Confirmit Model Builder v3.2.0

This is revision 2 of the Confirmit Model Builder User Guide published in February 2017. The information herein describes Confirmit Model Builder and its features as of Build nr. 3.2.0. New features may be introduced into the product after this date. Go to www.confirmit.com or check “News” on the Customer Extranet for the latest updates.

Table of Contents

What's New in this Revision?

The following changes have been made in revision 2 of the Confirmit Model Builder User Guide:

- The text and illustration in the Model List Details section is edited (see The Models List Details on page 3 for more information).
- The text in the Category List section is edited (see The Category List on page 9 for more information).
- The Uncaptured Attributes section is added to the General Information chapter (see The Uncaptured Attribute on page 14 for more information).
- The text in the Logical Operators section is edited (see The Logical Operators on page 14 for more information).
- The Model Overview Tab section is added to the Creating a New Model chapter (see The Model Overview Tab on page 20 for more information).
- The Design Tab section is added to the Creating a New Model chapter (see The Design Tab on page 22 for more information).
- The Adding Comments to Expressions section is added to the Creating a New Model chapter (see Adding Comments to Expressions on page 24 for more information).
- The Uncategorized Comments section is added to the Creating a New Model chapter (see Uncategorized Comments on page 29 for more information).
- The text in the Launching the Model section is edited (see Launching the Model on page 31 for more information).

Note: The general layout and language in this document is continually being corrected, adjusted and improved to ensure the user has the best possible source of information. Only NEW information and details of functionality that has changed since the previous issue are listed here - minor corrections to the text and document layout are not listed.

Important
We need your feedback so we can improve this document and provide you with the information you require. If you have any comments or constructive criticism concerning the content or layout of this documentation, please send an email to documentation@confirmit.com. Please include in your email the section number and/or heading text of the section to which your comment applies.

1. What is Model Builder?

Confirmit Model Builder is an integral part of the Confirmit Genius text analytics application.

Model Builder is the application where you build the categorization model, which is used to analyze your text in Text Analytics.

The categorization model is built using a hierarchical tree structure and the nodes of the tree contain expressions, which are built up of keywords and Boolean operators. These expressions determine the resulting categorization of your text.

For example, a shop could ask customers what they thought of the service they received while making their most recent purchase. A respondent's reply could read: "Your employees were very knowledgeable and helpful, but the floor was dirty and the shop was rather untidy." The meaning of this sentence and its clauses is obvious to a person who knows the language, and this sentence could result in a “staff knowledge” category (with keyword “knowledgeable”) and a “staff helpfulness” category (with keyword “helpful”), the “shop cleanliness” attribute (with keyword “dirty”) and “tidiness” attribute (with keyword “untidy”). However if you have several thousand such replies, all of them slightly different, extracting the categories by manual analysis becomes a laborious and inaccurate job. Analysis by computer then becomes the only feasible option. After automatic processing, the several thousand replies will provide the shop with a set of data they can use as a basis for changes such that the customer experience, and hopefully thereby customer retention, is improved.

Model Builder addresses the categorization process for Text Analytics. There is also, of course, the sentiment process, which adds the positive, neutral or negative sentiment analysis to the overall verbatim and also each of the categories, which further enriches your analysis. Note that Model Builder does not deal with the sentiment process – this happens in the Genius Text Analytics processing task from Horizons.

During the analysis part of the process, each word in each response is compared against a database (the model) of words and groups of words that could be used in responses to the question. The model must therefore contain all the terms that are to be searched for in the responses. Model Builder is the application that assists you with creating and compiling the model.

As the vocabulary used in different situations can differ considerably, the model must be tuned for the specific customer domain. For example, the retail branch, call centers, consumer electronics branch, hotel branch and vehicle workshop branch etc. will all need different models. And even within branches, different specializations will need to be fine-tuned; for example a hardware store will need a different model than a chemist. Template building blocks are available that can be used to speed-up the process.

1.1. Logging In

When you go to the Model Builder URL, you are presented with the login page.

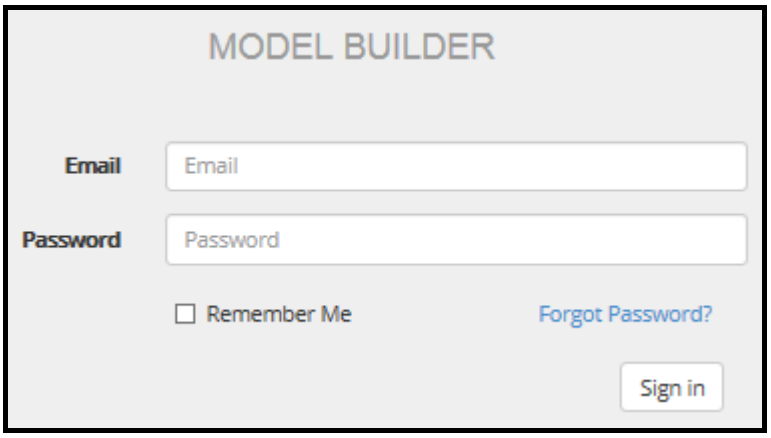

Figure 1 The Model Builder login page

Type your email address and password into the fields and click **Sign in**. In the event you have forgotten your password or you are a new user and do not yet have one, click **Forgot Password**. You are taken through the new password routine so you can reset your password (see Forgotten Password on page 2 for more information).

The Model List page opens. This shows all models that have been built within your customer account.

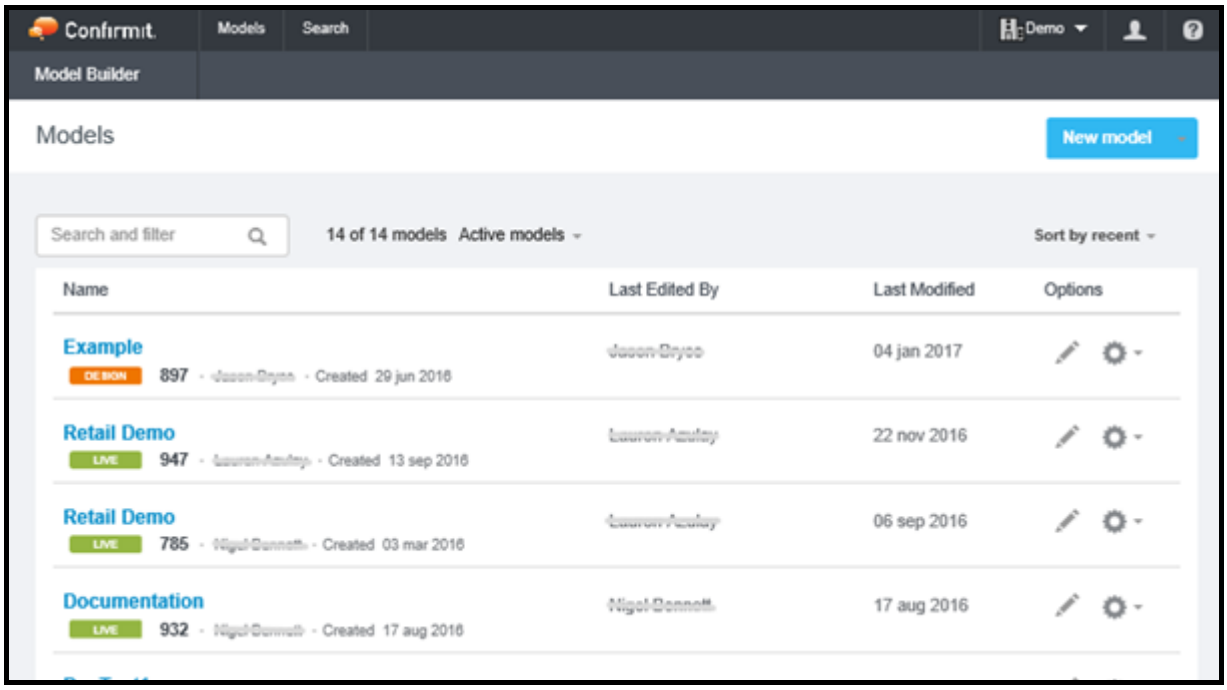

Figure 2 Example of the model list for a customer

Here you can select an existing model to work with, or create a new model.

Later when you have completed your work with Model Builder, close and log out of the application by going to the **User** menu and selecting **Log Out** (see The User Menu on page 7 for more information). This closes Model Builder correctly and returns you to the Log In page.

1.1.1. Forgotten Password

Your Model Builder password prevents unauthorized people from using your account and interfering with your models. You must therefore ensure your password is difficult to guess and is known only to you. This of course can create problems - if you forget your password then you will not be able to use Model Builder. The login page therefore includes the possibility for you to reset your password should you forget it, and also create a password the first time you log in to Model Builder. To do so:

1. Click the **Forgot Password?** link below the login page.

   The dialog shown below opens.

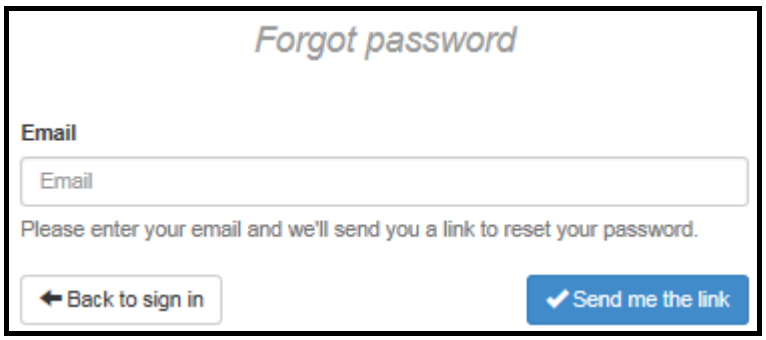

Figure 3 The "Forgot password" dialog

2. Type your email address into the field and click **Send me the link**.

   A message is sent to Confirmit. Assuming your email address is registered, an email will be sent to you containing a link and instructions for how you can reset your password.

Note: The reset password link is only valid for one hour. If you do not reset your password within the hour then the link will expire and you must repeat the procedure to be sent a new link.

1.1.2. Inactive-Screen Lock

For security reasons, when the Model Builder application has been inactive for 30 minutes it will be locked and you will have to log back in to continue. If you have not logged back in within three hours, Model Builder will close and any work that has not been saved will be lost. If you log back in within the three-hour period, Model Builder will reopen at the place you were last working.

Warning
Save your work at regular intervals, and ensure you save before you leave Model Builder unattended.

1.2. The Models List Details

When you open the Models page, each model that you have access to is displayed as one row in the list. To create a new model, click **New model** (see How to Create a New Model on page 19 for more information). To create a new model from an import, click on the drop-down arrow next to **New Model** then select **Import model** (see Importing a Model on page 16 for more information).

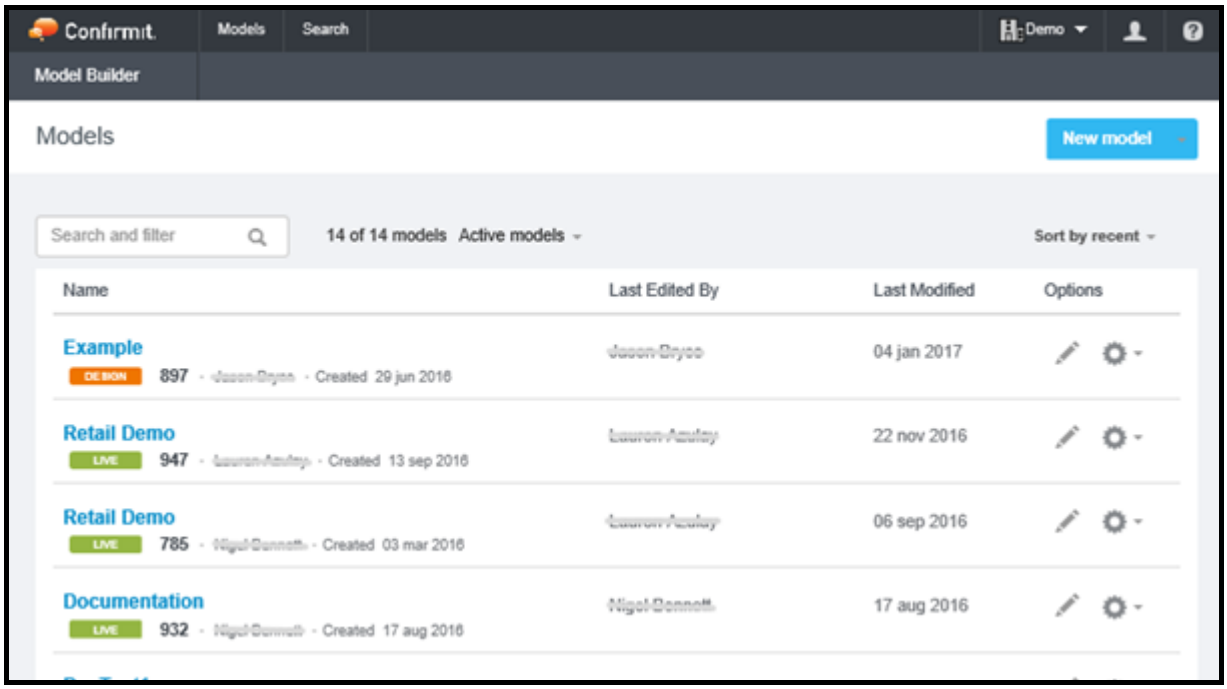

Figure 4 Example of the Models list

In the event you have many models available, you can use the Search field to find the model you wish to work with. The search looks at both the model names and numbers. You can type a text string into the Search field to filter the model list and display only those models with names that include the entered characters anywhere in the model name. Or you can search by model number. If for example you search for 15, Model Builder will return all models with a 15 anywhere in the name or in the number. Note that the search field is not case-sensitive. To re-display the full list, clear any characters from the field.

You can sort the model list by the most recent accessed, the newest (most recently created), and alphabetically by the model name. The button then indicated the sort selected. Click **Sort by...** to open the drop-down.

Click on the blue model name link to open the Design page for that model. The details displayed for each model include:

- **Model name -** this is the access link to the model. Click on the link to open the Design page for that model.

