



THE
POWER
TO KNOW.

SAS® Enterprise Content Categorization 12.1

Administrator's Guide



The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2012.
SAS® Enterprise Content Categorization 12.1: Administrator's Guide. Cary, NC:
SAS Institute Inc.

SAS® Enterprise Content Categorization 12.1: Administrator's Guide

Copyright © 2012, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

For a hard-copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a Web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government Restricted Rights Notice: Use, duplication, or disclosure of this software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19, Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st electronic book, August 2012

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at support.sas.com/publishing or call 1-800-727-3228.

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Contents

About This Book	5
Audience	5
Prerequisites	5
Conventions	6
 What's New in SAS Enterprise Content Categorization Studio 12.1	7
Changes to the SAS Content Categorization Studio 12.1 Documentation	7
 1 About SAS Enterprise Content Categorization Studio	9
1.1 What Is SAS Enterprise Content Categorization Studio?	9
1.2 Benefits to Using SAS Enterprise Content Categorization Studio	11
1.3 How Do the Collaborative Features Work in SAS Enterprise Content Categorization Studio?	12
1.4 Using Collaborative Operations	13
1.5 Architecture	14
 2 Setting Up a Project on Your Server	15
2.1 Overview of Setting Up a Project on Your Server	15
2.2 Installing and Configuring the Server	16
2.2.1 Install the Server	16
2.2.2 Change the Size of the .cnf File for Linux	16
2.3 Installing the _tk240 Utility	17
2.4 Setting Up ODBC Data Sources	18
2.4.1 Driver Types	18
2.4.2 Installing the ODBC Drivers	18
2.4.3 Configuring Data Sources	19
2.4.3.A Credentials	19
2.4.3.B Configuring ODBC Data Sources on UNIX	20
2.4.3.C Configure MySQL Data Sources	22
2.4.3.D Configure Microsoft SQL Server Data Source	25
2.4.3.E Download and Configure Oracle ODBC Drivers	32
2.5 Defining Users	36
2.6 Logging In	36
2.7 Quick Start Guide Using _tk240	37

3 Using the Interface	39
3.1 Collaboration and the User Interface	39
3.2 The Menu Bar	40
3.3 The Server Operation Windows	40
3.3.1 Accessing Server Permissions	40
3.3.2 Access and Use the Project Server Permissions Window	41
3.3.3 Use the Set Server Permissions for a Selected Node Window	43
3.3.4 The Set User Permission Windows	45
3.4 The Options Window: Shared Projects Folder	46
3.5 The Miscellaneous Windows	48
3.5.1 The Repository Login Window	48
3.5.2 The Select a Directory Window	49
3.5.3 The Select a Project Window	50
3.5.4 The Change Password Window	51
3.5.5 The Upload to Server Window	52
 4 Administering a Collaborative Project	 55
4.1 Overview of Administering a Collaborative Project	55
4.2 Create a Project on a Local Machine	57
4.3 Upload a Project	58
4.4 Access a Remote Project	60
4.5 Setting Permission Levels in a Project	62
4.5.1 Understanding Permissions	62
4.5.2 Overview of Project Permission Levels	62
4.5.3 The Permission Levels	63
4.5.4 Setting, Changing, and Deleting Permission Levels	64
4.5.4.A Set Permissions at the Project Level	64
4.5.4.B Set Permissions for Individual Nodes	65
4.5.4.C Changing Project Level Permissions	67
4.5.4.D Delete Project Level Permissions	67
4.6 Remove a Project from the Server	68
 Reference Section	 71
 A Using the _tk240 Utility	 73
A.1 Overview of the _tk240 Utility	73
A.2 Command-Line Parameters	74
A.3 Command-Line Operations	76
A.4 Performing Database Operations	80

A.4.1 Initialize the SAS Enterprise Content Categorization Studio Database	80
A.4.2 Update the SAS Enterprise Content Categorization Studio Database	81
A.4.3 Add Support for Multiple Databases	81
A.4.4 Initialize the Testing Document Repository	82
A.4.5 Lock and Check Out a Project	82
A.4.6 Commit and Unlock a Project	82
A.4.7 Clear Database Locks	83
A.5 Performing Server Operations	84
A.5.1 Add Users	84
A.5.2 Add Users with Microsoft SQL Server with Windows Authentication Mode	84
A.5.3 Modify User Permissions and Passwords	85
A.5.4 Delete a User	85
A.5.5 Show All Users	86
A.5.6 Upload a Project Using _tk240	86
A.5.7 Open a Remote Project	87
A.5.8 Delete a Remote Project Using _tk240	87
A.5.9 List Remote Projects	88
A.6 Performing General Operations	88
A.6.1 Rebuild a Project	88
A.6.2 Locate Missing Concepts in a Project	88
A.6.3 Run a Project Check	89
A.7 Set the Path to the SAS License	89
B The Program Files	91
B.1 Overview of the Program Files	91
B.2 Program File Directives	91
B.3 Configuration Example	92
C Recommended Reading	95
D Glossary	97
Index	99

About This Book

Audience

SAS Enterprise Content Categorization Studio is SAS Content Categorization Studio for multiple users. This product includes support for LITI concepts and collaborative projects.

This book is written for administrators who perform some, or all, of these tasks:

- Database administrator who use _tk240, which is part of SAS Enterprise Content Categorization Servers
- Project administrators who assign user permissions for collaborative projects
- Database administrators who are also project administrators

SAS Enterprise Content Categorization Studio enables you to use this software with other SAS products. This manual focuses on tasks that define and configure the collaborative operations for SAS Enterprise Content Categorization Studio.


Prerequisites

Here are the prerequisites for using the collaborative operations for SAS Enterprise Content Categorization Studio:

- SAS Enterprise Content Categorization Studio loaded onto your machine
- Appropriate server permission for all users, assigned by the database and project administrators
- Read *SAS Content Categorization Studio: User's Guide*

Conventions

This manual uses the following typographical conventions:

Convention	Description
TGM_ROOT	The root directory where SAS Enterprise Content Categorization Studio is installed, typically the following: Windows: C:/Program Files/SAS/SAS Content Categorization Studio UNIX: /opt/SAS_collab-server
-admin -regular	The code examples for _tk240 are shown in a fixed-width font.
<admin_password>	The code examples that are enclosed in brackets display information that is replaced by your system values.
OK button	The labels for user interface controls are shown in a bold, sans-serif font.
Top	The names of taxonomy nodes appear in a fixed-width font.
www.sas.com	The hypertext links are shown in a light blue, fixed-width font, and are underlined.
	The Question Mark button accesses <i>SAS Content Categorization Studio: User's Guide</i> in PDF format.

What's New in SAS Enterprise Content Categorization Studio 12.1

The new features for SAS Enterprise Content Categorization Studio enable you to do the following:

Operations that are specific to SAS Enterprise Content Categorization Studio are enabled by the setinit file for this product. This license is the SAS installation data file (SID file) that is included in the Software Order E-mail (SOE) that you received.

Changes to the SAS Content Categorization Studio 12.1 Documentation

See the changes to SAS Content Categorization Studio documentation:

Documentation	12.1 Product	Tasks and 5.2 Product References
<i>SAS Content Categorization Studio: Installation Guide</i>	SAS Content Categorization Studio	Install the single user or the enterprise version of SAS Content Categorization Studio that you purchased. The enterprise version automatically installs support for collaborative features and LITI concepts.
	SAS Enterprise Content Categorization Studio	
<i>SAS Content Categorization Studio: User's Guide</i>	SAS Content Categorization Studio	Create a SAS Content Categorization Studio project, test, and upload the project to SAS Content Categorization Server. This guide is written for a single user and is a companion book for <i>SAS Enterprise Content Categorization Studio: User's Guide</i> .
<i>SAS Enterprise Content Categorization Studio: Administrator's Guide</i>	SAS Enterprise Content Categorization Studio with collaborative operations.	Configure your server for collaborative operations. (In the 5.2 release, this book was <i>SAS Content Categorization Collaborative Server: Administrator's Guide</i> .)

Documentation	12.1 Product	Tasks and 5.2 Product References
<i>SAS Enterprise Content Categorization Studio: User's Guide</i>	SAS Enterprise Content Categorization Studio with collaborative operations and LITI concepts capabilities.	See the cell above and use this guide to understand how collaborative operations work. Use the second part of this guide to write LITI rules and to upload these rules to SAS Content Categorization Server. (In the 5.2 release, LITI rules were explained in <i>SAS Contextual Extraction Studio: User's Guide</i> .)
<i>SAS Enterprise Content Categorization Servers: Administrator's Guide</i>	Download any, or all, of the following: <ul style="list-style-type: none"> - SAS Content Categorization Server - SAS Content Categorization Collaborative Server - SAS Content Categorization Java API - SAS Content Categorization Python API - SAS Document Conversion Server and Java API 	Install, configure, and use SAS Content Categorization Server, SAS Content Categorization Collaborative Server, and SAS Document Conversion Server. You can also upload .li files using this product. In the 5.2 release, the information in this book was found in the following manuals: <ul style="list-style-type: none"> - <i>SAS Content Categorization Server: Administrator's Guide</i> - <i>SAS Content Categorization Collaborative Server: Administrator's Guide</i> - <i>SAS Document Conversion: Developer's Guide</i>
<i>SAS Content Categorization Single User Servers: Administrator's Guide</i>	Download any, or all, of the following: <ul style="list-style-type: none"> - SAS Content Categorization Server - SAS Content Categorization Java API - SAS Content Categorization Python API - SAS Document Conversion Server and Java API 	Install, configure, and use SAS Content Categorization Server and SAS Document Conversion Server In the 5.2 release, the information in this book was found in the following manuals: <ul style="list-style-type: none"> - <i>SAS Content Categorization Server: Administrator's Guide</i> - <i>SAS Document Conversion: Developer's Guide</i>.

Chapter: 1

About SAS Enterprise Content Categorization Studio

- *What Is SAS Enterprise Content Categorization Studio?*
- *Benefits to Using SAS Enterprise Content Categorization Studio*
- *How Do the Collaborative Features Work in SAS Enterprise Content Categorization Studio?*
- *Using Collaborative Operations*
- *Architecture*

1.1 What Is SAS Enterprise Content Categorization Studio?

In most organizations it is necessary to obtain information about, and from, data that is created internally and externally. This process is expedited when a team of subject matter experts work together to create a single project on a server. The collaborative operations are one of the features that distinguish SAS Enterprise Content Categorization Studio from SAS Content Categorization Studio.

Using the collaborative server and an intuitive, Windows interface, multiple users can access a single project residing on a server. At the same time, these users can work on a cached version on their local machines and commit the changes to the server as necessary. Users with various skill sets and levels of expertise develop a taxonomy. These users write rules for the categories that

classify data and the concepts that extract entities. These users can also test these rules by using sample documents.

You and your team can use the collaborative operations to create, maintain, and revise a taxonomy of topics and entities according to your areas of expertise. An administrator can specify the various levels of access for all users.

Easy permission setting

The database administrator adds users to the project and sets their permission levels. Easy-to-use interfaces within the project enable the project administrator to set permissions for these users both across the taxonomy and at the individual category and concept level.

Easy distributed workflow

Processes are distributed across various areas of expertise. For example, taxonomists develop categories. Subject matter experts later write the rules to define these categories.

Easy version management

Use the revision log to see, and revert to, earlier versions of your taxonomy, categories, and concepts.

Levels of permissions

Grant users permissions to work on a project according to their level of expertise. For example, grant some users permissions to read and write rules. All users added to the database have Read-Only permissions.

1.2 Benefits to Using SAS Enterprise Content Categorization Studio

SAS Enterprise Content Categorization Studio provides users with the following add-on benefits to SAS Content Categorization Studio:

Use the expertise of a wide range of subject matter experts

SAS Enterprise Content Categorization Studio adds features that enable multiple subject matter experts to work together on one SAS Content Categorization Studio project. These benefits include permission setting, and the shared projects folders that enable two or more developers to work on a single project on one machine.

Control access to the project

The database and project administrators set permissions at database, project, and node levels to restrict access, where appropriate.

Automatic project updates

Use the Options window to automate the updates for your project.

Enable two or more users to build separate projects on one machine

Two, or more developers, create projects on one machine using the cached version in the shared projects folders.

The following benefits for SAS Content Categorization Studio extend to SAS Enterprise Content Categorization Studio:

Empower subject matter experts and taxonomists by providing a simple, visual interface where you build a taxonomy, define rules, and test

SAS Content Categorization Studio includes easy-to-use Windows interfaces that enable you to build large, complex, and hierarchical taxonomies. Specify your own rules, test them, and generate `.mco`, `.concepts`, and `.li` files that are applied by SAS Content Categorization Server to input documents.

Develop metadata for your information

SAS Content Categorization Studio uses advanced linguistic technologies to identify metadata in, and about, your documents.

Improve the business value of information technology and the corporate data that it manages

SAS Content Categorization Studio creates `.mco`, `.concepts`, and `.li` files that automate the classification and extraction of entities from input documents during real time using SAS Content Categorization Server.

Save money on information retrieval and organization costs

All of the information created by, or within, your organization can be classified and retrieved. Find related information, whether you know the exact terms that you are seeking.

1.3 How Do the Collaborative Features Work in SAS Enterprise Content Categorization Studio?

Use SAS Enterprise Content Categorization Studio to enable multiple subject matter experts to develop one project, while controlling access according to expertise. These experts use SAS Enterprise Content Categorization Studio as explained in the following paragraphs:

SAS Enterprise Content Categorization Studio is a Windows application that anyone can use to develop taxonomies that classify and extract the information found in your organization. Interactively identify the data that you need without using a programming language.

SAS Enterprise Content Categorization Studio enables users to easily create taxonomies, write rules, and test these rules against a variety of testing sets. You can upload the output `.mco`, `.concepts`, and `.li` files to SAS Content Categorization Server where they are applied to input documents.

You can also use the `_tk240` command-line program to build the binary files. These files are the `<language>.mco` file for categories, `<language>.li` file for LITI concepts, and the `<language>.concepts` file for regular concepts. All of these files are used by SAS Content Categorization Server. For more information, see Appendix A: *Using the _tk240 Utility*.

