Shemdros web service

Installation manual

Abstract

The shemdros web service uses a locally installed Emdros engine to access linguistic databases. This document provides an installation manual for the complete RESTful web service.

Project	SHEBANQ
Author	Henk van den Berg
Organization	DANS
Date	April 08, 2014
Audience	System administrators
Version	1.0
Status	Final

Version History

Version	Date	Summary
0.1	April 08, 2014	Initial draft
1.0	July 14, 2014	Final 1.0 version

Table of Contents

1 Installing Emdros

Emdros is a text database engine for analyzed or annotated text. We will use Emdros version 3.4.0 from May 13, 2014. The official release can be found at

http://sourceforge.net/projects/emdros/files/emdros/

Download the source file

http://sourceforge.net/projects/emdros/files/emdros/3.4.0/emdros-3.4.0.tar.gz/download

to a suitable directory.

1.1 Prerequisites

Some development tools are required for building from source.

```
# yum install perl gcc make kernel-headers kernel-devel -y
```

```
# yum groupinstall "Development Tools"
```

```
# yum install zlib-devel
```

For optional integration with Python we will need python-devel

```
# yum install python-devel
```

For integration with Java we best have an official JDK-distribution from Oracle, if the JRE from openjdk will work is untested.

SWIG is used to integrate the Emdros API with several programming languages.

```
# yum install swig
```

MySql should be installed for configure options:

```
# yum install mysql mysql-server
```

```
# yum install mysql-devel
```

1.2 Building and installing Emdros

Create a directory to contain the source distribution.

```
# mkdir /opt/emdros-src
# cd /opt/emdros-src/
```

Copy the source distribution emdros-3.4.0.pre12.tar.gz to this folder. Untar.

```
# tar xfzv emdros-3.4.0.tar.gz
```

Change to the unpacked source folder.

```
# cd emdros-3.4.0/
```

Installation of emdros is detailed in the file INSTALL in this folder. The following instructions are based on information from this file.

On zandbak01 Emdros was build with

```
# ./configure --prefix=/opt/emdros --with-sqlite3=local --with-mysql=yes --
with-swig-language-java=yes --with-swig-language-python=yes --with-jdk-
dir=/usr/java/jdk1.7.0_51/
```

Changing the value of parameter –with-jdk-dir may be necessary.

```
# make
```

```
# make install
```

Emdros is now installed in /opt/emdros.

1.3 Setting up system variables

It can be nice to be able to issue Emdros commands, i.e. `mql`, from the command line. Let's set some global variables.

```
# vim /etc/profile.d/emdros.sh
```

Add the following:

```
EMDROS_HOME=/opt/emdros
export EMDROS_HOME
PATH=$EMDROS_HOME/bin:$PATH
export PATH
```

After reboot the system now is ready:

```
# mql --version
mql from Emdros version 3.4.0.pre02
```

The build-in testing program should end successfully:

```
# emdftry
```

```
...
Backends tested:
                - SQLite 2.X.X
                - SQLite 3.X.X
SUCCESS: All tests completed successfully.
```

Same for …

```
# mqltry
```

Same for …

```
# mqllingtry
```

1.4 Setting up SWIG java bindings

Set symbolic links to `jemdros.jar`, `libjemdros.so` and `libharvest.so`.

```
# mkdir /usr/local/lib/emdros
# cd /usr/local/lib/emdros/
# ln -s /opt/emdros-src/emdros-3.4.0/SWIG/java/jemdros.jar jemdros.jar
# ln -s /opt/emdros-src/emdros-3.4.0/SWIG/java/.libs/libjemdros.so libjemdros.so
# ln -s /opt/emdros-src/emdros-3.4.0/harvest/.libs/libharvest.so libharvest.so
```

Compile and run `TestEmdros.java`:

```
# mkdir /opt/emdros-java
# cd /opt/emdros-java
# cp /opt/emdros-src/emdros-3.4.0/SWIG/java/TestEmdros.java .
# javac -cp /usr/local/lib/emdros/jemdros.jar:. TestEmdros.java
# java -cp /usr/local/lib/emdros/jemdros.jar:. TestEmdros
```

With the last line a test program is run and should return:

```
Linux
Connection established.
...
All tests completed successfully.
```

1.5 Setting up SWIG python bindings

Find out which is the Python site-packages directory:

```
# cd /opt/Emdros-src/emdros-3.4.0.pre02/SWIG/python/
# python dir.py
```

```
/usr/lib/python2.6/site-packages
```