- **Model state –** the model is either in 'Design' mode or 'Live' mode. When it is in Design mode, it has not yet been launched and cannot be referenced in Horizons. Once the model has been launched, it is in Live mode and can therefore be used in Horizons.
- **Model ID -** when the model is created it is automatically given an identification number. This model ID is used when setting up Genius in Horizons.
- **Created date -** the date the model was created.
- **Last Edited By -** the user name of the person who last edited the model.
- **Last Modified -** the date the model was last modified.
- **View/Edit icon -** has the same effect as clicking the blue Model name link - see above.
- **Cog-wheel icon -** opens the **More Options** menu, enabling you to duplicate the model, archive the model, and open the list of categories.

1.3. The Menus

Model Builder holds a number of menus in a bar across the top of the screen and on the various pages in the application. These menus and tools provide access to the functionality.

1.3.1. Models

Click to return to the Models page, in which all the models to which you have access are listed.

1.3.2. Search

This opens the Search page.

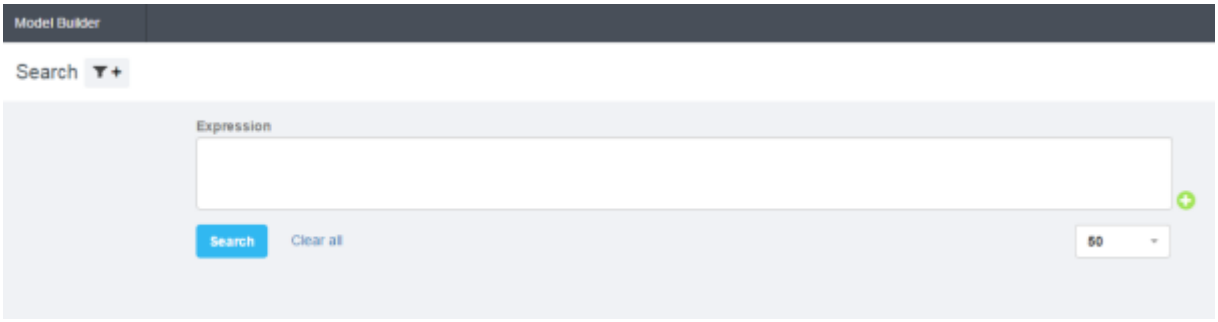

Figure 5 The Search page

Use this page to test proposed keywords on your verbatims before you add the keywords to your expressions. Adding keywords can sometimes be quite complex, in that you need to test which words and combinations of words achieve the most correct hits. There will be occasions where keywords merely return 'noise', but you will need to assess whether or not the hits-to-noise ratio is acceptable. The Search page can help you to decide this for individual words without having the results cluttered by the hits for all the other words in the expression.

Tip!
If you have a large expression, you can expand the search box by dragging the bottom right-hand corner.

Click on the green **+** icon at the side of the expression box to open up other options for testing. You have the choice of adding an AND, OR or AND NOT box. Click on **Search** to get your results (use **Ctrl-Enter** as a shortcut to clicking the **Search** button). .

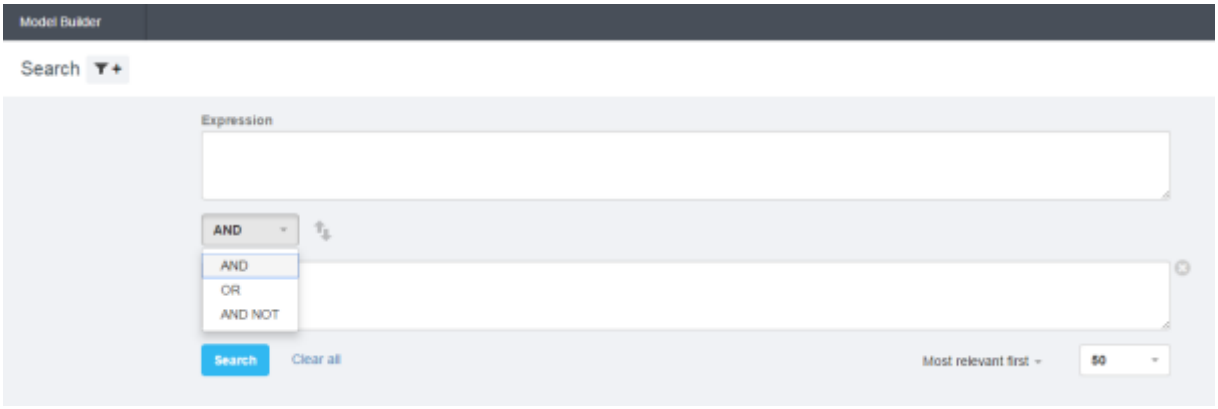

Figure 6 Adding an AND, OR or AND NOT box

Once you have a set of search results, you can choose how they are sorted. This allows you to look at different samples so you can test the accuracy of your results. Choose between 'Most relevant first', 'Least relevant first', 'Newest first' or 'Oldest first'.

Figure 7 Sorting the search results

You can also flip the 2 search boxes around by clicking on the up and down arrow icons next to the drop-down list. This will only affect the AND NOT functionality, and it gives a good opportunity to test your expressions in the opposite order without having to copy the expressions into the reverse boxes.

You can open the filter options so that you can filter by a particular survey or by time period.

1. Click the filter icon beside the Search page title to open the filter.

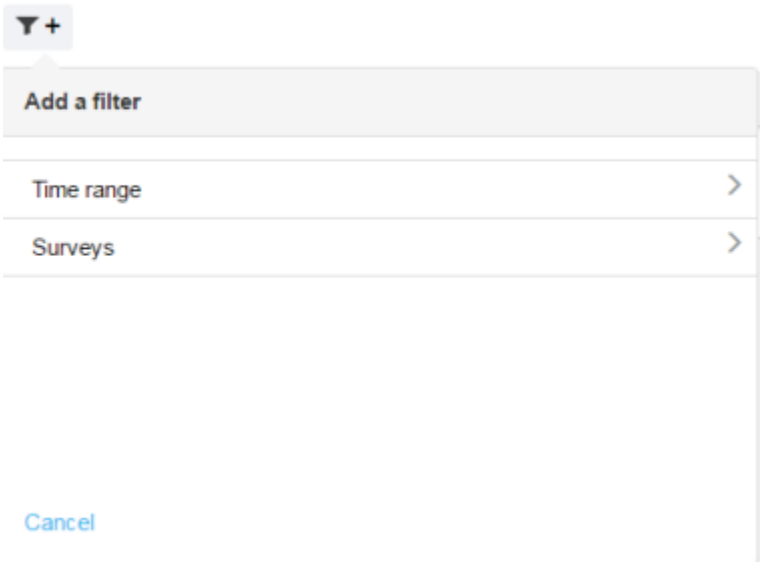

Figure 8 The filter overlay

2. Select one or more surveys to filter your results. You can also choose the question within the survey if that question has had the Text Analytics folder applied to it.

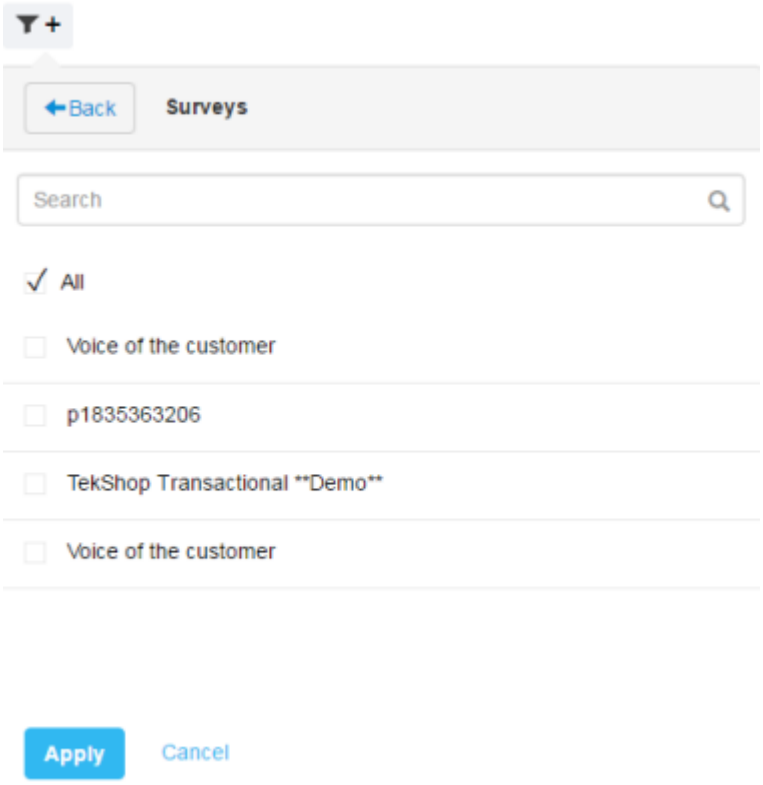

Figure 9 Selecting the surveys

3. Click **Apply** or select **Time Range**.

   Under **Time range** you can choose from a list of options (last 30 days, last 60 days, last 90 days or last 120 days) or select a custom range.

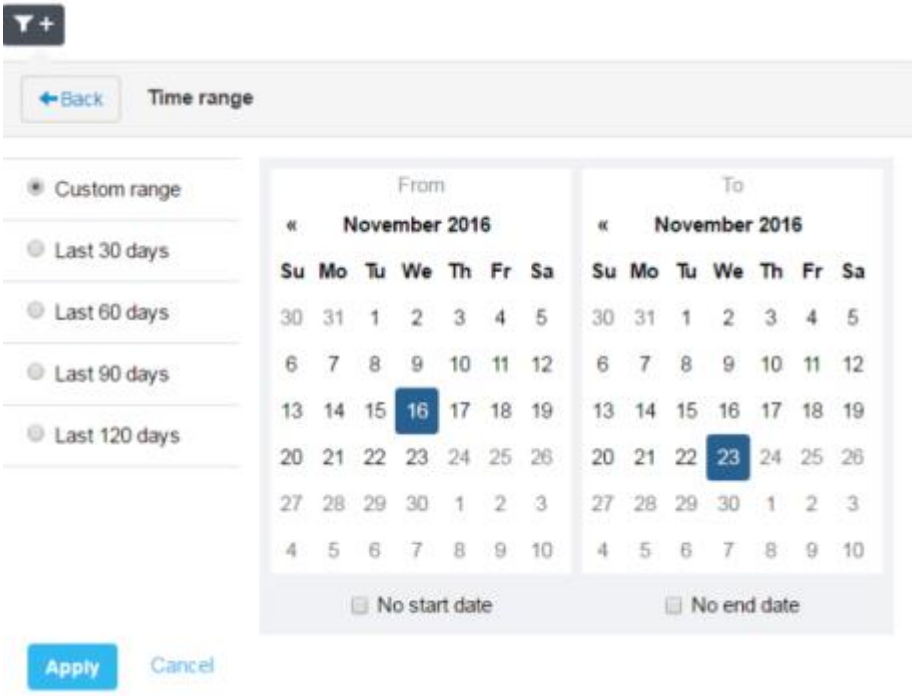

Figure 10 The Time range overlay

4. On completion, click **Apply**.

   Once you have applied the filter you can perform your search and you will only see results from the selected surveys, questions and / or date range.

1.3.3. Company Name

Your company account name is displayed next to the **User** menu icon. In the event you have administrator access to more than one customer, click the account name to open a list of the customers to which you have access, and select the customer you wish to work with. The list of models associated with that customer and to which you have access then opens.

1.3.4. The User Menu

The User menu  , located towards the right end of the main toolbar, provides access to commands applicable to you personally as the user.

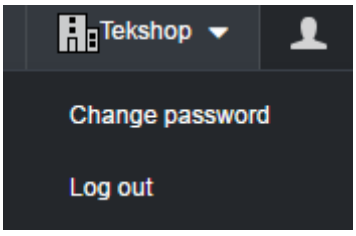

Figure 11 The User menu

- **Change password -** you can change your password from the **User** menu. To do this, click on **Change password**. You will be asked for your current password, your new password, and you will then have to repeat your new password to ensure that it is as you intended it to be.

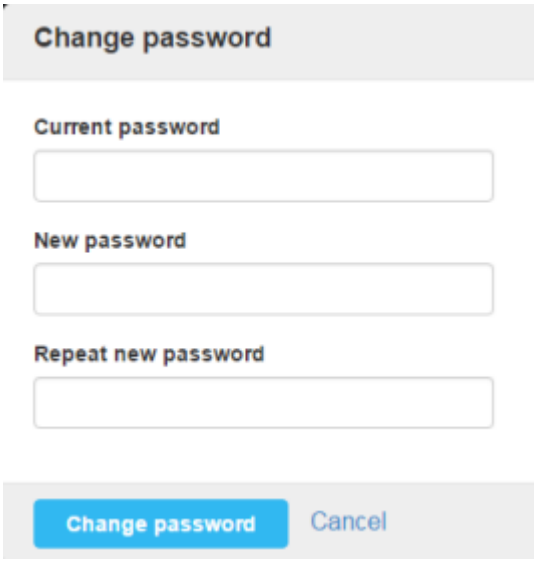

Figure 12 The Change password overlay

- **Log out -** click to log out of Model Builder and return to the login page (see Logging In on page 1 for more information). The next time you log in you will be taken to the Models list. In the event you have access to more than one company, you will be taken to the Models list for the company you were last working with when you logged out.

1.3.5. The More Options Menu

The **More options** icon , located towards the right end of each model row in the Model List, opens the More Options menu.

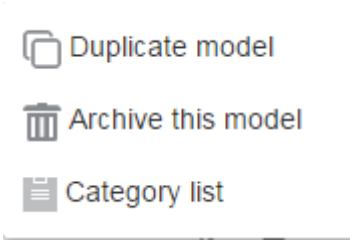

Figure 13 The More Options menu

The options in this menu are:

- **Duplicate model** - if you need to create a new model that you know will be similar to an existing model, you can create a copy of the existing model and edit the copy, thus saving much time.
- **Archive** - when you no longer need to use a model and you wish to remove it from the Model List, you can archive it (see Archiving Models on page 8 for more information).
- **Category list** - opens a dialog box that gives you access to the category list (with sub-categories and attributes) and the associated category and attribute IDs (see The Category List on page 9 for more information).

