



THE  
POWER  
TO KNOW.

# **SAS® Content Categorization Collaborative Server 5.2 User's Guide**



The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2011.  
*SAS® Content Categorization Collaborative Server 5.2: User's Guide*. Cary, NC: SAS Institute Inc.

## **SAS® Content Categorization Collaborative Server 5.2: User's Guide**

Copyright © 2011, 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, July 2011

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](http://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

---

<b>About This Book .....</b>	<b>1</b>
Audience .....	1
Prerequisites .....	1
Conventions .....	2
 <b>What's New in SAS Content Categorization Collaborative Server 5.2 .....</b>	 <b>3</b>
 <b>1 About SAS Content Categorization Collaborative Server .....</b>	 <b>5</b>
1.1 What Is SAS Content Categorization Collaborative Server? .....	5
1.2 Benefits to Using SAS Content Categorization Collaborative Server .....	6
1.3 How Does SAS Content Categorization Collaborative Server Work with SAS Content Categorization Studio? .....	7
1.4 Using Collaborative Operations .....	8
1.5 Architecture .....	10
 <b>2 Setting Up Your Data Source .....</b>	 <b>11</b>
2.1 Overview of Setting Up a Project on Your Server .....	11
2.2 Configuring ODBC Data Sources .....	12
2.2.1 Credentials .....	12
2.2.2 Configure a MySQL Data Source .....	13
2.2.3 Configure a Microsoft SQL Server Data Source .....	15
2.2.4 Configure an Oracle Data Source .....	21
2.3 Logging In .....	23
 <b>3 Using the Interface .....</b>	 <b>25</b>
3.1 How Collaborative Management Affects the User Interface for SAS Content Categorization Studio .....	25
3.2 The Menu Bar .....	26
3.3 The Standard Toolbar .....	28
3.4 The Options Window .....	29
3.5 The Server Operation Windows .....	31
3.5.1 Overview of the Server Operation Windows .....	31
3.5.2 Server Operations Available for Individual Taxonomy Nodes .....	32
3.6 The Miscellaneous Windows .....	33

---

3.6.1 The Repository Login Window .....	33
3.6.2 The Select a Directory Window .....	35
3.6.3 The Select a Project Window .....	36
3.6.4 The Change Password Window .....	37
3.6.5 The Upload Test Docs Window .....	37
3.6.6 The Download Test Docs Window .....	39
3.6.7 The Enter Comment Window .....	40
3.6.8 The RevisionLog Screens .....	40
3.6.9 The Revert to Older Version Window .....	43
3.6.10 The Import Category (or Concept) Window .....	44
3.6.11 An Example of Status Windows .....	46
<b>4 Getting Started .....</b>	<b>47</b>
4.1 How to Begin Collaborative Work .....	47
4.2 Cached Project .....	48
4.3 Open a Project .....	48
4.4 Keeping Projects Up-to-Date .....	50
4.4.1 Overview of Keeping Projects Up-to-Date .....	50
4.4.2 Specify Options .....	51
4.4.3 Manually Accessing Server Operations .....	52
4.5 Understanding Your User Permission Level .....	53
4.6 Working with a Cached Project .....	54
4.7 Collaborative Changes .....	54
4.7.1 Overview of Collaborative Changes .....	54
4.7.2 Modifying the Taxonomy .....	55
4.7.3 Changing the Rules or Definitions .....	55
4.8 Server Operations .....	56
4.8.1 Understanding the Server Operations .....	56
4.8.2 Using the Server Operations .....	59
4.8.2.A Check the Server Status for Single Nodes .....	59
4.8.2.B Checking the Status of the Taxonomy .....	62
4.8.2.C Removing Taxonomy Tree Messages .....	63
4.8.3 Update from the Server .....	63
4.8.4 Commit Changes to the Server .....	65
4.8.5 Using the Revision Logs .....	67
4.8.5.A Overview of Using the Revision Logs .....	67
4.8.5.B Revision Logs .....	68
4.8.5.C Revert to an Older Version .....	70

---

4.8.6 Import Categories or Concepts from a Repository .....	72
4.9 Save a Project .....	75
<b>5 Other Collaborative Operations .....</b>	<b>77</b>
5.1 Overview of Other Collaborative Operations .....	77
5.2 Uploading or Downloading Testing Files .....	77
5.2.1 Overview of Shared Test Files .....	77
5.2.2 Before Uploading or Downloading Testing Files .....	78
5.2.3 Upload Testing Files .....	78
5.2.4 Download Testing Files .....	79
5.3 Change Your Server Password .....	81
5.4 Remove a Project from Server Window .....	81
<b>Reference Section .....</b>	<b>83</b>
<b>A Recommended Reading .....</b>	<b>85</b>
<b>B Glossary .....</b>	<b>87</b>
<b>Index .....</b>	<b>89</b>



---

# About This Book

---

## Audience

SAS Content Categorization Collaborative Server is SAS Content Categorization Studio with operations that enable the following users to work together collaboratively on one project:

- Taxonomists who develop the categories and concepts that comprise the taxonomy for your enterprise
- Linguists and other persons who write the category rules and concept definitions
- Persons responsible for testing and analyzing the testing results for the rules specified for the categories and concepts

You could be assigned one of these functions, or all of them.

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

## Prerequisites


Here are the prerequisites for using SAS Content Categorization Collaborative Server:

- SAS 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 Content Categorization Studio is installed, typically the following:  <b>Windows:</b> C:/Program Files/SAS/SAS Content Categorization Studio <b>UNIX:</b> /opt/SAS_collab_server
<b>OK</b> 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.
<a href="http://www.sas.com">www.sas.com</a>	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 Content Categorization Collaborative Server 5.2

---

New and enhanced features in SAS Content Categorization Collaborative Server include the following:

- Support for Oracle
- Support for SAS Contextual Extraction Studio concepts
- Improved support for Microsoft SQL Server
- Improved user account management



---

# *Chapter: 1*

## About SAS Content Categorization Collaborative Server

---

- *What Is SAS Content Categorization Collaborative Server?*
- *Benefits to Using SAS Content Categorization Collaborative Server*
- *How Does SAS Content Categorization Collaborative Server Work with SAS Content Categorization Studio?*
- *Using Collaborative Operations*
- *Architecture*

### 1.1 What Is SAS Content Categorization Collaborative Server?

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 SAS Content Categorization Studio project. The collaborative operations are what distinguish SAS Content Categorization Collaborative Server 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. You, as an administrator specify the various levels of access for all project users, upload projects to the server, and perform other administrative operations.

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 might develop the categories, while subject matter experts write rules.

#### Levels of permissions

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

## 1.2 Benefits to Using SAS Content Categorization Collaborative Server

SAS Content Categorization Collaborative Server 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 Content Categorization Collaborative Server adds the 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 to various components.

#### 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 are for SAS Content Categorization Studio. They are included with SAS Content Categorization Collaborative Server:

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 make it easy to build large, complex, and hierarchical taxonomies. Specify your own rules, test, and generate `.mco` and `.concepts` 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` and `.concepts` 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. You can find related information, whether you know the exact terms that you are seeking.

## 1.3 How Does SAS Content Categorization Collaborative Server Work with SAS Content Categorization Studio?

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

SAS 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 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` and `.concepts` files to SAS Content Categorization Server where they are automatically applied to input documents.

## 1.4 Using Collaborative Operations

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

- Create a collaborative, customized SAS 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.

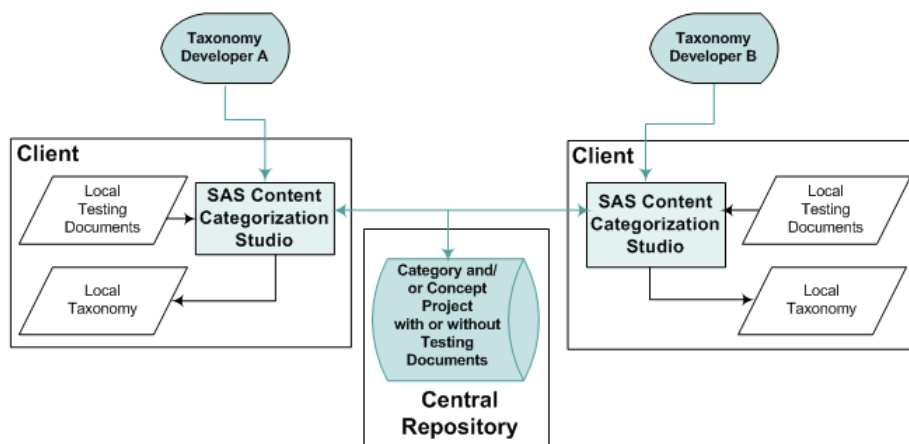
In summation, these features provide the benefits of collaboration. These operations enable your team to balance necessary control with optimal flexibility in order to meet your organization's project development requirements.

---

## 1.5 Architecture

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

*Figure 1-1 SAS Content Categorization Collaborative Server Architecture*



A SAS 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 Your Data Source

---

- *Overview of Setting Up a Project on Your Server*
- *Configuring ODBC Data Sources*
- *Logging In*

### 2.1 Overview of Setting Up a Project on Your Server

After the database administrator sets up data sources on the server, you can create an ODBC data source on your local machine. This chapter is for those users whose administrator does not set up their data sources. This data source points to the SAS Content Categorization Collaborative Server database. After you complete this process, you can access the SAS Content Categorization Studio project on the server.

The term *administrator* is used to refer to both the database and the project administrator. Database administrators are system administrators, while project administrators have upload permissions and other permissions that you as a regular user do not.

For more information, see *SAS Content Categorization Collaborative Server: Administrator's Guide*. Refer to the chapter used by administrators to set up data sources, if necessary.

---

## 2.2 Configuring ODBC Data Sources

### 2.2.1 Credentials

There are several credentials that supply the information required to create your data source:

Name

User

Password

Database: You can either specify a name or use the default entry.

Select one of the following sections, depending on the type of ODBC data source that you are creating:

- Section 2.2.2 *Configure a MySQL Data Source* on page 13
- Section 2.2.3 *Configure a Microsoft SQL Server Data Source* on page 15
- Section 2.2.4 *Configure an Oracle Data Source* on page 21

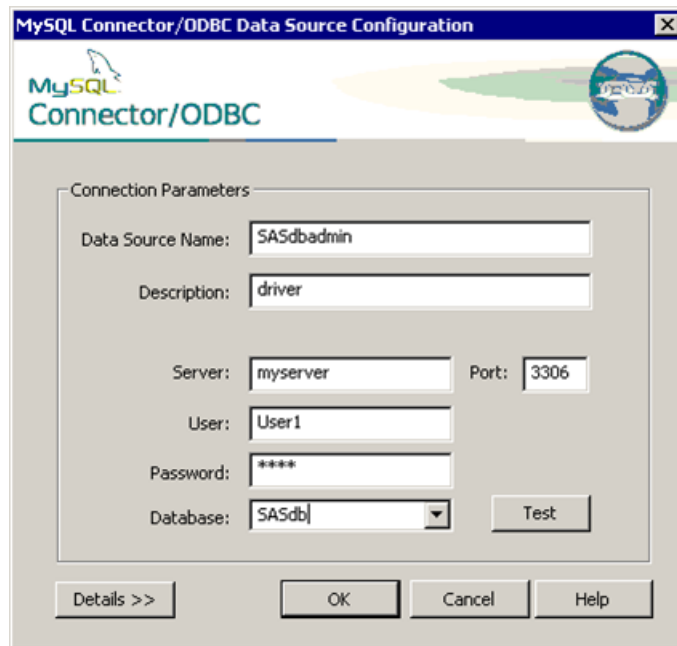
---

## 2.2.2 Configure a MySQL Data Source

If you are working with a MySQL ODBC driver, configure this data source.

To configure your MySQL data source, complete these steps:

- 1 Open the MySQL Connector/ODBC page. Use this page to point the ODBC driver to the server:



The screenshot shows the 'MySQL Connector/ODBC Data Source Configuration' dialog box. It has a title bar with the text 'MySQL Connector/ODBC Data Source Configuration' and a close button. Below the title bar is a header area with the MySQL logo and the text 'Connector/ODBC'. The main area is titled 'Connection Parameters' and contains several input fields: 'Data Source Name' with the value 'SASdbadmin', 'Description' with the value 'driver', 'Server' with the value 'myserver', 'Port' with the value '3306', 'User' with the value 'User1', 'Password' with the value '\*\*\*\*', and 'Database' with a dropdown menu showing 'SASdb'. There is a 'Test' button next to the 'Database' dropdown. At the bottom of the dialog are four buttons: 'Details >>', 'OK', 'Cancel', and 'Help'.

- 2 Enter the name of the data source, selected by the administrator, into the **Data Source Name** field. For example, type SASdb.
3. (Optional) Enter the description into the **Description** field. For example, type driver.
4. Enter the server name into the **Server** field. For example, type myserver.
5. (Optional) The default entry, shown in the right pane of this interface is entered into the **Port** field.

- 
6. Enter the user name, entered by the administrator to the database, into the **User** field.

---

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

---

7. Enter your password into the **Password** field. The password for the administrator is the administrator's password on the server.
8. Enter the name of the database that you are connecting to into the **Database** field. For example, you might type `tk240db`.
9. Click **Test**. If the connection is successful, the Connector/ODBC window appears with a message stating that the connection is successful.



