

Inhaltsverzeichnis

1	Hinweise	1
2	Voraussetzungen	2
3	Überblick	2
4	Installation	3
4.1	WissGrid-Installationsprogramm	3
4.2	Installationskomponenten	5
4.3	Zusätzliche Komponenten	9
5	Mögliche Probleme bei der Installation	13
6	Eingabehinweise	14
6.1	iRODS - Installationsangaben	14
6.2	Fedora Installationsangaben	15
6.3	OAI Provider - Beispiel für <code>Identify.xml</code>	17
6.4	Fedora GSearch - Ausführliche Einstellungen	18
6.5	Pazpar2 - Einstellungen mit Solr für Debian	22
6.6	Apache HTTP Server - Debian Proxy Einstellungen	24

Installationsanleitung

1 Hinweise

Diese Installationsanleitung (README) liegt nach der Installation im Verzeichnis `doc/` des Installationsverzeichnis, und sie ist zudem auf der WissGrid Web-Seite abrufbar unter: D-3.5.3 - <http://www.wissgrid.de/publikationen/deliverables/wp3.html>.

- Wenn das Datenbankmanagementsystem, z.B. Postgres, im Rahmen der iRODS-Installation geladen und installiert wird, dann ist die Installation des iRODS-Servers

sehr zeitaufwändig (>30 Minuten). Bei Nutzung einer vorliegende Datenbank-Installation (Postgres, mysql, ...) verläuft die Installation wesentlich schneller

- Bei der Installation der Komponenten ist die angegebene Reihenfolge einzuhalten, weil z.B. die Web-Service Komponente (wissgridservices) installiert und gestartet sein muss, damit bei der iRODS Installation, Code für Service-Aufrufe generiert werden kann.
- Das System wurde unter Tomcat 6.x getestet, unter Tomcat 7.x gab es Probleme mit Fedora (für weiteres wird auf die Fedora-Webseite verwiesen). Im folgenden wird die Installation mit Tomcat 6.x beschrieben, wobei Tomcat ausreichend Heap-Speicher zugewiesen werden muss.
- Für den Einsatz des Apache HTTP Server genügt es, wenn man diesen vom Linux Distributions Repository installiert. Dabei muss man beachten, dass 'mod_proxy' und 'mod_proxy_ajp' zusätzlich installiert werden.

2 Voraussetzungen

1. Mindestens 1024MB Arbeitsspeicher.
2. Java JRE in Version 1.6.x
3. Tomcat in Version 6.x
4. In Konsole/Terminal testen, ob die Entwicklungstools 'g++', 'yacc', 'flex' auf dem vorliegenden System installiert sind, ggf. müssen diese nach-installiert werden (ohne diese Tools scheitert die Installation).

Test, z.B. durch Abfragen der Version:

```
g++ -v
yacc -V
flex -V
```

3 Überblick

Die Installation erfolgt in zwei Etappen:

1. Ausführen des WissGrid-Installationsprogramms, welches einzelne Parameter abfragt und die Installationskomponenten (wissgridservices, Fedora, iRODS, Änderungs-/Erweiterungsdateien für iRODS, gSOAP, Solr) in ein Installations Verzeichnis kopiert (siehe 4.1).

Dieses Verzeichnis wird nachfolgend Bezeichnet mit: **Installationsverzeichnis** oder **Installations-Pfad**.

2. Anschließend müssen die Installationskomponenten (iRODS, Fedora, ...) installiert werden. Dieser Schritt wird im wesentlichen über Shell-Skripte gesteuert (siehe 4.2). Bei den Verzeichnissen, in welche diese Komponenten installiert werden, wird von Home-Verzeichnis oder Installationspfad der Komponente gesprochen (IRODS_HOME oder IRODS-Installationspfad).

Bei einer Testinstallation ist es empfehlenswert alle Komponenten (iRODS, Fedora, ...) auf einem Host zu installieren. Es ist aber auch möglich und sinnvoll, eine verteilte Installation vorzunehmen. Die lokale Installation erfolgt entsprechend den Ausführungen in dieser README. Für die verteilte Installation muss der Installer auf allen Hosts ausgeführt werden, wobei nur die Installationsskripte der Komponenten, die auf dem betreffenden Host von Interesse sind, ausgeführt werden. Bei der verteilten Installation sind zudem die in Abschnitt 5.5 der Dokumentation¹ beschriebenen Anpassungen an den Konfigurationsdateien zu beachten.

4 Installation

Zunächst sind die folgenden Umgebungsvariablen zu setzen. Dies ist nicht nur für den Betrieb wichtig, sondern vereinfacht auch die Installation, da über die Umgebungsvariablen auch die Pfade in den zugehörigen Eingabemasken des Installationsprogramms korrekt vorbelegt werden.

- `export IRODS_HOME=<iRODS-Installationspfad>`
- `export FEDORA_HOME=<Fedora-Installationspfad>`
- `export SOLR_HOME=<Solr-Installationspfad>`
- `export CATALINA_HOME=<Tomcat-Installationspfad>`
- `export JAVA_OPTS="-server -Xms768M -Xmx1024M -XX:MaxPermSize=768M"`
- `PATH=$PATH:$IRODS_HOME/clients/icommands/bin:
$IRODS_HOME/`

Es bietet sich an, die Umgebungsvariablen in der Datei `~/.profile` zu setzen, womit sie auch nach einem erneuten Systemstart korrekt gesetzt sind (Ausdrücke in `<..>` müssen auf lokale Gegebenheiten angepasst werden, z.B. `export IRODS_HOME=/opt/iRODS`). Damit die Änderungen auch übernommen werden, ist der Aufruf `source ~/.profile` auf der Kommandozeile erforderlich.

