Genesys Knowledge Management User Guide

Genesys Engage Digital (eServices) 8.1.4

12/29/2021

Table of Contents

Genesys Knowledge Management User's Guide

This section describes the use of Genesys Knowledge Management functionalities, including:

Use this section in conjunction with eServices 8.1 Knowledge Manager Help.

The use of Genesys Content Analyzer, an optional enhancement to Knowledge Management, is described in Genesys Knowledge Management: Content Analyzer.

Functionality

Knowledge Management functionalities fall into the following four groups:

- **Categories/standard responses/field codes.** A system of categories, organized in a tree structure, provides the means of organizing *standard responses,* which are prewritten responses to interactions. Field codes provide a way to particularize the standard response to individual interactions. Category trees are also integral to the classification functionality of Genesys Content Analyzer (see the third item in this list). You use Knowledge Manager to create category trees, and to create and edit the standard responses and the field codes that they can contain.
- **Screening rules.** Screening rules perform pattern matching on incoming interactions. The results of the pattern matching are then available for use in subsequent steps in routing and in interaction workflows. You use Knowledge Manager to create and edit the screening rules.
- **Genesys Content Analyzer.** This optional functionality uses natural language processing to analyze incoming interactions and assign them to categories in a category tree. Content analysis uses "models," which are statistical representations of category trees. Models are produced by "training" on a collection of precategorized text objects (e-mails and other types). Knowledge Manager controls the training process and displays information about models. For details see See Genesys Knowledge Management: Content Analyzer.

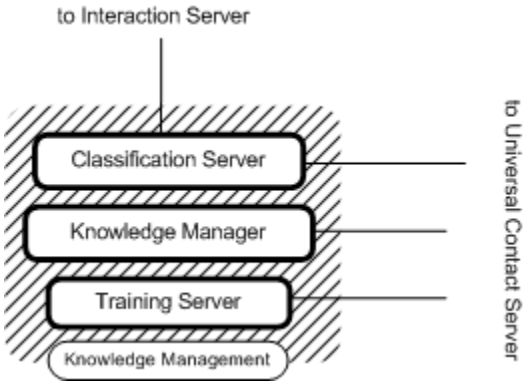
- **FAQ.** With Content Analyzer, you can convert your category structure and standard responses into an FAQ list. You can either post the resulting FAQ list as text on your web site or use it as the source for an automatic question-answering facility.

Components

Knowledge Management consists of the following components:

- **Classification Server** applies screening rules when triggered to do so by a routing strategy. In Genesys Content Analyzer, Classification Server also applies models to categorize incoming interactions. Both screening rules and models are stored in the Universal Contact Server database.
- **Training Server** trains the system to recognize categories. It is active only in the Genesys Content Analyzer.
- **Knowledge Manager** is the user interface.

"Knowledge Management and Other Relevant Components" shows the Knowledge Management components.



Knowledge Management and Other Relevant Components

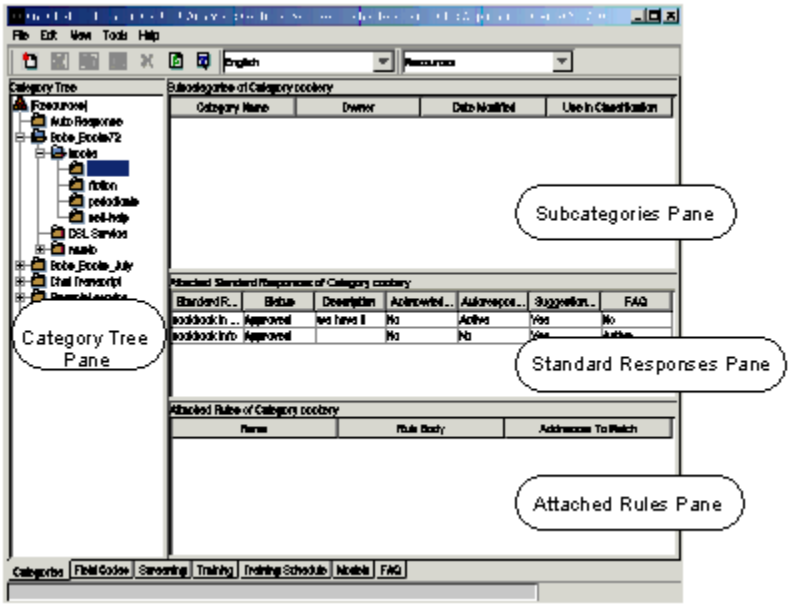
Knowledge Management Overview

This section provides an overview of Knowledge Management, including:

- User Interface
- Introduction to Category Structure
- Using Basic Knowledge Manager Objects

Knowledge Management User Interface

The Knowledge Manager user interface is a window with three (basic Knowledge Manager) or seven (Genesys Content Analyzer) tabs. The figure "Knowledge Manager Window" shows the interface with seven tabs. It also shows the four panes Category Tree, Subcategories, Standard Responses, and Attached Rules.



Knowledge Manager Window

> ## Important
> All Knowledge Manager screenshots in this guide show the Genesys Content Analyzer version. This differs from the version without Genesys Content Analyzer, even on tabs that are common to both. For example, without Genesys Content Analyzer, there is no `Tools` menu.

The table "Knowledge Manager Tabs" briefly describes each tab and what it does. Note that the first three tabs are relevant to all users of Knowledge Manager, whereas the last four tabs are relevant only to users of Genesys Content Analyzer. For details see eServices 8.1 Knowledge Manager Help.

Knowledge Manager Tabs

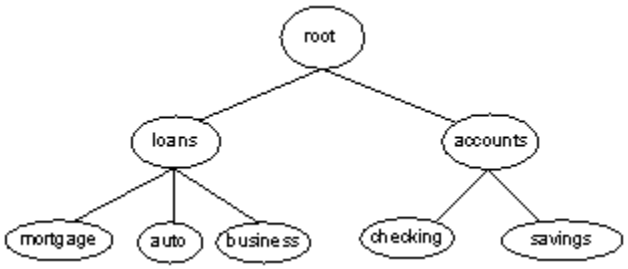
Relevance	Tab	Description
General	Categories	Displays and gives access to category trees and the standard responses and screening rules associated with categories. Provides access to the `Standard Response Editor`.
	Field Codes	Displays the set of field codes and provides access to the `Field Code Editor`
	Screening	Displays the set of screening rules and provides access to the `Screening Rules Editor`
Genesys Content Analyzer only	Training	Displays the set of training objects and provides access to the `Mail Editor`
	Training Schedule	Displays the start time, status, and other information about training objects
	Models	Displays information about models
	FAQ	Displays information about FAQ objects

Important

In most Knowledge Manager windows, lists of items can be sorted by clicking on the column headings. The major exceptions are lists of test results on the `Models` tab. See Testing Models and Using and Rating Models.

Introduction to Category Structure

In general terms, a *category* is a unit of knowledge. Categories are organized in a tree structure; "Example Category Tree" shows an example.



Example Category Tree

Genesys eServices uses category trees in the following ways:

- **Standard Responses.** The category tree is the means of organizing and providing access to the library of standard responses. Each standard response must be associated with one category. One category can have zero or many standard responses associated with it.
- **Classification and Routing.** Genesys Content Analyzer can classify an incoming e-mail in terms of the category tree. You can use the resulting classification in three ways:
  - To supply a standard response as an acknowledgment or an automatic response to an e-mail.
  - To suggest a standard response to the agent handling the e-mail.
  - To route the e-mail.

Categories with no associated standard responses may be of use in grouping other categories together.

Note these definitions:

- A *terminal* category is one that has no subcategories: a leaf on the category tree.
- A *nonterminal* category is one that has subcategories.
- *Child* is another term for subcategory. For example, in "Example Category Tree", `savings` is a child of `accounts`, and `accounts` has the two children `checking` and `savings`.

A category tree is specific to a tenant and a language. Each tenant/language pair can have multiple category trees. Select tenant and language using the drop-down lists near the top of the Knowledge Manager window.

Without Content Analyzer, language is simply a label. You can design different sets of screening rules (for example) for different languages within a single tenant. But the screening rules operate the same way regardless of which language they are grouped under. With Content Analyzer, choice of language

affects the way the system operates; for example, when classifying interactions. Language and Dictionary Names provides more information on this topic.

Category membership is inherited. That is, if Category 1 includes Categories 10 and 11, and Category 10 includes Categories 100 and 101, then Category 1 also includes Categories 100 and 101.

Using Basic Knowledge Manager Objects

The basic Knowledge Manager objects are categories, standard responses, and screening rules. The table "Task Flow for Basic Knowledge Manager Objects" presents an overall task flow for creating and managing these objects, and for importing and exporting them.

Task Flow for Basic Knowledge Manager Objects

Objective	Related Procedures and Actions
1. Create a category tree.	Creating a category tree.
2. Create and manage standard responses.	To create and manage standard responses, open the Standard Response Editor and use the General tab. 1. Filling out the HTML part tab to create an HTML version of the standard response. 2. Filling out the Additional tab to specify possible uses and other attributes of the standard response. 3. The Attachments Tab to add and manage attachments. 4. The History Tab for version control: 5. Searching for a standard response 6. Standard Responses for use with SMS Gateways
3. Create field codes and use them in standard responses.	Creating Field Codes. 1. Inserting field codes into a standard response. 2. Examples of the use of field codes.
4. Create and manage screening rules.	Creating and managing a Screening Rule 1. Testing screening rules on the UCS database and Testing screening rules on specially-created text. 2. Finding a Screening Rule.
5. Import and export Knowledge Manager	1. Exporting Knowledge Manager objects.

Objective	Related Procedures and Actions
objects.	2. Importing Knowledge Manager objects.

Using Categories and Standard Responses

This section describes the creation and editing of categories and standard responses. It also provides some explanation of how they work.

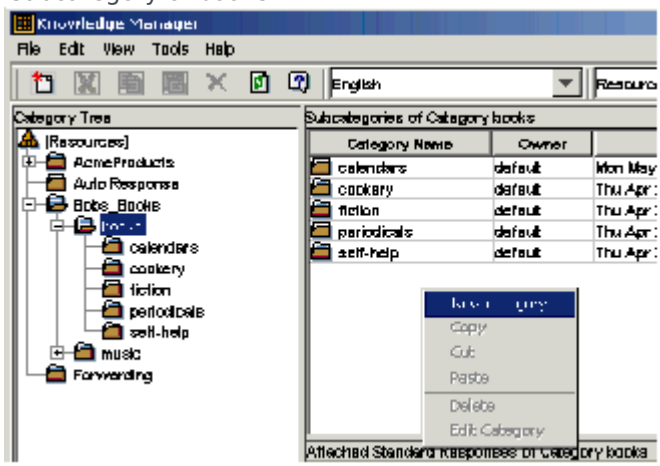
See the following procedures:

Creating a Category Tree

You create categories on the `Categories` tab of Knowledge Manager.

Purpose: To create a category tree for organizing and providing access to standard responses (and to additional functions in Genesys Content Analyzer).

1. Decide whether you want a new root category or a subcategory of an existing category:
   a. For a root category, select the tenant, which is the topmost node in the category. Be sure that you also select the correct language.
   b. For a subcategory of an existing category, select the existing category on the Category Tree pane.
2. Select `New category` from the `File` menu (or right-click in the Category Tree pane or the Subcategories pane). See the figure "Creating a New Category," where the new category being created will be a subcategory of books.



Creating a New Category

Characters Allowed in Object Names

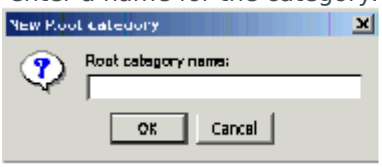
> **Important**
>
> Names of categories, like those of all Knowledge Manager objects, can consist only of alphanumeric characters (A–Z, a–z, 0–9), plus the characters shown in "Characters Allowed in Object Names".

Name	Character
Hyphen	-
Exclamation point	!
Number sign, pound	#
Dollar sign	$
Caret	^
Asterisk	*
Underscore	_
Curly brackets	{ }
Angle brackets	< >
Period, full stop	.

Names must also be no more than 64 characters long; if the object is to be imported, the limit is 58 characters.

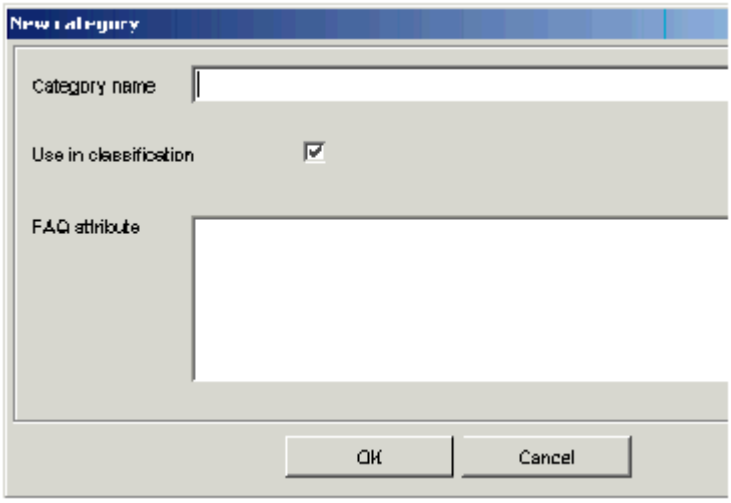
3. Selecting `New category` produces different dialog boxes, depending on whether your new category is a root or a subcategory.

   a. The New root category dialog box, shown in "New Root Category Dialog Box," requires that you enter a name for the category:



New Root Category Dialog Box

   b. The `New category` dialog box, shown in "New Category Dialog Box," creates a subcategory. Enter a name for the category.



New Category Dialog Box

The `Use in Classification` check box enables you to choose whether Genesys Content Analyzer uses the category in classification. You may wish to have some categories that are used only for organizing other categories or standard responses, not for classification.

In the `FAQ attribute` box you can enter the text of a question to which a standard response associated with this category can serve as the answer. See FAQ Objects.

You can also edit an existing nonroot category by right-clicking it and selecting `Edit category`.

Next Steps

- Read further information below about "Copying, Pasting, and Deleting" categories and category trees," and about "Considerations in Designing a Category Tree."
- Create standard responses.

Copying, Pasting, and Deleting

You can copy, paste, edit, and delete categories and category trees and their associated standard responses. Any of these operations on a category also applies to all of its subcategories.

If a category tree is used by a training object, you cannot delete the entire tree unless you first delete the training object. You can, however, delete any nonroot category from the tree (you cannot delete the root category).

Considerations in Designing a Category Tree

Without Genesys Content Analyzer, the sole function of categories is to organize the library of standard responses. With Genesys Content Analyzer, additional considerations come into play.

Creating Standard Responses

Creating Standard Responses includes:

- Creating or Editing a Standard Response—Also shown in the video lower down on this page.
- Language and Dictionary Names
- Filling Out the HTML part Tab
- Filling Out the Additional Tab
- The Attachments Tab
- The History Tab

You create standard responses on the `Categories` tab of the Knowledge Manager User Interface. If you want to use field codes in a standard response you must create them on the Field Codes tab.

You can create two versions of a standard response, one in plain text format and one in HTML format, by filling out the HTML part tab. When E-mail Server uses the standard response to create an e-mail (for example, when generating an acknowledgment), it creates a multipart e-mail that includes both plain text and HTML versions. Then the settings of the e-mail client that receives the e-mail determine which version displays.

You should be aware that e-mail clients may display multipart e-mails in varying ways. For example, if Microsoft Outlook has AutoPreview turned on, the preview may show the plain text version whereas the full display shows the HTML version. For this reason you should be careful that the plain text and HTML versions have identical content.

> **Important**
>
> To create and edit an HTML version of a standard response, you must have Internet Explorer 5.5 or later on the same host as Knowledge Manager.

Here is a video demonstrating the creation of a standard response.

Link to video

Procedure: Creating or Editing a Standard Response

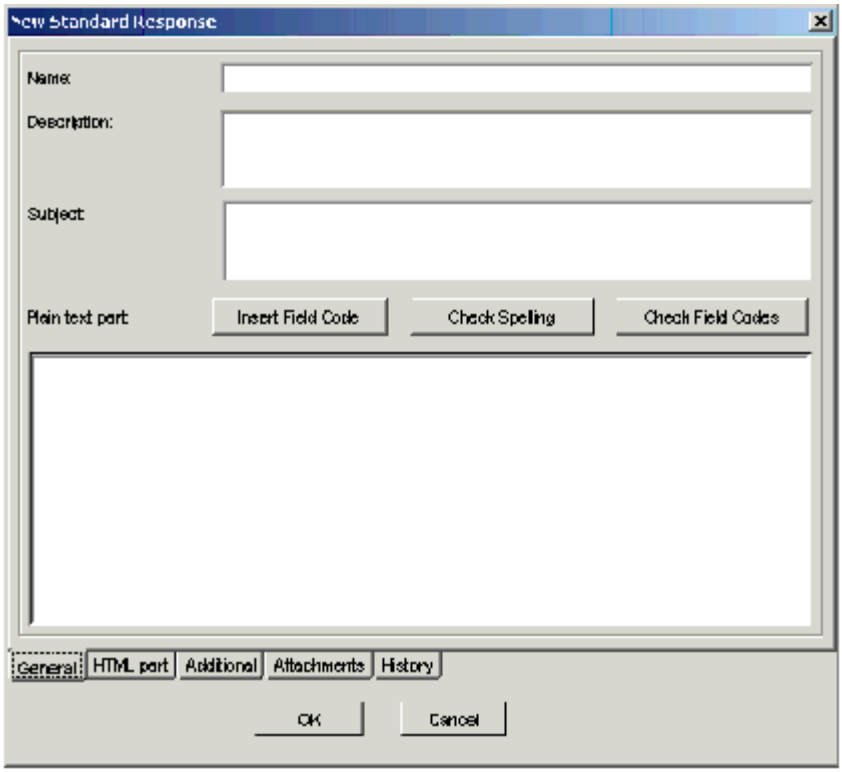
Purpose: To create prewritten content that can be used in responding to interactions.

Prerequisites

- You must first create a category tree.
- If you want to use field codes in your standard responses, you must first create the field codes.

1. On the Category Tree pane, select the category that you want to add a standard response to.
2. On the Standard Responses pane, right-click and select `New Standard Response.`
   The Standard Response Editor displays, as shown in the figure "New Standard Response."

> **Important**
>
> The title of the Standard Response Editor dialog box is New Standard Response if you are creating a standard response and Edit Standard Response if you are editing an existing standard response. The contents are otherwise identical.



New Standard Response

The display opens on the `General` tab.

3. Enter a name (required) and description (optional) for the standard response. Remember that the name can use only alphanumeric characters (A–Z, a–z, 0–9), hyphen, underscore, and space.
4. Enter a subject for the standard response. What you enter here appears as the Subject line in an e-mail (such as acknowledgement or autoresponse) generated from this standard response. You can also specify a Subject line in any IRD strategy object that has a Format tab (see the *Universal Routing 8.1 Reference Manual*). If you do, this overrides any Subject line that is specified for the standard response in Knowledge Manager.
5. Enter text for the plain text version of the standard response, using the `Insert Field Code` button to insert field codes. If you have not yet created field codes, you can continue creating the standard response, then return to it and insert field codes after you have created them.
6. Click `Check Field Codes` to show the result of applying (rendering) the field codes in this standard response. The values rendered are taken from a collection of generic properties that exists for this purpose.
7. Click `Check Spelling` to check spelling for the language selected in the adjacent drop-down list. See Language and Dictionary Names for more information on this.

Next Steps

- Read further information about the Language and Dictionary Names.

- Go on to Filling out the HTML Part Tab.

Language and Dictionary Names

For each language with spelling checker support, Knowledge Manager maintains a number of files in the \lex subdirectory of the Knowledge Manager installation directory. This section explains several features of the following file types:

- <language_name>_spllchk.pro file—Properties file, contains settings for options such as case sensitivity. Also points to the dictionary and user dictionary files.
- userdic_<language_name>_spllchk.tlx file—User dictionary file, contains words that you add by selecting Add in the Spelling dialog box.

For a language to display in the Check Spelling drop-down list, it must have both of the following:

- <language_name>_spllchk.pro file
- A Business Attribute, of type Language, whose Attribute Value name is <language_name>, exactly matching the name of the .pro file.

For example, the file name BrazilianPortuguese_spllchk.pro shows that the name of the Business Attribute for this language must be BrazilianPortuguese, not Brazilian Portuguese or BrazPort or anything else.

Renaming a Language

It is not possible to rename an existing language Business Attribute. You can, however, alter the language name that displays in the Check Spelling drop-down.

For example, if you build a German-language category tree, you may want the name of the language to appear as the German word "Deutsch"; rather than the English word "German."

So you can create a Business Attribute called Deutsch and select it when creating your category tree. But when it comes to checking the spelling, Genesys Knowledge Management uses the name "German" as shown by the filename German_spllchk.pro. With the .pro file and the Business Attribute having nonmatching names, German will not be available on the drop-down list.

You can rectify this situation by renaming the dictionary file:

1. Locate the file German_spllchk.pro and rename it Deutsch_spllchk.pro.
2. Refresh the view in Knowledge Manager (select Refresh from the View menu or click the refresh icon).

Saving a User Dictionary

If you have customized your spelling checking by adding items to a user dictionary, you will want to avoid overwriting your user dictionary when installing a new version of Knowledge Manager (or reinstalling the existing one).

Saving an existing user dictionary

1. Make a copy of the existing `userdic.tlx` file, located in the existing Knowledge Manager installation directory.

> **Important**
>
> Prior to release 7.6.1, Knowledge Manager had spell checking for English only, so there was only one user dictionary file, named userdic.tlx. This section describes saving the contents of this English user dictionary.

2. Install or reinstall Knowledge Manager.

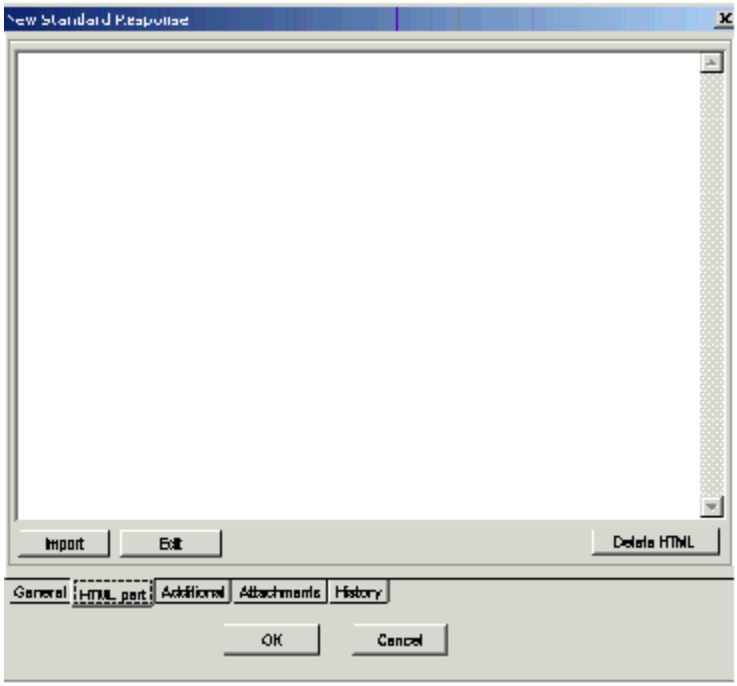
3. Locate the new `userdic_English_spllchk.tlx` (in the new Knowledge Manager installation directory) and replace its contents with the contents of `userdic.tlx`.

Procedure: Filling out the HTML Part Tab

Prerequisites

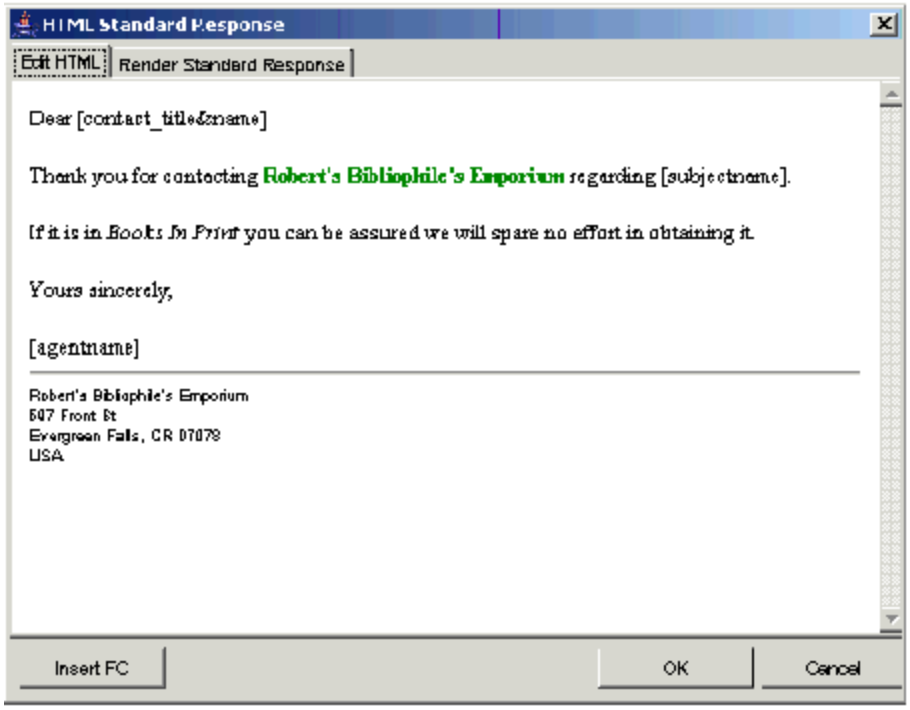
For the HTML version of a standard response, you must first create the standard response as a standalone HTML file, using whatever authoring tool or text editor you prefer.

1. Start on the HTML part tab, as shown in "HTML Part Tab."



HTML Part Tab

2. On the HTML part tab, click Import and browse to the HTML file that you want to import, then click OK. The HTML Standard Response dialog box appears, as shown in "HTML Standard Response."

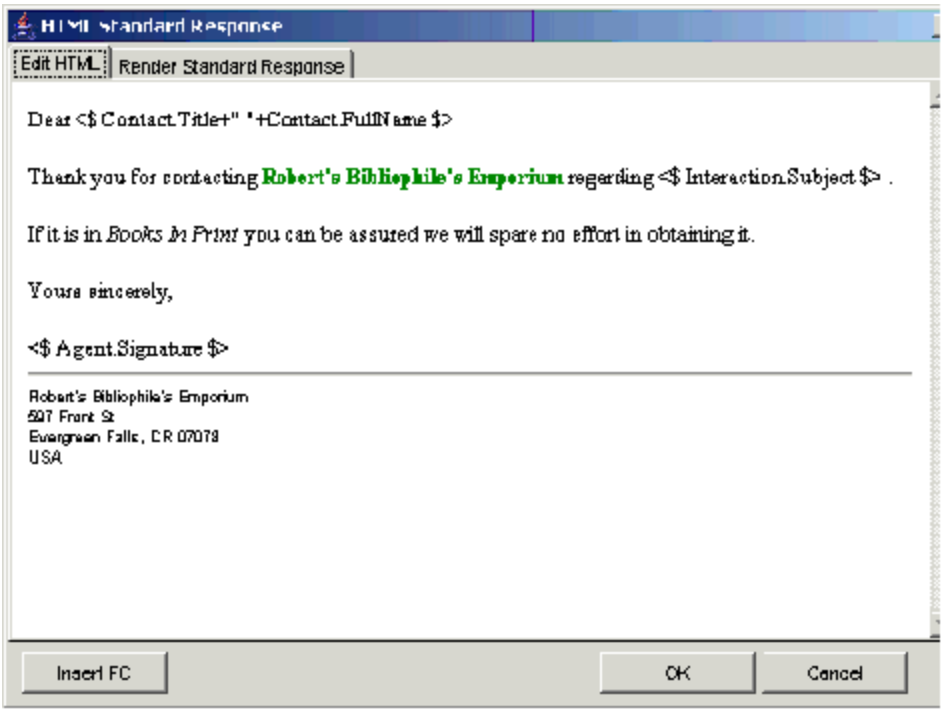


HTML Standard Response

> **Important**
>
> In the HTML version of a standard response, links to resources that are used in the content, such as image files, must use absolute URLs; embedded graphics or relative links are not supported. Also, these resources must be available on the web through an HTTP server at the time that the standard response containing this HTML is sent to the customer.

This dialog box opens with its `Edit HTML` tab. On this tab you can do the following:

- Insert, delete, copy, and paste text. Pasted text retains its attributes. Inserted text takes on the attributes of its insertion point.
- Insert field codes. Note that the imported file in "HTML Standard Response" includes bracketed expressions like `[subjectname]` to indicate where field codes should be inserted. Using the `Insert FC` button to insert field codes produces the result shown in "HTML Standard Response with Field Codes Inserted."



HTML Standard Response with Field Codes Inserted

> ## Important
> Generally, plain text part and HTML part should have the same semantic content from the call center point of view. These parts should contain the same field codes and differ, perhaps, only in the content that requires HTML representation, such as pictures or rich text capabilities. If both parts are present, it will depend on the email clients how such MIME (email) content is displayed. For example, if Microsoft Outlook has AutoPreview turned on, the preview may show the plain text version whereas the full display shows the HTML version. For this reason you should be careful that the plain text and HTML versions have identical content.

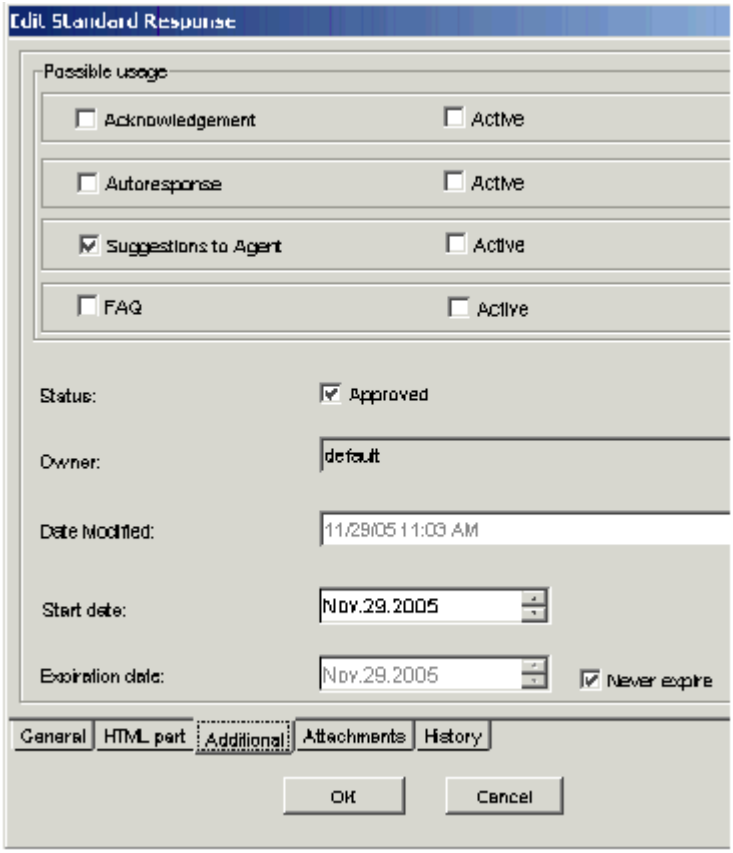
Go to the `Render Standard Responses` tab to see the standard response with the field codes rendered in the same way as when you click `Check Field Codes` on the `General` tab, as described in Inserting field codes into a standard response.

Next Steps

- Go on to the Additional tab.

Procedure: Filling out the Additional Tab

1. You must specify the possible uses of the standard response by using the check boxes on the `Additional` tab, shown in "Standard Response, Additional Tab."



Standard Response, Additional Tab

The possible uses are:

a. Acknowledgment—The standard response may be sent to acknowledge receipt of an incoming interaction.

b. `Autoresponse` —The standard response may be used as an automatic response to an incoming interaction.

c. `Suggestions to Agent` —The standard response may be offered to agents as suggested wording to use in their own replies to interactions.

d. FAQ—The standard response may supply the answer for an item in an FAQ object. This use type applies only with Genesys Content Analyzer.

- For each usage type, you must also specify this standard response as Active or not. Each category may have multiple standard responses of each usage type, but only one standard response of a given usage type can be Active. For more information on this point, see "The Meaning of Active" below.
- Specify the status using the Approved check box. Only Approved standard responses can appear in Routing objects.
- Use the other check boxes and fields on this tab to specify the owner, date modified, start date, and expiration date if any.

If a standard response's expiration date has been reached, it has the following effects:

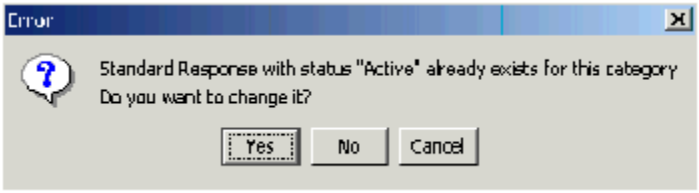
- The standard response is not shown in IRD, so it cannot be used in a new or modified strategy.
- If this standard response was saved in a strategy before the expiration date was reached, E-mail Server does not send the standard response, but returns an error message.

Next Steps

- Read further information about "The Meaning of Active" below.
- Use The Attachments Tab to add an attachment to the standard response.
- Maintain multiple versions of the standard response on the The History Tab.
- Search for a Standard Response.
- Create Field Codes to use in your standard responses.