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

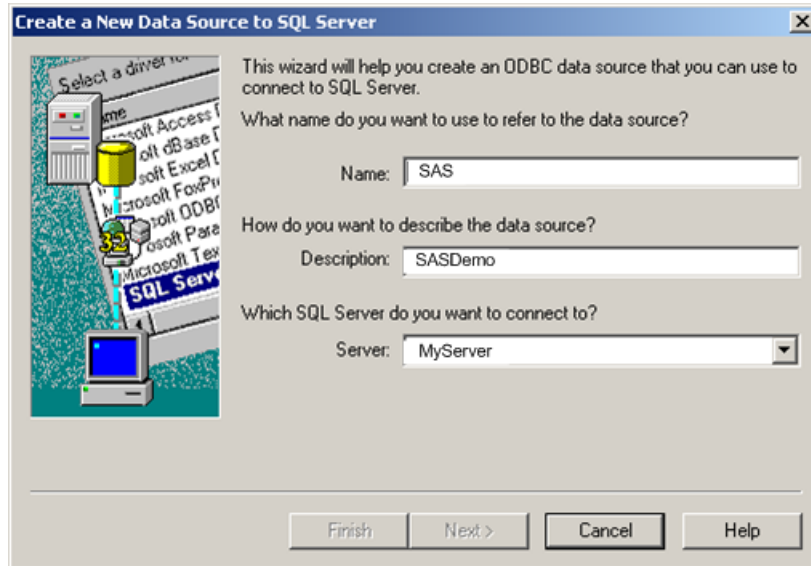
---

### 2.2.3 Configure a Microsoft SQL Server Data Source

If you are working with Microsoft SQL Server, configure this data source.

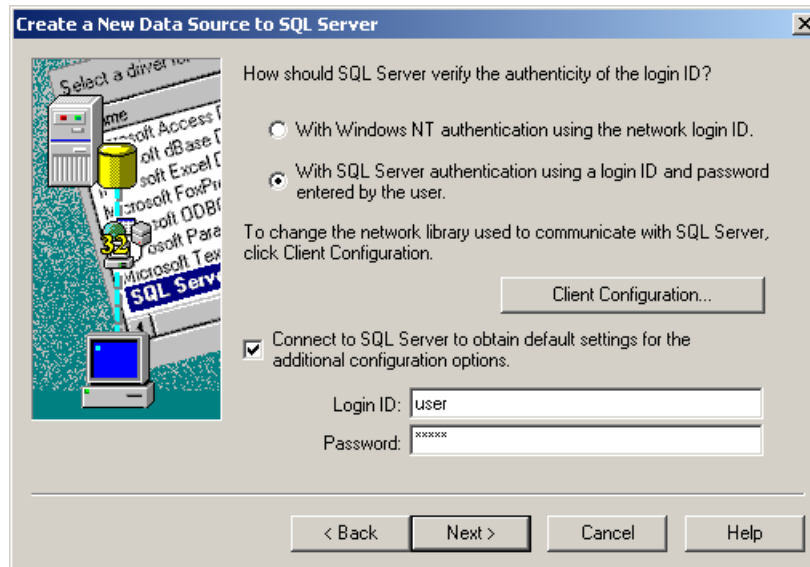
To configure your Microsoft SQL Server ODBC data sources, complete these steps:

1. Open the Create a New Data Source to SQL Server wizard.



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

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



- (Either Windows NT authentication or SQL Server authentication is OK.) Select **With SQL Server authentication using a login ID and password entered by the user.**

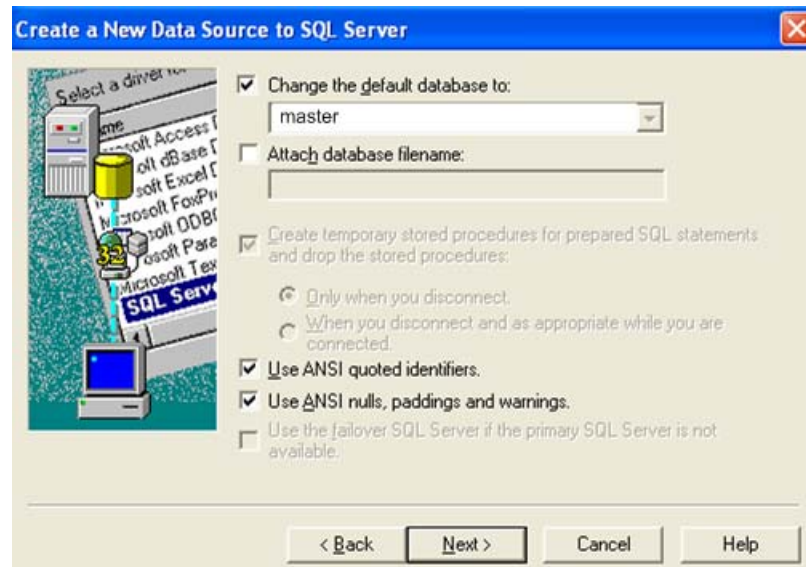
---

**Note:** Step 7 and Step 8 apply only if SQL Server authentication is selected.

---

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

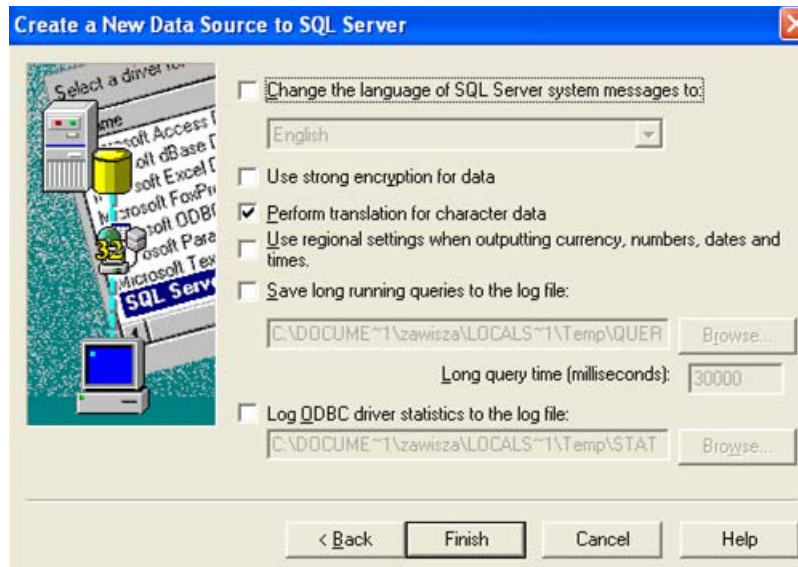
- 
9. Click **Next** and the next Create a New Data Source to SQL Server page appears.



Any of the default settings are OK.

10. Click **Next**.

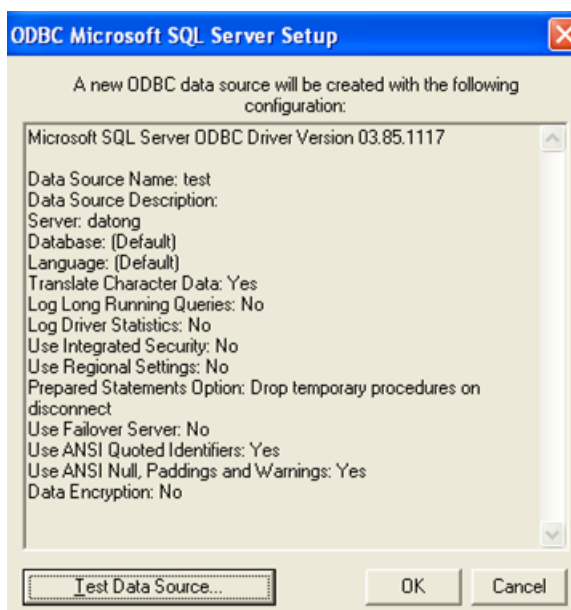
- 
11. Click **Next** and the next **Create a New Data Source to SQL Server** page appears.



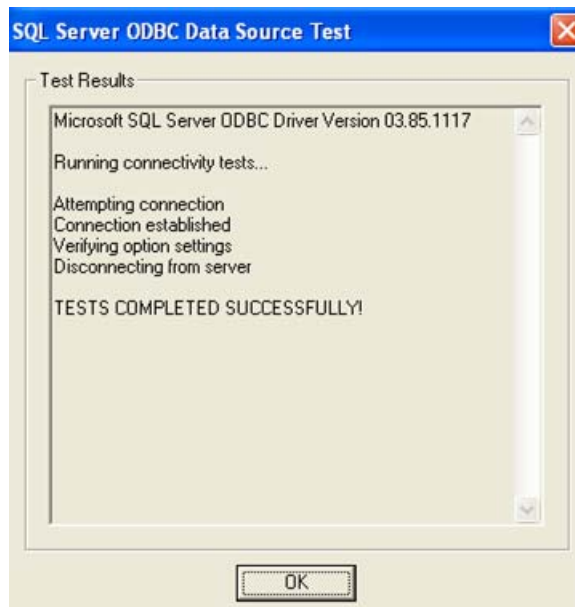
Any of the default settings are OK.



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



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



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

---

## 2.2.4 Configure an Oracle Data Source

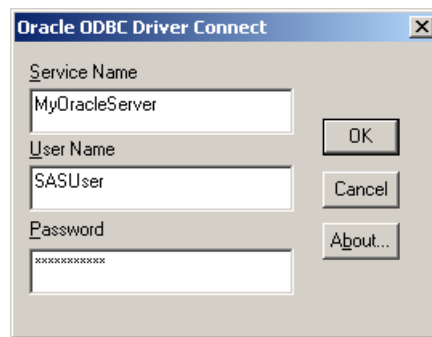
If you are working with an Oracle ODBC Driver, configure this data source.

To configure your Oracle ODBC data sources, complete these steps:

1. The Oracle ODBC Driver Configuration window appears.

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

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



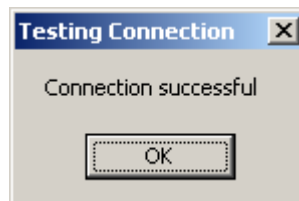
7. The fields in this window are automatically filled in for you with the exception of the **Password** field.

---

**Hints:** The **Server Name** field contains the same name as the entry in the **TNS Service Name** field.  
The **User Name** field is specified by the database administrator.

---

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



9. Click **OK**.
10. Click **OK** to close the Oracle ODBC Driver Connect window.

---

## 2.3 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



---

# 3

## Using the Interface

---

- *How Collaborative Management Affects the User Interface for SAS Content Categorization Studio*
- *The Menu Bar*
- *The Standard Toolbar*
- *The Options Window*
- *The Server Operation Windows*
- *The Miscellaneous Windows*

### 3.1 How Collaborative Management Affects the User Interface for SAS Content Categorization Studio

Collaborative operations are enabled by a server. For this reason, some of the operations that are visible in the SAS Content Categorization Studio user interface are not accessible until you set up the server. The collaborative operations enable multiple users with different permission levels to work with the same projects on one server.

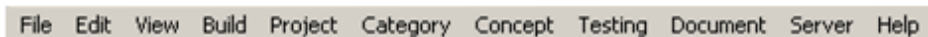
This chapter explains the interface components that regular users access. However, regular users are granted various levels of permissions by the administrator. These permission levels affect your access to the project and the operations that you perform. This chapter does not describe any of the components that are limited to administrative users. This chapter assumes the highest level of access for regular users, unless otherwise stated.

---

## 3.2 The Menu Bar

Many of the collaborative operations for SAS Content Categorization Studio 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*



The table below describes the collaborative-only commands that are available in the SAS Content Categorization Studio menus for regular users. For all other operations, see *SAS Content Categorization Studio: User's Guide*.

Table 3-1: Collaborative Operations

Menu	Operation	Description
File	Open Remote Project	Access the existing projects on the server using the Select a Project window that appears.
	Remove Project From Server	Select a project to delete in the Select a Project window that appears. For more information, see Section 3.6.3 <i>The Select a Project Window</i> on page 36.
	Repository Login	Log in to the server where the collaborative projects reside using the Repository Login window that appears. For more information, see Section 3.6.1 <i>The Repository Login Window</i> on page 33.
	Change Repository Password	Change the password that you use to access the server in the Change Password window that appears. For more information, see Section 3.6.4 <i>The Change Password Window</i> on page 37.
<b>Note:</b> The Repository Login window appears if you are not logged in to the server before selecting one of the operations above.		
Edit	Select <b>Edit --&gt; Options</b> and the Options window appears with six collaborative operations. For more information, see Section 3.4 <i>The Options Window</i> on page 29.	



Table 3-1: Collaborative Operations (Continued)

Menu	Operation	Description
Server	Status	Learn the status of the selected node in relation to the server project using the messages that appear in the Taxonomy window. For more information, see Section 4.8.2.B <i>Checking the Status of the Taxonomy</i> on page 62.
	Update	Download the current version of the project, if you select either the Categorizer or Concepts node in the taxonomy. If you select a category or concept node, download the current version of the rule. Use this server operation to obtain changes from a project that is stored on the server. Any uncommitted syntax changes are overwritten during this process. For more information, see Section 4.8.3 <i>Update from the Server</i> on page 63.
	Commit	Update the project with your local changes using this operation. During this process, SAS Content Categorization Studio places a lock on the selected node to prevent another developer from committing changes during this operation. For more information, see Section 4.8.4 <i>Commit Changes to the Server</i> on page 65.
	Revision Log	Open one of the following two RevisionLog windows in <i>Notepad</i> . The first window lists only rule or definition changes for the selected node. The second window provides information about deleted and renamed nodes and rule or definition changes for the selected branch in the taxonomy. For more information, see Section 4.8.5 <i>Using the Revision Logs</i> on page 67.
	Revert to Older Version	Use the Revert to Previous Version window that enables you to replace the rule or definition with an earlier version. Use the Revision Log window to determine the version number that meets your requirements. For more information, see Section 4.8.5.C <i>Revert to an Older Version</i> on page 70.
	Upload Test Files	Load all of the test files to the server for the project that is open on your machine. This operation makes the testing files available to all of the permissioned users for this project. For more information, see Section 5.2.3 <i>Upload Testing Files</i> on page 78.