1.4. Archiving Models

To remove a model from your model list, click in the cogwheel (on the Model List page or next to the name of the model in the Design page) and select **Archive this model**; the model will be archived.

To view all archived models, click on the **Active models** link next to the search box on the Model List page, then choose **Archived models**.

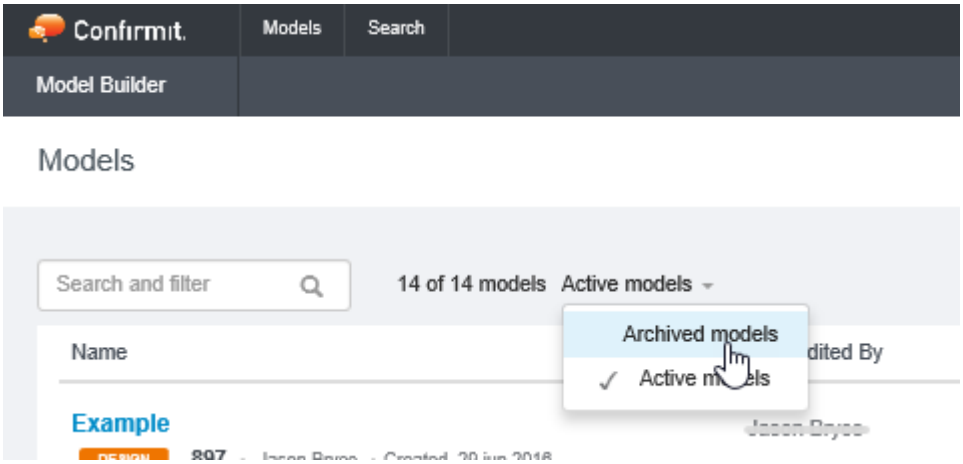

Figure 14 Selecting the Archived Models list

From the Archived Models page you can **View** the archived model or **Restore** it. Restoring an archived model puts it back into the Active Model list page, where you can continue to use and edit it.

1.5. The Category List

The **Category List** menu item is located in the **More options** menu on the Model List page and next to the model name on the Design page. This menu item opens a dialog box that lists the categories in the model (with sub-categories and attributes) and the associated category and attribute IDs. You can copy this list from the dialog box into an Excel® spreadsheet and save it as a tab delimited text file for use in the hierarchy table in Database Designer.

Category and sub-category IDs start with a ‘c’ and attribute IDs start with an ‘a’. Note that there are 3 tabs in the Category list pop up. Depending on how your project is set up and how you are reporting on your data, you will need to choose the correct list:

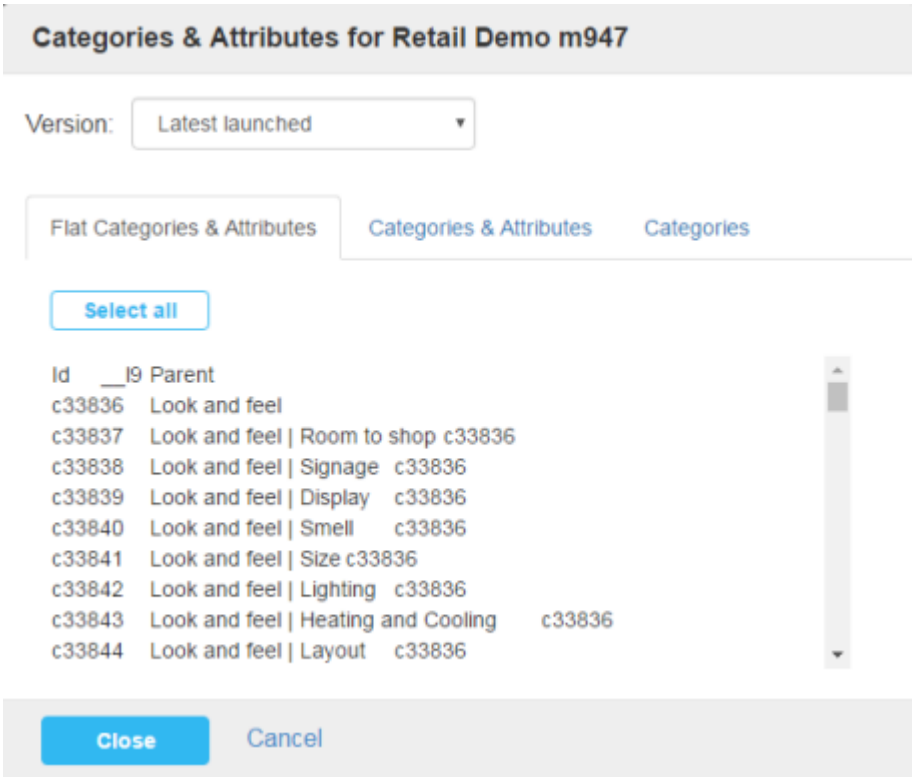

Figure 15 Example of categories and attributes for a model

- **Flat Categories and Attributes -** use this category list when you have a hierarchical model and you are going to be using the Text Analytics standard report template for your report. It has the names of the categories, sub-categories and attributes flattened out, with pipe character (|) in between. The report uses the pipe character to parse the information so that you still have the hierarchical structure where needed, but you can also have the full category / sub-category / attribute name in the charts (you need this for the context).
- **Categories and Attributes -** use this category list if you have a report is not the standard template (i.e. a bespoke report) or if your report was set up before the template was developed.
- **Categories -** this category list is a legacy option and only to be used if you have an older, flat model.

If you are at all unsure of which list to use, contact your Program Manager or Account Manager for advice.

Use the 'Latest launched' version of the Category List.

Important
Ensure you use the 'Categories and Attributes' tab to get your category and attribute IDs. The 'Categories' tab is for use in older projects that use a flat category structure and that have not yet been transferred to the hierarchical model.

1. Click **Copy all** to select the full list, then use the **Ctrl+C** keys on your keyboard to copy the list.
2. Paste your list into an Excel spreadsheet.

   The copy/paste routine will put the content into the appropriate columns, ready for loading into your Database Designer table in Horizons.
3. Save the excel spreadsheet in tab delimited text format.

   This file will be used to reference your hierarchy in Horizons in a Database Designer table. Refer to the Genius chapter of the Authoring User Guide for further information.

2. General Information

This chapter explains the basic principles of the process and defines the terminology.

2.1. Where do you Start

In the knowledge that we are seeking, the subjects are typically your customers or consumers, and the objects are typically elements of your products and services. The relationships are the customers' views and opinions; their sentiments. The first step in creating a model is to determine which objects you are interested in. Objects can be grouped into categories, and the objects within each category are arranged as a hierarchy. The depth of the hierarchy for a category depends on the level of granularity needed to determine the knowledge that you can act upon, and also on the volume of texts that will match the category.

For example, the ordering channel for a product could be categorized as "Online", "Call-center" or "Store". "Staff" could be divided into "Managers", "Associates", "Cashiers" etc. Remember that you will need to sort these keywords from your customers' verbatims, so you will need to create sub-categories based on the terminology that your customers use. The temptation might be to refine the hierarchy to many levels, but often the consumer does not easily distinguish things to the same detail so it may not be worthwhile spending time defining a complex model.

On the other hand, not refining the model sufficiently will mean losing out on valuable intelligence. For example, having a single category of "Flight" does not allow you to separate out analysis of different aspects of the flight, such as legroom, cabin service, menu choice, entertainment, etc. These are items that customers are likely to express opinions about, so it makes sense to build these hierarchy levels into the categorization model.

At the same time as you are thinking about the objects that are being discussed in the texts, and how they can be classified, you should also think about the way opinions about the objects can be categorized. For example, what opinions could be expressed about the staff in a store, or about the flight experience? How helpful were the staff in the store, or how knowledgeable were they? How comfortable were the seats, how extensive was the choice of in-flight entertainment etc? These are Attributes and they need to be carefully thought through such that they can be segmented into useful groupings that can typically be determined from analysis of textual content. While we might be interested in consumers' opinions of the change in color of our logo, they are unlikely to express an unsolicited opinion about it. If we were to ask them for an explicit opinion using an open-ended question, then we already have the context and we do not need to look for it in the text.

2.2. Asking the Right Questions

One way to approach building your categorization model and associated attributes is to think about how you would ask structured questions. With text analytics, you are effectively asking questions of the open-ended text, and the answers are in the structured output you get in the categorization model and the sentiment analysis of the opinions in the text.

Using an open-ended question instead of a series of closed questions means you are not explicitly asking about a particular aspect that you are interested in, such as the friendliness of the cabin staff. However you can imagine asking that question of the open-ended text responses and using that thought to help build your categorization model. This can also help to segment which questions you want to ask explicitly in a closed-question format, and which questions you can leave to be asked by the analysis of open-ended text responses.

In a survey, the way an open-ended question is worded impacts the answers. If you say to a customer "Tell us why you like the product", then you will get answers about aspects of the product but not necessarily opinions, such as "The price", "The design", "The range of add-on modules available", etc. This is fine as long as you have previously determined from a closed question that the customer actually likes the product, as you can then derive actionable insight from this.

If you are considering analyzing unsolicited comments, for example from social media or call-center notes, then ideally you should analyze these comments alongside your survey comments. In this case your survey question should be less loaded, for example "Please provide any further comments you have about the product". Your text analysis will then ask the questions of the solicited response text in the same way as it is asking questions of the unsolicited text from support or social media interactions.

2.3. The Elements of a Categorization Model

The Genius categorization model is built in a tree hierarchy structure. Some basic definitions:

- **Category -** an object that you have defined within the business (see Categories and Sub-Categories on page 12 for more information).

- **Sub-category -** a further refinement of the category.
- **Attributes** – The elements of opinion. These are applied to categories and sub-categories, and have their own expressions (see Attributes on page 13 for more information).

2.4. Categories and Sub-Categories

The categories and sub-categories are the main parts of the categorization tree. The categories describe the main touch points of the business, such as 'Staff', 'Store' and 'Product', and these can be broken down into sub-categories that further define these categories. For example “Staff” could be broken down into 'Management', 'Associates', 'Trainees' and 'Other Staff', while 'Product' could be broken down into the different products that you offer, or sub-categories based on groups of products.

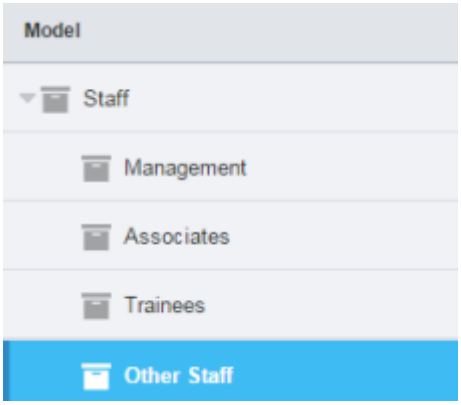

Figure 16 Example of categories and sub-categories

Expressions are added onto the child nodes, that is the sub-categories that have no further children (in this case Management, Associates, Trainees and Other Staff). This is where you add all the keywords that will add verbatims into that category. Parent nodes, that is the categories that contain leaf or child nodes (Staff in the above example), do not have their own expressions. Their expressions are made up of the expressions of the child nodes joined with an OR Boolean operator.

You can easily see whether your categories or sub-categories contain an expression by looking at the icon in front of the name. If the icon is colored, there is an expression associated with it. If the icon is clear, there is no expression.

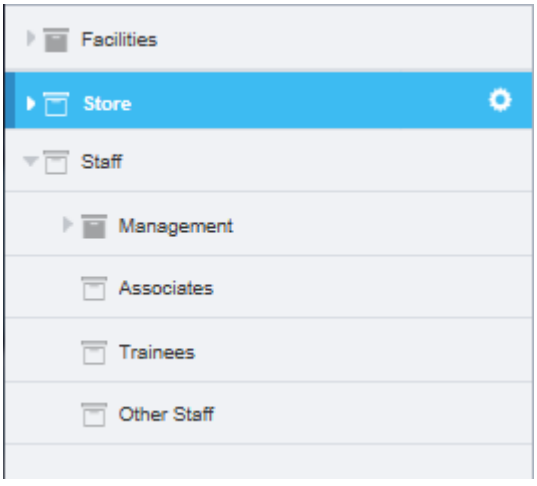

Figure 17 Example of complete and incomplete categories

In the example above, Management has an expression, but Associates, Trainees and Other Staff do not. Staff is therefore also shown as incomplete because of the incomplete nodes beneath it.

Note: If you add a sub-category to a category that already has an expression, then you will be asked if you want to move the expression to the sub-category you are creating. If you decline this invitation then the expression will be deleted from the parent category.

Warning
To avoid disappointment, save your work at regular intervals, and ensure you save before you leave Model Builder unattended (see Inactive-Screen Lock on page 3 for more information).

2.5. Attributes

Attributes are the elements of the categorization model that you wish to measure; your customers' opinions. These could be for example Product knowledge, Friendliness, Speed of service of the Staff, Location, Opening hours, Cleanliness of the Store, Availability, Reliability, Variety of the Product etc.

These attributes can only be applied to the lowest node of the tree. Attributes are joined to their category / sub-category nodes with an AND Boolean operator. So the expression that Genius will use for determining if text is referencing the friendliness of management could be:

```
(manager OR managers OR management OR "mgr" OR supervisor OR "in
charge") AND (friendly OR friendliness OR "not friendly"~3)
```

By analyzing only those verbatim texts that match this expression, Genius can then use its sentiment algorithm to determine the opinion expressed by the customers regarding the friendliness of the manager.

