

## HAT Installation Guide

It is advised that the Cent O/S VM no longer be used, please use the Ubuntu image

**Download HAT VM from test server** <http://137.149.157.11/files/Ubu-CS482-Mar102011.ova>

- Install into VirtualBox 4.0 or greater
  - username: islandora
  - password: islandora
- use sudo su to change to root user
- change ProxyRequests value in /etc/apache2/sites-available/islandora on line 59 to On
  - sudo vim /etc/apache2/sites-available/islandora
- Disable the old proxy module, enable the new with commands:
  - a2dismod
    - proxy
  - a2enmod
    - proxy\_http
- Restart apache
  - /etc/init.d/apache2 restart

**Checkout latest HAT repository from** <http://github.com/posif/HAT>

- Copy HAT/iiv into /usr/local/fedora/tomcat/webapps/iiv
- Change file permissions to ownership by fedora
  - sudo chown -R fedora:fedora /usr/local/fedora/tomcat/webapps/iiv
- Install the iiv\_annotation module
  - copy HAT/iiv\_annotation to /var/www/html/drupal/sites/all/modules/iiv\_annotation on VM
  - sudo chown -R www-data:www-data /var/www/html/drupal/sites/all/modules/iiv\_annotation
  - log into the drupal site at <http://192.168.56.101/user>
    - username: admin
    - password: islandora
  - using the administrative menu, select "Site Building-->Modules"
    - Under the "Fedora Repository" fieldset there will be a IIV\_Annotation option, enable it, scroll to the bottom and submit
  - Be certain to set the Drupal permissions for the module
- Repeat module installation steps to install the guardian\_solr\_config module
  - Update the Solr Search Block admin page to output results from the guardian\_solr\_config module by visiting [http://192.168.56.101/admin/settings/islandora\\_solr\\_search](http://192.168.56.101/admin/settings/islandora_solr_search)
    - Set Configuration Package to Guardian Output

Checkout the latest from from <http://github.com/posif/islandora>

- Copy posif islandora files into /var/www/html/drupal/sites/all/modules/islandora

Grab the latest IIV copy from <http://github.com/posif/iiv>

- follow readme and install war to /usr/local/fedora/tomcat/webapps

## Update newspapers:viewerSdep-issueCModel in fedora

Go to <http://192.168.56.101:8080/fedora/admin> (default credentials are fedoraAdmin fedoraAdmin)

Search for newspapers:viewerSdep-issueCModel

Open the data object newspapers:viewerSdep-issueCModel

For the datastream METHODMAP, edit and replace with

```
<fmm:MethodMap xmlns:fmm="http://fedora.comm.nsdlib.org/service/methodmap"
name="methodmap">

<fmm:Method operationName="getViewer" wsdlMsgName="getViewerRequestMsg"
wsdlMsgOutput="getViewerResponseMsg">
  <fmm:DefaultInputParm defaultValue="$PID" label="fedora object pid" paramName="PID"
passBy="VALUE" required="true"></fmm:DefaultInputParm>
  <fmm:DefaultInputParm defaultValue="newspapers:issueCModel" label="content model"
paramName="CMODEL" passBy="VALUE" required="true"></fmm:DefaultInputParm>
  <fmm:DefaultInputParm defaultValue="JP2" label="content model" paramName="DSID"
passBy="VALUE" required="true"></fmm:DefaultInputParm>

<fmm:UserInputParm defaultValue="" paramName="uid" passBy="VALUE" required="false"></
fmm:UserInputParm>
  <fmm:UserInputParm defaultValue="" paramName="drupal_uid" passBy="VALUE"
required="false"></fmm:UserInputParm>
  <fmm:UserInputParm defaultValue="" paramName="drupal_domain" passBy="VALUE"
required="false"></fmm:UserInputParm>

<fmm:UserInputParm defaultValue="" paramName="search" passBy="VALUE"
required="false"></fmm:UserInputParm>
  <fmm:MethodReturnType wsdlMsgName="getViewerResponseMsg" wsdlMsgTOMIME="text/
html"></fmm:MethodReturnType>
</fmm:Method>
</fmm:MethodMap>
```

