Paper 103-2013

Unleashing the Power of Unified Text Analytics to Categorize Call Center Data

Arila Barnes, Jared Peterson, Saratendu Sethi, SAS Institute

ABSTRACT

Business analysts often want to take advantage of text analytics to analyze unstructured data. With that in mind, SAS is delivering a new web-based application that is designed to put the power of SAS® Text Analytics into the hands of the analyst. This application combines the power of SAS® Text Miner and SAS® Content Categorization in a single user interface that enables users to automatically create statistical and rule-based models from their domain knowledge. This paper demonstrates how a business analyst in a call center environment can identify emerging topics, generate automatic rules for those topics, edit and refine those rules to improve results, derive insights through visualization, and deploy the resulting model to score new data.

INTRODUCTION

HP reports more than 2.5 billion customer transactions per year,[1] and health insurer Humana's provider call center handles more than 1 million calls per month.[2] The sheer volume of data makes it cost prohibitive to rely on humans alone to analyze the information in the call center agents' notes. Companies large and small are eager to find common customer issues early while keeping costs down. SAS Text Analytics solutions such as SAS Text Miner and SAS Content Categorization are invaluable tools for tackling such problems. SAS continually strives to provide easier and speedier access to text analysis of unstructured data, addressing both the volume and the variety of big data business problems. Last year, SAS Text Miner debuted the Text Rule Builder in support of active learning.[3] This year SAS is introducing these powerful features in a new Unified Text Analytics Interface (UTAI), which is aimed at the business analyst who wants to understand unstructured data in a more automated and meaningful way.

The SAS UTAI combines the sophisticated term-clustering statistical methods of SAS Text Miner with the rule-based natural language processing techniques of SAS Content Categorization. Although those solutions are usually used by experts, the new web application provides a single, convenient interface that enables both business analysts and experts to interactively discover topics and build categorization models so that they can better respond to the problem of automatically detecting consumer issues in a timely and efficient manner.

This paper uses a call center scenario for a fictitious online printing company to guide you through the process of topic discovery, rule generation, and model refinement and deployment. This example also shows how to use Boolean rules to tune the model and how to use the DS2 procedure to deploy this model. The paper also demonstrates how to view the results of the analysis in SAS® Visual Analytics.

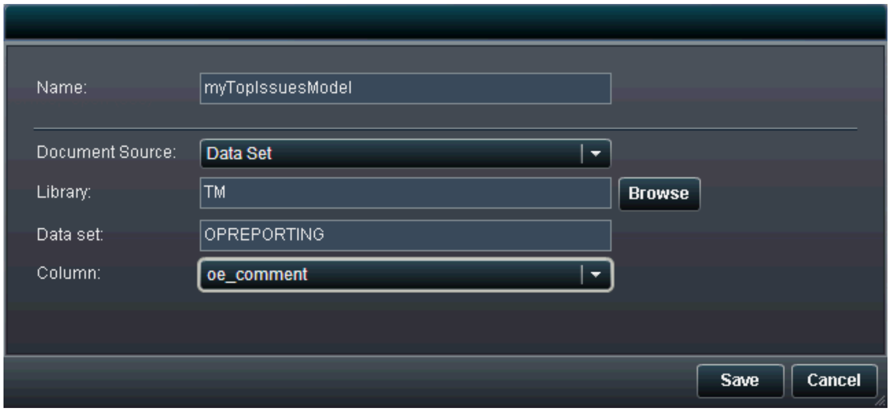
You can easily apply the process outlined in this paper to similar issues in other industries whenever data include unstructured text that contains valuable information. Patient electronic records, physician notes, police records, and insurance claims are just a few examples in which potentially valuable information is available only in unstructured text.

EXAMPLE DATA

The scenario in this paper analyzes about 15,000 free-form survey responses from the call center of an online printing company. The goal is to find customers' most common complaints about the company's products. This information enables you to classify future calls more accurately so that a call center agent can better resolve a customer issue on the first try. If the data are originally stored in a directory on the file system, the UTAI application automatically converts the files to text and loads them into a SAS data set library. The document conversion feature supports several common document formats: HTML, PDF, and Microsoft Office formats.