1.4 Using Collaborative Operations

The ability of multiple developers to work together depends on collaborative operations:

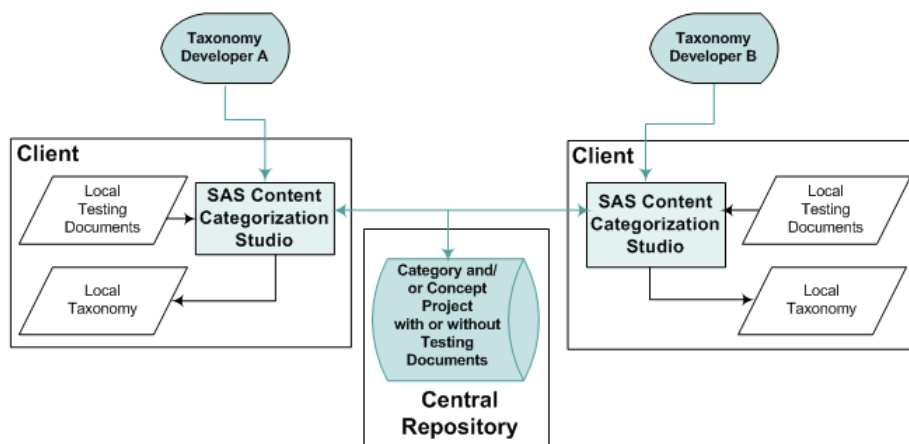
- Create a collaborative, customized SAS Enterprise Content Categorization Studio project using categories and concepts that form a taxonomy residing on a server. This project can be accessed by multiple developers using a single machine that has different shared projects folders, or by users who are each working on their own local machines.
- Set project access levels for each of the individual developers and teams if you are an administrator. These permissions, or access levels, can be set at the project, and at the category and concept levels. For example, choose to restrict development permissions for specific taxonomy nodes to a single individual or group. You can simultaneously grant this person or group wider access to other categories and concepts.
- Track modifications to the project using a revision log for both a category rule and a concept definition. You can also use the information in this log file to revert to an older version of the rule or definition.
- Simplify the task of keeping a local project up-to-date with its counterpart on the server.
- Make changes to the taxonomy when a project is up-to-date with the server. This feature ensures that the modifications, additions, and deletions that are made by one developer are not overwritten by another.
- Automate commit operations for all of the changes to the taxonomy.
- Automate a number of server and cached project update processes and set up separate shared projects folders using the Options window. This process enables multiple developers to work on individual cached projects that are located on one machine.
- (Optional) Upload testing documents to the server to enable other developers to use these documents to reproduce your results.

In summation, these features provide the benefits of collaboration. These operations also enable your team to balance necessary control with optimal flexibility in order to meet your organization's project development requirements.

1.5 Architecture

Two, or more, subject matter experts, or developers, can work together to create one SAS Enterprise Content Categorization Studio project.

Figure 1-1 SAS Enterprise Content Categorization Studio Architecture



A SAS Enterprise Content Categorization Studio project that is used for collaborative work, resides on a remote server. The server enables multiple subject matter experts to work together on the same project. This project is cached on their local machines. A project that resides on a server is called a remote project.

2

Setting Up a Project on Your Server

- *Overview of Setting Up a Project on Your Server*
- *Installing and Configuring the Server*
- *Installing the _tk240 Utility*
- *Setting Up ODBC Data Sources*
- *Defining Users*
- *Logging In*
- *Quick Start Guide Using _tk240*

2.1 Overview of Setting Up a Project on Your Server

When your organization chooses to run SAS Enterprise Content Categorization Studio in the collaborative mode, you as the database administrator, set up data sources on the server. You also add users and specify their permission on the database. This chapter is written primarily for database administrators, including those who might also function as a project administrator in a small company.

The supported database versions are:

- MySQL: Version 5.0 or higher
- Microsoft SQL Server: 2005, or newer
- Oracle: 10 or newer

Sections of this chapter also contain information about the tasks that end users perform. These users can set up their own data sources that access the project

running on the server. However, this chapter assumes that administrators often set up data sources for users.

For more information about the tasks that apply to both types of users, see Section 2.4 *Setting Up ODBC Data Sources* on page 18.

2.2 Installing and Configuring the Server

2.2.1 Install the Server

Install a SQL-based database server on your server in order to set up a repository for SAS Enterprise Content Categorization Studio. Both the SAS Enterprise Content Categorization Studio application and _tk240 support MySQL, Microsoft SQL Server, and Oracle.

If you are running Microsoft SQL Server, the ODBC drivers are pre-installed on Windows. However, if you want to connect to a Microsoft SQL Server database using the ODBC driver for UNIX and did not install this driver, a third-party solution is required.

The address to install MySQL is below:

<http://dev.mysql.com/downloads/mysql>

Make the following changes to the server configuration file to ensure that SAS Enterprise Content Categorization Studio runs smoothly. These changes are important if you plan to develop classifier concepts with 100,000, or more, entries.

2.2.2 Change the Size of the .cnf File for Linux

To change the size of data in the `my.cnf` file, complete these steps:

1. Copy the `my-xxx.cnf` file from `/usr/share/mysql` to `/etc` as `my.cnf`.

The size of this `.cnf` file can vary from small to large, depending on your requirements.

2. Locate the file. The location depends on the SQL installation. For example, see the following path:

```
cp /usr/share/sqlfilename /etc/my.cnf
```

or:

```
cp /usr/share/sqlfilename /etc/mysql/mysql.cnf
```

3. Change the following default setting in the `my.cnf` file:

```
max_allowed_packet = 1M
```

to:

```
max_allowed_packet = 32M
```

For classifier definitions of one million or more entries, the value is 96M.

4. Save the file.
5. Restart MySQL using the following command line:

```
/etc/init.d/mysql restart
```

2.3 Installing the _tk240 Utility

Before you can configure your ODBC data sources, install `_tk240`. The `_tk240` utility is a command line program that enables you to set up the collaborative mode. This program is shipped with SAS Enterprise Content Categorization Servers. Download this `.zip` file onto your computer and extract its files into the location of your choice.

Note: To specify the path to the SAS license, see Section A.7 *Set the Path to the SAS License* on page 89.

2.4 Setting Up ODBC Data Sources

2.4.1 Driver Types

Install and configure the ODBC driver connectivity software on the server and client systems that are used to access the project repositories.

SAS Enterprise Content Categorization Studio uses three types of ODBC data sources. The first two configurations are for the administrator's machine. The third configuration is for regular users.

- The first ODBC driver points to the default database and is used by the database administrator to create the SAS Enterprise Content Categorization Studio database on the server.
- The second ODBC driver points to the SAS Enterprise Content Categorization Studio database. Database administrators use this driver to add users to this database.
- Regular users can create an ODBC data source on a local machine that points to the SAS Enterprise Content Categorization Studio database.

For more information, see Section 2.4.3 *Configuring Data Sources* on page 19 and Section 2.7 *Quick Start Guide Using _tk240* on page 37.

2.4.2 Installing the ODBC Drivers

SAS Enterprise Content Categorization Studio and _tk240 use ODBC drivers to connect to database repositories.

Note: To install the ODBC software, make sure that your privileges are set to perform an installation.

You can download the MySQL ODBC drivers for either Windows or UNIX, by going to the following location:

<http://dev.mysql.com/downloads/connector/odbc>.

Microsoft SQL Server, Windows: preinstalled

Microsoft SQL Server, UNIX: third-party solution required

Oracle: Complete these steps:

1. Go to <http://www.oracle.com/technetwork/database/features/instant-client/index-097480.html>.
2. Download the Database Instant Client for your architecture.

2.4.3 Configuring Data Sources

2.4.3.A Credentials

Use the following credentials to configure each of the three data sources that are required to run SAS Enterprise Content Categorization Studio and _tk240. For more information, see:

- Section 2.4.3.C *Configure MySQL Data Sources* on page 22
- Section 2.4.3.D *Configure Microsoft SQL Server Data Source* on page 25
- Section 2.4.3.E *Download and Configure Oracle ODBC Drivers* on page 32

There are several credentials that supply the information required to create the three different types of data sources:

Name

User

Password

Database: You can either specify a name or use the default entry. For more information, see Table 2-1 on page 21.

Examples of these credentials are provided for each of the following three data sources that you create:

- Administrators with database permissions should see Step 1 on page 37 to initialize a SAS Enterprise Content Categorization Studio database on the server using these credentials.
- Administrators with database permissions should see Step 2 on page 37 to add users to the SAS Enterprise Content Categorization Studio database.

-
- Regular users see the example of credentials shown in Step 4 on page 38 to open a remote project after you create the third data source.

2.4.3.B Configuring ODBC Data Sources on UNIX

Administrators can use _tk240 to set up database connectivity. If you, as an administrator, choose to configure through UNIX, all ODBC data sources are configured using an `odbc.ini` file. To locate this file, see the path specified by the `ODBCINI` environment variable or see your home directory. If the `odbc.ini` file is stored in `/etc`, it is not writeable by regular users. To create a local `.odbc.ini` file, create a file with this name in your home directory. For more information about the file format, see the following link:

<http://www.iodbc.org/index.php?page=docs/faq#iodbc237887083>

You can also use the graphical data source administrative tool if this tool is installed on your system.

Use the sample `odbc.ini` file displayed below to either create or to modify your `odbc.ini` file:

Example 2-1: A Local odbc.ini File

```
;
; odbc.ini
;
[ODBC Data Sources]
data_source_name1 = ODBC Driver for MySQL
data_source_name2 = ODBC Driver for MySQL

[data_source_name1]
Driver           = /usr/lib/odbc/libmyodbc.so
Server           = server_name
Port             = 3306
Database         = database_name

[data_source_name2]
Driver           = /usr/lib/odbc/libmyodbc.so
Server           = server_name
Port             = 3306
Database         = _tk240_database_name
```

Note: The example above uses the `MySQL` database. If you are connecting to Microsoft SQL Server or Oracle, specify the appropriate name.

Use the second example to create your third data source as well as your second data source.

As shown in Example 2-1 above, you create each data source on a separate line that appears below `[ODBC Data Sources]`. This section of the `odbc.ini` file lists all of the ODBC data sources. Below this section, you specify each data source name followed by separate lines detailing the `Driver`, `Server`, `Port`, and `Database` information.

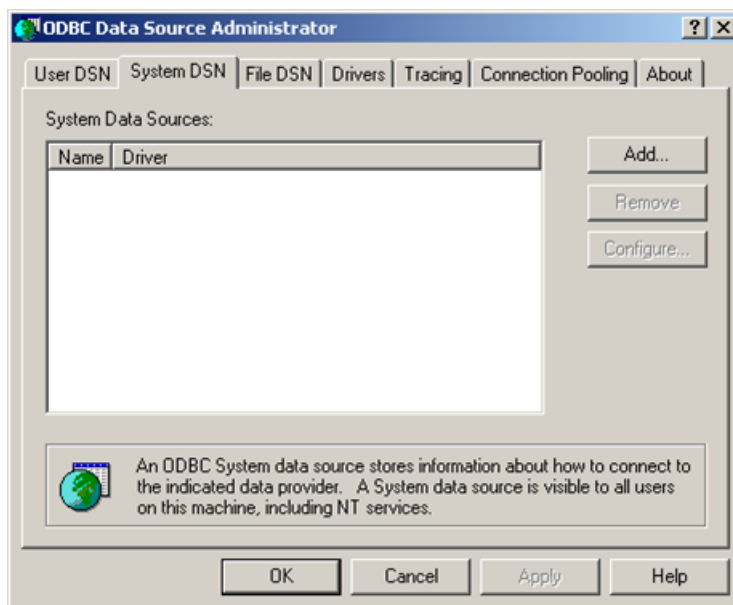
Table 2-1: Data Source Parameters

Parameter	Description
Driver	Specify the name and location of the driver library that is installed as part of the ODBC driver. For example, the name of this file might be <code>libmyodbc.so</code> .
Server	Specify the name of the machine where the database is running. Use <code>127.0.0.1</code> instead of <code>localhost</code> if you are using MySQL Server running on <code>localhost</code> .
Port	Specify the number of the port to connect to the database. MySQL Server default: 3306 SQL Server default: 1433 Oracle default: 5560 These default values apply unless your database is configured to use another port.
Database	Specify the name of the database to access using this data source.

2.4.3.C Configure MySQL Data Sources

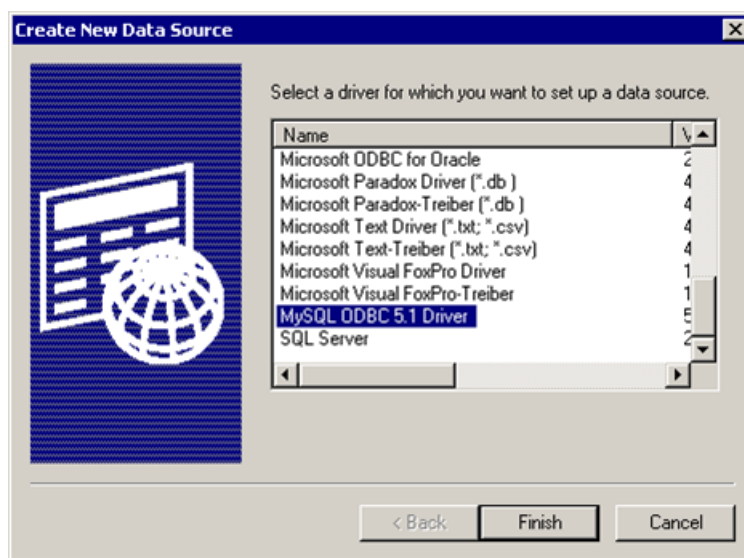
To configure data sources on a Windows machines, complete these steps:

1. Select **Control Panel --> Administrative Tools --> Data Sources (ODBC)** and the ODBC Data Source Administrator wizard appears.



2. (Administrators) Select the **System DSN** tab.
(Regular users) Select the **User DSN** tab.

-
3. Click **Add**. The Create New Data Source window appears.



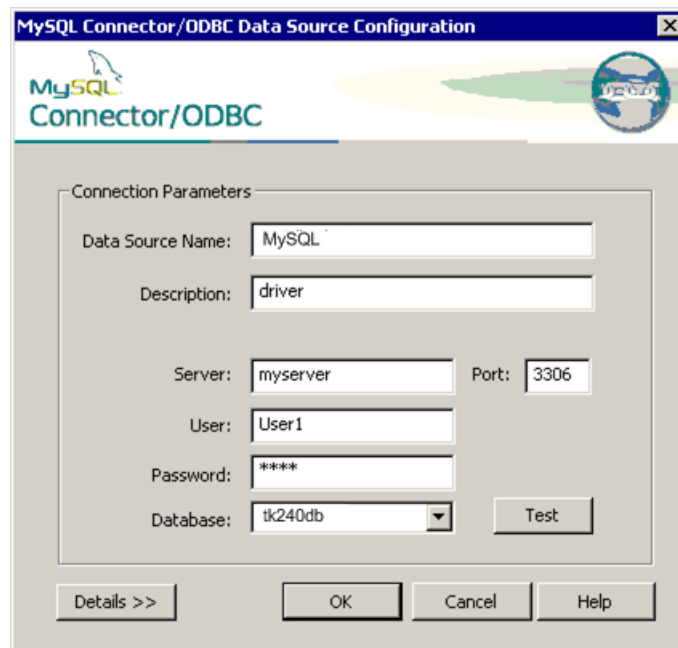
4. Use the scroll bar to locate the appropriate driver. If you select MySQL, see Step 5 below.

If you are using Microsoft SQL Server, see Step 5 on page 26.

If you are using Oracle, see Step 5 on page 33.

Note: Version 5.1 of the MySQL ODBC driver is used for example purposes only.