For the datastream WSDL, edit and replace with-

```
<wsdl:definitions xmlns:http="http://schemas.xmlsoap.org/wsdl/http/"
xmlns:mime="http://schemas.xmlsoap.org/wsdl/mime/" xmlns:soap="http://
schemas.xmlsoap.org/wsdl/soap" xmlns:soapenc="http://schemas.xmlsoap.org/wsdl/
```

```

soap/encoding" xmlns:this="viewerSdep" xmlns:wsdl="http://schemas.xmlsoap.org/
wsdl/" xmlns:xsd="http://www.w3.org/2001/XMLSchema" name="viewerSdep"
targetNamespace="viewerSdep">
  <wsdl:message name="getViewerRequestMsg">
    <wsdl:part name="PID" type="xsd:string"></wsdl:part>
    <wsdl:part name="CMODEL" type="xsd:string"></wsdl:part>
    <wsdl:part name="DSID" type="xsd:string"></wsdl:part>
    <wsdl:part name="uid" type="xsd:string"></wsdl:part>
    <wsdl:part name="drupal_uid" type="xsd:string"></wsdl:part>
    <wsdl:part name="drupal_domain" type="xsd:string"></wsdl:part>
    <wsdl:part name="search" type="xsd:string"></wsdl:part>
  </wsdl:message>
  <wsdl:message name="getViewerResponseMsg">
    <wsdl:part name="RESPONSE" type="xsd:string"></wsdl:part>
  </wsdl:message>
  <wsdl:portType name="viewer_portType">
    <wsdl:operation name="getViewer">
      <wsdl:input message="this:getViewerRequestMsg"></wsdl:input>
      <wsdl:output message="this:getViewerResponseMsg"></wsdl:output>
    </wsdl:operation>
  </wsdl:portType>
  <wsdl:service name="viewer_service">
    <wsdl:port binding="this:viewer_binding" name="viewer_port">
      <http:address location="http://localhost:8080/iiv/viewer.jsp"></http:address>
    </wsdl:port>
  </wsdl:service>
  <wsdl:binding name="viewer_binding" type="this:viewer_portType">
    <http:binding verb="GET"></http:binding>
    <wsdl:operation name="getViewer">
      <http:operation location="?pid=(PID) &amp;cmodel=(CMODEL) &amp;dsid=(DSID)
&amp;uid=(uid) &amp;drupal_uid=(drupal_uid) &amp;drupal_domain=(drupal_domain)
&amp;search=(search)"></http:operation>
      <wsdl:input>
        <http:urlReplacement></http:urlReplacement>
      </wsdl:input>
      <wsdl:output>
        <mime:content type="text/html"></mime:content>
      </wsdl:output>
    </wsdl:operation>
  </wsdl:binding>
</wsdl:definitions>

```

## Update newspapers:viewerSdep-pageCModel in fedora

Go to <http://192.168.56.101:8080/fedora/admin> (default credentials are fedoraAdmin fedoraAdmin)