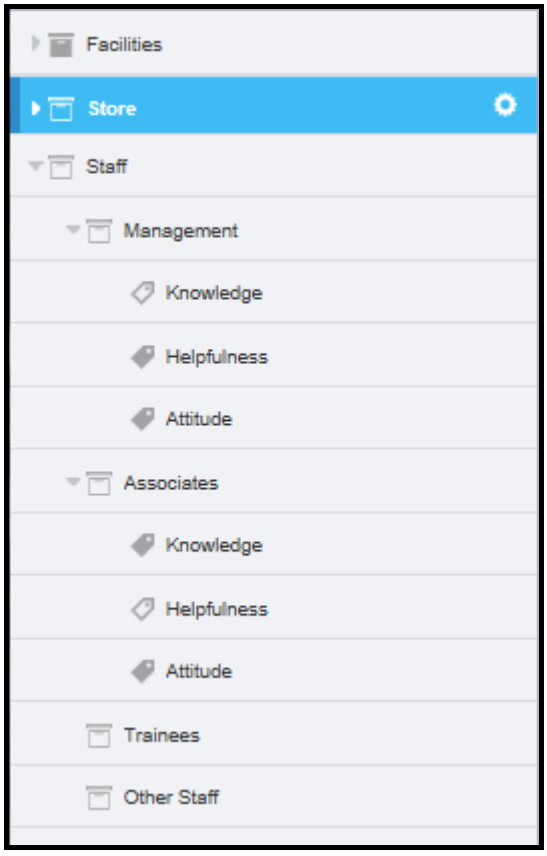

Figure 18 Example of categories, sub-categories and attributes

The example above shows the attributes for Management and Associates. Some of the attributes do not yet have expressions, so the icons are shown as empty. The "incomplete" status for the icons is passed back up to the parent node (Staff) so that you can see that there is an empty attribute even when the hierarchy is closed.

2.6. The Uncaptured Attribute

When setting up the model, you have the option to automatically generate attributes for uncaptured verbatims. This option is also available on the Model Overview page (see The Model Overview Tab on page 20 for more information). If you select this option, once you have at least one attribute added to the category tree an additional attribute **Uncaptured** appears in the tree. This automatically shows you all of the verbatims that have been captured by the sub-category expression ("Sales" in the example illustration) and excludes the attributes that have already been set up ("Helpfulness" and "Knowledge" in the example illustration). The Uncaptured attribute lists examples of other keywords that can be added to existing attribute expressions; it suggests new attributes based on what respondents are talking about regarding the function.

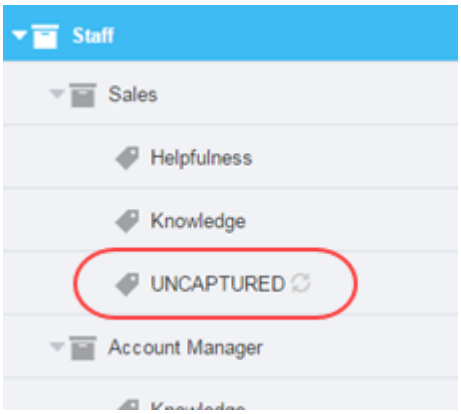

Figure 19 The Uncaptured attribute

The Topic Discovery words are useful in helping to understand whether you are missing keywords for your current attributes or whether you need to define some more attributes based on the frequency of the topics in the Topic Discovery word list (see Topic Discovery on page 30 for more information).

2.7. The Logical Operators

You can use logical operators in the search field when you write search expressions. The logical operators combine search words in your expression.

Logical AND

The AND operator will combine two (or more) search words. For example if you want to search for posts containing both of the car manufacturers Audi and Mercedes, your search expression will be (audi AND mercedes). Only posts where both words occur will be returned in the search result. Note that the order in which the words appear, and their distance apart, will not be taken into consideration.

Logical OR

The OR operator will return posts where any of the given search words are represented. For example if you want to search for posts that mention either of the phone companies O2 or Vodafone, your search expression will be (o2 OR vodafone). Any posts containing either O2 or Vodafone will then be returned.

Logical AND NOT

The AND NOT operator is used to exclude posts that contain the specified search words. For example if you want to search for all posts that mention iPhone, but exclude posts that mention Samsung, your search expression would be (iphone AND NOT samsung).

Note: Logical operators are always written in UPPER CASE. To make your search expressions easier to read we suggest that you always write your search words in lower case. For example both "AUDI AND MERCEDES" and "audi AND mercedes" will return the same results, but the latter is easier to read.

Proximity operator

The proximity operator (~) is useful when you are searching for two words in a verbatim, but you are not certain about the order in which they might appear or how far apart they are. In a long verbatim the two words may appear but could also be unrelated. Use the proximity operator therefore to specify a maximum number of words between your two search words. For example if you want to find the posts where car manufacturers Audi and Mercedes are close to each other, you would use the proximity operator. Your search expression could then be (“audi mercedes”~2). In this example the expression will return only those posts where there are a maximum of two words between the two names.

Special characters

Some characters have special functions (operators) and cannot be searched-for like normal characters. However they can be used inside quotation marks "", or they can be escaped. They are escaped if they have backslash (\) before them or quotation marks around them ("").

Example: \+ OR "+"

	Special characters	Special characters
+	)	“
-	{	~
&&	}	*
||	[	?
!	]	:
(	^	\

Parenthesis

Parenthesis can make or break your expressions. They are very useful for simplifying and making the expressions readable, but incorrect usage will give incorrect hits.

Ensure that multiple words have quotation marks around them - "multiple words".

Wildcards

The * wildcard ensures that all words that start or end with the preceding or following text characters will be captured. For example, categor* will find words such as category, categories, categorisation, categorization, , and *common will find words such as common and uncommon, etc. Note that you cannot have a * within quotation marks; for example "all categor*" will not work!

The ? can be used to wildcard a single character in a word. For example both *Categorization* in US English and *Categorisation* in UK English could be captured by using the keyword (categori?ation).

Regular expressions

Regular expressions can be used as part of the expressions. Model Builder supports regular expressions through Lucene (currently Lucene 6.2.1.).

Go to the links below:

```
https://lucene.apache.org/core/6_2_1/core/org/apache/lucene/util/automaton/RegExp.html
```

or

```
https://www.elastic.co/guide/en/elasticsearch/reference/current/query-dsl-regexp-query.html#regexp-syntax
```

for help on writing these expressions.

Contact your Account Manager if you want to learn more from our Analytics team.

3. Importing a Model

You can import a new model into Confirmit from a specially formatted Excel file. To find the Import functionality, click on the drop-down on the **Create model** button. You will need to follow the wizard in order to be able to import your model structure into Model Builder.

3.1. Download the Excel Template

To download an Excel template:

1. In Model Builder, click on the drop-down on the **Create model** button.

   The Create New - import overlay opens.

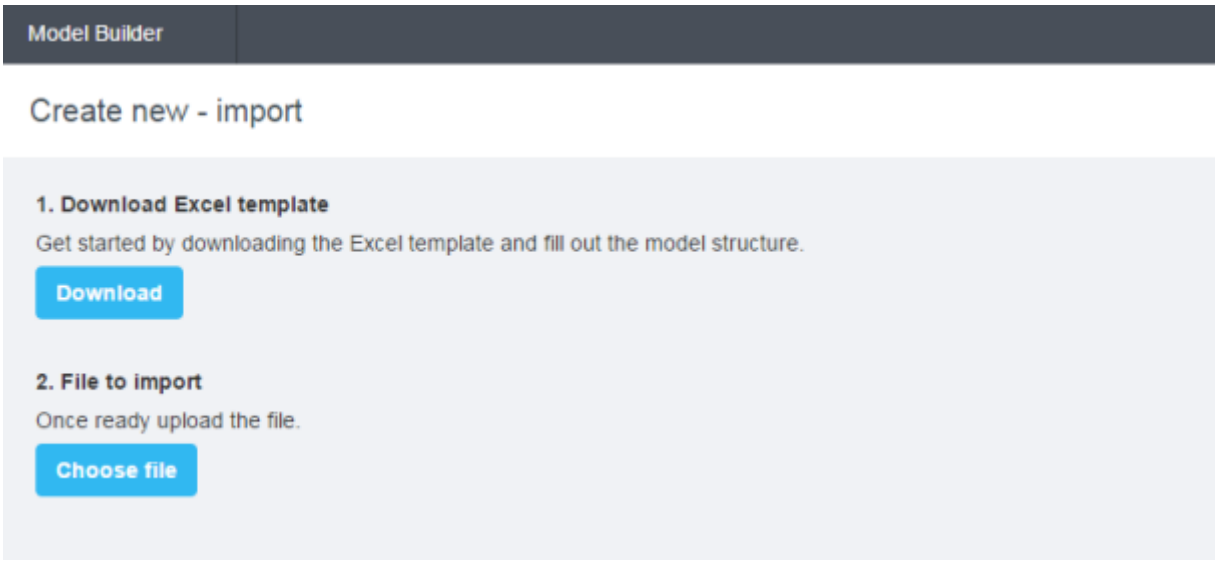

Figure 20 The Create New - import overlay

2. Click **Download**.

   The Excel template downloads.

3. Open the template to see an example of how the file must be formatted for your categorization model.

   You can use the downloaded template as the basis of your file.

Techshop					
Category	Sub-category	Attribute	Description (will not be imported)	Example keywords (will not be imported)	Expression
Staff					
	Sales				sales OR "account manager"
		Knowledge			knowledge OR detail
		Professionalism			professional*
		Communication			email OR communication OR language
	Services				services OR tam OR tss OR "technical services"
		Speed of service			speed
		Knowledge			knowledge OR detail
		Professionalism			professional*
	Technical Support				"technical support" OR "tech support" OR support
		Speed of service			speed
		Knowledge			knowledge OR detail
		Professionalism			professional*

Figure 21 Example of an Excel template

Add your model definition to the spreadsheet in the required format (don't forget to add a name in the first column and row). Ensure you follow the format, with category, sub-category and attributes on different rows. You can also add in an expression, which will be imported into the model. Do not leave any line spaces between the Category header and where you start adding category names.

4. Save your file as an .xlsx file.

3.2. Importing your File

1. On the Create new - import overlay, click **Choose file**.
2. Find and select the .xlsx file that you have previously saved

   The selected file is indicated in the overlay.

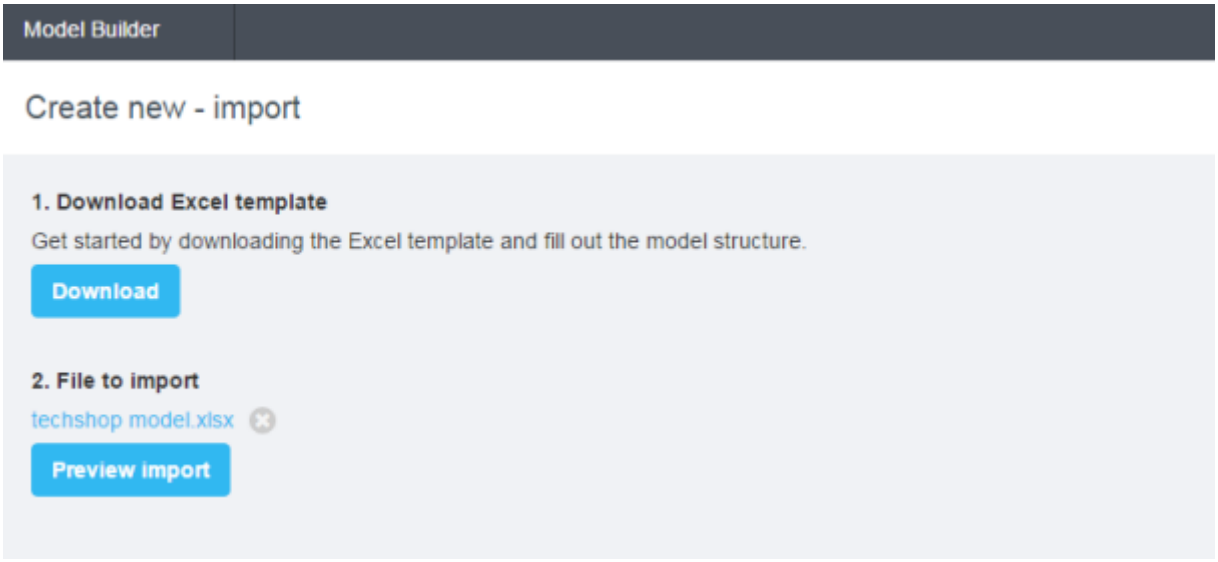

Figure 22 Selecting the file to import

3. Click **Preview import**.

   The preview page opens.

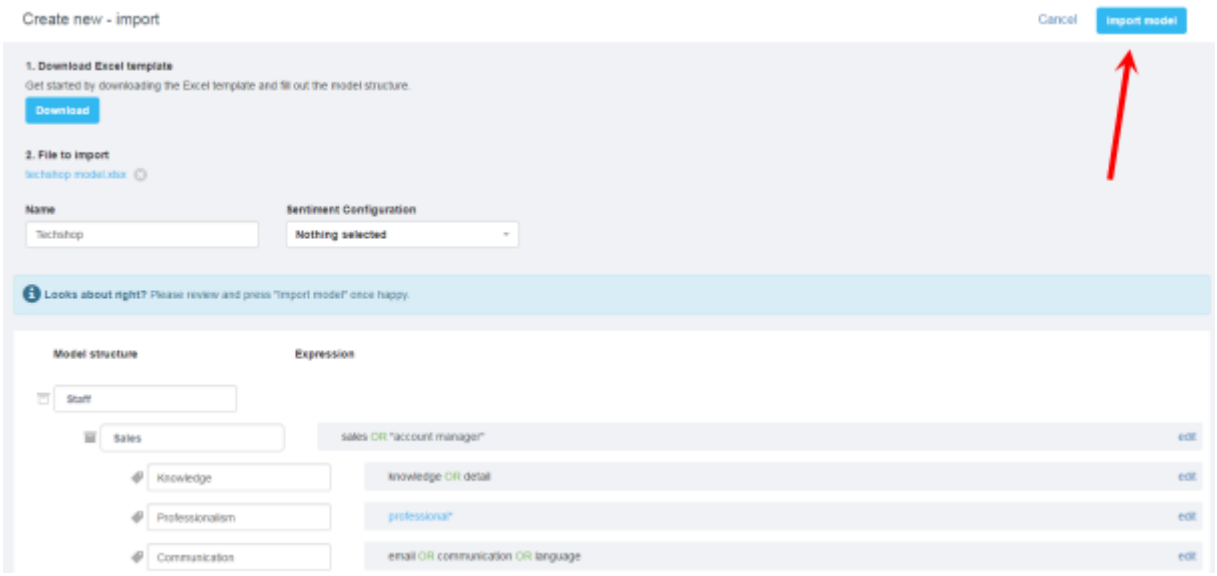

Figure 23 Previewing the import file

4. Review the file. If it is correct, click **Import model** (arrowed).

   The model is added to Model Builder. Don't forget to launch the model once you are ready to use it in Horizons.