-
5. Click **Finish** and the MySQL Connector/ODBC page appears. Use the MySQL Connector/ODBC wizard to point the ODBC driver to the server:



6. Enter the name of the data source into the **Data Source Name** field. For example, administrators might specify `SASdbadmin` for the first database and `SASdb` for the second data source. Regular users can use the database name selected by the administrator. For example, they can enter `SASdb`.
7. (Optional) Enter the description into the **Description** field. For example, type `driver`.
8. Enter the server name into the **Server** field. For example, type `myserver`.
9. (Optional) The default entry, shown in the right pane of this interface, is entered into the **Port** field.
10. Enter the name of the person making the connection, into the **User** field. Administrators enter their administrative user name as entered on the

server. Regular users specify the user name entered for them by the administrator who added users to the database.

11. Enter your password into the **Password** field. The password for the administrator is the administrator's password on the server.
12. Enter the name of the database that you are connecting to in the **Database** field. The administrator specifies `mysql` the first time she or he uses this section. The second time, the administrator like all regular users who are using this section for the first time, specifies the database name. For example, you might type `tk240db`.
13. Click **Test**. If the connection is successful, the Connector/ODBC window appears with a message stating that the connection is successful.

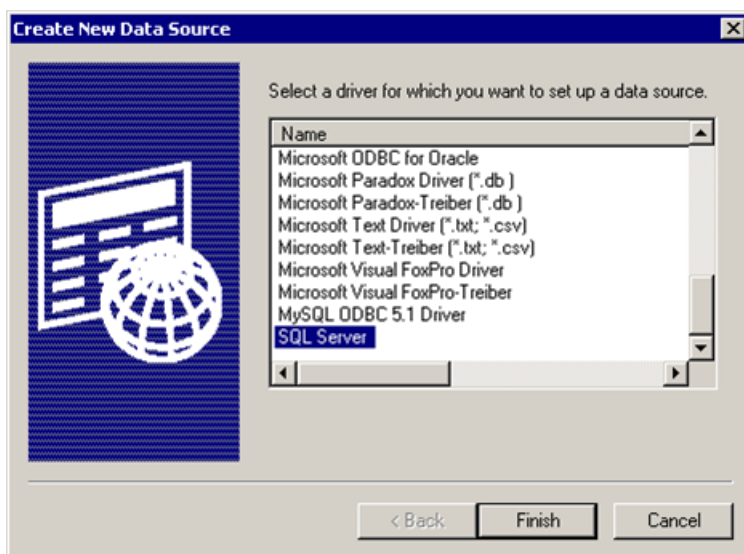


14. (Optional) Click **Details** in the MySQL Connector/ODBC window to set various types of flags.
15. Click **OK**.
8. Click **OK** in the MySQL Connector/ODBC window.

2.4.3.D Configure Microsoft SQL Server Data Source

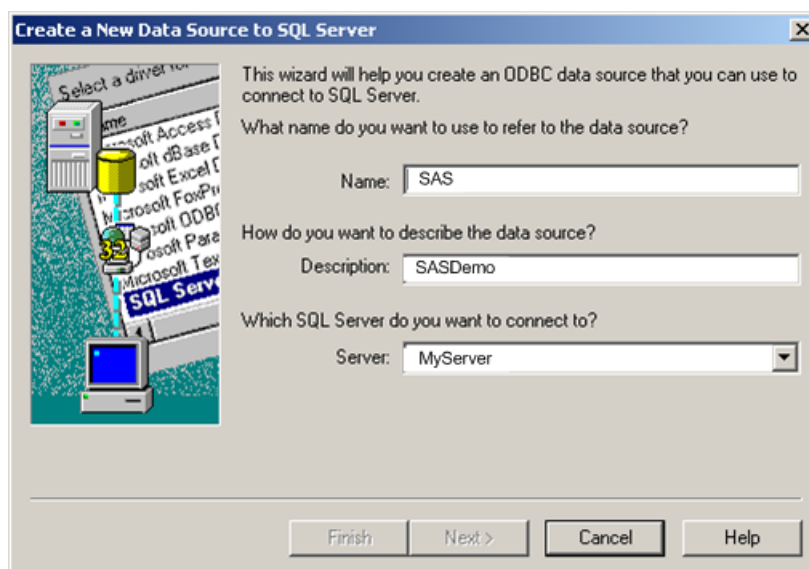
Use the following steps to configure the SAS Enterprise Content Categorization Studio database using Microsoft SQL Server. These steps continue from Step 4 on page 23:

-
5. (Optional) If you select `SQL Server` in the Create New Data Source page, click **Finish**.



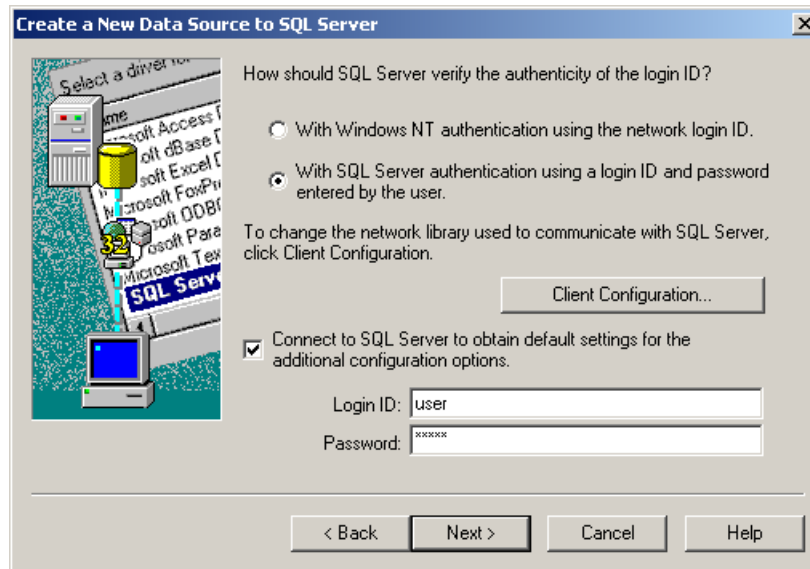
Note: SQL Native Client is also available on some machines. This SQL selection is also valid.

The first screen of the Create a New Data Source to SQL Server wizard appears.



6. Enter the name of your data source into the **Name** field. For example, type SAS.
7. (Optional) Enter the descriptive information for your database into the **Description** field. For example, type SASDemo.
8. Enter the name of the server into the **Server** field. For example, type MyServer.

-
9. Click **Next** and the next Create a New Data Source to SQL Server page appears.

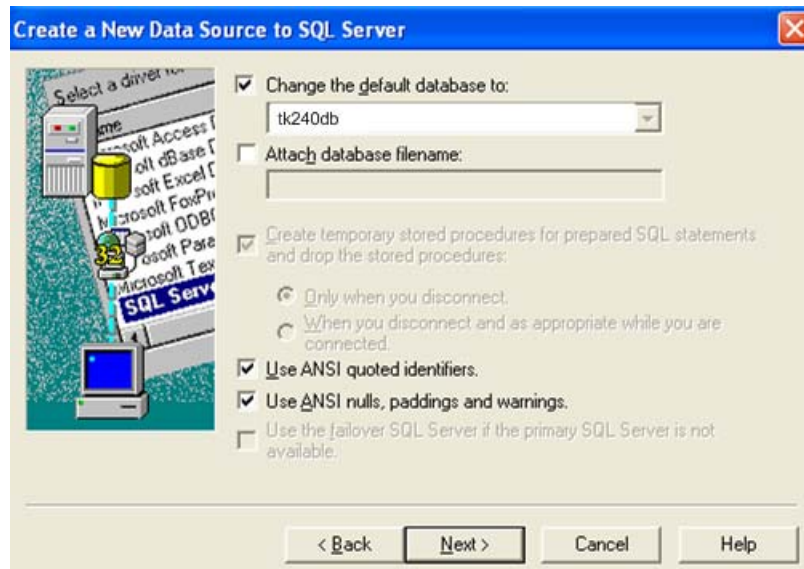


10. Select either **With Windows NT authentication using a login ID** or **With SQL Server authentication using a login ID and password entered by the user**.

Note: Step 11 and Step 12 apply only if SQL Server authentication is selected.

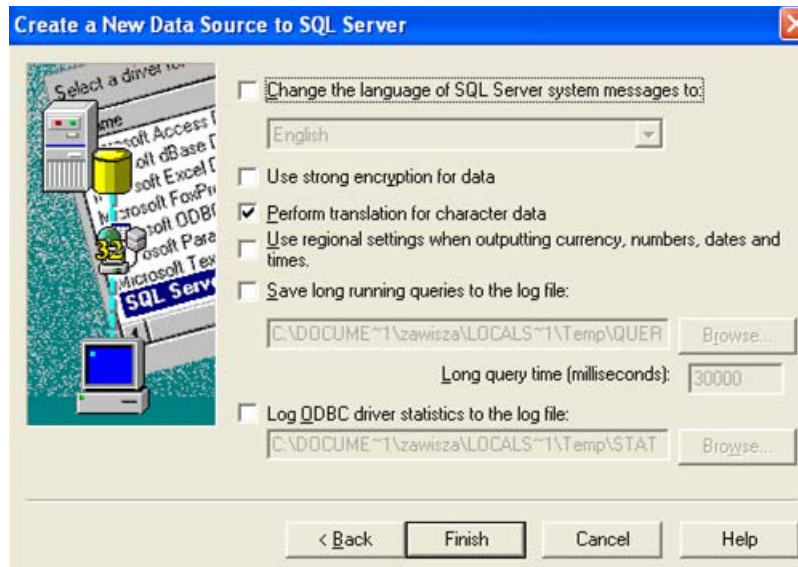
11. (Optional, but recommended) Enter your name into the **Login ID** field.
12. (Optional, but recommended) Enter your password into the **Password** field.

-
13. Click **Next** and the next Create a New Data Source to SQL Server page appears.



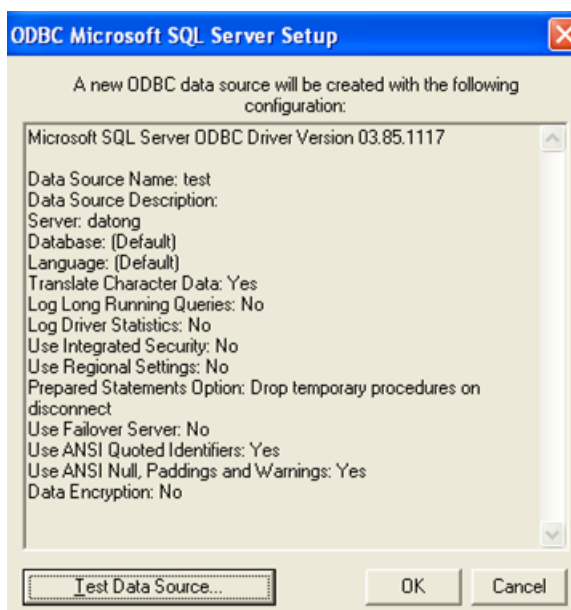
14. The administrator specifies `mysql` the first time the data source is created. The next time, the administrator like all regular users (who are using this section for the first time) specifies the database name. For example, type `tk240db`.
15. Select any of the other default settings.

-
16. Click **Next** and the next **Create a New Data Source to SQL Server** page appears.

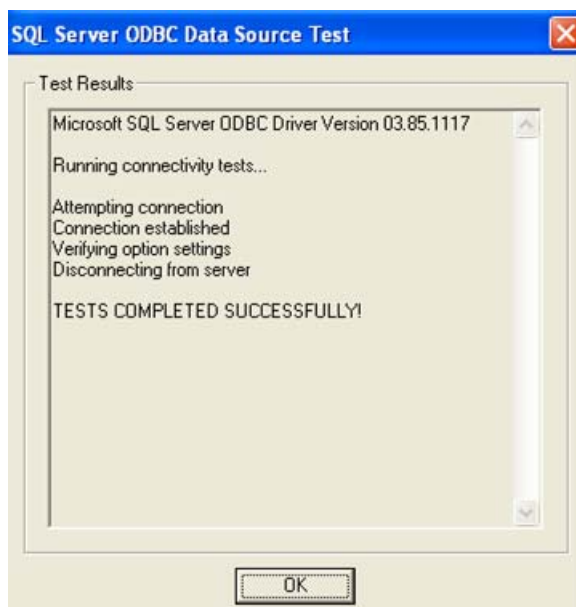


Any of the default settings are OK.

-
17. Click **Finish**. The ODBC Microsoft SQL Server Setup page appears, displaying the configuration that you defined for the new ODBC data source.



-
18. (Optional) Click **Test Data Source** and a window similar to the SQL Server ODBC Data Source Test window appears.



19. Click **OK**, if you see the SQL Server ODBC Data Source Test window.
20. Click **OK** in the ODBC Microsoft SQL Server Setup page.

2.4.3.E Download and Configure Oracle ODBC Drivers

You can download Oracle ODBC drivers to your machine.

To download an Oracle ODBC driver, complete these steps.

1. Go to the Oracle Web site and download the Instant Client package (direct link: <http://www.oracle.com/technetwork/database/features/instant-client/index-097480.html>).
2. Choose the desired architecture in the Web page that opens from this link.
3. Click Sign In/Register for Account and log in.
4. Accept the license agreement.

-
5. Download both the Instant Client – Basic and Instant Client – ODBC packages.

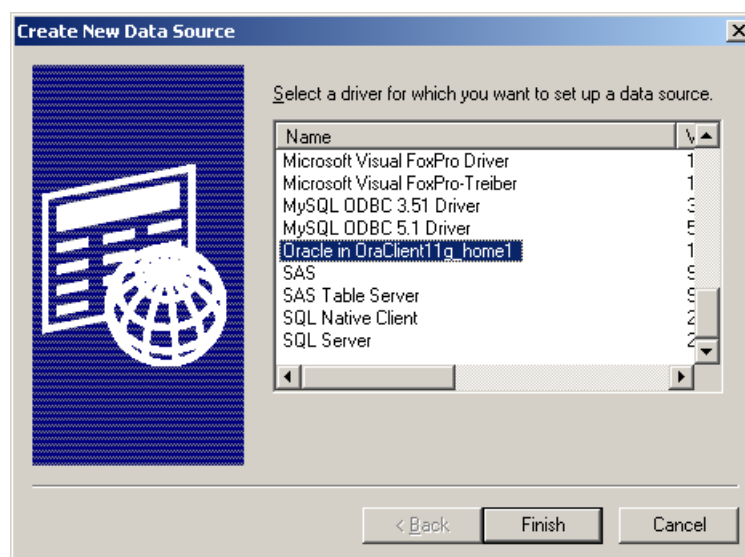
Note: Make sure that both client package files are located in the same directory.

6. Extract the client packages.
7. Navigate to the subdirectory where the packages were extracted.
8. Run `odbc_install`. (On Windows, run `odbc_install.exe`.)

ODBC support for Oracle is configured after you run the install command.

To configure your Oracle ODBC drivers to set up your data source, complete these steps. These steps continue from Step 5 on page 17:

5. Select the ODBC driver that you installed. The exact name of the driver depends on the version of Oracle Instant Client that you install.



6. Click **Finish**.

The Oracle ODBC Driver Configuration window appears.