Table 3-1: Collaborative Operations (Continued)

Menu	Operation	Description
	Download Test Files	Download the testing files that are stored on the server to your machine. For more information, see Section 5.2.4 <i>Download Testing Files</i> on page 79.
Help	User's Guide	Click to see a PDF version of the <i>SAS Content Categorization Studio: User's Guide</i> where you can find general SAS Content Categorization Studio information.  <b>Note:</b> If you also purchased <i>SAS Concept Extraction Studio: User's Guide</i> , this pdf link also appears in the Help menu.
	About	Open the About SAS Content Categorization Studio window that displays version, licensing, and dating information.

## 3.3 The Standard Toolbar

The standard toolbar is located below the menu bar. Some of the buttons are specific to server operations.




Display 3-2: Standard Toolbar Buttons



---

See the table below for an explanation of each of these buttons.

Table 3-2: Standard Toolbar Buttons

Icon	Button		Description
	Server Update	Project-wide update	Select either the Categorizer or Concepts node and click this button. The SAS Content Categorization Studio status window appears asking if you want to overwrite or preserve the local changes to your project. If you do <i>not</i> want to preserve your local changes, your entire project is overwritten by the project on the server. If you want to preserve your local changes, updates are downloaded for the components that you have not modified.
		Local update	Highlight a node in the taxonomy and click this button to see a message telling you whether the selected category or concept is up-to-date with its counterpart on the server.
<b>Note:</b> In either case, status messages appear in the taxonomy tree.			
	Server Status		Select a node in the Taxonomy window and click this button. Messages appear to the right of the nodes in the taxonomy tree providing information about local changes and whether the local copy is up-to-date with the project on the server. For more information, see Table 3-4 on page 32.
	Server Commit		Click this button to commit your local changes to the server. The Enter Comment window appears, unless this component is disabled in the Options window or the category or concept is already up-to-date with its counterpart on the server. Use the Enter Comment window that appears, unless you select <b>Skip comments on commit</b> in the Options window. Write notes to track the changes to your project.

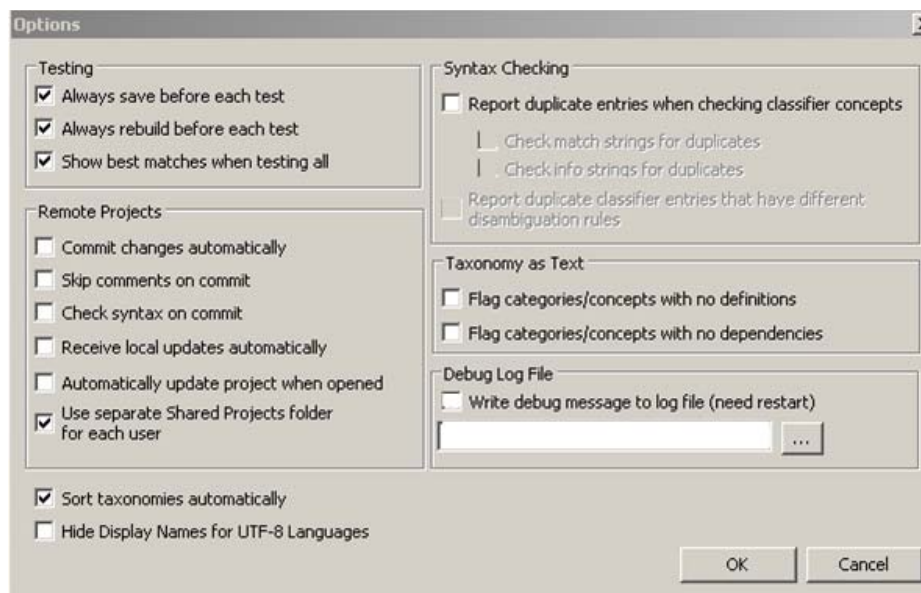
## 3.4 The Options Window

The Options window enables you to set installation-wide settings for collaborative work. (These specifications apply across all collaborative projects, unless you make different selections.) Use the six **Remote Projects**

selections in this interface to simplify the process of keeping your local project up-to-date with the project on the server.

When you make choices using the Options window, they remain selected as the default operations until you uninstall and reinstall SAS Content Categorization Studio or reset these operations. Options are installation-specific. Project Settings are project-specific.

To open the Options window, select **Edit --> Options**.



Select any of the following collaborative operations that are listed under the **Remote Projects** heading:

Table 3-3: Options Window Components

Component	Description
<b>Commit changes automatically</b>	Automatically commit rule and definition changes to the project on the server, after you make a change and select a different taxonomy node.
<b>Skip comments on commit</b>	Commit changes to the server without tracking the reasons for these revisions.

Table 3-3: Options Window Components (Continued)

Component	Description
<b>Check syntax on commit</b>	SAS Content Categorization Studio automatically checks the syntax of the definition or rule before committing it to the server.
<b>Receive local updates automatically</b>	Automatically update the category rules and concept definitions from the server whenever a change is committed by another user.
<b>Automatically update project when opened</b>	Automatically download any changes committed to the server for the taxonomy and its categories and concepts. This operation provides an alternative to the Server Update operation. For more information, see Table 3-1 on page 26.
<b>Use separate Shared Projects folder for each user</b>	(Default) Enable two or more users to use the same machine to build projects in SAS Content Categorization Studio. These users can maintain separate local (cached) projects. These projects are stored in the Shared Projects folder that is created on your local machine.

## 3.5 The Server Operation Windows

### 3.5.1 Overview of the Server Operation Windows

The server operations enable multiple developers to work together, collaboratively, on a project. You can access the server operations for your project through the **Server** menu or by right-clicking on a taxonomy node. For information about the server operations that are available from the project name node, see Table 3-4 on page 32. For information about the server operations that are available from the individual taxonomy nodes, see Section 3.5.2 *Server Operations Available for Individual Taxonomy Nodes* on page 32.

### 3.5.2 Server Operations Available for Individual Taxonomy Nodes

Some of the server operations described in Table 3-1 on page 26 and Table 3-2 on page 29 are also available when you right-click on a taxonomy node. Server operations are available for all of the nodes in the Taxonomy pane with the exception of the `Top` node.

The type of node that you select in the Taxonomy pane determines the server operations that are available. For example, if you select the project node, the available operations affect the entire taxonomy. If you select an individual taxonomy node instead, the available operations affect only the selected node.

*Display 3-3: Server Operations*



A drop-down menu appears. The available server operations are specific to the type of node that you select. See the table below for all of the available operations for regular users from the various nodes in the taxonomy.

Table 3-4: Server Operations for Taxonomy Nodes

Server Operation	Project Name	Language	Categorizer or Concepts Extractor	Top	Category or Concept
Server Status	X	X	X		X
Server Update	X	X	X		X
Server Commit	X	X	X		X

Table 3-4: Server Operations for Taxonomy Nodes (Continued)

Server Operation	Project Name	Language	Categorizer or Concepts Extractor	Top	Category or Concept
Revision Log			X		X
Revert to Older Version					X
Import Category from Repository				X	X
Import Concept from Repository				X	X
<b>Note:</b> The server operations, <b>Upload Test Files</b> and <b>Download Test Files</b> are available only in the <b>Server</b> menu.					

## 3.6 The Miscellaneous Windows

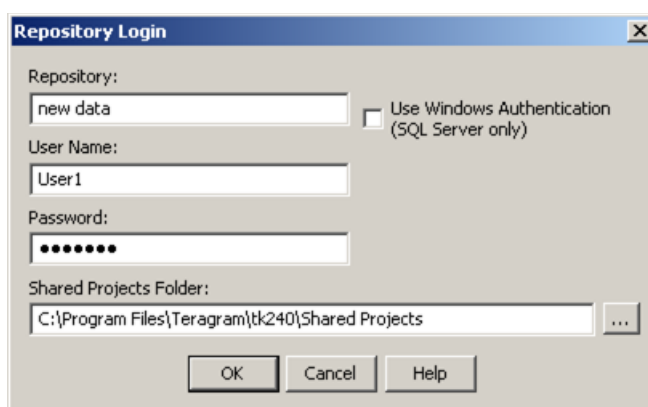
### 3.6.1 The Repository Login Window

Log in to the server before you work on a collaborative project. If you do not log in and you select any of the following operations from the **File** drop-down menu, the Repository Login window automatically appears:

- **Open Remote Project**
- **Remove Project from Server**
- **Repository Login**
- **Change Repository Password**

---

Display 3-4: 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.
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.
4. Enter your password into the **Password** field.

---

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

---

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 3.6.2 *The Select a Directory Window* below.
6. Click **OK**.



---

## 3.6.2 The Select a Directory Window

Use this section in cases where your operating system restricts what can be stored in the Program Files directory. For example, use this section if you are running Windows Vista/2008 Server/7.

In any of these cases, use the Select a Directory dialog box to locate either a remote project, or a cached copy of a SAS Content Categorization Studio project.


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

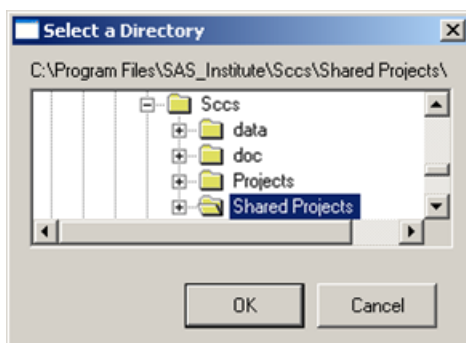
1. Select **File --> Open Remote Project**. The **Repository Login** window appears. See Figure 3-4 on page 34.
2. Use Step 1 to Step 4 on page 34.

---

**Hint:** These fields are automatically filled in for you, unless this is the first time you are using this window.

---

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



4. Select the Shared Projects folder or a cached project in another location on your machine.
5. Click **OK**.

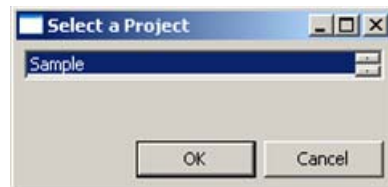
---

### 3.6.3 The Select a Project Window

Use the Select a Project window to locate a project on the server. You can open the project on your local machine, or delete the project.

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

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



2. Highlight the project in the Select a Project window that you want to open. For example, select `Sample`.
3. Click **OK**.

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 open and running on your machine during the removal process.

---

3. Click **Yes**.

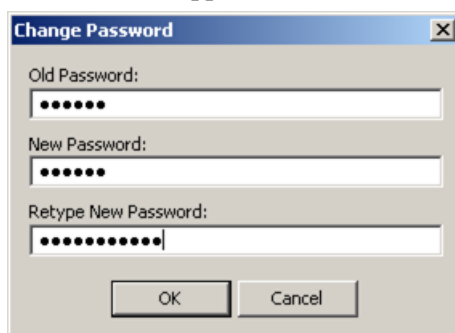
---

### 3.6.4 The Change Password Window

Use the Change Password window to specify a new password.

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

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



2. Enter the password assigned to you by the database administrator 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**.

### 3.6.5 The Upload Test Docs Window

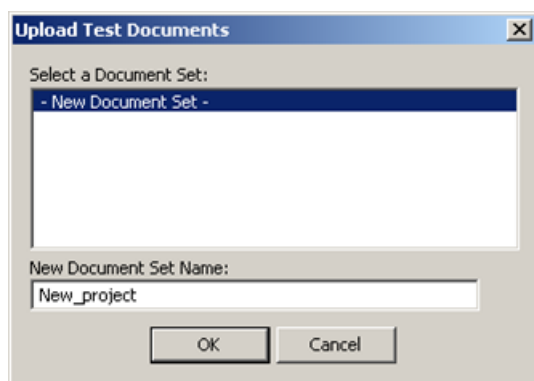
Testing documents that are stored on the server enable other subject matter experts to test their rules and definitions against the same set of texts that you use. Other users can also validate the results that you see.

A user who has, at a minimum, Read, Write, and Change Taxonomy permissions can use the Upload Test Documents window. This operation enables the user to transfer a set of testing documents from a local machine to the server. The SAS Content Categorization Studio project is accessible to other permissioned users here. This subject matter expert can also use the Download Test Documents window to access test files uploaded to the server by another user.

---

To open and use the Upload Test Documents window, complete these steps:

1. Open the Data window and set the path to the testing documents using the **Testing Path** field, unless it is already entered.
2. Open the Testing window and select **Server --> Upload Test Files**. The Upload Test Documents window appears.



Any uploaded sets of documents are listed under the **Select a Document Set** heading. If no testing documents have been uploaded to the server, you see - New Document Set -.

3. Type the name of the new set of texts that you are uploading into the **New Document Set Name** field.
4. Click **OK** and a SAS Content Categorization Collaborative Server confirmation window appears.