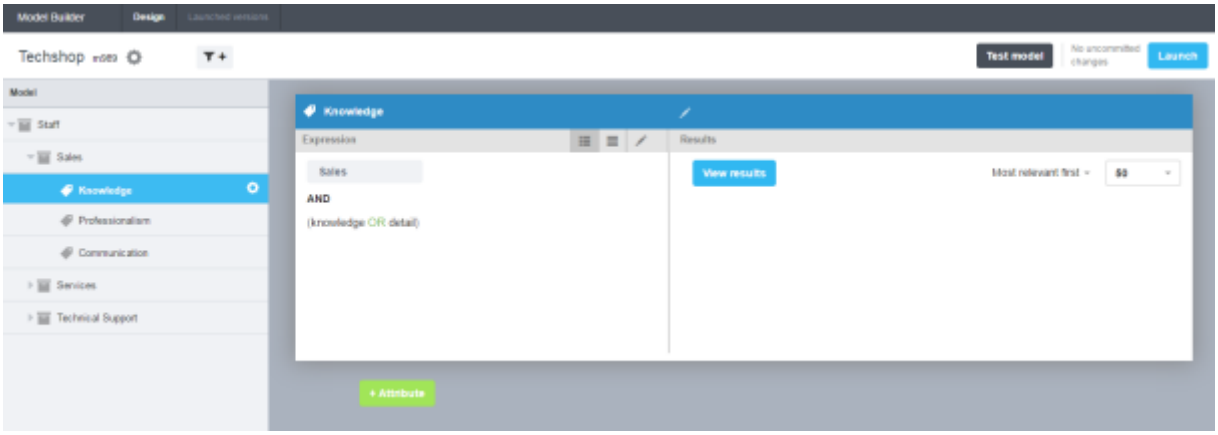

Figure 24 The file in Model Builder

4. Creating a New Model

A categorization model must be built around how your customers are speaking about your business; borrowing a model from another business will not work. So before you start to build the categorization structure you will need to have a fairly detailed understanding of what your customers are saying about you. Based on this you can then identify the major themes or categories that define your business structure, and decide where Text Analytics will provide value by reporting sentiment about each category.

When creating a new model it is best to start simple and to build up more complex categories as you explore your verbatim data. Having identified the main categories, you must then identify the sub-categories. Here you should start with simple sub-categories; you can always create further sub-categories as you find them.

You can then create the expression for each sub-category.

Identify what customers are saying about these categories that you want to be able to measure; these will be your Attributes that the analysis will look for.

Create the expressions for the Attributes.

Warning
To avoid disappointment, save your work at regular intervals and ensure you save before you leave Model Builder unattended (see Inactive-Screen Lock on page 3 for more information).

4.1. How to Create a New Model

1. In the Models page, click **New Model**.

   The Create Model overlay opens.

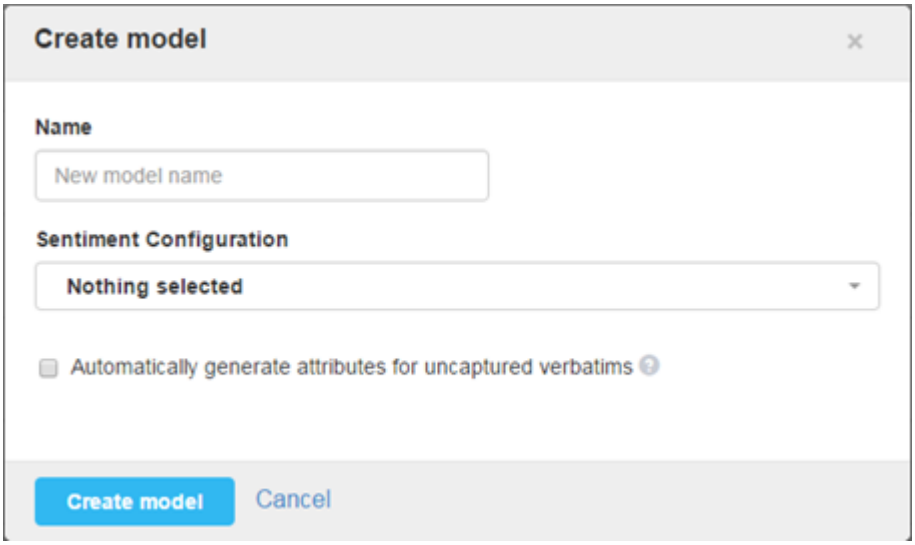

Figure 25 The Create Model overlay

2. Type in a name for your new model.
3. Select the sentiment configuration you want to use.

   The options available to you are in the drop-down list. You will either have the choice of the General English sentiment configuration or perhaps a configuration that has been adjusted for your business. If in doubt, use the General English configuration and speak to your Account Manager if you feel the sentiment needs some adjustment.

4. Decide whether you want to automatically generate attributes for uncaptured verbatims for your model.

   This helps to find other attributes mentioned in the comments for a category or sub-category (see The Uncaptured Attribute on page 14 for more information).

5. Click **Create model**.

   The model is created and given a model ID number, and the Add Category overlay opens.

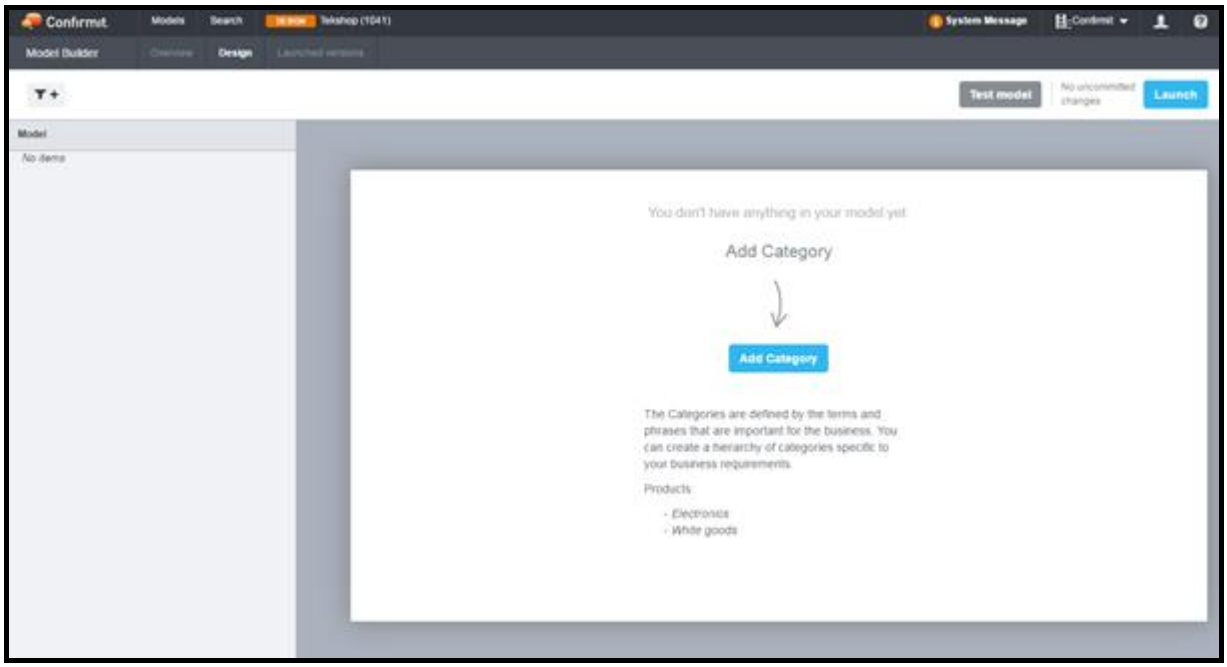

Figure 26 The Add Category overlay

6. Click **Add Category**.

   The Create a new Category overlay appears.

Figure 27 The Create a new Category overlay

7. Type the category name into the field, then click **Create**.

   The category pane opens. You can now create your expression for this category or choose to create a sub-category if you want to refine your model further. If this category is to be a leaf node - the last level in the category hierarchy before you add attributes - then enter the required expression in this overlay.

Warning
To avoid disappointment, save your work at regular intervals and ensure you save before you leave Model Builder unattended (see Inactive-Screen Lock on page 3 for more information).

4.2. The Model Overview Tab

The Overview tab contains all the background information about your model.

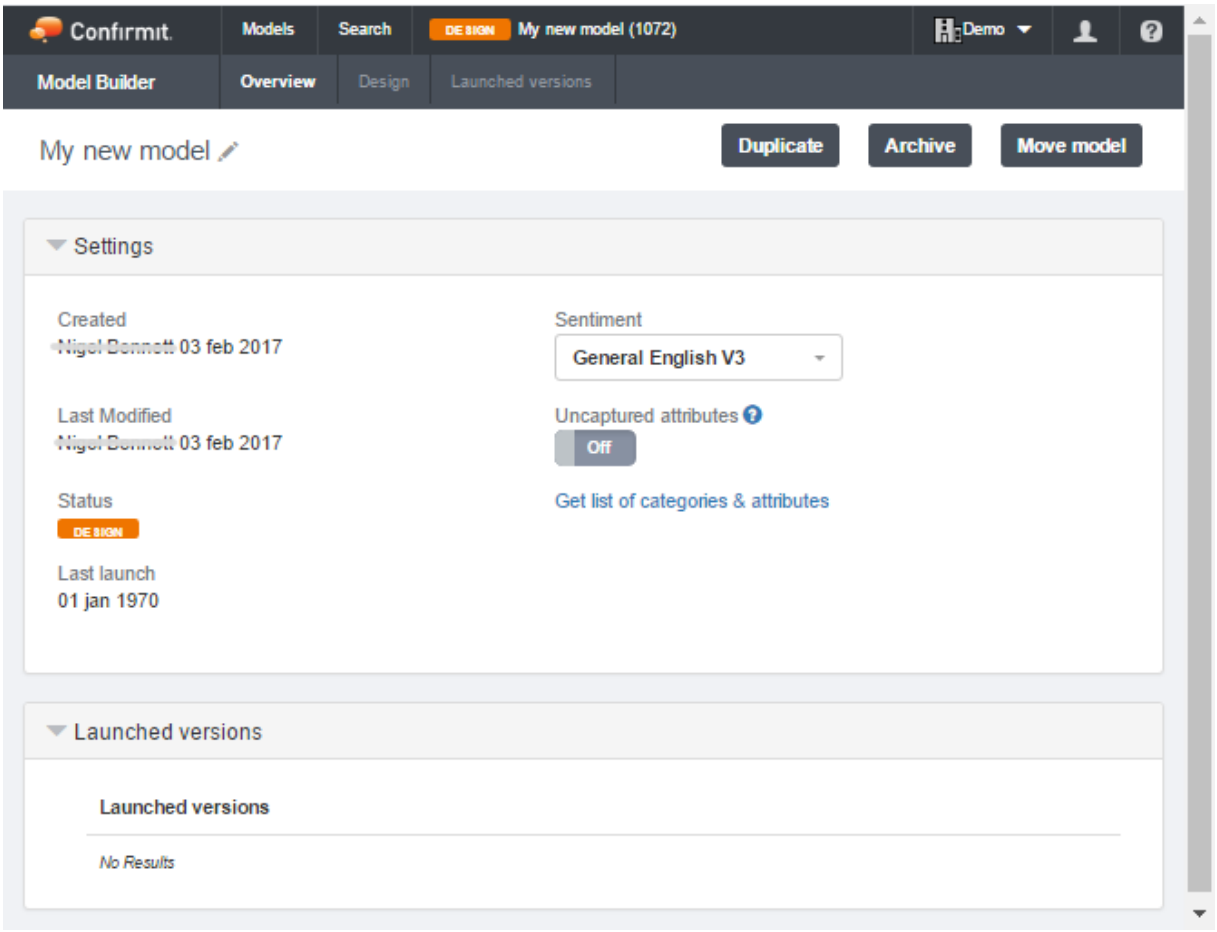

Figure 28 Example of the Model Overview tab

Use this tab to:

- Edit your model name (click on the edit icon next to the model name).
- See who created and who last modified your model, with the dates.
- **Status –** shows an orange **Design** button if the model has not yet been saved and marked as live (this also means that it cannot be used in Horizons) and a green **Live** button if the model has been saved as a version. The first time it is saved as a version, it is set to Live. Note that you can only use a model in Horizons if it has the green Live button.
- **Last launch –** the date that this model was last saved as a version.
- **Sentiment –** you can select the sentiment configuration from the drop down menu. If you do not know which configuration to use, speak to your Account Manager.
- **Uncaptured attributes –** an on/off switch. If set to **On**, uncaptured attributes are generated automatically in your model. If set to off, the Uncaptured attributes folder is not generated.
- **Get list of categories and attributes list –** displays the list of categories, sub-categories, attributes and their associated IDs, along with the parent hierarchy reference, which you need to copy and paste into a Horizons database designer table.
- **Launched versions -** a list of saved versions of the model, along with any notes that you have written when saving the version. Click on the name and it will take you into that saved version, so that you can test the model and view the results of the expressions. From this screen you can duplicate this version of the model, which copies that version and gives it a new model ID.

The buttons towards the upper-right corner of the screen are:

- **Duplicate –** makes a copy of the model and gives the duplicate a new model ID. You are redirected to the new model immediately, and the new model is called ‘Copy of ...’

- **Archive –** archiving a model takes it out of your Model list and puts it into your Archived models list. To view the Archived models list, click on the Active models drop-down on the Model List page (see Archiving Models on page 8 for more information). From the Archived models list you can restore the model back into your Active models list.
- **Move Model –** if you have access to more than one account you can move a model from one account to the other. Moving the model removes it from the originating account and puts it into the destination account.

4.3. The Design Tab

The Design tab is where you make any changes to the hierarchy and expressions in the model.