The Meaning of Active

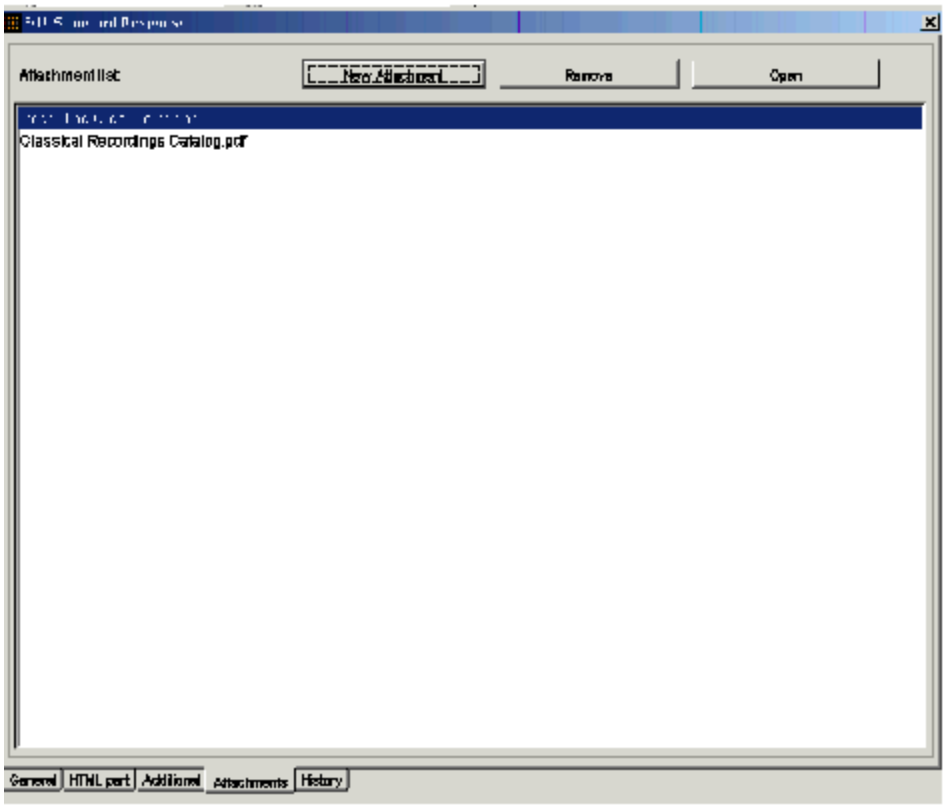
- Purpose of Active: There are times when the system must select, without immediate user input, a standard response of a given usage type for a given category. If there are multiple standard responses of a single usage type for the category in question, the system selects the one that is designated as Active .
- Changing the Active standard response: If you attempt to select Active for a Standard Response (either a new one or an existing one), and there is already an Active Standard Response with that usage type for that category, Knowledge Manager offers to take the previously Active Standard Response out of Active status, displaying the message shown in "Changing the Active Standard Response."



Changing the Active Standard Response

The Attachments Tab

The Attachments tab displays a list of attachments to the standard response, as shown in "The Attachments Tab."



The Attachments Tab

On this tab you can do the following:

- To add an attachment, click New Attachment. Browse to the file that you want to attach, then select it.
- To remove an attachment, select it, then click Remove.
- To view an attachment, select it, then click Open (you cannot double-click on the attachment name to view it).

> **Important**
>
> This opens the attachment for viewing only. You cannot save any changes that you make to the attachment.

The History Tab

On this tab you can do the following:

- To save multiple versions of the standard response, click `Save to new version.`
- To choose the version that appears on the Standard Response pane, select the version and click `Restore.`

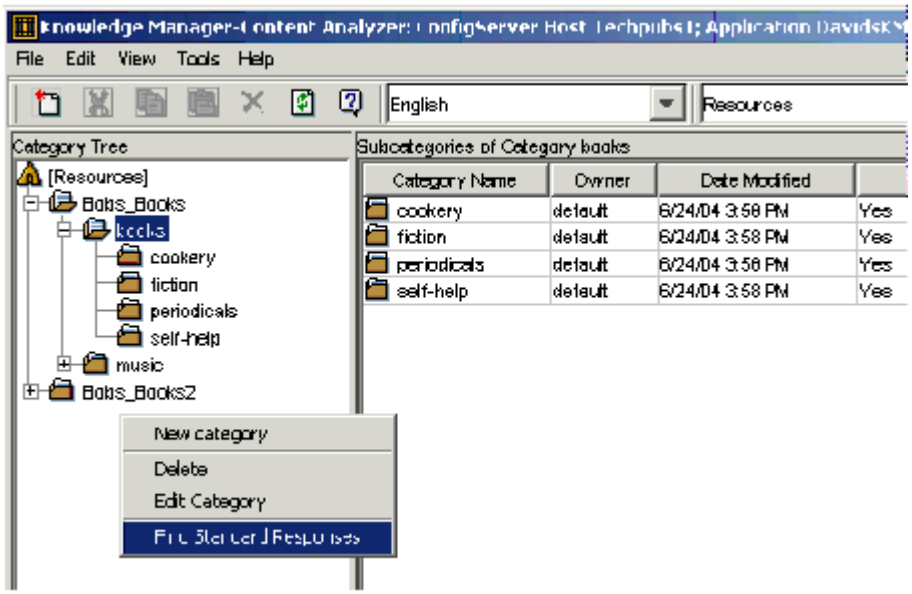
Restoring a version restores only the Text and Description parts of the standard response. For example:

1. Version 1 of a standard response has an expiration date of December 10, 2010.
2. Create a Version 2 with `Never expire` selected (no expiration date).
3. Restore Version 1.
4. The restored version has `Never expire` selected. The former expiration date of December 10, 2010 is not restored.

Procedure: Searching for a Standard Response

You can search for the standard responses that are associated with a selected category and its subcategories.

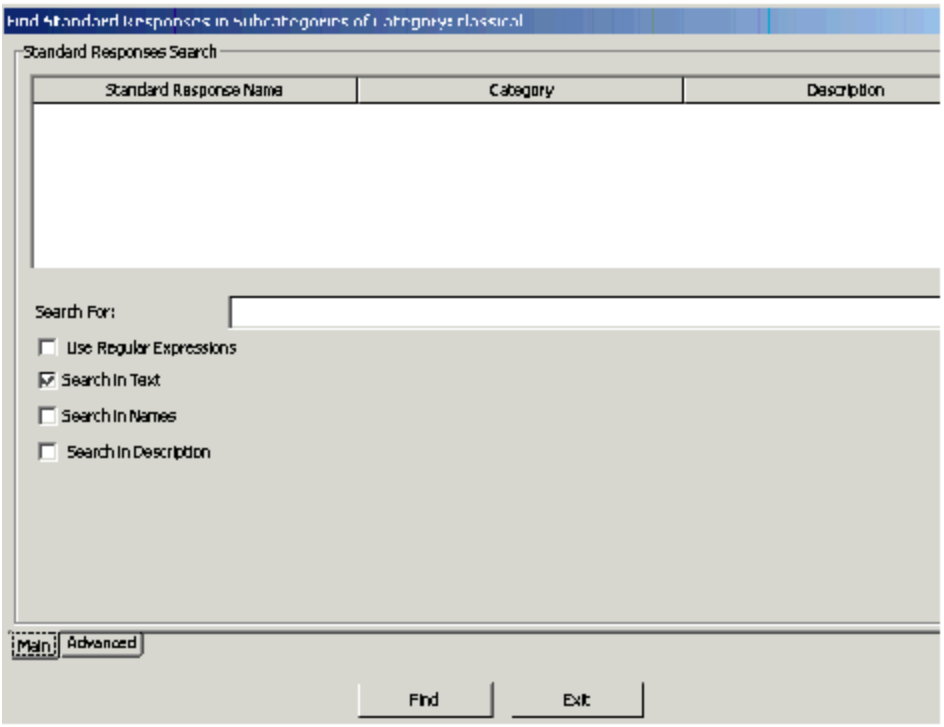
1. Select a category on the Category Tree pane.
2. Right-click anywhere in the Category Tree pane and select `Find Standard Responses` from the shortcut menu, as shown in "Find Standard Responses".



Find Standard Responses

The search function searches for standard responses associated with the selected category and all of its subcategories.

- The `Find Standard Responses in Subcategories of Category: <name>` dialog box appears, as shown in "Find Standard Responses Dialog, Main Tab"

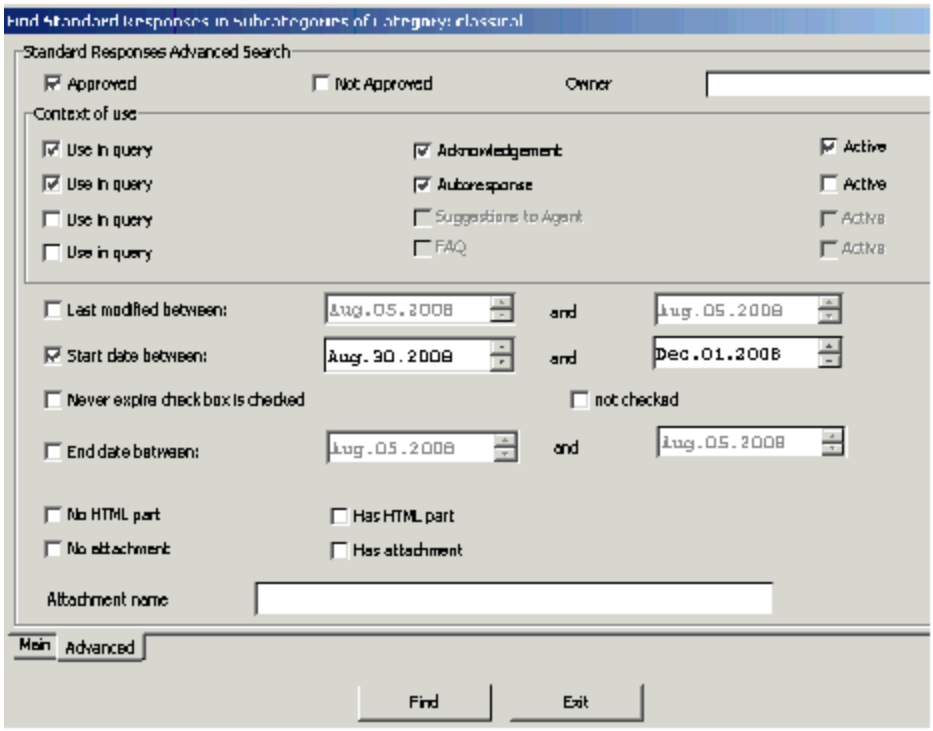


Find Standard Responses Dialog, Main Tab

- Enter a string to search for in the `Search For` field (if there is nothing in this field, the system reports `Empty search pattern`). Select `Use Regular Expressions` if you want the search to treat the string as a regular expression. For more information see Regular Expressions.
- Select check boxes to search in the text of the standard response, its name, its description, or any combination. The `Search in Text` box is selected by default (if none of these check boxes is selected, the system reports `No Standard Responses have been found`).

> **Important**
>
> When the Search in Text box is selected, Knowledge Manager searches both plain text and HTML versions.

- Go to the Advanced tab, shown in "Find Standard Responses Dialog, Advanced Tab", to select further attributes, in any combination.



Find Standard Responses Dialog, Advanced Tab

- Click `Find`. The dialog box displays the name, category, and description of all standard responses found. Click `Exit` to close the dialog box.

Standard Responses for SMS Gateways

Standard Responses can be used to carry the body of a Short Message Service (SMS) message that E-mail Server sends using an SMS gateway. To accomplish this you must use a routing strategy that includes a `CreateSMS` object and specifies certain attached data. For details on this strategy configuration see the "Multimedia Objects" section of the "Interaction Routing Designer Objects" chapter of the Universal Routing 8.1 Reference Manual.

The form of the special standard response differs according to the requirements of the gateway that you are using. This section provides examples of standard responses that can be used with three available gateways.

Clickatell

For the Clickatell gateway, create a standard response with the following as its body:

```
api_id:1234

user:Name

password:Secret

from: <$AttachedData("OrigSMSNumber")$>

to: <$AttachedData("DestSMSNumber")$>

text: <$AttachedData("SMSText")$>
```

Where:

- `1234` should be replaced with the api_id that you received upon registering for the service.
- `Name` should be replaced with the user name that you created when registering for the service.
- `Secret` should be replaced with the password that you created when registering for the service.
- The user data `"OrigSMSNumber"` contains the number of the originating SMS device.
- The user data `"DestSMSNumber"` contains the number of the recipient SMS device.
- The user data `"SMSText"` contains the text of the SMS to send (limited to 160 characters).

You can add other available parameters to the body. For information about what parameters are available for this gateway, do as follows:

1. Go to `http://www.clickatell.com/`.
2. Select `Developers`, then `SMTP`.

SMS Gateway for Mdaemon and sms2email.com

For the SMS Gateway for Mdaemon or the sms2email.com gateway, create a standard response with the following as its body:

```
<$AttachedData("SMSText")$>
```

Where the user data "SMSText" contains the text of the SMS to send (limited to 160 characters).

For more detailed information on Mdaemon:

1. Go to http://www.achab.com/.
2. Click SMS Gateway for MDaemon, then Features -Outbound SMS.

For more detailed information on sms2email.com:

1. Go to http://www.sms2email.com/.
2. Click Developer Info, then Email to SMS Gateway, then how-to guide.

Field Codes

Although field codes are used mostly in standard responses, they are the most complex and powerful aspect of standard responses, so they are described in this separate section, including:

- Field Codes Overview
- Field Code Variables
- Field Codes in Knowledge Manager
- Field Code Anatomy
- Field Code Examples

Field Codes Overview

The main use of field codes is to particularize standard responses, in a manner similar to the Mail Merge feature in word processors.

For example, you can use the field code `<$Contact.FirstName$>` in a response beginning `Dear <$Contact.FirstName$>`, which you send to dozens of recipients. In each message, `<$Contact.FirstName$>` is replaced by the first name of the addressee of the message (the contact) as listed in the Universal Contact Server database.

More generally, a "field code" is a formula that you insert into an outgoing text object, such as an e-mail that E-mail Server generates when triggered to do so by a routing strategy object.

The most common type of such text object is a standard response (triggered by an Autoresponse or Acknowledgement object), but you can also insert field codes into other types, such as chat transcripts, SMS messages, and forwarded or redirected e-mails. In some cases the only place you can insert a field code is in the Subject line using the `Format` tab in a strategy object.

The following is a complete list of the strategy objects that can use field codes either in a standard response or in the Subject line:

- Acknowledgement
- Autoresponse
- ChatTranscript
- Create EmailOut
- Create Notification
- Create SMS
- Forward
- RenderMessageContent

The following is a complete list of the strategy objects that can use field codes only in the Subject line:

- Redirect
- Reply from External Resource
- Send

When a text object containing such a formula is processed, the following happens:

1. The formula performs an operation, which produces a result.
2. The result replaces the field code in the text object.

This process of performing an operation and substituting its result is called "rendering."

> **Important**
> Field codes can be used in outgoing text objects only.

A complete reference list of field codes is available in the Genesys eServices Field Codes Reference Manual.

Field Code Variables

Using Field Code Variables includes:

- System Variables
- Custom Variables
- Using Your Own Data in Standard Responses

Using UCS Data in Standard Responses: System Variables

In the example given in Field Codes Overview, the `Contact.FirstName` retrieves a piece of data about the interaction. The ability to access interaction data is perhaps the most frequent use of field codes. Although field code formulas can be very complicated, many simply retrieve a single piece of data, such as a contact's name.

You access Universal Contact Server data using predefined variables, called "system variables."

These variables access three predefined objects. Each object has a name and a set of properties. In the example, `Contact` is an object and `FirstName` is one of its properties. The system variable `Contact.FirstName` retrieves the value of the `FirstName` property of the `Contact` object.

In similar fashion, there is a system variable for each `object+property` pair. The objects and properties that you can use in field code formulas are described in the following sections.

Interaction

This object represents the particular interaction being worked on, such as an inbound e-mail. These are its properties:

- `Id`
- `DateCreated`
- `Subject`
- `ToAddress`
- `FromAddress`
- `AttachedData`
- `TimeZone`

Contact

This object represents the contact associated with the interaction being worked on. These are its properties:

- `Id`
- `Title`
- `FirstName`
- `LastName`
- `FullName`

- `PrimaryPhoneNumber`
- `PrimaryEmailAddress`

Agent

This object represents the agent working on the interaction. These are its properties:

- `FirstName`
- `LastName`
- `FullName`
- `Signature`

> **Important**
>
> Automated responses use the default agent. Create the default agent as a Person object just like any other in Configuration Manager. Then select this Person on the Automated Reply Agent screen of the E-mail Server configuration wizard (or set this Person as the value of the autobot-agent-login-name option in the E-Mail Processing section of the E-mail Server Java application). Since this is the Person who the automated response appears to be from, you may want to name it after your company or institution.

Custom Variables

In addition to the system variables, you can use Knowledge Manager to create custom variables. Custom variables have the following properties:

- Their values are assigned by strategy objects.
- Therefore, standard responses that use field codes containing custom variables must have the usage type `Autoresponse` or `Acknowledgment`.

For an example of the use of a custom variable in a standard response, see Procedure, Using a custom variable. For a complete description of the Routing objects that can use custom variables, see the Universal Routing 8.1 Reference Manual.

> **Important**
>
> The names of custom variables must begin with an alphabetic character or underscore, and the remainder of the name must consist only of alphanumeric characters or underscores. This differs from the requirements for the names of other Knowledge Manager objects, which may also contain hyphen and space. For example, `5-usercode` is not an acceptable name for a custom variable, but it is acceptable as the name of a screening rule or category.

Using Your Own Data in Standard Responses

It is possible to incorporate data that you keep external to Universal Contact Server into your standard responses (including automated responses). This data could include case numbers, account information, and so on. Remember that attached data always consists of key-value pairs.

Incorporating external data into standard responses is a two-step process:

1. Retrieve the external information and add it to the interaction as attached data. One place to do this is in a routing strategy (see Interaction Routing Designer Help).
2. Now that you have attached the data to the interaction, you can use the `AttachedData` property of the Interaction object to access the data and incorporate it into your standard response. The `AttachedData` property requires one argument, which is the key name. The result of the following formula is the value associated with the `OrderStatus` attached-data key:

```
<$Interaction.AttachedData("OrderStatus")$>
```

Field Codes in Knowledge Manager

Using Field Codes includes:

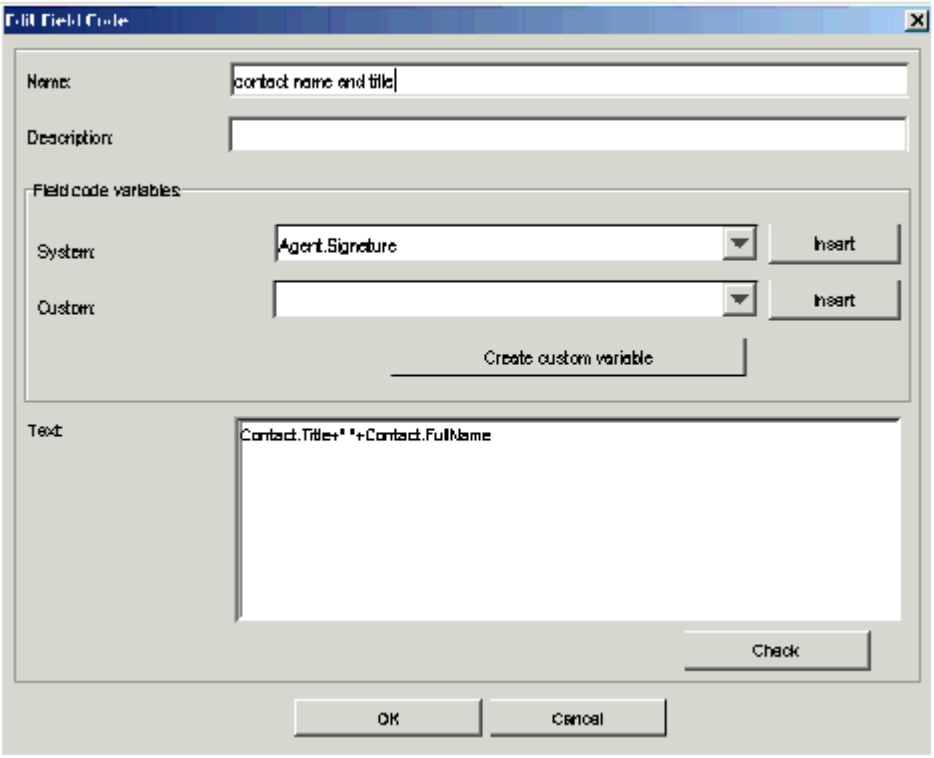
- Procedure, Creating Field Codes
- Procedure, Inserting Field Codes into a Standard Response

Knowledge Manager separates the task of creating field codes from the task of creating standard responses. This allows you to create complex field codes that include multiple objects, formulas, and constants (see Using Complex Field Codes for an example). You can then use these complex field codes in multiple standard responses.

> **Important**
>
> Standard Responses that are intended for use in FAQ objects should not contain field codes. FAQ objects contain no means of rendering field codes. See also See FAQ Objects.

Procedure: Creating Field Codes

1. On the Field Codes tab of Knowledge Manager, right-click and select New Field Code. The New Field Code dialog box (also called Edit Field Code) opens, as shown in "Edit Field Code".



Edit Field Code

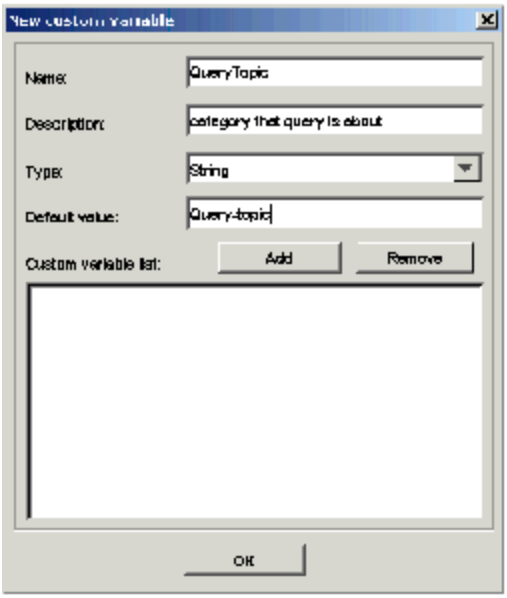
The Text field shows the field code as you create and edit it.

- Enter a name and description for the field code. Name is the only required field in this dialog box.
- The Field code variables section includes two fields:

  a. System. Click the down arrow to display a list of all system variables.

  b. Custom. Click the down arrow to display a list of all custom variables that you have created.

To create a new custom variable, click Create custom variable. The New custom variable dialog box opens, as shown in "New Custom Variable".



New Custom Variable

- Enter a name (required) and description (optional). Observe the restrictions (see See The names of custom variables must begin with an alphabetic character or underscore, and the remainder of the name must consist only of alphanumeric characters or underscores. This differs from the requirements for the names of other Knowledge Manager objects, which may also contain hyphen and space. For example, 5-usercode is not an acceptable name for a custom variable, but it is acceptable as the name of a screening rule or category.) on custom variable names.
- Select a type (`String` or `Integer`).
- Enter a default value (required). This value is rendered when you click `Check Field Codes` in the Standard Response Editor (see Step 4 in Creating and Editing Standard Responses).
- Click `Add`.

  You can also use this dialog box to edit and delete existing custom variables. Click OK to return to the `New Field Code` dialog box.

- Select system and custom variables from the drop-down lists, then click `Insert` to insert them into the `Text` field.
- Enter any other desired text in the `Text` field. This text must conform to the rules described in Field Code Overview: Field Code Syntax.
- Click Check to verify that the field code is well-formed (that is, that it has no typographical errors, missing parentheses, and so on).

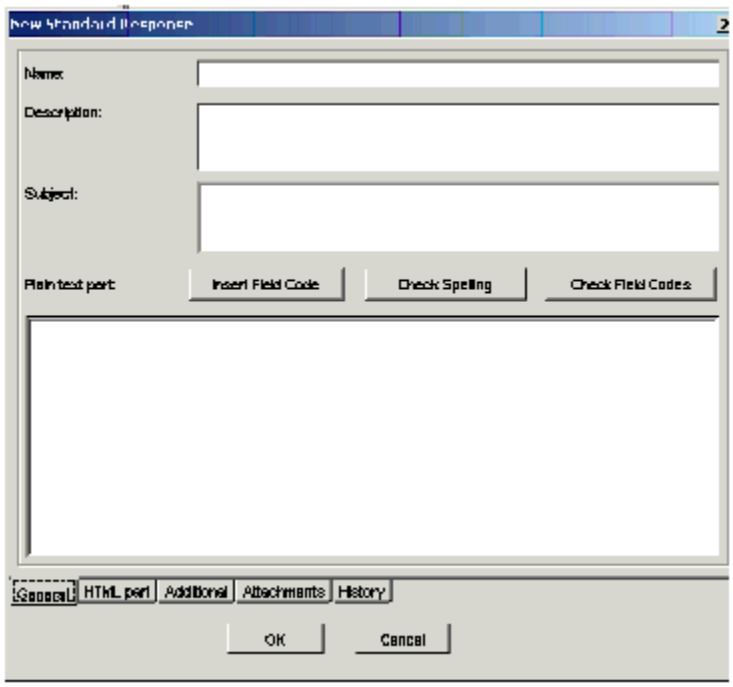
Next Steps

- Use field codes in a standard response.

Procedure: Inserting Field Codes into a Standard Response

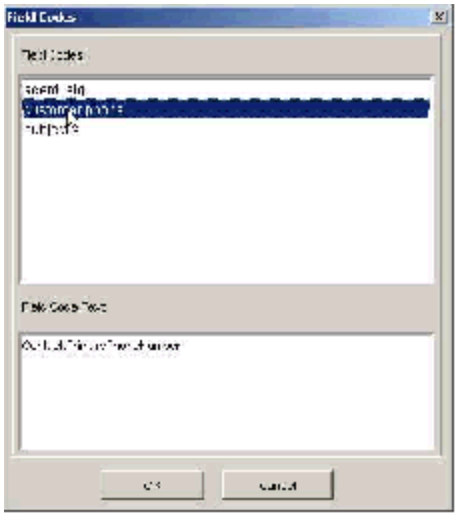
1. On the Categories tab, do one of the following:
   a. Double-click a standard response. The Edit Standard Response dialog box opens.
   b. Right-click and select New. The New Standard Response dialog box opens.

   These two dialog boxes are identical apart from their titles. The figure "New Standard Response" shows New Standard Response.



New Standard Response

- Click Insert Field Code to display a list of all the field codes that you have created. The list appears in the dialog box shown in "List of Field Codes."



List of Field Codes

- Select a field code and click OK to insert it, together with its required delimiters <$ $>, into a standard response. This returns you to the `Edit Standard Response` (or `New Standard Response`) dialog box.
- Click `Check Field Codes` to see the standard response with the field codes rendered. The values rendered come from a collection of generic properties that exists for this purpose.

Next Steps

- Examine the use of field codes in Field Code Examples.

Field Code Anatomy

In addition to system variables such as `Contact.FirstName`, field codes may contain formulas. This section provides an outline of formula usage. For details on many of these topics, see the information provided in "Field Codes in Standard Responses," in the *eServices 8.1 Reference Manual*.

You must always delimit field codes by using `<$ … $>`. If you type a field code directly into the body of a standard response, then you must enter the delimiters yourself. If you select from the list of field codes in Knowledge Manager, then the delimiters are added automatically.

The text that appears inside the delimiters is a formula. Field code formulas are very similar to formulas in other applications, such as Microsoft Excel.

A *formula* is a sequence of one or more operands (such as numbers and text strings), separated by operators (such as + and -).

For example, in the following formula, 2 and 3 are operands and + is an operator:

```
<$2 + 3$>
```

Operands can be values that do not change (constants), or values that vary based on the context. In the previous formula, all the operands are constants, so the formula always evaluates to `5`. The next formula, on the other hand, evaluates to a different value for each agent who uses it:

```
<$Agent.Signature$>
```

Field Code Syntax

To summarize field code syntax:

- As stated previously, a field code must be delimited by `<$ ... $>`.
- Alphabetic strings, whether constants in formulas or elsewhere in a field code, must be enclosed in double quotes.
- Numeric constants require no special treatment.
- You must use special characters for some purposes. For example, for your field code to render with a line break, you cannot simply type a carriage return. Instead, you must insert the expression `\n`. A list of these special characters is available.

HTML in Field Codes

With special configuration, field codes can contain HTML markup; for example, you could have a field code <$my.agent.signature$> defined as

```
Sam Agent<BR />
Acme Products<BR />
29 Exterior Blvd<BR />
Springfield, CX 09090<BR />
```

To enable this, you must use the Java property `-Dsrl-field-code-allow-html=true`, in one of the following ways:

- Add it to the `JavaArgs` section of `ContactServerDriver.ini`
- Add it as an argument to the startup command line in `contactServer.sh`.

See also:

- Operator Precedence
- Functions

Operator Precedence

If you use more than one operator in a formula, the order in which they are evaluated depends on their relative *precedence* (higher precedence operators are evaluated first). For example, multiplication (*) has a higher precedence than addition (+), so that the formula below evaluates to `14`, not `20`:

```
<$2 + 3 * 4$>
```

You can use parentheses to override the default precedence. The formula below evaluates to `20`:

```
<$(2 + 3) * 4$>
```

For a complete list of operators and their relative precedence, see “Operator Precedence” in Chapter 3 of the eServices 8.1 Reference Manual.

Data Types

Operands of several different types may appear in formulas:

- Number
- String (text)
- Date/time
- Boolean (true/false)
- Object (Contact, Interaction, and Agent)

Each data type behaves differently in formulas, and the operators have different meanings when you use them with different data types. For example, the + operator means “add” when used with numbers, but “concatenate” (paste together) when used with strings. This formula evaluates to *Uncle Sam Wants You*

```
<$"Uncle Sam " + "Wants You"$>
```

In addition, some operators cannot be used with some data types at all. For example, you cannot use the multiplication (*) operator on two strings.
All formulas, regardless of their final data type, are converted to strings before being merged into your standard response. This conversion follows a set of default rules that depend on the data type. For example, the default rules for numbers round them off to integers. This formula causes 2 to be inserted into your standard response, even though the real result is `2.25`:

```
<$9 / 4$>
```

You can use the `Text` function (see below) or `format operator`:) to override the default formatting.

Either of the following formulas inserts `2.25` into your standard response:

```
<$Text(9 / 4, "#.##")$>
```

```
<$(9 / 4):"#.##"$>
```

For a detailed list of data types and how you can use them, see “Data Types” in the eServices 8.1 Reference Manual.

Functions

When composing formulas, you can use many built-in functions. *Functions* are predefined formulas that perform calculations using values, called *arguments,* which you supply. To use a function, write its name, followed by an opening parenthesis, the arguments for the function separated by commas, and a closing parenthesis.

Function arguments may be of any data type, although individual functions may place restrictions on their arguments. Function arguments may be constants or formulas. The `Length` function, for example, takes a single string argument and returns its length in characters. This formula evaluates to `13`:

```
<$Length("Hello, world!")$>
```

As another example, the`Date` function takes individual date components (year, month, day, and so on), and constructs a date/time value. The formula below evaluates to `1965-11-23 09:03:10`:

```
<$Date(1965, 11, 23, 9, 3, 10)$>
```

Functions may act as arguments to other functions. The `WeekdayName` function takes a single date/time argument and returns the day of the week as a string. The formula below evaluates to `Tuesday`:

```
<$WeekdayName(Date(1965, 11, 23, 9, 3, 10))$>
```

This formula evaluates to `7`:

```
<$Length(WeekdayName(Date(1965, 11, 23, 9, 3, 10))$>)
```

Detailed descriptions of all available functions are provided in the Genesys eServices Field Codes Reference Manual.

Using Objects

All object/property pairs are also available in the `Variables` drop-down menu in the Knowledge Manager `Field Code Editor`.
Object properties can be of any data type. `Agent.FullName`, for example, is a string, but `Interaction.DateCreated` is a date/time.
The data type of an object property can even be another object. For example, `Contact.EmailAddresses` yields another object called a `ContactEmailAddressList`. In cases such as this, you can access the properties of the resulting object by entering a period (.), followed by the property name, just as before. For example, the formula below evaluates to the number of e-mail addresses assigned to the contact:

```
<$Contact.EmailAddresses.Count$>
```

Some object properties require arguments just as functions do. For these properties, write the

arguments, enclosed in parentheses after the property name, just as before.

For example, the ContactEmailAddressList object has a property named `Exists`, which you can use to test whether a particular e-mail address is assigned to a contact. The data type of this property is Boolean (true/false), and it takes one argument, the e-mail address to test. For example:

```
<$Contact.EmailAddresses.Exists("samd@acme.com")$>
```

Detailed descriptions of all objects and their properties are provided in the Genesys eServices Field Codes Reference Manual.

Field Code Examples

This section presents examples of the use of field codes.

- Using a custom variable
- Using a complex variable

Procedure: Using a Custom Variable

Purpose: This is a simple example of the use of a custom variable in a standard response.

Prerequisites
This example assumes a category tree that includes categories called Cookbooks, Mysteries, and Reference.