Oracle ODBC Driver Configuration

Data Source Name: SASdb

Description: SASDemo

TNS Service Name: MyOracleServer

User ID: SASUser

Buttons: OK, Cancel, Help, Test Connection

Tabs: Application, Oracle, Workarounds, SQLServer Migration

Application Tab Options:

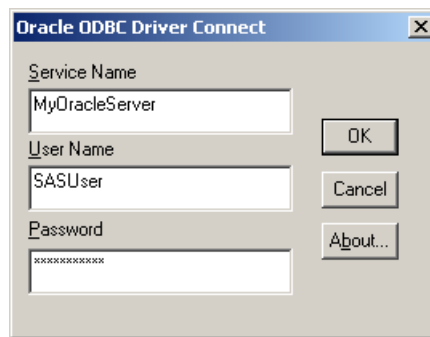
- Enable Result Sets ☒
- Enable Query Timeout ☒
- Read-Only Connection ☐
- Enable Closing Cursors ☐
- Enable Thread Safety ☒

Batch Autocommit Mode: Commit only if all statements succeed

Numeric Settings: Use Oracle NLS settings

7. Enter the name of the data source into the **Data Source Name** field. For example, type SASdb.
8. (Optional) Enter the information about your database into the **Description** field. For example, type SASDemo.
9. Enter the server name into the **TNS Service Name** field. For example, type MyOracleServer.
10. Enter the user name assigned by the database administrator into the **User ID** field. For example, type SASUser.

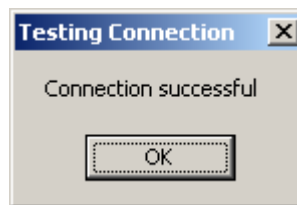
-
11. Click **OK**. The Oracle ODBC Driver Connect window appears.



12. The fields in this window are automatically filled in for you with the exception of the **Password** field.

Hints: The **Server Name** field contains the same name as the entry in the **TNS Service Name** field.
The **User Name** field contains the same name entered by the database administrator into the **User ID** field.
See Step 10 on page 24.

13. Click **OK**. If the connection is successful, the Testing Connection window appears with this message.



14. Click **OK**.
15. Click **OK** to close the Oracle ODBC Driver Connect window.

2.5 Defining Users

After the SAS databases are configured, but before you add user accounts to a new project, define administrative and regular users. For more information, see Section A.5.1 *Add Users* on page 84.

To make a new user a project administrator, assign them upload permissions for the project. For more information, see Section 4.5 *Setting Permission Levels in a Project* on page 62.

Database administrators can act as project administrators:

- Access every project on the server.
- Project administrators can also upload projects to the server.
- Add, change, and delete users in the database.

After regular users are added to a project by an administrator, their permission levels for the project are set using the SAS Enterprise Content Categorization Studio interface. This permission level setting controls their access to the project. By default, a regular user has Read-Only permissions.

2.6 Logging In

After the connections to the remote server are set up from the local host, a user can log in from any host that they are set up to use. Three pieces of information are required to log in to the remote repository:

- the data source name
- your user name
- your password

2.7 Quick Start Guide Using _tk240

Administrators use _tk240 to enable the collaborative mode of SAS Enterprise Content Categorization Studio.

Note: Replace the code that is enclosed in angled brackets (<>) with values from your system.

To use _tk240, complete these steps:

1. You, as the database administrator, initialize the SAS Enterprise Content Categorization Studio database using the following command line example:

```
_tk240 -initialize_database  
      -driver <datasource_1>  
      -user <admin_username>  
      -pass <admin_password>  
      -verbose
```

2. You, as the database administrator, add one or more project administrators to the project:

```
_tk240 -add_user  
      -driver <datasource_2>  
      -user <admin_username>  
      -pass <admin_password>  
      -username <pam>  
      -password <pam34>  
      -regular  
      -can_upload  
      -verbose
```

Notes: If you use SQL server with Windows Authentication, replace -user and -pass with -windows_authentication in this step and all of the following steps. Replace with -login_name <DOMAIN\username> in Step 2 above and Step 4 on page 38.

The user name of a new project administrator, who has upload privileges, is specified using the <username> parameter. For example, pam has upload permission in this example. This user is required to use

the specified password. For example, specify `pam34`, to access this project. For more information, see Section A.5.1 *Add Users* on page 84.

3. (Optional) Continue using Step 2 above, reiteratively, until all of the project administrators are added to your project.
4. You, as the database administrator, add regular users to the project. For more information, see Section A.5.1 *Add Users* on page 84.

```
_tk240 -add_user -driver <datasource_2>
-user <admin_username>
-pass <admin_password>
-username <tom>
-password <tom22>
-regular
-verbose
```

The user name of the new regular user is specified using the `<username>` parameter. For example, specify `tom`. This user is required to use a password to access the project because his password is specified using the `<pwd>` parameter. In this example, `tom` uses the password `tom22` to access the project.

Note: You cannot use a period (`.`), an at sign (`@`), or an e-mail address when you create a user name for MySQL Server.

This user can access a project on the server as a regular user and download that project. This user cannot upload a project to the server, because upload permissions are not specified. If upload permissions are specified, this regular user becomes a project administrator.

5. (Optional) Continue using Step 4 above, reiteratively, until you have added all of the regular users to your project.

3

Using the Interface

- *Collaboration and the User Interface*
- *The Menu Bar*
- *The Server Operation Windows*
- *The Options Window: Shared Projects Folder*
- *The Miscellaneous Windows*

3.1 Collaboration and the User Interface

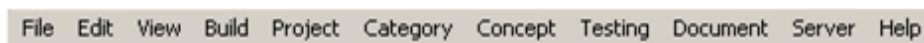
Collaborative operations are enabled by a database server. For this reason, some of the operations that are visible in the SAS Enterprise Content Categorization Studio user interface are not accessible until you set up the server. The collaborative operations enable you as an administrator to upload files and to set user permissions.

This chapter describes the operations that are available for an administrative user. For information about the collaborative operations that are available for regular users, see the *SAS Enterprise Content Categorization Studio: User's Guide*.

3.2 The Menu Bar

Many of the collaborative operations are located in the menu bar. Some of the menu commands are also available on the standard toolbar, and a few can also be accessed by right-clicking on the nodes in the Taxonomy window.

Display 3-1: Menu Bar



This is the command that is available only for project administrators in the menus:

File --> Upload Project to Server

Place a copy of the project onto the server. This operation can be performed only if you are a project administrator. When you select this operation, the Upload to Server window appears. For more information, see Section 3.5.5 *The Upload to Server Window* on page 52.

3.3 The Server Operation Windows

3.3.1 Accessing Server Permissions

Server permissions enable you to set the level of access for each user.

Display 3-2: Server Operations



When you right-click on a taxonomy node, a menu appears. The **Server Permissions** operation is available from individual nodes in the taxonomy.

Hint: You can also access **Local Permission**, and **Project Permissions** in the **Server** menu.

3.3.2 Access and Use the Project Server Permissions Window

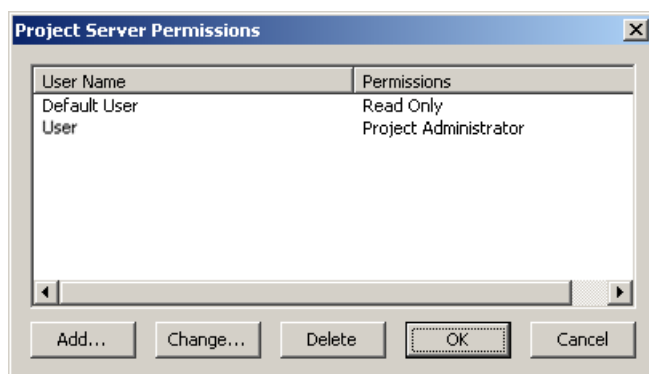
Specify, and see a list of users and their permission levels for your entire project using the Project Server Permissions window.

To access the Project Server Permissions window, complete these steps:

1. Create a collaborative project. For more information, see Section 4.2 *Create a Project on a Local Machine* on page 57 and Section 4.3 *Upload a Project* on page 58.
2. Right-click on the project name node. For example, right-click on SAS_Demo. Select **Server Permissions** from the drop-down menu that appears.



The Project Server Permissions window appears.



3. Use the components of the Project Server Permissions window as specified below:

User Name

See a list of all of the project users for this project.

Permissions

See the permission level assigned to each user.

Add

Assign a user with his or her own permission level. The Set User Permission window appears. For more information, see Section 3.3.4 *The Set User Permission Windows* on page 45.

Change

Click after you select a user in the Project Server Permissions window. The Set User Permission window appears. For more information, see Section 3.3.4 *The Set User Permission Windows* on page 45.

Delete

Click after you select a user. A SAS Content Categorization Studio confirmation window appears that asks you to confirm your decision to remove the permission setting for the selected user.

3.3.3 Use the Set Server Permissions for a Selected Node Window

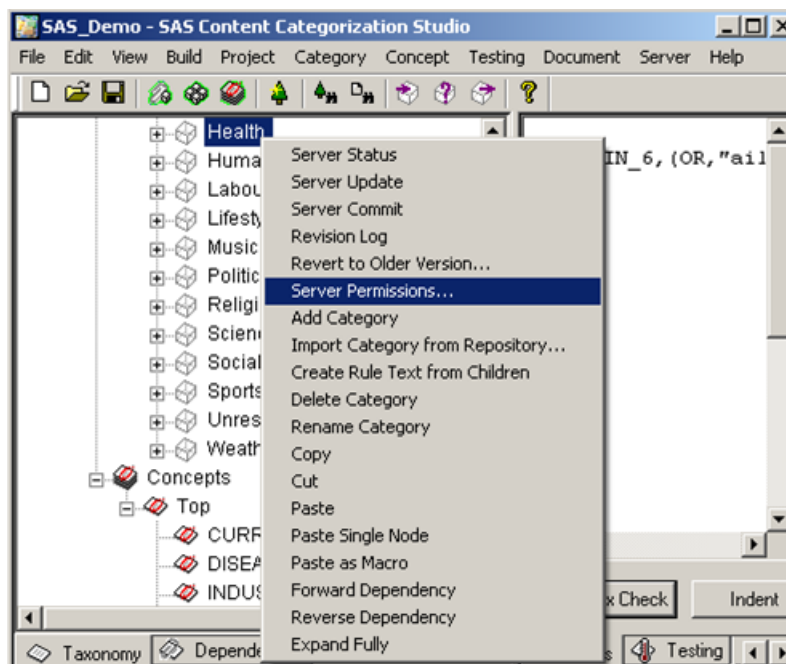
Use the Server Permissions for (the selected category or concept node) window to specify users and to set their access levels for the selected category or concept. The permission levels that you specify apply to the selected node only and override the access level that you set for the project.

For example, you might specify that `User1` has `Read` and `Write` privileges for the project. You might also specify a `Read-Only` permission setting for the `Health` node. In this case, `User1` can read the rule for the `Health` category, but cannot make any changes to this rule.

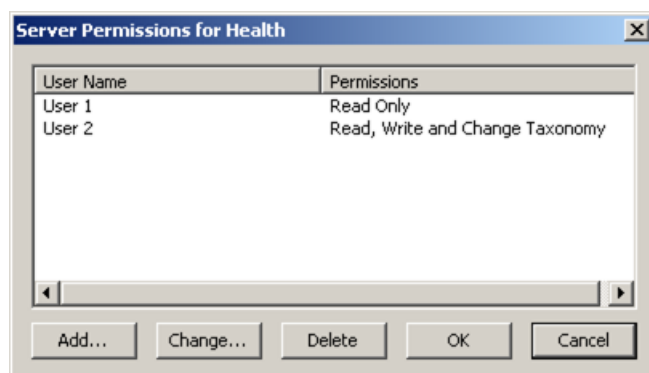
To access the Server Permissions for a selected node window, complete these steps:

1. Create a collaborative project. For more information, see Section 4.2 *Create a Project on a Local Machine* on page 57 and Section 4.3 *Upload a Project* on page 58.

2. Right-click on a taxonomy node. For example, right-click the `Health` category node. Select **Server Permissions** from the drop-down menu that appears.



The Server Permissions for (the selected category or concept name) window appears.



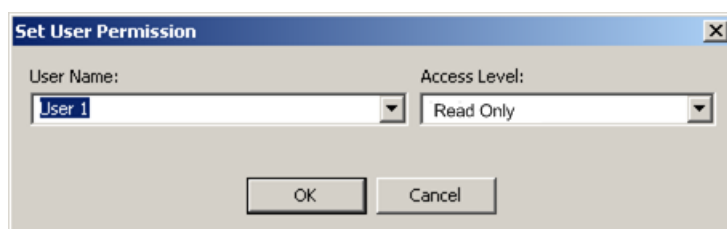
-
3. Use the various components of this window to set user permissions for the selected nodes. For more information, see Section 3.3.2 *Access and Use the Project Server Permissions Window* on page 41.

3.3.4 The Set User Permission Windows

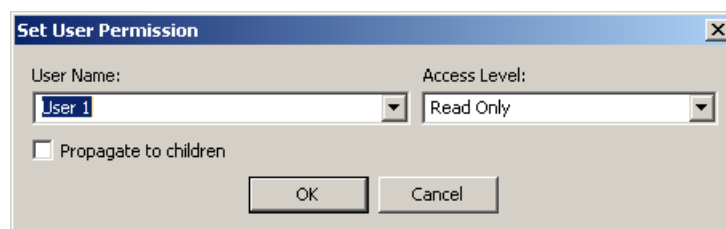
Only an administrator can change user permissions from the default setting of `Read Only` that is set at the command line when each user is added to the project.

There are two windows for this operation:

Display 3-3: Set User Permissions at the Project Level



Display 3-4: Set User Permissions at the Category or Concept Node Level



To use a Set User Permission window, complete these steps:

1. Click ▼ to the right of the **User Name** field to select a name from the drop-down list of users added to the project. For more information, see Section A.5.1 *Add Users* on page 84.
2. Click ▼ to the right of the **Access Level** heading to specify the user's new permission level:

-
- No Access (project level, only): This user cannot access this project.
 - Read Only: Read, but do not make any changes to the project.
 - Read and Write Rules: Read the project and write rules and definitions.
 - Read, Write, and Change Taxonomy: Develop any part of this project, including the taxonomy.
 - Project Administrator (project level, only): This user has full permissions to access this project, which includes adding, deleting, and renaming nodes in the taxonomy. The project administrator can also set permission levels for the project users.
3. (Optional, and only available at the individual node level) Select **Propagate to children**. When you make this selection, the user has the specified permission level for the child nodes.
 4. Click **OK** to save these settings.

For more information, see Section 4.5 *Setting Permission Levels in a Project* on page 62.

