CONVERSION SOFTWARE REGISTRY (CSR): INSTALLATION MANUAL

Version 0.1

REVISION HISTORY			
Revision	**Date**	**Author**	**Notes**
0.1	09/22/2010	MO+RK+PB+KM	Initial version of the document.

The Conversion Software Registry (CSR - http://isda.ncsa.uiuc.edu/NARA/CSR) was created at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign.

The research and development of CSR has been supported by a National Archive and Records Administration (NARA) supplement to NSF PACI cooperative agreement CA SCI-9619019. The main contributions to designing, developing, testing and documenting the CSR came from Kenton McHenry, Rob Kooper, Michal Ondrejcek and Peter Bajcsy. This document represents a current description of the on-going research and development of CSR and hence the document is updated on a regular basis.

CONTENTS

INTRODUCTION

The Conversion Software Registry (CSR) system is a web-based service for finding software to convert a file form one format to another. The CSR database stores information specifically about software and its supported input and output formats. In addition to being a general web service the CSR also serves as the source of information for configuring NCSA Polyglot, a system designed to execute file format conversions automatically across third party software. The CSR database is available for scientific communities and instances of it can be installed and run locally. This manual describes the step by step installation of the CSR system.

In this document we assume that the underlying OS is a running Linux Ubuntu Server (preferred) or the Linux Ubuntu (GUI) system 10.04. We also assume either dedicated computer hardware or a virtual machine (VM) running in a computer cloud. The installation process for other Linux flavors (Fedora, Red Hat, Debian etc.) is similar to the installation process for the Ubuntu Linux flavor, and hence the installation could be followed without major changes. This document does not cover installation of the Ubuntu OS server nor topics related to setting up and running different instances of VMs.

CHAPTER 2. CSR SYSTEM REQUIREMENTS

The CSR system is a MySQL open source database with a web-based interface written in the PHP scripting language. The minimum software requirements for running the CSR system are:

- Ubuntu Linux, preferred server version 10.04 or later
- Apache HTTPD server with PHP module
- MySQL
- Apache Subversion (SVN) version control system

The CSR scripts, data model, and data are provided in the accompanying ZIP file.

APACHE HTTPD SERVER WITH PHP

Apache is an open source, secure, efficient, and extensible HTTP server. It features support for HTTPS, virtual hosting, CGI, SSI, IPv6, easy scripting and database integration, request/response filtering, many flexible authentication schemes, and more. PHP is an HTML-embedded scripting language to write dynamically generated pages which exists as a module for the apache web server.

http://httpd.apache.org
http://www.php.net

MYSQL

MySQL is a fast, stable, multi-user, multi-threaded SQL database server.

http://www.mysql.com

SVN SUBVERSION CONTROL SYSTEM

Subversion or SVN is a version control system which allows many programmers to work on a common set of source code files.

http://subversion.apache.org

CHAPTER 3. CSR INSTALLATION

TERMINOLOGY

The Manual is written for the Ubuntu Server OS and hence utilizes primarily the command line interface. In the case of Ubuntu Desktop Edition, a user can access the command line interface by launching a Terminal from the "Applications" menu:

Applications -> Accessories -> Terminal

choosing the "Accessories" sub-menu and clicking on "Terminal".

The syntax for an example command is given below:

```
rm filename
```

where a command, in this example, "rm" appears in constant width type. The variable part (here, filename) appears in *constant width italic* type. The command "rm" followed by the name of a file removes the file from the current directory.

PRE-INSTALLATION

Any user installing a new instance of the CSR database needs an svn username, svn password, account name and account password from the personnel at the Image Spatial Data Analysis (ISDA) group at NCSA

in order to successfully install the database. Please, request the information below from isda@ncsa.illinois.edu:

svn_username and *svn_password*
account_name and *account_password*

We will provide the necessary information via e-mail or other communication channels.

Next you will need to make sure you have administrative access to your machine by typing sudo ("super-user do") command. The command is executed in the user shell, not a root shell, and will gives you temporary root (super-user) privileges. Not all users are allowed to use the sudo command. Please, make sure that the system administrator of your Linux machine has assigned the privileges to you so that you can type:

```
sudo –s
```

You will be prompted for your own user password. Typing it does not show any substitute characters on the screen like seen within GUI applications. This is a Unix security feature. Take extra care to log out before walking away from your computer since other users can use the command again without being asked for a password within a certain period of time set in the system configuration.

Next we will insure that our system is completely up to date. The command "apt-get" calls Ubuntu's Advanced Packaging Tool (APT) for the installation and updating of software packages. A GUI Update Manager also provides updates and upgrades of the software packages:

System -> Administration -> Update Manager

Go back to the command line interface type following command:

```
apt-get update
```

followed by:

```
apt-get dist-upgrade
```

The system is now ready for the CSR to be installed.

SOFTWARE INSTALLATION

The following packages will be installed (http://packages.debian.org):

- mysql-server-5.1 - This package contains all the infrastructure needed to setup system databases.
- apache2 – Package containing the web server.
- libapache2-mod-php5 - Provides the PHP5 module for the Apache 2 webserver.
- php5-cgi - CGI binary interpreter built for use in Apache 2.
- php5-mysql - Provides modules for MySQL database connections directly from PHP scripts. It includes the generic "mysql" module which can be used to connect to all versions of MySQL and the pdo_mysql module for use with the PHP Data Object extension.
- subversion - Includes the Subversion client (svn), tools to create a Subversion repository (svnadmin), and the svn server to make a repository available over a network (svnserver).