1. In Knowledge Manager:
    a. Create a custom variable called QueryTopic (see See Creating field codes).
    b. Create a field code called Query_Topic that consists of the variable QueryTopic.
    c. Create a standard response of type Autoresponse called AUTO that includes the sentence Thank you for your inquiry about <$ Query_Topic $>.
    d. Create the following screening rules:
    e. Cookbook: RegExFind("cook") || RegExFind("recipe") || RegExFind("food") || RegExFind("cuisine")
    f. Mystery: RegExFind("murder") || RegExFind("crime") || RegExFind("case of the") || RegExFind("detective")
    g. Reference: RegExFind("dictionar") || RegExFind("encyclopedia") || RegExFind("almanac")

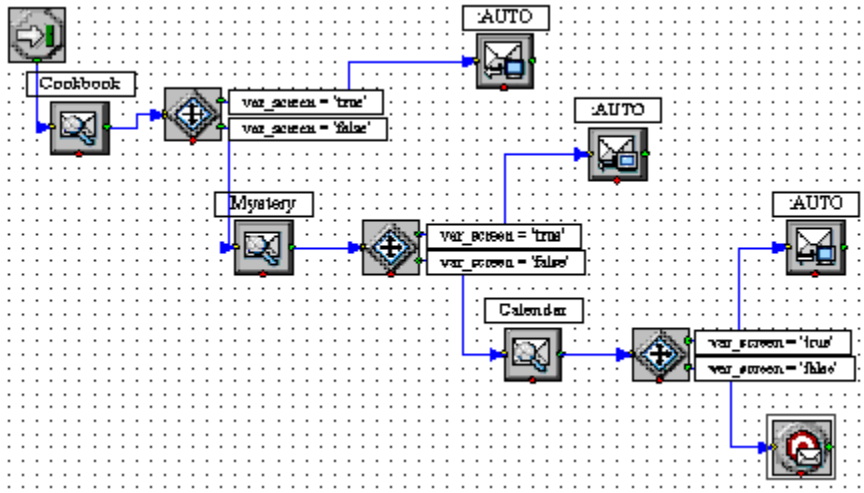
8. In Interaction Routing Designer, create a strategy that applies these screening rules one after the other, assigning a different value to the custom variable for each screening rule:
    a. Create a variable called var_screen.
    b. Create a strategy. Start the strategy with a Screen object. On the General tab of the Screen object, select the Cookbook rule.
    c. On the Result tab, click Assign values of the key-value pairs. Then under Output values select var_screen for Variable and enter ScreenRuleMatch for Key from output.
    d. Connect the Screen object to a Generic Segmentation object. Create two segments: var_screen = true and var_screen = false.
    e. Connect an Autoresponse object to the top green port (the one corresponding to true) of the Segmentation object. In the Autoresponse, select the Select standard response radio button and select AUTO in the associated drop-down list.
    f. Still in this Autoresponse, go to the General tab and in the Field Codes area (bottom of the tab) click the New icon, enter QueryTopic under Key, and enter cookbooks under Value.
    This will generate an e-mail that includes the sentence *Thank you for your inquiry about cookbooks.*
    g. Return to the Generic Segmentation and connect a new Screen object to its second green port (the one corresponding to false).
    h. On the General tab of the new Screen object, select the Mystery rule. On the Result tab, click Assign values of the key-value pairs. Then under Output values select var_screen for Variable and enter ScreenRuleMatch for Key from output .

i. Proceed as in Steps d–f: Connect this `Screen` object to a new `Generic Segmentation`, again with segments for `var_screen = true` and `var_screen = false`.

j. As in Step e, connect a new `Autoresponse` object to the green port for `true`, select the `AUTO` standard response, and enter `QueryTopic` under `Key`. This time enter `mysteries` under `Value`.

k. Return to the second `Generic Segmentation`'s green port for `false` and repeat Steps g–j, creating a third `Screen` object and `Generic Segmentation`. In the `Screen` object, select the `Reference` rule; in the `Segmentation` object, set the custom variable to `reference` .

The figure "Strategy Using Custom Variable" shows the strategy as described. The single standard response AUTO generates three e-mails, each with a different word filling the blank in *Thank you for your inquiry about ___*.



Strategy Using Custom Variable

Next Steps

- Examine an example of a Complex Field Code.
- Go on to create and manage Screening Rules.

Using a Complex Field Code

The following is an example of a complex field code:

```
<$ If (Time() - Interaction.DateCreated > 14, “Please accept our apologies for not having replied sooner. ”, “”) $>
```

This field code inserts a tardiness apology if more than 14 days have elapsed since the interaction first entered the system. It uses the function `If`, which has these properties:

- Its syntax is `If (Boolean, TrueResult, FalseResult)`
- If `Boolean` evaluates to `True`, it returns the second argument.
- If `Boolean` evaluates to `False`, it returns the third argument.

In this example the three arguments of `If` are as follows:

1. `Time() - Interaction.DateCreated > 14` A formula that returns `True` if the difference between the date created and the current system time is more than 14 days. (The result of a mathematical operation on dates is given in days.)
2. `“Please accept our apologies for not having replied sooner. ”` A text string apologizing for tardiness, inserted if the formula evaluates to `True`.
3. The null string: if the reply is not late (the formula evaluates to `False`), nothing is inserted in it.

Screening Rules

Screening rules scan an interaction and try to match either a destination address, a regular expression, or both. Screening is performed by Classification Server when it is triggered by a `Screen` object in a routing strategy.

A screening rule can optionally be associated with a category.

> **Important**
>
> Screening, like classification (see Genesys Knowledge Management: Content Analyzer), can operate on any interaction that has text somehow associated with it, whether as the body of the interaction (e-mail, chat), or otherwise (as user data, for example). In practice, it is expected that most interactions which are screened or classified will be e-mail messages; therefore, the terms *e-mail* and *message* are used interchangeably here, to refer to these interactions. In fact, whatever is said here about e-mail applies to any interaction that has associated text.

Screening Rules topics include:

- Screening Rules Use
- Screening Rules Reference
- Examples of Screening Rules
- Screening for Sentiment and Actionality

Screening Rule Use

Use the Screening Rule Editor to compose and test Screening Rules.

You can also search for Screening Rules.

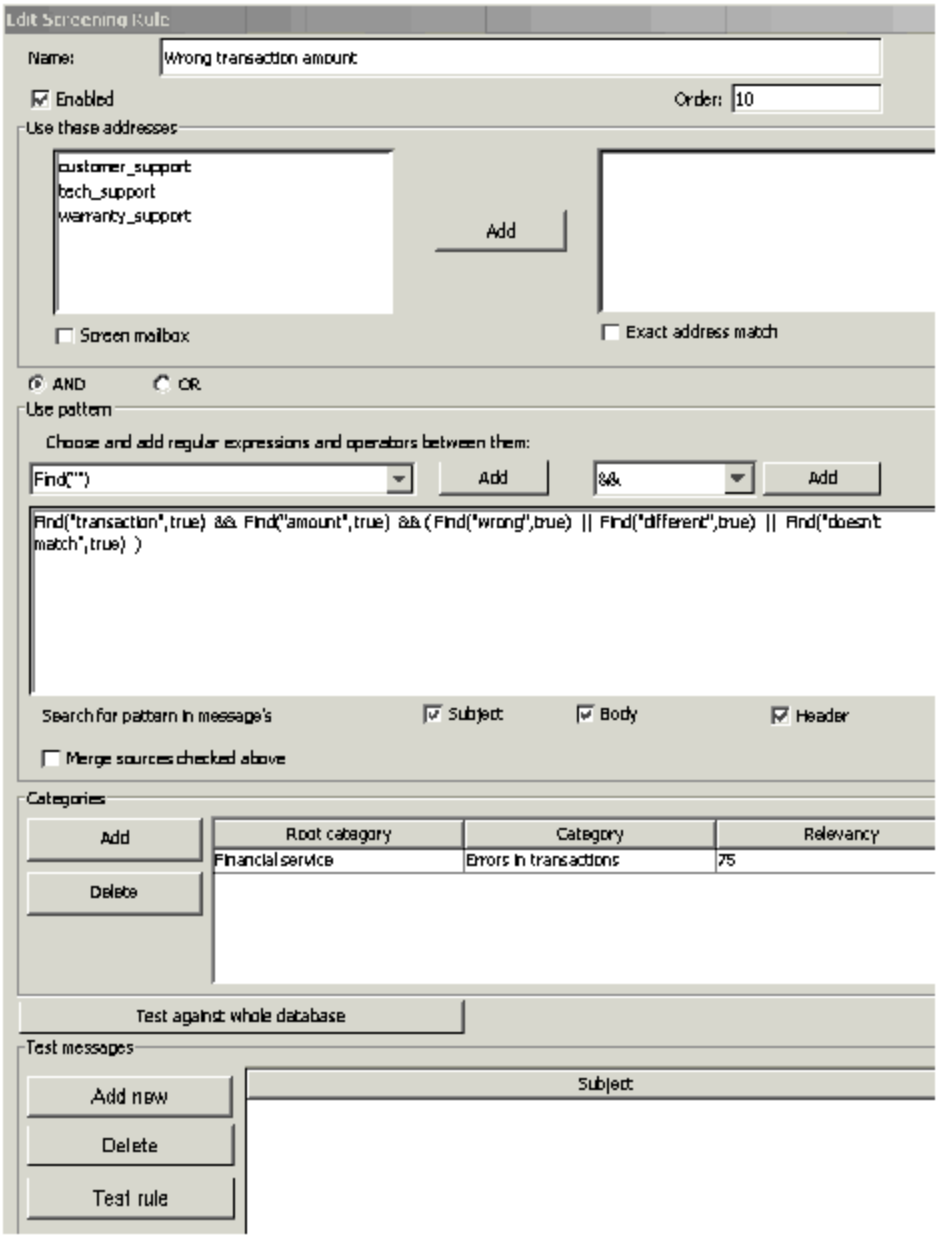
Procedure: Creating a Screening Rule

Purpose: To create or edit a screening rule.

1. Go to the Screening tab.
2. Do one of the following:

- Select New from the File menu.
- Right-click anywhere on the Screening tab and select New.
- Select an existing screening rule and double-click or select Edit from the File menu.

The Screening Rule Editor opens, as shown in "Screening Rule Editor." If you are creating a new screening rule, the title of the dialog box is New Screening Rule; if you are editing an existing rule, its title is Edit Screening Rule.



Screening Rule Editor

- Create or edit a name for the rule, observing the limitations on Knowledge Manager names. Names of categories, like those of all Knowledge Manager objects, can consist only of alphanumeric characters (A–Z, a–z, 0–9), plus the characters shown in Procedure: Creating a New Category, "Characters Allowed in Object Names."
- Use the Enabled check box to enable and disable the rule. The rule must be enabled in order to be available when you add a Screen object to a routing strategy. However, once a strategy includes a Screen object that uses a particular screening rule, disabling the rule does not prevent the strategy from using the rule.

- Use the Order box to specify the order in which you want this rule to apply with respect to other screening rules. This ordering applies only in a Multiscreen routing object in which All rules is selected.
- In the Use these addresses area, select an address from the left-hand window, then click Add to copy it to the right-hand window.
  - With the Exact address match box cleared, the rule looks for messages having that address as a substring of their destination address. For example, xyz@domainname.com matches abc.xyz@domainname.com and xyz@domainname.com.
  - With the Exact address match box selected, the screening rule looks for messages having that exact address as a destination. For example, xyz@domainname.com matches xyz@domainname.com but not abc.xyz@domainname.com. This match is not case sensitive.

> **Important**
>
> The e-mail addresses listed in the left-hand window of Use these addresses are defined in this configuration database object: Business Attributes > EmailAccounts > Attribute Values > Annex > general > address.
>
> Instead of moving addresses from the left-hand window to the right-hand, you can directly type an address in the right-hand window.
>
> You may leave the right-hand window empty, in which case the rule ignores the address in matching.

- Still in the Use these addresses area, select Screen mailbox to make the rule match the POP box from which the e-mail entered the eServices system, rather than the To field of the e-mail itself. The difference is that each e-mail enters the system from exactly one mailbox, while the To field can contain multiple addresses.

> **Important**
>
> For this feature to work as expected, the E-mail Server enable-same-mail-from-mailboxes option must be set to true. With this setting, E-mail Server creates a separate interaction for each address in the To field (that is, for each mailbox that it pulls the e-mail from when it creates the interaction).

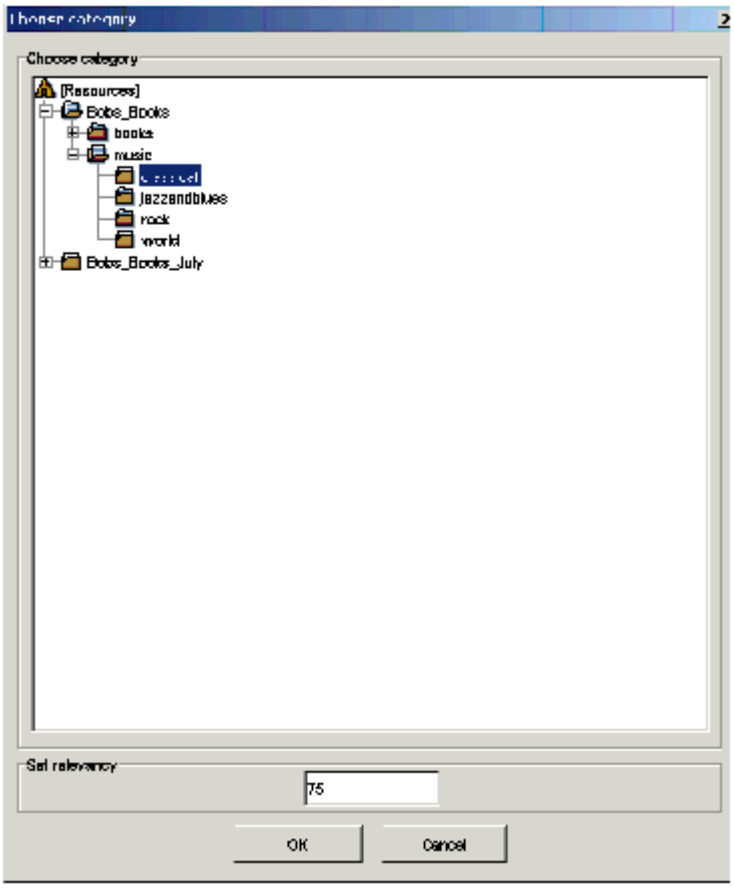
- Select AND or OR radio buttons:
  - AND —Match the addresses selected in Use these addresses *and* match the pattern defined in Use pattern In the Use pattern area, compose the rule, using the drop-down lists for functions and operators. See Step 9.
  - OR —Match the addresses selected in Use these addresses *or* match the pattern defined in Use pattern. In the Use pattern area, compose the rule, using the drop-down lists for functions and operators. See Step 9.
- In the Use pattern area, compose the rule, using the drop-down lists for functions and operators. See Functions and Arguments for an explanation of how to construct rules.
- Use the check boxes to have the screening rule apply to the message body, subject, header, or any

combination. You must select at least one.

> With multiple check boxes selected, there are two ways that a screening rule can behave. Use the `Merge sources checked above` checkbox as one way to control this; see Screening Rules Reference: Subject, Body, and Header for an explanation.

- In the `Categories` area, associate the screening rule with a category:

  a. Click Add. The Choose category dialog box appears, as shown in Choose Category.



Choose Category

> When an interaction matches this screening rule, the results are similar to classifying it using Genesys Content Analyzer: it receives a category name and a confidence level. The confidence level indicates that the system is x percent confident that this interaction belongs in this category. With Content Analyzer, the system assigns the category and confidence level. But with a screening rule, it is up to you, the user, to decide what category the

interaction belongs to and how confident you are of that categorization.

Do so as follows:

- Select the category you want to associate the screening rule with.
- Set the relevancy in the `Set relevancy` box.

Next Steps

- Test the Screening Rules.

Testing a Screening Rule

There are two ways to test a screening rule: on the UCS database, and on specially-created text.

Procedure: Testing screening rules on the UCS database

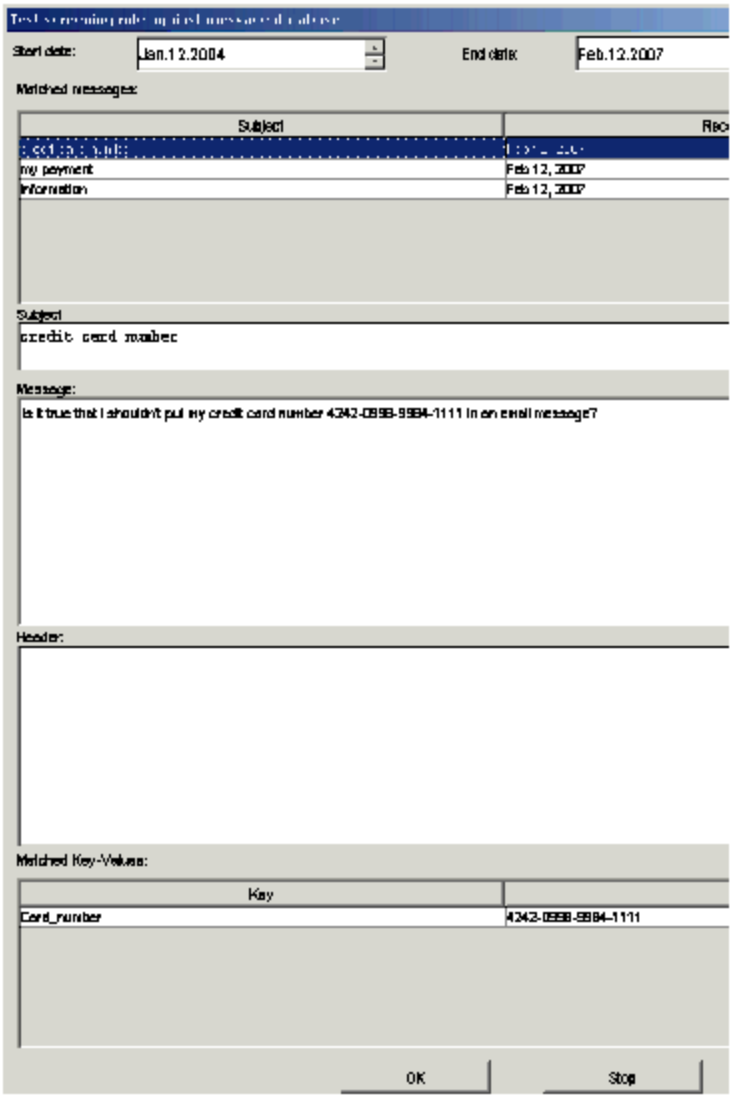
Purpose: To test a screening rule on all interactions in the UCS database that relate to the current tenant.

1. Click `Test against whole database` near the bottom of the `Screening Rule Editor`. The `Test screening rule against message database` dialog box appears, as shown in "Test Screening Rule Against Message Database Dialog Box."
2. Adjust the contents of the `Start date` and `End date` boxes to cover the desired timespan, then click `Find`.

> If there are a great many interactions and the testing process is taking too long, you can click `Stop`.

The results of the test display as follows:

- The `Matched messages` area lists all interactions that match the rule.
- The `Subject`, `Message`, and `Header` boxes show the text of the subject, body, and header of the interaction that is selected in the `Matched messages` area.
- The `Matched Key-Values` area shows all keys in the interaction's user data whose values match the rule. In the example in "Test Screening Rule Against Message Database Dialog Box," the screening rule created a key-value pair to hold the matched string, in this case a series of sixteen digits that could be a credit card number.



Test Screening Rule Against Message Database Dialog Box

Next Steps

- Test a rule on text that you create.

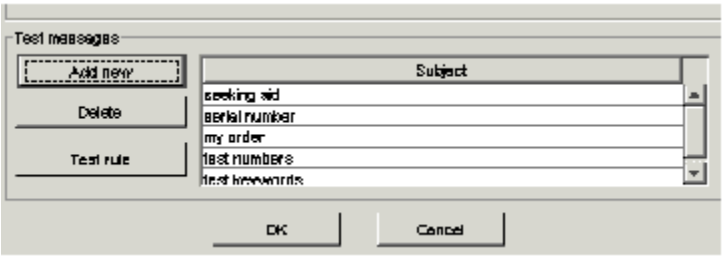
Procedure: Testing screening rules on specially-created text

Purpose: To test a screening rule on text that you create for that specific purpose.

Summary

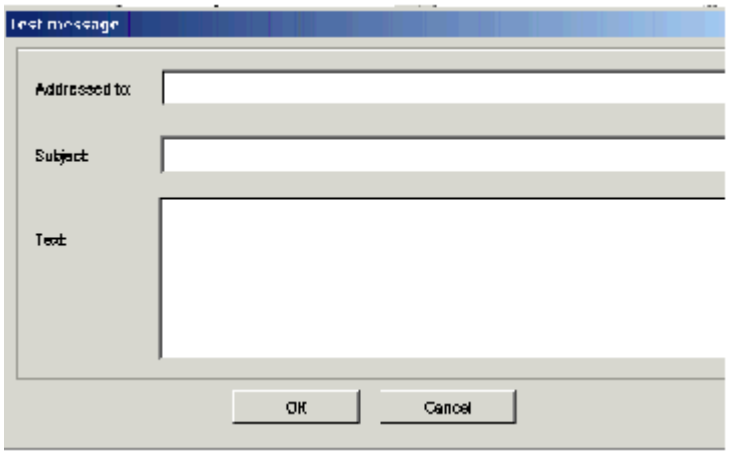
At the bottom of the `Screening Rule Editor` is the `Test messages` area, shown in "Test Messages

Area." Use this area to create messages that you can test screening rules on.



Test Messages Area

1. Click Add new to open the Test message dialog box, in which you can enter the address, subject, and body of a test message as shown in "Test Message Dialog Box."



Test Message Dialog Box

Click OK to return to the Screening Rule Editor.

- Manage existing test messages, as follows:
  - Double-click an existing test message to edit it.
  - Select a test message and click Delete to delete it.
- To test a rule:
  a. Select the rule on the Screening tab, then open the Screening Rule Editor..
  b. Select a test message from the list in the Test messages area.
  c. Click Test rule to apply the rule to the selected message.
  d. A message window appears, reporting whether the selected message matches the rule.

Next Steps

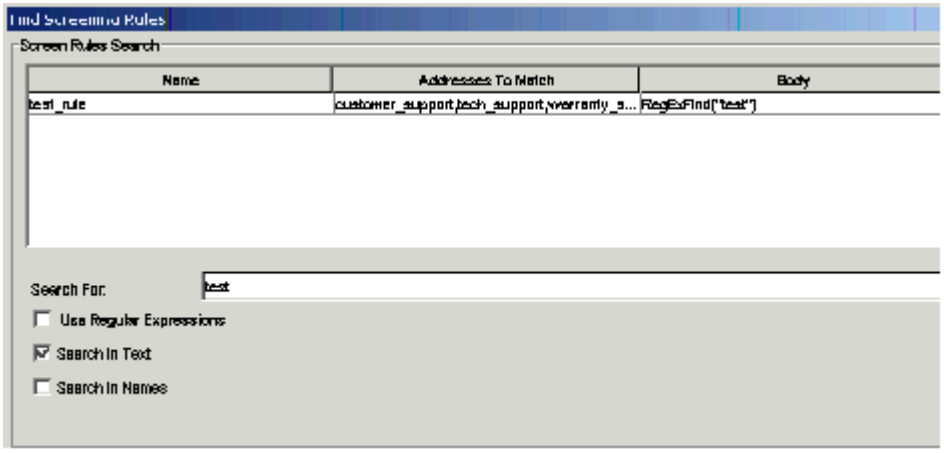
- Learn how to Search for Screening Rules.

> **Important**
>
> The test messages created by each user are stored separately, and are not synchronized between multiple Knowledge Manager instances. Test messages that you create are stored on your local machine under the default Documents folder of the logged-in Windows User folder, in the `UserInfo` file. This file is binary and must not be edited manually.

Finding a Screening Rule

Procedure: Finding a rule

1. Right-click anywhere on the Screening tab, then select Find Screening Rules. A dialog box with this name displays, as shown in "Find Screening Rule."



Find Screening Rule

2. Enter a string to search for in the Search For field.
3. Select Use Regular Expressions if you want the search to treat the string as a regular expression.
4. Select check boxes to search in the screening rule's text, name, or both.
5. Click Find. The dialog box displays the name, addresses to match, and body of all screening rules found.

Next Steps

- Learn more about how screening rules work:
  - Screening Rules Reference
  - Screening Rules Examples
  - Sentiment and Actionability
- Learn how to import and export Knowledge Manager objects:
  - Importing Knowledge Manager objects
  - Exporting Knowledge Manager objects

Screening Rule Reference

This section provides reference information on the following topics relating to screening rules:

- What they check: What Screening Rules Check
- The functions, arguments, and operators used in them: Functions and Arguments
- A pattern builder to help formulate them: Pattern Builder
- Regular expressions to use in them: Regular Expressions

What Screening Rules Check

Screening rules check the following parts of an interaction, depending on what you select in the Screening Rule Editor, and on the settings in the IRD screening objects:

- The subject, if you select that check box.
- The body, if you select that check box.
- The header, if you select that check box. See also "Subject, Body, and Header" below, on how screening rules behave if two or more of the preceding are selected.
- The destination address, if you have put anything in the right-hand box in the `Use these addresses` area of the `Edit Screening Rule` dialog box.
- The value of any key in the user data, if both of the following are true:
    - In the Multiscreen or Classify strategy object, you select a key in the `User data key if specified` drop-down list under `Get screened data from.`
    - In the Screening Rule Editor, you select the `Body` check box in the `Use pattern` area of [(see See]]. Use the check boxes to have the screening rule apply to the message body, subject, header, or any combination. You must select at least one. on See Use the check boxes to have the screening rule apply to the message body, subject, header, or any combination. You must select at least one.).

  User data is first associated with the interaction by the media server when it creates that interaction. As an example, E-mail Server associates the following user data with the interaction:

    - FirstName (from Contact information)
    - LastName (from Contact information)
    - Mailbox (value of the `address` option in the `pop-client` section of the E-mail Server Application object)
    - To (MIME header field)
    - Subject (truncated to 512 characters)
    - FromAddress (personal part of From header field)
    - FromPersonal (e-mail address part of From header field)

- All Header fields (except Received, Return-Path, X-MIMETrack, Subject, Sender, From, To, Cc, Bcc) prefixed by `Header_`
- All parent attached data (originally created by E-mail Server) which can be inherited; that is, all parent attached data:
  - not starting with `Header_`
  - not starting with _ (underscore)
  - not equal to `GEM_Failure`
  - not equal to `GEM_FailureMsg`
  - not equal to `GEM_FailureArgs` l

User data may then be added or modified by a routing strategy.

Subject, Body, and Header

If you select more than one of the Subject, Body, and Header areas, a screening rule can behave in the following two ways:
The default behavior is for the rule to apply to each area in turn; for example, with `Subject` and Body selected, the rule applies first to the Subject, then to the Body.

The alternative behavior is for all selected items to first be concatenated so that the rule applies to all at once. There are two ways to achieve this alternative behavior.

- To enforce it for all screening rules, set the `subject-body-header` option for both Knowledge Manager and Classification Server to `true`.
- To enforce it for a particular screening rule:
  a. Leave subject-body-header set to false.
  b. Open the rule in the Screening Rule Editor.
  c. Select the `Merge sources checked above` check box.

> **Important**
> Setting subject-body-header to `true` overrides any selection of the Merge sources checked above check box for a particular rule.

Functions and Arguments

Functions

Screening rules can use three basic functions:

- `Find("<text>")`, where `<text>` is a text string. It returns the result `true` if the interaction contains the exact string between quotes, ignoring case.
- `RegExFind("<regular expression>")`, where `<regular expression>` is a regular expression (see Regular Expressions). It returns the result `true` if the interaction contains any string that matches the regular expression between quotes.
- `RegExMatch("<regExp>")`, where `<regular expression>` is a regular expression. It returns the result `true` only if the entire content of the interaction matches the regular expression between quotes.

> **Important**
>
> RegExFind and RegExMatch are the same except that RegExFind looks for a match anywhere in the body of the interaction, whereas RegExMatch demands that the entire body of the interaction match the regular expression.

Arguments

All functions have one required argument, which must appear between double quotation marks, as represented above (`<text>`) or (`<regular expression>`). This required argument can be followed by one or two optional arguments, depending on the function. The full form of each function, including all arguments, is as follows:

```
Find("<text>", <IgnoreCase>)
RegExFind("<regular expression>",<"key">,<IgnoreCase>)
RegExMatch("<regular expression>",<IgnoreCase)
```

IgnoreCase

The IgnoreCase argument must be a Boolean value (*true* or *false*). All three functions ignore case in searches unless you include the IgnoreCase argument with a value of `false`.
For example:

- `Find("pacific")` finds *Pacific* and *pacific.*
- `Find("Pacific",false)` finds *Pacific* but not *pacific.*

You can also substitute `true` for `false`—for example, `Find("Pacific",true)`—which means that case is ignored. So `Find("Pacific",true)` is the same as `Find("Pacific")`.

Key

The key argument must be a string. If this argument is present, the system creates a key-value pair with the following characteristics:

- The key name is the string specified by the key argument, prefixed by `ScrKey_`.
- The value is the material that the screening rule matches.

The system then adds this key-value pair to the interaction's attached data.
For example, `RegExFind("[A-Z]\d\d\d","ID_code",false)`:

1. Finds strings consisting of a capital letter followed by three digits (see Regular Expressions) .
2. Attaches to the interaction a key-value pair called `ScrKey_ID_code` whose value is A123, X005, M999, or whatever the function found in this interaction to match the regular expression.

Screening Rule Operators

Operators

Operators are of two types:

- Binary operators join two functions.
- Unary operators operate on a single function.

`&&` is the binary operator "and". For example,

```
Find(“interest rate”) && Find(“APR”,false)
```

matches a message only if it includes both "interest rate" and "APR."

`||` is the binary operator "or." For example,

```
Find(“station wagon”) || Find(“convertible”)
```

matches any message that includes either "station wagon" or "convertible" (or "Station Wagon" or "station Wagon" or "Convertible").

`!` is the unary operator "not." For example,

```
!Find(“windows”)
```

matches any message that does not include the word "windows."

You can combine `!` with a binary operator. For example,

```
Find(“bird”) && !Find(“goose”)
```

matches any message that includes "bird" but does not include "goose."

Operator Precedence

`p && q || r` is parsed as `(p && q) || r`. For example, consider:

```
Find(“debt”) && Find(“income”) || Find(“profit”)
```

To paraphrase, this screening rule is basically “find X or find Y,” where X is "debt" and "income," and Y is "profit."
It matches both "debt exceeds income" and "profits are fantastic".

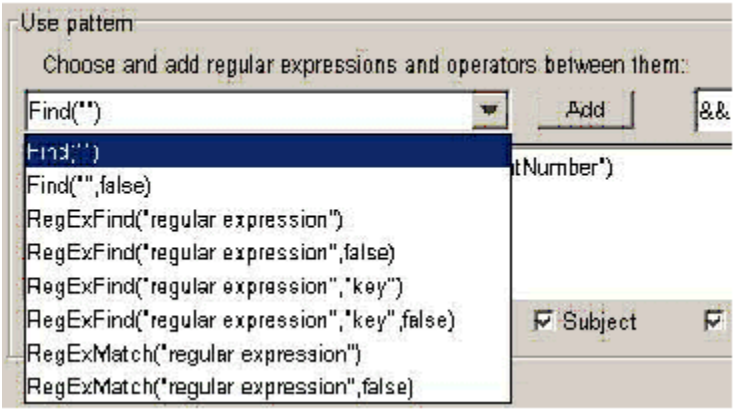
You can modify the default precedence by the explicit use of parentheses; for example:

```
Find(“debt”) && (Find(“income”) || Find(“profit”))
```

This screening rule is basically “find X and find Y,” where X is "debt" and Y is either "income" or "profit."
It matches both "debt exceeds income" and "debts impact profit."

Pattern Builder

Knowledge Manager includes a pattern builder that offers the choice of each function type in all possible forms, with and without optional arguments, for a total of eight, as shown in "Regular Expressions in Knowledge Manager."



Regular Expressions in Knowledge Manager

After you select a form and click `Add`, you must put text between the quotation marks. More specifically, you must:

- For `Find`, put text between the empty quotation marks.
- For `RegExFind` and `RegExMatch`, substitute your desired text for `regular expression` and/or `key`.

Regular Expressions

A regular expression stands for, not one particular character string, but a class of character strings. For example, suppose that you want to find all interactions with U.S. Zip codes in them. U.S. Zip codes are five-digit numbers, so you could in theory write about 9,000 screening rules (`Find("00000")`, `Find("00001")`, `Find("00002")`, and so on).

Fortunately, you can use the special symbol `\d`, which stands for any digit, to write a screening rule using a regular expression: `RegExFind("\d\d\d\d\d")`. This screening rule matches any sequence of five digits.
There are often several different ways of writing the same regular expression.

For instance, two items separated by a hyphen and enclosed in square brackets denotes a range of which the two items are endpoints. So `[a-d]` matches a, b, c, or d, and `[5-8]` matches any digit between 5 and 8; hence `\d` is the same as `[0-9]`.

> **Important**
>
> In general usage, apart from Genesys eServices, regular expressions are case sensitive. However, in Knowledge Manager regular expressions are not case sensitive unless you add `,false` as described above.

The table "Elements of Regular Expressions" lists some of the most commonly-used elements of regular expressions:

Elements of Regular Expressions

Symbol	Meaning	Example
.	Any character, including space	`b.t` matches *bat, bet, bit,* and *but.*
\d	Any digit	`\d\d` matches any pair of digits from 00 to 99.
\s	Space	`\d\s\d` matches 1 0, 5 9, and so on.
*	Zero or more instances of the preceding expression	`o*f` matches *oof, of,* and *f.* `me.*d` matches *med, mead,* and *meed.*
+	One or more instances of the preceding expression	`bre+d` matches *bred, breed* and*breeed.*

Symbol	Meaning	Example
?	Zero or one instances of the preceding expression	`c?rude` matches *rude* and *crude.*
`{x}`	X instances of the preceding expression	`st.{2}k` matches *steak, stork* , and *stink.*
^	Any character except the following	`s[^e]t` matches *sat, sit,* and *sot,* but not *set.*
`[]`	Any characters or ranges within the brackets	Any characters: `b[aeiou]at` matches *boat* but not *brat.* Any range(s): `[0-9]th` matches *5th, 6th, 7th* `[a-z]` matches any lowercase letter; `[A-Z]` matches any uppercase letter.
\	Turns off the special meaning of the following symbol	`\*` matches the character * (asterisk);`\.` matches the character . (period or full stop).
|	Or	`[b|p]ig` matches *big* and *pig.* Do not be confused: `|` means *or* in regular expressions, but `||` means *or* as one of the Operators used in screening rule formulas.

