CoreFlow DOWNLOAD AND INSTALLATION INSTRUCTIONS

The latest stable release of CoreFlow is available for download at
http://pawsonlab.mshri.on.ca/CoreFlow/download.php.

Installation Instructions

INSTALLATION REQUIREMENTS

- ✔ Apache HTTP server: http://httpd.apache.org/ - on both web and application servers
- ✔ PHP version 5 or later: http://php.net/ with MySQL support (php-mysql) - web and application servers
- ✔ MySQL DBMS version 5 or later: http://www.mysql.com/ - web and application servers
- ✔ Perl: http://perl.org with the following modules: - web and application servers
  - DBI
  - DBD::mysql
  - XML::XPath
  - XML::Parser
  - If required by above XML modules, install Expat XML parser: http://expat.sourceforge.net/
- ✔ R statistical software: http://www.r-project.org/ - application server
- ✔ Python version 2.4 or later: http://python.org/ with MySQLdb module installed - application server
- ✔ BioPython: http://biopython.org/wiki/Main_Page - application server
- ✔ Optional: Install git from http://git-scm.com/. For more information on the use of git for management of your analysis scripts developed in task playgrounds please refer to the YouTube videos on http://pawsonlab.mshri.on.ca/CoreFlow.
- ✔ **The application server must be configured to allow execution of shell scripts from the runR.pl

INSTALLATION PROCEDURE

INSTALLING CoreFlow FOR THE FIRST TIME

- CoreFlow can be installed on any Unix/Linux/Mac platform
- Database and server code can be installed either on the same hardware node, or on separate nodes, according to your preference.

1. Download and extract coreFlow_web_2013_01.tgz into your web server directory.

   The unpacked directory contains:
   - A folder named CoreFlow – main website directory

- A database dump file, named mol_bio_dump.sql, and
- A copy of the Installation Instructions in PDF format

2. Copy the R subfolder from the unpacked CoreFlow directory onto your application server. The application server will run the R, Perl, Python/BioPython, etc. scripts that are attached in each data analysis task.

3. Define the storage location of R CGI files on the (possibly different) application server:

   In .config.xml, edit the following section:

   <R>
   <cgi_dir>http://your_app_server_name/R_directory/</cgi_dir> <!-- e.g. This may be "http://localhost/CoreFlow/R/" -->
   <temp_dir>**full_path_to_any_temp_directory**</temp_dir> <!-- e.g. **/tmp/CoreFlow/temp**; if this directory does not exist, create it, AND change permissions to a+rwx -->
   </R>

4. Edit your Apache httpd.conf file as follows (if the web server and application server are on different machines, perform the configuration on both machines):

   - Define ServerName, DocumentRoot and ServerAlias for your CoreFlow setup directory.
   - Set LD_LIBRARY_PATH to point to your MySQL directory.
   - It is helpful to designate an ErrorLog file, e.g. /my_install_dir/my_coreflow-error.log.

   Example:
   (replace the highlighted directives according to your server's specifications)

```
<Directory /var/www/html/CoreFlow>
    IndexIgnore .config.xml
    DirectoryIndex index.php
    Options Indexes FollowSymLinks
AddType application/x-httpd-php .php        # MUST HAVE THIS TO VIEW PHP PAGES!!!!
    Options ExecCGI
     AddHandler cgi-script .cgi .pl .py      # DO NOT INCLUDE .PHP HERE!!!
     AllowOverride AuthConfig
     Order allow,deny
     Allow from all
  </Directory>
```

5. On your database server, create an empty MySQL schema that will be populated in the next step:

```
mysql> create database mol_bio;
```

   (replace *mol_bio* with a schema name of your choice)

6. Use the database dump file **mol_bio.sql** to add tables to the newly created schema. At the command prompt, type:

```
xterm> mysql –u root –p mol_bio < mol_bio.sql
```

7. Create two new MySQL users – a web user, e.g. '*coreflow_web*', and an app user, e.g. '*coreflow_app*':

```
mysql> create user coreflow_web identified by 'web_passwd';

mysql> create user coreflow_app identified by 'app_passwd';
```

8. Grant SELECT, INSERT and UPDATE privileges on database *mol_bio* to the new users:

```
mysql> GRANT SELECT, CREATE TEMPORARY TABLES, LOCK TABLES ON mol_bio.* TO
'coreflow_web'@'your_web_server_name_or_ip_address';

mysql> GRANT SELECT, INSERT, UPDATE, DELETE, ALTER, DROP, LOCK TABLES, CREATE
TEMPORARY TABLES ON mol_bio.* TO 'coreflow_app'@'your_web_server_name_or_ip_address';

mysql> flush privileges;
```

9. In the CoreFlow folder, edit .config.xml:

```
<config>
  <db> <!-- related to the database mysql server -->
  <host>YOUR DATABASE SERVER HOSTNAME</host>
    <database>mol_bio</database>
    <user passwd="app_passwd">coreflow_app</user>
    <user_public passwd="web_passwd">coreflow_web</user_public>
...
</config>
```

For users who are allowed to insert/delete/update (Write) into DB or eXecute files, add their workstations' IP addresses to the following line:

<allow_wx_IPs>(**user1\.workstation\.IP\.address**)|(**user2\.workstation\.IP\.address**)|...
</allow_wx_IPs>

THERE MUST BE **NO SPACES** IN THIS LINE!!!

10. Copy .config.xml from your CoreFlow web server directory into the R subdirectory on the application server

UPGRADING TO A NEWER VERSION OF CoreFlow

1) Make a copy of your current .config.xml file
2) Download the .tgz and the mysql_dump files from the 'Download' section of the public CoreFlow website
3) Erase your current CoreFlow installation directory and extract the downloaded archive
4) Replace the database tables with the latest downloaded mysql dump.
5) Copy back your saved .config.xml file into the extracted CoreFlow directory