CANANOLAB 1.4

Installation Guide

July 11, 2008

Table of Contents

Introduction to caNanoLab

Welcome to the cancer Nanotechnology Laboratory (caNanoLab) 1.4 Installation Guide. caNanoLab is a data sharing portal designed to allow researchers to share information on nanoparticles including the composition of the particle, the functions (e.g. therapeutic, targeting, diagnostic imaging) of the particle, the characterizations of the particle from physical (e.g. size, molecular weight) and in vitro (e.g. cytotoxicity, immunotoxicity) nanoparticle assays, and the protocols of these characterizations.

As of release 1.1, the caNanoLab domain model has been caGrid enabled. In other words, a caNanoLab grid data service can be deployed and registered with a caGrid index server, allowing sharing of public nanoparticle information across the caGrid. In release 1.4, the caNanoLab grid service has been updated to use caGrid 1.2. For more information, see https://cabig.nci.nih.gov/workspaces/Architecture/caGrid.

Targeted Developer

The caNanoLab application development is best suited for an experienced Java developer who has some familiarity with the following J2EE and related technologies:

Important Background Knowledge

- Unix/Linux environment, Windows XP environment or Mac OS environment (Configuring environment variables; Installing Ant, JDK, Apache Tomcat and JBoss servers)

- Ant build scripts

- J2EE web application development using the Struts framework, Servlet/JSP's, JavaScript

- J2EE middle-ware technologies such as n-tier service-oriented architecture and software design patterns

- caGrid 1.2 infrastructure (for understanding how a caNanoLab grid data service functions)

General System Requirements

The following open source technologies power a caNanoLab web application:

Open Source Technologies	• Java Software Development Kit (JDK) version 5.0 http://java.sun.com/javase/downloads/index_jdk5.jsp
	• JBoss version 4.0.5 http://labs.jboss.com/jbossas/downloads
	• Jakarta Ant version 1.6.5 http://archive.apache.org/dist/ant/binaries/
	• MySQL version 5.0.x http://dev.mysql.com/downloads/mysql/5.0.html

The caNanoLab web application has been tested within NCICB against JBoss servers (version 4.0.5) hosted on Windows XP and RedHat Linux systems, and against MySQL 5.0.27 databases hosted on RedHat Linux systems and MySQL 5.0.45 on Windows XP systems. Prior to release 1.2.1, the caNanoLab web application had been tested against Oracle 9i databases hosted on Sun Solaris systems, and Oracle 10g XE database hosted on Windows XP.

Download each of the tools listed in the bulleted list and follow the installation instructions provided with each respective product for your environment. Assistance from a MySQL database administrator is expected to properly configure the MySQL database. For MySQL database configuration and maintenance, it is also helpful to download the MySQL 5.0 suite of GUI tools at http://dev.mysql.com/downloads/gui-tools/5.0.html.

| **Grid Service** | Setting up a caNanoLab grid service is optional. It is recommended that you install the grid service only after you successfully install the web application and become familiar with submitting and searching data through the web application. |

| **Grid Service Downloads** | In release 1.4, there are no separate technology downloads required for installing the caNanoLab grid service, and most have been packed within the caNanoLab source distribution file. |

Obtaining the caNanoLab Source

| **Source Code** | The caNanoLab web application source code and the caNanoLab grid service deployment war file are distributed as a zip file named `caNanoLab_1.4.zip` from the NCICB GForge site at http://gforge.nci.nih.gov/frs/?group_id=69 under release 1.4 |

Database Technology

Assumptions and Requirements	The caNanoLab source distribution `caNanoLab_1.4.zip` has been downloaded from the NCICB GForge site. A MySQL database has been set up on a system (local or remote) with an administrator (or root) account. You should be familiar with the process of creating InnoDB tablespaces, databases creating users and assigning privileges to users.
For a Previous Installation	If you have installed caNanoLab release 1.3 against an MySQL database, or release 1.2 or release 1.1.1 or release 1.1 against an Oracle database, and have associated production data in these schemas and you would like to continue to use the same data for caNanoLab release 1.4 in MySQL, review the following database initialization steps, then go directly to *Data Migration* on page *4*.
For a New Installation	If you are installing the caNanoLab application for the first time or want to install a new schema for release 1.4, follow the steps below to set up the required MySQL schema objects and the seed data for release 1.4.