Here are some other points to keep in mind:

- Space is just another character. The regular expression `savings account` contains a space, and so it does not match the string *savingsaccount.*
- Word boundaries are not considered. The regular expression `read` matches not only *read,* but also *reader, ready, spread, bread,* and so on.
- Use parentheses to group parts of regular expressions together. For example, `RegExFind("(\d{3}\.){2}")` puts `\d{3}\.` in parentheses so that the number-of-instances item `{2}` applies to the all of `d{3}\.`, not just to `\.` This expression matches any group of three digits plus period plus any three digits plus period (for example, 198.351.). Further examples are provided in Examples of Screening Rules.
- Regular expressions make use of many more special characters and operators than those listed in the table "Elements of Regular Expressions." Much documentation on regular expressions is available on the Web. Because Genesys Knowledge Management uses Java classes for regular expressions, it is best to consult documents describing the particular version of regular expressions used in Java.

Examples of Screening Rules

This section provides examples of screening rules.

Credit Card Number

To find text that includes a typical credit card number, you need to match a sequence of four groups of four digits, each group separated by -(hyphen):

```
\d\d\d\d\-\d\d\d\d\-\d\d\d\d\-\d\d\d\d
```

> **Important**
>
> This regular expression also works without the \ (backslash) before the hyphens. However, it is better practice to write \- for the character hyphen, because the hyphen also has a special use in range expressions like [a-z].

Or if you want to allow for the possibility that some people will omit the hyphens, use? to make the hyphen optional:

```
\d\d\d\d\-?\d\d\d\d\-?\d\d\d\d\-?\d\d\d\d
```

You could also use the repetition notation to shorten each \d\d\d\d to \d{4}.

North American Phone Number

North American phone numbers consists of ten digits, grouped into two groups of three and one of four. There are a number of ways for the groups to be separated:

203-555-1234

(203) 555-1234

(203)555-1234

203 555-1234

203.555.1234

The following regular expression matches all of the above:

```
(\d\d\d|\(\d\d\d\))[\s\.\-]?\s*\d\d\d[\-\.]\d\d\d\d
```

The table "Phone Number Regular Expression" analyzes this regular expression.

Phone Number Regular Expression

Symbols	Meaning	Remarks
\d\d\d	Three digits	
\d\d\d|\(\d\d\d\)	Three digits, or three digits enclosed in parentheses	\ turns off the special meaning of the character (
[\s\.\-]?	Space or period or hyphen or zero	Any one of the items enclosed in square brackets, either once or not at all
\s*	Zero or more spaces	
\d\d\d	Three digits	
[\-\.]	Hyphen or period	Note again the need to use \
\d\d\d\d	Four digits	

Telltale Words

To screen for interactions from dissatisfied customers, you might try a regular expression like the following:

```
(not\s([a-z]+\s)*(pleased | satisfied)) | unhappy | complain
```

The first part of this expression matches *not* followed by zero or more words followed by *pleased* or *satisfied*; for example, *not* very pleased, not satisfied, not at all satisfied (but it also matches strings like *can not believe how pleased I am*). The rest matches the single words "unhappy" and "complain."

Screening for Sentiment and Actionability

You can use Genesys Knowledge Management to analyze interactions that have been brought into the system by Genesys Social Messaging Management.

Genesys supplies a sample that analyzes these interactions for:

- The sentiment expressed—Positive, negative, or neutral.
- Actionability—Whether the interaction calls for attention from an agent.

To use this sample, import the file `SentimentAndActionabilityScreeningRules.kme`, which is located in the `KnowledgeManagerHome\SentimentModel` directory. See the section Importing Knowledge Management Objects.

This file provides examples of screening rules for detecting sentiment and actionability, plus category trees containing categories that are assigned to interactions that match the rules.

If you have Genesys Content Analyzer, you can use it to analyze sentiment and actionability, as described in Analyzing Sentiment and Actionability with Content Analyzer.

For more information on Genesys Social Messaging Management, see the eServices Social Media Solution Guide.

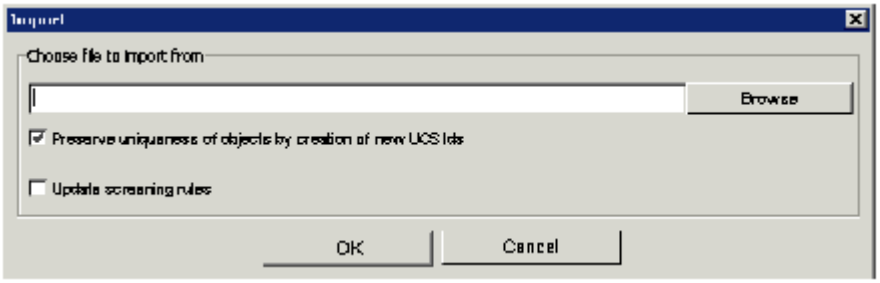
Importing Knowledge Management Objects

You can import categories, standard responses, field codes, screening rules, training objects, and models (training objects and models are restricted to Genesys Content Analyzer; see Genesys Content Analyzer for more information.

Use exported files as backups or to transfer objects between environments.

> **Important**
> At this time you cannot import or export FAQ objects.

1. On any tab, select `Import`. The `Import` dialog box opens, as in "Import Dialog Box."



Import Dialog Box

2. Browse to the file that you want to import, or enter its path and name.

The check box `Preserve uniqueness of objects by creation of new UCS Ids` controls whether the imported records receive new database IDs. Genesys strongly recommends that you leave this check box selected; otherwise, the imported records keep their old IDs and there is a risk of creating uniqueness conflicts. It is only safe to keep old IDs when you are importing into an empty database. One reason to keep the old IDs would be to preserve compatibility with non–Knowledge Manager components (such as a routing strategy) that need to refer to them.

> **Important**
> If the Preserve uniqueness check box is cleared and an imported object comes in with an ID that is identical with an existing object's ID, the import process

cancels.

The check box `Update screening rules` controls whether imported screening rules overwrite existing screening rules with the same name. If this check box is not selected, screening rules are treated like all other Knowledge Management objects, as described in the preceding paragraph. If the check box is selected, screening rules are treated differently from all other objects: If the names match, the imported screening rule replaces the existing one.

Important

If the imported rule's name does not match any existing rule but its database ID happens to match that of an existing rule, then Knowledge Manager creates a new ID for the imported rule.

- If a root category being imported has the same name as an existing category, Knowledge Manager asks you to change the name of the category being imported.

If other objects have the same name as existing objects, Knowledge Manager appends _<hms> to the name of the imported object. <hms> is a timestamp where h is the hour (using a 12-hour clock),m is the minutes, ands is the seconds. Each unit may be one or two digits; there is no padding. For example, if at 4:25:07 PM you import a screening rule called `Sales`, and there is also an existing rule called `Sales`, the new name of the imported rule is `Sales_4257`.

Important

This adds between four and seven characters to the name of the object. You should be especially careful of this if any imported object's name is more than 58 characters long: the added characters may produce a new name that violates the 64-character limit on names of Knowledge Management objects. Importing may fail on objects with names that are too long.

- Click `OK`. All objects in the selected file are imported.
- A message appears reporting on any errors or warnings encountered, see "Import Results Message."

Import Results Message

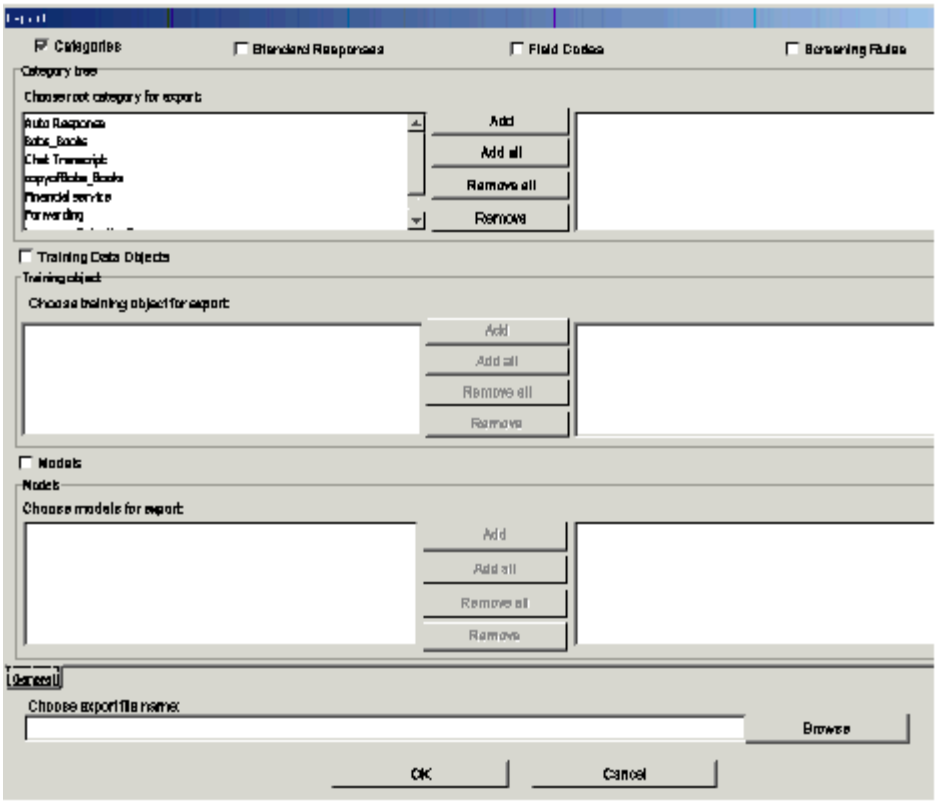
Exporting Knowledge Manager Objects

You can export categories, standard responses, field codes, screening rules, training objects, and models. Training objects and models are restricted to Genesys Content Analyzer. Use exported files as backups or to transfer objects between environments.

> **Important**
> At this time you cannot import or export FAQ objects.

Procedure: Exporting Knowledge Manager objects

1. On any tab, select `Export`. The `Export` dialog box opens, as shown in "Export Dialog Box."



Export Dialog Box

2. Select the object(s) to export, using the check boxes at the top of the dialog box.

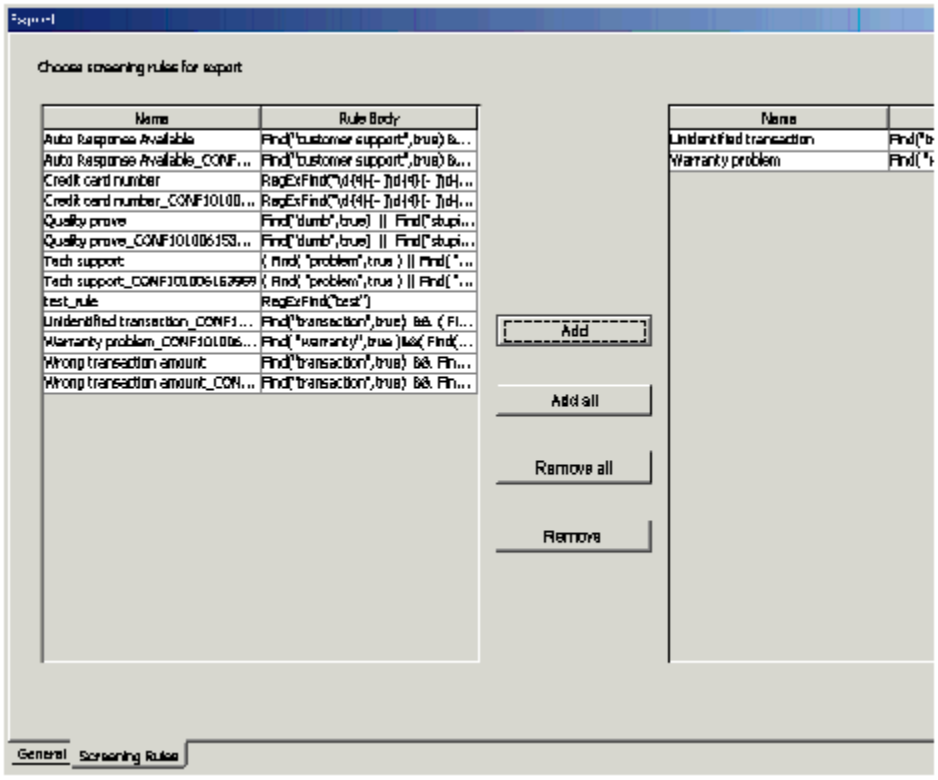
   - Categories—Exporting categories is a prerequisite for exporting standard responses, training objects, or models.
   - Standard Responses—Knowledge Manager exports the standard responses belonging to the category tree that you selected.
   - Field Codes—No selection is possible. You must export all or none.
   - Screening Rules—If you select this checkbox, a Screening Rules tab appears towards the bottom of the dialog box. You must make further selections on this tab, as described in See Make further selections for screening rules: On the Screening Rules tab, move rules from the list of all rules (left-hand field) to the list of selected rules (right-hand field).

- Make further selections for category trees, training objects, and models.

  - Each type of object has two fields in the Export dialog box: one on the left listing available objects, another on the right listing the objects selected for export. Use the `Add/Add all/Remove all/Remove` buttons to move objects between the two fields.
  - The category structures that are selected (that appear in the right-hand category field) determine which training objects and models appear in the Training object and Models areas.

- Make further selections for screening rules: On the `Screening Rules` tab, shown in "Export: Screening Rules

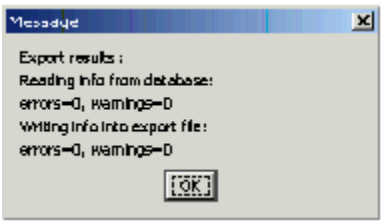
Tab," move rules from the list of all rules (left-hand field) to the list of selected rules (right-hand field).



Export: Screening Rules Tab

> **Important**
>
> When selecting objects to export, keep in mind that importing will import all objects in the exported file. You cannot select which objects to import.

- Enter a name for the exported file and click OK. The file is created in the directory containing the Knowledge Manager executable.
- A message appears reporting on any errors or warnings encountered; see "Export Results Message."



Export Results Message

Knowledge Management Response Times

This section describes some typical response times for Knowledge Manager tasks. These figures are for a machine with two Pentium 4 processors and 1 GB of RAM.

- Loading and refreshing a category tree takes 4–6 seconds per 2000 categories (each category having one standard response).
- Copying and pasting a category tree takes approximately 100 seconds per 1000 categories.
- Deleting a category tree takes approximately 100 seconds per 1000 categories.

Knowledge Management Roles

Starting with eServices 8.0.1, you can define different levels of access to Knowledge Manager objects. For example, you might create a role called Supervisor which is allowed to create new standard responses but not to delete or approve them.

You define Roles in Genesys Administrator.

> **Important**
>
> Other aspects of eServices configuration can be done either in Configuration Manager or Genesys Administrator, but Genesys highly recommends using only Genesys Administrator for configuring Roles and Tasks.

Logically, a Role is a mapping between `Person` or `Access Group` objects on the one hand, and permitted tasks on the other. The `Person` and `Access Group` objects are created by you to suit your needs, while the tasks are predefined by eServices.

Most tasks are defined as the possible application of three actions (create, modify, and delete) to five KM objects (category, standard response, field code, screening rule, and training object); there are several more tasks such as Approve Standard Response. The complete list is given in the table Example Set of Roles.

> **Important**
>
> To enable a Person to *modify* standard responses, you must assign both the View Standard Response and the Modify Standard Response tasks to them.
>
> To enable a Person to *approve* standard responses, you must assign them (a) the Approve Standard Response task and (b) either the View Standard Response or the Create Standard Response task.

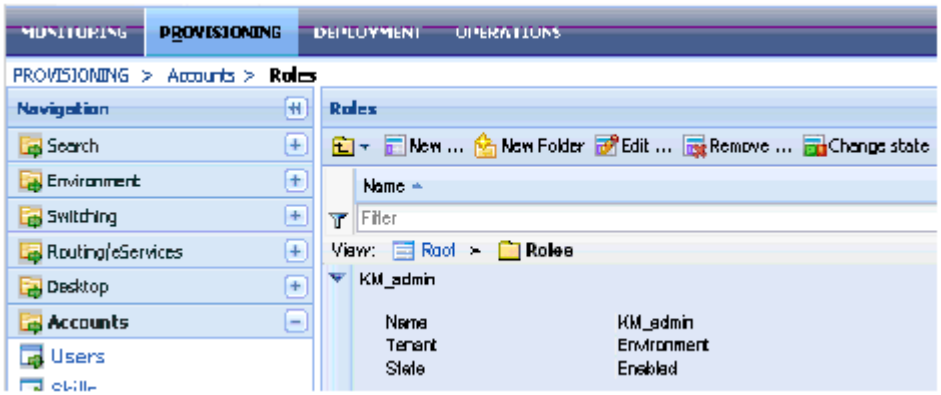
Defining an Access Role

Prerequisites

- The Knowledge Manager `disable-rbac` option has its default value `false`.
- Genesys Administrator is running.
- You have imported the Knowledge Manager metadata template into Knowledge Manager Application object.

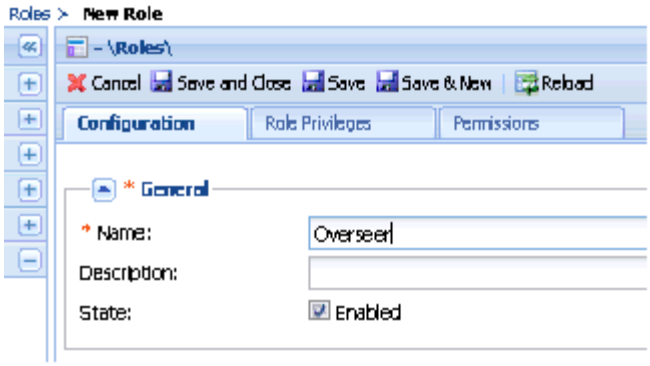
> **Important**
> For information on using Genesys Administrator, see the Genesys Administrator Help.

1. In Genesys Administrator, navigate to `PROVISIONING > Accounts > Roles`, as shown in "Roles in Genesys Administrator." Be sure that the correct tenant shows in the drop-down list.



Roles in Genesys Administrator

2. Click `New`. On the Configuration tab:
   a. In the General section, enter a name, as shown in "Configuration: General."

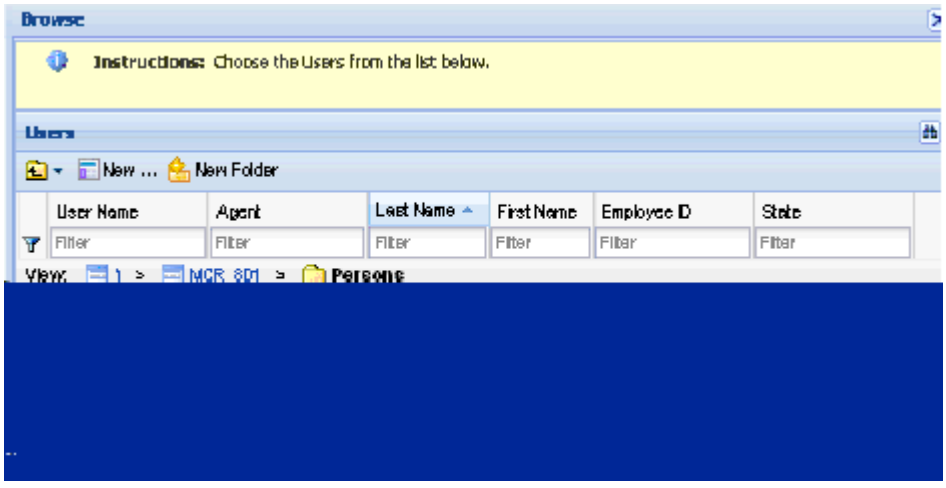


Configuration: General

b. In the Members section, click Add for either Users or Access Groups, as shown in "Configuration: Members."

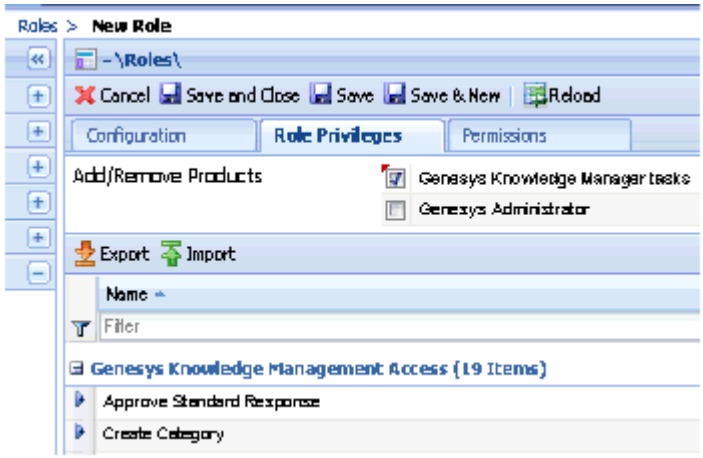
Configuration: Members

c. In the resulting Browse window, select Users or Access Groups to add to the list of members. The figure "Browse for Users" shows the Browse window for Users.



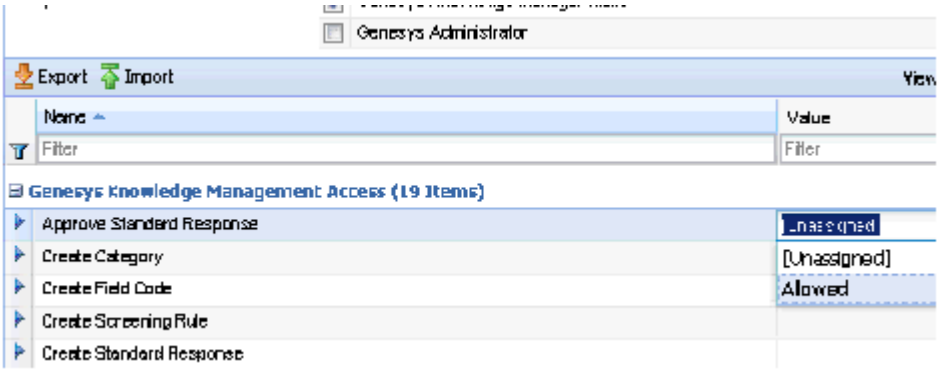
Browse for Users

4. On the `Role Privileges` tab, select the `Genesys Knowledge Manager tasks` box, as shown in "Knowledge Management Tasks." If this box does not appear, you must import the metadata template into your Knowledge Manager Application.



Knowledge Management Tasks

5. A list of tasks appears. To enable this new role to perform one of them, select the task and change its value to `Allowed`, as shown in "Assigning a Task."



Assigning a Task

> **Important**
>
> The value [Unassigned] means that this Role cannot perform this task. But if multiple Roles are assigned to the same User or Access Group, this setting is overridden if another Role sets this privilege as Allowed.

- On the `Permissions` tab, add Users or Access Groups, as described in Genesys Administrator Help.

Examples of Knowledge Management Roles

The table "Example Set of Roles" lists all of the tasks provided for Knowledge Manager and provides an example of allocating task permissions among four different Roles called Manager, Administrator, Supervisor, and Agent. This is only an example; you can define any number of roles, with any names.

Example Set of Roles

Task	Manager	Administrator	Supervisor	Agent
Create Standard Response	X	X		
View Standard Response	X	X	X	X
Modify Standard Response	X	X	X	
Delete Standard Response	X	X		
Approve Standard Response	X			
Create Category	X	X		
Modify Categories	X	X	X	
Delete Categories	X	X		
Create Screening Rule	X	X		
Modify Screening Rule	X	X	X	
Delete Screening Rule	X	X		
Create Field Code	X	X		
Modify Field Code	X	X	X	

Task	Manager	Administrator	Supervisor	Agent
Delete Field Code	X	X		
Create Training Object	X	X		
Modify Training Object	X	X	X	
Delete Training Object	X	X		
Import Knowledge Management Objects	X			
Export Knowledge Management Objects	X			

Genesys Knowledge Management: Content Analyzer

Genesys Content Analyzer is an option to Genesys E-mail, requiring an additional license. It adds natural language processing technology to Genesys Knowledge Management.

The Knowledge Manager Overview section provides an outline of the structure of the Knowledge Manager user interface.

Models

Genesys Content Analyzer applies a classification model—a statistical representation of a category tree—to an incoming interaction and produces a list of the categories that the interaction is most likely to belong to. Each likely category is assigned a percentage rating indicating the probability that the interaction belongs to this category.

> **Important**
>
> Classification, like screening, can operate on any interaction that has text somehow associated with it, whether as the body of the interaction (e-mail, chat), or otherwise (as user data, for example). In practice, it is expected that most interactions that are screened or classified will be e-mail messages; therefore this Guide uses the terms *e-mail* and *message* interchangeably to refer to these interactions. In fact whatever is said here about e-mail applies to any interaction that has associated text.

Training Objects

The process of creating a model is called *training.* Training operates on a *training object,* which is a category tree plus a set of text objects, with each text object assigned to one category in the tree.

Import and Export

You can import and export training objects and models. This is also a means of copying training objects.

Components

Genesys Content Analyzer does not have components as such. Rather, it adds functionality to the components of Genesys Knowledge Management:

- It activates Training Server, which has no function in the basic Genesys Knowledge Management but is required for training models.
- It enables Classification Server to categorize incoming interactions using models.
- It enables Knowledge Manager to control the creation of training objects, classification models, and FAQ objects.

The following topics also deal with Genesys Content Analyzer:

- Training
- Testing Models
- Using and Rating Models
- Design and Use Considerations
- Language Detection Model
- Analyzing Sentiment and Actionability with Content Analyzer
- FAQ Objects
- Typical Response Times

Training

This topic describes part of the functionality of Genesys Content Analyzer.

Training consists of the following steps:

1. Creating a training object.
2. Scheduling and running the training.

These steps include various actions, such as:

- Create a New Training Object
- Add Outbound Messages
- Copy a Training Object
- Add More Text Objects to the Training Object
- Add Uncategorized E-Mails Using the TO Data Analyzer
- Schedule Training
- Content Analyzer Language
- Large Training Objects

All of these actions are performed by Training Server according to settings you provide using Knowledge Manager, which also displays the results.

Creating a Training Object

A training object combines a category tree and a set of text objects, with each text object assigned to one category in the tree (categorized). The text objects are typically e-mails, but you may choose to have the set of text objects also include the standard responses associated with the category tree.

Training scans the text objects and forms a statistical model of the words and phrases that tend to occur in each category. This is why you may want to have the training scan standard responses: they are very likely to include many words and phrases diagnostic of their category. This is also the reason that more text objects is better: it increases the sample size, which increases the accuracy of the model.

There are five possible sources of categorized text objects:

- E-mails that agents have assigned to categories, as described in this section.
- Text objects (in the form of e-mails) that you create in Knowledge Manager.
- Responses that you use when creating standard responses.
- Objects that you use when creating other training objects.

- E-mails that you categorize when adding uncategorized e-Mails using the TO Data Analyzer.

Once you have completed creating a category tree, you can create a training object.

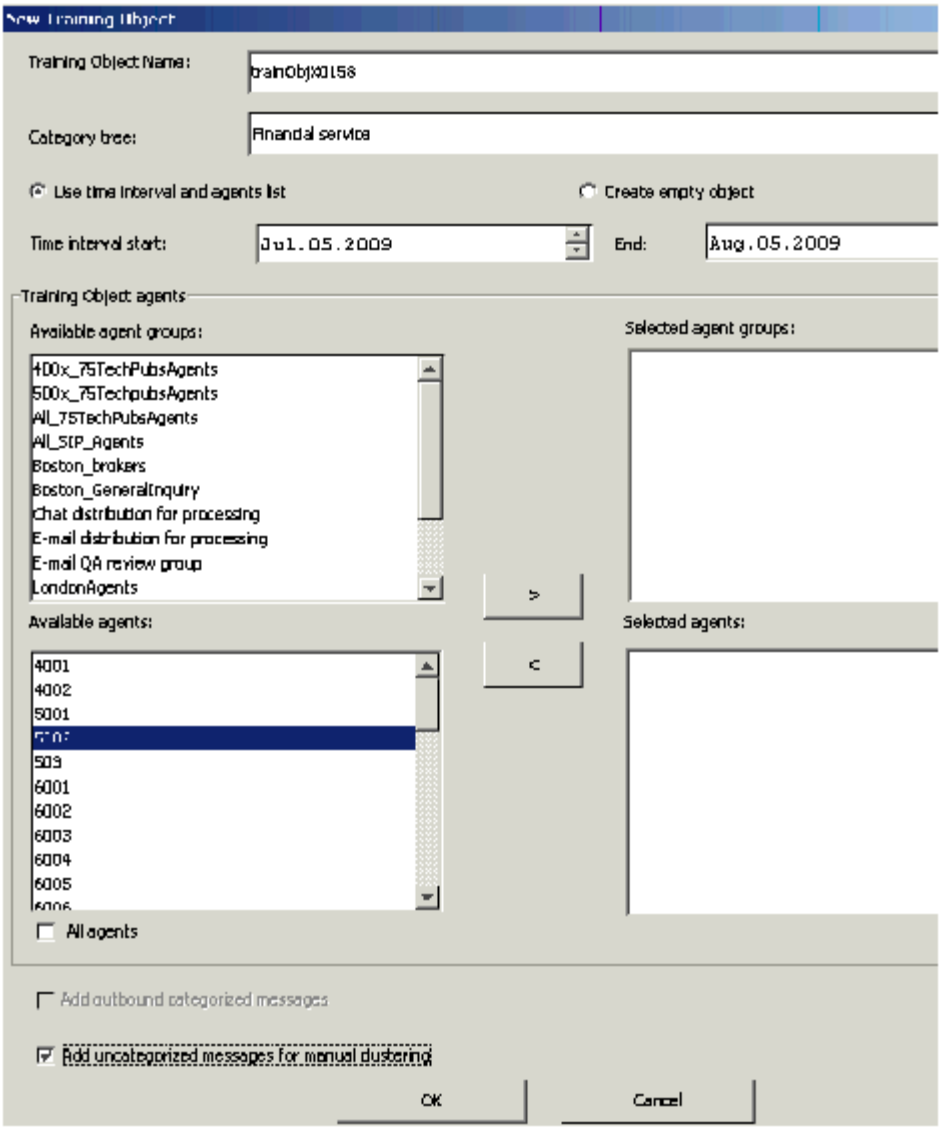
Procedure: Creating a New Training Object

This topic describes part of the functionality of Genesys Content Analyzer.

Purpose: To create an object that can be used to produce a classification model.

Prerequisites: A category tree.

1. In Knowledge Manager, select a tenant and language. This determines which category trees are available. See Notes on Language for some considerations relevant to choosing a language.
2. On the Training tab, select `File > New`, click the `New` icon, or right-click anywhere and select `New Training Object` on the shortcut menu. The New Training Object dialog box appears, as shown in "New Training Object Dialog Box."



New Training Object Dialog Box

3. Enter a name for the object (see Language and Dictionary Names for restrictions on the names of Knowledge Manager objects) and select a category structure (category tree).
4. You now have two alternatives, depending on whether you already have a collection of e-mails that have been categorized:
   a. If you do not already have a collection of categorized e-mails, select the `Create empty object` radio button.
   b. If you already have categorized e-mails, the categories were assigned to each e-mail by the agent who handled the interaction. In adding categorized e-mails to a training object, you can simply add all e-mails categorized by any agent at any time, or you can make the following adjustments:

- Use only the e-mails that were categorized during a certain time period. You may want to exclude earlier times when agents were still learning to use the category system. To do this, set dates in the `Time interval start` and End fields.

> **Important**
>
> In defining the time interval, be careful to avoid these common errors: (1) identical start and end dates, resulting in zero e-mails in the training object; (2) too large a time span, which can result in omission of the categorization of the latest e-mails in the set; (3) time span of only a few days, which may lead to omission of the results for the last date.

- Use only the e-mails that were categorized by selected agents or groups (some agents or groups may be especially reliable in their choice of categories). To do this, use the > and < buttons to move agent names from the `Available agent groups` window or `Available agents` window to the `Selected agent groups` window or the `Selected agents` window.

> **Important**
>
> Agent names appear in alphabetical order. For agent groups to appear in the `Available agent groups` box, they must be configured in the Tenant that you are creating this training object in. To move individual agents, you must first clear the `All agents` check box. There is a limitation on the number of SQL request parameters when creating a new training object.

Note: The limitation on the number of SQL request parameters becomes relevant when you are dealing with a large number of agents. For example, you will need a correspondingly large number of instances of the parameter `Interaction.OwnerId = xxxxxx`. The limitation varies with the RDBMS you are using, but it is approximately 2,100.

- The `Add outbound categorized messages` check box is enabled only if Knowledge Manager is operating with UCS 7.5. See further explanation below.
- Select the `Add uncategorized messages for manual clustering` check box if you want to add uncategorized e-mails to the root category. You can then use the TO Data Analyzer to categorize them, as described in Adding e-mails to a category on the Main tab.

> **Important**
>
> The uncategorized e-mails added by this step are from the time interval and/or agents that you designated in the previous step. If you made no designations, the

> uncategorized e-mails are from your entire UCS database.

Next Steps

- If you are operating Knowledge Manager with UCS 7.5, review Add Outbound Messages.
- To add more text objects, continue with Adding More Text Objects to the Training Object and Adding Uncategorized E-Mails Using the TO Data Analyzer.
- To proceed with training, continue with Schedule Training.

