schema

Segments Info

Statistics

Last Modified: -

Num Docs: 0

Max Doc: 0

Heap Memory: 0

Usage: -

Deleted Docs: 0

Version: 2

Segment Count: 0

Optimized:

Current:

Instance

CWD: /Users/vivianamarquez/solr-6.6.1/server

Instance: /Users/vivianamarquez/solr-6.6.1/server/solr/demo

Data: /Users/vivianamarquez/solr-6.6.1/server/solr/demo/data

Index: /Users/vivianamarquez/solr-6.6.1/server/solr/demo/data/index

Impl: org.apache.solr.core.NRTCachingDirectoryFactory

Replication (Master)

	Version	Gen	Size
Master (Searching)	0	1	71 bytes
Master (Replicable)	-	-	-

Healthcheck

Ping request handler is not configured with a healthcheck file.

Documentation

Issue Tracker

IRC Channel

Community forum

Solr Query Syntax

7. Indexar archivos de ejemplo

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/post -c demo example/exampledocs/books.json
java -classpath /Users/vivianamarquez/solr-6.6.1/dist/solr-core-6.6.1.jar -Dauto=yes -Dc=demo -Ddata=files org.apache.solr.util
SimplePostTool example/exampledocs/books.json
SimplePostTool version 5.0.0
Posting files to [base] url http://localhost:8983/solr/demo/update...
Entering auto mode. File endings considered are xml,json,jsonl,csv,pdf,doc,docx,ppt,pptx,xls,xlsx,odt,odp,ods,ott,otp,ots,rtf,rt
m,html,txt,log
POSTing file books.json (application/json) to [base]/json/docs
1 files indexed.
COMMITting Solr index changes to http://localhost:8983/solr/demo/update...
Time spent: 0:00:00.527
vivianas-mbp:solr-6.6.1 vivianamarquez$
```

, and Current is . The Instance section shows the same configuration. The Replication (Master) section shows a table with columns Version, Gen, and Size, containing rows for Master (Searching) and Master (Replicable). The Healthcheck section shows the same message: 'Ping request handler is not configured with a healthcheck file.'"/>

Solr Admin

Apache Solr -

Solr Tutorial - Getting started with Solr 6

Use original UI

Solr

Dashboard

Logging

Core Admin

Java Properties

Thread Dump

demo

Overview

Analysis

Dataimport

Documents

Files

Ping

Plugins / Stats

Query

Replication

Schema

Segments Info

Statistics

Last Modified: less than a minute ago

Num Docs: 4

Max Doc: 4

Heap Memory: -1

Usage: -

Deleted Docs: 0

Version: 6

Segment Count: 1

Optimized:

Current:

Instance

CWD: /Users/vivianamarquez/solr-6.6.1/server

Instance: /Users/vivianamarquez/solr-6.6.1/server/solr/demo

Data: /Users/vivianamarquez/solr-6.6.1/server/solr/demo/data

Index: /Users/vivianamarquez/solr-6.6.1/server/solr/demo/data/index

Impl: org.apache.solr.core.NRTCachingDirectoryFactory

Replication (Master)

	Version	Gen	Size
Master (Searching)	1505572692913	2	6.78 KB
Master (Replicable)	1505572692913	2	-

Healthcheck

Ping request handler is not configured with a healthcheck file.

Documentation

Issue Tracker

IRC Channel

Community forum

Solr Query Syntax

The screenshot shows the Solr Admin interface in a web browser. The left sidebar contains navigation links: Dashboard, Logging, Core Admin, Java Properties, Thread Dump, and a dropdown menu with demo selected. The main content area is divided into two panels. The left panel, titled 'Request-Handler (qt)', shows the '/select' handler with various query parameters: q (set to '*:*'), fq (empty), sort (empty), start (0), rows (10), fl (empty), df (empty), Raw Query Parameters (key1=val1&key2=val2), wt (set to 'json'), and indent (checked). The right panel displays the JSON response for the query. The response includes a 'responseHeader' with status 0, QTime 38, and params. The 'response' object contains 'numFound': 4, 'start': 0, and 'docs' with two book entries. The first book has id '978-0641723445', cat 'book', name 'The Lightning Thief', author 'Rick Riordan', series 'Percy Jackson and the Olympians', sequence 1, genre 'fantasy', inStock true, price 12.5, pages 384, and version 1578707391922307072. The second book has id '978-1423103349', cat 'book', name 'The Sea of Monsters', author 'Rick Riordan', series 'Percy Jackson and the Olympians', sequence 2, genre 'fantasy', inStock true, price 6.49, and version 1578707391922307072.

8. Cómo borrar una colección, reiniciar y parar Solr

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr create -c demo2
```

Copying configuration to new core instance directory:
/Users/vivianamarquez/solr-6.6.1/server/solr/demo2

Creating new core 'demo2' using command:
<http://localhost:8983/solr/admin/cores?action=CREATE&name=demo2&instanceDir=demo2>

```
{
  "responseHeader":{
    "status":0,
    "QTime":1106},
  "core":"demo2"}
```

```
vivianas-mbp:solr-6.6.1 vivianamarquez$
```

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr delete -c demo2
```

Deleting core 'demo2' using command:
<http://localhost:8983/solr/admin/cores?action=UNLOAD&core=demo2&deleteIndex=true&deleteDataDir=true&deleteInstanceDir=true>

```
{
  "responseHeader":{
    "status":0,
    "QTime":83}}
```

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr restart
Sending stop command to Solr running on port 8983 ... waiting up to 180 seconds to allow Jetty process 5020 to stop gracefully.
Waiting up to 180 seconds to see Solr running on port 8983 [N]
Started Solr server on port 8983 (pid=5570). Happy searching!
```

```
vivianas-mbp:solr-6.6.1 vivianamarquez$
```

```
vivianas-mbp:solr-6.6.1 vivianamarquez$ bin/solr stop
Sending stop command to Solr running on port 8983 ... waiting up to 180 seconds to allow Jetty process 5570 to stop gracefully.
```