Search for newspapers:viewerSdep-pageCModel

Open the data object newspapers:viewerSdep-pageCModel

For the datastream METHODMAP, edit and replace with

```
<fmm:MethodMap xmlns:fmm="http://fedora.comm.nsdlib.org/service/methodmap"
name="methodmap">

<fmm:Method operationName="getViewer" wsdlMsgName="getViewerRequestMsg"
wsdlMsgOutput="getViewerResponseMsg">
  <fmm:DefaultInputParm defaultValue="$PID" label="fedora object pid" paramName="PID"
passBy="VALUE" required="true"></fmm:DefaultInputParm>
  <fmm:DefaultInputParm defaultValue="newspapers:pageCModel" label="content model"
paramName="CMODEL" passBy="VALUE" required="true"></fmm:DefaultInputParm>
  <fmm:DefaultInputParm defaultValue="JP2" label="content model" paramName="DSID"
passBy="VALUE" required="true"></fmm:DefaultInputParm>

<fmm:UserInputParm defaultValue="" paramName="uid" passBy="VALUE" required="false"></
fmm:UserInputParm>
  <fmm:UserInputParm defaultValue="" paramName="drupal_uid" passBy="VALUE"
required="false"></fmm:UserInputParm>
  <fmm:UserInputParm defaultValue="" paramName="drupal_domain" passBy="VALUE"
required="false"></fmm:UserInputParm>

<fmm:UserInputParm defaultValue="" paramName="search" passBy="VALUE"
required="false"></fmm:UserInputParm>
  <fmm:MethodReturnType wsdlMsgName="getViewerResponseMsg" wsdlMsgTOMIME="text/
html"></fmm:MethodReturnType>
</fmm:Method>
</fmm:MethodMap>
```

For the datastream WSDL, edit and replace with-

```
<wsdl:definitions xmlns:http="http://schemas.xmlsoap.org/wsdl/http/"
xmlns:mime="http://schemas.xmlsoap.org/wsdl/mime/" xmlns:soap="http://
schemas.xmlsoap.org/wsdl/soap" xmlns:soapenc="http://schemas.xmlsoap.org/wsdl/
soap/encoding" xmlns:this="viewerSdep" xmlns:wsdl="http://schemas.xmlsoap.org/
wsdl/" xmlns:xsd="http://www.w3.org/2001/XMLSchema" name="viewerSdep"
targetNamespace="viewerSdep">
  <wsdl:message name="getViewerRequestMsg">
    <wsdl:part name="PID" type="xsd:string"></wsdl:part>
    <wsdl:part name="CMODEL" type="xsd:string"></wsdl:part>
    <wsdl:part name="DSID" type="xsd:string"></wsdl:part>
    <wsdl:part name="uid" type="xsd:string"></wsdl:part>
    <wsdl:part name="drupal_uid" type="xsd:string"></wsdl:part>
    <wsdl:part name="drupal_domain" type="xsd:string"></wsdl:part>
    <wsdl:part name="search" type="xsd:string"></wsdl:part>
  </wsdl:message>
  <wsdl:message name="getViewerResponseMsg">
    <wsdl:part name="RESPONSE" type="xsd:string"></wsdl:part>
  </wsdl:message>
  <wsdl:portType name="viewer_portType">
```

```

    <wsdl:operation name="getViewer">
      <wsdl:input message="this:getViewerRequestMsg"></wsdl:input>
      <wsdl:output message="this:getViewerResponseMsg"></wsdl:output>
    </wsdl:operation>
  </wsdl:portType>
  <wsdl:service name="viewer_service">
    <wsdl:port binding="this:viewer_binding" name="viewer_port">
      <http:address location="http://localhost:8080/iiv/viewer.jsp"></http:address>
    </wsdl:port>
  </wsdl:service>
  <wsdl:binding name="viewer_binding" type="this:viewer_portType">
    <http:binding verb="GET"></http:binding>
    <wsdl:operation name="getViewer">
      <http:operation location="?pid=(PID) &amp;cmodel=(CMODEL) &amp;dsid=(DSID)
&amp;uid=(uid) &amp;drupal_uid=(drupal_uid) &amp;drupal_domain=(drupal_domain)
&amp;search=(search)"></http:operation>
      <wsdl:input>
        <http:urlReplacement></http:urlReplacement>
      </wsdl:input>
      <wsdl:output>
        <mime:content type="text/html"></mime:content>
      </wsdl:output>
    </wsdl:operation>
  </wsdl:binding>
</wsdl:definitions>

```

## Solar Multiword Index Creation

Solr indexing allows efficient processing of strings in the wordsxy file.