Add Outbound Messages

This topic describes part of the functionality of Genesys Content Analyzer.

The `Add outbound categorized messages` check box, shown in the figure "New Training Object Dialog Box," behaves differently depending on the version of UCS that Knowledge Manager is operating with.

- If Knowledge Manager is operating with UCS 7.6 or later, the checkbox is disabled. The new training object includes both of the following (as delineated by any time interval and agent list that you set):
  - Categorized inbound e-mails
  - Uncategorized inbound e-mails whose parent is a categorized outbound e-mail.
- If Knowledge Manager is operating with UCS 7.5, the check box is enabled, and has the following effects:
  - Not selected (the default): The training object includes only categorized messages that are inbound and are first in their thread.
  - Selected: The training object includes all categorized messages, both inbound and outbound.

Selecting this check box may be helpful in a situation in which agents do not categorize inbound e-mails, but do categorize the outbound e-mails that they generate in reply. In such a situation you may have an unacceptably small number of categorized e-mails unless you add categorized outbound e-mails. However, note that selecting this check box adds all outbound categorized messages, and those outbound messages that are replies may have content that is not very relevant to the category of the original e-mail.

Copying a Training Object

This topic describes part of the functionality of Genesys Content Analyzer.

There are two ways to copy a training object: as an exact copy, or with selected text filtered out.

Procedure: Making an exact copy of a training object

To make an exact copy of a training object, first export the training object, then import it.

1. In the Export dialog box:

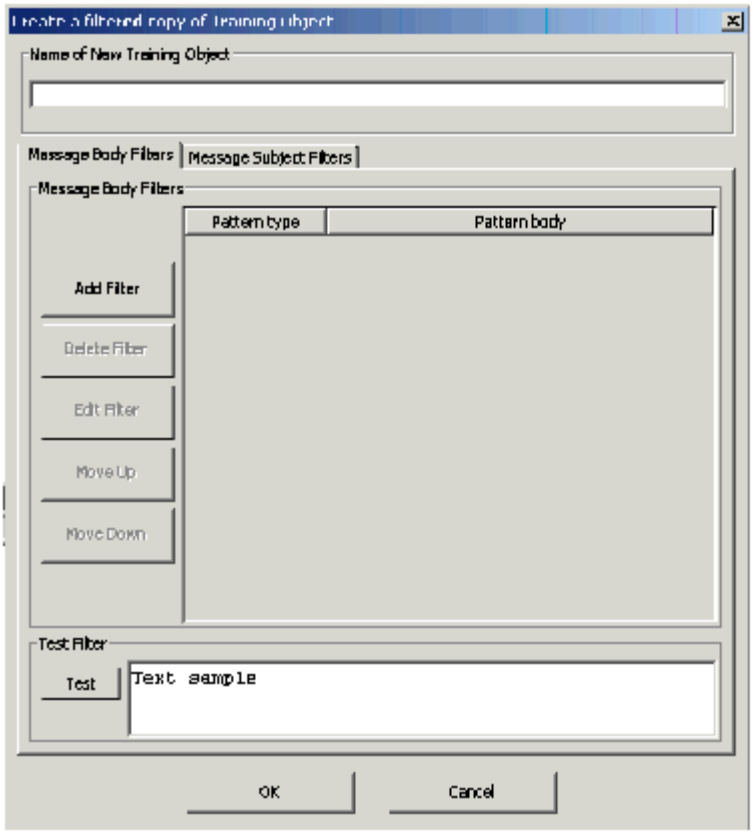
- Select the category tree.
- Select the training object.

- When importing, select new database IDs (see Importing Knowledge Manager Objects). Knowledge Manager asks you to modify the name of the category tree being imported.

Procedure: Making a filtered copy of a training object

To make a copy which filters out text that you specify, proceed as follows.

1. On the left-hand pane of the `Training` tab, select the root `Training Objects` node.
2. On the right-hand pane, highlight and right-click the training object that you want to make a filtered copy of.
3. Select `Create filtered copy of Training Object` from the resulting drop-down list.



Create a Filtered Copy of Training Object

4. On the resulting `Create filtered copy of Training Object` dialog box, shown in "Create a Filtered Copy of Training Object," proceed as follows:
   a. Enter a name for the new filtered copy of the training object.
   b. Create filters, and test them if you wish.

      This works identically with the `Text Preprocessing` tab of the `Model Training Schedule` dialog box, except that here you create separate filters for the body and subject of messages, using the `Message Body Filters` and `Message Subject Filters` tabs.

   c. Click OK to save the new filtered copy of the training object.

Adding More Text Objects to the Training Object

This topic describes part of the functionality of Genesys Content Analyzer.

Regardless of whether you have previously-categorized e-mails that you included in your training object, you can add more text objects to it in the following ways:

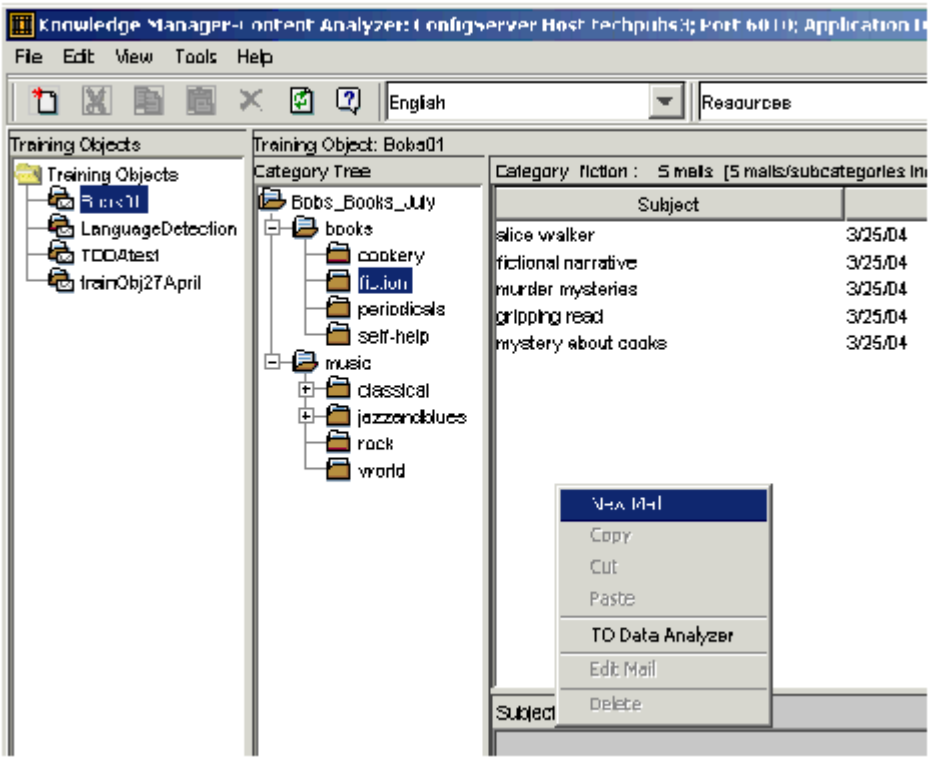
- Create new e-mails manually.
- Add standard responses or other training objects.
- Add uncategorized e-mails using the TO Data Analyzer.

Creating new e-mails manually

This topic describes part of the functionality of Genesys Content Analyzer.

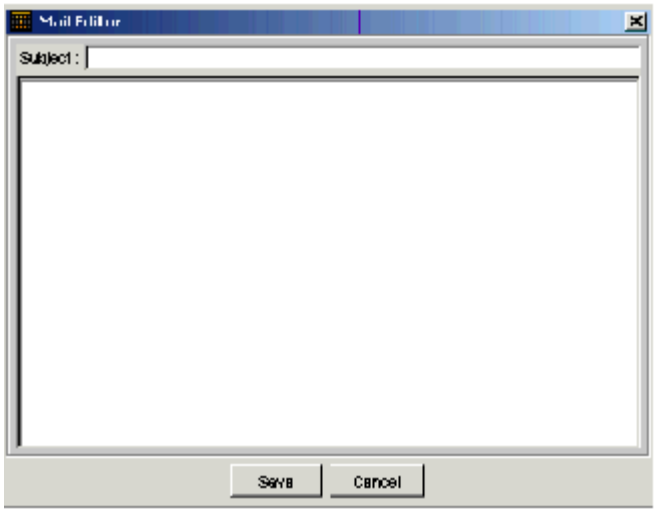
Purpose: To increase the accuracy of a model produced by a training object by adding new manually-created e-mails to the training object.

1. Close the `New Training Object` dialog box. On the left-hand pane of the `Training` tab, select the training object. This produces a three-pane view, shown in "Add E-mail to a Training Object."



Add E-mail to a Training Object

2. On the center pane, select the category that you want to add an e-mail to. The figure "Add E-mail to a Training Object" shows a user about to add a new e-mail to the `fiction` category.
3. On the right-hand pane, right-click and select `New Mail` from the shortcut menu. The `Mail Editor` opens, as shown in "Mail Editor."



Mail Editor

4. Enter text into the `Subject` and Body fields, then click Save.

   E-mails that you add in this way are stored as being handled by the agent called `default`.

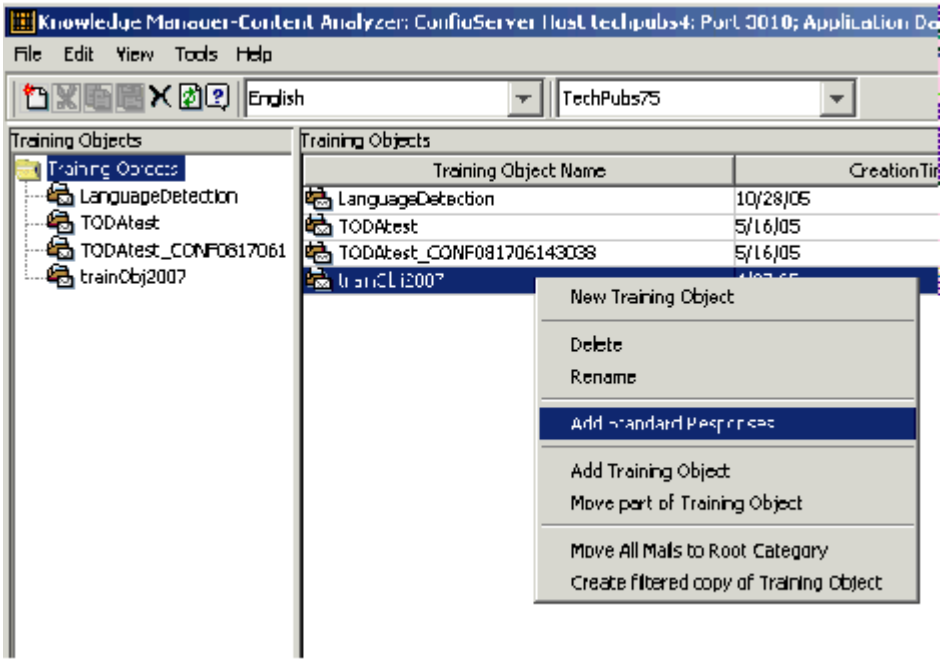
> **Important**
>
> You can also edit existing e-mails by selecting an e-mail in the right-hand pane, right-clicking, and selecting Edit Mail.

Adding standard responses or other training objects

This topic describes part of the functionality of Genesys Content Analyzer.

Purpose: To increase the accuracy of a model produced by a training object by adding either standard responses or other training objects to the training object.

1. To add standard responses, go to the left-hand pane of the `Training` tab and select the root `Training` Objects node. This produces a two-pane view, as shown in "Add Standard Responses to a Training Object."



Add Standard Responses to a Training Object

2. Right-click a training object on the right-hand pane, then choose `Add Standard Responses` from the shortcut menu. The figure "Add Standard Responses to a Training Object" shows a user about to add standard responses to the `trainObj2007` training object.
3. To add a training object, repeat Steps 1 and 2, but choose `Add Training Object` from the shortcut menu.

> **Important**
>
> You can add a training object to another training object only if both training objects use the same category tree.

Adding Uncategorized E-Mails Using TO Data Analyzer

This topic describes part of the functionality of Genesys Content Analyzer.

If you have a sizeable database of uncategorized e-mails, the TO Data Analyzer can help you to:

- Group the uncategorized e-mails and build a category tree for them.
- Assign the uncategorized e-mails to existing categories.

In the most basic terms, the TO Data Analyzer does the following:

1. It shows you all uncategorized e-mails one at a time, in an order determined by criteria that you set (if you set no criteria, it shows them in the order in which they were created).
2. As each e-mail displays, you choose whether to include it in a holding area.
3. You then assign the e-mails in the holding area to a category. This can be an existing category or a new one that you create.

The TO Data Analyzer contains the following tabs:

- Main—Displays uncategorized e-mails one at a time for you to categorize
- Search Criteria—Sets the criterion determining the order that uncategorized e-mails are displayed in
- Indexing—Displays information on co-occurrence of words in uncategorized e-mails

An example of using TO Data Analyzer is available.

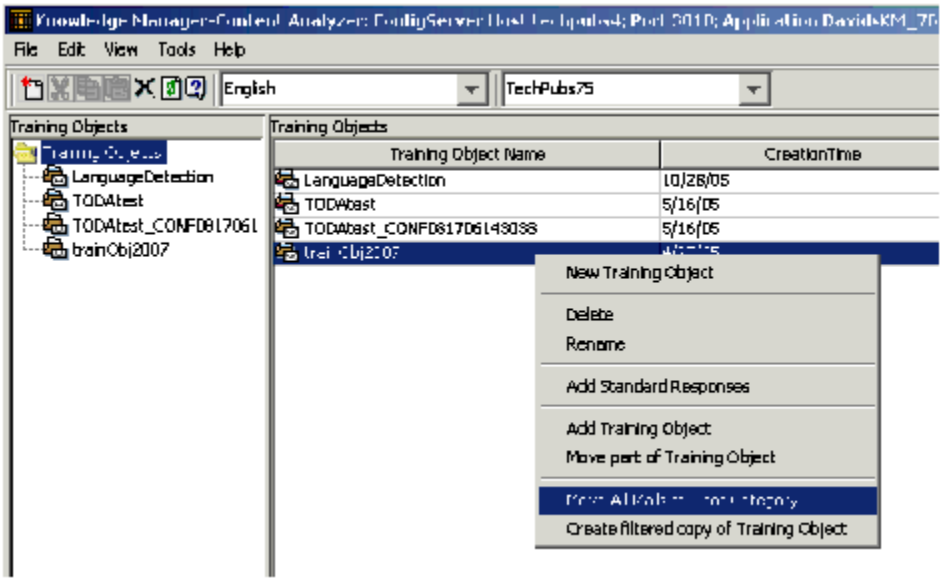
Procedure: Adding e-mails to a category on the Main tab

This procedure uses the TO Data Analyzer to search for uncategorized e-mails that are a likely fit for a selected category.

Prerequisites

For uncategorized e-mails to be available to the TO Data Analyzer, they must be associated with the root category of the category tree in the training object. You perform this association in either of the following ways:

- When you create the training object, select the `Add uncategorized messages for manual clustering` check box as described in Step 6 of Procedure: Creating a new training object.
- In an existing training object, you can move all e-mails to the root category (before you do this it is advisable to make a backup copy of the training object):
  a. On the left-hand pane, select the root `Training Objects` node to give the two-pane view.
  b. On the right-hand pane, right-click a training object and select `Move All Mails to Root Category`, as shown in "Move All E-Mails to the Root Category."

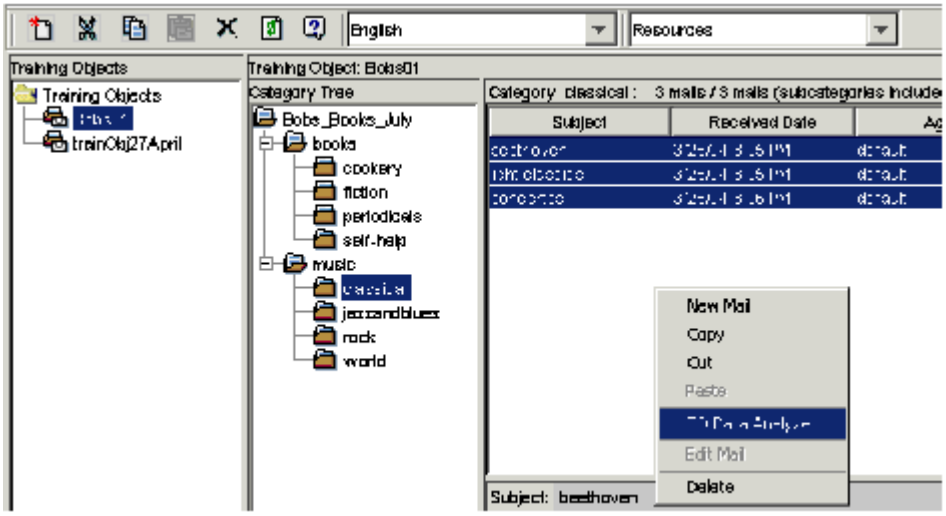


Move All E-Mails to the Root Category

Procedure

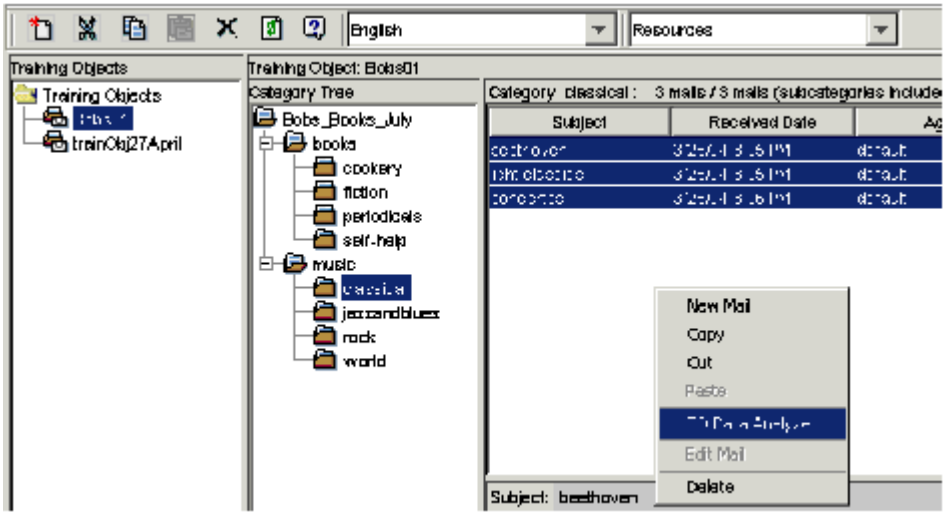
1. On the `Training` tab, select a training object.

2. On the middle pane, select a category that you want to add e-mails to.
3. On the rightmost pane, select the e-mails already assigned to this category.
4. Still on the rightmost pane, right-click and select `TO Data Analyzer`, as shown in "Start the Training Object Data Analyzer."



Start the Training Object Data Analyzer

5. The `Uncategorized Message Clustering` dialog box appears with its `Main` tab showing, as shown in Uncategorized Message Clustering."



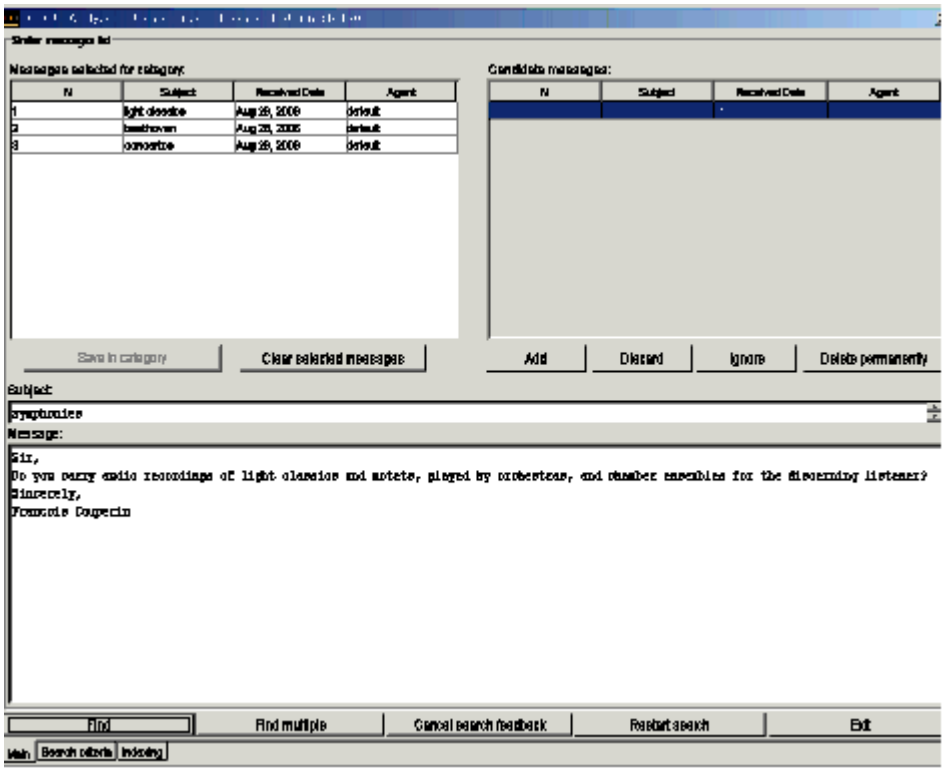
Uncategorized Message Clustering

Note that the e-mails that you selected in Step 3 appear in the `Similar messages list.` These e-mails are now part of the criteria that TO Data Analyzer uses to choose e-mails to display.

6. Click `Find.` The system displays a candidate e-mail as shown in "Message Clustering: First Candidate":
   - The `Candidate messages` box shows information about the e-mail. The `N` is number, relevant only

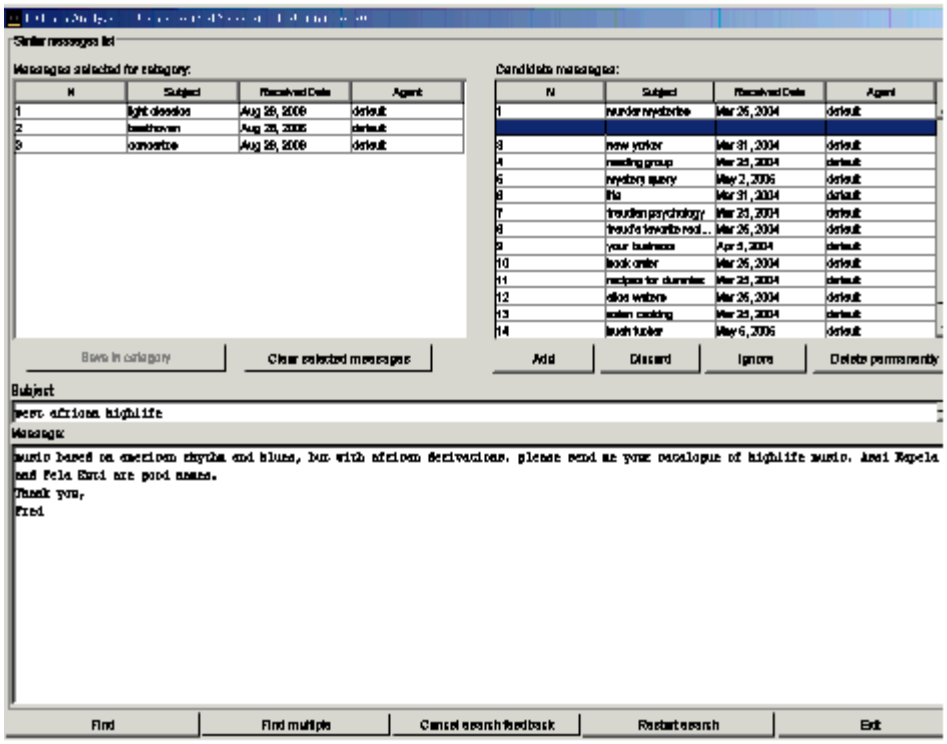
for the Find multiple command, described below.

- The Subject box shows the subject of the e-mail.
- The Message box shows the e-mail body.



Message Clustering: First Candidate

As an alternative to Find, click Find multiple. The system displays the 15 e-mails that are the best candidates for the category, as shown in "Message Clustering: Multiple Candidates."



Message Clustering: Multiple Candidates

7. Use the buttons at the bottom of the dialog to select one of the following actions for this candidate (or candidates, if you clicked `Find multiple` and then used SHIFT-click or CTL-click to select multiple candidates):
   - Add—Add to the selected category and adjust the search accordingly (see Search Criteria Tab for more explanation).
   - `Discard`—Omit from the selected category and adjust the search accordingly.
   - `Ignore`—Omit from the selected category but do not adjust the search.
   - `Delete permanently`—Remove from the training object permanently. This e-mail will not appear again.

> **Important**
>
> Delete permanently is useful for deleting garbage e-mails, that is,e-mails that you do not want to use anywhere in the training object. With Discard and Ignore, in contrast, the e-mail is no longer considered in the current search for the selected category, but it remains in the training object.

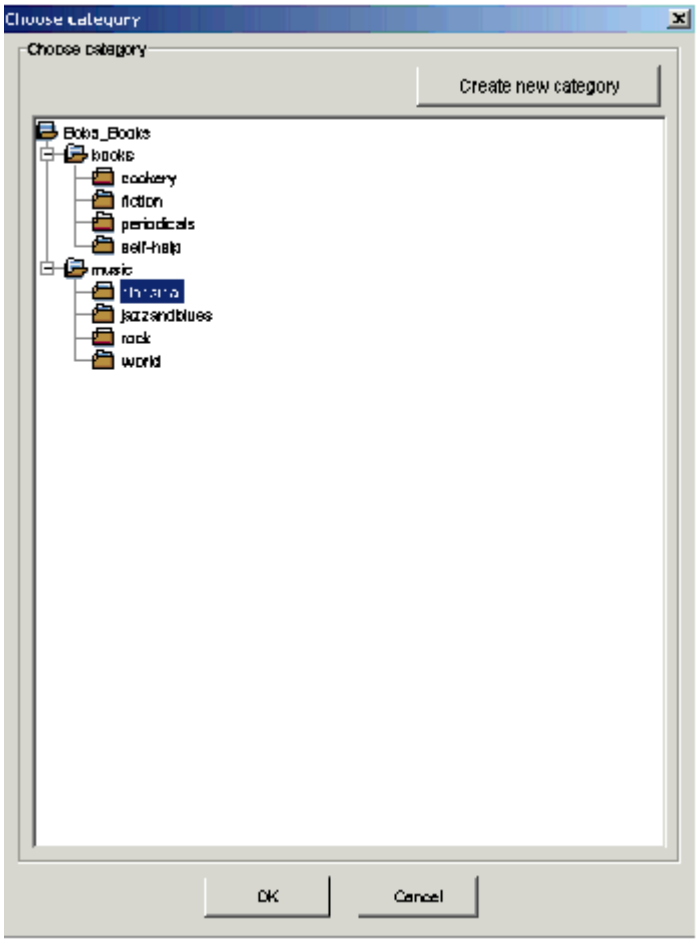
8. Click `Find` again, and the next candidate e-mail displays in the `Subject` and `Message` boxes. Decide whether to add, discard, or ignore this e-mail.
9. Continue in the same way, repeating Step 8. Knowledge Manager continues to display e-mails in the order determined by the criteria that you have set. When there are no more uncategorized e-mails, {

`Mail not found }` displays in the `Message` box.

10. When you are done adding e-mails to the `Similar messages list`, click one of the following:

    `Save in category`—This adds the e-mails to a category, as explained in the next step.
    `Clear selected messages`—This clears all e-mails from the `Similar messages list`. You can start again by clicking `Find`.

11. If you click `Save in category`, the `Choose category` dialog box appears, as shown in "Choose Category Dialog Box."



Choose Category Dialog Box

The `Choose category` dialog box displays the category tree that you are working with. If you selected a category in Step 2 above, that category is selected in this dialog box, but you can select another category (but not the root category). You can also click `Create new category` to add a new category which becomes a subcategory of the category selected in the `Choose category` dialog box. Click OK to add the e-mails in the `Similar`

messages list to the selected category.

Next Steps

- Optionally, use the Search Criteria tab to specify the order that Knowledge Manager shows you the uncategorized e-mails (this page).
- Study the example of using the TO Data Analyzer.

Search Criteria Tab

This topic describes part of the functionality of Genesys Content Analyzer.

When you click Find, Knowledge Manager shows you all uncategorized e-mails one at a time, in an order determined by criteria that you set using TO Data Analyzer's Search Criteria. If Knowledge Manager finds that an e-mail is a good fit for the criteria, this section says that it "gives priority" to that e-mail.

> **Important**
> Knowledge Manager gives priority to e-mails that meet the criteria best. Those that are a bad fit for the criteria are not excluded; they are simply put later in the list.

If you set no criteria, Knowledge Manager displays the e-mails in the order in which they were created.

Similar Messages List

Knowledge Manager gives priority to e-mails that resemble the e-mails in the Similar messages list. E-mails move to the Similar messages list in two ways:

- You select them before right-clicking TO Data Analyzer, as in the procedure described in the previous section.
- You click Add when they display in the Uncategorized message clustering dialog box.

The previous section instructs you to select all of the e-mails in the category that you are interested in. Of course, if you think that some e-mails in the category are not a good fit, you should not bring them to the Similar messages list.

Text Length

Use the Min and Max boxes to set limits on the size (number of characters) of e-mails. One use of this is to exclude very long e-mails, which would take you too long to read in the Message box.

Include and Exclude Text

Enter text in these boxes to adjust the way that Knowledge Manager assigns priority. E-mails that include text that matches the Include text box receive higher priority. E-mails that include text that matches the Exclude text box receive lower priority. These boxes are especially useful when you are starting out with nothing in the Similar messages list.

What you enter in these boxes is literal text, not regular expressions.

> **Important**
>
> E-mails that include text that matches the Exclude text box are not, in fact, excluded. They are simply moved towards the end of the list.

Refining or Resetting the Search

Each time you click Find, Knowledge Manager presents the e-mail that best fits the criteria that you have set. It continues to do this until it has presented all of the uncategorized e-mails. Knowledge Manager keeps track of which e-mails it has presented. If you alter the criteria, it then presents the best fit "from among those it has not yet presented." If you alter the criteria and want Knowledge Manager to start scanning the e-mails from the beginning again, you must click Restart Search before clicking Find.

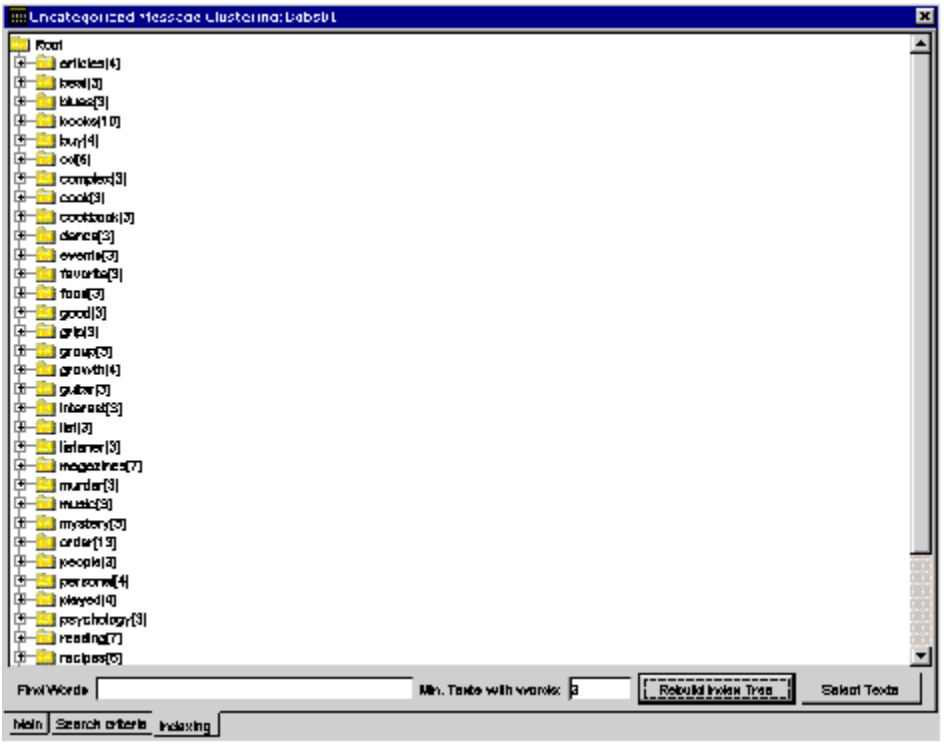
It is important to understand that you alter the search criteria each time you click Add or Discard (this is in addition to the possibility of you changing the contents of the Include text, Exclude text, and Text length boxes). When you click Add or Discard, you confirm or reject Knowledge Manager's guess as to which e-mail best fits the criteria, and Knowledge Manager uses your confirmation or rejection to adjust the criteria.

If, after going through a number of candidate e-mails, you decide that you are on the wrong track, you can click Restart Search, and Knowledge Manager restarts its search from the beginning, using only the criteria supplied by the contents of the Uncategorized message clustering dialog box (similar messages, text length, include/exclude text), and discarding all of your preceding Add/Discard input. However, any e-mails that you have added to the Similar messages list remain there after you click Restart Search.

Indexing Tab

This topic describes part of the functionality of Genesys Content Analyzer.

This tab, shown in Indexing Tab, displays information on cooccurrence patterns of words in uncategorized e-mails.



Indexing Tab

The tab displays, in tree form, a list of the words that occur in all uncategorized e-mails (except Stop Words).