3.4 The Options Window: Shared Projects Folder

For the purpose of improving performance, collaborative projects are cached locally on your computer in the Shared Projects folder. The default path is shown below:

```
Shared Projects\<repository_name>\<user_name>\  
    <project_name>
```

To change this location, deselect the **Use separate Shared Projects folder for each user** check box in the Options window. If you deselect this check box, the default path is shown below:

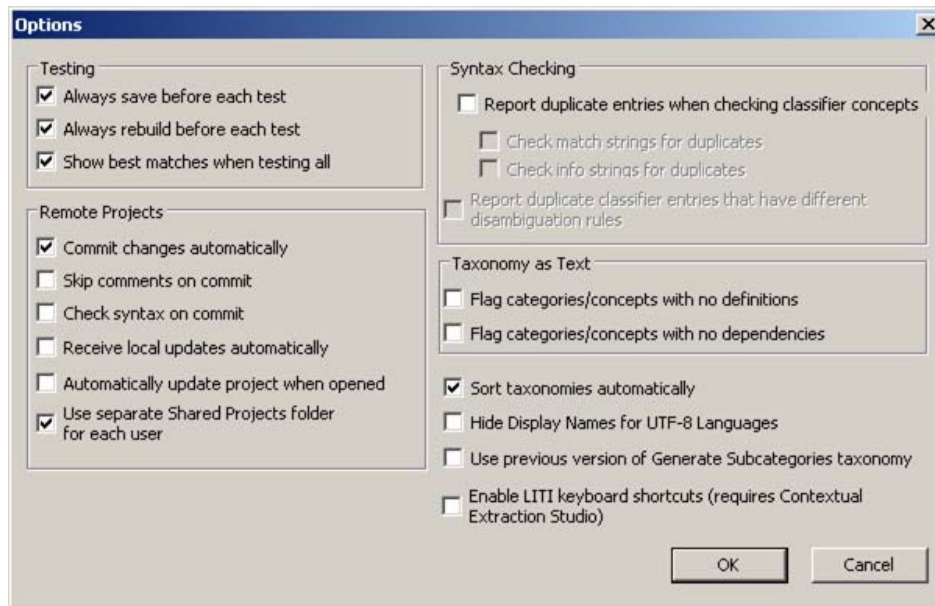
```
Shared Projects\<project_name>
```

Note: SAS recommends that you keep the **Use separate Shared Projects folder for each user** check box selected.

Use the Options window to save a local copy of a project that you want to delete on the server. To create a local copy for each user on a machine, select the **Use separate Shared Projects folder for each user** operation in the Options window.

To access and use the Options window, complete these steps:

1. Select **Edit --> Options**.



2. Leave the default selection **Use separate Shared Projects folder for each user** selected.
3. Click **OK** to close this window and to save any changes that you made.

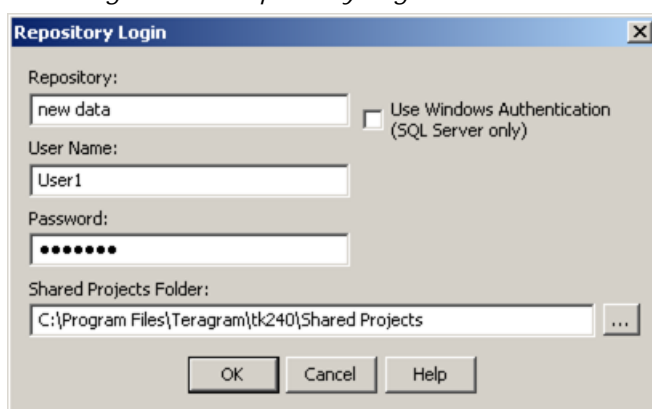
The local copy is stored in the location specified in the **Shared projects folder** field of the Repository Login window. For more information, see

3.5 The Miscellaneous Windows

3.5.1 The Repository Login Window

Log in to the server before you work on a collaborative project. If you are not working in a collaborative project and you select any server operations from the **File** menu, the Repository Login window automatically appears:


Figure 3-5: Repository Login Window



To use the Repository Login window, complete these steps:

1. Enter the name of the ODBC data source into the **Repository** field. For example, type `new data`.
2. (Optional) Select **User Windows Authentication (SQL Server only)** to use the Windows user account principal token to connect. If you select this check box, the **User Name** and **Password** fields are grayed.
3. Enter your user name into the **User Name** field.

Hint: If you previously accessed this server, the **Repository**, **User Name**, and **Shared Projects Folder** fields are all filled in.

4. Enter your password into the **Password** field.
5. (Optional) Click  under the **Shared Projects Folder** heading and the Select a Directory window appears. Use this window to select a location for the shared projects folder. For more information, see Section 3.5.2 *The Select a Directory Window* below.

Note: The default path for the Shared Projects Folder varies depending on the version of Windows. For Windows XP and Server 2003, this path is located in the Program Files directory hierarchy. For all other versions of Windows, this path is located in your user directory.

6. Click **OK** to close this window.


3.5.2 The Select a Directory Window

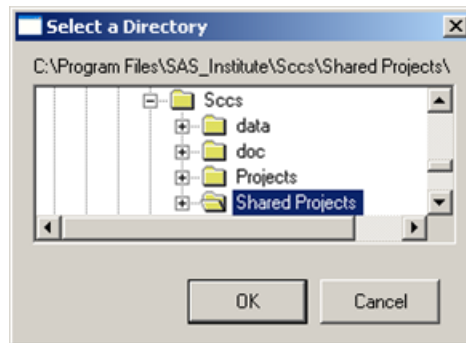
Use the Select a Directory window to locate the Shared Projects folder.

Note: SAS recommends that you do not modify the default Shared Projects specification.

To access and use the Select a Directory window, complete these steps:

1. Select **File --> Open Remote Project**. The **Repository Login** window appears. See Figure 3-5 on page 48.
2. Use Step 1 on page 48 to Step 4 on page 49.

-
- (Optional) Click  under the **Shared Projects Folder** heading and the Select a Directory window appears.



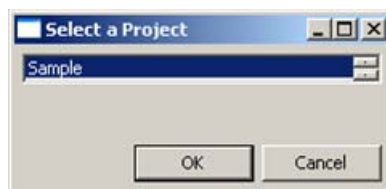
- Select the SharedProjects folder or a cached project in another location on your machine.
- Click **OK**.

3.5.3 The Select a Project Window

Use the Select a Project window to locate a project on the server that you want to access or delete.

To access and use the Select a Project window to access a project, complete these steps:

- Select **File --> Open Remote Project**. The Select a Project window appears.



- Highlight the project in the Select a Project window that you want to access. For example, select `Sample`.

-
3. Click **OK**. This project appears in the user interface and the Select a Project window closes.

To remove a project, complete these steps:

1. Select **File --> Remove Project From Server**.
2. Use Step 2 through Step 3 above. A SAS Content Categorization Studio confirmation window appears.



Note: The project that you want to remove from the server cannot be running on your machine during the removal process.

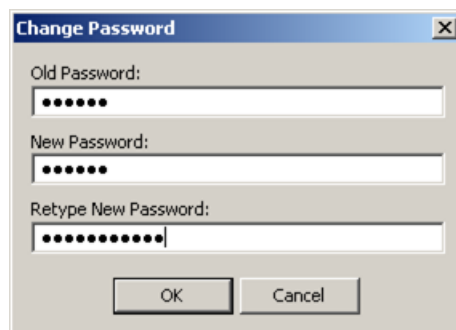
3. Click **Yes** to remove this project.

3.5.4 The Change Password Window

Use the Change Password window to specify a new password.

To access and use the Change Password window, complete these steps:

-
1. Select **File --> Change Repository Password** and the Change Password window appears.

A dialog box titled "Change Password" with a close button (X) in the top right corner. It contains three text input fields: "Old Password:" with six dots, "New Password:" with six dots, and "Retype New Password:" with ten dots. At the bottom are "OK" and "Cancel" buttons.

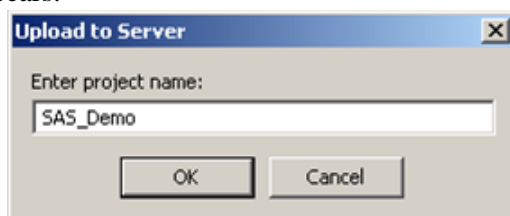
2. Enter your password into the **Old Password** field.
3. Enter your new password into the **New Password** field.
4. Re-enter the new password into the **Retype New Password** field.
5. Click **OK**.

3.5.5 The Upload to Server Window

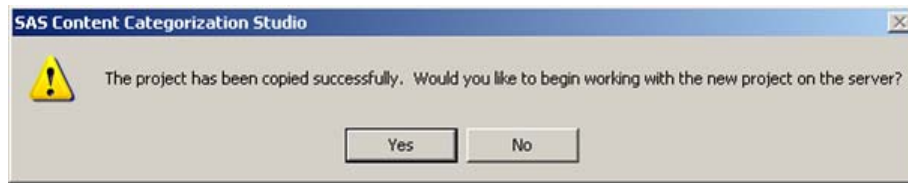
Upload operations are restricted to project or database administrators. Use the Upload to Server window to upload a project from a local workstation to the server. After the project is created on a local machine and uploaded to the server, this project is accessible to other users.

To upload a project, complete these steps:

1. Create a project on your local machine. For more information, see the *SAS Content Categorization Studio: User's Guide*.
2. Select **File --> Upload Project to Server** and the Upload to Server window appears.

A dialog box titled "Upload to Server" with a close button (X) in the top right corner. It contains a text input field labeled "Enter project name:" with the text "SAS_Demo" entered. At the bottom are "OK" and "Cancel" buttons.

-
3. Enter the name of the project that you are uploading to the server into the **Enter project name** field. For example, type `SAS_Demo`.
 4. Click **OK**. A SAS Content Categorization Studio confirmation window appears.



5. Click **Yes**.

4

Administering a Collaborative Project

- *Overview of Administering a Collaborative Project*
- *Create a Project on a Local Machine*
- *Upload a Project*
- *Access a Remote Project*
- *Setting Permission Levels in a Project*
- *Remove a Project from the Server*

4.1 Overview of Administering a Collaborative Project

You, as the project or database administrator, set up a project that is developed by subject matter experts working together. Collaboration between project developers is possible only when a project resides on a server with cached versions of the project on one or more workstations. Working at their local machines, regular users can collaborate on one project.

In order to administer a collaborative project, you grant the permission levels that are necessary to access the project at its various levels. After SAS databases are configured, but before you add user accounts to a new project, define administrative and regular users.

Only administrators can upload a project to the server, or remove a project from the server, and set permissions. To make a regular user a project administrator, assign this person upload permissions. For more information, see Section 4.5 *Setting Permission Levels in a Project* on page 62.

For information about using many of the same collaborative operations with the `_tk240` command line see Appendix A: *Using the _tk240 Utility*.

The collaborative operations for SAS Enterprise Content Categorization Studio also enable an administrator to specify permission levels for each of the subject matter experts working on the taxonomy. This is true whether these experts are defining categories and concepts, or writing rules and definitions. These permission, or access, levels can be applied across the taxonomy as well as to specific categories and concepts. For example, all taxonomists might be granted Read permissions for the entire taxonomy. Some taxonomists might be assigned Write permissions for the categories and concepts representing their area of subject matter expertise.

The operations to set up the appropriate data sources, run `_tk240`, initialize the database, and add users are performed before projects can be uploaded to the server. For more information, see Chapter 2: *Setting Up a Project on Your Server*.

A copy of the project is automatically saved to your local machine. This project is stored in the *Shared Projects* folder that is automatically created by SAS Enterprise Content Categorization Studio. When you log in to the repository, you access your local version of the selected project.

Warning: Not all of the new features for the 12.1 release, such as the terms generated by automatic rule generation, and synonym lists. You can test, but not upload an Excel file to the collaborative server.

4.2 Create a Project on a Local Machine

You, as an administrator, upload a project to a server where the project becomes a collaborative project. For more information about creating a local project, see *SAS Content Categorization Studio: User's Guide*.

A user with administrative permissions creates the initial project, because both permission setting and the capability to upload projects to the server are limited to this user type. However, the initial project does not require more than the project name node.

Display 4-1: New Project



After you upload a project, you change this project from a regular to a collaborative project.

To create a collaborative project, complete these steps:

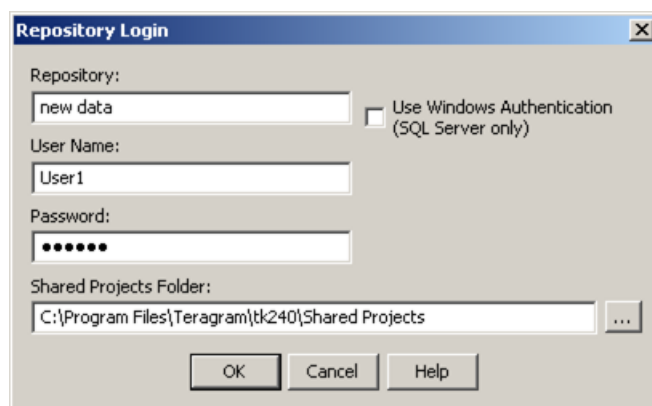
1. **Upload a project to the server:** Access by multiple users cannot begin until you upload a project to the server. For more information, see Section 4.3 *Upload a Project* below.
2. **Open the remote project:** An administrator accesses the remote project on the server before she or he can set the permission levels for other users. For more information, see Section 4.4 *Access a Remote Project* on page 60.
3. **Set permission levels at the project level:** Initially all non-administrative users have Read-Only permissions that prevent them from performing any project development tasks. For more information about permission levels set on the server, see Section A.5.1 *Add Users* on page 84. Also see Section 4.5 *Setting Permission Levels in a Project* on page 62.

4.3 Upload a Project

After a project is created, you, the database or project administrator, upload the project to the server where collaborative work can begin. The directions in this section apply to the user interface. For more information about the upload process using _tk240, see Section A.5.6 *Upload a Project Using _tk240* on page 86.

To upload a project to the server, complete these steps:

1. Access a project.
2. Select **File --> Upload Project to Server** and the Repository Login window appears:



The screenshot shows a 'Repository Login' dialog box with the following fields and controls:

- Repository:** A text field containing 'new data'.
- User Name:** A text field containing 'User1'.
- Password:** A text field with masked characters (dots).
- Shared Projects Folder:** A text field containing 'C:\Program Files\Teragram\tk240\Shared Projects' with a browse button (three dots) to its right.
- Use Windows Authentication (SQL Server only):** An unchecked checkbox.
- Buttons:** 'OK', 'Cancel', and 'Help' buttons at the bottom.

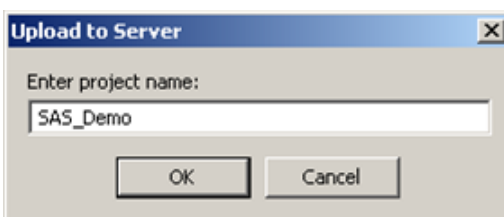
Hint: If you have already accessed the server, the Upload to Server window appears instead. For more information, see Step 7 and Section 3.5.5 *The Upload to Server Window* on page 52.

3. Enter the name of the ODBC data source into the **Repository** field.
4. (Optional) Select **User Windows Authentication (SQL Server only)** to use the Windows user account principal token to connect. If you select this check box, the **User Name** and **Password** fields are grayed.

-
5. Enter your user name into the **User Name** field.
 6. Enter your password into the **Password** field.
-

Hint: The **Shared Project Folder** field is not applicable until you access an uploaded project.