CREATING A NEW PROJECT

The UTAI provides a convenient wizard that enables you to quickly set up a project to explore unstructured data. To create a new project, click the "New Project" icon to start the New Project wizard. In the wizard (shown in Display 1), type a name for your project in the **Name** field. To choose the data that you want to explore, select the document type from the **Document Source** list, where you can either browse for files or select a SAS data set from a library. If you select a data set, you can focus on a particular column by selecting it from the **Column** list. Display 1 shows that this scenario focuses on the *oe_comment* column in the OPREPORTING data set in the TM library. The system is automatically preconfigured with smart defaults for natural language processing for your text. An algorithm randomly splits the specified data into a training set and a testing set for use in this model-building exercise.



Name: myTopIssuesModel

Document Source: Data Set

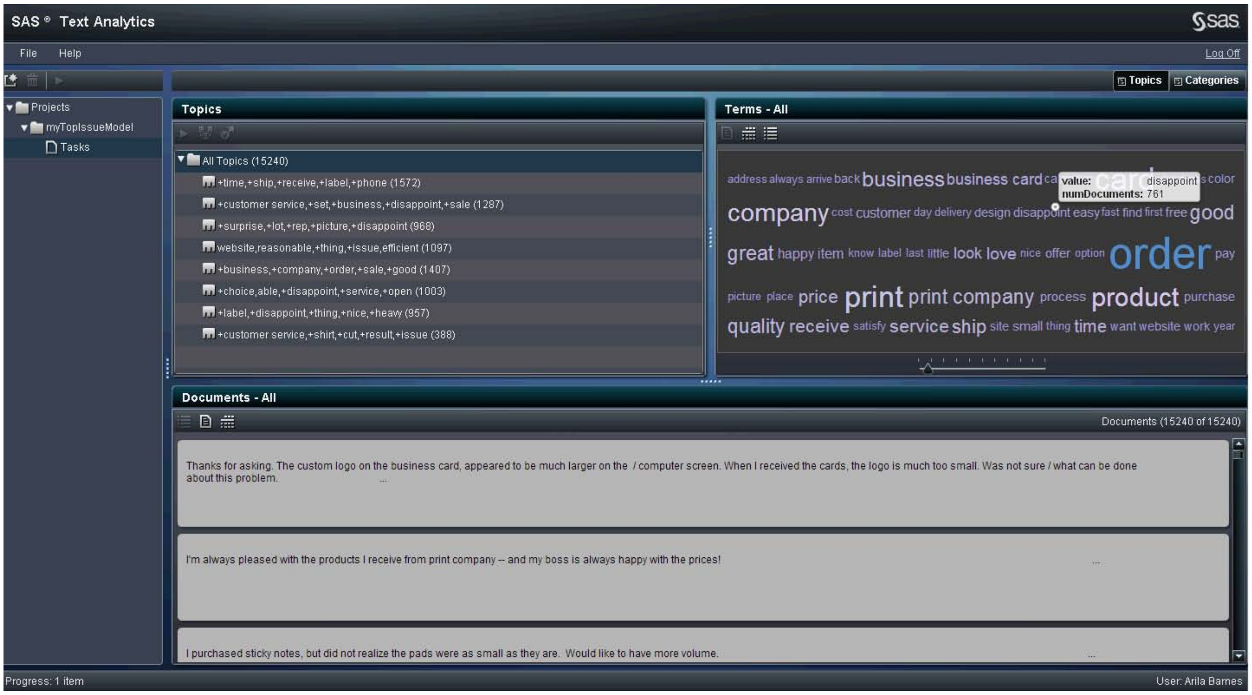
Library: TM

Data set: OPREPORTING

Column: oe_comment

Display 1. New Project Wizard

After you create the project, you can run it in the background. Behind the scenes, the software automatically converts the files to text if needed, loads them into data sets, analyzes them for part-of-speech disambiguation and entity extraction, and runs SAS Text Miner to discover topics from the raw document text.


SAS ® Text Analytics
sas
File
Help
Log Off
Topics
Categories
Projects
myTopIssueModel
Tasks
Topics
Terms - All
Documents - All
address always arrive back business business card ca value: card disappoint s color
company cost customer day delivery design disappoint easy fast find first free good
great happy item know label last little look love nice offer option order pay
picture place price print print company process product purchase
quality receive satisfy service ship site small thing time want website work year

All Topics (15240)

- +time,+ship,+receive,+label,+phone (1572)
- +customer service,+set,+business,+disappoint,+sale (1287)
- +surprise,+lot,+rep,+picture,+disappoint (968)
- website,reasonable,+thing,+issue,efficient (1097)
- +business,+company,+order,+sale,+good (1407)
- +choice,able,+disappoint,+service,+open (1003)
- +label,+disappoint,+thing,+nice,+heavy (957)
- +customer service,+shirt,+cut,+result,+issue (388)

value: disappoint
numDocuments: 761

Documents (15240 of 15240)

Thanks for asking. The custom logo on the business card, appeared to be much larger on the / computer screen. When I received the cards, the logo is much too small. Was not sure / what can be done about this problem. ...