5. Click **OK**.

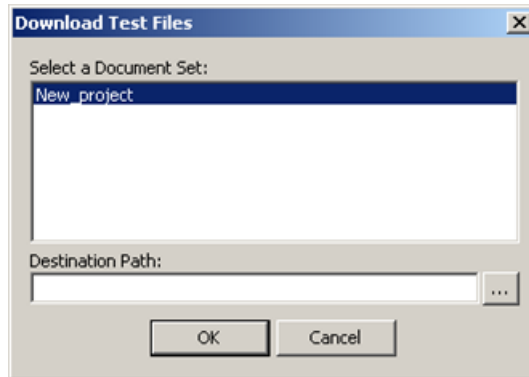
---


### 3.6.6 The Download Test Docs Window

If you have at a minimum, Read, Write, and Change Taxonomy permissions, you can download testing documents from the server to your machine. Use these files to test your taxonomy. When you use shared test documents, you can validate the testing results of another user.

To open and use the Download Test Files window, complete these steps:

1. Open the Testing window and select **Server --> Download Test Files**. The Download Test Files window appears.



2. Select a file under the **Select a Document Set** heading. For example, choose New\_project.
3. Click  to the right of the **Destination Path** field and the Select a Directory window appears.



4. Select the folder location for the uploaded testing documents.

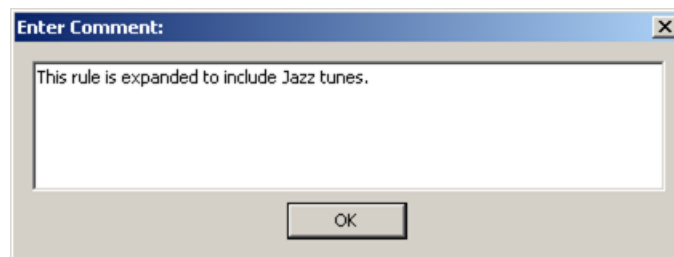
- 
5. Click **OK**. This path appears in the **Destination Path** field.
  6. Click **OK** in the Download Test Files window.

### 3.6.7 The Enter Comment Window

Use the Enter Comment window to write notes about changes to a category or a concept, for either a taxonomy change or a **Server Commit** operation. These notes appear in the RevisionLog windows.

To open and type notes into the Enter Comment window, complete these steps:

1. Make a change to your project at either the taxonomy, or individual node, level.
2. Select **Server --> Commit** and the Enter Comment window appears.



3. Type your comments in the blank field. For example, type *The rule is expanded to include Jazz tunes.*
4. Click **OK**.

Your notes on the changes appear in the RevisionLog windows. For more information, see Section 3.6.8 *The RevisionLog Screens* on page 40.

### 3.6.8 The RevisionLog Screens

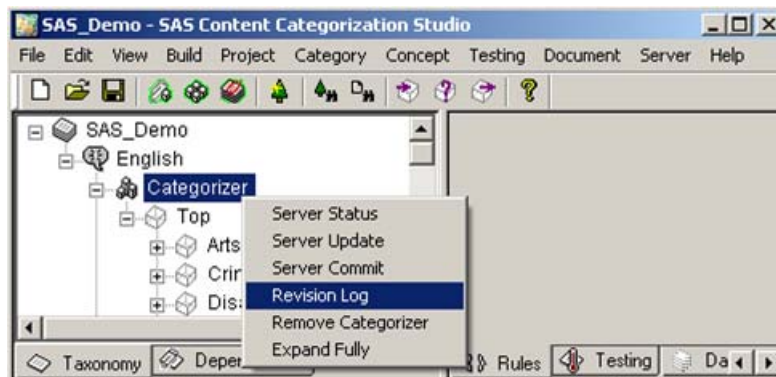
Use both of the RevisionLog screens for reference purposes or to revert to an earlier version of a category or concept. There are two types of RevisionLog screens that make it possible for you to track changes after they are committed to the server:

- Select a taxonomy node and use the revision log operation to see only the rule or definition changes for the selected category and concept.

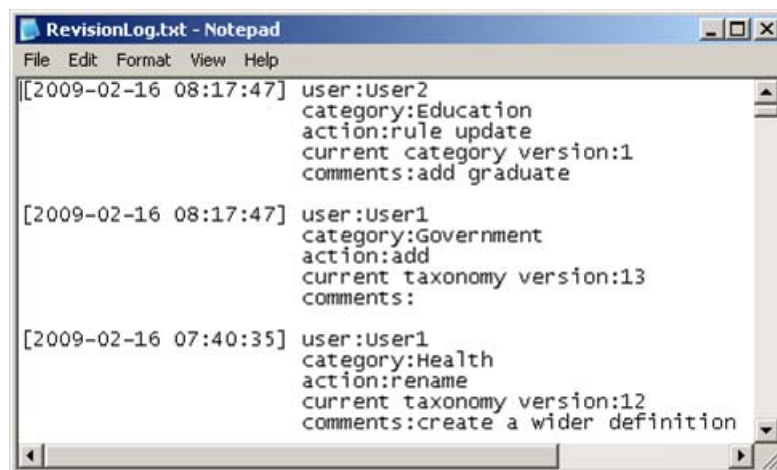
- 
- Select either the `Categorizer` or `Concepts` node. Use the **Revision Log** operation to see the changes to all of the nodes that comprise one branch of the taxonomy. In other words, see the changes for either the categorizer or the concepts branch and only for one language.

In both cases the revision logs are comprehensive. In other words, they track all of the changes, whether for a rule, a definition, or a selected branch of the taxonomy. An example is provided for the `Categorizer` node. Similar results are displayed for the `Concepts` node. If you instead, select an individual category or concept node, only the results for the selected node are displayed.

To open the `RevisionLog` window for a taxonomy branch, right-click on, the `Categorizer` node in the `Taxonomy` window, and select the **Revision Log** operation.



The RevisionLog screen appears displaying the changes made to all of the nodes in the selected taxonomy branch.



Use the components of the RevisionLog screen to learn about the changes made to this branch of the taxonomy:

Table 3-5: RevisionLog Window Components

Component	Description
date and time	Date in YYYY-MM-DD and hh-mm-ss formats.
user	Name of the user making the changes appears here.
category	Name of the category appears here.
action	Type of change that was made appears here.
current taxonomy version	Count of the saved changes made for the referenced node. For example, the Education category rule was updated making the current category version: 1.
comments	Comments entered in the Enter Comment window, if any.



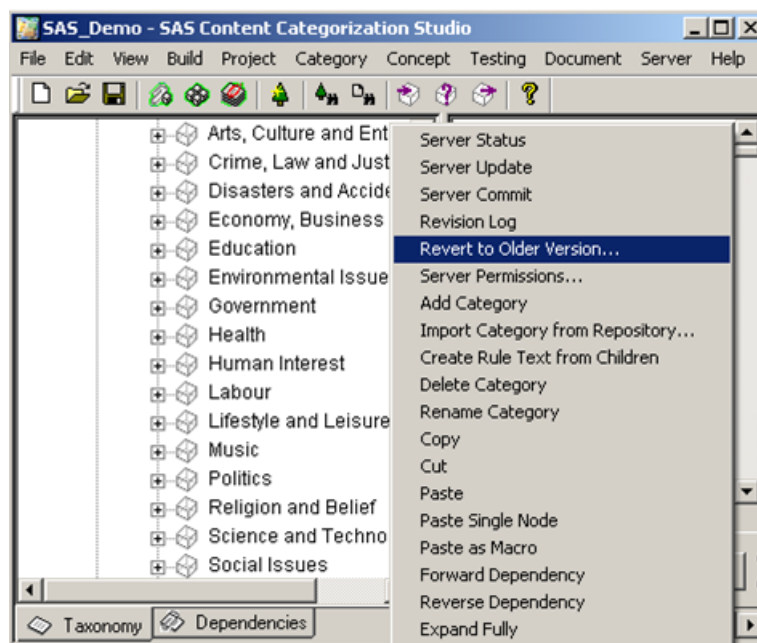
---

### 3.6.9 The Revert to Older Version Window

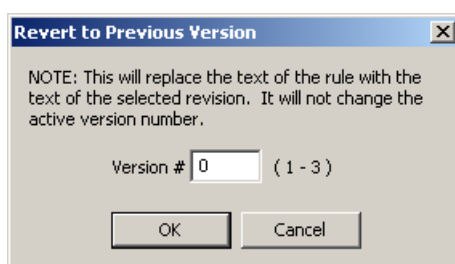
Use the Revert to Older Version window to replace the selected rule or definition with an older version of this rule or definition.

To replace a rule or definition, complete these steps:

1. Right-click on a category or a concept node and select **Revert to Older Version** from the drop-down menu that appears.



The Revert to Previous Version window appears.



2. Enter the version number in the **Version #** field. The numbers in parentheses (O) represent the range of changes made to the selected node. For example, these numbers indicate the range of versions 1–3.
3. Click **OK**.

### 3.6.10 The Import Category (or Concept) Window

A project, or database, administrator can use the Import Concept or Import Category windows to add children to the selected node. Choose to download concepts and categories from a project on the server into the current project using these import operations. For example, if you created a project with a concept that you also want to use in your current project, use this import operation.

When you import a concept or a category, this node appears as a child of the selected node in the taxonomy. For example, if you want your concept to be a top level concept, select **Top**. If you select a child of the **Top** node, the imported node becomes a child of that node.

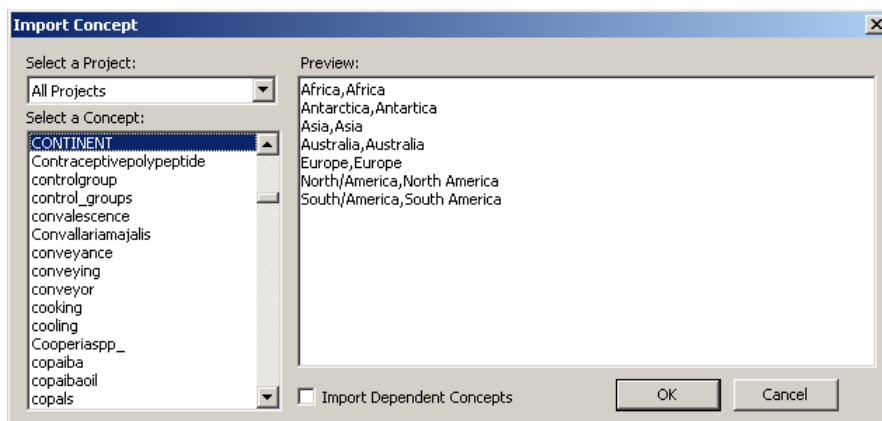
The example provided is for a concept. Adapt these steps to import a category.


To open and use the Import Concept window, complete these steps:

1. Right-click on a node in the concepts branch of the taxonomy. This node becomes the parent of the concept that you import. For example, if you want your concept to be a top level concept, select **Top**.



- 
2. Select **Import Concept from Repository** in the drop-down menu that appears. The Import Concept window appears:



3. (Optional) Click  to the right of the **Select a Project** field and a drop-down menu displaying a list of all of the SAS Content Categorization Collaborative Server projects residing on the server appears. If you select one of these projects, the display of concepts that you can import is limited to this project.

---

**Hint:** By default, select `All Projects` in the **Select a Project** field is selected. This is true although `All Projects` does not appear in the **Select a Project** field. This operation enables you to see all of the concepts for all of the SAS Content Categorization Studio projects on the server.

---

4. See the definition for the selected concept in the **Preview** screen. If there is an error with the database see `Error receiving preview for concept`.
5. Click the **Import Dependent Concept** check box, located below the **Preview** pane, to import any dependent concepts for the selected concept.
6. Click **OK**.

---

### 3.6.11 An Example of Status Windows

When you perform an operation that has significant consequences, a SAS Content Categorization Studio confirmation window might appear. For example, after you delete the permission level for a user, a SAS Content Categorization Studio confirmation window appears.



To confirm this operation, click **Yes**.

---

# 4

## Getting Started

---

- *How to Begin Collaborative Work*
- *Open a Project*
- *Keeping Projects Up-to-Date*
- *Understanding Your User Permission Level*
- *Working with a Cached Project*
- *Collaborative Changes*
- *Server Operations*
- *Save a Project*

### 4.1 How to Begin Collaborative Work

This chapter explains how to set up, and begin working within, a SAS Content Categorization Collaborative Server project. Use this chapter after you read Chapter 3: *Using the Interface* and before Chapter 5: *Other Collaborative Operations*.

The process of two or more developers working together on one project, whether their permission settings are equivalent or different, is defined as *collaboration*. A SAS Content Categorization Studio project that is used for collaborative work, resides on a remote server. This server enables multiple subject matter experts to work together on the same project that is cached on their local machines. The project that resides on the server is called a *remote project*.

Multiple developers can specify the categories and concepts that comprise a taxonomy, write, and modify category rules and concept definitions, while revising a project. By committing changes from local machines to the server, and updating the local project from the server, the project is kept up-to-date.

---

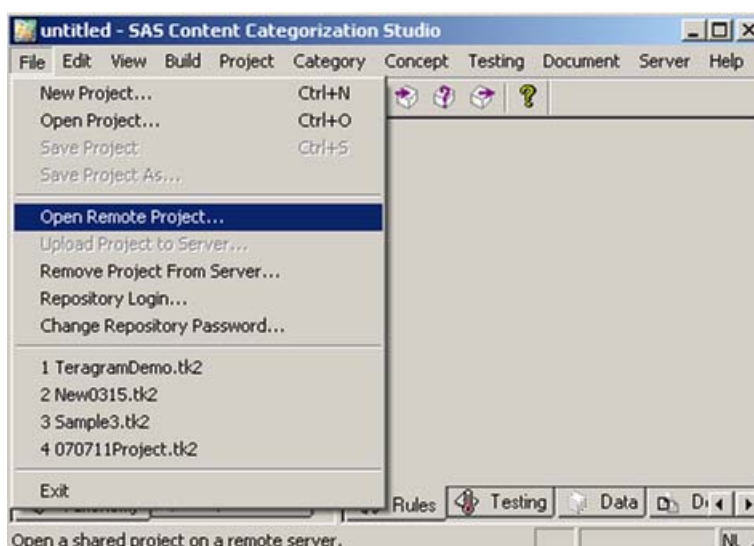
## 4.2 Cached Project

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 Content Categorization Studio. When you log in to the repository, you open your local version of the selected project. You can simplify the process of synchronizing the local project with the version stored on the server by using the installation-specific settings that are available in the Options window. For more information, see Section 3.4 *The Options Window* on page 29 and Section 4.6 *Working with a Cached Project* on page 54.