4.1 WissGrid-Installationsprogramm

Beim Auslieferungspaket `WissGrid_D3.5.3-grid-repository-fedora-irods-installer.jar` handelt es sich um ein ausführbares Java-Archiv, welches den ersten o.g. Installationsschritt durchführt. Der Aufruf erfolgt durch:

¹siehe: `<Installationspfad>/doc/WissGrid-D3.5.3-grid-repository-fedora-irods-Dokumentation.pdf`

```
java -jar WissGrid_D3.5.3-grid-repository-fedora-irods-installer.jar
```

Der Installer umfasst die folgenden Schritte (bzw. Seiten):

1. Allgemeine Infoseite
2. Informationen (diese README)
3. Lizenzseite und Bestätigung
4. Auswahl eines temporären Installationsverzeichnis (**<Installations-Pfad>**)
5. Benutzerdaten zu iRODS (u.a. **IRODS_HOME**)
6. Benutzerdaten zu Fedora (u.a. **FEDORA_HOME**)
7. Sonstige Benutzerdaten (u.a. **Tomcat_HOME**, **SOLR_HOME**)
8. Fortschrittsanzeige zur Installation
9. Informationen zum weiteren Vorgehen
10. Abschlussseite

Nach Abschluss des Installationsprogramms, ist im Installationsverzeichnis (gemäß Angabe in Schritt 4) folgende Struktur vorhanden, wobei diese README auch im Unterverzeichnis 'doc/' zu finden ist:

```
<Installations-Pfad>/
  Lizenz.txt
  doc/
    WissGrid_D3.5.3-grid-repository-fedora-irods-README.pdf (diese)
    WissGrid_D3.5.3-grid-repository-fedora-irods-Dokumentation.pdf
  licenses/
    Lizenztexte
  res/
    irods2.5.tgz
    irods3.0.tgz
    wissgridservices.war
    fcrepo-installer_WissGrid-AP3-Deliverable3.5v0.5.0.jar
  irodsChangeset/
    do.sh
  ERA/
    info.txt
    gsoap_2.8.0.zip
    gsoap.sh
    irods_wissgridservices_MS.c
    irods_wissgridservices_MS.h
```

```

Makefile
microservices.header
microservices.table
subsequentFedoraIngest.r
testsetup.sh
tomcat-users.xml
webservices/
  info.txt
wissgrid.irb
wissgrid.re
scripts/
  cleanUpIrods.sh
  config.sh
  fedora.sh
  irods.sh
  irodsChanges.sh
  solr.sh
  wissgridservices.sh
solr/
  ...
solr_sources/
  ...
wissgridservices.war

```

Im Verzeichnis `res/` liegen die Installationsprogramme für iRODS und Fedora, sowie die WAR-Datei der Service-Komponente.

Die Anpassungsdateien für iRODS, die gSOAP-Distribution und Shell-Skripte zum Steuern der gSOAP-Installation (`gsoap.sh`) und generieren von Code für Service-Aufrufe (`do.sh`) liegen im Verzeichnis `res/irodsChangeset/`.

4.2 Installationskomponenten

Die folgenden Schritte sind auszuführen, nachdem das WissGrid Installationsprogramm erfolgreich abgeschlossen wurde. Die einzelnen Schritte werden von Skripten gesteuert, wobei jeweils Hinweise zum aktuellen und den anschließenden Schritten gegeben werden. Begonnen wird mit dem Skript `<Installations-Pfad>/res/scripts/setEnv`.

1. Umgebungsvariablen setzen

Wenn die Umgebungsvariablen noch nicht gesetzt wurden (vgl. 4), dann kann man das mit dem Skript `<Installations-Pfad>/res/scripts/setEnv` nachholen. Das Skript setzt die Umgebungsvariablen (`IRODS_HOME`, `FEDORA_HOME`, `CATALINA_HOME`, `SOLR_HOME`), wie sie im WissGrid-Installationsprogramm angegeben wurden.

Aufruf: `source setEnv`

2. iRODS-Installation

Hinweis: Bei dieser Auslieferung handelt es sich um eine prototypische Entwicklung, deren Eignung für den Produktivbetrieb nicht gewährleistet werden kann.

Wenn man bereits einen iRODS-Server installiert hat, können die WissGrid-betreffenden Änderungen auch daran vorgenommen werden (siehe 4. iRODS Veränderungen). In diesem Fall kann dieser Schritt (2. iRODS-Installation) vollständig übersprungen werden.

- (a) Shell-Skript `<Installations-Pfad>/res/scripts/irods.sh` aufrufen und Ausgabe bzw. Hinweise beachten (siehe 6.1).

Aufruf: `./irods.sh`

- (b) optional: Wenn man mehrere iRODS-Accounts oder -Server verwenden will, kann man dafür verschiedene `irodsEnv`-Dateien² anlegen, z.B. `.irodsEnv_user1` für Benutzer `user1`. Ein Wechsel zwischen den Accounts ist dann durch Setzen der Variable `irodsEnvFile=~/.irods/.irodsEnv_user1` möglich. Anschließend ist ein Aufruf von `iinit` erforderlich, um das für die `iCommands` hinterlegte Passwort anzupassen. Die in dem besagten Verzeichnis liegende Standard `Env`-Datei kann als Muster verwendet werden.

Am Ende dieses Installationsschritts wird man aufgefordert, folgende Dateien zu Editieren:

- (a) In `$(IRODS_HOME)/server/config/irodsHost` muss das Mapping von `localhost` auf die zugehörige `<aktuelle-IP-Adresse>` festgelegt werden. Das geschieht, indem am Ende folgende Zeile hinzugefügt wird, wobei `<local-ip-address>` durch die aktuelle IP-Adresse zu ersetzen ist):

```
localhost <local-ip-address>
```

- (b) In der Datei `$(IRODS_HOME)/server/config/server.config` muss dem Attribut `reRuleSet` `wissgrid` hinzugefügt werden, so dass in der betreffenden Zeile mindestens folgendes steht:

```
reRuleSet wissgrid,core.
```

- (c) In der Datei `$(IRODS_HOME)/config/config.mk` muss dem Attribut `MODULE` die Eigenschaft `webservices` und `ERA` hinzugefügt werden, so dass in der betreffenden Zeile folgendes steht:

```
MODULES= webservices properties ERA
```

3. Installation der Web-Service-Komponente

Shell-Skript `<Installations-Pfad>/res/scripts/wissgridservice.sh` aufrufen und Hinweise beachten.

Aufruf: `./wissgridservice.sh`

4. iRODS Veränderungen für WissGrid vornehmen

Shell-Skript `<Installations-Pfad>/res/scripts/irodsChanges.sh` aufrufen und Hinweise beachten.