I'm always pleased with the products I receive from print company -- and my boss is always happy with the prices! ...

I purchased sticky notes, but did not realize the pads were as small as they are. Would like to have more volume. ...

Progress: 1 item

User: Arila Barnes

Display 2. Topics Discovered Automatically

To explore the results, click **Tasks** under the **myTopIssuesModel** project in the left navigation pane. The UTAI suggests topics as shown in Display 2.

EXAMINING TOPICS

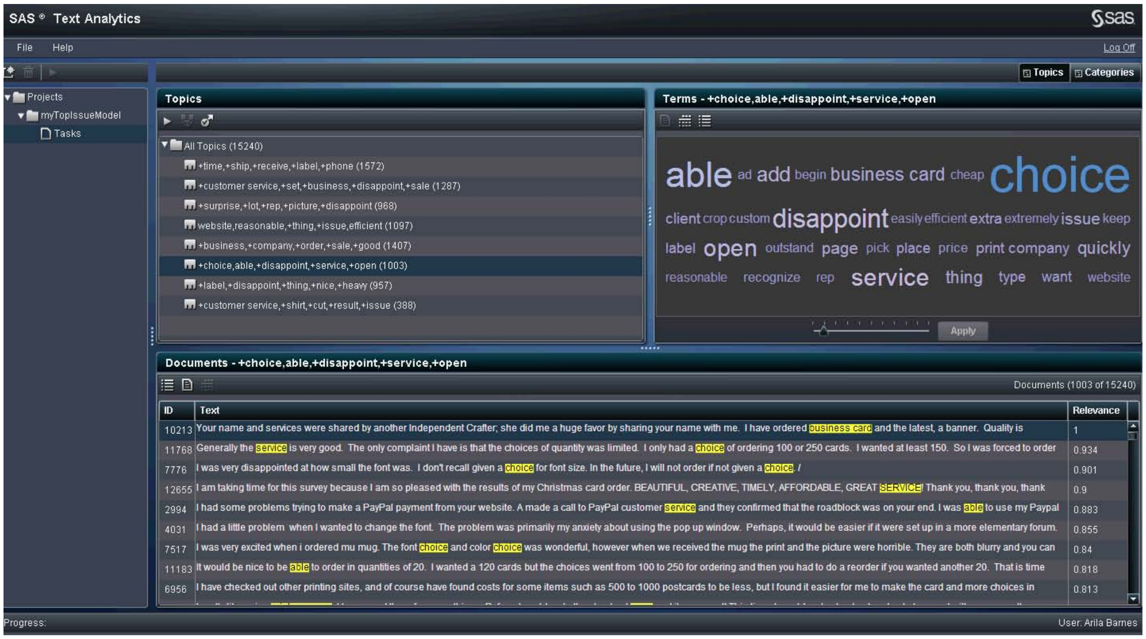
You can quickly browse the terms and their associated documents and get a feel for their importance. By default, each generated topic is named by using the five terms that occur most frequently in that set of documents. For each topic, the number of matching documents is displayed.

When you select a topic in the **Topics** pane, you can view the results in the following ways:

- A phrase cloud visualization, as shown in the upper right pane in Display 3. If the phrase cloud visualization is not displayed, you can display it by clicking the icon.

- A table view, as shown at the bottom of Display 3. The table view is designed for expert users and is not described in this paper. If the table view is not displayed, you can display it by clicking the (▦) icon.
- A concept map view, which is not shown in Display 3. For more information about the concept map, see the section "CONCEPT MAP VISUALIZATION." If the concept map view is not displayed, you can display it by clicking the ⁂ icon.