## 4.3 Open a Project

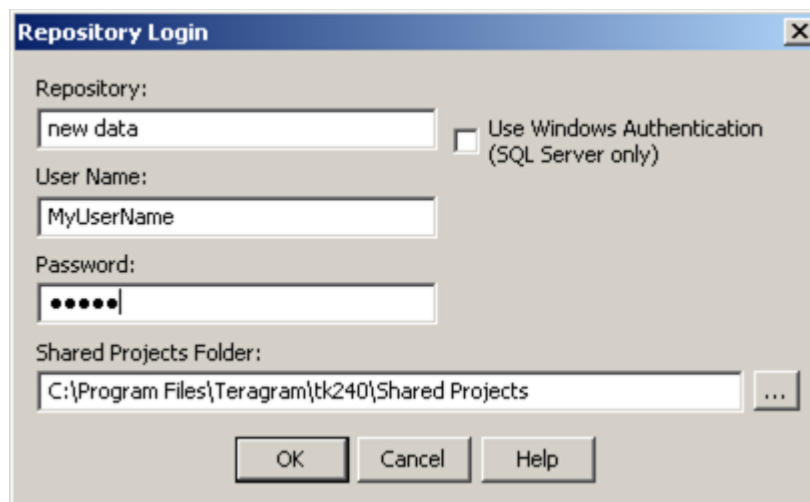
To begin working in your collaborative projects, complete these steps:

1. Select **Start --> Programs --> SAS Content Categorization Studio --> SAS Content Categorization Studio**.



The untitled SAS Content Categorization Studio user interface appears.

- 
2. Select **File --> Open Remote Project**. The Repository Login window appears.




1. Enter the name of the ODBC data source into the **Repository** field.
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.
4. Enter your password into the **Password** field.

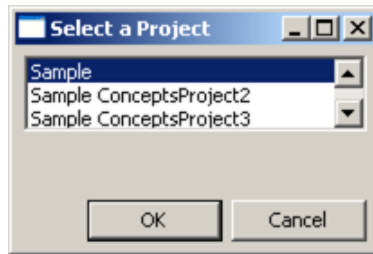
---

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

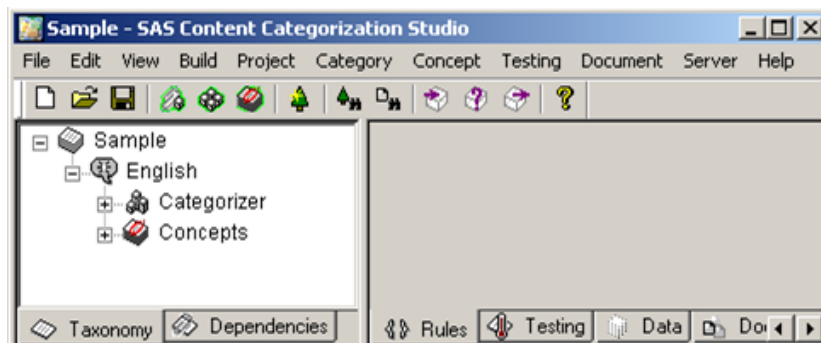
---

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 3.6.2 *The Select a Directory Window* below.

- 
- Click **OK**. The Select a Project window appears.



- Select a project. For example, choose Sample.
- Click **OK**. The selected project appears in the interface.



## 4.4 Keeping Projects Up-to-Date

### 4.4.1 Overview of Keeping Projects Up-to-Date

You can keep your local project up-to-date with the project on the server. You can either automate this process, or choose to perform these operations manually.



---

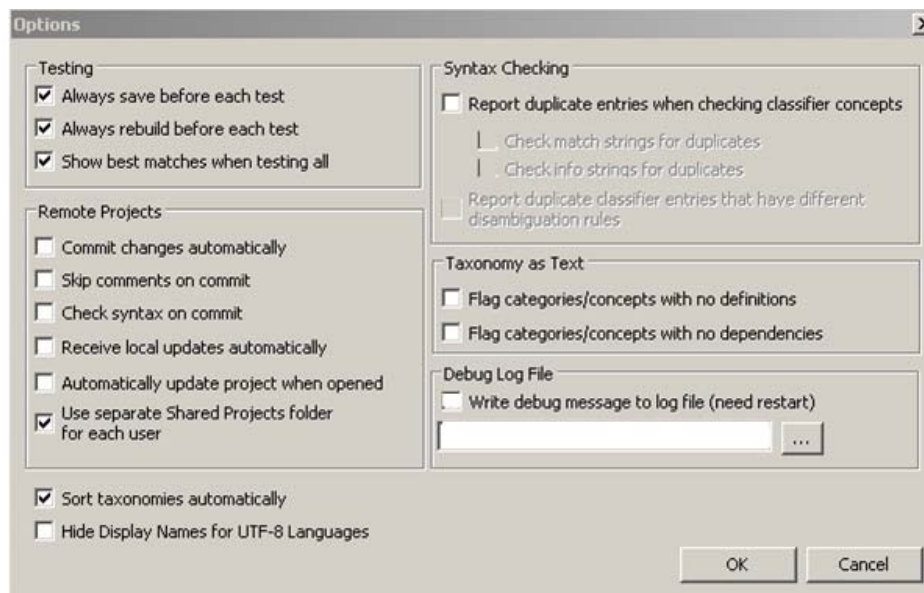
## 4.4.2 Specify Options

The Options window enables you to set installation-wide settings for collaborative work. Use any of the six **Remote Projects** selections in this window to simplify the process of keeping your local project up-to-date with the project on the server. These operations are automatically performed by SAS Content Categorization Studio to save you time.

When you make choices using the Options window, they remain selected as the default operations until you uninstall and reinstall SAS Content Categorization Studio or make new selections. Options are installation-specific.

To specify your installation-specific options, complete these steps:

1. Select **Edit --> Options**. The Options window appears.



2. Select any of the following collaborative operations that are listed under the **Remote Projects** heading:

**Commit changes automatically**

---

Automatically commit rule and definition changes to the project on the server. This process occurs after you edit a rule or definition and select a different taxonomy node.

**Skip comments on commit**

Commit changes to the server without tracking the reasons for these revisions.

**Check syntax on commit**

Automatically verify the syntax of the definition or rule before the syntax is committed to the server.

**Receive local updates automatically**

Automatically update the category rules and concept definitions from the server whenever a change is committed by another user.

**Automatically update project when opened**

Ensure that any changes made to the taxonomy, categories, and concepts are automatically downloaded to your project when it is opened.

**Use separate Shared Projects folder for each user**

(Default) Enable two or more users to use the same machine to build projects in SAS Content Categorization Studio. These users can maintain separate local (cached) projects. These projects are stored in the Shared Projects folder that is created on your local machine.

**3. Click OK.**

You can also perform some of these operations manually. For more information, see Section 4.4.3 *Manually Accessing Server Operations* below.

### 4.4.3 Manually Accessing Server Operations

Manually use the server operations that are available when you right-click on the nodes in the Taxonomy window. Similar operations are also available in the **Server** drop-down menu, but because they are located in the **Server** menu, the word *server* is dropped:

**Server Status**

See the status of your taxonomy.

**Server Update**

---

Obtain any changes committed by other users and overwrite any local changes made to a single category, or concept, by using its counterpart on the server.

**Server Commit**

Update the category rule or concept definition on the server with the local rule or definition that you develop.

**Revert to Older Version**

Select an earlier version of the rule or definition and restore this statement as the current category rule or concept definition.

## 4.5 Understanding Your User Permission Level

The key to successful collaboration is keeping each user's cached project current with the changes that are made according to the expertise of other users.

Permissions ensure that all users have access to the project that are consistent with their level of expertise. By default, all users are assigned Read-Only permissions. After you are assigned a higher permission level by the administrator, you can work on the project.

The levels of permissions are ordered from least to most access:

**Read Only**

Users can see the SAS Content Categorization Collaborative Server project. These users are unable to make any changes to the project.

**Read and Write Rules**

Subject matter experts can read the project and develop category rules and concept definitions for the taxonomy.

**Read, Write, and Change Taxonomy**

Users see the project, develop rules and definitions, and modify the taxonomy.

For more information, see Section 4.5 *Setting Permission Levels in an Open Project* on page 56.

---

## 4.6 Working with a Cached Project

After a project is uploaded to the server and opened on your machine, a local copy is automatically stored in the Shared Projects folder on your hard drive. This cached project is the project that appears in the user interface. For example, the Sample project shown in Step 8 on page 50.

The cached version of the SAS Content Categorization Collaborative Server project contains the changes that you make to the project, unless you select another operation or cache location.

As you modify the project, it is important to keep your local project up-to-date with the project on the server. For more information, see Section 4.8 *Server Operations* on page 56.

However, there are reasons for saving changes only to the local project. To see an overview of these purposes for the Save As operation, see Section 4.9 *Save a Project* on page 75.

## 4.7 Collaborative Changes

### 4.7.1 Overview of Collaborative Changes

When you are working on a collaborative project, each developer ensures that all of their changes are stored in the project repository. This makes it possible for other developers, working on the same project, to access the updated project. The two major types of changes stored in the repository are taxonomy and syntax changes:

Taxonomy changes are defined as the addition, deletion, and renaming of categories and concepts. These changes are automatically committed to the server. However, they are permitted only if your cached project has a taxonomy structure that is identical to the project on the server.

Syntax changes are defined as the changes that are made to a category rule or to a concept definition. These changes are committed to the server either singularly, meaning one rule at a time, or collectively where all of the changed rules are committed at the same time.

---

## 4.7.2 Modifying the Taxonomy

The taxonomy structure is modified whenever you choose to add, delete, or rename a category or a concept. Each of these taxonomy changes requires the server to be up-to-date. Only users with the `Read`, `Write`, and `Change Taxonomy` permission level can perform these operations. If the taxonomy is not up-to-date, update the local project that is cached on your machine. For more information, see Section 4.8.3 *Update from the Server* on page 63.

---

**Note:** The symbolic links component is limited to SAS Content Categorization Studio projects that are not collaborative. You can, however, create dependencies (and dependent concepts can be uploaded with the imported concept) for SAS Content Categorization Studio collaborative projects. For more information, see the *SAS Content Categorization Studio: User's Guide*.

---

When you make taxonomy changes, these node revisions are automatically committed to the taxonomy.

## 4.7.3 Changing the Rules or Definitions

Modify rules or definitions for categories and concepts, respectively, if you are a user with `Read` and `Write Rules` permissions. Unlike taxonomy changes, you manually commit these modifications to the server. For more information, see Section 4.8.4 *Commit Changes to the Server* on page 65. However, you cannot change concept types. For example, you cannot make a classifier into a grammar concept.

When you select **Receive local updates automatically** in the **Remote Project** section of the Options window, the server automatically updates your category rules and concept definitions. This automatic operation is performed whenever a change is committed to the server by another user.

---

## 4.8 Server Operations

### 4.8.1 Understanding the Server Operations

As you begin developing and making changes to your project, you use the server operations to keep your local project up-to-date with the project residing on the server.

These commands check the status of your local copy against the project residing on the server. You can update your project, commit your changes to the server, see detailed revision logs, and return to an earlier version of the project using these operations. You can also use server operations to prevent conflicting changes that might corrupt the collaborative project. Like permissions, server operations can affect one node, or the entire project.

When you save your project, a cached version of the current project is saved only to your local machine. It is stored in the Shared Projects folder on the local machine. Multiple Shared Projects folders allow several developers to create separate projects on this machine.

The operations for server operations are listed in the table below:.

Table 4-1: Server Operations

Menu Operation	Description for Category and Concept Level Operations	Description for Taxonomy or Branch Level Operations
Server Status	<p>See the relationship between the local and server nodes:</p> <ul style="list-style-type: none"> <li>- <b>Up to date:</b> The local version is identical to the project on the server.</li> <li>- <b>Local Changes:</b> The cached, or local, copy has changes that are not committed to the server.</li> <li>- <b>Out of date:</b> Another user made syntax changes to the server project.</li> <li>- <b>Conflict:</b> The syntax for the selected node has been changed at both the local and server levels.</li> <li>- <b>Deleted:</b> This node is not in the server project.</li> <li>- <b>OK{ }:</b> The parent node is up-to-date with the project on the server, but one or more (the number in the curly braces { }) of the child nodes might not be up-to-date.</li> </ul>	<p>See the correlation between your cached project, or the selected branch of your taxonomy, and the project on the server. This operation also checks to see whether a language was added to the project on the server. The nodes in the taxonomy, or selected branch, display one of the following messages:</p> <ul style="list-style-type: none"> <li>- <b>OK:</b> This message appears for either of the top two nodes when these nodes are synchronized with the project residing on the server.</li> <li>- <b>Up to date:</b> The cached and server versions are identical.</li> <li>- <b>Local Changes:</b> Your local project has changes that are not committed to the server.</li> </ul>
Server Update	<p>Update a single category or concept with its counterpart on the server. This operation deletes any local changes. A SAS Content Categorization Studio status window gives you the opportunity to reject the download. When the update is complete, a confirmation message appears in the Taxonomy window.</p>	<p>Download the current project in its entirety, or as one branch, from the server, deleting any local changes. This operation checks to see whether a language was added to the project. If not, the new language is added to the local project. When the update process is complete, an <b>Up to date</b> message appears in the Taxonomy window.</p>
<b>Note:</b> For more information, see Section 4.8.2 <i>Using the Server Operations</i> on page 59.		