4.3.1. Creating the Categorization Hierarchy

You can now start to create your categorization tree. From your first category, you have 3 options:

1. You can click on the **+ Category** button at the bottom of the expression pane to add another category.

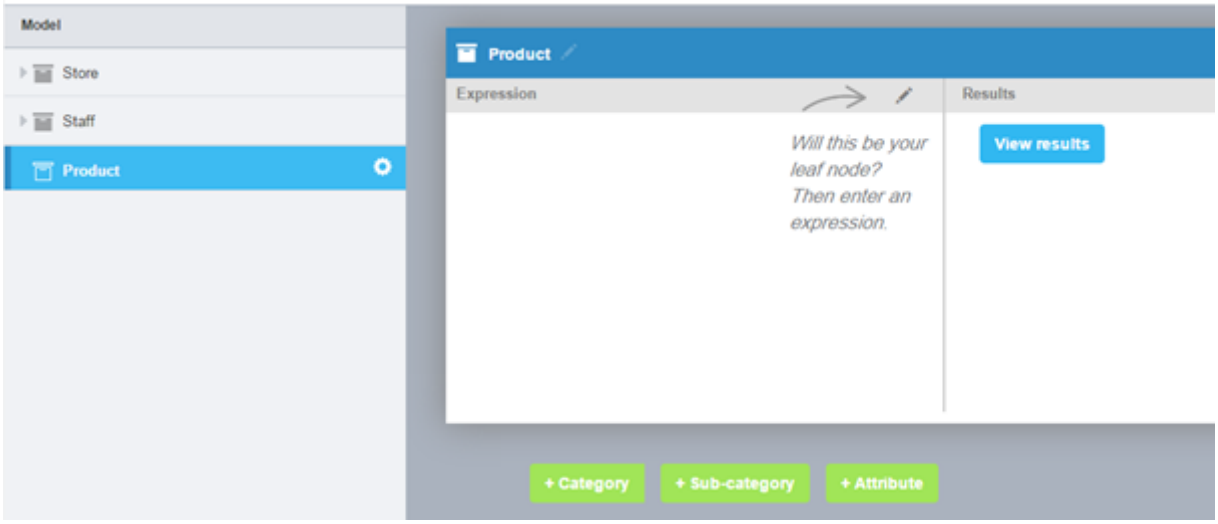

Figure 29 To add a Category

2. Click **+ Sub-category** and **+ Attribute** to add to add a Sub-category and an Attribute respectively.

Build up your categorization tree based on how your model needs to look, with Categories and Sub-categories describing the touch points in the business and Attributes describing the emotions or measures of those touch points.

4.3.2. How to Edit a Category Name

To edit the names of the categories, sub-categories or attributes:

1. Click on the **Edit** icon in the blue bar.

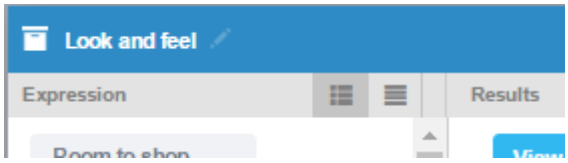

Figure 30 Click the Edit icon in the blue bar

The text field opens.

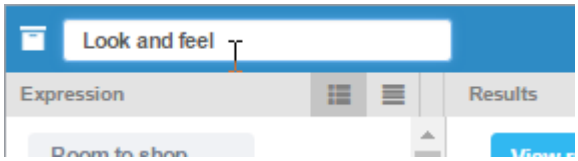

Figure 31 The text field opens

2. Edit the name as required, then press **Return** or click out of the field to close it.

4.4. Writing an Expression

Your expressions are made up of the keywords that describe your category, sub-category or attribute. Click on the **Edit** icon to start writing your expression.

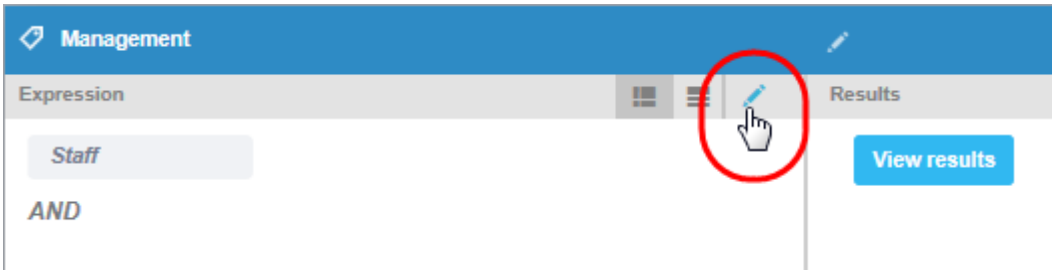

Figure 32 Starting the expression

You will need to build up the expression by testing the keywords against your verbatims from the survey to ensure that they are collecting the correct verbatims for that category. When you start out, you can test keywords in the expression pane by clicking on **View results**. If you find you are getting verbatims that shouldn't be in that category then you will need to refine your expression by using more targeted keywords or proximity indicators and AND NOT operators. You will also need to add in misspellings to ensure that misspelled words and words that are written in text speak are included in the analysis.

Once you have a longer expression, the process can be simplified by using the Search page (see Search on page 4 for more information). Here you can test keywords, test the proximity of the words to one another, and test the effect of using an AND NOT on your expression. It is recommended to go through a good set of results in order to validate that the expression is bringing through a good percentage of accurate hits. You should aim to reduce noise (unwanted hits) as much as possible, but this is not always 100% possible. You will need to decide how much noise is acceptable for each category.

When writing expressions in Model Builder, you should save the expression at regular intervals and between adding new keywords as any changes will be lost if you leave Model Builder for 30 minutes with no interaction. This is especially valid if using a separate tab with your Search page opened.

You can use the different AND and OR boxes to separate parts of your expression and to format it for easier reading. You might like to keep certain concepts or words and misspellings together so that it is easier to read and edit afterwards.

AND boxes will always be associated with an OR box. This means that at least one of the keywords in the AND box will need to be present in the verbatim as well as at least one of the keywords from the OR box. You can have a separate OR box that will not have the AND box association, if you click on the button for a different OR box. The gray line on the left side shows you the association.

The global AND NOT box towards the bottom of the pane is used when you have some keywords that need to be excluded from the entire expression. You can also add local AND NOTs in the other OR and AND boxes.

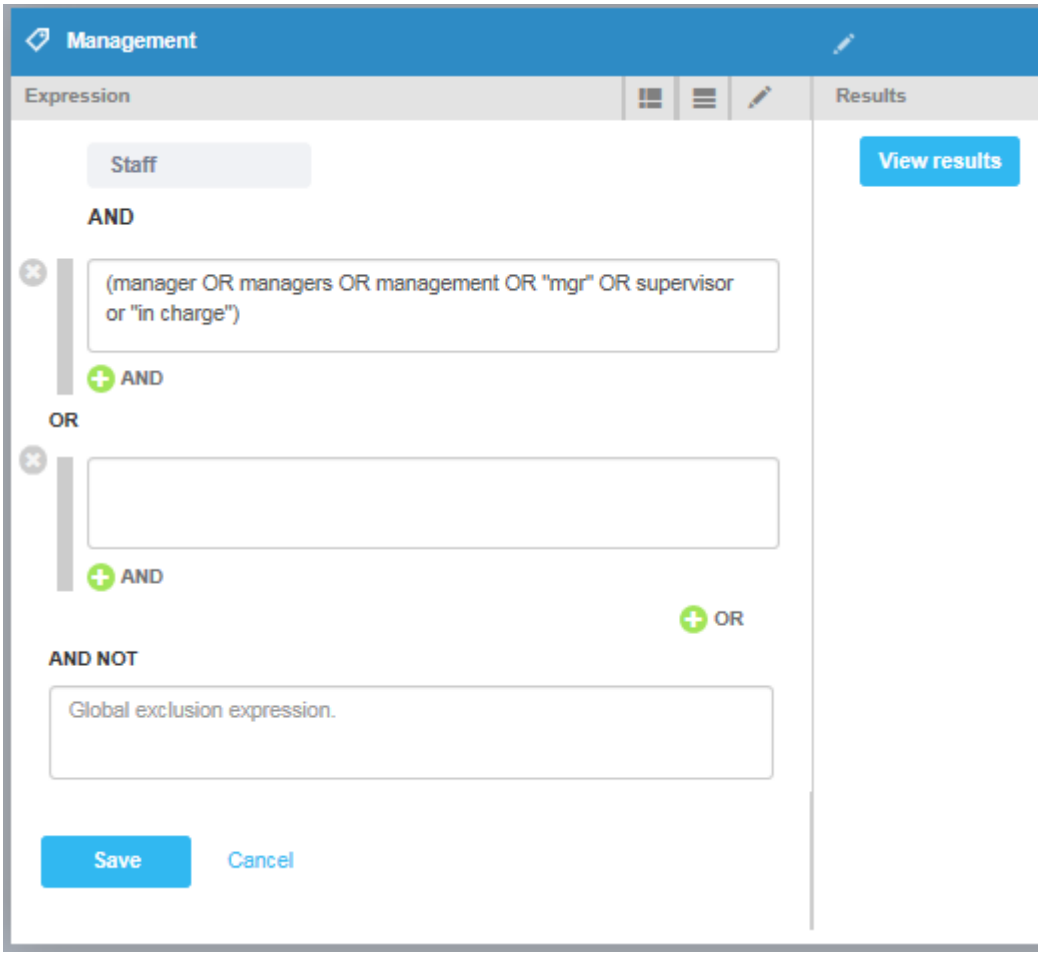

Figure 33 The AND, OR and AND NOT fields

Warning
Save your work at regular intervals, and ensure you save before you leave Model Builder unattended.

4.4.1. Adding Comments to Expressions

You can add comments into an expression to provide information for future reference. These comments will not be included in the analysis.

To add a comment:

Type the characters ***/**** at the beginning of the comment text, and the characters ****/*** to close the comment.

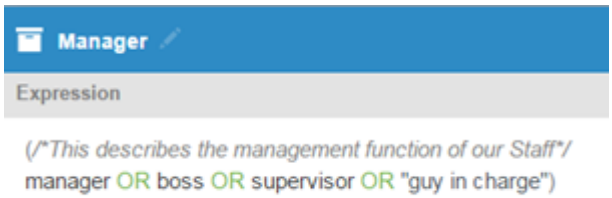

Figure 34 Example of a expression comment

4.4.2. Error Messages

As you type your expression, it is validated in the background. In the event you make a mistake and your expression becomes invalid, if you are still in the expression box the sides will go red. If you leave the expression box and the expression is not valid, an error message will be displayed and you will not be able to save your expression until the errors have been rectified. You will need to ensure that the parentheses match up, that all the Boolean operators are in capitals and that the required formatting is correct. Checking your expression in an application such as Notepad++ can help to verify the parentheses are matched.

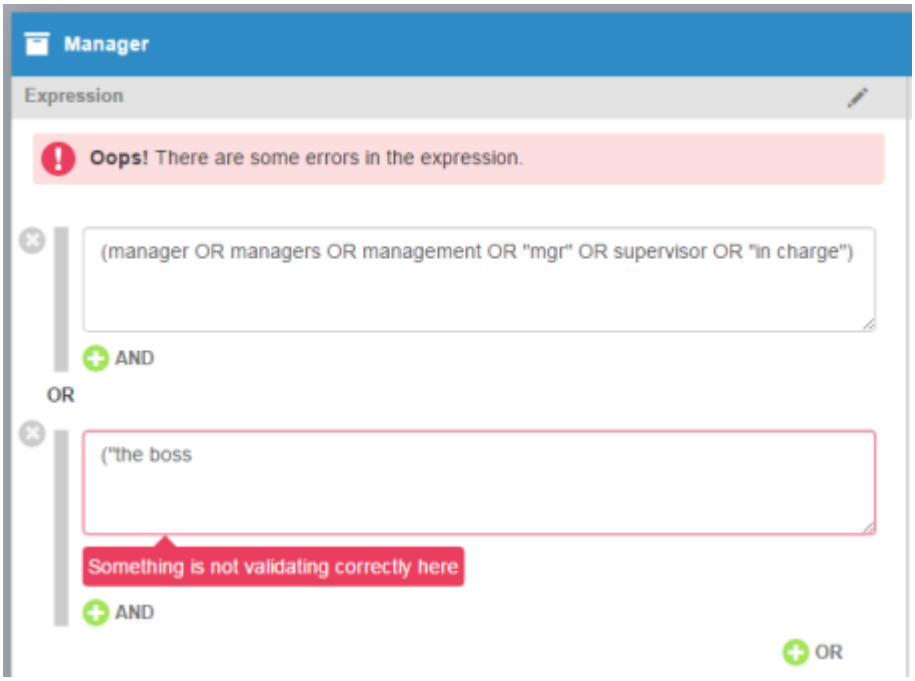

Figure 35 Example of an error message

4.4.3. The Survey Filter

If you have Genius Text Analytics on more than one survey, the results that you see when clicking on 'View results' are from all of the survey variables that have the Genius folder attached for analysis. To view results from a reduced set of surveys, use the survey filter.

1. Click the **Filter** button to open a list of the survey names and related Project IDs that have had Genius Text Analytics added to the open text variables.

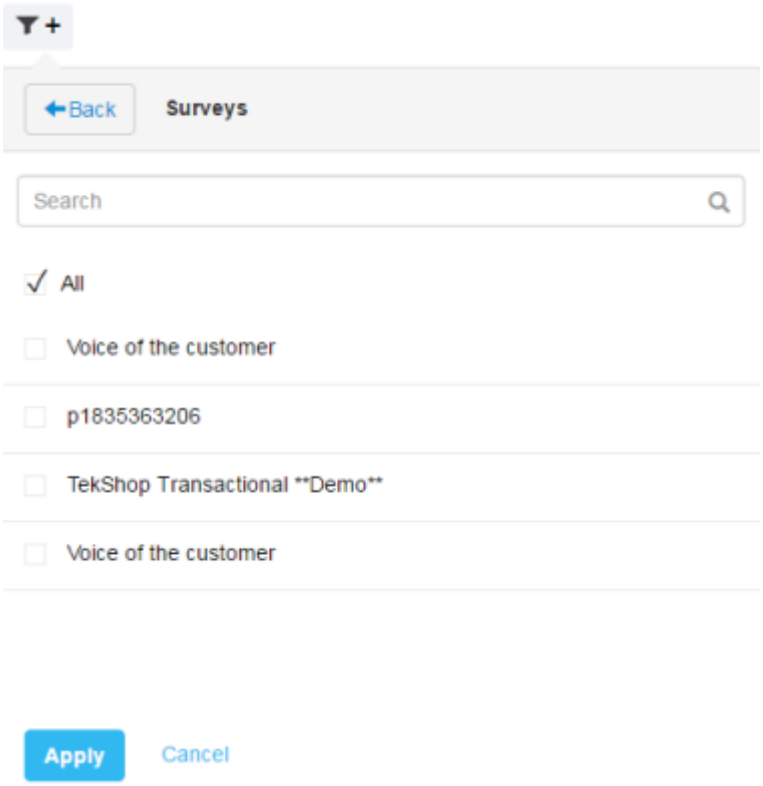

Figure 36 Example of the survey filter

2. Select the survey(s) you wish to use and click **Apply** or select **Time Range**.

   Under **Time range** you can choose from a list of options (last 30 days, last 60 days, last 90 days or last 120 days) or select a custom range.