Aufruf: `./irodsChanges.sh`

²Enthält Verbindungsdaten, Benutzername und Default-Einstellungen, wie z.B. dem aktuellen Arbeitsverzeichnis auf iRODS.

5. Fedora-Installation

Shell-Skript `<Installations-Pfad>/res/scripts/fedora.sh` aufrufen und Hinweise beachten (zur Fedora-Installation siehe 6.2).

Aufruf: `./fedora.sh`

6. Solr-Installation

Für die Solr-Installation werden zwei alternative Wege beschrieben. 6a beschreibt eine schnelle und ebenso einfache Installation, welche durch ein Shell-Skript unterstützt wird und im wesentlichen eine Solr-Instanz anlegt und in Tomcat registriert. Bei der zweiten Alternative, welche 6b beschreibt, wird eine Solr-Instanz aus den Solr-Sourcen erstellt, nachdem diese auf die WissGrid-Bedürfnisse angepasst wurden. Dieser Weg wäre für eine bestehende Solr-Instanz zu beschreiten - jedoch sei auch an dieser Stelle nochmals auf den prototypischen Charakter der Entwicklung hingewiesen.

(a) Solr-Demo-Instanz anlegen

Shell-Skript `<Installations-Pfad>/res/scripts/solr.sh` aufrufen und Hinweise beachten.

Aufruf: `./solr.sh`

(b) Solr-Sourcen anpassen, Solr bauen und in Tomcat registrieren

i. Laden der Solr-Distribution und anschliessend entpacken. Der Entpackungs-ort wird im folgenden mit `<SOLR_DIR>` bezeichnet.

Download (in Version 1.4.1): <http://lucene.apache.org/solr/>

ii. `<Installations-Pfad>/res/solr/ContentStreamBase.java` nach `<SOLR_DIR>/src/common/org/apache/solr/common/util/` kopieren.

iii. `<Install.-Pfad>/res/solr/jargon-core-2.4.1-SNAPSHOT.jar` nach `<SOLR_DIR>/lib/` kopieren.

iv. Solr bauen, d.h. in `<SOLR_DIR>/` folgende Aufrufe absetzen

```
ant clean
ant dist
```

v. `SOLR_HOME` (Kontext-Pfad) anlegen und anpassen

Dieses Verzeichnis (z.B. `/opt/solr/testinstance/solr`) bildet das Kontext-Verzeichnis für die Solr-Webanwendung. Hier werden Konfigurationsdateien abgelegt und die eigentlichen Index-Daten gespeichert. Das Verzeichnis `<SOLR_HOME>` sollte nicht mit dem Installationspfad `<SOLR_DIR>` verwechselt werden.

- `<SOLR_DIR>/example/solr/` nach `/opt/solr/testinstance/` kopieren.
- `<SOLR_DIR>/dist/apache-solr-1.4.1.war` nach `<SOLR_HOME>` kopieren.

vi. `<SOLR_HOME>/conf/solrconfig.xml` anpassen

- `<dataDir>`-Element auf `<SOLR_HOME>/data` setzen, z.B.:

```

<dataDir>
  ${solr.data.dir:/opt/solr/testinstance/solr/data}
</dataDir>

```

Hinweis:

Wenn Solr nicht startet (Test mit Aufruf <http://localhost:8080/solr> und in der Tomcat Log-Datei (<tomcat>/logs/catalina.out) darauf hingewiesen wird, dass <SOLR_HOME> nicht gefunden werden kann, dann sollte die Erweiterung der JAVA_OPTS Umgebungsvariablen helfen:

```

export JAVA_OPTS="$JAVA_OPTS
-Dsolr.solr.home=/opt/solr/testinstance/solr"

```

In ~/.profile aufnehmen und Änderung mit source ~/.profile aktivieren.

- Element <requestHandler name="/update/extract"...> editieren und folgendes ändern bzw. ersetzen:

```

<lst name="defaults">
  <str name="fmap.content">attr_content</str>
  <str name="lowernames">true</str>
  <str name="uprefix">attr_</str>
  ...

```

Hinweis:

- fmap.content bildet Content auf ein dynamisches Feld attr_content ab.
- uprefix hängt Elemente den Zusatz attr_ vornan, wenn sie nicht direkt einem Feld zugeordnet sind. Damit werden sie vom dynamischen Feld attr_* behandelt.

vii. <SOLR_HOME>/conf/schema.xml anpassen

- Folgende Elemente bei den Feld-Definitionen hinzufügen:

```

<!-- wissgrid-fields ->
<field name="pid" type="text" indexed="true"
  stored="true" required="false"/>
<field name="dsid" type="text" indexed="true"
  stored="true" required="false" />
<field name="timestamp" type="text" indexed="true"
  stored="true" required="false" />
<field name="dslocation" type="text" indexed="true"
  stored="true" required="false" />

```

- Das dynamische Feld attr_* bestimmt, wie mit nicht explizit definierten Feldern umgegangen werden soll, z.B. mit Content (siehe 6(b)vi).

```

<dynamicField name="attr_*" type="textgen"
  indexed="true" stored="false"

```



```
multiValued='true'/'>
```

Hier wird der Content bspw. indexiert, nicht aber gespeichert. Der Content ist damit nicht über den Index verfügbar.

viii. Tomcat Context Fragment erstellen

Dazu wird eine Datei `<SOLR_HOME>/temp/solr.xml` erstellt:

```
<?xml version='1.0' encoding='utf-8'?>
<Context debug='0' crossContext='true'
  docBase=
    '/opt/solr/testinstance/solr/apache-solr-1.4.1.war'>
  <Environment name='solr/home' type='java.lang.String'
    value='/opt/solr/testinstance/solr' override='true'/'>
</Context>
```

In 6(b)x wird Solr über diese Datei auf Tomcat installiert.

ix. Bibliotheken kopieren:

- `<SOLR_HOME>/lib/` anlegen
- Kopieren aller Dateien aus `<SOLR_DIR>/contrib/extraction/lib/` nach `<SOLR_HOME>/lib/`
- Kopieren der Datei `<SOLR_DIR>/dist/apache-solr-cell-<vers.>.jar` nach `<SOLR_HOME>/lib/`

x. Installieren von Solr in Tomcat

`<SOLR_HOME>/temp/solr.xml` nach `<tomcat>/conf/Catalina/localhost/` kopieren (siehe 6(b)viii). Tomcat nimmt diese Datei automatisch auf und installiert darüber Solr. Bei erfolgreicher Installation wird die Datei automatisch gelöscht.

xi. Die Installation kann nun mit <http://localhost:8080/solr> getestet werden.