Initializing the MySQL Database

Follow these steps to initialize your MySQL database system:

Step	Action
1	Extract the `caNanoLab_1.4.zip` to a location on your local system, for example, `C:\caNanoLab_1.4`. This location is referred as `<CANANOLAB_SOURCE>` throughout the document. Verify that the following eight SQL scripts exist in the directory `<CANANOLAB_SOURCE>/db/1.4`:

- `mysql_seed.sql`

- `create_database_and_user.sql`

- `app_schema.sql`

- `insert_common_lookup.sql`

- `csm_schema.sql`

- `app_priming_data.sql`

- `csm_priming_data.sql`

- `app_csm_priming_data.sql`

Note: The scripts `create_database_and_user.sql` create a new 1.4 schema ***canano*** and a new database user ***cananolab_app*** with required privileges and a password ***go!234***. If you'd like to create a different database user, be sure to modify this script and make a note of the information for later use

Step	Action
	in setting up the caNanoLab application.
	User administration and schema privileges can also be managed by the MySQL Administrator GUI as a part of the MySQL 5.0 suite of GUI tools.

2	Log into the MySQL database as the database administrator (root) and go to the directory `<CANANOLAB_SOURCE>/db/1.4` and execute the script `mysql_seed.sql` to set up a schema named ***canano*** and its required seed data.
	
	Example: On a Windows XP system hosting a MySQL 5.0.45 database, you would issue the following commands at the DOS prompt (assuming the ***root*** account has password ***rootpass***):
	
	`C:\>cd C:\caNanoLab_1.4\db\1.4`
	
	`C:\caNanoLab_1.4\db\1.4\>mysql –h localhost –u root –prootpass <mysql_seed.sql`

Verification   Once the MySQL database has been created, either through new setup or through data migration (described in the next section), verify that the following numbers of database objects are created:

- Tables 63

Example: Issue the following query at the MySQL prompt, logging in as ***cananolab_app***:

```
mysql> select count(*) from information_schema.tables
where table_schema='canano' and table_type='BASE TABLE';
```

Data Migration

NOTE:	
	If you are installing caNanoLab 1.4 for the first time or installing a new caNanoLab release 1.4 schema, you can skip this section.

This installation guide only discusses the steps for migrating from release 1.3 in MySQL to release 1.4 in MySQL. If you have previously installed caNanoLab release 1.2.1 in MySQL, refer to the release 1.3 installation guide http://gforge.nci.nih.gov/docman/view.php/69/10380/caNanoLab_1.3_Installation_Guide.pdf for instructions migrating to release 1.3 first. If you have previously installed caNanoLab release

1.2, release 1.1.1, or release 1.1 against an Oracle database and have production data associated with these schemas that you wish to keep, you must perform a data migration to release 1.2.1 in MySQL first before migrating to release 1.3.   Please see the release 1.2.1 installation guide
http://gforge.nci.nih.gov/docman/index.php?group_id=69&selected_doc_group_id=2490&language_id=1 for details.

Follow these steps to complete the required data migration from release 1.3 in MySQL to release 1.4 in MySQL:

Step	Action
1	Extract the `caNanoLab_1.4.zip` to a location on your local system, for example, `C:\caNanoLab_1.4`. This location is referred as `<CANANOLAB_SOURCE>` throughout the document. Verify that the following SQL script exists in the directory `<CANANOLAB_SOURCE>/db/1.4`: • `pre_migration.sql` and the following SQL scripts exist in `<CANANOLAB_SOURCE>/db/1.4/data_migration`: • `migration_general.sql` • `assign_particles_to_source_groups.sql` • `assign_associated_data_to_public_group.sql` • `image_contrast_agent_ncl.sql` • `image_contrast_agent_stanford.sql` • `image_contrast_agent_washu.sql` **Note**: Scripts `image_contrast_agent_*.sql` are specialized scripts for migrating image contrast agents in NCL, WashU and Stanford, respectively. If you institute has image contrast agents in release 1.3 that need to be migrated into release 1.4, please notify the caNanoLab technical team. Because there are significant database schema differences between release 1.3 and release 1.4, specialized migration scripts may be necessary.