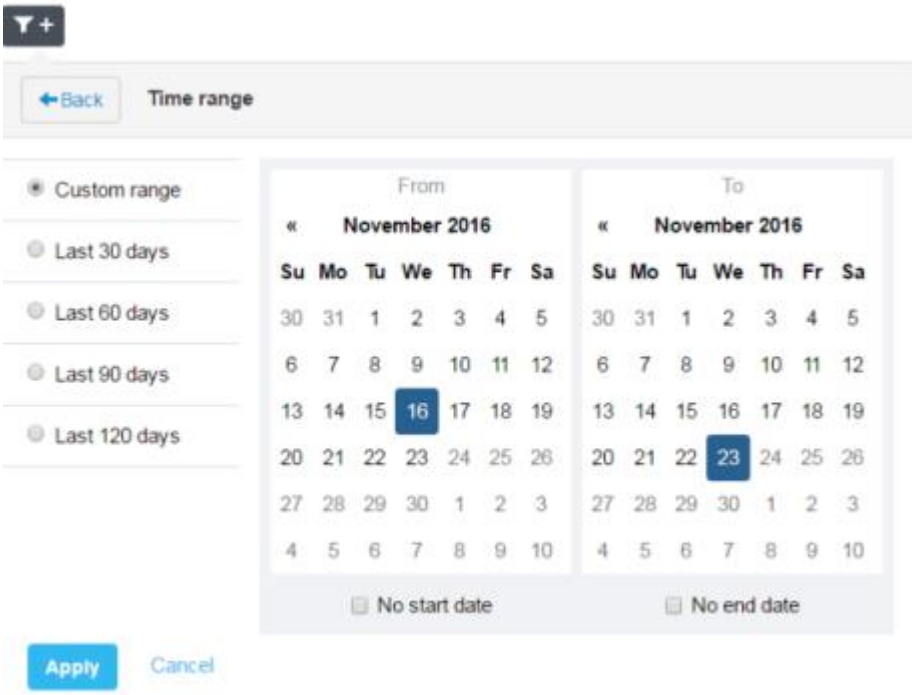

Figure 37 The Time range overlay

3. On completion, click **Apply**.

   Once you have applied the filter you can perform your search and you will only see results from the selected surveys, questions and / or date range.

   The filter that has been applied is indicated next to the filter button. This will remain visible for the duration of the session or until you cancel the filter by clicking on the (**X**) to remove it.

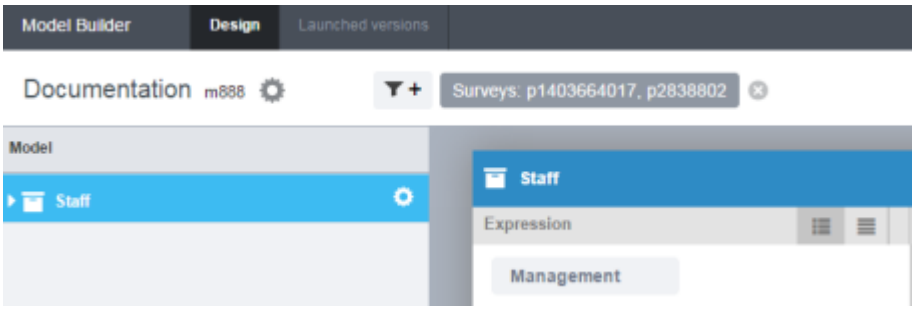

Figure 38 A filter applied

Warning
Save your work at regular intervals and ensure you save before you leave Model Builder unattended (see Inactive-Screen Lock on page 3 for more information).

4.5. How to Edit an Existing Expression

To edit an existing expression:

1. Click on the **Edit** icon in the expression pane.

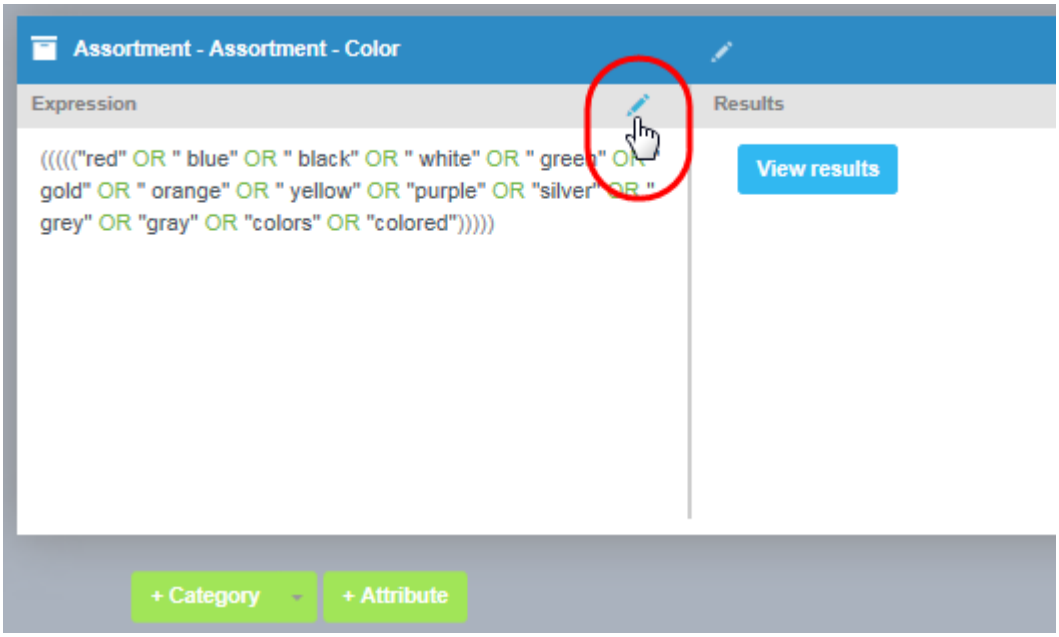

Figure 39 Click the Edit icon

This opens the Edit Expression pane for editing.

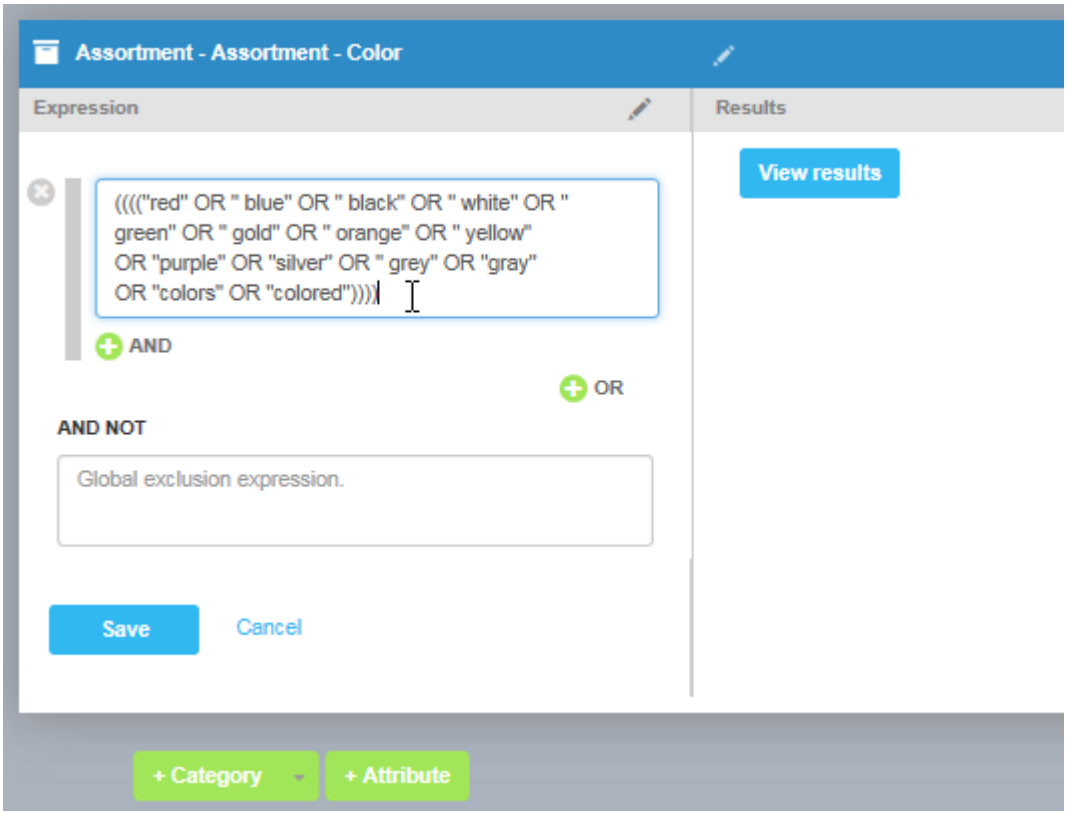

Figure 40 The Edit Expression pane opens

You can now click into the expression and add and take away keywords and add additional operators as required.

4.6. How to Test an Expression

To test whether your expression is working well, click **View results** in the right pane opposite your expression.

You will see that the keywords from the expression have been highlighted in yellow. This will help you to assess why the verbatim was captured and whether it is correct or not for the category, sub-category or attribute that you defined.

Once you have clicked **View results**, you also have the option to filter your result set by using the search bar at the top of the right pane.

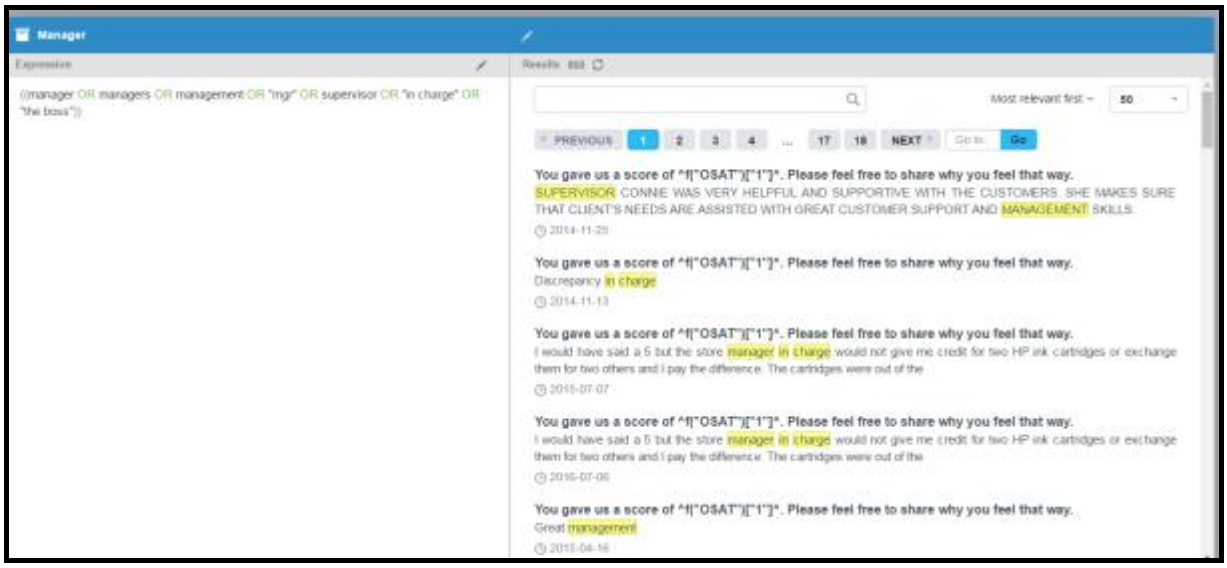

Figure 41 Testing an expression

If you wish to update your search results, click the **Refresh** icon at the top of the results page.

You have the option to view your results by 'Most relevant first', 'Least relevant first', 'Newest first' and 'Oldest first' so that you can test a good sample of the results.

4.7. Testing the Model

You can test the model against your verbatim set at any time. To run a test, click the **Test Model** button towards the upper right corner of the Model Builder screen.

This will run all the available verbatims through the entire model, and once it has completed running it will show as the number of hits against each category.

A new folder will also be placed at the bottom of your categorization tree, called 'Uncategorized comments'. This folder contains all the verbatims that have not been categorized by the rest of the model. You can go into this folder to look through the results to see whether there are any misspellings or keywords that need to be added to the existing model or whether you need to add any further categories, sub-categories or attributes.

4.8. Uncategorized Comments

The Uncategorized Comments folder appears after you have tested your model against all the verbatims from the surveys that are currently saved in Model Builder, or the filtered set of verbatims if you've selected to filter by a specific survey or question.