7. Davis-Installation

(a) Gemäß:

- <https://projects.arcs.org.au/trac/davis/wiki/HowTo/Install>
- <https://projects.arcs.org.au/trac/davis/wiki/HowTo/Configuration>

Die Installation des WissGrid-Repository ist hier abgeschlossen. Nachfolgend wird die Installation von Fedora-Plug-Ins und zusätzlichen Diensten beschrieben, welche den Funktionsumfang des Repository erweitern und deren Installation daher empfohlen wird.

4.3 Zusätzliche Komponenten

1. ORE Provider

Diese Komponente kann unter folgender Adresse heruntergeladen werden:

<http://sourceforge.net/projects/oreprovider/files/oreprovider/oreprovider%200.4%20Beta%202/oreprovider-0.4b2.war/download>

Die offizielle Online Dokumentation findet man unter:

<http://oreprovider.sourceforge.net/examples.html>

Es handelt sich um eine WAR-Datei, die zur Installation in das `webapp`-Verzeichnis des Tomcat-Servers gelegt werden muss. Die Konfiguration erfolgt über die Property-Datei `$CATALINA_HOME/webapps/oreprovider/WEB-INF/classes/ore.properties`, in welcher folgende Einstellungen vorgenommen werden müssen:

- `fedora.server = http://localhost:8080/fedora`
- `rem.autocreation = true`

2. OAI Provider

Die Komponente kann unter folgender Adresse Herunterladen werden:

<http://downloads.sourceforge.net/fedora-commons/oaiprovider-1.2.2.zip>

Die offizielle Online Dokumentation findet man unter:

<https://wiki.duraspace.org/display/FCSVCS/OAI+Provider+Service+1.2.2>

Vor der Installation der Komponente muss eine Datenbank angelegt und ein Benutzer mit vollen Zugriffsrechten darauf erstellt werden. Hier beispielhaft für PostgreSQL beschrieben:

- Erstellung der Datenbank *oaiPMH*:
`/path/to/postgres/pgsql/bin/createdb -h localhost oaiPMH`
- Definition eines Benutzers *oaiAdmin*, Kennwort *oaiAdmin*: (kein *Superuser*!)
`/path/to/postgres/pgsql/bin/createuser -h localhost oaiAdmin`
- In `/path/to/postgres/pgsql/data/pg_hba.conf` kann nun das Recht zum Zugriff auf die Datenbank gesetzt werden, indem folgende Zeile hinzugefügt wird:
`host oaiPMH "oaiAdmin" "IP-ADRESSE" "255.255.255.255" md5.`
- Danach erstellt man eine Verbindung zum *PostgreSQL*-Server und setzt das Kennwort für den Benutzer *oaiAdmin*:
`/path/to/postgres/pgsql/psql -h localhost -d template1 -E
template1=# ALTER USER "oaiAdmin" PASSWORD 'oaiAdmin'`
- Um zu prüfen, ob das Kennwort gesetzt wurde, dient das folgende `SELECT`, wobei das Kennwort MD5-kodiert angezeigt werden sollte:
`SELECT username,password FROM pg_shadow;`

Nach der Intergration der WAR-Datei in Apache Tomcat müssen folgende Werte in `$CATALINA_HOME/webapps/oaiprovider/WEB-INF/classes/proai.properties` geändert werden (hier mit beispielhaften Werten):

- `proai.cacheBaseDir = /path/to/tomcat/webapps/oaiprovider/cache`
- `proai.sessionBaseDir = /path/to/tomcat/oaiprovider/sessions`

- `proai.schemaDir = /path/to/tomcat/webapps/oaiprovider/schemas`
- `proai.db.url = jdbc:postgresql://localhost/oaiPMH`
- `proai.driverClassName = org.postgresql.Driver`
- `proai.db.username = oaiAdmin`
- `proai.db.password = oaiAdmin`
- `org.postgresql.Driver.ddlConverter = proai.util.
PostgresDDLConverter`
- `org.postgresql.Driver.backslashIsEscape = true`
- `driver.fedora.identify = http://localhost:8080/oaiprovider
/Identify.xml`
- `driver.fedora.md.format = oai_doc`
- `driver.fedora.md.format.oai_dc.dissType = info:fedora/*/oai_dc`
- `driver.fedora.md.format.oai_dc.about.dissType = info:fedora/*/oai_dc`
- `driver.fedora.md.format.oai_dc.about.dissType = info:fedora/*/DC`

Es ist zu erkennen, dass PostgreSQL als Datenbank gewählt wurde. Bei Verwendung eines anderen Datenbankserver sind die Einstellungen entsprechend zu ändern.

Nun muss die Datei `Identify.xml` in `$CATALINA_HOME/webapps/oaiprovider` erstellt werden. Ein Beispiel dazu ist im Abschnitt 6.3 angegeben.

3. Fedora GSearch

GSearch kann unter folgender Adresse geladen werden:

<http://sourceforge.net/projects/fedora-commons/files/services/3.1/genericsearch-2.2.zip/download>

Die zugehörige Dokumentation ist zu finden unter:

<https://wiki.duraspace.org/display/FCSVCS/Generic+Search+Service+2.2>

Auch bei GSearch sind Anpassungen bzw. Einstellung vorzunehmen. Nachfolgende Auflistung zeigt einen Teil der Konfigurationsmöglichkeiten, eine ausführliche Beschreibung folgt im Abschnitt 6.4:

- (a) Bezüglich des Apache Tomcat Servers, unterhalb des Verzeichnisses `$CATALINA_HOME/webapps/fedoragsearch/WEB-INF/classes/`:
 - `log4j.xml`
 - Alle `demo...`-Dateien in `config/rest/` wurden umbenannt. Die Namen werden später in der Property-Datei korrigiert. Auch in diesen Dateien soll die Zeile `CONFIGPATH/rest/demoCommon.xslt` mit dem richtigen Pfad und Dateinamen ersetzt werden.
 - `config/fedoragsearch.properties`
 - `config/index/INDEXNAME/index.properties`
 - `config/index/INDEXNAME/indexInfo.xml`
 - `config/repositroy/REPOSITROYNAME/repositroy.properties`
 - `config/repositroy/REPOSITROYNAME/repositroyInfo.xml`

- `config/updater/UPDATERNAME/update.properties`
- (b) Beim Apache Solr Server unter `$SOLR_HOME`³:
- Ordner `data/gsearchindex`
 - `conf/solrConfig.xml`
 - `conf/schema.xml`
- (c) Die Konfigurationsdatei `fedora.fcfig` (Fedora Common Konfigurationsdatei) unter `$FEDORA_HOME/server/config` muss überarbeitet werden.
- (d) Es wird empfohlen, alle Solr- und Lucene-Bibliotheken durch aktuelleren zu ersetzen (durch Versionen ab 2.9.x bzw. 1.4.x). Zudem müssen dann auch die Versionsbezeichner in folgenden Dateien angepasst werden:
- `library.properties` unter `$CATALINE_HOME/webapps/fedoragsearch/WEB-INF/lib`
 - `runRESTClient.sh` und `runSOAPClient.sh` unter `$CATALINA_HOME/webapps/fedoragsearch/client`

4. Pazpar2

Es wird bei dieser Installation empfohlen, den Anweisung der Linux Distribution zu folgen. Die Anwendung kann unter folgender Adresse herunterladen:

<http://ftp.indexdata.dk/pub/pazpar2>

Die offizielle Online Dokumentation findet man unter:

<http://www.indexdata.com/pazpar2/doc/>

<http://www.indexdata.com/pazpar2/doc/pazpar2.pdf>

Vor der Installation sind aktuelle Versionen der Bibliotheken `libyaz4` und `libyazpp` zu laden und installieren.

Zur Konfiguration müssen folgende Dateien angepasst und/oder bestimmte Konfigurationsdateien aktivieren⁴ werden. Ausführliche Beispiele und Anweisungen sind im Abschnitt 6.5 zu finden:

- `/etc/pazpar2/server.xml`
- `/etc/pazpar2/services-available/default.xml`
- `/etc/pazpar2/services-available/solr.xml`
- Erstellen Symbolischer Links von beiden letzten Dateien in `/etc/pazpar2/services-enabled`

5. Apache HTTP Server²

Zum Beispiel für Debian als `'root'`:

```
$ apt-get install apache2 apache2-common
```

Das Ziel ist, eine einheitliche URL-Adressierung für alle installierten Komponenten zu erreichen. Aus diesem Grund sollen einige "Module" nach der Installation vom Server aktiviert werden, wie `mod_proxy`, `mod_proxy_ajp` und `mod_rewrite`.² Ein Konfigurationsbeispiel für Debian ist im Abschnitt 6.6 zu finden.

³Hier gehen wir davon aus, dass Sie nur einen Instanz von **Apache Solr Server** installiert haben. Eine Alternative Einleitung wird noch im Abschnitt 6.4 beschrieben.

⁴Gemäß der Anweisung von Linux Distribution

5 Mögliche Probleme bei der Installation

1. **Postgres und unixODBC:** Das iRODS Installations-Skript lädt bei entsprechender Auswahl den Postgres-Server und unixODBC Bibliothek nach. Bei Test Installationen war die hinterlegte Adresse wiederholt nicht erreichbar, so dass die Installation abbrach. Um das Problem zu lösen, können die Dateien `postgresql-9.0.3.tar.gz` und `unixODBC-2.2.12.tar.gz` manuell geladen werden.

Damit die geladenen Sourcen gefunden werden können, muss von der Muster-Datei `$IRODS_HOME/config/installPostgres.config.template` eine Kopie angelegt werden (`installPostgres.config` im selben Verzeichnis), und es sind folgende Angaben, entsprechend den lokaler Gegebenheiten, anzupassen.

```
# (Ziel) Pfad für Postgres-Sourcen
$POSTGRES_SRC_DIR = '/opt/postgres';

# Postgres (Ziel) Installationsverzeichnis
$POSTGRES_HOME = '/opt/postgres/pgsql';

# Wo liegen die Dateien (die oben geladenen)
$DOWNLOAD_DIR = '/home/userXYZ/Downloads/postgres';

$POSTGRES_SOURCE = 'postgresql-9.0.3.tar'; # bzw. .tar.gz
$ODBC_SOURCE = 'unixODBC-2.2.12.tar'; # bzw. .tar.gz
```

2. **Berechtigungen:** Sollten beim Start der Server oder im Betrieb Schwierigkeiten auftreten, so könnte dies durch fehlende Schreib-Rechte des Benutzers verursacht werden. Z.B. schreibt iRODS, wenn nicht anders konfiguriert, in ein Verzeichnis `'$IRODS_HOME/Vault/'` (Standard iRODS-Ressourcen). Wenn die Rechte zum Schreiben in diesem Verzeichnis fehlen, verursacht das einen Fehler.
3. Ein gescheiterter Ingest oder die Fehlfunktion des Tomcat-Servers im Rahmen eines Ingests, mit der Fehlermeldung `OutOfMemory Error: PermGen Space` in der Tomcat Log Datei `catalina.out`, kann seine Ursache darin haben, dass Tomcat nicht mit ausreichend Heap-Speicher gestartet wurde.