Step	Action

2	Log into the MySQL database as the database administrator (root) and go to the directory `<CANANOLAB_SOURCE>/db/1.4` and execute the script `pre_migration.sql.` Then go the directory `<CANANOLAB_SOURCE>/db/1.4/data_migration` and execute the following scripts in the order specified:
	
	    • `migration_general.sql`
	
	    • `assign_particles_to_source_groups.sql`
	
	    • `assign_associated_data_to_public_group.sql`
	
	    • `image_contrast_agent_X.sql`
	
	Example: On a Windows XP system hosting a MySQL 5.0.45 database, you would issue the following commands at the DOS prompt (assuming the ***root*** account has password ***rootpass***):
	
	`C:\>cd C:\caNanoLab_1.4\db\1.4`
	`C:\caNanoLab_1.4\db\1.4>mysql -h localhost -u root -prootpass <pre_migration.sql`
	
	`C:\caNanoLab_1.4\db\1.4>cd data_migration`
	`C:\caNanoLab_1.4\db\1.4\data_migration>mysql -h localhost -u root -prootpass <migration_general.sql`
	`C:\caNanoLab_1.4\db\1.4\data_migration>mysql -h localhost -u root -prootpass <assign_particles_to_source_groups.sql`
	`C:\caNanoLab_1.4\db\1.4\data_migration>mysql -h localhost -u root -prootpass <assign_associated_data_to_public_group.sql`
	`C:\caNanoLab_1.4\db\1.4\data_migration>mysql -h localhost -u root -prootpass <image_contrast_agent_ncl.sql`

After data migration, refer to the Verification section on page 4 to verify that the migration has been successful. The same number of schema objects should be generated as described in that section.

caNanoLab Web Application

Assumptions and Requirements

The database has been installed and verified, as described on pages 4 through 7. Ant has been installed. The JBoss application server has been installed on a system (local or remote) and can be started at a designated port. JBoss install directory is referred as `<JBOSS_HOME>` in the document. The JBoss application server host URL is referred as `<APP_SERVER_URL>`, for example, cananolab-dev.nci.nih.gov:19080. It is assumed that the default configuration is used for deploying caNanoLab web archive files. For example, in JBoss 4.0.5, the default configuration is located at the directory `<JBOSS_HOME>/server/default`. The caNanoLab web archive file shall be deployed at the directory `<JBOSS_HOME>server/default/deploy`.

It is possible to configure Apache server to interface with the JBoss server and set up a virtual host for the caNanoLab application, if you need assistance, please contact NCICB Application Support (info provided at the end).

Installation and Deployment

Follow these steps to install and deploy caNanoLab.

Step	Action
1	Set up an environment variable `JBOSS_HOME` to point to the JBoss installation directory. **Note:** Either double back slashes \\ or a single forward slash / should be used as the file separator if working on Windows platform, for example, `JBOSS_HOME=C:/jboss-4.0.5.GA` (`C:\jboss-4.0.5.GA` would fail).
2	Edit the Ant build properties file `build.properties` at `<CANANOLAB_SOURCE>` by specifying values for the following three properties: a. `file.repository.dir`: a directory on the system that hosts the JBoss application server for storing uploaded files, for example, `C:/caNanoLab/fileupload`. **Note:** This directory should be writable by the user that starts the JBoss server, and this directory should be created prior to starting the application. Either double back slashes \\ or a single forward slash / should be used as the file separator if working on Windows platform. b. `application.owner`: the owner of the caNanoLab release 1.4 installation instance, for example, `NCL`. c. `grid.indexserver`: the grid index server from which the caNanoLab grid data services can be automatically discovered, for example, http\://cagrid-index.nci.nih.gov\:8080/wsrf/services/DefaultIndexService, the NCICB's production caGrid index server.

Step	Action

3	Execute the Ant build script `build.xml` located at `<CANANOLAB_SOURCE>` with the target `deploy-application`.
	
	Example: Issue the following commands to execute the Ant script:
	
	   `C:\>cd C:\caNanoLab_1.4`
	   `C:\caNanoLab_1.4>ant deploy-application`
	
	Successful execution of the Ant script creates two deployable web archive war files in the directory `<CANANOLAB_SOURCE>/build`: `caNanoLab.war` and `upt.war`, and places these two wars to the JBoss deploy directory, for example, `<JBOSS_HOME>/server/default/deploy`.