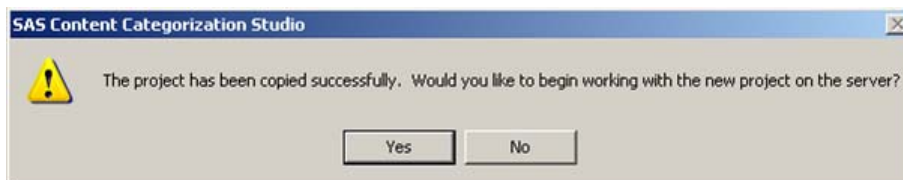
7. Click **OK** and the Upload to Server window appears:



The name of the collaborative project that you are working on is automatically entered for you. For example, see `SAS_Demo` in the **Enter Project Name** field. You can also choose to rename the project by entering a new name into the **Enter project name** field.

The project on the server can have the same name as the local project that you upload, or you can choose a different name. However, two or more collaborative projects cannot have the same name if they are both located on the same server.

8. Click **OK** and a SAS Content Categorization Studio confirmation window appears:



9. Click **Yes** to begin working in this project.

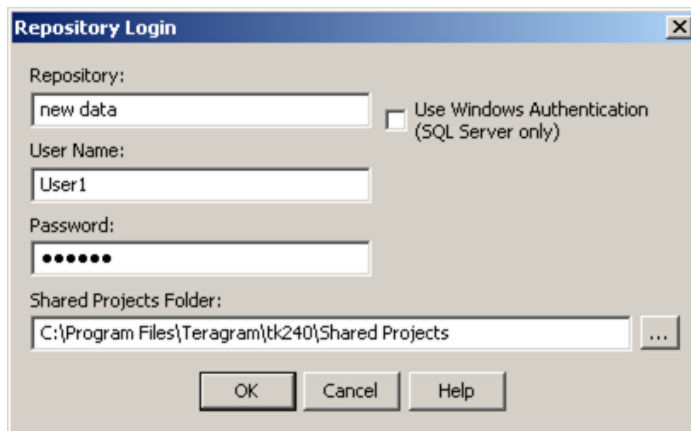
Hint: For large projects, this operation might take several minutes to complete.

4.4 Access a Remote Project

Before you can set user permissions, or work on a project, access a remote project.

To access a remote project, complete these steps:

1. Go to **File --> Open Remote Project** and the Repository Login window appears:

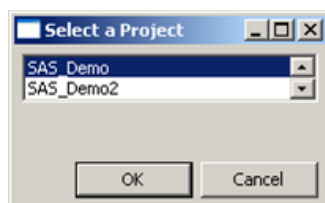


2. Enter the name of the ODBC data source into the **Repository** field.
3. (For Microsoft SQL Server only) Select **User Windows Authentication (SQL Server only)** to use the Windows user account principal token to connect. If you select this check box, the **User Name** and **Password** fields are grayed.
4. Enter your user name into the **User Name** field. If you want to log in using a different user name, enter your new user name as specified by the administrator. For more information, see Section A.5.1 *Add Users* on page 84.

-
5. Enter the password assigned to you in to the **Password** field.

Hints: If you previously accessed the server, the **Repository**, **User Name**, and **Shared Projects Folder** fields are all filled in.
In *Windows Vista* and *Windows 7* a preference is set so that data is not stored in the program files directory hierarchy.

6. The **Shared Projects Folder** field automatically displays the path to the Shared Projects folder after you access a project on the server. This folder stores the local, cached versions of the projects that you develop. For more information, see Section 3.5.2 *The Select a Directory Window* on page 49.
7. Click **OK**. The Select a Project window appears:



8. Select the project that you want to access. For example, select SAS_Demo.

Note: Select a project in order to begin working in the user interface.

9. Click **OK**.

SAS Enterprise Content Categorization Studio connects to the database using ODBC drivers.

4.5 Setting Permission Levels in a Project

4.5.1 Understanding Permissions

Administrators, only, can use the permissions windows to specify access levels for each project developer. By default, each regular user who is added to the project has Read-Only permissions until this permission level is changed by an administrator. For more information, see Section A.5.1 *Add Users* on page 84.

Each user, with the exception of an administrator, can have different permission levels specified for individual categories and concepts. These permissions can also differ from the project permission level for this user.

For example, you could specify Read and Write Rules at the project level for User1, but Read-Only permissions for the Education category. This user could write rules for every category and concept in this taxonomy, with the exception of the Education category.

4.5.2 Overview of Project Permission Levels

After users are added to a project, they are collectively known as regular users. The default permission setting for all users is Read Only. This setting enables users to see, but not to make any changes to, a project.

Server permissions for each of the users that are developing a project can be specified at both the project and individual node levels. The project administrator can specify a general access level for the entire project. This administrator can also choose to set different permission settings for one or more of the taxonomy nodes.

For example, if a user is granted Read, Write and Change Taxonomy project-level permissions, this user can read, and write rules for any node in the taxonomy. This user can also change the taxonomy by adding or deleting categories and concepts. However, if the permission level for the same user is restricted to Read Only for one category, this user cannot write the rule. For more information about the various levels of permission setting, see Section 4.5.3 *The Permission Levels* below.

4.5.3 The Permission Levels

This section provides an overview of all six levels of permission setting. Permission levels that are set at individual category and concept levels override the permission levels set at the project level. However, this statement is true only for the category or concept where these permissions are reset. Use the information in the table below to understand these settings before applying them in your project.

Table 4-1: Permission Levels for Collaborative Users

Setting	Description	Permission Level is set at:		
		Local or Individual Node	Whole Project	Server Level
Application users				
No Access	This user has no access to this project.		X	
Read Only	(Default) This user can read, but cannot make any changes to, the project.	X	X	
Read and Write Rules	This user can read this project and write category rules and concept definitions.	X	X	
Read, Write, and Change Taxonomy	This user can read the project, write rules or definitions for categories or concepts, and change the taxonomy by renaming, adding, and deleting categories or concepts.	X	X	
Project Administrators				
The project administrator has full project permissions and can add or delete users to the project.			X	X
Data Administrators				
The database administrator has access to all of the projects on the server In addition, the database administrator can add and delete users at the server level. (In some organizations, this administrator can also function as the project administrator.)				X
Note: Different users, with the exception of administrators, can have different permission levels for every project.				

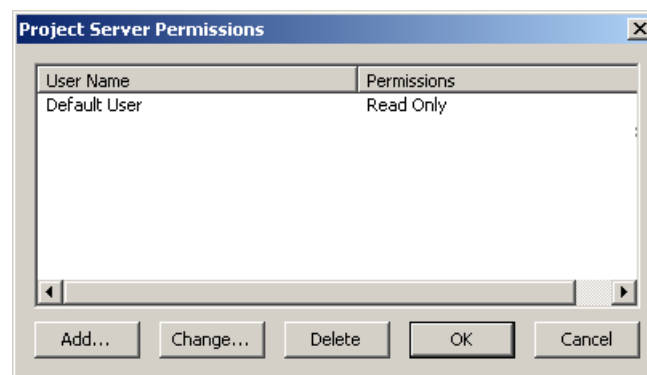
4.5.4 Setting, Changing, and Deleting Permission Levels

4.5.4.A Set Permissions at the Project Level

Use the Project Server Permissions window to specify new permission levels for individual users. By default, all new users are part of the `Default User` class with the permission level `Read Only`.

To add users to the project and change their permission level, complete these steps:

1. Go to **Server --> Project Permissions** and the Project Server Permissions window appears.



By default, the Project Server Permissions window initially lists only the `Default User` class. This group of users has the default permission setting of `Read Only`. These permissions are reset, one user at a time, in order for the members of this group to develop this project.

2. Select a user. For example, choose `Default User`.

-
3. Click **Add** and the Set User Permission window appears:



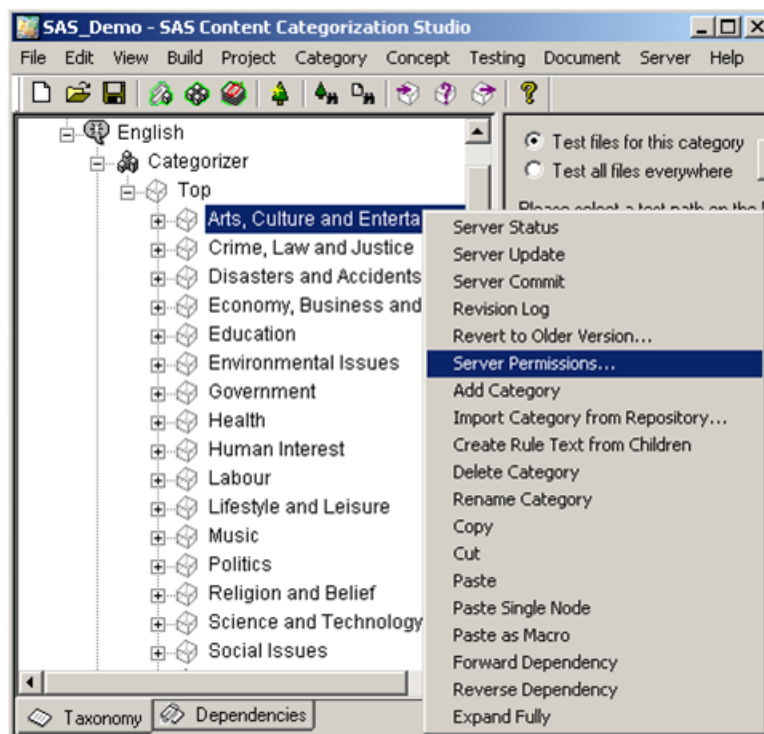
4. (Optional) Click ▼ to the right of the **User Name** field and select another user. By default, the name of one of the default users is displayed in the **User Name** field. For example, choose **User 1**.
5. Click ▼ to the right of the **Access Level** field and select another permission level. For example, choose **No Access**.
6. Click **OK** and the Set User Permissions window closes.
7. Click **OK** and the Project Server Permissions window closes.

4.5.4.B Set Permissions for Individual Nodes

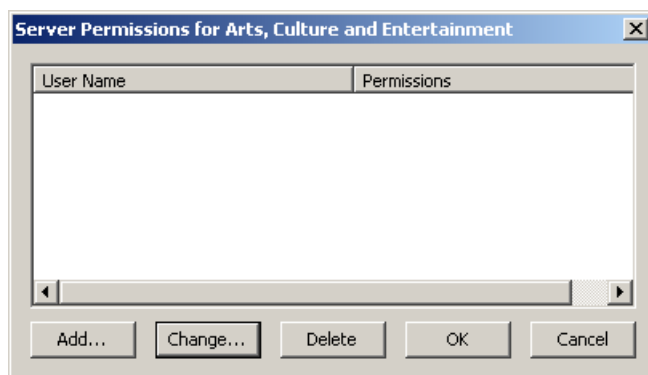
Specify permissions at the individual category or concept level to override the project level setting. You can also assign child nodes the same permission levels as their parent nodes. This process is similar to setting access levels to the project.

To set permission levels for a parent and its children, complete these steps:

1. Right-click on a node and select **Server Permissions** from the drop-down menu that appears:



The Server Permissions for (selected node name) window appears.



2. Click **Add** and the Set User Permission window appears.



3. (Optional) Click **Propagate to children**.
4. Follow Step 5 through Step 7 on page 65.

4.5.4.C Changing Project Level Permissions

You can change a user's permission level by using Step 1 on page 64 through Step 7 on page 65. The exception is Step 3 on page 65 where you select **Change** instead of **Add**.

4.5.4.D Delete Project Level Permissions

To delete a user from a project, complete these steps:

1. Follow Step 1 and Step 2 on page 64.
2. Click **Delete** and a SAS Content Categorization Studio confirmation window appears:



3. Click **Yes** to close the SAS Content Categorization Studio confirmation window.
4. Click **OK** in the Project Server Permissions window and this window closes.

When you remove a user from the project, this user no longer appears in the Project Server Permissions window. This user retains `Default User` `Read Only`

permissions until you change this user's permission level. If you want to remove the user from all of these interfaces, delete the user at the _tk240 level. For more information, see Section A.5.4 *Delete a User* on page 85.

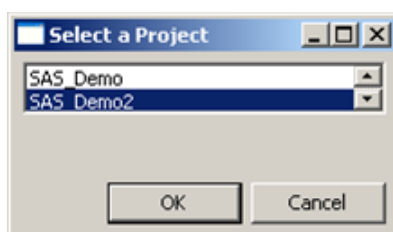
4.6 Remove a Project from the Server

Only an administrator has the permission level necessary to delete a project from the server. This operation cannot be performed on a project that is currently in use. Use the steps in this section to remove a project using the user interface, or see Section A.5.8 *Delete a Remote Project Using _tk240* on page 87.

When you remove a project from the server, you permanently delete this project.

To remove a project from the server, complete these steps:

1. Select **File --> Remove Project From Server**. The Select a Project window appears.



2. Select the project that you want to permanently delete from the server.
3. Click **OK**. Two successive SAS Content Categorization Studio windows appear.



-
4. If you are sure that you want to remove the project, click **OK** to perform this operation and to close the SAS Content Categorization Studio window.



5. Click **OK** to close the SAS Content Categorization Studio window.

Reference Section

- Appendix A: *Using the _tk240 Utility on page 73*
- Appendix B: *The Program Files on page 91*
- Appendix C: *Recommended Reading on page 95*
- Appendix D: *Glossary on page 97*

Appendix: A

Using the _tk240 Utility

- *Overview of the _tk240 Utility*
- *Command-Line Parameters*
- *Command-Line Operations*
- *Performing Database Operations*
- *Performing Server Operations*
- *Performing General Operations*
- *Set the Path to the SAS License*

A.1 Overview of the _tk240 Utility

The _tk240 utility is a command-line program that complements the functionality of the SAS Enterprise Content Categorization Studio application for Windows. This utility is used only by administrators, not regular users. Supported platforms for the _tk240 utility include Windows and all major UNIX distributions. For example, use Linux, Solaris, FreeBSD, and others. For more information, see the *SAS Content Categorization Studio: User's Guide*.

The _tk240 utility program is shipped in a package with a name that is similar to the example shown below:

```
SAS_ConCat_Collab_Servers_<arch>.exe  
sas_cc_collab_servers_<arch>.tar.gz
```

This package contains the SAS Enterprise Content Categorization Studio data directory that is required by _tk240. This directory is the root directory of the extracted package for both Windows and UNIX operating systems.

There are several types of operations that you, as an administrator, can perform using _tk240:

-
- Perform database operations. For more information, see Section A.4 *Performing Database Operations on page 80*.
 - Access remote projects on the server and perform server operations. For more information, see Section A.5 *Performing Server Operations on page 84*.
 - Perform project-specific operations. For more information, see Section A.6 *Performing General Operations on page 88*.

The parameters for these operations are described in Section A.2 *Command-Line Parameters* below and Section A.3 *Command-Line Operations on page 76*.

