

# Solr Indexing Quick Guide

This guide will help you get up and running with a Fedora 4 instance whose updates are automatically indexed in a [Solr](#) repository. This guide glosses over many details and should be considered a starting point for using this feature. The document assumes a [POSIX](#) operating system with [cURL](#), a text editor, [Java](#), and a download of [Apache Solr 4.10.3](#).

## Versions

- [Fedora Webapp Plus 4.7.0](#)
- [Solr 4.10.3](#)
- [Karaf 4.0.5](#)

## Install and Start Fedora 4

### Assumptions

- Fedora 4 is running on port 8080 at context "fcrepo" (with JMS events published at port 61616)
- Your Fedora instance has the [transform service](#) enabled. Since the transform service is not available in the core Fedora webapp, you will likely need to use the [Fedora Webapp Plus](#).

### Verify

1. You should be able to view Fedora in a web browser at the following URL: <http://localhost:8080/fcrepo/rest>

## Install, Configure and Start Solr

### Download Solr

```
wget http://archive.apache.org/dist/lucene/solr/4.10.3/solr-4.10.3.tgz
tar -xzf solr-4.10.3.tgz
```

The location of your untarred Solr installation will be hereinafter referenced as `$SOLR_HOME`.

### Update Solr schema

```
wget https://raw.githubusercontent.com/fcrepo4-exts/fcrepo4-vagrant-base-box/master/config/schema.xml
cp schema.xml $SOLR_HOME/example/solr/collection1/conf/
```

### Start Solr

```
cd $SOLR_HOME/example
java -jar start.jar
```

### Verify

- Navigating to the following URL in a web browser should show the Solr administrative interface: <http://localhost:8983/solr/>

## Install and Start Karaf

### Download Karaf

```
wget http://archive.apache.org/dist/karaf/4.0.5/apache-karaf-4.0.5.tar.gz
tar xvzf apache-karaf-4.0.5.tar.gz
```

The location of your untarred Karaf installation will be hereinafter referenced as `$KARAF_HOME`.

### Start Karaf

```
cd $KARAF_HOME
./bin/karaf
```

## Verify

After running the command above

- you should be presented with some ASCII art in your terminal and
- you should be put into the Karaf client shell, such as:

```

      _ _
     / // _ _ _ _ _ _ _ \ _ _
    / ,< / _ _ `/_ _ _ _ `/_ _
   / || | / _// / / / _// _ _
  /_/_|_|_\_,_/_/_\_,_/_/_
Apache Karaf (4.0.2)

karaf@root(>
```

```
karaf@root( )>
```

- **Note**, to exit the Karaf client shell, type: `CTRL-D`. This will stop the Karaf server, ending the indexing process.
- To run Karaf as a system service please refer to the [Karaf Service Wrapper](#) documentation.

## Install, Configure and Start Fedora Camel Toolbox

## Install Toolbox

In the Karaf client shell type the following:

```
feature:repo-add mvn:org.fcrepo.camel/toolbox-features/4.6.2/xml/features
feature:install fcrepo-service-activemq
feature:install fcrepo-indexing-solr
```

## Verify - Toolbox Installation

Still in the Karaf client shell, the following command

```
feature:list|grep fcrepo
```

should result in both the `fcrepo-camel` and `fcrepo-indexing-solr` features being in the Started state

fcrepo-camel	4.4.3		Started	fcrepo-camel-4.4.3
fcrepo-indexing-solr	4.6.2	x	Started	toolbox-features-4.6.2
fcrepo-ldpath	4.6.2		Started	toolbox-features-4.6.2
fcrepo-service-ldcache-file	4.6.2		Started	toolbox-features-4.6.2
fcrepo-marmotta-osgi	4.6.2		Started	toolbox-features-4.6.2

## Verify - LDPPath

1. You should be able to create and navigate to an existing Fedora resource in the web browser, for example <http://localhost:8080/fcrepo/rest/collection>
2. Assuming the resource is named "collection", you should be able to verify that the LDPATH service is enabled by navigating to the following URL in a web browser: <http://localhost:9086/dpath/collection>
  - a. You should see a JSON document such as:

```
[{"extent":[],"references":[],"prev":[],"altLabel":[],"type":["http://fedora.info/definitions/v4/repository#Container","http://fedora.info/definitions/v4/repository#Resource","http://www.w3.org/ns/ldp#Container","http://www.w3.org/ns/ldp#RDFSSource"],"narrowMatch":[],"relation":[],"accrualMethod":[],"notation":[],"id":["http://localhost:8080/fcrepo/rest/collection"],
...
"lastModifiedBy":["bypassAdmin"],"prefLabel":[],"alternative":[],"label":[],"accessTo":[],"createdBy":["bypassAdmin"],"hiddenLabel":[],"comment":[],"accessRights":[],"sameAs":[]}]
```

## Configure Toolbox

The main configuration of the `fcrepo-indexing-solr` feature is found at: `$KARAF_HOME/etc/org.fcrepo.camel.indexing.solr.cfg`

You will need to make updates to this configuration file if any of the following are true:

- Your Solr is deployed at a URL different than the one detailed earlier in this document
- Your Fedora is deployed at a URL different than the one detailed earlier in this document
- Your Fedora has Authorization enabled, e.g. WebAC

For configuration details, please refer to the documentation found at the [fcrepo-indexing-solr](#) github page. If you updated the `$KARAF_HOME/etc/org.fcrepo.camel.indexing.solr.cfg` file, it is quite likely that you will also need to update the `$KARAF_HOME/etc/org.fcrepo.camel.ldpath.cfg` file, particularly the sections related to Fedora location and authorization. Please refer to the [fcrepo-ldpath](#) page on github for configuration details.

## Success

You should now be able to create/update/delete resources in your Fedora repository, and subsequently see them in your Solr index!

- <http://localhost:8983/solr/#/collection1/query>

## Resources

For debugging purposes, you may want to inspect the logs of the various applications:

- Fedora log (unless configured otherwise): `/var/log/tomcat8/catalina.out`
- Solr log: `$SOLR_HOME/example/logs/solr.log`
- Karaf log: `$KARAF_HOME/data/log/karaf.log`