Set 4 symbolic links in the directory found by dir.py:

```
# ln -s /opt/emdros-src/emdros-3.4.0.pre02/SWIG/python/.libs/_EmdrosPy.so
/usr/lib/python2.6/site-packages/
# ln -s /opt/emdros-src/emdros-3.4.0.pre02/SWIG/python/EmdrosPy.py
/usr/lib/python2.6/site-packages/
# ln -s /opt/emdros-src/emdros-3.4.0.pre02/SWIG/python/EmdrosPy.pyc
/usr/lib/python2.6/site-packages/
# ln -s /opt/emdros-src/emdros-3.4.0.pre02/SWIG/python/EmdrosPy.pyo
/usr/lib/python2.6/site-packages/
```

Run `test.py`:

```
# python test.py
```

A minimal test program is run and returns:

```
string
{ 1-5, 7-10 }
Query returned: no DB errors:  True  no compiler errors:  True
Query returned: no DB errors:  True  no compiler errors:  True
```

1.6 A SQLite database for testing

The wivu dataset in SQLite and a collection of queries can be copied from zandbak01. The database and queries can be used for testing. On zandbak01 they are in

```
/data/emdros/wivu/s3/bhs3
```

and

```
/data/emdros/wivu/queries/
```

2 Installing shemdros

Shemdros is a Restful web service wrapping the Emdros engine. It interfaces with Emdros by way of the SWIG java binding for the mql-API. The experimental renderObjects-API has no SWIG binding and is called by command line invocations from java.

2.1 Deploying shemdros

Shemdros source code is at https://github.com/dans-er/shemdros . At this moment we describe installation of shemdros version 1.0.

In the directory `/opt/shemdros` on zandbak01 we find everything we need.

The directory `/opt/shemdros` is used as the home for shemdros, containing the war-file for deployment under Tomcat, the `shemdros.xml`, which can be copied to the Tomcat directory `/usr/share/tomcat6/conf/Catalina/localhost` for easy deployment.

The current deployment is regulated with Spring application context which, after deployment, can be found at

`/usr/share/tomcat6/webapps/shemdros/WEB-INF/classes/application-context.xml`.

Logging is regulated with

`/usr/share/tomcat6/webapps/shemdros/WEB-INF/classes/logback.xml`, and logs to

`/var/log/shemdros/shemdros.log`.

After this the service at `localhost:8080/shemdros` should be made accessible in Apache configuration.

2.2 Configuration of shemdros

With the configuration file

`/usr/share/tomcat6/webapps/shemdros/WEB-INF/classes/application-context.xml`.

several options can be set.

2.2.1 Environment

The environment takes two constructor values, the location of libjemdros.so and libharvest.so which we set with symbolic links under 1.4.

```
<bean class="nl.knaw.dans.shemdros.core.Environment" name="environment">
        <constructor-arg value="/usr/local/lib/emdros/libjemdros.so"/>
        <constructor-arg value="/usr/local/lib/emdros/libharvest.so"/>
</bean>
```

2.2.2 Databases

Required is at least one database under the name 'default'. Multible databases can be set.

```
<bean class="nl.knaw.dans.shemdros.core.Database" name="defaultDB">
        <constructor-arg value="default"/>
        <property name="outputKind" value="1"/>
        <property name="charset" value="3"/>
        <property name="backendKind" value="4"/>
        <property name="hostname" value="localhost"/>
        <property name="username" value=""/>
        <property name="password" value=""/>
        <property name="initialDB" value="/data/emdros/wivu/s3/bhs3"/>
        <property name="maxPoolSize" value="10"/>
</bean>
```

See the API for EmdrosEnv for meaning and possible values of the properties: http://emdros.org/progref/current/1160.html

2.2.3 JsonFiles

The option to set stylesheet files known as JsonFiles for the renderObjects api is no longer used.

2.2.4 EmdrosFactory

The EmdrosFactory bean puts all together:

```
<bean class="nl.knaw.dans.shemdros.core.EmdrosFactory" name="emdrosFactory">
        <property name="databases">
                <list>
                        <ref bean="defaultDB"/>
                        <ref bean="otherDB"/>
                </list>
        </property>
</bean>
```

2.3 Testing shemdros

After installation and deployment a webbrowser targeted at ../shemdros/version should give:

```
Shemdros-version: 1.0
Shemdros-build: 2014-07-07 14:43
MQL-version: mql from Emdros version 3.4.0
```