4	Edit the Ant build properties file `deploy.properties` at `<CANANOLAB_SOURCE>` by specifying values for the following four properties: a. `database.host:` the host name of the server hosting the database, for example, `localhost` b. `database.port:` the database port number, for example, `3306` c. `database.user:` the database user name used in the caNanoLab application. Please refer to page 4 . d. `database.password:` the password for the database user specified above.

5	Execute the Ant build script `build.xml` located at `<CANANOLAB_SOURCE>` with the target `prepare-csm-conf`.
	
	Example: Issue the following commands to execute the Ant script:
	
	   `C:\>cd C:\caNanoLab_1.4`
	   `C:\caNanoLab_1.4>ant prepare-csm-conf`
	
	Successful execution of the Ant script sets up the required CSM configurations in the JBoss server. The caNanoLab web application relies on a modified version of NCICB CSM (Common Security Module) version 3.1 for user authentication and role-based authorization. For more information, see the CSM documentation at http://gforge.nci.nih.gov/frs/?group_id=12.
	
	Note: If you are using the JBoss server you set up before from release 1.3, you may choose to skip step 5, be sure to change references of the caNanoLab schema from ***cananolab*** to ***canano*** in the files `<JBOSS_HOME>/server/default/conf/login-config.xml` and `<JBOSS_HOME>/server/default/deploy/mysql-ds.xml`.

Step	Action

6	When deploying both the caNanoLab application and the caNanoLab grid service to the same JBoss server, we recommend increasing the JBoss JVM heap size to 1G bytes and permanent generation (permgen) memory space to 512M bytes by updating the file `<JBOSS_HOME>/bin/run.conf` with the following `JAVA_OPTS`:
	
	`JAVA_OPTS="-Xms512m -Xmx1024m -XX:PermSize=256m -`
	`XX:MaxPermSize=512m`
	` -Dsun.rmi.dgc.client.gcInterval=3600000 -`
	`Dsun.rmi.dgc.server.gcInterval=3600000"`
	
	This configuration file is read when JBoss server starts.

7	When deploying the caNanolab application and the grid service in a production environment, we also recommend updating the default logging behavior of the JBoss server by turning off the unnecessary loggings thus reducing file system requirements for server log files. Please locate the file `<JBOSS_HOME>/server/default/conf/log4j.xml` and add to the beginning of the section `Limit categories` the following text: ``` <!-- ================ --> <!-- Limit categories --> <!-- ================ --> <!-- edited for caNanoLab --> <category name="org.hibernate"> <priority value="WARN"/> </category> <category name="org.springframework"> <priority value="WARN"/> </category> <category name="org.globus"> <priority value="WARN"/> </category> <category name="gov.nih.nci"> <priority value="ERROR"/> </category> <category name="org.displaytag"> <priority value="ERROR"/> </category> <category name="org.directwebremoting"> <priority value="ERROR"/> </category> <category name="net.sf.ehcache"> <priority value="INFO"/> </category> <category name="com.mchange.v2"> <priority value="INFO"/> </category> <category name="org.jboss"> <priority value="INFO"/> </category> <category name="CSM"> <priority value="ERROR"/> ```

Step	Action

	``xml
	</category>
	<category name="Authorization">
	    <priority value="ERROR"/>
	</category>
	 <!-- end of edit for caNanoLab -->
	``

Verification   Once the deployment war files have been deployed and the JBoss application server is correctly configured, you can now start the JBoss application server, which in turn starts the caNanoLab application.

Open the URL http://<APP_SERVER_URL>/caNanoLab/ (for example, http://cananolab-dev.nci.nih.gov:19080/caNanoLab). You should see a disclaimer page. Click on **CLICKING HERE** to go to the Welcome/Login page.

User Provisioning Tool (UPT)

Creating User Accounts	Before users can log in to the caNanoLab application to submit and search data, you must first create their user accounts and assign them to the pre-defined user groups with the pre-defined roles to access the pre-defined protection groups. The caNanoLab application makes use of the NCICB's User Provisioning Tool (UPT), a separate web application, for user account management. The concepts of users, groups, roles, protection groups are defined according the CSM/UPT principles. See the CSM documentation at http://gforge.nci.nih.gov/frs/?group_id=12 for details on these concepts and the use of the UPT tool.