The index tree consists of folder icons, each labeled with a word, with the number of occurrences (number of e-mails it occurs in) in square brackets. These words can be called *head words.*

Each head word folder expands to a list of the words (also folders) that cooccur with the head word—that is, that occur together with the head word in one or more e-mails. Each cooccurring word is followed by square brackets containing two numbers: the number of e-mails this word occurs in, and a ratio. This ratio is the rate of occurrence with this head word divided by rate of occurrence in whole corpus. Indexing Tab Example provides an example.

Indexing Tab Example

Among the information displayed in this example is the following:

- *magazines* occurs in seven uncategorized e-mails.
- *articles* occurs in three of those seven e-mails, which is 4.4 times as often as it occurs in the entire corpus of uncategorized e-mails.
- Of the three e-mails containing *magazines* and *articles,* two also contain *newsstand.* This is 13.7 times as often as *newsstand* occurs in the entire corpus.

This indicates that the words *articles* and *newsstand* are highly likely to occur together, which means e-mails that contain both words are good candidates for grouping together in a category. If you select *newsstand*, then click `Select texts`, the display switches to the `Main` tab, showing that all e-mails that contain *magazines, articles,* and *newsstand* have been put in the `Candidate messages list.`

At the bottom of the tab are the following:

- Two boxes for filtering the words that are displayed:
  - `Find Words` —Restrict the words displayed. More on this below.
  - `Min. Texts with words` —The word must occur in at least this number of e-mails to be displayed in the list.
- Two buttons that initiate actions:
  - `Rebuild Index Tree` —Rebuild the tree to apply the filters that you set in the `Find Words` and `Min. Texts with words` boxes.
  - `Select Texts` —Select a word in the index tree, then click this button to put all e-mails containing this word in the `Candidate messages` list.

Use the `Find words` box to restrict the words displayed. Enter a single word to display only that word and the words that occur with it. Enter multiple words to specify which cooccurring words to start the list with. The figure "Find Words = “mystery” " shows the result of entering *mystery* in the `Find words` box, then clicking `Rebuild Index Tree.`

Find Words = “mystery”

The figure "Find Words = “mystery reading” " shows the result of entering "mystery reading" in the `Find words` box: the index tree shows only the head word *mystery* and the cooccurring word *reading.*

Find Words = “mystery reading”

Example of Using TO Data Analyzer

This topic describes part of the functionality of Genesys Content Analyzer.

This section presents an example of using the TO Data Analyzer to add a category and build subcategories for it.

Prerequisites

The example makes the following assumptions:

- Part of your business deals with DSL service.
- Your category tree does not represent this DSL service sector.
- You have a collection of uncategorized e-mails, some of which deal with DSL service.

Procedure

1. On the Categories tab, add a DSL service category.
2. Create a training object using this category tree.
3. On the Training tab, select the DSL service category and open the TO Data Analyzer.
4. In the Include Text box, enter DSL.
5. Click Find repeatedly, browsing through the uncategorized e-mails and looking for common themes.
6. As you do this, you find a number of e-mails inquiring about the status of a DSL service order. You decide they should go in a subcategory that you will call “DSL Shipping Status.”
7. Add one of these e-mails to the Similar messages list.
8. Add the word shipping to the Include Text box to refine the criteria.
9. Click Restart Search, then Find. This starts the search from the beginning with the revised criteria.
10. Continue, clicking Find, then clicking Add for e-mails that deal with DSL shipping status and Discard for others.
11. When you have enough e-mails on the Similar messages list (between seven and 30), click Save in Category.
12. In the Choose category dialog box, select the DSL service category, then click Create new category.
13. In the New category dialog box, enter DSL Shipping Status in the Category name box, then click OK.
14. Back in the Choose category dialog box, click OK to save the new category and its associated e-mails.
15. Click Restart Search and clear the Include Text box of everything except DSL.
16. Start again from Step 5, looking for another common theme that you can use as a subcategory.

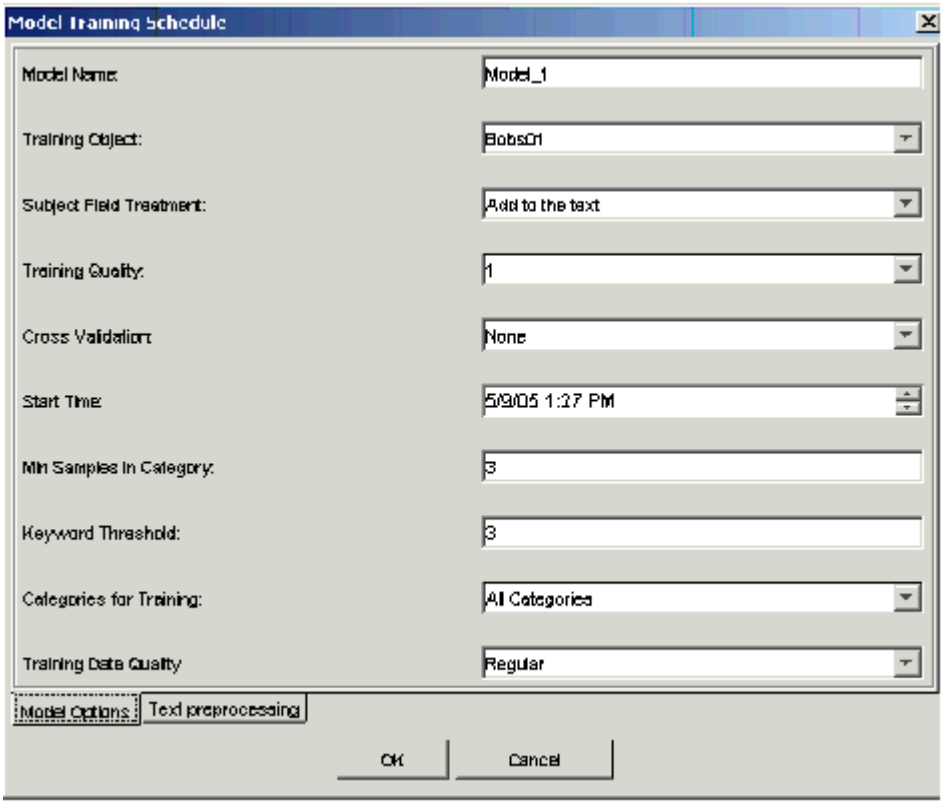
Schedule Training

This topic describes part of the functionality of Genesys Content Analyzer.

When you have a training object with enough e-mails, you are ready to schedule training.

There are the following two options:

- You can schedule training that uses an existing scheduled job as a template. This is a convenient way to change the time that an existing job is scheduled to run. To do this:
  a. On the `Training Schedule` tab, right-click the existing job that you want to use as a template.
  b. From the context menu, select `New Training Job (Use this Job as Template)`.
- You can create a new training job from scratch. To do this, on the `Tools` menu of the `Training` tab, select `Schedule New Model Training`.
  Both options bring up the `Model Training Schedule` dialog box. With option 1, the fields of the dialog box are populated with values copied from the existing job that was used as a template. With option 2, the fields are populated with default values, as shown in "Model Training Schedule: Model Options Tab." This dialog box, which has two tabs, opens on the `Model Options` tab.



Model Training Schedule: Model Options Tab

Procedure: Schedule training using the Model Options tab

This topic describes part of the functionality of Genesys Content Analyzer.

Purpose: To specify how and when a training object will be processed to produce a model.

Prerequisites

- A training object containing a sufficient number of e-mails or other text objects. When to Train provides suggestions about judging whether there are enough text objects.

1. `Model Name` —Enter a name for the model that will result from the scheduled training. Creating a Category Tree explains restrictions on the names of Knowledge Manager objects.
2. `Training Object` —Select a training object.
3. `Subject Field Treatment` —Select from the following treatments of the `Subject` field of e-mails:

- `Ignore` —Training does not consider the content of the `Subject` field
- `Add to the text` —Training considers the content of the `Subject` field.
- `Add with double weight` —Training gives the content of the `Subject` field twice as much importance as the content of the e-mail body.

- `Training Quality` —Draft is the lowest quality, 6 is the highest. Note the following:
  - Training time increases as you move from Draft quality to level 3 quality. But once the quality goes above 3, there is not much difference in training time.
  - Genesys recommends that you use Draft quality only when you want to obtain a preliminary reading of the model's quality estimation. For production, use quality 2–6.
- `Cross Validation` —Select either no cross-validation, or cross-validation that splits the data into 3, 5, or 10 sets. Cross-Validation provides an explanation.

  If you select cross-validation, training produces an accuracy rating for the model along with the model itself. This has the advantage of not requiring an extra testing step, but it increases the training time.

- `Start Time` —Enter a start time or select a unit (day, month, hour, minute) and change its value using the up and down arrows. Because training can use a large proportion of system resources, you will probably want to schedule it for nonpeak hours.

> Important

> Be sure to set a time later than the present moment.

- `Min Samples in Category` —Enter the minimum number of text objects that a category must have in order to be included in training. Categories with no or few text objects make poor subjects for training.
- `Keyword Threshold` —Enter the minimum number of text objects that a keyword must occur in for that keyword to be considered in training.

  A relatively high value for this setting can reduce training time, but it can also reduce quality. What counts as a high or low value for this setting depends on the total size of the training object. For example, if a training object has 5 to 10 text objects per category, a high keyword threshold might be 2 or`3`. If a training object has 30 to 50 text objects per category, a high keyword threshold might be `20`.

- `Categories for Training` —Select `All Categories` or `Terminal Categories Only`. A "terminal category" is one that contains no subcategories. It may be that a category tree uses nonterminal categories only or mostly for organizing the terminal categories. In this case few or no text objects are associated with the nonterminal categories, and there is little to be gained by including the nonterminal categories in training.
- `Training Data Quality` —Select `Regular` unless you know that the training object contains many wrongly categorized text objects. If it does, select `Unreliable` to set the categorization algorithm to run in an altered mode that gives better results with this type of data.

Next Steps

- Optionally, remove superfluous or misleading text from the training object (next section).
- Once the model is trained, test it. See Testing Models.

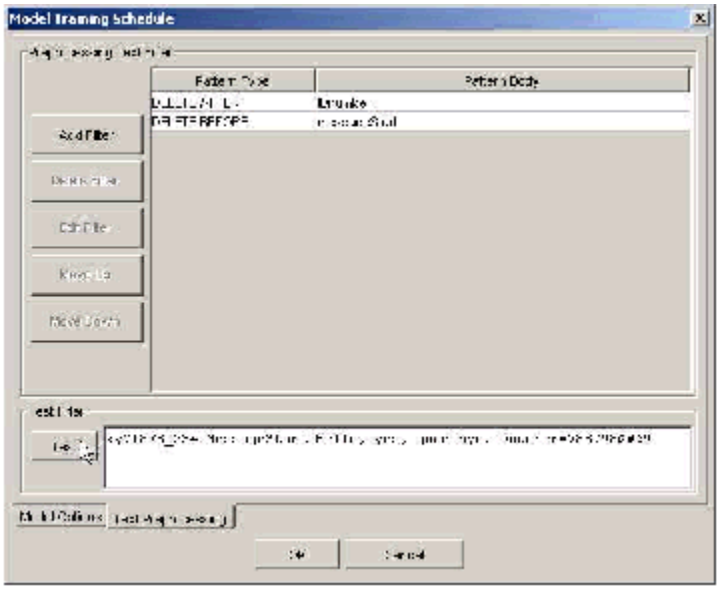
Creating and Testing Filters

This topic describes part of the functionality of Genesys Content Analyzer.

Text Preprocessing Tab

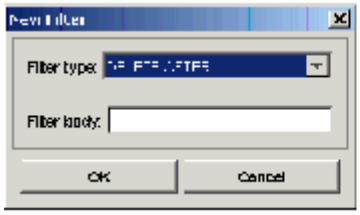
The `Text Preprocessing` tab, shown in "Model Training Schedule: Text Preprocessing Tab," enables you to remove extraneous text from the text objects of a training object. From this tab, you can create filters (patterns) that search for text and perform various deletion operations. This can be helpful when the e-mails that you want to use for training contain significant amounts of text that has both of these characteristics:

- It is predictable enough in content to be identifiable by a regular expression.
- It is irrelevant or misleading for classification purposes.



Model Training Schedule: Text Preprocessing Tab

1. Click `Add filter.` The `New Filter` dialog box appears, as shown in "New Filter Dialog Box."

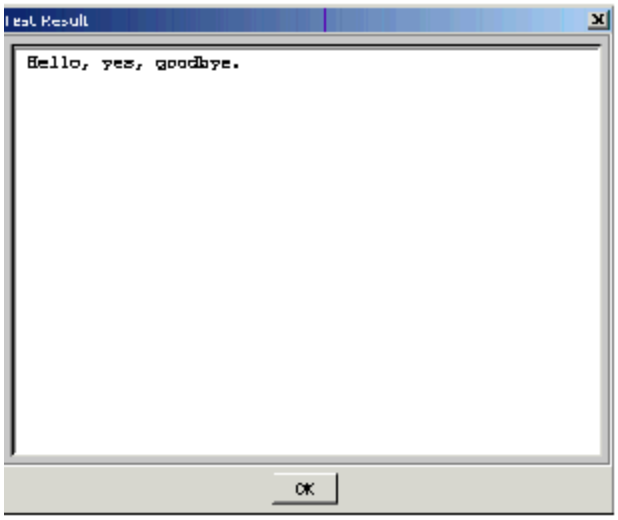


New Filter Dialog Box

2. Choose a type from the `Filter type` drop-down list. The filter type specifies the action to take; for example, delete all text up to and including the matched text. Filter Types below provides descriptions. Filter type is called `Pattern Type` on the main `Text Preprocessing` tab.
3. Enter text in the `Filter body` box. The filter body contains the text to match, as either a literal string or a regular expression. Filter body is called `Pattern Body` on the main `Text Preprocessing` tab.
4. Click 0K.

> The figure "Model Training Schedule: Text Preprocessing Tab" above shows an example using two filters. The first deletes the text "IDnumber=" and anything following it. The second deletes the text "messageStart" and anything preceding it.

- Continue by testing the filter: enter sample text in the window in the `Test Filter` area.
- Click `Test.` A new window displays the result of applying all filters, in order. The figure "FIlter Test Result" shows the result of the test on the text shown in "New Filter Dialog Box."



FIlter Test Result

Filter Types

The following is a list of the available filter types:

- `DELETE AFTER` —Search for a match to the pattern body, then delete all text after and including the matching text.
- `DELETE BEFORE` —Search for a match to the pattern body, then delete all text before and including the matching text.

- `DELETE ALL IF FIND` —Search for a match to the pattern body, then delete the entire e-mail that includes the matching text.
- `DELETE ALL IF NOT FIND` —Search for a match to the pattern body, then delete the entire e-mail if it does not include the matching text.
- `DELETE PATTERN` —Search for a match to the pattern body, then delete only the text that matches the pattern.

Text Preprocessing Examples

Examples

This topic describes part of the functionality of Genesys Content Analyzer.

The table "Examples of Preprocessing Filters" displays simple examples of text-preprocessing filters.

Examples of Preprocessing Filters

Pattern Type	Pattern Body	Input Text	Test Result
DELETE AFTER	finch	one two finch three four	one two
DELETE BEFORE	finch	one two finch three four	three four
	[Mm]essage_?[Ss]tart	x897 message_Start one two three	one two three
DELETE ALL IF FIND	finch	one two finch three four	footnote a.
		one two three four	one two three four
	internal\d\d	one two three internal36 four	
DELETE ALL IF NOT FIND	finch	one two finch three four	one two finch three four
	finch	one two three four	
DELETE PATTERN	f.*ch\s	one two finch three four	one two three four
		one two fach three four	one two three four

a. If you test this filter, the resulting window contains the message `TEXT HAS BEEN DELETED.` In actual use of `DELETE ALL IF FIND` or `DELETE ALL IF NOT FIND`, the entire text object is deleted from the training object.

Tables "Preprocessing Filters Example" and "Results of Testing the Example" present a more complex example using all five filter types.

The table, "Preprocessing Filters Example" lists the filters used in the example.

Preprocessing Filters Example

Filter Number	Pattern Type	Pattern Body
1	DELETE BEFORE	MessageStart
2	DELETE AFTER	IDnumber=
3	DELETE ALL IF FIND	internal\d\d
4	DELETE ALL IF NOT FIND	nihil_obstat
5	DELETE PATTERN	company

The table "Results of Testing the Example" shows an example of input text and the results of applying the filters from "Preprocessing Filters Example" to it.

Results of Testing the Example

Input Text	Test Result
x88_2 MessageStart nihil_obstat: Hello, companyyes, good-bye.IDnumber=7989	nihil_obstat: Hello, yes, good-bye.

The results in the table "Results of Testing the Example" come about as follows:

1. Filter 1 deletes the text "x88_2 MessageStart."
2. Filter 2 deletes the text "IDnumber=7989."
3. Filter 3 does nothing (it finds a match for nihil_obstat).
4. Filter 4 does nothing (it fails to find a match for internal/d/d).
5. Filter 5 deletes "company."

Notes on Language

This topic describes part of the functionality of Genesys Content Analyzer.

Selecting a Language

The first step in creating a training object is to select a tenant and language. The language that you choose has special relevance in the following two cases:

- If you want to build a model that classifies according to language, you must use a tree whose language is specified as `unknown`. First you must add this language attribute in `Configuration Manager > Business Attributes > Languages`.
- Selecting English activates a lexical analyzer that is specific to English. If you are operating in a language other than English, you should not select English because the English lexical analyzer will hinder the training.

Selecting any language other than English activates a default lexical analyzer. You can also create a lexical analyzer that is specific to any language you choose.

Lexical Analyzer

The function of a lexical analyzer is to convert input text (such as the text of an e-mail) to an array, either of words or of stems.

Content Analyzer includes the following:

- Language-specific lexical analyzers for English and Japanese. The latter is available with the Content Analyzer - Japanese edition.
- A default lexical analyzer that is much simpler than the English analyzer.
- Sample code that you can use to create your own lexical analyzer for a language of your choice.

Language-Specific Analyzers

The language-specific analyzers convert words to *stems,* delete digits and special characters, and segment the result into an array, omitting *stop words.* To explain two key terms:

- A *stem* is a basic string that is shared by a family of words; for example, *see, seen,* and *seeing* all have the same stem, as do `allow`, `allows`, and `allowance`.
- *Stop words* are words that are so common that there is little to be gained in searching for them or listing their occurrences. Examples for English are *the, a, an, of, to, is,* and so on. The system does not consider stop words when performing classification, and stop words do not appear on the Indexing Tab of the TO Data Analyzer. The installation packages for Genesys Content Analyzer install a file of stop

words for English, called `English.stop`, in the home directories of Knowledge Manager and Training Server. This is a simple text file containing a list of words separated by carriage return. You can create other files of stop words for other languages.

Default Analyzer

The default analyzer is less robust than the language specific ones: it does not distinguish stems, and it does not ignore stop words. It considers any sequences of alphabetic characters (a–z, A–Z) to be words, and all other characters to be word separators. For this reason, this analyzer requires significantly more training to work satisfactorily with languages such as Arabic, Chinese, or Korean, which do not indicate separations between words.

> Here is an example contrasting the English lexical analyzer with the default lexical analyzer: Take the phrases *I find your service disappointing* and *I find your service to be a disappointment*. The English lexical analyzer is smart enough to know that *disappointing* and *disappointment* share the stem *disappoint*. It also ignores the word *a*. So if it is trained to classify the first phrase as negative, it does not have to conduct a separate analysis of the second phrase to know that it is also negative. The default lexical analyzer is not as smart at picking out similarities in words and phrases, so it would conduct separate classifications of *disappointing* and *disappointment*. It would also waste time considering that *a* might be a marker of negative sentiment. That is why the amount of training needed to obtain the same level of precision is greater with the default lexical analyzer than with the English one.

Analyzer Code Sample

This sample is a class that implements the LexicalAnalyzer and Serializable interfaces. The LexicalAnalyzer interface includes two methods:

```
public interface LexicalAnalyzer {
  public String getLanguage();
  public String[] convert(String text);
}
```

- `public String getLanguage()` returns the name of the language that this lexical analyzer applies to.
- `public String[] convert(String text)` converts text to words or stems.

You can add one or more lexical analyzers for languages of your choice. To do this, you must prepare a Java class that implements the LexicalAnalyzer interface with the two methods just described.

The lexical analyzer example is located in `<KnowledgeManagerHome>\LexicalAnalyzerExample`. `<KnowledgeManagerHome>` is normally something like `C:\Program Files\GCTI\eServices 8.1.4\Knowledge Manager`. The source code is in `LexAnalyzerTest.java`. To adapt it to a language of your choice, use the following procedure.

1. Select a name for the language that your analyzer will apply to.
2. Adapt the `LexAnalyzerTest` class to the target language, changing the name of the class and

substituting the language name that you selected for the name used in the example (`English09`). For the purposes of this description, suppose you rename the class `MyLexAnalyzer`.

3. Compile the `MyLexAnalyzer` class using the following command:

   ```
   javac -classpath "gcengine.jar" MyLexAnalyzer.java
   ```

   The `gcengine.jar` file is located in the `LexicalAnalyzerExample` directory.

4. Copy the resulting `MyLexAnalyzer.class` file to the home directories of Knowledge Manager, Training Server, and Classification Server.

> **Important**
>
> The stop word file must be in the UTF-8 format (prior to release 7.6, stop word files required the ANSI format).

Content Analyzer Japanese

Genesys Content Analyzer – Japanese is a lexical analyzer for Japanese, available as an extra option. To use it, contact your Genesys representative to purchase a license, then proceed as follows:

- Locate the `license.dat` file and copy it to `<KnowledgeManagerHome>\LexicalAnalyzerGLA\lang`. Overwrite the dummy `license.dat` file that is already there.
- Add a language called `Japanese_GLA` to `Configuration Manager > Business Attributes > Languages`.

Large Training Objects

This topic describes part of the functionality of Genesys Content Analyzer.

If your training object is very large (over 50,000 e-mails), training may consume considerable memory and time. To reduce this consumption without impacting quality, follow these recommendations when you Scheduling training:

- Set `Cross Validation` to `None`.
- Set `Keyword Threshold` above 25.
- Set `Min Samples in Category` above 25.
- Set `Training Quality` below 4. A level of 3 or 4 is adequate for production use.

You should also allocate memory as follows:

- Ensure that the host machine of Training Server has at least 4 GB of RAM for Solaris, or 2 GB of RAM for Windows.
- In the .sh or ProcessParameters.ini file, change the parameter `-Xmx800m` as follows:
  - On Windows, change to `-Xmx1400m`. This is enough for a training object of about 40,000 e-mails, the maximum recommended size on this platform.
  - On Solaris, change to `-Xmx3000m`. This is enough for a training object of about 100,000 e-mails, the maximum recommended size on this platform.

> For large training objects, these recommendations supersede those in the "Knowledge Manager" section on Administration General Recommendations.

A successful test has been done with the following parameters:

> Host: Solaris, Enterprise 450 Model 4300 with 4000 MB RAM
>
> Training object: 100,000 e-mails in 1,000 categories
>
> Cross Validation: None
>
> Keyword Threshold: 25
>
> Min Samples in Category: 25
>
> Training Quality: 3

The expected computational time is between 12 and 18 hours.

Note that the model produced has no quality ratings because you set `Cross Validation` to `None`. Genesys strongly recommends against using cross-validation on such large training objects. To obtain quality ratings for the model, build an additional small training object and test the model model on a training object.

Testing Models

This topic describes part of the functionality of Genesys Content Analyzer.

It includes:

- Cross-Validation
- Testing a Model on a Training Object
- Testing a Model on Composed Text
- Testing a Model on Uncategorized Messages

There are four methods of testing a model. The following two methods test the model's accuracy and produce ratings of it, as described in Reading and Understanding the Ratings:

- If you select cross-validation when you schedule training, Training Server produces accuracy ratings along with the model.
- You can select Testing a Model on a Training Object.

The following two methods show which category the model assigns to selected text objects, but does not test the accuracy of that categorization:

- You can test the model on text that you compose: Testing a Model on Composed Text.
- You can apply the model to the uncategorized messages of a training object.

Cross-Validation

This topic describes part of the functionality of Genesys Content Analyzer.

In cross-validation, Training Server follows these steps:

1. It builds one model using all of the data.

2. It divides the data into *x* partitions, where *x* = 3, 5, or 10.

3. It builds a number of partial models: as many as there are partitions, each one using a different combination of *x* -1 partitions.

> For example, if the data is divided into the three partitions A, B, and C, Training Server builds model X using partitions A and B, model Y using partitions A and C, and model Z using partitions B and C.

4. It tests each of these partial models against the partition that it omitted when it was built.

> In the example, it tests model X against partition C, model Y against partition B, and model Z against partition A.

5. It aggregates the results of all these tests and presents them as the rating of the entire model.

These ideas underlie the concept of cross-validation:

- The best way to test a model is to apply it to data that was not used in building the model.
- A model built using most of the data is usefully similar to the model built using all of the data, so the results of testing (for example) all possible 90-percent models are a good indication of the quality of the 100-percent model.

Because cross-validation adds to the time required to build a model, you may not want to select cross-validation for very large training objects or for objects for which you selected training quality level 6.

Testing a Model on a Training Object

This topic describes part of the functionality of Genesys Content Analyzer.

It includes:

- Setting Up Training Objects for Testing
- Procedure: Creating Two New Training Objects
- Procedure: Extracting Random Text Objects
- Procedure: Testing a Model on a Training Object

You can test a model on a training object. This process applies the model to the texts in the training object and compares the resulting classification with the classification in the training object itself. The training object must use the same category tree as the model you are testing on it.

> **Important**
> It is not advisable to test a model on the training object that generated it: the results will be unrealistically favorable.

Setting Up Training Objects for Testing

You may want to create a new training object just for the purpose of testing. Use the same category tree but different text objects. There are two ways to do this:

- Starting with a collection of categorized e-mails, create two new training objects. See the figure "Creating two new training objects" on this page.
- Starting with an existing training object, create a second training object using randomly-chosen text objects from the first. In more detail:

a. Move five percent of the text objects, randomly selected, from one training object to another.
b. Train a model on the first training object.
c. Test the model on the second training object.

This is similar to using cross-validation with two partitions. Procedure: Extracting random text objects describes this process in detail.

Creating Two New Training Objects

This topic describes part of the functionality of Genesys Content Analyzer.

Prerequisites

- A large collection of categorized e-mails from a relatively long period of time.

1. Divide your collection into two parts, either according to date received or according to the agent that handled the interaction.

   The two parts may be mutually exclusive or not. For example, one part could be interactions from January through June of last year. and the other could be from July through December of last year. Or one part could be all interactions from last year, and the other part could be interactions from November and December of last year.

- Create two training objects, one using each part of the collection.
- Build a model on one training object.
- Test the model on the other training object, as described in Testing a model on a training object.

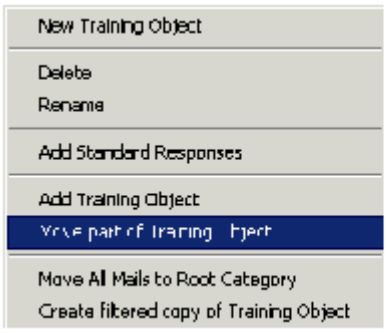
Procedure: Extracting Random Text Objects

This topic describes part of the functionality of Genesys Content Analyzer.

Prerequisites

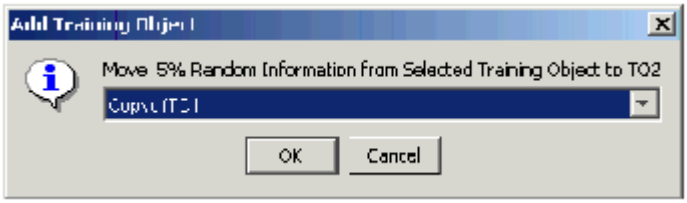
- This example assumes that you have a large training object called T01.

1. Make a copy of T01, calling it CopyofT01.
2. Create an new empty training object, called T02, using the same category tree as T01.
3. On the Training tab, left hand pane, select the root Training Objects node to give the two-pane view.
4. On the right-hand pane, right-click T02. Select Move part of Training Object, as shown in "Move Part of Training Object."



Move Part of Training Object

5. In the resulting Add Training Object dialog box, select CopyofT01, as shown in "Add Training Object Dialog Box."



Add Training Object Dialog Box

6. Knowledge Manager randomly selects five percent of the text objects in `CopyofT01`, copies them to `T02`, and deletes them from `CopyofT01`.
7. Train a model on `CopyofT01`, then test it on `T02`, as described in Procedure: Testing a Model on a Training Object.

Procedure: Testing a Model on the Training Object

This topic describes part of the functionality of Genesys Content Analyzer.

1. On the Training tab, select Tools > Schedule Model Testing.
   The Model Testing Schedule dialog box opens, as shown in "Model Testing Schedule."

Model Testing Schedule

2. Select a testing object—that is, select a training object to use.
3. Select a model to test.
4. Enter a start time.
5. Click OK.

   If the results are good and if your two training objects include some non-overlapping items, you can merge the two objects, by adding one to the other:

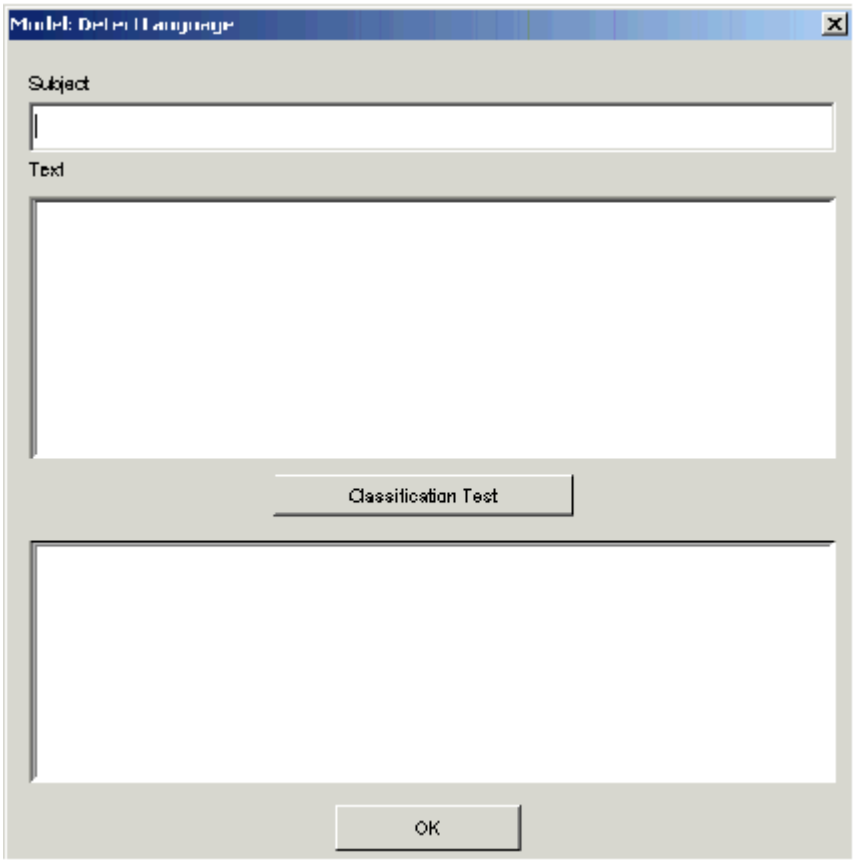
6. On the two-pane view of the Training tab (see the figure "Add Standard Responses to a Training Object)," select one training object on the right-hand pane.
7. Right-click and select Add Training Object from the shortcut menu.
8. In the dialog box that opens, select the other training object from the drop-down list, then click OK.

Procedure: Testing a Model on Composed Text

This topic describes part of the functionality of Genesys Content Analyzer.

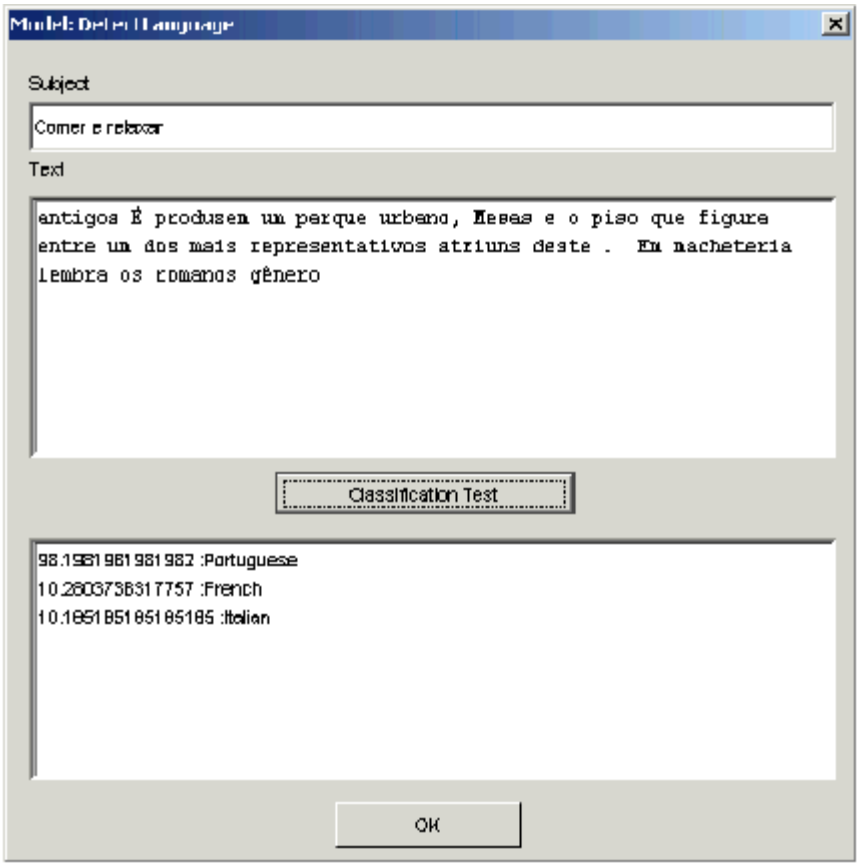
Purpose To test a model by seeing how it classifies a text object that you write for the purpose.