NOTE: An equivalent index could be made a little more efficiently by using the basic OCR

NOTE: Careful if copy/pasting; double quotes may be messed up.

1. Stop fedora. (/etc/init.d/fedora stop)
2. Setup to index WORDSXY as dsm.WORDSXY
  - a. Add to (demo)FoxmlToSolr.xslt (\$FEDORA\_HOME/tomcat/webapps/fedoragsearch/WEB-INF/classes/config/index/gsearch\_solr/demoFoxmlToSolr.xslt) :
 

```

<xsl:for-each select="foxml:datastream[@ID='WORDSXY']">
  <field>
    <xsl:attribute name="name">
      <xsl:value-of select="concat('dsm.', @ID)"/>
    </xsl:attribute>
    <xsl:value-of select="exts:getDatastreamText($PID, $REPOSITORYNAME,
@ID, $FEDORASOAP, $FEDORAUSER, $FEDORAPASS, $TRUSTSTOREPATH,
$TRUSTSTOREPASS)"/>
  </field>
</xsl:for-each>

```
  - b. Update schema.xml (\$FEDORA\_HOME/gsearch\_solr/solr/conf/schema.xml):
    - i. Add type to schema.xml:
 

```

<fieldType name="wordsxy_text" class="solr.TextField">

```

```

<analyzer>
  <charFilter class="solr.HTMLStripCharFilterFactory"/>
  <tokenizer class="solr.StandardTokenizerFactory"/>
  <filter class="solr.StandardFilterFactory"/>
  <filter class="solr.LowerCaseFilterFactory"/>
  <!--Eliminate preceding number, quotes, and right angle
bracket; may not be necessary, or in fact useless given the change
to Standard Tokenizer/Filter...-->
  <filter class="solr.PatternReplaceFilterFactory"
pattern="([0-9]*(\u0022|(\u0026amp;\u003B))(\u003E|(\u0026gt;\u003B))
)" replacement="" replace="all" />
  <filter class="solr.TrimFilterFactory"
updateOffsets="true"/>
</analyzer>
</fieldType>

```

ii. Add field to schema.xml:

```

<field name="dsm.WORDSXY" type="wordsxy_text" indexed="true"
stored="true" multiValued="true" termVectors="true"
termPositions="true" termOffsets="
true"/>

```

3. Update solrconfig.xml (\$FEDORA\_HOME/gsearch\_solr/solr/conf/solrconfig.xml):

a. Add search component:

```

<searchComponent name="TVC"
class="org.apache.solr.handler.component.TermVectorComponent"/>

```

b. Add request handler:

```

<requestHandler name="guardian_words" class="solr.SearchHandler">
  <lst name="defaults">
    <str name="fl">PID, dc.title, type, score</str>
    <str name="fq">PID:guardian*</str>
    <str name="qf">PID^0.5 dc.title dsm.WORDSXY^2.0</str>
    <str name="pf">dsm.WORDSXY^4.0</str>
    <str name="mm">100%</str>
    <str name="ps">20</str>
  </lst>
  <arr name="last-components">
    <str>TVC</str>
  </arr>
</requestHandler>

```

4. Start fedora (/etc/init.d/fedora start)

5. Actually index the files

- Go to <islandora-url>:8080/fedoragsearch/rest
- Hit "update index"
- Enter the "Foxml files path:" field (/usr/local/fedora/data/objectStore)
- Hit "updateIndex from Foxml files"
  - This may take a while; should return to the same screen.

6. Modify configuration of "Islandora Solr Client" in Drupal

- Change to use the "proper" Solr instance: localhost:8080/solr
- Change to use the new request handler (requires submission of previous step)

- c. Change "facet fields" to:  
dc.title,dsm.WORDSXY
- 7. Test Islandora Solr Client:
  - a. Search for "certain gentlemen took up the movement" (guardian:18901216-00s-003-001),  
or some other phrase...