In release 1.4, as a part of the database seed data, a default user group `Public` has been created and has been assigned role `R` to protection groups `protocol`, `nanoparticle` and `report`, i.e. assigned access to search protocols, nanoparticles and reports.

During the caNanoLab application start up, two more default users groups are auto-generated: `<application.owner>_Researcher`, `<application.owner>_PI`, where the value for `<application.owner>` is specified in the file `build.properties` prior to building the application as described on page 7. Automatic roles are assigned for these two default user groups at start up as follows:

1. `<application.owner>_Researcher` is assigned role `CURD` to protection group `sample`, i.e. assigned access to submit samples.

2. `<application.owner>_PI` is assigned role `CURD` to protection groups `sample`, `protocol`, `nanoparticle`, and `report`, i.e. assigned access to submit samples, protocols, nanoparticles and reports.

NOTE:	In release 1.4, a user must be assigned as a caNanoLab administrator to see the `ADMINISTRATION` menu item in the application to log into the UPT tool.

UPT Example   The following steps illustrate an example use of the UPT tool to create a new user, to assign the user to be a caNanoLab administrator, and to assign the user to a pre-defined user group.

Step	Action
1	Launch the UPT tool at http://<APP_SERVER_URL>/upt and login as *superadmin/password*. Use *caNanoLab-upt* as the application name when prompted at the UPT log in. **Note:** The user **superadmin** was created as a part of the database setup.
2	Once logged into the UPT tool, follow these steps: a. Select **User > Create a New User**. b. Create a new user account named **admin**. c. Select **Application > Select an Existing Application**; click **Search**. d. Select **caNanoLab** from the application list. e. Click **View Details**, then **Associated Admins**. f. Assign this user to be an administrator for the caNanoLab application. g. Click **Update Association** to commit the change.
3	Log out of the UPT tool and log back in as **admin/admin**. Use **caNanoLab** as the application name when prompted at the UPT log in. **Note:** When a new user account is created, its initial default password is the same as its login name. Given that the UPT tool is lacking a function for users to manage their own passwords, users can only update their passwords within the caNanoLab application. See the Notes below on how to update user passwords.
4	Select **User > Select an Existing User**, and click **Search**. Select **admin** from the User list. a. Click **View Details**, then **Associated Groups**. b. Select **<application.owner>_PI** from the pre-defined group list and assign it to the user. c. Click **Update Associations** to commit the change.

Follow similar steps to create other application user accounts and to assign them to different users groups, as appropriate.

NOTES:	• The initial passwords for user accounts are set to be the same as their user accounts. When a user first logs into the caNanoLab application at http://<APP_SERVER_URL>/caNanoLab, he/she would be prompted to change the initial password. Users can update their passwords at any time within the caNanoLab application. • Only users assigned as caNanoLab administrators and assigned to the `<application.owner>_PI` group can delete information from caNanoLab. • Publicly available data can be browsed through Browse caNanoLab links on the home page without logins. • Users logged into the application have to be assigned to one of the predefined user groups to be able access data. • A new sample source entered during sample creation is automatically created as a new user group, and the new sample is auto-assigned role `R` to the newly created user group. Using the UPT tool, one can assign different roles to different protection groups for these new user groups. • A user can be assigned to multiple user groups, and roles and protection groups are assigned to user groups.

For more information about how to use the UPT tool for managing user accounts, contact NCICB Application Support at ncicb@pop.nci.nih.gov and request that the caNanoLab technical team give you a demonstration of the UPT tool in the context of the caNanoLab application.

caNanoLab Grid Data Service

Assumptions and Requirements

The JBoss server has been downloaded, installed and configured correctly, and the release 1.4 caNanoLab application has been correctly installed. If you had started the JBoss server to test the caNanoLab application installation, please shut it down now. In release 1.4, the caNanoLab grid service is configured to be deployed to the Tomcat embedded in the same JBoss server used for deploying the caNanoLab application. Users are no longer required to install a separate Tomcat server or to download and install Globus Toolkit.

In order for the grid service to successfully register with the NCICB production index server, the server and the port number hosting the JBoss server must be publicly accessible.

Installation and Deployment Follow these steps to create and deploy a caNanoLab grid data service.