1. On the `Models` tab, go to the left-hand pane and select the root `Models` node.
2. On the right-hand pane, select the model that you want to test, then right-click and select `Test` from the shortcut menu. A dialog box appears, titled `Model: <modelname>`, as shown in "Model Test Dialog Box."



Model Test Dialog Box

3. Enter text in either or both of the `Subject` and `Text` boxes, then click `Classification Test`.
4. Results display in the lowest box. The figure "Model Test Results" shows the results of a test on the `DetectLanguage` model that is supplied with Genesys Content Analyzer.



Model Test Results

The results are in the form of a list of categories, each category preceded by the rating of the confidence with which the system assigns the test text to that category.

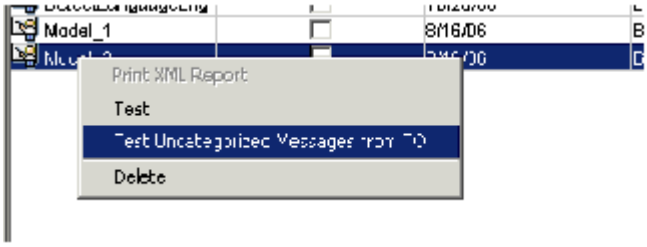
Testing a Model on Uncategorized Messages

This topic describes part of the functionality of Genesys Content Analyzer.

Prerequisites

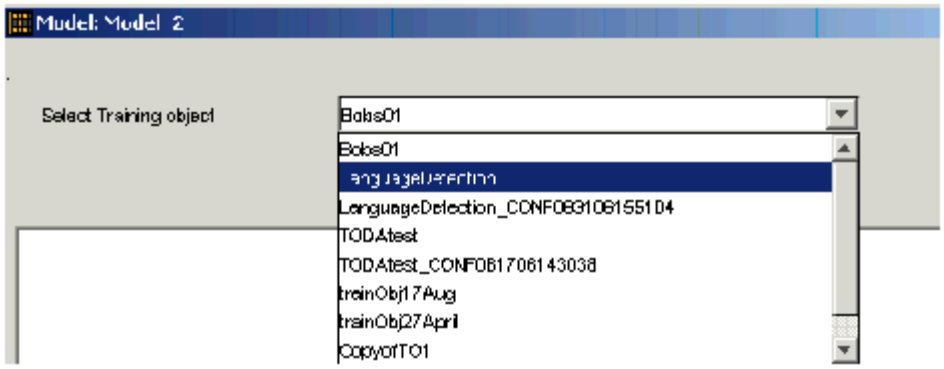
- There must be uncategorized messages in the training object's root category. You accomplish this by doing either of the following:
  - Assign uncategorized messages to the root category when creating the training object. See Step 6 of Procedure: Creating a New Training Object.
  - Move all text objects to the root category after creating the training object. See the Prerequisites for the Procedure: Adding e-mails to a category on the Main tab.

1. On the `Models` tab, go to the left-hand pane and select the root `Models` node.
2. On the right-hand pane, right-click the model that you want to test, then select `Test Uncategorized messages from TO`, as shown in "Test Uncategorized Messages."



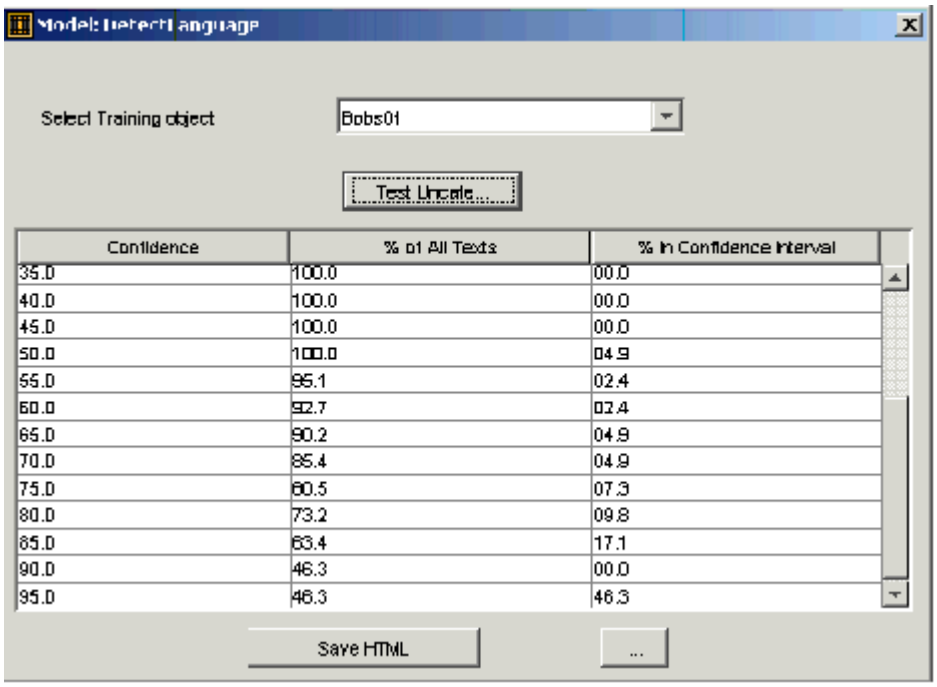
Test Uncategorized Messages

3. In the resulting window, select a training object from the drop-down list, as shown in "Select a Training Object." Be sure to select a training object that contains a good number of uncategorized messages.



Select a Training Object

4. Click `Test Uncategorized texts`. The results are displayed as in "Results of Test Uncategorized Messages."



Results of Test Uncategorized Messages

The results show the following for each confidence level:

- `% of All Texts:` The percentage of texts that were classified above this level of confidence.
- `% in Confidence Interval:` The percentage of texts that were classified with a level of confidence between this level and the next higher level on the scale.

These results tell you how well the model does, according to its own internal metric, at assigning new texts to some category or other. They do not evaluate the accuracy of these category assignments. To save the results as an HTML file, click Save HTML, provide a name for the file, then click Save.

Using and Rating Models

This topic describes part of the functionality of Genesys Content Analyzer.

It includes:

- Reading and Understanding the Ratings
- Average Results Subtab
- Category Confusion Subtab
- Results by Category Subtab
- Correct in Top N Subtab

You can have many models, but you can use only one at a time for classification. You designate the model to use in classification by setting it as `Active` on the `Models` tab.

Improving the Results

If the results of testing a model are unsatisfactory, there are several things you can do to try to produce an improved model.

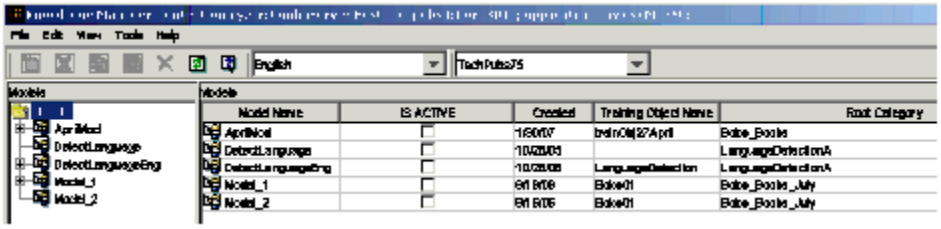
- Add more data to the training object
- Analyze your category tree. Are some categories never or seldom used? Are a few categories so general that they absorb most e-mails, leaving little for other categories? See Design and Use Considerations for more on this topic.

Reading and Understanding the Ratings

This topic describes part of the functionality of Genesys Content Analyzer.

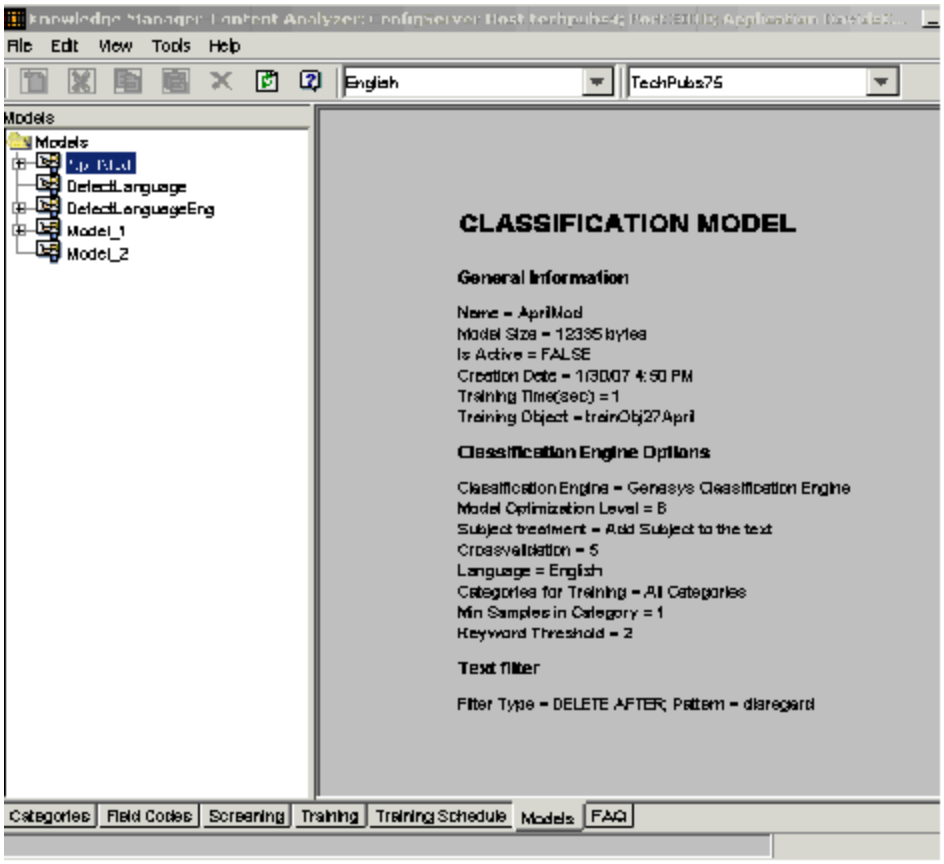
The Models tab displays a browserlike tree structure on its left-hand pane. You can use the structure as follows:

- Select the root Models node to display a list of summary information about all models, as shown in "Models Tab: Root Node Selected." The IS ACTIVE column consists of check boxes; select one check box to select the model that will be active in classification.



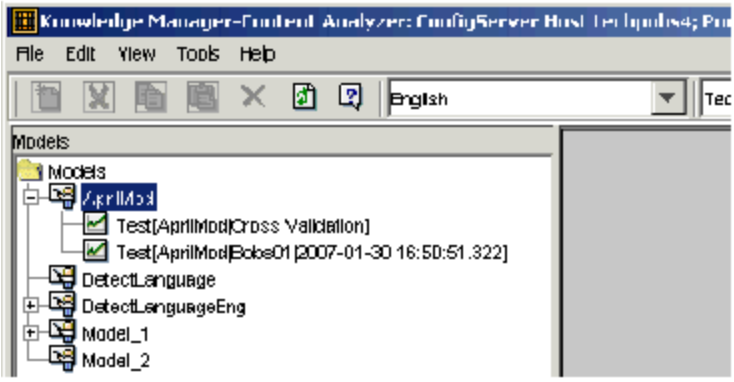
Models Tab: Root Node Selected

- Select a model node to display detailed information about the model, as shown in "Models Tab: Model Node Selected."



Models Tab: Model Node Selected

- If a model has been tested (Testing Models), its node can expand to display ratings nodes, as shown in "Models Tab: Ratings Nodes Displayed."



Models Tab: Ratings Nodes Displayed

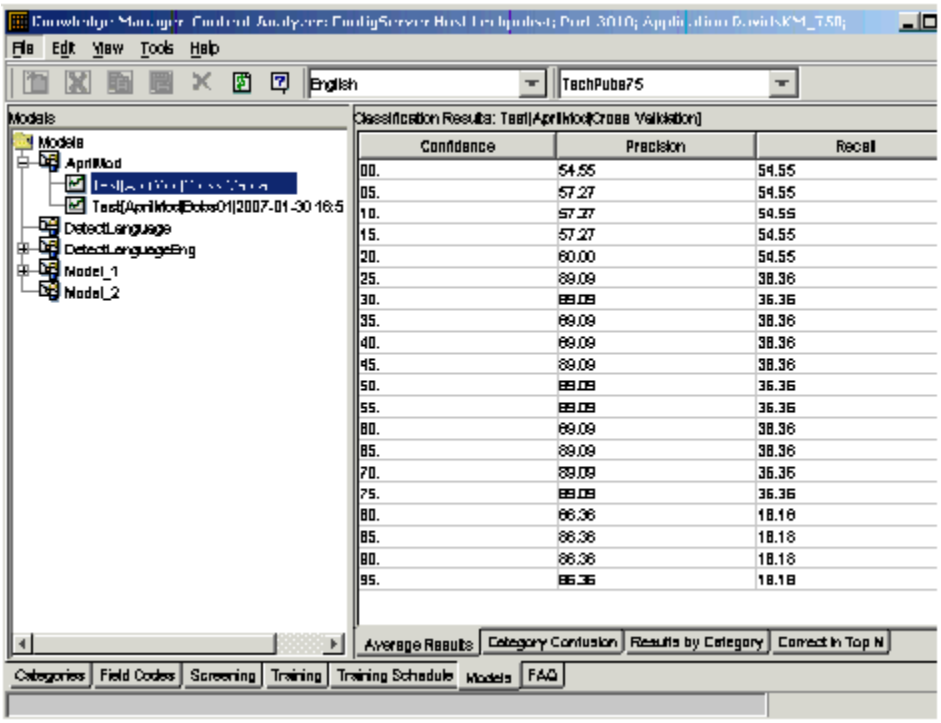
The ratings node label has the form `Model <modelname>[<ratingsource>]`, where `<ratingsource>` is either `'CrossValidation` or `on<testObjectName> Testing Object/ Time=<date time>`. For example, in Models Tab: Ratings Nodes Displayed, the Model `New04` has two sets of ratings. One is from cross-validation applied during generation of the model, and the other is from testing the model on the training object `Bobs01`.

- Select a model's ratings node to display its ratings on four subtabs, as described in the sections Average Results, Category Confusion, Results by Category, and Correct in Top N.

Average Results Subtab

This topic describes part of the functionality of Genesys Content Analyzer.

The Average Results subtab, shown in the figure "Average Results Subtab," rates how well the model classifies, averaged across all categories.



Models Tab: Average Results Subtab

The `Confidence` rating corresponds to the attribute of the same name that you set in the IRD objects `Classify` and `Classification switch`.

To understand `Precision` and `Recall`, consider several possible ways of looking at the performance of a model. If your model attempts to assign a certain number of items to a particular category X, you can make the following counts:

"a" = the number of items the model correctly assigns to X

"b" = the number of items the model incorrectly assigns to X

"c" = the number of items the model incorrectly rejects from X (that is, items that the model should assign to X but does not)

From these quantities, you can calculate the following performance measures:

- Precision = "a" /("a" + "b")
- Recall = "a" /("a" + "c")

Generally, for increasing precision you pay the price of decreasing recall. That is, the model assigns an item to a category only when it is very sure that the item belongs. But by insisting on being very sure, it runs the risk of rejecting items that really do belong in the category.
Here is another way to look at it. Suppose that at a confidence rating of 50, precision = 70 and recall = 80. Then, for text T and category C, the statements in the table "Example of Precision and Recall" hold.

Example of Precision and Recall

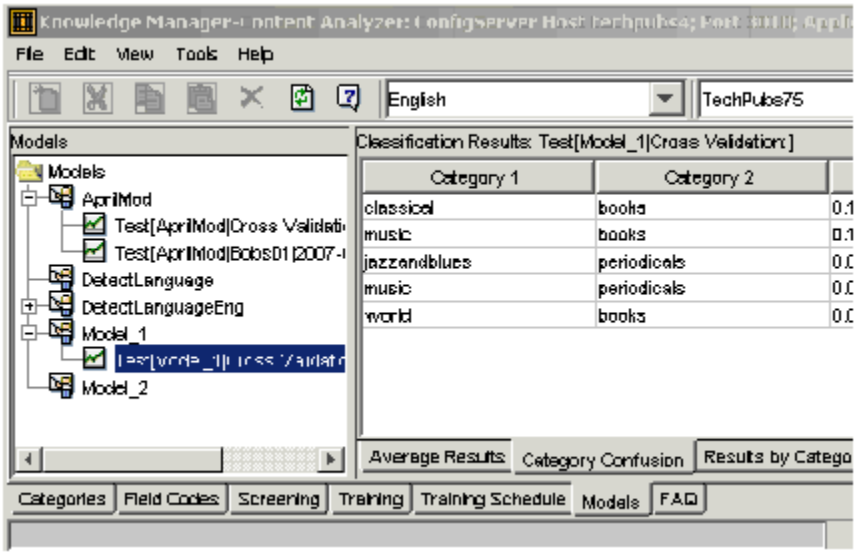
If you know that	then you can infer that	this percent of the time
T belongs to C,	The model will classify T as C, with confidence of over 50%,	80%
The model classifies T as C, with confidence of over 50%,	T does belong to C,	70%

Use the Average Results ratings as an assessment of the overall quality of the model. See also Applying the Ratings.

Category Confusion Subtab

This topic describes part of the functionality of Genesys Content Analyzer.

This subtab, shown in "Models Tab: Category Confusion Subtab," lists up to 10 pairs of categories that the model is likely to confuse.



Models Tab: Category Confusion Subtab

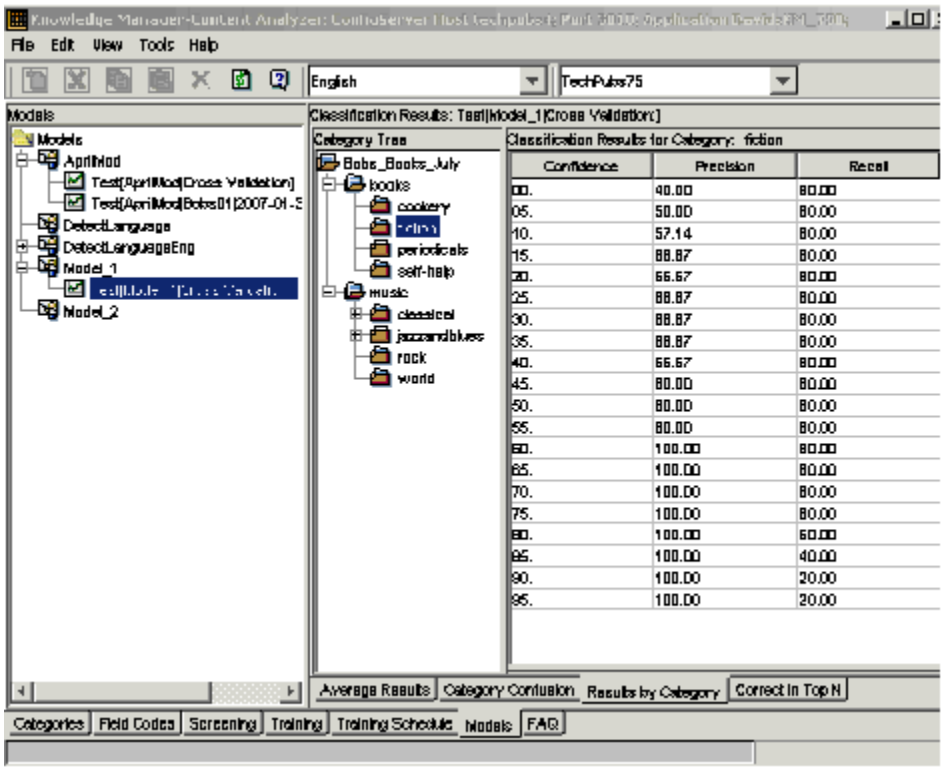
The `Confusion` column gives the probability (between 0 and 1) of confusion between the categories in the `Category 1` and `Category 2` columns. A rating of 0.5 would mean total confusion: the model cannot distinguish A from B. A rating of 1.0 would mean that the model always calls A B and always calls B A —a complete reversal.

If a pair of categories has a rating of over 0.2 and both categories have more than three or four members, you should consider modifying them. You can modify them in either of two directions:

- Merge them (decide that they are so similar they amount to a single category).
- Further differentiate them by adding more highly contrasting e-mails to them in the training object.

Results by Category Subtab

This topic describes part of the functionality of Genesys Content Analyzer. This subtab, shown in Models Tab: Results by Category Subtab, displays the same ratings as the two preceding subtabs, but for a single category. A central pane displays the category tree; select a category to display `Confidence, Precision, Recall,` and `Confusion` on the right-hand pane.



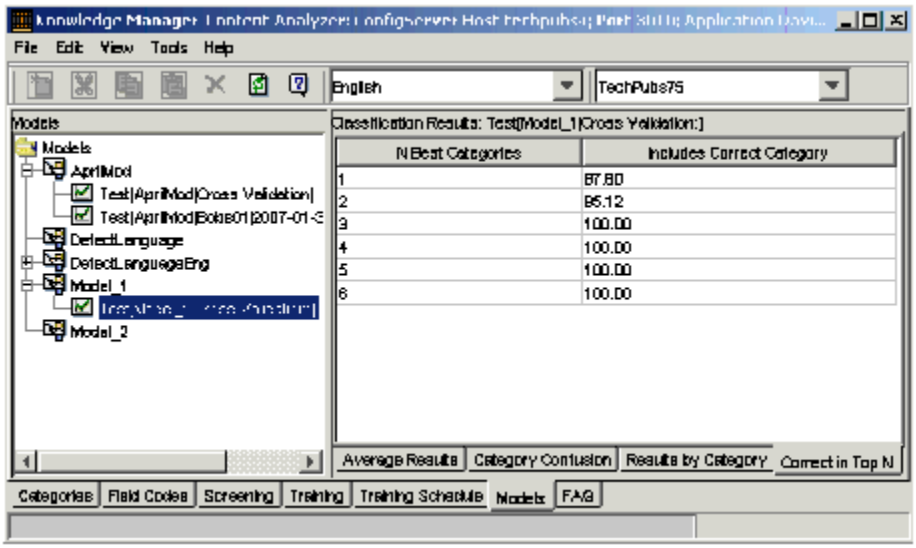
Models Tab: Results by Category Subtab

Use the `Results by Category` ratings to help you set the confidence level in IRD objects, as described in Applying the Ratings.

Correct in Top N Subtab

This topic describes part of the functionality of Genesys Content Analyzer.

When a model classifies a text object, it returns a list of categories and the probability that the object belongs to them. Ranking the returned categories with the highest probability first, how likely is it that the correct category appears within the top two, the top three, and so on? Ratings of this likelihood are displayed on the `Correct in Top N` subtab, shown in "Models Tab: Correct in Top N Subtab."



Models Tab: Correct in Top N Subtab

Look at the first row for a good indication of the overall quality of the model. Its general meaning is that a single classification attempt using this model will be correct for this percentage of categories.

The other rows can help you further assess model quality. A very accurate model may have a 95 per cent probability that the correct category appears in the top two or three. For a less accurate model, you may have to go down to the top five or six to achieve 95 per cent coverage.

You can also use this rating to advise agents how many categories to look at when choosing a standard response. If there is a 95 per cent probability that the right category is in the top three, you can advise agents to examine only the top three categories.

See also Applying the Ratings.

Reporting on the Ratings

This topic describes part of the functionality of Genesys Content Analyzer.

You can produce a report on the ratings of a model, either by directly printing it or by generating an HTML file. To obtain a report, first select a model's ratings node to display the ratings. Then do one of the following.

- Print directly:

a. On the right-hand pane, right-click and select `Print.`

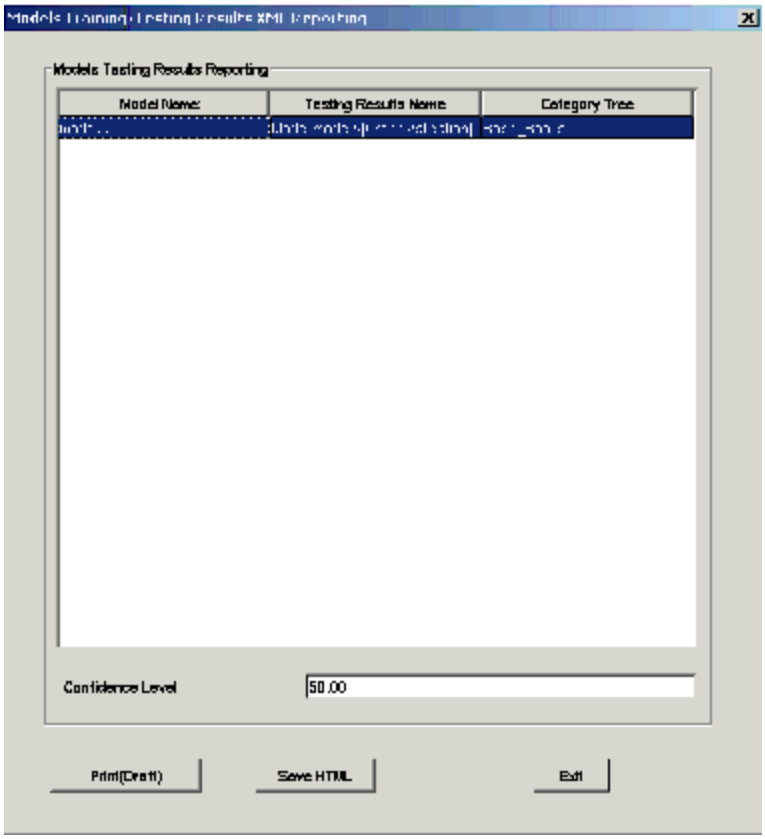
b. Proceed through Page setup and Print dialog boxes.
This report contains the figures from the `Average Results` and `Correct in Top N` tabs.

- Use the following procedure to produce an HTML file.

Procedure: Producing an HTML Report on Ratings

1. On the left-hand pane, right-click and select `Print XML Report.`

The `Models Training/Testing Results XML Reporting` dialog appears, as shown in "Models Training/Testing Results XML Reporting Dialog Box."



Models Training/Testing Results XML Reporting Dialog Box

- Select the model(s) that you want to report on.
- Set a confidence level. This determines the way that results by category are displayed in the report; for details, see Item "5. "Results for Categories," below.
- Do one of the following:
  - Click `Print(Draft)`. This produces a printout of the HTML file.
  - Click `Save HTML`, then choose a filename and location for the report.

The resulting report has the following structure:

1. Introductory material, including definitions of precision, recall, and an additional measure called F1, which is a kind of averaging (more precisely, the harmonic mean) of precision and recall.
2. Model name and information, as it appears when you select the model on the `Models` subtab.
3. Microaverage Table, which reproduces the statistics from the `Average Results` subtab.
4. Correct category In Top N Categories, which reproduces the statistics from the `Correct in Top N` subtab.
5. Results for Categories, which reproduces the statistics from the `Results by Category` subtab. It does this by listing the following for each category:
   - Name

- Precision at the confidence level that you set when producing the report.
- Recall at the confidence level that you set when producing the report.
- F1 averaging for the precision and recall in the preceding two items.
- The top two categories likely to be confused with this category, with their confusion ratings.

Design and Use Considerations

This topic describes part of the functionality of Genesys Content Analyzer.

Getting started with Genesys Content Analyzer requires four basic steps. This section provides information and advice on these steps, as follows:

1. Create a category tree and standard responses. See Design.

2. Create a training object using the tree. Add text objects (e-mails and other objects) to the training object. See Design.

3. Train a new model using the training object. See When to Train.

4. Test the model and use the resulting ratings. See Applying the Ratings.

Design

This topic describes part of the functionality of Genesys Content Analyzer.

In designing your category trees and standard responses, remember that they will have two very different groups of “users”—that is, agents and training.

- Training uses the categories plus categorized e-mails to generate models.
- Agents use categories in two quite different ways:
    - They use the categories to find standard responses.
    - They give feedback on the category/standard response system, essentially indicating, “Yes (no), the standard response of this category is (is not) a good match with this e-mail,” affirming that this e-mail should/should not be tagged with this category. This tagging becomes one of the attributes of the interaction as it is stored in the Universal Contact Server database.

> **Important**
>
> Agents can use standard responses without giving feedback, but if they do not give feedback you cannot collect enough categorized emails to be useful for training. You then have to create e-mail manually from Knowledge Manager’s `Training` tab.

Given the importance of high-quality feedback, you may want to designate a special group of agents for this purpose: define the categorizing of interactions as one of their main duties. Remember: the more categorized e-mails you have and the more accurate the categorization is, the more likely the system is to produce accurate models

In designing your category trees and standard responses, keep in mind the following:

- Do not create too many categories. Many categories allows for many standard responses, and if there are large numbers of standard responses agents are likely to use some responses very little or not at all. This creates the following chain of causation:
    a. There are very few e-mails tagged with a particular category.
    b. The system cannot train for that category.
    c. The system cannot suggest that response.
    d. That category and its response continue to be used very little.

    In short, excess categories are likely to not be used.
- Try to make categories sufficiently distinct. If two or more standard responses apply to very similar situations, training has difficulty producing a model that can tell them apart.
- Avoid categories/responses that are too general, like “Not enough information.” Agents will use only one

or two such general responses and ignore any others, with two undesirable results:

- Training has a hard time producing a good model because the e-mails it uses have a huge variety of content.
- The system is unable to include the unused categories/responses in training, because there are very few e-mails tagged with those categories in the database.

When to Train

This topic describes part of the functionality of Genesys Content Analyzer.

When does your training object have enough categorized text objects to make training worthwhile? Here are some possible situations and comments on them.

Uniformly Low Feedback

In this situation, all categories have a small amount of feedback (less than about 12 text objects per category). This object is not fully ready for training. You can still try training a model, but you should be aware that the results probably will not be very good.

Unbalanced Feedback: Mostly Low

In this situation, all categories except a small group have a small amount of feedback (less than about 12 text objects per category). The small group (one to five categories) may have several hundred or even thousands of feedback objects per category. You can train a model, but the resulting model will mostly return the categories from the small group. A situation of this type may have these causes:

- It may be an accurate reflection of the situation. For example, your company may sell 25 products but just three of them may account for 90 per cent of its business.
- It may reflect shortcomings in the system:
  - Agents may not use standard responses properly.
  - The standard responses and/or the category tree may be poorly designed.

To determine which of these causes obtains, inspect your category tree, standard responses, and agents' use of them. If the situation arises because of shortcomings in the system, consider doing the following:

- Bring some balance into the training object by deleting some of the text objects associated with categories of the high-feedback group.
- Modify the low-feedback categories.

Unbalanced Feedback: Mostly High

In this situation, some categories have a small amount of feedback (less than about 12 text objects per category), but a significant number (over 50) of categories have a large amount of feedback (over 30–50 text objects or more per category). This is a rather common situation. You can train a model and it will work acceptably on the high-feedback categories. But consider modifying the low-feedback categories.

Uniformly High Feedback

In this situation, almost all categories have significant feedback (over 50 text objects per category).

This is the best situation. It means that agents are frequently using almost all standard responses. You can train the model and it should perform well on all categories.

Applying the Ratings

This topic describes part of the functionality of Genesys Content Analyzer.

Ratings of models, described in Testing Models, have several possible uses.

Assessing Overall Quality

To assess the overall quality of a model, look at the ratings on the Average Results Subtab and the Correct in Top N Subtab.

Identifying Confusing Categories

To identify categories that may be too similar and/or have insufficient feedback, use the ratings on the Category Confusion Subtab. Consider modifying any pair of categories that has a confusion rating over 0.2.

Using in Routing

To decide where to set the `Confidence` level in an IRD `Classify` object (see "eServices Objects" in the "Interaction Routing Designer" chapter of the Universal Routing 8.1 Reference Manual, use the ratings on the Average Results Subtab and the Results by Category Subtab. The way that you use these ratings differs according to the task that you are performing at this step of the strategy:

- If you are using the classification to choose an autoresponse or acknowledgment: Set a relatively high `Confidence` value, one at which `Precision` is 85 or higher and `Recall` is 5 or higher.
- If you are using the classification to choose standard responses as suggestions to the agent: Set a `Confidence` value of 1. Even a very low `Precision` rate (for example, 15) is safe because an agent will make the final decision on whether to use the standard response. Also the lower the `Confidence` value, the more categories are returned, and:
  - The higher the probability that the correct category is among them.
  - The more categories the agent can provide feedback on.
- If you are using the classification to determine where the interaction goes in the next step of the strategy: Set the `Confidence` value at the point where `Precision` is approximately equal to `Recall`.

Language Detection Model

As part of Content Analyzer, Genesys provides a model that classifies e-mails as English, French, German, Italian, Portuguese, Russian, or Spanish.
To import this model and its training object, use the following procedure.

Importing a language classification model

1. Select unknown as the language. If there is no such language you must create one in Configuration Manager (see Notes on Language).
2. Select the Import command.
3. Click Browse and navigate to <KnowledgeManagerHome>\LanguageModel\lang.kme. <KnowledgeManagerHome> is normally C:\Program Files\GCTI\eServices 8.1.0\Knowledge Manager.
4. Click OK.

The training object consists of seven categories, one for each language. Each category contains a number of text objects in its language. You can add more text objects to these categories as well. This could be especially valuable if you have a collection of text objects (such as e-mails) whose subject matter relates to your business. After you add text objects, you must train a new model to take advantage of the added data.

You can also add other languages to the model, as follows:

Procedure: Adding more languages to the model

1. On the Categories tab (still with unknown selected as the language), add a category for the new language to the LanguageDetection category tree.
2. On the Training tab, select the LanguageDetection training object, then select the new language category in the training object.
3. Add text objects in the language.
4. Train a new model that includes the new language.