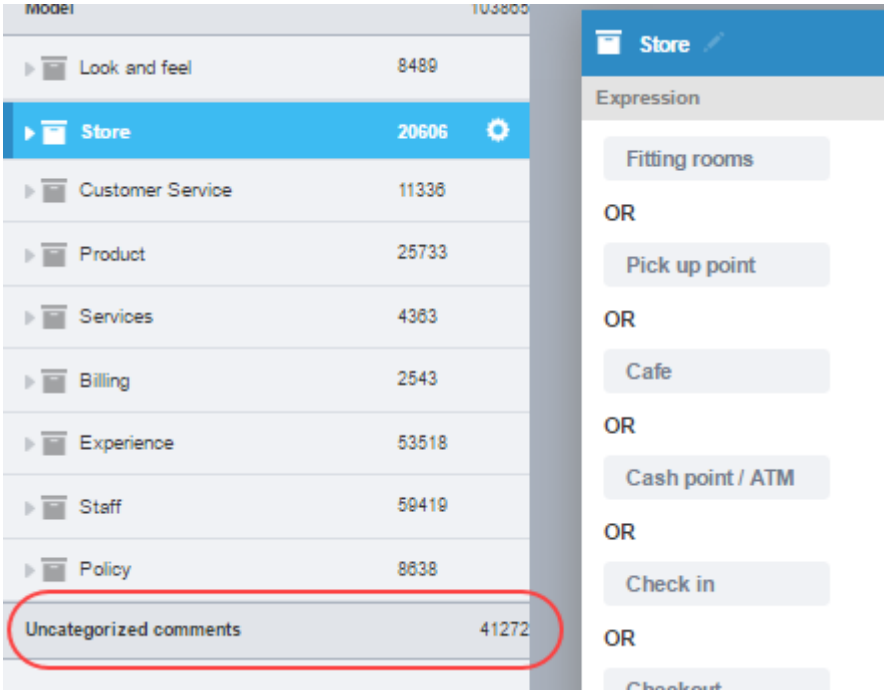

Figure 42 The Uncategorized Comments folder

The folder contains all comments that have not been captured by the existing model. You can explore these comments to find new categories, sub-categories and attributes or add to existing categories, sub-categories and attributes to improve your model.

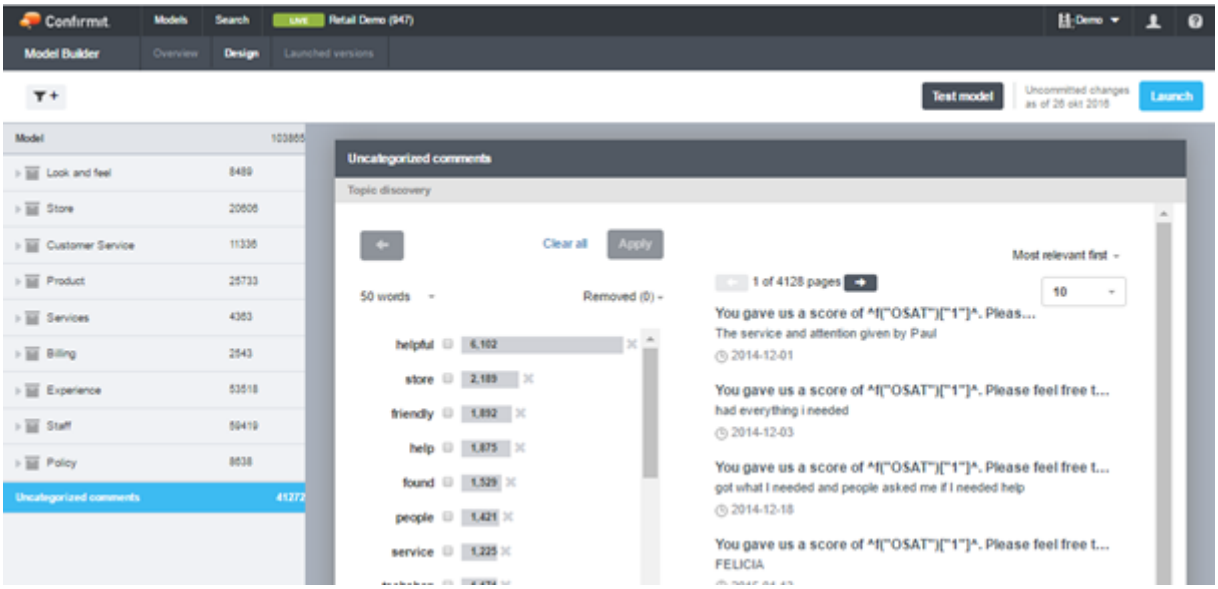

Figure 43 The Uncategorized comments

4.8.1. Topic Discovery

Topic Discovery is a feature that allows you to see the most frequently used words in your Uncategorized Comments. By default the top 50 words are shown, along with their frequency of appearance, in the left pane. Use the drop-down to select between the top 10, 50, 100 or 200 words. All comments related to the listed words are shown in the results pane on the right. Again, use the drop-down to select between the top 10, 50, 100 or 200 comments.

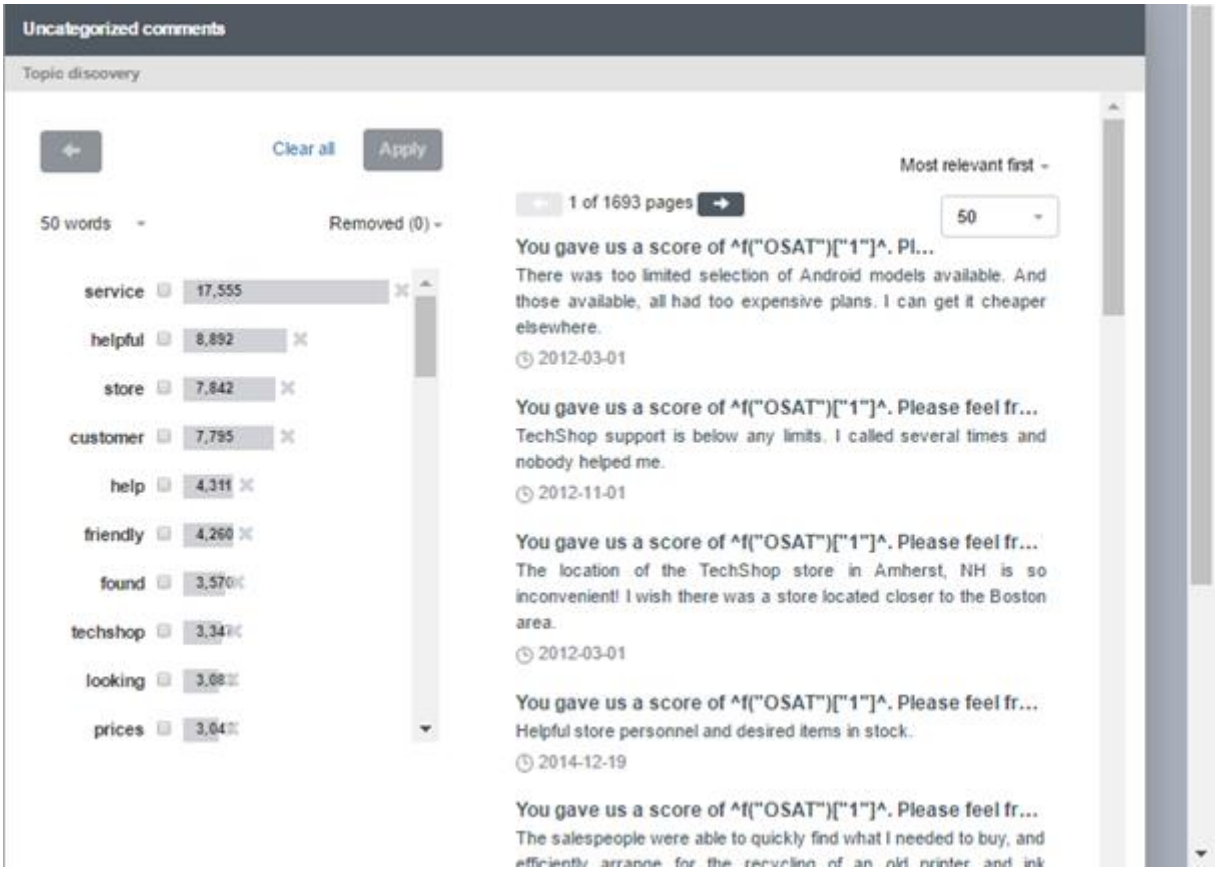

Figure 44 The Topic discovery pane

To explore and filter the results further:

- If there are words that you do not want to be included in the top words list, for example brand names or company names, remove them from the list by clicking on the **x** next to the frequency count. If you later decide you need to re-include them in the list, click the **Removed** button to display a list of removed words, then click **Add back** for the appropriate word(s).
- To display all the verbatims that include a particular keyword, check the checkbox for the word or words and click **Apply**. The verbatims that include the selected keyword(s) are then listed in the pane on the right, with the keywords highlighted. A new list of the most used words associated with the now-displayed verbatims is then presented in the left pane. You can repeat this process as many times as you wish to focus in on particular combinations of keywords.

  You can use these words to help you to update existing category expressions or to create new categories.

Click **Clear all** to take one step back in the sequence of filters.

4.9. Launching the Model

Model Builder supports versioning of the categorization models. This is to keep models consistent so that you have a benchmark for comparisons with previous periods. Maintenance of a model is important, and by using versioning you can save up all your changes and launch the complete version of the model on a particular date, enabling you to keep track of when the changes were made.

For a model to be used in Horizons, it must first be launched, thereby creating a version of the model. You will see that the label changes from Design to Live. Once launched the model is given a version number, and any edits that are made to the model after it has been launched are made in the 'working version'. The working version will need to be launched again (getting a new version number) before the changes are used in Horizons. Any changes to the structure of the model (for example if you have added any new categories, sub-categories and attributes) will need to be updated in the Horizons Database Designer table. Refer to the Genius User Guide for more details.

The **Launch** button is located towards the right side of the Design page for the Model. Update text next to the **Launch** button tells you whether you have uncommitted changes and the date that the first set of changes were made. Once you launch the next version of the model, this update will reflect whether you have made any further changes.

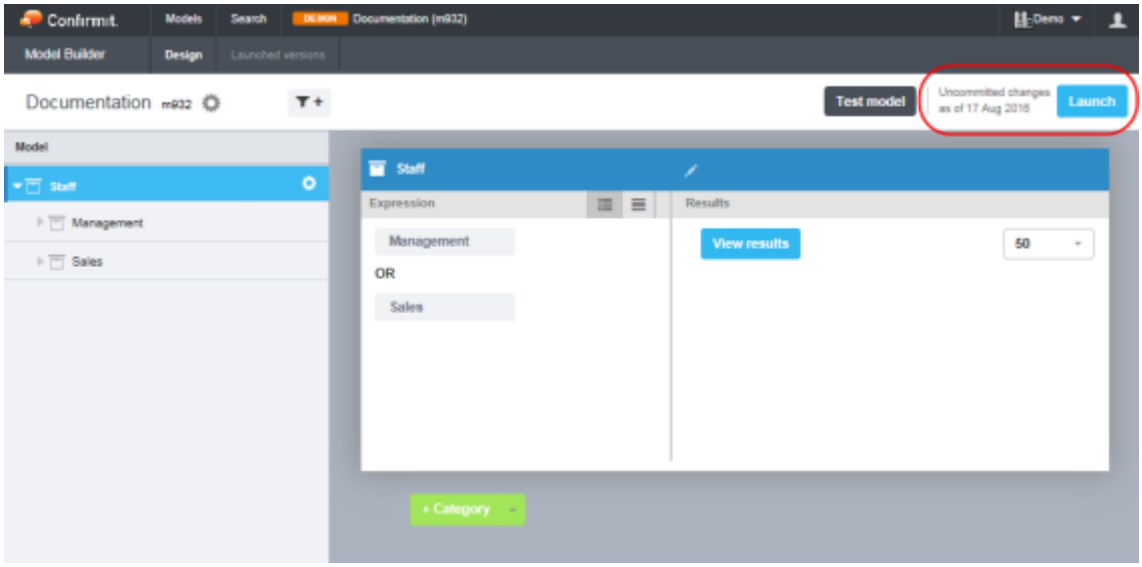

Figure 45 The Launch button on the Design page

To launch your model:

1. Go to the Design page for the model you wish to launch.
2. Select the **Launch** button in the upper-right corner of the page.

   The Launch model overlay appears.

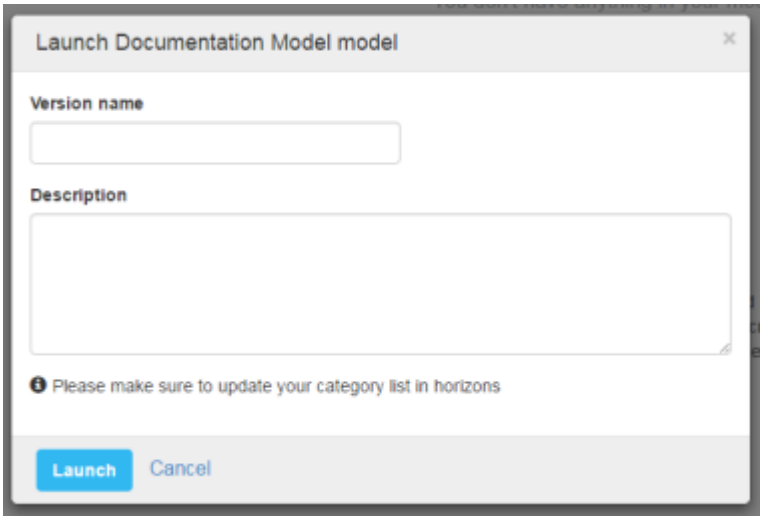

Figure 46 The Launch model overlay

3. Add a name for the version as appropriate (we recommend using the model name and a date), and a description for future reference, then click **Launch**.

   The model is “locked” (it can no longer be edited) and is released for use.
4. If you have made changes to the structure of the model (that is, you have added or deleted any categories, sub-categories or attributes), you will need to update these in the model hierarchy (in Database Designer) in Horizons.

Note: In the event one or more categories or attributes are empty then a message will be displayed informing you of this.

The next time changes are made to the model, a new “working version” will be created automatically and the changes will be made to that version. That version will then remain the working version until it is launched, whereupon it will be locked and a new working version will subsequently be created.

4.10. Viewing your Launched Versions

Once you have launched your model, you can view that model in the ‘Launched versions’ tab (next to the Design tab on the secondary navigation tab). All your launched versions are available here. You cannot edit these versions, but you can still test the model and see the results.

You can edit the description and name of the version from the cogwheel on the main launched versions page.

5. Putting it All Together

Now that you have a categorization model that is live (i.e. you have launched it), you can set up the Genius folders in Horizons. Refer to the Genius chapter in the Authoring User Guide for further details on how to achieve this.

Once you have set this up and run the Text Analytics task the first time, you will see your verbatims in Model Builder and the survey name and PID will be referenced in the survey filters.

Note: If you are running Text Analytics on a lot of projects, you will not see all your projects in the survey filter. Once there are more than 100 projects that have been run on the same account, the system will only show you the 100 project ID's with the highest number of verbatims.

Index

T

U

V

W