Step	Action
1	If not done already, set up an environment variable `JBOSS_HOME` to point to the JBoss server install direction as in step 1 on page 8.
	If not done already, edit the Ant build properties file `build.properties` at `<CANANOLAB_GRID_SOURCE>` as in step 2 on page 9.
	Note: the value of the property `application.owner` is used to construct the grid location name that is shown as a part of the public-browse location drop-down list within the caNanoLab application.
2	Edit the Ant build properties file `grid.properties` at `<CANANOLAB_GRID_SOURCE>` by specifying values for the following properties: a. `grid.service.hostname:` virtual host name for the server hosting JBoss, for example, `cananolab-dev.nci.gov` b. `grid.service.port:` port number configured for the Tomcat service embedded in JBoss, for example, `19080` c. `country:` the country in which the caNanoLab grid service is running, for example, `US`. d. `city:` the city at which the caNanoLab grid service is running, for example, `Rockville`. e. `zipcode:` the zip code at which the caNanoLab grid service is running, for example, `20852`. f. `state:` the state at which the caNanoLab grid service is running, for example, `MD`. g. `street:` the street at which the caNanoLab grid service is running, for example, `2115 East Jefferson Street`. h. `affiliation:` the affiliation of the person who is responsible for maintaining the caNanoLab grid service, for example, `NCICB`. i. `email:` the email of the person who is responsible for maintaining the caNanoLab grid service, for example, ncicb@pop.nih.gov j. `first.name:` the first name of the person who is responsible for maintaining the caNanoLab grid service. k. `last.name:` the last name of the person who is responsible for maintaining the caNanoLab grid service. **Note**: Properties c through g are grid service metadata that are required for the grid index server to visually map the location of a grid service in the caGrid portal, and are useful during grid service auto-discovery. Properties h through k are optional, but useful information for others to look up the grid service.

Step	Action

3	Execute the Ant build file `build.xml` located at `<CANANOLAB_GRID_SOURCE>` with the target `deploy-grid`. For example,
	
	`C:\>cd C:\caNanoLab_1.4`
	`C:\caNanoLab_1.4>ant deploy-grid`
	
	Successful execution of the Ant build script generates a war file `wsrf-canano.war` at the `<CANANOLAB_GRID_SOURCE>/build` directory and places the war file to the JBoss deploy directory, for example, `<JBOSS_HOME>/server/default/deploy`.

4	Execute the Ant build script `build.xml` located at `<CANANOLAB_SOURCE>` with the target `prepare-globus-lib`. For example,
	
	`C:\>cd C:\caNanoLab_1.4`
	`C:\caNanoLab_1.4>ant prepare-globus-lib`
	
	Successful execution of the Ant build script extracts eight jar files for deploying Globus to the JBoss server at `<JBOSS_HOME>/server/default/lib`.

5	If not done already, follow step 6 and 7 on pages 10 and 11 to configure appropriate memory and logging settings for the JBoss server.

Verification   Once the grid service war file has been deployed and you have correctly configured JBoss server with required Globus libraries, you can now start the JBoss server, which in turn starts the caNanoLab application and the caNanoLab grid service. To verify the creation of the caNanoLab grid data service, open the URL http://<APP_SERVERL_URL>/wsrf-canano/services/cagrid/CaNanoLabService, for example, http://cananolab-dev.nci.nih.gov:19080/wsrf-canano/services/cagrid/CaNanoLabService. You should see a page that says:

    cagrid/CaNanoLabService

    Hi there, this is an AXIS service!

It takes about ten minutes before the grid service is successfully registered in the index server. Wait a short while and go into the caGrid portal at http://cagrid-portal.nci.nih.gov/web/guest/home. You should see your grid service showing up on the portal map. This completes the release 1.4 installation.

Note: Your grid service name will not show up in the **Location** drop-down list on the home page of the caNanoLab application because the grid service is considered local to the caNanoLab application installation. The drop-down list shows all other registered caNanoLab grid services.

For more information and help on caGrid service configuration and deployment, refer to the caGrid wiki page: http://www.cagrid.org.

If you need further assistance on setting a caNanoLab grid service, contact NCICBIIT Application Support.

Contacting Application Support

NCICB Application Support	http://ncicb.nci.nih.gov/NCICB/support ncicb@pop.nci.nih.gov Telephone: 301-451-4384 Toll free: 888-478-4423