A.2 Command-Line Parameters

Use this section to understand and apply the various parameters that you specify when you use the `_tk240` command line. Before you use this chapter, see the table below:

Table A-1: `_tk240` Command-Line Parameters

Parameter	Description
<code><admin _pwd></code>	Specify the password for the administrative user with permission to access the database on the server. For more information, see Section A.4 <i>Performing Database Operations on page 80</i> .
<code><admin _user></code>	Specify the name of the administrative user with permission to perform database operations. For more information, see Section A.4 <i>Performing Database Operations on page 80</i> .

Table A-1: _tk240 Command-Line Parameters (Continued)

Parameter	Description
<data_path>	<p>Specify the full path to the SAS Enterprise Content Categorization Studio data directory that is the root directory of the extracted package file.</p> <p>The data directory is the top-level folder, not the subdirectory named data that is located within the top-level directory of the package. This top-level folder stores the language data files.</p> <p>On a Windows system, the data directory is located here:</p> <p style="padding-left: 40px;">C:\Program Files\Teragram\SAS Content Categorization Collaborative Server</p> <p>On a UNIX system, the data directory is located here:</p> <p style="padding-left: 40px;">sas_cc_collab_servers_linux32/collab_server/ dataC:\Program Files\Teragram\tk240</p>
<missing_ tk2_file _path>	Specify the path to the .tk2 file that contains missing concepts. See <data_path> above before specifying this path. For more information, see Section A.6.2 <i>Locate Missing Concepts in a Project on page 88</i> .
<odbc_ driver _name>	Enter the name of the ODBC data source. For example, specify tk240db. For more information, see Section 2.4 <i>Setting Up ODBC Data Sources on page 18</i> .
<pwd>	Enter the password for a regular user. For more information, see Section A.5 <i>Performing Server Operations on page 84</i> .
<project _name>	Upload the project using this name. Otherwise, it is uploaded, opened, or deleted using the name specified in the SAS Enterprise Content Categorization Studio file.
<tk2_file _path>	<p>Enter the path to the SAS Enterprise Content Categorization Studio project file that you want to work with. This can be a relative path. See the following example:</p> <p style="padding-left: 40px;">Projects/Test/Test.tk2</p> <p>Alternatively, it can be an absolute path. See the following example:</p> <p style="padding-left: 40px;">C:\Program Files\Teragram\tk240\Projects\ Test\Test.tk2</p>

Table A-1: _tk240 Command-Line Parameters (Continued)

Parameter	Description
<user>	Enter the name of a regular user here. For more information, see Section A.5 <i>Performing Server Operations on page 84</i> .
<path>	Specify the relative path to the <code>setinit</code> file for the SAS license.

A.3 Command-Line Operations

Use this section to understand and apply the various operations that you specify when using the `_tk240` command line. The table below lists and explains the operations that apply to command-line operations in the following three subsections of this chapter:

- Section A.4 *Performing Database Operations on page 80*
- Section A.5 *Performing Server Operations on page 84*
- Section A.6 *Performing General Operations on page 88*

Table A-2: _tk240 Command-Line Operations

Operation	Description
<code>-add_multiple_database_support</code>	Enable support for multiple databases on a single server.
<code>-add_user</code>	Add a user to the database.
<code>-admin</code>	Make this user an administrative user with full database permissions. Note: The <code>-admin</code> and <code>-regular</code> options are mutually exclusive. (<code>-regular</code> is explained below.)
<code>-can_upload</code>	Give a regular user permission to upload projects to the server. When this user uploads a SAS Enterprise Content Categorization Studio project to the server, this person becomes a project administrator. However, this person is only an administrator for the uploaded project.

Table A-2: _tk240 Command-Line Operations (Continued)

Operation	Description
-change_user	Change the specified permission level or password for this user. For example, you might want to change the user type to enable a regular user to become a project administrator for upload purposes.
-check_project	Check the project for errors.
-clear_locks	Clear any orphan locks, regardless of the type, from the database. For more information, see Section A.4.7 <i>Clear Database Locks</i> on page 83.
-commit_and_unlock_project	Commit changes and unlock a SAS Enterprise Content Categorization Studio collaborative project.
-concepts_fpat	Produce a concepts binary file that is optimized for best performance with SAS Enterprise Content Categorization Studio. This parameter prioritizes concept matching over build time. This selection is also equivalent to selecting Matching Speed in the Optimize For setting on the Concepts tab of the Project Settings window.
-data	Specify the path to the SAS Enterprise Content Categorization Studio data directory here. For more information, see Table A-1 <i>_tk240 Command-Line Parameters</i> on page 74 that describes the <data_path> parameter.
-database	Enter a name for a database that is not the default database. Before you specify the -database operation, see Section A.4.3 <i>Add Support for Multiple Databases</i> on page 81.
-delete_user	Delete the selected user from the database.
-delete_remote_project	Delete a SAS Enterprise Content Categorization Studio project from the server.
-driver	Specify the name of the ODBC data source that is used to connect to the database.
-files	Specify the names of the files appearing in the file list.
-host	Specify this operation for a MySQL user. By default, two user accounts are created. These accounts are <user>@% and <user>@localhost. If the -host parameter is specified, accounts <user>@host1, <user>@host2, and others, are created instead.

Table A-2: _tk240 Command-Line Operations (Continued)

Operation	Description
-initialize_database	Initialize the database for SAS Enterprise Content Categorization Studio.
-initialize_doc_repository	Initialize the testing document repository.
-list_remote_project	List all of the remote projects located on the server.
-lock_and_checkout_project	Lock the project, and check it out, to prevent other users from making any changes to this project.
-login_name	Define the Windows login name that is used to derive the database user name. Specify the login name in the DOMAIN\username form. Note: This parameter is used with Windows Authentication for SQL Server databases.
-missing	Locate the missing concepts project.
-noserverrole	Prevent administrative users from being added to the sysadmin server role. This is also true when regular users become administrative users. Note: This parameter is used with Windows Authentication for SQL Server databases.
-nosysdelete	Use this operation to prevent a user record in the system database from being deleted when you delete a user from the database.
-open_remote_project	Open a project from the server if you have a subdirectory named Shared Projects in the <data_path> directory.
-pass	(Optional) If Windows authentication is not specified, enter the password for the administrator.
-password	If Windows Authentication is not specified, enter the password for the new account.
-project	Enter the path to the SAS Enterprise Content Categorization Studio project.

Table A-2: _tk240 Command-Line Operations (Continued)

Operation	Description
-rebuild	<p>Choose one of the following operations:</p> <ul style="list-style-type: none"> - all: Rebuild both the categories and concepts in the project, if applicable. - cat: This operation is equivalent to Build Rule-based Categorizer in the SAS Enterprise Content Categorization Studio graphical user interface. - concepts: This syntax is equivalent to Compile Concepts in the SAS Enterprise Content Categorization Studio user interface. <p>Note: For more information, see the <i>SAS Content Categorization Studio: User's Guide</i>.</p>
-regular	<p>Read-Only permissions are granted to this user until another permission level is specified using the Set User Permission window. For more information, see Section 3.3.4 <i>The Set User Permission Windows on page 45</i>.</p> <p>Note: The -admin and -regular options are mutually exclusive.</p>
-save_ram	_tk240 clears out some additional memory space during the category build process and uses less overall memory, at the expense of speed.
-setinit <setinit_path>	Specify the path to the setinit file that contains the SAS license.
-show_users	Print a list of all of the project users.
-thresh	Output only the first n results.
-upload _project	Upload a project from your local machine to the server.
-update_ database	Update the SAS Enterprise Content Categorization Studio database without losing any of your data.
-use_old_ shared_ projects	<p>Store the locally cached copy of the project in: Shared Projects\<project_name> instead of: Shared Projects\<odbc_driver_name>\<user> \<project_name>.</p>
-user	(Optional) If Windows Authentication is not specified, enter the user name of an existing administrative account.

Table A-2: _tk240 Command-Line Operations (Continued)

Operation	Description
-username	(Optional) If Windows Authentication is not specified, enter the name of the new user that you are adding to this project.
-verbose	Run in verbose mode. This operation is useful for debugging purposes.
-windows_authentication	Use Windows authentication to connect to a Microsoft SQL Server database. Note: This parameter is used with Windows Authentication.

A.4 Performing Database Operations

A.4.1 Initialize the SAS Enterprise Content Categorization Studio Database

Initialize the SAS Enterprise Content Categorization Studio database tables, before you as an administrator, perform any collaborative operations. Initialize the database using the following command line:

```
_tk240 -initialize_database
      -driver <odbc_driver_name>
      -user <admin_user>
      [-pass <admin_pwd>]
      [-verbose]
      [-database <db_name>]
      [-windows_authentication]
```

If [-windows_authentication] is specified, it is not necessary to specify the user, password, or the database name.

If **-database** is specified, the database is created under a name that is different from the default name tk240db. The **-database** operation is specific to this command line.

Note: Before you specify `-database`, see Section A.4.3 *Add Support for Multiple Databases on page 81*.

A.4.2 Update the SAS Enterprise Content Categorization Studio Database

To gain the benefits of periodic SAS Enterprise Content Categorization Studio database schema and table content updates and new features, without losing any existing data, specify the following command line:

```
_tk240 -update_database  
      -driver <odbc_driver_name>  
      -user <admin_user>  
      [-pass <admin_pwd>]  
      [-verbose]  
      [-windows_authentication]
```

The `-update_database` operation is specific to this command line.

A.4.3 Add Support for Multiple Databases

Enable multiple database support for SAS Enterprise Content Categorization Studio. By default, the SAS Enterprise Content Categorization Studio database is named `tk240db`. Add multiple database support using the following command line:

```
_tk240 -add_multiple_database_support  
      -driver <odbc_driver_name>  
      -user <admin_user>  
      [-pass <admin_pwd>]  
      [-verbose]  
      [-windows_authentication]
```

The `-add_multiple_database_support` operation is specific to this command line.

A.4.4 Initialize the Testing Document Repository

Initialize the repository to enable the testing document repository. This repository stores sets of test documents that are associated with a project along with that project.

```
_tk240 -initialize_doc_repository  
      -driver <odbc_driver_name>  
      -user <admin_user>  
      [-pass <admin_pwd>]  
      [-verbose]  
      [-windows_authentication]
```

The **-initialize_doc_repository** operation is specific to this command line.

A.4.5 Lock and Check Out a Project

Put a lock on the project after you open the project to ensure that no other user can make changes to this project. This operation is similar to opening a remote project.

```
_tk240 -lock_and_checkout_project project  
      -data <data_path>  
      -driver <odbc_driver_name>  
      -user <admin_user>  
      [-pass <admin_pwd>]  
      [-verbose]  
      [-windows_authentication]
```

The **-lock_and_checkout_project** operation is specific to this command line. For more information, see Section A.4.6 *Commit and Unlock a Project* below and Section A.5.7 *Open a Remote Project on page 87*.

A.4.6 Commit and Unlock a Project

Commit all outstanding changes to a project and release any locks that you, as the administrator, currently hold. This operation is used after you specify the **-lock_and_checkout_project** operation.

```
_tk240 -commit_and_unlock_project project  
      -data <data_path>  
      -driver <odbc_driver_name>
```

```
-user <admin_user>  
[-pass <admin_pwd>]  
[-verbose]  
[-windows_authentication]
```

The **-commit and unlock project** operation is specific to this command line. For more information, see Section A.4.5 *Lock and Check Out a Project* above.

A.4.7 Clear Database Locks

Run the `_tk240` command line to clear any existing orphan locks. When you commit the changes to either a taxonomy, or to an individual node, SAS Enterprise Content Categorization Studio creates locks in the database. These locks ensure that other users do not try to make changes while the commit operation is in progress. The three types of locks are the taxonomy, category, and concept locks.

Under normal circumstances, these locks are deleted when the commit operation is completed. On rare occasions, however, it is possible for orphan locks to remain in the database. These orphan locks make it impossible for users to update the project. Use the following command line to clear any orphan locks:

```
_tk240 -clear_locks [taxonomy|category|concept]  
-driver <odbc_driver_name>  
-user <admin_user>  
[-pass <admin_pwd>]  
[-verbose]  
[-windows_authentication]
```

The **-clear_locks** operation is specific to this command line.

Caution: This command should be run carefully, because it deletes all of the locks that are currently in the database. If valid locks are deleted, data corruption could result.

A.5 Performing Server Operations

A.5.1 Add Users

Create both a system database account and a SAS Enterprise Content Categorization Studio user record for the new user. If a system database account for the user already exists, only a SAS Enterprise Content Categorization Studio user record is created. The user name of the new user is specified as <user>. The user password is specified as <pwd>.

```
_tk240 -add_user -driver <odbc_driver_name>
      -user <admin_user>
      [-pass <admin_pwd>]
      -username <user>
      [-password <pwd>]
      -admin|-regular
      [-can_upload]
      [-host host1 -host host2 ...] [-verbose]
      [-noserverrole]
```

The -add_user operation is specific to this command line.

Note: If you are using the Windows authentication mode with Microsoft SQL Server, see Section A.5.2 *Add Users with Microsoft SQL Server with Windows Authentication Mode* below.

A.5.2 Add Users with Microsoft SQL Server with Windows Authentication Mode

If you are using Microsoft SQL Server with Windows Authentication mode, use -windows_authentication to create a user. The -user, -pass, -username, and -password parameters are not required in this case. The username for the new user is generated automatically from the -login_name parameter.

```
_tk240 -windows_authentication
      -login_name <DOMAIN\username>
```

This operation does not create a Windows login or a database account for the login. This operation does create a database user corresponding to the login. Specify that login with the **-login_name** parameter. If you use the **-admin** parameter, the user that is created is added to the sysadmin server role by default. To prevent this operation, specify the **-noserverrole** parameter.

A.5.3 Modify User Permissions and Passwords

Modify the permissions, or change the password, for an existing SAS Enterprise Content Categorization Studio user. For example, you might want to change the user type to make a regular user a project administrator by enabling upload capabilities. Use the following command line to perform these operations:

```
_tk240 -change_user -driver <odbc_driver_name>
      -user <admin_user>
      [-pass <admin_pwd>]
      -username <user>
      [-password <pwd>]
      [-admin|-regular]
      [-can_upload]
      [-verbose]
      [-noserverrole]I
```

The **-change_user** operation is specific to this command line. This operation is similar to the **-add_user** operation. If you are using Microsoft SQL Server with Windows Authentication, specify the **-windows_authentication** and the **-login_name <DOMAIN\username>** parameters. It is not necessary to specify the **-user**, **-pass**, **-username**, or **-password** parameters with the Windows Authentication mode.

A.5.4 Delete a User

Remove a SAS Enterprise Content Categorization Studio user from the collaborative database by using the following command line:

```
_tk240 -delete_user
      -driver <odbc_driver_name>
      -user <admin_user>
      [-pass <admin_pwd>]
      -username <user>
```

```
[-verbose]
[-nosysdelete]
```

Note: If `-nosysdelete` or Windows Authentication is specified, this operation does not erase the user's account on the system. Only the SAS Enterprise Content Categorization Studio user record is deleted.