You can do this for any language supported by E-mail Server (E-mail Server supports all languages that are supported by the version of JRE that is supplied with Genesys eServices). However, Genesys has not tested any language other than those listed above.

Analyzing Sentiment and Actionability with Content Analyzer

This topic describes part of the functionality of Genesys Content Analyzer.

It includes:

- Sentiment
- Actionability

You can use Genesys Content Analyzer to analyze the sentiment and actionability of interactions that have been brought into the system by Genesys Social Messaging Management. Genesys supplies samples which demonstrate these capabilities.

Sentiment

To deploy the sentiment sample, use the following procedure.

1. In Configuration Manager or Genesys Administrator, create a language called English_Sentiment.
2. With Knowledge Manager set to that language, import the file `EnglishSentiment.kme`, which is located in the `<KnowledgeManagerHome>\SentimentModel` directory.

This provides:

- A model `SentimentSampleModel` for analyzing sentiment.
- The training object `Sentiment` that created that model.
- A category tree `SentimentDetection` that contains the categories to assign to interactions as a result of the analysis.

Actionability

To use the actionability sample, import the file `Actionability.kme`, which is located in the `<KnowledgeManagerHome>\ActionabilityModel` directory.

This provides:

- A model `Actionability` for analyzing actionability.
- The training object `Actionability` that created that model.
- A category tree `Actionability` that contains the categories to assign to interactions as a result of the analysis.

Next Steps

You can use the sample training objects to produce new models, improving the quality by making adjustments such as:

- Altering the settings such as those for quality level. See Step 4 of the Procedure: Scheduling training using the Model Options tab and Cross-Validation.
- Using the Mail Editor to edit the content of the messages in the training object. See Step 3 of Creating New E-mails Manually.
- Using the Mail Editor to add more sample messages to the training object.

Genesys also provides sample screening rules for detecting sentiment and actionability.

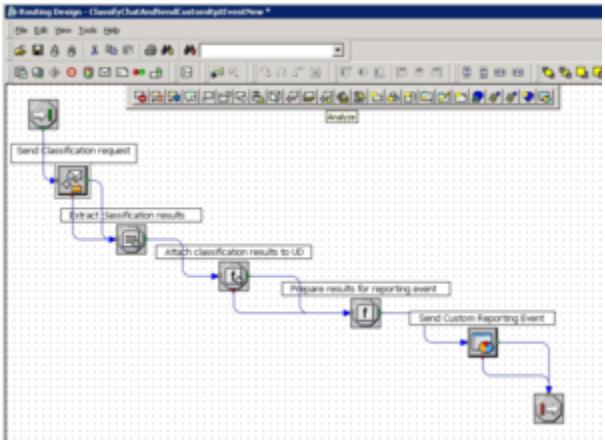
For more information on Genesys Social Messaging Management, see the *eServices Social Media Solution Guide,* available on the eServices product page.

Classifying Chat Transcripts

> **Important**
> This page applies only to environments that use Classification Server 8.5.3 or higher.

Chat conversations typically do not contain much contextual information prior to routing. Consequently, for reporting purposes, managers cannot determine which customer issues are being discussed in chat.

However, using Classification Server and Interaction Routing Designer (IRD), you can create a business process to classify chat transcripts post-interaction, enabling you to accurately categorize the conversation and improve reporting data.



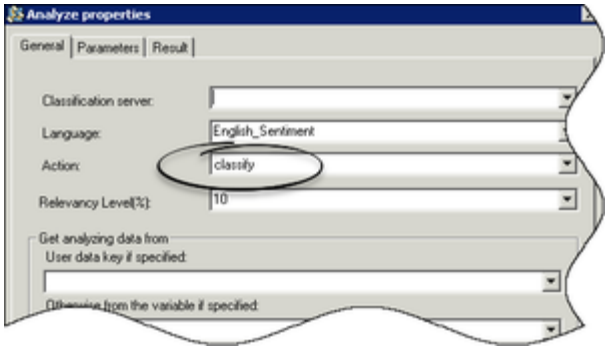
A sample subroutine in Interaction Routing Designer that classifies chat content and sends a customized reporting event with results to Interaction Server.

How to implement

In IRD, use the Multimedia object **Analyze** to classify chat transcripts using Classification Server.

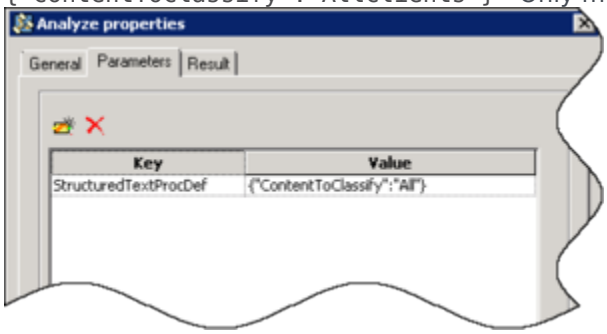
Simplified instructions on how to implement the **Analyze** object in IRD are provide below. Refer to Universal Routing documentation for more detailed information.

1. From the **Multimedia** toolbar, click the **Analyze** object and then click inside the Routing Design window workspace to insert the object.
2. Double-click the **Analyze** object to open its properties dialog box.
3. In the **General** tab, ensure the **Action** drop-down is set to **classify**.



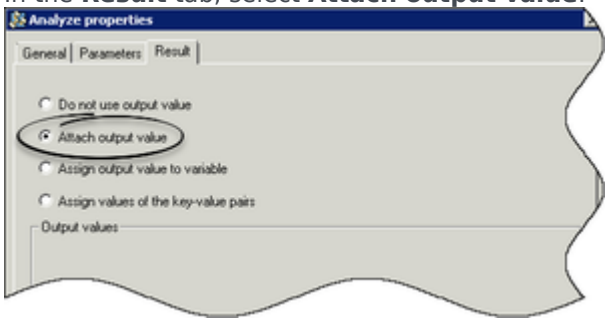
The General tab of the Analyze object.

4. In the **Parameters** tab, set one of the following values for **StructuredTextProcDef**:
   - `{"ContentToClassify":"All"}`—All messages from a chat session are classified.
   - `{"ContentToClassify":"AllAgents"}`—Only messages from agents are classified.
   - `{"ContentToClassify":"AllClients"}`—Only messages from clients are classified.



The Parameters tab of the Analyze object.

5. In the **Result** tab, select **Attach output value**.



The Result tab of the Analyze object.

6. Click **OK**.

FAQ Objects

This topic describes part of the functionality of Genesys Content Analyzer.

Taking a category tree and its associated standard responses as input, Knowledge Manager can produce an FAQ object. From this object Knowledge Manager can produce a .jar file, which can in turn be used to:

- Build a web application that accepts written requests and, using content analysis, returns a set of standard responses.
- Present the contents (or a selection from the contents) of the standard response library as answers to frequently-asked questions.

An FAQ object combines a category tree, a training object based on the tree, and, optionally, a model built from the training object. The model is required in order to build a web application.

FAQ objects allow you to include in your web application a means of gathering user feedback about the correctness of a returned standard response. The application then uses this feedback to update the confidence rating of that particular standard response. This functionality is exemplified in the FAQ sample in the Simple Samples that are installed along with Web API Server.

For a description of this sample and its source code, see the eServices 8.1 Web API Client Developer's Guide.

The section includes the following:

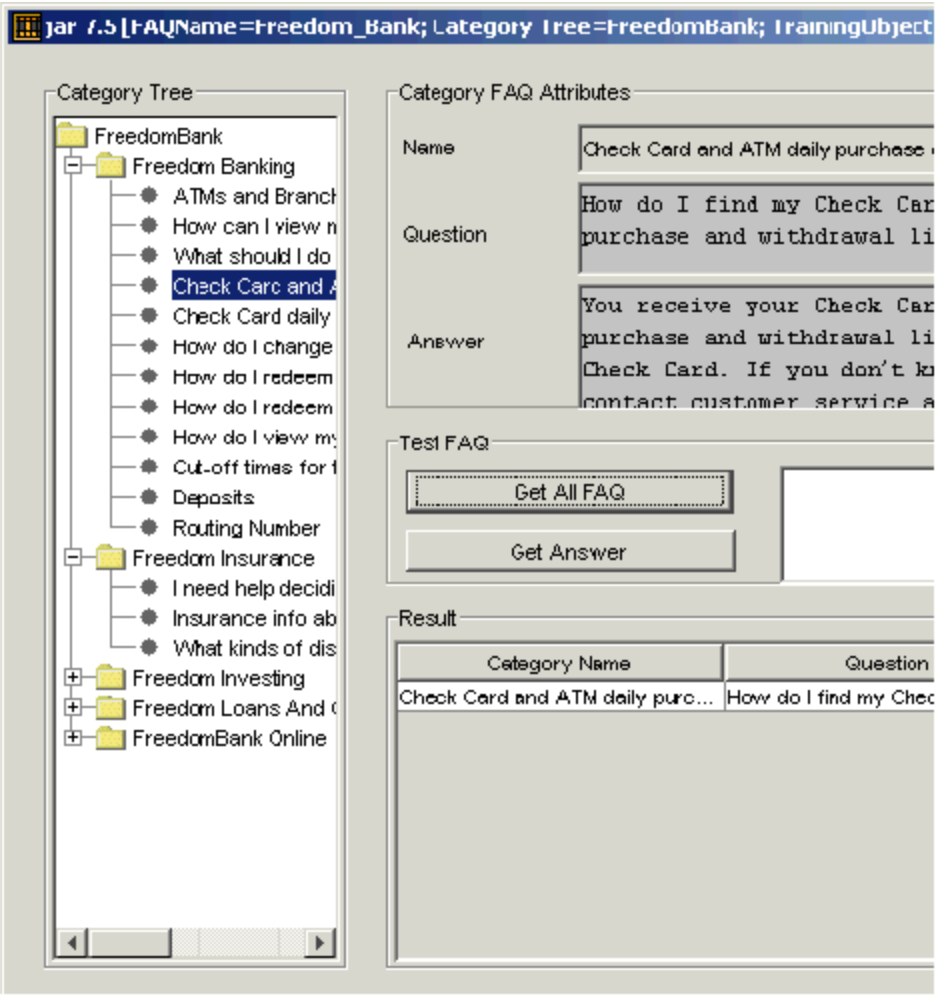
Sample FAQ .jar File

This topic describes part of the functionality of Genesys Content Analyzer.

The sample FAQ that is supplied with Knowledge Manager demonstrates the way that an FAQ object can present a question/answer list. This sample is also included in the Web API samples. (See the eServices 8.1 Web API Client Developer's Guide).

The filename of the sample is `faq_example.jar`. The installation places it in a directory called `FAQExample` in the Knowledge Manager home directory (normally `C:\Program Files\GCTI\eServices 8.1.0\Knowledge Manager`). To use this sample:

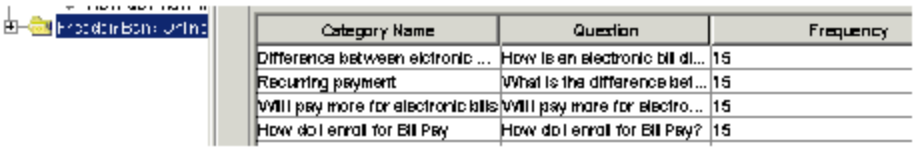
1. Use a text editor to open its batch file `unit_test.bat`.
2. Edit the line
   `set JAVAHOME=D:\jdk1.4\bin\java`
   so that it points to the location of Java in your environment.
3. Use the batch file to launch the sample. You should see the window shown in the figure "FAQ Sample."



FAQ Sample

This FAQ object works as follows:

- Click a category on the left-hand pane to see its name and associated question and answer displayed in the boxes at upper right.
  The main function of the nonterminal categories (represented as folders) is to organize the terminal categories (represented as grey disks), which are the main locus of questions and answers. Mostly the nonterminal categories do not have answers, and their `Category Question` box displays either a duplicate of the category name or a short description, rather than a question.
- Click `Get All FAQ` to display, in the `Result` area, a list of all categories contained in the selected category (and its subcategories), along with the question and frequency for each. An example is shown in "Get All FAQ for a Selected Category."



Get All FAQ for a Selected Category

The `Result` area lists all subcategories of the category that is selected on the left-hand pane, providing the following information:

- Category name
- The question associated with the category
- The frequency rating—that is, the number of text objects that are associated with this category in the training object that is part of the FAQ.
- To test the FAQ object, enter a sample question in the `Testing Question` box, then click `Get Answer`. The `Result` area displays the answers that the system provides for the sample question at the selected level of the tree. In addition to the category name, question, and frequency, the system also displays a confidence rating. Confidence takes the model that is part of the FAQ object and tells you how good that model is at assigning new questions to this category.

More About FAQ Objects

This topic describes part of the functionality of Genesys Content Analyzer.

To be used in an FAQ object, a category must have all of the following attributes:

- Answer: It must have a standard response of the FAQ usage type. This attribute is optional for nonterminal categories.
- Question: It must have an associated question, to which the standard response can serve as the answer.
- Selection: It must be selected for inclusion in the FAQ object.

You must select the categories to include because you cannot assume that all categories in the tree are suitable for use in an FAQ list. For example, your tree might include a category that exists only to provide the `standard response` "Your account is overdrawn." Or you might want to use a single category tree to produce multiple FAQ objects, with some categories selected in one FAQ object but not in others.

Requirements are slightly different for terminal categories (those without subcategories) and nonterminal categories (those with subcategories), as shown in "Required Attributes for FAQ Categories."

Required Attributes for FAQ Categories

Attribute	Nonterminal Category	Terminal Category
Answer	Optional	Required
Question	Required	Required
Selected	Required	Required

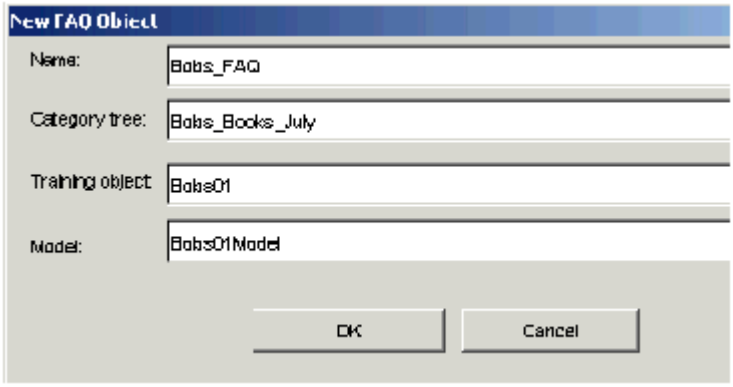
You can create answers and questions either on the `Categories` tab or the FAQ tab. But selection can be done only on the FAQ tab, after you generate the FAQ object.

Procedure: Creating an FAQ Object

This topic describes part of the functionality of Genesys Content Analyzer.

To create an FAQ Object, use the following procedure.

1. On the FAQ tab, do one of the following:
   a. Select File > New.
   b. On the left-hand pane, right-click and select New FAQ Object.

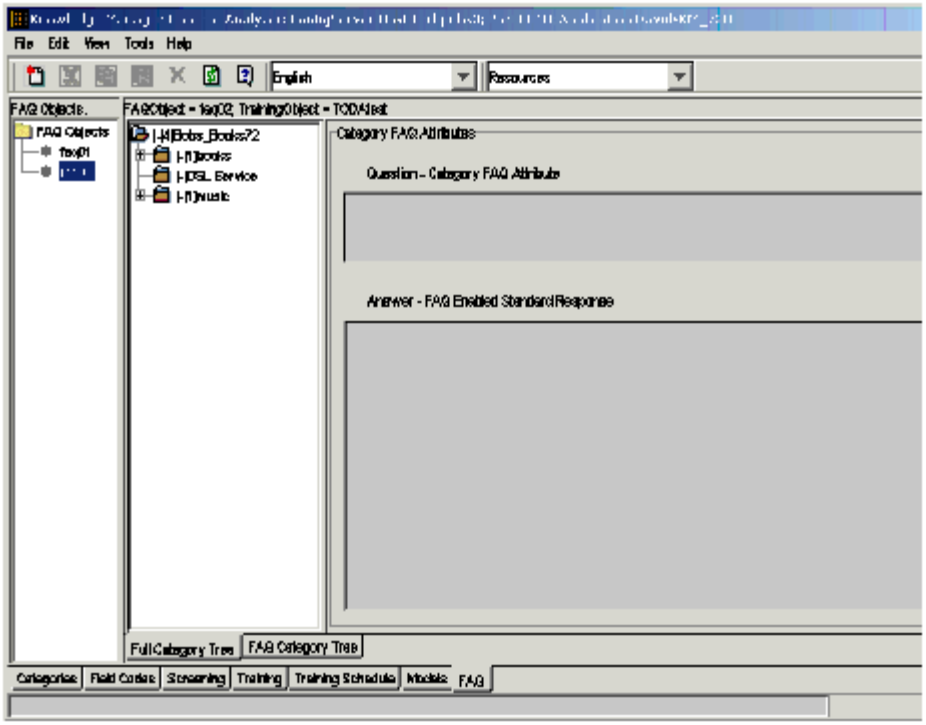


New FAQ Object Dialog Box

3. The New FAQ Object dialog box appears, as shown in "New FAQ Object Dialog Box." On it:
   a. Enter a name.
   b. Select a category tree, training object, and model.
      The category tree and training object are required. You can create an FAQ object without a model, but you will not be able to use it in conjunction with content analysis.

3. Click OK.

   The FAQ tab then appears as in the figure "FAQ Object."



FAQ Object

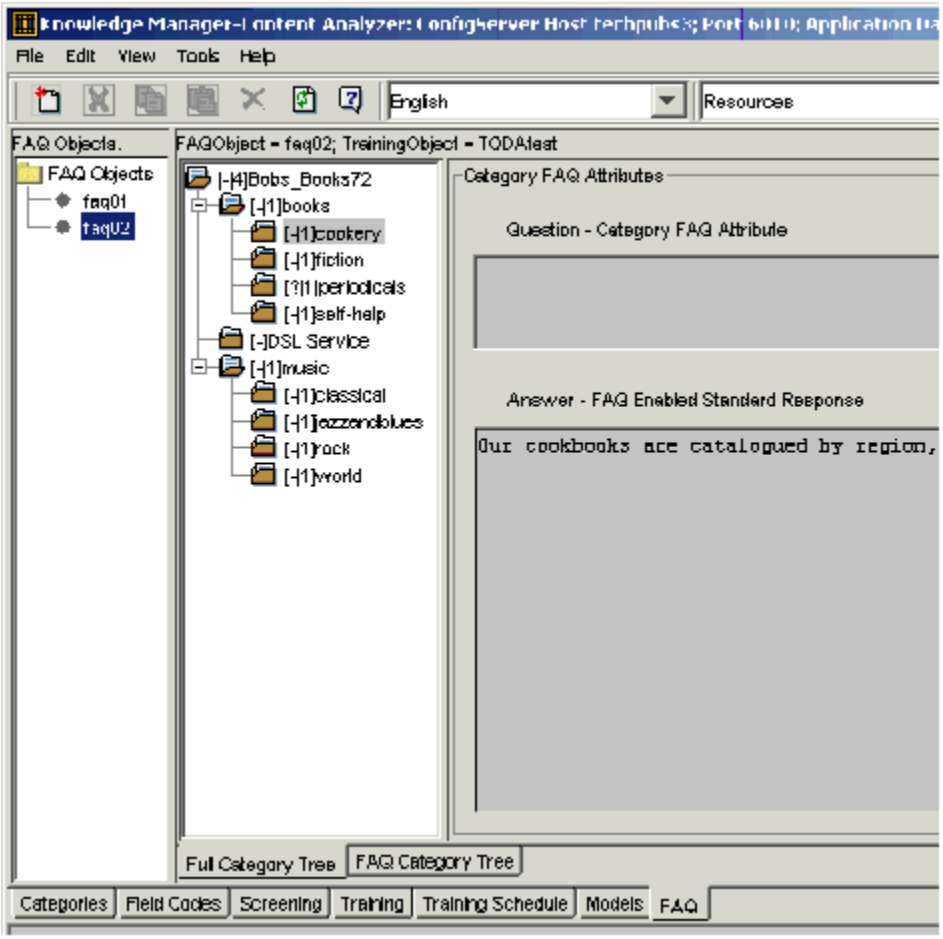
The FAQ tab contains the following panes:

- Left—Displays a list of all FAQ objects
- Center, with two subtabs:
  - Full Category Tree—Displays the entire category tree that serves as the source of the FAQ object that is selected in the left pane
  - FAQ Category Tree—Displays the selected FAQ object itself
- Right—These panes vary, depending on which of the center subtabs is selected.

Full Category Tree Subtab: Configuring the Category Tree

This topic describes part of the functionality of Genesys Content Analyzer.

Expanding the category tree on the center pane produces the details shown in "Center Pane of FAQ Object Tab."



Center Pane of FAQ Object Tab

The following information displays:

- The strip above the center and right-hand panes displays the names of the FAQ object, the training object, and the model if any.

- The right-hand panes display the question and answer (standard response) associated with the category that is selected on the center pane.

The display on the center pane requires more explanation.

On the center pane, each category name appears in the form
`[<statuscode>|<number>]<categoryname>`
<number> is the number of text objects (e-mails and other objects) associated with this category in the training object.
The status code combines the following information:

- Whether the category has the attributes that are required for it to be selected for use in the FAQ object.
- Whether the category is selected for use in the FAQ object.

Category Status in FAQ Objects lists the possible statuses:

Category Status in FAQ Objects

Status	Meaning
-	Category lacks the required attributes; therefore it cannot be selected.
?	Category has the required attributes but is not selected
+	Category has the required attributes and is selected

Answer

In the context of an FAQ object, an answer is a standard response that has the FAQ usage type selected and is specified as `Active`. You select this usage type on the `Additional` tab of the `New Standard Response` or `Edit Standard Response` dialog boxes. There are two ways to access these dialog boxes:

- On the `Categories` tab, as described in Creating Standard Responses .
- On the `FAQ` tab:
- To edit an existing standard response, right-click a category and select `Edit Response.`
- To create a new standard response, right- click a category and select `Set Response.`

Question

Create these questions on the `FAQ attribute` tab of the `New category` or `Category` dialog boxes. As with the usage type of a standard response, there are two ways to access these dialog boxes:

- On the Categories tab, as described in Creating a Category Tree .
- On the FAQ tab, right-click a category and select Set/Edit Question.

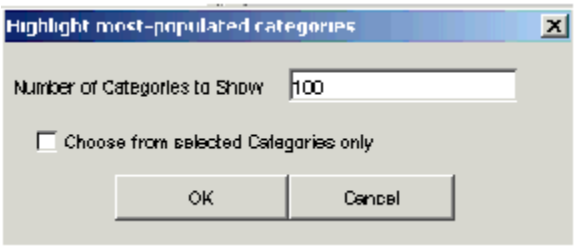
Selection

To select a category for inclusion in the FAQ object, right-click it and select:

- Select to select only this category.
- Select With Children to select this category and all subcategories under it.

Most-Used FAQs

Right-clicking anywhere on the Full Category Tree subtab and selecting Highlight Top Categories opens the Highlight most-populated categories dialog box, as shown in "Highlight Most-Populated Categories Dialog Box."

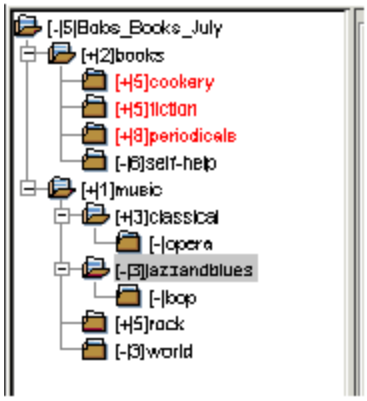


Highlight Most-Populated Categories Dialog Box

When you click OK, the dialog box closes, and the center pane highlights, in red, the top "X" categories, where:

- "X" is the number in the Number of Categories to Show box.
- "top" means having the largest number of associated text objects in the training object.

Select the Choose from selected Categories only check box if you want the calculation to consider only the categories that you have selected for inclusion in the FAQ object (the ones with + status).
The figure "Three Most-Used FAQs" shows an example.



Three Most-Used FAQs

The figure "Three Most-Used FAQs" shows the result of the following settings:

- `Number of Categories to Show` is set to 3.
- `Choose from selected Categories only` is selected.

If the `Choose from selected Categories only` check box were not selected, the highlighted categories would be `periodicals, self-help,` and `Bobs_Books_July.`

> **Important**
>
> If multiple categories are in a tie for inclusion in the top X, the categories that come first (highest) in the listing are the ones highlighted. In the figure "Three Most-Used FAQs" there are three categories with five text objects, but the top three has room for only two of them. So *cookery* and *fiction* are highlighted, but not *rock*, which is further down the list.

Shortcut Menu

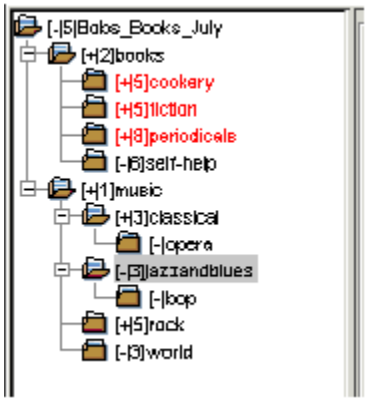
The complete list of commands that appear when you right-click a category is as follows:

- `Select`—Select the category.
- `Select With Children`—Select the category and all subcategories under it.
- `Unselect`—Unselect the category.
- `Unselect With Children`—Unselect the category and all subcategories under it.
- `Expand Tree`—Self-explanatory.
- `Collapse Tree`—Self-explanatory.
- `Highlight Top Categories`—Open the `Highlight most-populated categories` dialog box.

- `Set/Edit Question`—Create or edit a question for the category.
- `Set Response`—Create a new standard response for the category. For the response to appear in the FAQ object, you must go to the `Additional` tab and select `Active` FAQ usage.
- `Edit Response`—Edit an existing standard response for the category.

FAQ Category Tree Subtab: Viewing and Testing

After you have finished configuring the category tree, you can display the result on the FAQ Category Tree subtab, as shown in "FAQ Category Tree Tab."



FAQ Category Tree Tab

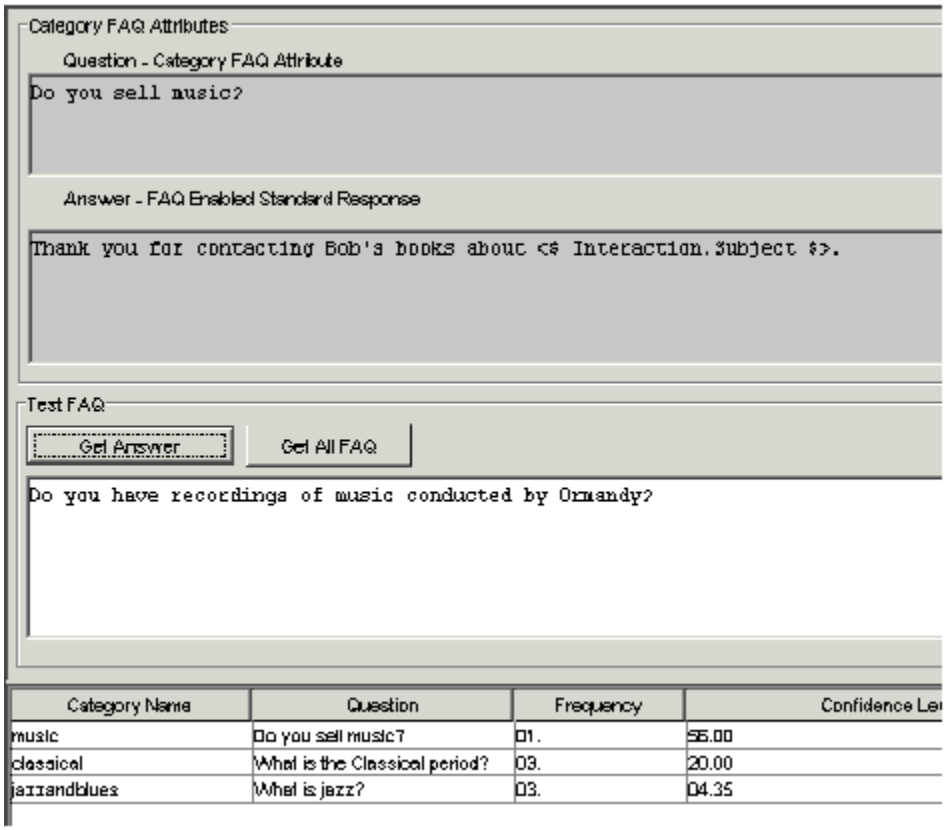
The FAQ Category tree is a preview of the FAQ object using a format similar to the sample described previously in Sample FAQ .jar File.
Selecting a category on the FAQ Category Tree subtab produces the following results:

- The Category FAQ Attributes area displays the question and answer for the category that is selected on the center pane.
- In the Test FAQ area, click Get All FAQ to display a list of all categories contained in the selected category (and its subcategories), along with the question and frequency for each.

> **Important**
>
> Text in the text box has no effect on this action. Categories that have questions but no answers do not appear on this list.

In the Test FAQ area, enter a sample question in the text box, then click Get Answer to display the answers that the system provides for the sample question at the selected level of the tree. (If the FAQ

object does not include a model, this button is dimmed.) An example is shown in "Test FAQ: Get Answer."



Test FAQ: Get Answer

The figure "Test FAQ: Get Answer" shows the result of the question "Do you have recordings of dance music conducted by Ormandy?" with the `Music` category selected, as in the figure "FAQ Category Tree Tab," for the category tree in question.

The bottom pane lists all answers supplied by the system, showing the category name, question, frequency (meaning the number of text objects associated with the category in the training object, also shown as <number> on the `Full Category Tree` tab display), and confidence rating, as described in Reading and Understanding the Ratings.

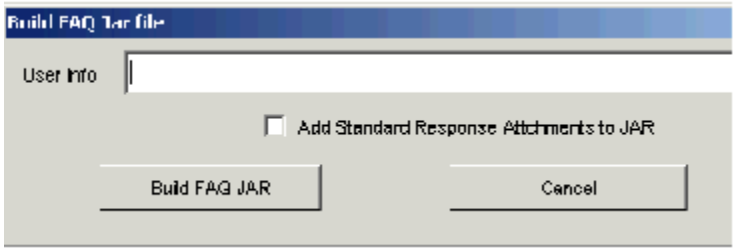
Procedure: Generating an FAQ.jar File

This topic describes part of the functionality of Genesys Content Analyzer.

1. Select an FAQ object on the left-hand pane, then do one of the following:

- Select `Tools > Build FAQ JAR file.`
- Right-click on the left-hand pane and select `Build FAQ JAR file.`

The `Build FAQ Jar file` dialog box appears, as shown in "Build FAQ Jar File Dialog Box."



Build FAQ Jar File Dialog Box

- In the `User Info` text box, enter the text that you want to appear in the title bar of the window that displays the FAQ object when you test it (see the procedure that immediately follows).
- Select `Add Standard Response Attachments to JAR` if the answers in your FAQ object have attachments that you want to include in the FAQ.jar file.
- Click `Build FAQ JAR.`
- The `JAR File Chooser` dialog box appears. Enter a file name.
- To test the resulting `FAQ.jar` file, make a copy of the `unit_test.bat` file that accompanies the FAQ example.
- Edit it so that it targets the .jar file that you built instead of `faq_example.jar.`

> Running the .bat file displays your FAQ object in the same way as the FAQ example. Notice that the title bar of the FAQ object window displays the text that you entered in the `User Info` text box of the `Build FAQ Jar file` dialog box.

Next Steps

- You can now use the `FAQ.jar` file to create web applications.

Typical Response Times

This topic describes part of the functionality of Genesys Content Analyzer.

It includes some typical response times for Genesys Content Analyzer. For other functions of Knowledge Manager see the discussion in Typical Response Times.

Unless otherwise stated, these figures are for a machine running Windows 2000 with two Pentium 4 processors and 1 GB of RAM.

- Deleting a training object takes approximately 4 seconds per 1,000 e-mails.
- Copying e-mails from one training object to another takes approximately 8 seconds per 1,000 e-mails.
- Creating a model (training time) naturally varies with the number of categories, number of e-mails, selected training quality, and selected cross-validation. As one example, for a training object containing 76 categories and 73,000 mails, with training quality set to level 1 and no cross-validation, training time is approximately 29 minutes. This is on a host running Windows 2000 with one 600 MHz processor and 1 GB of RAM.
- Cross-validation may increase training time significantly. The table "Increase of Training Time with Cross-Validation" shows, for selected cross-validation levels, the factors of increase of cross-validation over no cross-validation.

Increase of Training Time with Cross-Validation

Cross-Validation Level	Factor
3	1.9–2.9
5	4.0–4.7
10	8.0–9.0

For example, training a model at cross-validation level 3 takes between 1.9 and 2.9 times as long as the same model with no cross-validation.

- Classification performance depends on the size and nature (level of training quality) of the model. The table "Classification Performance" shows some examples, all of which use a test object that contains 3,726 text objects.

Classification Performance

Model	Host Machine	Classification Rate
72,734 text objects Size = 285 KB Quality = 1 Cross validation with split to three sets	Two Pentium 3 processors 512 MB RAM Windows operating system	31 objects classified per second
72,734 text objects Size = 309 KB Quality = 3 Cross validation with split to 10 sets	Two Pentium 3 processors 1 GB RAM Windows operating system	28 objects classified per second
72,734 text objects Size = 309 KB Quality = 3 Cross-validation with split to 10 sets	Four 350 MHz processors 4 GB RAM Solaris operating system	15 objects classified per second

- An FAQ object can process 30–50 classification requests per second on a model that contains 500–1,000 categories.