To install these packages type the following command at the command line (not this is all in one line):

```
apt-get -y install mysql-server-5.1 apache2 libapache2-mod-php5 subversion php5-mysql php5-cgi
```

The installation, downloading, unpacking, and installation usually takes a couple of minutes. During the installation a package configuration window appears to set the MySQL administrative “root” user password. Set the password (you will be prompted to repeat it) and remember it for further use.

You might see the message “Could not reliably determine the server's fully qualified domain name” on the screen while starting the apache2 web server. If this occurs the server will default to localhost (127.0.1.1). This will not influence the operation of the web server.

We now navigate to the web directory:

```
cd /var/www
```

From here remove the default index.html created by the apache installation. This file contains the default web page when a browser accesses this directory. The main page of the CSR directory is index.php and to make sure that it is called first we delete the default index.html:

```
rm index.html
```

We can now install the CSR HTML and scripts from the svn repository. The main svn server hosted by the ISDA group at NCSA is isda.ncsa.uiuc.edu. The checkout is achieved by typing:

```
svn co https://isda.ncsa.uiuc.edu/svn/nara/trunk/CSR .
```

You will be asked to Reject/Accept the svn server certificate temporarily or permanently. Type t or p to choose one of the options respectively. The svn repository is protected by a login password. In order to check out the CSR files from ISDA NCSA repository one has to obtain a username and password as

described in the previous section. Log into svn with the provided user name and password and complete the checkout.

At this stage we can test the Apache server with the CSR php scripts (no database access yet). Open your preferred browser on this machine and type "http://localhost". The main blue-ish CSR web page should appear with the message 'Could not connect: Access denied'. If this page does not appear try to re-start the Apache server by typing in the terminal:

```
/etc/init.d/apache2 restart
```

POPULATING THE CSR DATABASE

So far we have installed the server and downloaded the php scripts from svn. In this section we will populate CSR database with the needed data model and data. To do this issue the following command:

```
mysql -u root -p -h localhost
```

You will be asked for MySQL root password set previously during the server installation. Once logged in the prompt changes to "mysql>". You are now operating within the mysql environment. The mysql server will send each SQL statement you now type to the server to be executed. We create a database called "csr" by issuing the following command:

```
create database csr;
```

Next we add accounts using the "grant" command:

```
grant all on csr.* to ' account_name '@'%' identified by 'account_password';
```

In this example the account's name is "account_name" with a password "account_password". These values should be substituted with a name and password that you choose. The created account has full privileges to configure the database. The account uses the "%" wildcard for the host allowing connections from any host and not only the localhost.

We can now exit the mysql environment by typing:

```
quit
```

The database has been created on your server and you are ready to copy the database tables from the ISDA server. The following set of commands will download the content for the CSR database and load it into your MySQL database:

```
cd /tmp
```

```
wget http://isda.ncsa.uiuc.edu/NARA/csr_nara.sql.gz
zcat csr_nara.sql.gz | mysql -u root -p -h localhost csr
```

You will be asked again for the MySQL password since you are entering a mysql server environment. Note that uncompressing the zip file and populating the database with the data might take several minutes without any terminal feedback.

As a final installation step we recommend to reboot the system. This is only needed when new files were installed during the initial update (dist-upgrade):

```
reboot
```

CHAPTER 4. CSR UPDATES

Use the CSR JIRA reporting system at http://jira.ncsa.uiuc.edu) and/or the ISDA web pages at http://isda.ncsa.uiuc.edu/NARA/CSR/ to check whether you have the most up-to-date version of CSR.

CHAPTER 5. MISCELLANEOUS

BUG REPORTS AND BUG FIXES

We appreciate your feedback on the CSR functionality and usability. If you find bugs, please, report them to http://jira.ncsa.uiuc.edu or use the link from the CSR pages. Use JIRA also for suggestions and improvements related to this manual. People can file bug reports and improvement suggestions into the JIRA system at ISDA NCSA.

ACKNOWLEDGMENTS

This research was supported by a National Archive and Records Administration (NARA) supplement to NSF PACI cooperative agreement CA #SCI-9619019.

Appendix A. Summary of commands

Normal installation part, assuming a running Linux Ubuntu system 10.04.

Become root:

```
sudo -s
```

Update installation files:

```
apt-get update
apt-get dist-upgrade
```

Install mysql-server, apache httpd, php and svn:

```
apt-get -y install mysql-server-5.1 apache2 libapache2-mod-php5
subversion php5-mysql php5-cgi
```

Install CSR files:

```
cd /var/www
rm index.html
 svn co https://isda.ncsa.uiuc.edu/svn/nara/trunk/CSR .
```

Create a user for CSR:

```
mysql -u root -p -h localhost
create database csr;
grant all on csr.* to ' account_name '@'%' identified by
'account_password';
quit
```

Copy database from ISDA and load into MYSQL:

```
cd /tmp
wget http://isda.ncsa.uiuc.edu/NARA/csr_nara.sql.gz
zcat csr_nara.sql.gz | mysql -u root -p -h localhost csr
```

Final reboot:

```
reboot
```