Die Zuweisung des Heaps erfolgt beim Tomcat-Start durch Setzen der Umgebungsvariable `JAVA_OPTS`. Nachfolgend ist eine beispielhafte Konfiguration angegeben, wobei zu prüfen ist, ob diese für die eigenen Anwendungsfälle ausreichend ist (`xms` setzt die Startgröße und `xmx` die maximale Größe des Heap).

```
JAVA_OPTS="-server -Xms768M -Xmx1024M -XX:MaxPermSize=768M"
```

6 Eingabehinweise

6.1 iRODS - Installationsangaben

Im wesentlichen sollten die Standardwerte gewählt werden. Wenn die erste Aufforderung mit “no“ beantwortet wird, müssen lediglich der Zielpfad für die neue Postgres-Instanz und ein Postgres Passwort für den Benutzer angegeben werden. Die folgende Ausgabe ist verkürzt, der Installer liefert weitere Informationen.

```
Include additional prompts for advanced settings [no]?
Build an iRODS server [yes]?
Make this Server ICAT-Enabled [yes]?
iRODS zone name [tempZone]?
iRODS login name [rods]?
Password [rods]?
Download and build a new Postgres DBMS [yes]?
New Postgres directory? /path/to/postgres           #Eingabe
New database login name [user]?
Password? aPassword                                # Eingabe
PostgreSQL version [postgresql-9.0.3.tar.gz]?
ODBC version [unixODBC-2.2.12.tar.gz]?
Port [5432]?
Include GSI [no]? yes                               # Eingabe, wenn
                                                    # GSI erwünscht
GLOBUS_LOCATION? /opt/globus                       # Eingabe, gemäß den
                                                    # Gegebenheiten
GSI Install Type to use? gcc64dbg                 # Eingabe
Save configuration (irods.config) [yes]?             # abgelegt in:
                                                    # $IRODS_HOME/config/

Start iRODS build [yes]?
...
```

6.2 Fedora Installationsangaben

Auch hier werden im wesentlichen Standardwerte verwendet und der Installer liefert ebenfalls weitergehende Informationen.

```
Installation type
Enter a value ==> custom # Eingabe

Fedora home directory
Enter a value [default is $FEDORA_HOME] ==>

Fedora administrator password
Enter a value ==> fedoraAdmin # Eingabe

Fedora server host
Enter a value [default is localhost] ==> # Auf default belassen,
# sonst stimmt die iRODS
# Konfiguration nicht mehr

Fedora application server context
Enter a value [default is fedora] ==> # Auf default belassen,
# sonst stimmt die iRODS
# Konfiguration nicht mehr

Authentication requirement for API-A
Enter a value [default is false] ==>

SSL availability
Enter a value [default is true] ==> false # Eingabe, SSL wird
# für erste Tests nicht
# benötigt

Servlet engine
Enter a value [default is included] ==> existingTomcat # Eingabe. Existiert
# eine Tomcat Instanz?

Tomcat home directory
Enter a value ==> /path/to/tomcat/ # Eingabe. Wird mit
# $CATALINA_HOME
# automatisch gesetzt

Tomcat HTTP port
Enter a value [default is 8080] ==>

Tomcat shutdown port
Enter a value [default is 8005] ==>

Database
Enter a value ==> included # Eingabe. Das Einbinden
# einer existierenden DB
# verkürzt die Installation

Enable FeSL AuthN
Enter a value [default is true] ==>
```

```

Enable FeSL AuthZ (Experimental Feature)
Enter a value [default is false] ==> # Alternative ist AuthZ
                                        # über XACML

Policy enforcement enabled
Enter a value [default is true] ==> # Um Rechte festlegen zu
                                        # können. Definitionen
                                        # liegen in:
                                        # '$FEDORA_HOME/'
                                        # data/../../default/'

Low Level Storage
Enter a value [default is akubra-fs] ==> # Es wird trotzdem das
                                        # iRODS-Akubra Modul
                                        # installiert

Enable Resource Index
Enter a value [default is false] ==> # Für erste Tests nicht benötigt,
                                        # verwaltet RDFs in Triplestore

Enable Messaging
Enter a value [default is false] ==> # Für erste Tests nicht benötigt,
                                        # externe Systeme lassen sich
                                        # hierüber integrieren.

Deploy local services and demos
Enter a value [default is true] ==> false # Eingabe

```


6.3 OAI Provider - Beispiel für Identify.xml

```
<Identify>
  <repositoryName>WissGrid Aggregator</repositoryName>
  <baseURL/>
  <protocolVersion>2.0</protocolVersion>
  <adminEmail>mailto:samadi@sub.uni-goettingen.de</adminEmail>
  <earliestDatestamp>2011-01-01T00:00:00Z</earliestDatestamp>
  <deletedRecord>no</deletedRecord>
  <granularity>YYYY-MM-DDThh:mm:ssZ</granularity>
  <description>

    <oai-identifier xsi:schemaLocation="http://www.openarchives.org/OAI/2.0
      /oai-identifier http://www.openarchives.org/OAI/2.0/oai-identifier.xsd">

      <scheme>oai</scheme>
      <repositoryIdentifier>wissgrid.org</repositoryIdentifier>
      <delimiter>:</delimiter>
      <sampleIdentifier>oai:wissgrid:2</sampleIdentifier>

    </oai-identifier>
  </description>
</Identify>
```

6.4 Fedora GSearch - Ausführliche Einstellungen

- Beispiel für log4j.xml:

```
<appender name="FILEOUT" class="org.apache.log4j.FileAppender">
  <param name="File"
    value="/path/to/fedora/server/logs/gsearch.log"/>
  ...
<logger name="dk.defxws.fgslucene" additivity="true">
  <level value="DEBUGLEVELL" />
  <appender-ref ref="FILEOUT"/>
</logger>

<logger name="dk.defxws.fgssolr" additivity="true">
  <level value="DEBUGLEVELS" />
  <appender-ref ref="FILEOUT"/>
</logger>

<root>
  <level value="DEBUGLEVELS" />
  <appender-ref ref="FILEOUT"/>
</root>
```

- In `$CATALINA_HOME/webapps/fedoragsearch/WEB-INF/classes/` sieht die Ordnerstruktur wie folgt aus:

```
config/
|-index/SolrIndex/conf
|-repositroy/WissGridRepo
|-rest/css
|-updater/BasicUpdaters
```