Table 4-1: Server Operations (Continued)

Menu Operation	Description for Category and Concept Level Operations	Description for Taxonomy or Branch Level Operations
Server Commit	<p>Commit a changed rule or definition to the server. The <code>Commit Successful</code> string appears when this operation is successful. If the local rule or definition is the same as the one on the server, the <code>Already Up to Date</code> message appears.</p> <p>Unless this operation is deselected in the Options window, the <code>Enter Comment</code> window appears when you commit a change to the server. For more information, see Section 3.6.7 <i>The Enter Comment Window</i> on page 40.</p>	<p>Automatically commit changes to the server, if your project is up-to-date with the project residing on the server. The <code>Commit Complete</code> string appears to the right of the selected <code>Categorizer</code> or <code>Concepts</code> node. The <code>Commit Successful</code> string appears to the right of each changed taxonomy node when this operation is successful.</p>
Revision Log	<p>Open a <code>RevisionLog</code> screen displaying the following information for each revision that is relevant to the selected node:</p> <ul style="list-style-type: none"> <li>- date</li> <li>- time</li> <li>- user</li> <li>- category or concept</li> <li>- action taken</li> <li>- current category or concept version number</li> <li>- comments entered when the change was committed</li> </ul> <p>For more information, see Section 4.8.5 <i>Using the Revision Logs</i> on page 67.</p>	



Table 4-1: Server Operations (Continued)

Menu Operation	Description for Category and Concept Level Operations	Description for Taxonomy or Branch Level Operations
Revert to Older Version	<p>Open the Revert to Previous Version window that enables you to select an older version of the project. When you select this operation, the rule text for the selected version of the changes replaces the rule for the selected category or concept. For more information, see Section 4.8.5.C <i>Revert to an Older Version</i> on page 70.</p> <p><b>Note:</b> See the RevisionLog window (see the table row above) to determine the revision number that you want to use to replace your current rule or definition.</p>	not applicable
Import Concept (Category) from Repository	<p>Open the Import Concept, or Import Category, window to copy a concept, or category, making the imported node a child of the selected node. For more information, see Section 4.8.6 <i>Import Categories or Concepts from a Repository</i> on page 72.</p>	<p>Select the Top node and choose this operation to open the Import Concept, or Import Category, window. Select a concept or category and make it a child of the Top node. For more information, see Section 4.8.6 <i>Import Categories or Concepts from a Repository</i> on page 72.</p>

## 4.8.2 Using the Server Operations

### 4.8.2.A Check the Server Status for Single Nodes

Right-click on a node, and use the **Server Status** operation in the menu that appears, to see how the currently selected rule corresponds to the matching rule on the server.

To check the server status, complete these steps:

1. Right-click on an individual category or concept node.

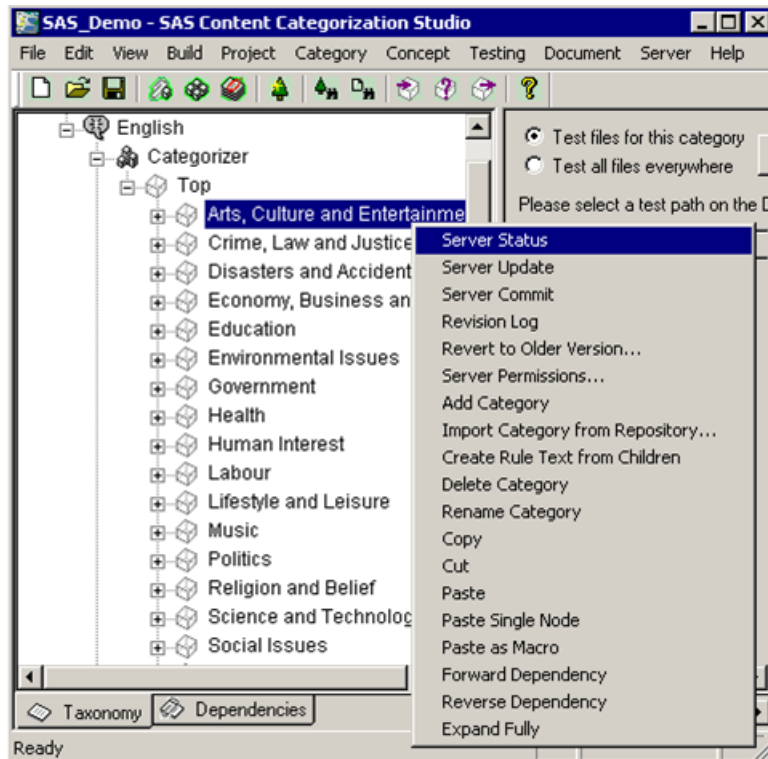
---

**Note:** If you click on the *Categorizer* or *Concepts* node, instead, you are checking the server status for the selected taxonomy branch.

---

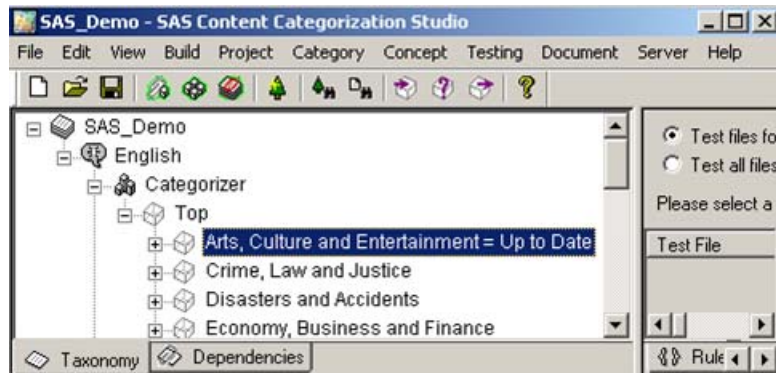
---

## 2. Select **Server Status**.



---

A message appears in the Taxonomy window, to the right of the selected node, displaying the server status for this node.



3. (Optional) If a displayed message reads, Out of Date, Local Changes, or Conflict, there is a difference between the rule in the local project and the server project. When you see one of these messages, you can use either of the following two operations to correct the difference:
  - Overwrite your local changes by updating the rule with the rule on the server. For more information, see Section 4.8.3 *Update from the Server* on page 63.
  - Overwrite the rule on the server by committing your local changes to the server. For more information, see Section 4.8.4 *Commit Changes to the Server* on page 65

For information about removing messages from the Taxonomy window, see Section 4.8.2.C *Removing Taxonomy Tree Messages* on page 63.

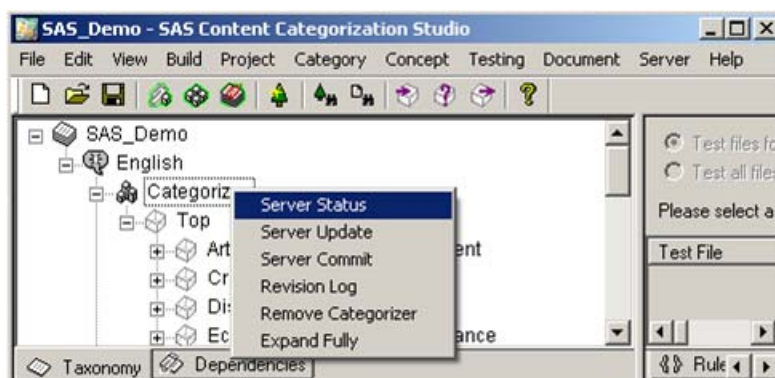
---


### 4.8.2.B Checking the Status of the Taxonomy

Check the status of all of the nodes in a specific branch of the taxonomy against the project on the server at any point during the process of project development.

To check the server status for a node, complete these steps:

1. Right-click on the **Categorizer**, or the **Concepts**, node and select **Server Status** from the drop-down menu that appears.



2. Check the messages that appear to the right of the **Categorizer** or **Concepts** node. If you see the **Deleted** message, the node is removed from the project on the server, but still appears in the taxonomy of the local project.
3. (Optional) To remove these messages from the Taxonomy window, click  .

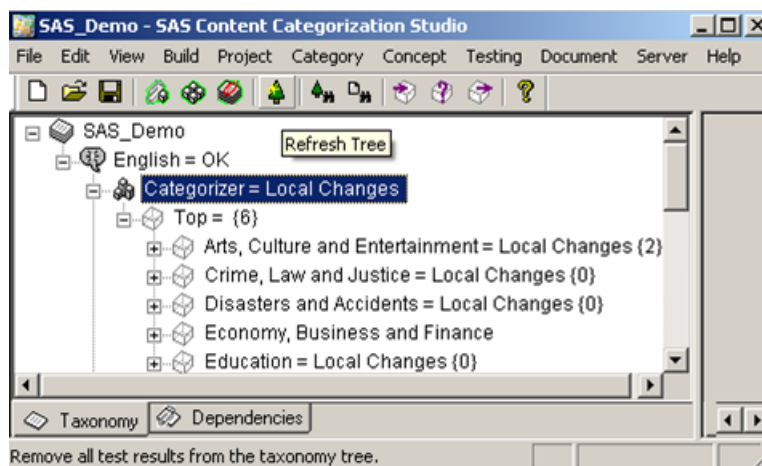
For more information and examples, see Section 4.8.2.A *Check the Server Status for Single Nodes* on page 59.

---

### 4.8.2.C Removing Taxonomy Tree Messages

You can remove the messages that appear in the taxonomy to refresh the tree. To remove the messages, complete this step:

Click  .



---

**Note:** The Revert to Previous Version operation does not change the rule on the server. The current rule text is replaced with the selected rule. To change the server version, a Server Commit operation is performed after the Revert to Previous Version operation.

---

### 4.8.3 Update from the Server

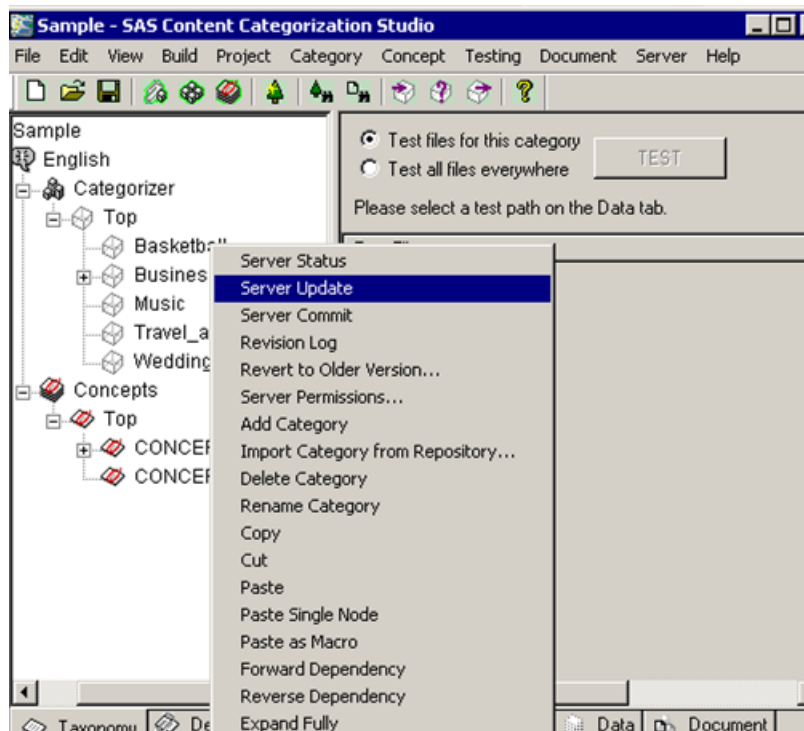
The Server Update operation modifies your local project to synchronize it with the version on the server. Use the Server Update operation when **Automatically update project when opened** is not selected in the Options window.

The Update Server operation is available only for individual taxonomy nodes.

---

To download individual node changes from the server to your local project, complete these steps:

1. Right-click on a category or concept node and select **Server Update**.



A SAS Content Categorization Studio status window appears.



2. Click **Yes**.

You can receive automatic updates from the server for individual categories and concepts. For more information, see Section 3.4 *The Options Window* on page 29.

---

## 4.8.4 Commit Changes to the Server

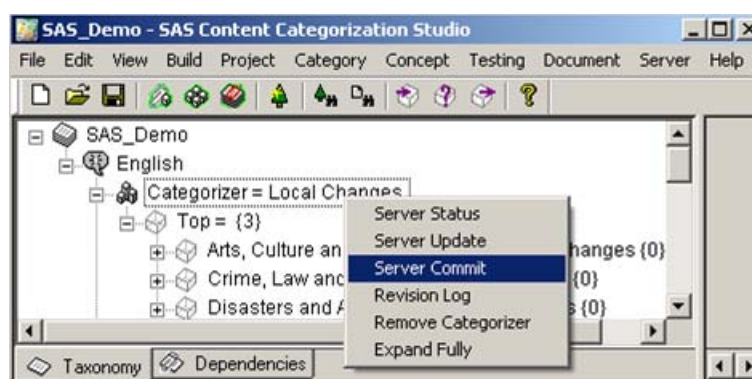
Commit each individual rule change to the server, update all of the changes to a single branch, or commit all of the changes to the project at one time. To perform any of these operations, right-click on a taxonomy node and select **Server Commit** from the menu that appears.