The **`-delete_user`** operation is specific to this command line. This operation is similar to the `-add_user` operation. If you are using Microsoft SQL Server with Windows Authentication, specify the `-windows_authentication` and the `-login_name <DOMAIN\username>` parameters. You do not need to specify the `-user`, `-pass`, `-username`, or `-password` parameters with the Windows Authentication mode.

A.5.5 Show All Users

Print a list of users with their permission levels using the following command line:

```
_tk240 -show_users
      -driver <odbc_driver_name>
      -user <user>
      [-pass <password>]
      [-verbose]
      [-windows_authentication]
```

The **`-show_users`** operation is specific to this command line.

A.5.6 Upload a Project Using `_tk240`

Upload a project from your local machine to the server using the following command line:

```
_tk240 -upload_project [<project_name>]
      -project <tk2_file_path>
      -data <data_path>
      -driver <odbc_driver_name>
      -user <user> [-pass <password>]
      [-windows_authentication]
```

The **`-upload_project`** operation is specific to this command line.

A.5.7 Open a Remote Project

Access a project if you have a subdirectory named `Shared Projects` in the `<data_path>` directory and you are either a regular user or an administrator for the remote project. Open the project using the following command line:

```
_tk240 -open_remote_project <project_name>
      -data <data_path>
      -driver <odbc_driver_name>
      -user <user>
      [-pass <password>]
      [-use_old_shared_projects]
      [-windows_authentication]
```

The `-open_remote_project` and the `-use_old_shared_projects` operations are unique to this command line. If you run one of these operations and there is no locally cached version of the project, the project is downloaded from the server. However, if there is a locally cached version, `_tk240` loads the cached version instead. When `-use_old_shared_projects` is specified, the locally cached copy of the project is stored in this location:

```
Shared Projects\<project_name>
```

instead of the following location:

```
Shared Projects\<odbc_driver_name>\<user>\<project_name>
```

Note: Specifying the `-use_old_shared_projects` operation is equivalent to unchecking **Use separate Shared Projects folder for each user** in the Options window.

A.5.8 Delete a Remote Project Using `_tk240`

Remove that project from the server, if you are an administrator. Delete the project using the following command line:

```
_tk240 -delete_remote_project <project_name>
      -driver <odbc_driver_name>
      -user <user>
      [-pass <password>]
      [-windows_authentication]
```

The `-delete_remote_project` operation is specific to this command line.

A.5.9 List Remote Projects

List all of the remote projects located on the server using the following command line:

```
_tk240 -list_remote_projects  
      -driver <odbc_driver_name>  
      -user <user>  
      [-pass <password>]  
      [-windows_authentication]
```

The -list_remote_project operation is specific to this command line.

A.6 Performing General Operations

A.6.1 Rebuild a Project

Rebuild a collaborative project using the following command line:

```
_tk240 -project <tk2_file_path>  
      -data <data_path>  
      -rebuild all|cat|concepts  
      [-concepts_fpat]  
      [-save_ram]
```

The -save_ram and -concepts_fpat operations are specific to this command line.

A.6.2 Locate Missing Concepts in a Project

Locate the missing concepts in a project by specifying the following command-line syntax:

```
_tk240 -project <tk2_file_path>  
      -data <data_path>  
      -missing <missing_tk2_file_path>
```

The -missing operation that is specific to this command line, enables you to compare your SAS Enterprise Content Categorization Studio project with

another SAS Enterprise Content Categorization Studio project containing the missing concepts.

The second project is specified by the `.tk2` file that is pointed to by `<missing_tk2_file_path>`. This command determines what concepts are in the missing project and stores this information in the `data` directory for the SAS Enterprise Content Categorization Studio project that is missing some of its concepts.

A.6.3 Run a Project Check

Check the syntax of the project files as they are generated by specifying the following command-line syntax. Specify this command-line operation when a process other than SAS Enterprise Content Categorization Studio is used to create the project files.

```
_tk240 -check_project project.tk2 -data datapath
```

The automatically generated script checks the SAS Enterprise Content Categorization Studio project to ensure there are no errors in the project.

Example A-1: Project Check

```
Checking.tk2 file syntax...PASS
Checking categorizer XML files...PASS
Checking concept XML files...PASS
Checking language project English
Checking for dependencies...PASS
Checking syntax for categories...PASS
Checking for English.concept directory...PASS
Checking syntax for concepts...PASS
Project syntax check succeeded.
```

The `-check_project` operation is specific to this command line.

A.7 Set the Path to the SAS License

This license is the SAS installation data file (SID file) that is included in the Software Order E-mail (SOE) that you received. The `setinit` file can be saved in any folder.

To specify the path using `_tk240`, add the following line to `_tk240`:

```
-setinit <setinit_path>
```

This line specifies the path to the file that contains the `setinit` information.

You can also set the environment variable `TG_SETINIT` and specify the path to the file that contains the `setinit` information.

Appendix: B

The Program Files

- *Overview of the Program Files*
- *Program File Directives*
- *Configuration Example*

B.1 Overview of the Program Files

This appendix covers the folders, files, tags, and directives that are specific to the collaborative features of SAS Enterprise Content Categorization Studio. This appendix also provides a configuration example.

B.2 Program File Directives

The table below describes the directives that are specific to the collaborative application of SAS Enterprise Content Categorization Studio.

Table B-1: *SAS Enterprise Content Categorization Studio* Directives

Directive	Description
See Table C-1 in the <i>SAS Content Categorization Studio: User's Guide</i> :	
<code>conceptname.n.def</code>	Specify the definition file for a user-created concept. The character <i>n</i> specifies the identification number on the server. Note: This value is different for stand-alone projects.
See Section B.2.1 .tk2 File Format and Table C-2 in the <i>SAS Content Categorization Studio: User's Guide</i> :	
<code>server_pid=<pid></code>	Specify the project identification number in the database.
<code>client-server</code>	Indicate that the project is a collaborative project.

Table B-1: *SAS Enterprise Content Categorization Studio* Directives

Directive	Description
<code>server_lpid=<lpid></code>	Specify the language project identification number in the SAS Enterprise Content Categorization Studio database. (Use only with collaborative projects.)
See information about category project settings and Table C-3 in the <i>SAS Content Categorization Studio: User's Guide</i> :	
<code>TaxonomyVersion</code>	Specify the current version of the category taxonomy on the server.
See information about categories and Table C-4 in <i>SAS Content Categorization Studio: User's Guide</i> :	
<code>LocallyModified</code>	(Optional) The presence of this tag, in the format <code><LocallyModified/></code> , indicates that the user made local changes. These changes are saved in the cached version on their desktop, but not yet committed to the server.
See information about concept Project Settings and Table C-5 in the <i>SAS Content Categorization Studio: User's Guide</i> :	
<code>TaxonomyVersion</code>	Specify the current version of the concepts taxonomy on the server.
See information about concepts and Table C-6 in the <i>SAS Content Categorization Studio: User's Guide</i> :	
<code>LocallyModified</code>	(Optional) See this tag, in the format <code><LocallyModified/></code> . When this tag occurs, the tag indicates that the user has made local changes that are saved in the cached version on their desktop. These changes are not yet committed to the server.

B.3 Configuration Example

The example below is a sample `.tk2` file that is intended for use with a collaborative project that contains a single English language project with both categories and concepts.

Example B-1: Single Language Project

```
project=News
server_pid=1
language=English
client-server
server_lpid=1
```

```
categorizer=English.directory.xml  
uses_concepts  
never_expand  
should_rebuild_categories  
concept=English.concept.xml  
should_rebuild_classifier  
should_rebuild_concepts
```

For more information, see Section C.4 *Configuration Examples* in *SAS Content Categorization Studio: User's Guide*.

Appendix: C

Recommended Reading

The following books are recommended:

- *SAS Enterprise Content Categorization Studio: User's Guide*: Deploy the collaborative operations and develop complex LITI concepts and their rules for SAS Content Categorization Studio. Use this book after you read *SAS Content Categorization Studio: User's Guide*.
- *SAS Content Categorization Studio: Installation Guide*: Install SAS Content Categorization Studio.
- *SAS Content Categorization Studio: User's Guide*: Create a SAS Content Categorization Studio project, test, and upload the output to SAS Content Categorization Server.
- *SAS Enterprise Content Categorization Servers: Administrator's Guide*: Install, configure, and use SAS Content Categorization Server, SAS Content Categorization Collaborative Server, and SAS Document Conversion Server. Upload .li files using this product.
- *SAS Content Categorization Single User Servers: Administrator's Guide*: Install, configure, and use SAS Content Categorization Server and SAS Document Conversion Server.

SAS offers instructor-led training and self-paced e-learning courses to help you get started with the SAS add-in, learn how the SAS add-in works with the other products in the SAS Enterprise Intelligence Platform, and learn how to run stored processes in the SAS add-in.

For more information about the courses available, see support.sas.com/training.

For a complete list of SAS publications, see the current SAS Publishing Catalog. To order the most current publications or to receive a free copy of the catalog, contact a SAS representative at

SAS Publishing Sales
SAS Campus Drive
Cary, NC 27513
Telephone: (800) 727-3228*

Fax: (919) 677-8166

E-mail: sasbook@sas.com

Web address: support.sas.com/pubs

* For other SAS Institute business, call (919) 677-8000.

Customers outside the United States should contact their local SAS office.

Appendix: D

Glossary

Branch

refers to either the category, or the concepts, section of the taxonomy tree. The first node in a branch is either the *Categorizer* or the *Concepts* node. If the project is built with more than one language, each language section is also referred to as a branch.

Categorization

refers to the process of concisely defining the subject matter of a document, in other words, the main idea or subject of the document.

Classifier

specifies a list-based set of terms that are extracted from your documents.

Concept

define an autonomous piece of information such as movie, book, title, and so on. Also see *Entity*.

Collaboration

work of two, or more, subject matter experts collaborating on one project. This project is located on a server with a cached copy in the shared projects folder on a local machine.

Definition

defines a concept. Also see *Rule*.

Document

refers to a printable page. Also see *Text*.

Node

refers to the visual representation of a concept or a category. Less frequently, this term is used to refer to the *Categorizer*, *Concepts*, *Top*, or another component of the taxonomy tree.

Regular user

refers to a collaborative user with one of several permission levels for a project. However, this user lacks either database or administrative privileges. The regular developer is added to the project and has his or her permission level set by either the database, or the project, administrator.

Remote project

resides on a server, instead of your local machine.

Rule

defines a category. This term can also be used to refer to a concept definition.

Shared Projects Folder

stores a cached copy of the project that is located on the server on your local machine. Individual cached copies of the project can be stored on one machine using different shared projects folders.

Syntax changes

edits to a category rule or to a concept definition are defined as changes to the syntax of the rule or definition.

Taxonomy

organizes a classification structure that can be either a flat or a hierarchical system.

Text

forms a written document, or a Web page. Also see *Document*.

User

defines a SAS Enterprise Content Categorization Studio developer. This term is usually used to distinguish a subject matter expert from an administrator. Also see *Developer*.

Index

.tk2 file	
sample	92
_tk240 command-line program	12
_tk240 utility	
defined	17

A

Access Level	
specify	65
types	45
add user	
syntax	38, 84
-add_multiple _database _support	
defined	76
-add_user parameter	
defined	76
-admin	
defined	76
administrative permission	
defined	36

B

binary file	
create	12

C

-can_upload	
defined	76
Category	
rules	63
change	
taxonomy permissions	63
Change Password interface	
usage	51

-change_user	
defined	77
-check_project	
defined	77
child nodes	
set permissions	66
-clear_locks	
defined	77
client-server	
defined	91
collaborative project	
set up	55
concept definition	
write	63
conceptname.n.def	
defined	91
-concepts_fpat	
defined	77
configure ODBC data sources	
ODBC.ini file	20
Connector/ODBC window	
usage	24
Create a New Data Source to SQL Server window	
open	27, 28, 29
create data sources	
credentials	19
Create New Data Source window	
open	23
usage	26
credentials	
create data sources	19

D

-data	
defined	77
data source name	
repository login	36
specify	24
-database	
option	77

database	
connect to	61
database administrator	
permission defined	63
Database parameter	
defined	21
database tables	
initialize	80
Default User	
class	64
Default User permission	
change	64
-delete_remote_project	
defined	77
-delete_user	
defined	77
-driver	
option	77
Driver parameter	
defined	21

E

Enter project name field	
usage	59

F

File	
command	40
options	40
Upload Project to Server	40

H

-host	
defined	77

I

initial project	
upload permissions	57
-initialize_doc_repository	
defined	78
-initialize_database	
option	78

L

language.concepts file	12
language.mco file	12
-list_remote_project	
defined	78
LocallyModified	
defined	92
-lock_and_checkout_project	
defined	78
log in	
information	36
-login_name	
defined	78

M

menu bar	40
-missing	
defined	78
my.cnf file	
default setting	17
max packet size	16
MySQL	
install	16

N

No Access option	
defined	63
-noserverrole	
defined	78
-nosysdelete	
defined	78

O

ODBC Data Source Administrator window	
open	22
ODBC data sources	
3 types	18
ODBC drivers	61
install	18
ODBC Microsoft SQL Server Setup window	
open	31, 32
odbc.ini file	
usage	20
Open Remote Project	
option	60
Options window	
usage	13

P

-pass	
option	78
-password	
defined	78
password	
repository login	36, 61
permission windows	
usage	62
permissions	
Category level	13
change taxonomy	63
Concept level	13

levels defined	63
override	63, 65
project level	13
read and write rules	63
Read Only	63
server level	63
setting	13, 56
view	42
Port parameter	
defined	21
-project	
defined	78
project	
delete	68
name	59
upload	68
project administrator	
permission defined	63
Project Server Permissions window	
open	41
usage	42, 64
user deleted	67

R

Read and Write Rules	
permission defined	46, 63
Read Only	
default permission level	62
default user setting	57, 64
permission defined	46, 63
Read, Write and Change Taxonomy	
permission defined	46, 62, 63
-rebuild	
defined	79
-regular	
defined	79
remote project	
open	57, 60
Remove Project From Server	
usage	68

repository	
remote	36
Repository field	
usage	60
Repository Login window	
open	60
usage	48
Revision Log	13

S

-save_ram	
defined	79
Select a Directory window	
usage	49
Select a Project window	
open	50, 68
server configuration	
change data size	16
server level	
permissions	63
server operations	13
Server parameter	
defined	21
Server Permissions	
node level	41, 44
server project	
name	59
server_lpid=	
defined	92
server_pid=	
defined	91
-setinit	
defined	79
Shared Projects folder	
cached copy	56
creation of	61
Shared Projects Folder field	
usage	61
-show_users	79
defined	79

T

taxonomy	
changes	13
window	40
TaxonomyVersion	
defined	92
Test Data Source button	
usage	32
testing documents	
upload	13
-thresh	
defined	79

U

-update_database	
defined	79
upload	
project	36, 57
Upload Project to Server	
File	40
option	58
usage	52
Upload to Server window	
open	59
usage	52
-upload_project	
defined	79
-use_old_shared_projects	
defined	79
-user	
option	79
User Name field	
usage	60
-username	
defined	80
username	
repository login	36

V

-verbose option	80
--------------------------	----

W

-windows_authentication defined	80
--	----