- Beispiel für `config/fedoragsearch.properties`:

```
fedoragsearch.soapBase = http://localhost:8080/gsearch/services
fedoragsearch.soapUser = fedoraAdmin
fedoragsearch.soapPass = fedoraAdmin
fedoragsearch.deployFile = /path/to/tomcat/webapps/fedoragsearch/
                           WEB-INF/classes/config/deploy.wsdd

#rest
fedoragsearch.defaultNoXslt = copyXml
fedoragsearch.defaultUpdateIndexRestXslt = UpdateIndexToHtml
fedoragsearch.defaultGfindObjectsRestXslt = GfindObjectsToHtml
fedoragsearch.defaultBrowseIndexRestXslt = BrowseIndexToHtml
fedoragsearch.defaultGetRepositoryInfoRestXslt = GetRepositoryInfoToHtml
fedoragsearch.defaultGetIndexInfoRestXslt = GetIndexInfoToHtml
```

```

#resultPage
fedoragsearch.maxPageSize = 100
fedoragsearch.defaultBrowseIndexTermPageSize = 25
fedoragsearch.defaultGfindObjectsHitPageSize = 10
fedoragsearch.defaultGfindObjectsSnippetsMax = 3
fedoragsearch.defaultGfindObjectsFieldMaxLength = 150
fedoragsearch.mimeTypes = text/plain text/html application/pdf
fedoragsearch.repositoryNames = WissGridRepo
fedoragsearch.indexNames = SolrIndex
fedoragsearch.updaterNames = BasicUpdaters

```

- Beispiel für config/index/SolrIndex/index.properties:

```

fgsindex.indexName = SolrIndex
fgsindex.operationsImpl = dk.defxws.fgssolr.OperationsImpl

fgsindex.defaultUpdateIndexDocXslt = FoxmlToSolr
fgsindex.defaultUpdateIndexResultXslt = updateIndexToResultPage
fgsindex.defaultGfindObjectsResultXslt = gfindObjectsToResultPage
fgsindex.defaultBrowseIndexResultXslt = browseIndexToResultPage
fgsindex.defaultGetIndexInfoResultXslt = copyXml

fgsindex.indexBase = http://localhost:8080/solr/GSearch
fgsindex.indexDir = /path/to/solr/cores/gsearch/data/index
fgsindex.analyzer = org.apache.lucene.analysis.standard.StandardAnalyzer

fgsindex.defaultQueryFields = dc.description dc.title dc.creator dc.identifier
dc.subject dc.date dc.type dc.coverage dc.publisher
dc.contributor dc.source pid PID dsid version
timestamp dslocation

fgsindex.defaultSortFields = PID,AUTO,true
fgsindex.untokenizedFields = fgs.contentModel fgs.ownerId PID
fgsindex.mergFactor = 10
fgsindex.maxBufferedDocs = 10

fgsindex.snippetBegin = <span class=\"highlight\">
fgsindex.snippetEnd = </span>

```

- Beispiel für config/index/SolrIndex/index.properties:

```

fgsrepository.repositoryName = WissGridRepo
fgsrepository.fedoraSoap = http://localhost:8080/fedora/services
fgsrepository.fedoraUser = fedoraAdmin
fgsrepository.fedoraPass = fedoraAdmin
fgsrepository.fedoraObjectDir = /path/to/iRODS/Vault/home/rods/objects
fgsrepository.fedoraVersion = 3.5

fgsrepository.defaultGetRepositoryInfoResultXslt = copyXml
fgsrepository.trustStorePath = /path/to/fedora/server/truststore
fgsrepository.trustStorePass = changeit

```

- Beispiel für config/rest/BrowseIndexToHtml:

```

<xsl:variable name="EQCHAR">
  <xsl:choose>

```

```

<xsl:when test="$INDEXNAME = 'ZebraIndex'"></xsl:when>
...
<xsl:text> </xsl:text>Index name:
  <select name="indexName">
    <xsl:choose>
      <xsl:when test="$INDEXNAME='SolrIndex'">
        <option value="SolrIndex" selected="true">SolrIndex
        </option>
      </xsl:when>
      <xsl:otherwise>
        <option value="SolrIndex" selected="true">SolrIndex
        </option>
      </xsl:otherwise>
    </xsl:choose>
  </select>
<xsl:text> </xsl:text>restXslt:
  <select name="restXslt">
    <option value="BrowseIndexToHtml">BrowseIndexToHtml
    </option>
    <option value="copyXml">no transformation</option>
  </select>
...

```

- Beispiel für config/updater/BasicUpdaters/updater.properties:

```

java.naming.factory.initial = org.apache.activemq.jndi.ActiveMQInitialContextFactory
java.naming.provider.url = tcp://localhost:61616
connection.factory.name = ConnectionFactory
topic.fedoraAPIM = fedora.apim.update
client.id = gsearch

```

- schema.xml muss angepasst werden, damit nach den neuen Indexkategorien indiziert werden kann:

```

<field name="pid" type="string" indexed="true" stored="true" required="true" />
<field name="dsid" type="string" indexed="true" stored="true" required="true" />
<field name="version" type="string" indexed="true" stored="true" required="true" />
<field name="dslocation" type="string" indexed="true" stored="true" required="true" />

```

- Wenn man bereits eine **Apache Solr Server** Instanz installiert hat, kann diese über die MultiCore-Funktion für den Einsatz mit **GSearch** umkonfigurieren. Dafür sind folgende Schritte vorzunehmen:

- Unter `$$SOLR_HOME/cores/<CORENAMEN>/conf` sind Konfigurationsdateien zu erstellen, z.B. `$$SOLR_HOME/cores/gsearch/conf`.
- Alle alten Konfigurationsdateien sind von `$$SOLR_HOME/conf` in den neuen Konfigurationsordner zu kopieren.
- Es ist noch eine Datei `server.xml`, unter `$$SOLR_HOME` zu erstellen, z.B.:

```
<?xml version="1.0" encoding="UTF-8" ?>
```

```
<solr present="true">
  <cores adminPath="amdin/cores">
    <core name="GSeach" intanceDir="cores/gsearch">
    <core name="iRODS" intanceDir="cores/irods">
  </cores>
</solr>
```

- Letztendlich müssen nochfolgende Dateien überarbeitet werden:
 - \$CATALINA_HOME/webapps/wissgridservices/WEB-INF/wissgrid.properties.
Die Adresse der **SOLR_URL** anpassen.
 - \$CATALINA_HOME/webapps/fedoragsearch/WEB-INF/classes/config/index/SolrIndex/index.properties. Hier müssen die Adresswerte (URLs) von **fgsindex.indexBase** und **fgsindex.indexDir** angepasst werden.