If you right-click on the **Categorizer** or **Concepts** node, you commit all of the changes for the selected branch. Right-click on the language node to commit all of the changes for the project. If you are building a project with more than one language, the changes are committed to the selected language branch. For more information and examples of the message that appear, see Table 3-4 on page 32.

Use this operation if you did not select **Commit changes automatically** in the Options window. For more information, see Section 4.8 *Server Operations* on page 56.

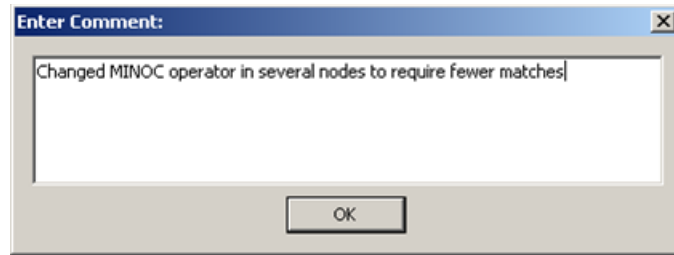
To commit changes to the server, complete these steps:

1. Right-click on a taxonomy node and select **Server Commit**.



---

The Enter Comment window appears, unless you selected **Skip comments on commit** in the Options window.



2. Enter an explanation for the changes. This comment is entered into the RevisionLog window so that all project developers can see the reason for any changes made to the project.
3. Click **OK**. Commit Complete and Commit Successful messages appear for the relevant nodes in the Taxonomy window.





---

## 4.8.5 Using the Revision Logs

### 4.8.5.A Overview of Using the Revision Logs

Use the RevisionLog window to access a history of the changes that were made to a specific taxonomy branch, or to an individual node. The type of RevisionLog window that appears depends on the node that you select in the taxonomy.

#### Taxonomy branch

Select the **Revision Log** operation using the *Categorizer* or *Concepts* node. The RevisionLog screen that appears enables you to see the changes that were made to all of the categories or concepts in this branch of the taxonomy.

#### Individual category or concept node

Select an individual category or concept node that enables you to see the history of the rule or definition changes for the selected node, only. For more information, see Section 4.8.5.B *Revision Logs* on page 68.

You can also use the revision logs to determine the rule or definition version that you want to use to replace the current rule or definition. For more information, see Section 4.8.5.C *Revert to an Older Version* on page 70.

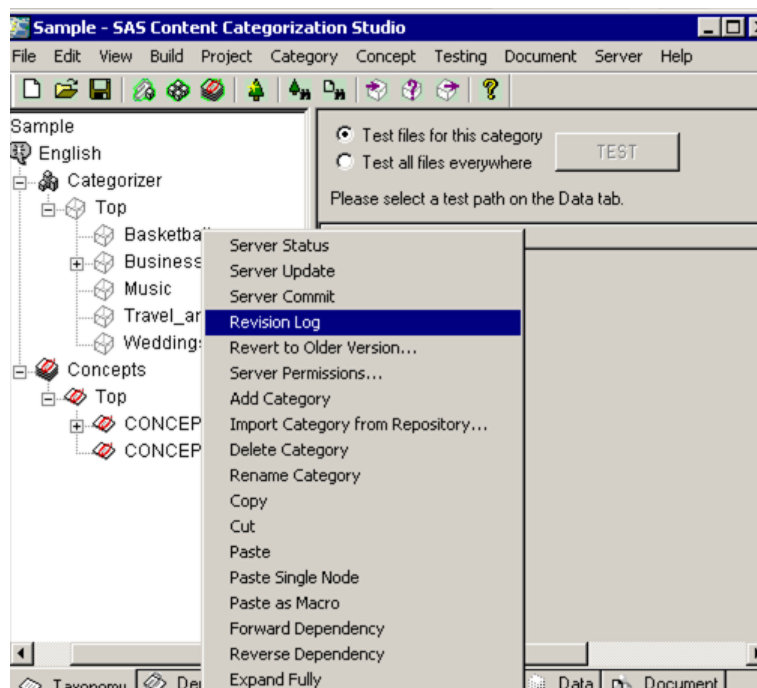
---

#### 4.8.5.B Revision Logs

Use revision logs to see the changes made to your taxonomy, or to a selected rule or definition. These screens also enable you to select a version that you can revert to.

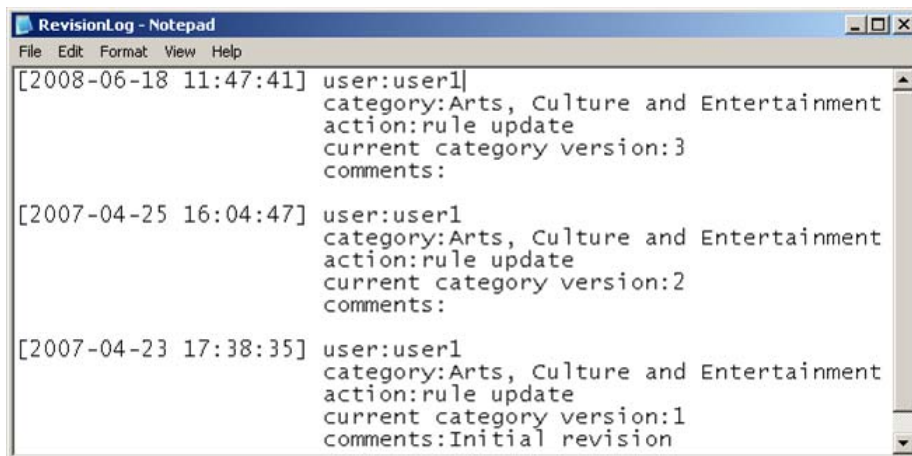
To open and use Revision Log screens for either an individual category or concept, or an entire taxonomy, complete these steps:

1. Right-click on a category or concept node to see only the log for this node. Select **Revision Log**.



---

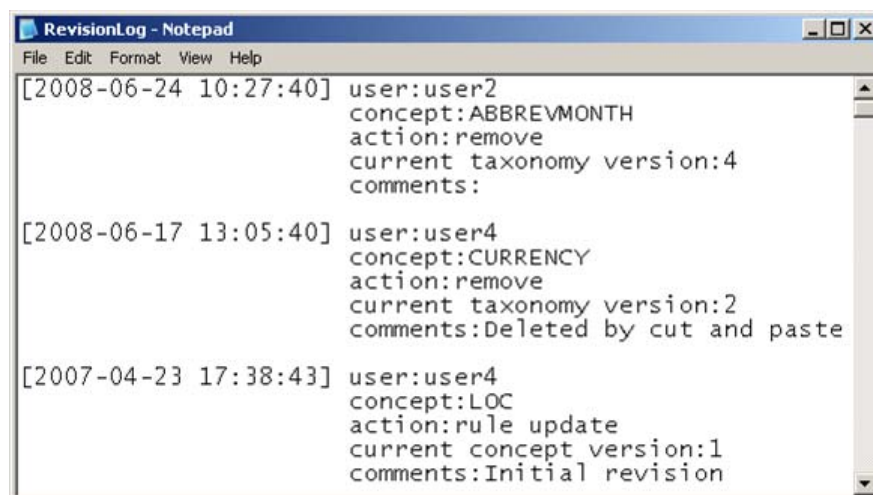
A RevisionLog screen appears.



2. Use Table 3-5 on page 42 to understand the information in this screen.
3. (Optional) If you want to revert to an older version of the rule, note this version number at this time.
4. Right-click on the `Categorizer` or `Concepts` node in the taxonomy to see the RevisionLog screen for all of the concepts or categories in the

---

selected branch. These modifications include the addition and deletion of the nodes in this branch and any changes to the rules or definitions.



5. Use Table 3-5 on page 42 to analyze the information in this screen.
6. (Optional) If you choose to revert to an older version of the taxonomy or the node, note the version number at this time.

#### 4.8.5.C Revert to an Older Version

Return to an older version of the selected category rule or concept definition after you choose a version using the RevisionLog screen. If you want to change several nodes within one branch, work through the changes node-by-node.

To revert to an older version of a rule or definition, complete these steps:

1. Complete Step 1 on page 68 through Step 3 on page 69.
2. Use the RevisionLog window to determine the revision number.



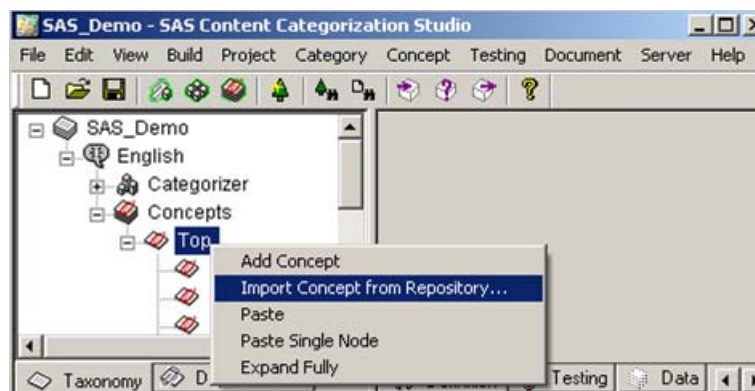
---

## 4.8.6 Import Categories or Concepts from a Repository

Add existing categories and concepts, with their definitions, to your taxonomy using the **Import Concept from Repository** operation, individually, or as a group. Each concept is imported with its definition and placed in your taxonomy as a child of the selected node. This section applies to both categories and concepts. If you are importing categories, modify the following directions as necessary.

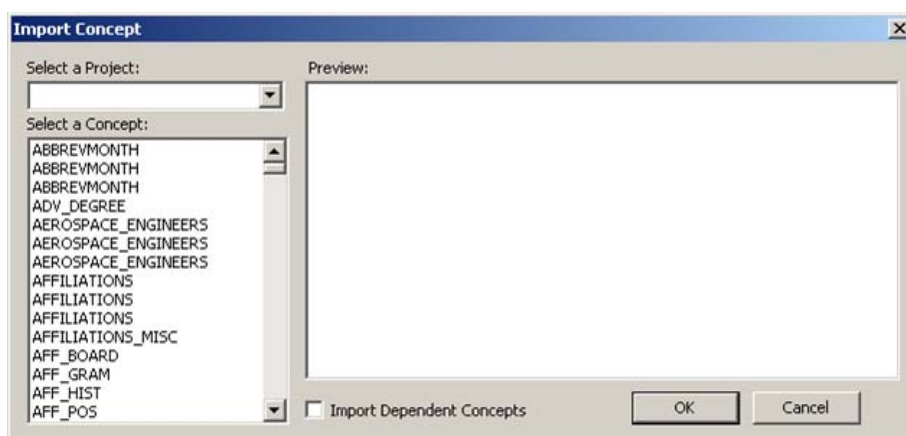
To import a single concept, or a group of concepts, complete these steps:

1. Right-click on the **Top** node, or select another concept node. This node is the parent of the imported concept. Select **Import Concept from Repository**.



---

The Import Concept window appears.



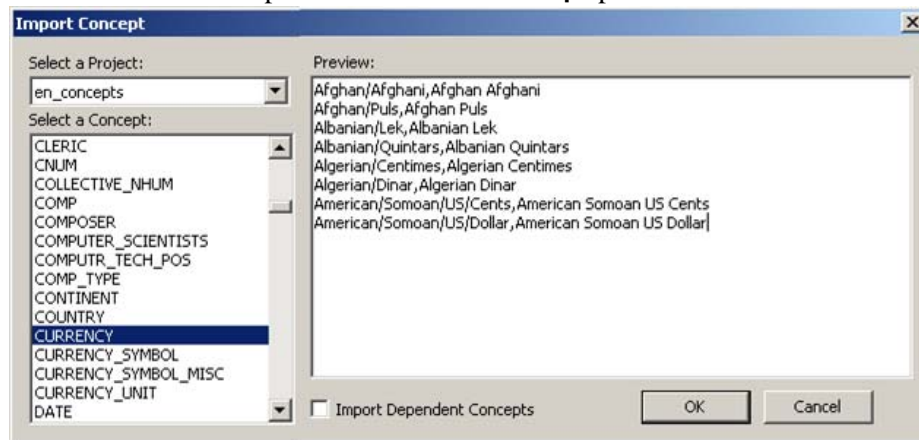
2. (Optional) Click ▼ to the right of the **Select a Project** field. Select a project to limit the display of concepts to those contained in that project. Select a project in the drop-down list that appears in the **Select a Project** field.

---

**Hint:** By default, all of the concepts for all of the projects on the server are displayed. This is true although `ALL concepts` is not displayed.

---

- 
3. Select a concept in the **Select a Concept** pane.



4. Check the appropriateness of the definition for the selected concept in the **Preview** screen.
5. (Optional) Select **Import Dependent Concept** if dependent concepts appear in this definition and you also want to import these concepts.

---

**Note:** The Import Dependent Concept operation does not apply to categories.

---

6. Click **OK**.
7. (Optional) Use these steps reiteratively to add multiple concepts to your project.



---

## 4.9 Save a Project

Save a project to your local machine by selecting **File --> Save Project**. When you make this selection, SAS Content Categorization Studio saves a copy of this project to the Shared Projects folder on your hard drive.

