



SEAD Virtual Archive: User Guide

Release V1.0 (November 2013)

SEAD Virtual Archive: User Guide

Steps to install:

1. [Prerequisites](#)
2. [Checkout Code](#)
3. [Installation](#)
4. [Running Services](#)

Pre-requisites:

Before setting up the Virtual Archive service, please complete the following tasks:

- **Creating an ACR account:**
 - Please request an account in Active Content Repository at <http://sead.ncsa.illinois.edu/acr/> or at <http://nced.ncsa.illinois.edu/acr/>
- **Build requirements:**
 - For building code please use **maven** (version 2.2.1 only currently) and **jdk** (1.7.x, does not work for 1.6)
 - Please setup latest version of [Apache Tomcat](#)
 - Once tomcat is setup, please copy [Hibernate jar](#) into the lib folder inside tomcat folder.
 - Increase the stack size as with JAVA_OPTS environment variable for jvm and to also allow encoding of slash and backslash as shown below.

```
export JAVA_OPTS="-Xms512m -Xmx1024m -XX:MaxPermSize=1024m  
-Dorg.apache.tomcat.util.buf.UDecoder.ALLOW_ENCODED_SLASH=true  
-Dorg.apache.catalina.connector.CoyoteAdapter.ALLOW_BACKSLASH=true"
```

Checkout Code

Note: Please note that the code in Subversion in bitternut.cs.indiana.edu/home/dquob/svn-repo/sead/dc_code has latest DC code (revision r6929) in the branches. The branches will be merged with the main code the week of Nov 5. You can directly download VA extension code from IU Github. The build command packages code by downloading dependencies from DC maven repository. So there is no need for building DC source code from subversion branches at this point.

Github (Preferred)

The code for Sead VA extension modules is available at Github: <https://github.com/Data2Insight/sead-virtual-archive>

The following steps can be used to download code:

Option #1: Download the zipped version of code.

Option #2: Use git commands to checkout the code with steps below.

```
git init

git remote add origin git@github.com:Data2Insight/sead-virtual-archive.git

git pull origin master
```

Subversion (Not needed for basic installation)

On Linux, one can use the following command to check out the code:

```
svn co svn+ssh://<username>@bitternut.cs.indiana.edu/home/dquob/svn-repo/sead
```

On Windows, please use the same url above in an svn client like [Tortoise client](#)

Installation

1. **Build Pre-requisites**
2. **Build Code**
3. **Install using tomcat**

1) Build Pre-requisites

- cd into the directory containing the code you checked out in the previous step.
- Before building the code, **please copy the settings.xml file** from maven/conf folder in the git repository into your maven folder.
- **Setting up Database:**
 - Please use the jar in `utils/initDB` folder to initialize the database. Command to initialize the db is "java -jar initDB.jar /path/to/database" eg: java -jar initDB.jar /tmp/database will create a derby database in the /tmp/database later needed for the UI.

Building backend workflow

```
cd SEAD-VA-extensions/dcs-integration/sead-workflow-integration
```

2 a) Build:

To clean and build code run:

```
mvn clean
```

```
mvn package (preferred mvn package -DskipTests, in case any tests fail at this point)
```

The above maven build generates a .war file in target folder.

3 a) Install: Please deploy this war in a servlet container. If you are using tomcat, please copy to webapp folder. When copying, please rename to `sead-wf.war`, since this is the value that is set as 'prefix' in `WEB-INF/classes/default.properties` file in the war. Please also change any values you want to change in `default.properties` file as needed.

The endpoints enabled would include

SIP Deposit Servlet at <http://localhost:8080/sead-wf/deposit/sip>

Query Servlet at <http://localhost:8080/sead-wf/deposit/squery>

Data stream Servlet at <http://localhost:8080/sead-wf/deposit/datastream>

Delete Servlet at <http://localhost:8080/sead-wf/deposit/del>

Building front-end UI

```
cd ..
```

```
cd SEAD-VA-extensions/dcs-access/sead-access-ui
```

2 b) Build:

mvn clean compile [gwt:compile](#) package

3 b) Install:

The above maven commands will clean, compile and package the UI code. It will generate a war file which also needs to be copied to tomcat webapp (rename war to sead-access.war when copying).

Setting up the configuration file:

During the installation process, sead-access.war file must have been copied into tomcat/webapp folder. Once deployed, please set values in the following configuration files:

1) tomcat/webapps/sead-access/sead_access/Config.properties looks like this:

```
accessServiceURL = http://localhost:8080/sead-wf/
bagItServiceURL = http://localhost:8080/bagit/acrToBag/
datastreamServletUrl = http://localhost:8080/sead-wf/datastream/
dbUrl = jdbc:derby:/tmp/database
tmp = /tmp/stage/
acrusername = your-acr-username
admins = admin-notification-email-address-here
```

In the above, please set your ACR username in acrusername and in admins, list one admin email address/ if there are multiple admins, list all admins' email addresses separated by semi-colon. eg: admin@indiana.edu; admin2@indiana.edu.

Also, please [install BagIt service](#) on localhost before moving to the next step.

Please change other URLs based on where the workflow service is hosted on your machine.

2) tomcat/webapps/sead-access/WEB-INF/classes/passwords.xml looks like this

```
<entities>
<entity>
<name>acr</name>
<password>your-acr-password</password>
</entity>
<entity>
```

```
<name>email-sender</name>
<password>email-sender-password</password>
</entity>
</entities>
```

Replace your-acr-password with your ACR password in this file. ACR account would have been created in the pre-requisite step.

email-sender-password needs to be replaced by the password of the email sender specified in the configuration file.

The UI will be accessible at localhost:8080/sead-access

Running the service:

Once installation is complete, you can access the UI at <http://localhost:8080/sead-access>

Faceted Browsing is available on the left pane under 'Data Search' tab. Data can be ingested into Virtual Archive using the 'Upload Data' tab which needs login.

Test Cases:

Admin Login:

To login, the database was initialized with admin username as 'seadva@gmail.com' and password as 'password'.

Creating new user:

The register button is used to register a new user. When logged in as 'Admin' (eg: seadva@gmail.com), an 'Administration' tab is visible, where the admin can select a role for a new user and approve the user.

Google Login:

A user can register using Google Login. If it is the first time a user logs in using Google, they need to be approved by the admin before they are able to ingest data sets into Virtual Archive.

Ingesting datasets:

In the 'Upload Data' tab, please login using admin credentials or any other login that was created. Now select "Nced Project" and click on "View Collections to Publish". 20 datasets in this case will be displayed that can be published.

For testing purposes, we recommend ingesting from the smaller collections like:

William Brownlie Alluvial Channel Data
Eel River Quads
Church and Rood Alluvial River Channel Regime Data
Experimental Study of Delta Erosion Due to Dam Removal
Eel River Steelhead Study
Eel River Flipchart

Once you click on one of these collections, the collection metadata is retrieved. You can view folder/collection on left pane and files within it on the right pane. Please click the checkbox on the top-level collection on the left pane, and all the files and subcollections within will be selected for ingest.

Now click on the 'Ingest Dataset' button. The ingest workflow will now start. You can view the status in the "Data Ingest Monitor" pane. A red colored notification icon that will appear on the top blue pane, to notify the user when the

workflow finishes executing.

For smaller datasets, depending on internet connectivity, the ingest workflow takes a minute or so.

Viewing datasets:

You can browse the collection and files under 'Data Search' tab. The data is also available under faceted search on the left pane, if there was sufficient metadata for the collection in ACR.

Development:

For developing code, please import any project (backend or UI or other jars) as a maven project into your IDE (eg: Eclipse, Idea IntelliJ).