6.5 Pazpar2 - Einstellungen mit Solr für Debian

- Beispiel für `/etc/pazpar2/server.xml`:

```
<?xml version="1.0" encoding="UTF-8"?>
<pazpar2 xmlns="http://www.indexdata.com/pazpar2/1.0">
  <threads number="50"/>
  <server>
    <listen port="8004"/>
    <include src="services-enabled/*.xml"/>
  </server>
</pazpar2>
```

- Beispiel für `/etc/pazpar2/services-enabled/default.xml`:

```
<timeout session="280" z3950_operation="240" z3950_session="360"/>
...
<icu_chain id="relevance" locale="de">
...
<icu_chain id="sort" locale="de">
...
<icu_chain id="mergekey" locale="de">
...
<icu_chain id="facet" locale="de">
...
...

```

- Beispiel für `/etc/pazpar2/services-enabled/solr.xml`:

```
...
  <timeout session="210" z3950_operation="180" z3950_session="240"/>
  ...

```

```

<set target="IP-ADRESSE:8080/solr/GSearch/" name="pz:name"
  value="solrGSearch"/>
<set target="IP-ADRESSE:8080/solr/iRODS/" name="pz:name"
  value="solriRODS"/>
...
<set name="pz:sru" value="solr"/>
<set name="pz:sru_version" value="1.4.1"/>
<set name="pz:maxrecs" value="500"/>
...
<set name="pz:cclmap:term" value="s=al t=r"/>
<set name="pz:cclmap:fulltext" value="s=al t=r"/>
<set name="pz:elements" value="*/>
<set name="pz:nativesyntax" value="xml;utf-8"/>
<set name="pz:xslt" value="solr-pz2.xsl"/>
...
<!-- Wissgrid special fields -->
<set name="pid" value="1=pid s=al t=r"/>
<set name="dsid" value="1=dsid s=al t=r"/>
<set name="version" value="1=version s=al t=r"/>
<set name="timestamp" value="1=timestamp s=al t=r"/>
<set name="dslocation" value="1=dslocation s=al t=r"/>
...
<icu_chain id="relevance" locale="de">
...
<icu_chain id="sort" locale="de">
...
<icu_chain id="mergekey" locale="de">
...
<!-- wissgrid specified metadata: start-->
<metadata name="pid" brief="yes" termlist="yes"/>
<metadata name="dsid"/>
<metadata name="version"/>
<metadata name="timestamp"/>
<metadata name="dslocation"/>
...

```

6.6 Apache HTTP Server - Debian Proxy Einstellungen

- Beispiel für die Konfigurationsdatei des vorinstallierten default Virtual Host:

ProxyPass /tomcat/ **http://localhost:8080/**

ProxyPassReverse /tomcat/ **http://localhost:8080/**

ProxyPass /manager/html **http://localhost:8080/manager/html**

ProxyPassReverse /manager/html **http://localhost:8080/manager/html**

ProxyPass /manager/status **http://localhost:8080/manager/status**

ProxyPassReverse /manager/status **http://localhost:8080/manager/status**

ProxyPass /fedora **http://localhost:8080/fedora**

ProxyPassReverse /fedora **http://localhost:8080/fedora**

ProxyPass /fedora/describe **http://localhost:8080/fedora/describe**

ProxyPassReverse /fedora/describe **http://localhost:8080/fedora/describe**

ProxyPass /fedora/objects **http://localhost:8080/fedora/objects**

ProxyPassReverse /fedora/objects **http://localhost:8080/fedora/objects**

ProxyPass /fedora/services **http://localhost:8080/fedora/services**

ProxyPassReverse /fedora/services **http://localhost:8080/fedora/services**

ProxyPass /fedoragsearch **http://localhost:8080/fedoragsearch**

ProxyPassReverse /fedoragsearch **http://localhost:8080/fedoragsearch**

ProxyPass /fedoragsearch/rest **http://localhost:8080/fedoragsearch/rest**

ProxyPassReverse /fedoragsearch/rest **http://localhost:8080/fedoragsearch/rest**

ProxyPass /fedoragsearch/services **http://localhost:8080/fedoragsearch/services**

ProxyPassReverse /fedoragsearch/services **http://localhost:8080/fedoragsearch/services**

ProxyPass /solr **http://localhost:8080/solr**

ProxyPassReverse /solr **http://localhost:8080/solr**

ProxyPass /solr/GSearch **http://localhost:8080/solr/GSearch**

ProxyPassReverse /solr/GSearch **http://localhost:8080/solr/GSearch**

ProxyPass /solr/iRODS **http://localhost:8080/solr/iRODS**

ProxyPassReverse /solr/iRODS **http://localhost:8080/solr/iRODS**

ProxyPass /oaiprovider **http://localhost:8080/oaiprovider**

ProxyPassReverse /oaiprovider **http://localhost:8080/oaiprovider**

ProxyPass /oreprovider **http://localhost:8080/oreprovider**

ProxyPassReverse /oreprovider http://localhost:8080/oreprovider

ProxyPass /wissgridservices http://localhost:8080/wissgridservices

ProxyPassReverse /wissgridservices http://localhost:8080/wissgridservices

ProxyPass /wissgridservices/apim http://localhost:8080/wissgridservices/apim

ProxyPassReverse /wissgridservices/apim http://localhost:8080/wissgridservices/apim

...

- Für dieses Beispiel müssen noch folgende Konfigurationsdateien angepasst werden:
 - \$CATALINA_HOME/webapps/fedoragsearch/WEB-INF/classes/config/fedoragsearch.properties
 - \$CATALINA_HOME/webapps/fedoragsearch/WEB-INF/classes/config/index/SolrIndex/index.properties
 - \$CATALINA_HOME/webapps/fedoragsearch/WEB-INF/classes/config/repository/WissGridRepo/repository.properties
 - \$CATALINA_HOME/webapps/wissgridservices/WEB-INF/wissgridservices.properties
 - \$CATALINA_HOME/webapps/oreprovider/WEB-INF/classes/ore.properties
 - \$CATALINA_HOME/webapps/oaiprovider/WEB-INF/classes/proai.properties
 - /etc/pazpar2/services-available/solr.xml