When making changes to a collaborative project, you should commit your changes to the server. However, in some circumstances, it might be better to save your project locally:

- Make changes offline
- Test syntax changes for rules and definitions
- Make project demonstrations
- Work off-site

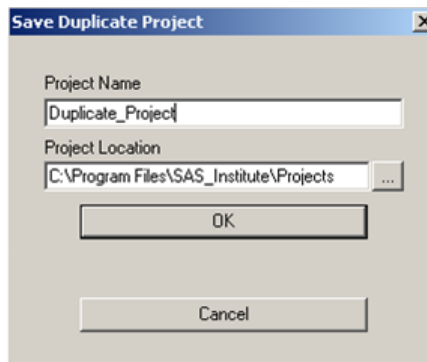
---

**Note:** When you choose to save your project locally, your changes are also committed to the server. To save your changes to the local machine only, use the Save As operation.

---

To save a project to your local machine, complete these steps:

1. Go to **File --> Save Project As** and the Save Duplicate Project window appears.



2. Enter the name of the duplicate project into the **Project Name** field.

- 
3. Enter the path to the location of your duplicate project into the **Project Location** field.
  4. Click **OK**.

When you select **File --> Save As**, the collaborative features for your SAS Content Categorization Studio project disappear.

---

# 5

## Other Collaborative Operations

---

- *Overview of Other Collaborative Operations*
- *Uploading or Downloading Testing Files*
- *Change Your Server Password*
- *Remove a Project from Server Window*

### 5.1 Overview of Other Collaborative Operations

This chapter provides information about collaborative operations that you can use as you build a project. These operations are not necessary to build a successful project, but they do facilitate collaborative development.

### 5.2 Uploading or Downloading Testing Files

#### 5.2.1 Overview of Shared Test Files

Sharing testing files enables multiple users who are working on different, or the same areas, of a taxonomy to see how category rules perform against the same input texts. Use the same set of testing files to test your category rules as another developer.

Identify rules that should be changed and documents that should fail, but might not. For example, a document on rose bushes should not match a rule for Rose Kennedy.

The operations to upload and download testing files to a server are restricted to users who have, at a minimum, *Read, Write and Change Taxonomy* permissions. With this access level, you can upload a set of testing files, within

---

their taxonomy structure, to the server where they can be downloaded by other permissioned users.

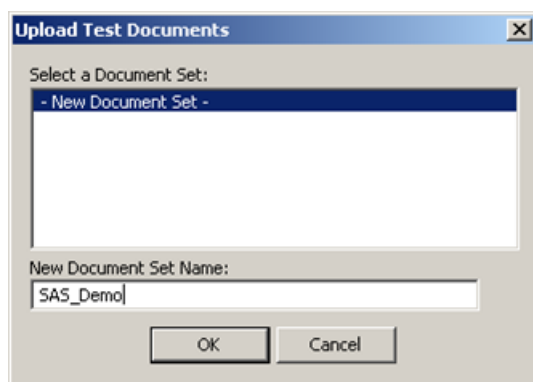
## 5.2.2 Before Uploading or Downloading Testing Files

Before you can use the upload and download operations, initialize the document repository where sets of testing documents can be associated with a project and stored. For more information, see Section A.4.4 *Initialize the Testing Document Repository* on page 76.

## 5.2.3 Upload Testing Files

To upload a set of testing files to the server, complete these steps:

1. Load a set of testing files into your SAS Content Categorization Studio project, by setting the **Testing Path** in the Data window. For more information, see *SAS Content Categorization Studio: User's Guide*.
2. Select **Server --> Upload Test Files** to open the Upload Test Documents window.



Type the name of the testing directory to upload into the **Select a Document Set Name** field.

- 
3. Click **OK** and a SAS Content Categorization Studio status window appears.



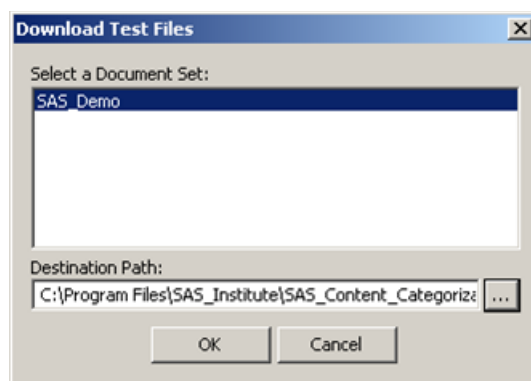
4. Click **OK**.

## 5.2.4 Download Testing Files

After a set of testing files has been uploaded to the server, a user with a minimum of *Read, Write, and Change Taxonomy* permissions can download this set.

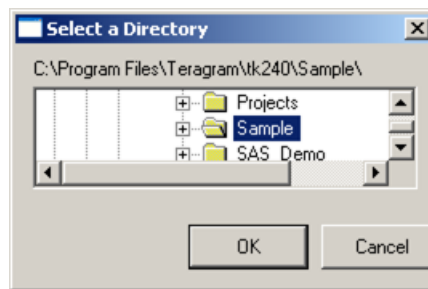
To download a set of testing documents, complete these steps:

1. Select **Server --> Download Test Files**. The Download Test Files window appears.



2. Select a testing folder in the **Select a Document Set** pane.

- 
3. Click  to the right of the **Destination Path** field. The Select a Directory window appears.



4. Select the folder to store these documents on your local machine.
5. Click **OK**. The path to the selected directory appears in the **Destination Path** field.
6. Click **OK**.

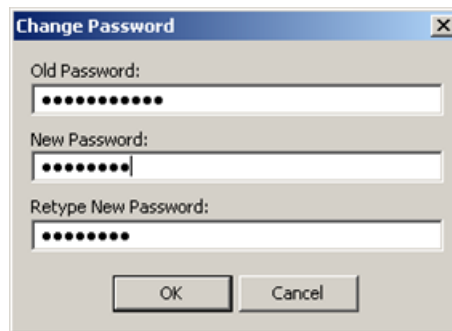
---

## 5.3 Change Your Server Password

You can change your server password during at any time during project development.

To change your password, complete these steps:

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



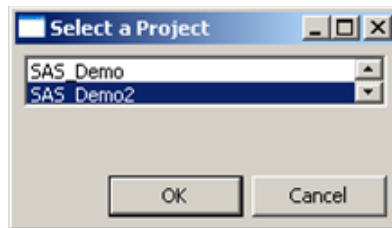
2. Enter your existing password into the **Old Password** field.
3. Enter your new password into the **New Password** field.
4. Enter your new password into the **Retype New Password** field.
5. Click **OK**.

## 5.4 Remove a Project from Server Window

When you remove a project from the server, you permanently delete this project. This is true unless a copy of the deleted project is stored on your workstation. The Remove Project from Server operation is irreversible.

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 open.



4. If you are sure that you want to remove the project, click **OK**.



5. Click **OK**.



---

# Reference Section

---

- Appendix A: *Recommended Reading on page 85*
- Appendix B: *Glossary on page 87*



---

# Appendix: A

## Recommended Reading

---

The following books are recommended:

- *SAS Content Categorization Collaborative Server: Administrator's Guide*: Create the data sources, upload projects, and specify users' permissions.
- *SAS Content Categorization Studio: User's Guide*: Create a SAS Content Categorization Studio project, test, and upload to SAS Content Categorization Server.
- *SAS Content Categorization Studio: Installation Guide*: Install SAS Content Categorization Studio.
- *SAS Content Categorization Studio: Quick Start Guide*: Quickly set up a Categories-only SAS Content Categorization Studio project.
- *SAS Content Categorization Server: Administrator's Guide*: Understand how SAS Content Categorization Server applies the `.mco` and `.concepts` files to input documents.
- *SAS Contextual Extraction Studio: User's Guide*: Build complex definition that include different types of rules. Use Boolean operators and regular expressions.
- *SAS Contextual Extraction Studio: Installation Guide*: Install SAS Contextual Extraction Studio.
- Use the language books for each language purchased to see the comprehensive list of part-of-speech tags that are available for grammar concepts.

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](http://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](mailto:sasbook@sas.com)

Web address: [support.sas.com/pubs](http://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: B

## 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**

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

**classifiers**

specify 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 working together 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*.

**developer**

defines a permissioned user who is often referred to as a project developer. Also see *User*.

**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**

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

---

# Index

---

## A

Already Up to Date	
status message .....	58
automatically update	
project .....	31, 52

## C

cached version	
defined .....	56
category	
add .....	55
changes .....	54
delete .....	55
rename .....	55
update .....	53, 57
category rule	
modify .....	47
status .....	59
change	
repository password .....	81
Change Password interface	
usage .....	37, 81
Change Repository Password	
File .....	26
changes	
category .....	54
collectively committed .....	54
concept .....	54
local .....	57
overwrite .....	29
server .....	54
singularly committed .....	54
taxonomy .....	54
types .....	54

---

check syntax	
on commit .....	31, 52
collaboration	
defined .....	47
Commit	
message .....	27
commit	
check syntax .....	31, 52
commit changes	
automatic .....	30, 51
Commit changes automatically	
usage .....	65
Commit Successful	
status message .....	58
concept	
changes .....	54
concept definition	
modify .....	47
status .....	59
Conflict	
taxonomy message .....	61
Connector/ODBC window	
usage .....	13
Create a New Data Source to SQL Server window	
open .....	15, 16, 17
create data sources	
credentials .....	12
credentials	
create data sources .....	12

## D

data source name	
repository login .....	23
specify .....	13
definition	
change .....	58
commit .....	58
Deleted	
taxonomy message .....	62



---

download	
test documents .....	79
Download Test Docs window	
usage .....	37
Download Test Files	
command .....	28
permission level .....	77
Server .....	28
Download Test Files window	
open .....	79

## E

edit	
command .....	26
options .....	26
Enter Comment window	
open .....	29
usage .....	40

## F

File	
Change Repository Password .....	26
command .....	26
Open Remote Project .....	26
options .....	26
Remove Project From Server .....	26
Repository Login .....	26

## H

Help	
command .....	28
options .....	28
history	
rule changes .....	67

---

## I

Import Category from Repository	
node level .....	33
Import Category window	
usage .....	44
Import Concept from Repository	
node level .....	33
usage .....	59, 72
Import Concept window	
open .....	73
usage .....	44, 59
Import Dependent Concept option	
usage .....	45
Import Dependent Concepts box	
usage .....	74

## L

licensing information	
About option .....	28
Local Changes	
taxonomy message .....	61
local changes	
overwrite .....	61
local update	
Server Update .....	29
log in	
information .....	23

## M

menu bar .....	26
messages	
clear .....	63

---

## O

ODBC Microsoft SQL Server Setup window	
open .....	19, 20
Open Remote Project	
File .....	26
Options window	
usage .....	9, 29, 30, 51
Out of Date	
taxonomy message .....	61

## P

password	
repository login .....	23
permissions	
Category level .....	8
Concept level .....	8
project level .....	8
setting .....	8, 47
Preview screen	
Import Concept window .....	74
project	
automatic update .....	31, 52
remote .....	47
update .....	55
project repository	
project changes .....	54
Project-wide update	
server .....	29

## R

Read and Write Rules	
permission defined .....	53
Read Only	
permission defined .....	53
Read, Write and Change Taxonomy	
permission defined .....	53, 55
upload and download permissions .....	77

---

receive updates	
automatic .....	31, 52, 55
Remove Project From Server	
File .....	26
usage .....	82
repository	
change password .....	81
remote .....	23
Repository Login	
File .....	26
Repository Login window	
usage .....	26, 33
Revert to Older Version	
defined .....	53
node level .....	33
option .....	71
usage .....	59
window .....	27, 43
Revert to Previous Version window	
open .....	43, 71
Revision Log .....	8
node level .....	33
option .....	41, 67
Server .....	27
usage .....	58
RevisionLog window	
components .....	42
options .....	40
usage .....	40
rule changes	
commit .....	54, 58
history .....	67
syntax changes .....	47

## S

SAS Content Categorization Studio	
User's Guide	
location	28
Save Duplicate Project window	
usage .....	75

---

Save Project	
usage .....	75
Save Project As	
usage .....	75
Select a Directory window	
usage .....	35
Select a Document Set	
usage .....	79
Select a Project field	
option .....	45
usage .....	73
Select a Project window	
open .....	36, 82
Server	
command .....	27
Download Test Files .....	28
options .....	27
Revert to Older Version .....	27
Revision Log .....	27
Upload Test Files .....	27
server	
changes .....	54
options .....	32
remote .....	47
up-to-date .....	55
Server Commit	
defined .....	53
node level .....	32
usage .....	29
server operations .....	8
defined .....	56
Server Status	
check .....	62
defined .....	52
messages .....	59
node level .....	32
Taxonomy window .....	60
usage .....	29
Server Update	
defined .....	52
local update .....	29
node level .....	32

---

overwrite changes .....	29
usage .....	64
Shared Projects folder	
cached copy .....	48
project copy .....	75
separate .....	31, 52
Skip comments on commit	
option .....	30, 52
standard toolbar .....	28
Status option	
usage .....	27
syntax changes	
rule changes .....	47

## T

taxonomy	
changes .....	8, 54, 55
window .....	26
taxonomy tree	
modify .....	55
Taxonomy window	
messages .....	66
remove messages .....	61
Test Data Source button	
usage .....	20
testing documents	
download .....	79
upload .....	9
TK240 dating	
About option .....	28

## U

update	
category .....	53, 57
Update option	
usage .....	27
Upload and Download Test Files	
availability .....	33

---

Upload Test Docs window	
usage .....	37
Upload Test Files	
option .....	78
permission .....	77
Server .....	27
username	
repository login .....	23

## V

version	
About option .....	28
number .....	59