If you select the topic **+choice,able,+disappoint,+service,+open** from the **All Topics** list, you see the phrase cloud that is shown in Display 3. This phrase cloud shows that "choice," "disappoint," and "business card" are potential issues to use for the classification model. You can control the number of terms displayed in the phrase cloud by moving the slider. Click **Apply** to see the terms highlighted in the **Documents** pane. Both the **Terms** pane and the **Documents** pane are updated based on selections in the **Topics** pane. When you interact with the options in the **Terms** pane, the **Documents** pane content is updated accordingly.


SAS ® Text Analytics
sas
File
Help
Log Off
Topics
Categories
Projects
myTopIssueModel
Tasks
Topics
Terms - +choice,able,+disappoint,+service,+open
able ad add begin business card cheap choice
client crop custom disappoint easily efficient extra extremely issue keep
label open outstand page pick place price print company quickly
reasonable recognize rep service thing type want website
Apply
Documents - +choice,able,+disappoint,+service,+open
Documents (1003 of 15240)
Progress:
User: Arila Barnes

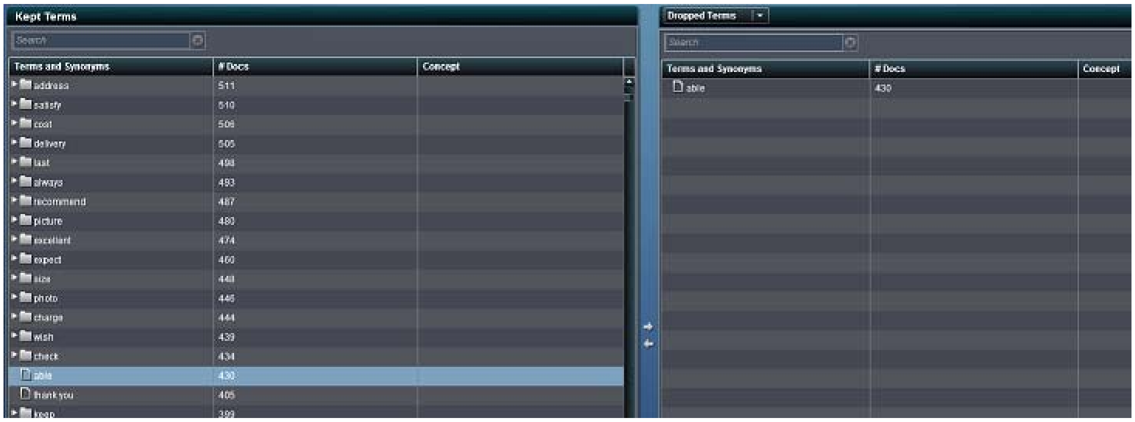
All Topics (15240)
- +time,+ship,+receive,+label,+phone (1572)
- +customer service,+set,+business,+disappoint,+sale (1287)
- +surprise,+lot,+rep,+picture,+disappoint (968)
- website,reasonable,+thing,+issue,efficient (1097)
- +business,+company,+order,+sale,+good (1407)
- +choice,able,+disappoint,+service,+open (1003)
- +label,+disappoint,+thing,+nice,+heavy (957)
- +customer service,+shirt,+cut,+result,+issue (388)

ID	Text	Relevance
10213	Your name and services were shared by another Independent Crafter; she did me a huge favor by sharing your name with me. I have ordered business card and the latest, a banner. Quality is	1
11768	Generally the service is very good. The only complaint I have is that the choices of quantity was limited. I only had a choice of ordering 100 or 250 cards. I wanted at least 150. So I was forced to order	0.934
7776	I was very disappointed at how small the font was. I don't recall given a choice for font size. In the future, I will not order if not given a choice. /	0.901
12655	I am taking time for this survey because I am so pleased with the results of my Christmas card order. BEAUTIFUL, CREATIVE, TIMELY, AFFORDABLE, GREAT SERVICE! Thank you, thank you, thank	0.9
2994	I had some problems trying to make a PayPal payment from your website. A made a call to PayPal customer service and they confirmed that the roadblock was on your end. I was able to use my Paypal	0.883
4031	I had a little problem when I wanted to change the font. The problem was primarily my anxiety about using the pop up window. Perhaps, it would be easier if it were set up in a more elementary forum.	0.855
7517	I was very excited when i ordered mu mug. The font choice and color choice was wonderful, however when we received the mug the print and the picture were horrible. They are both blurry and you can	0.84
11183	It would be nice to be able to order in quantities of 20. I wanted a 120 cards but the choices went from 100 to 250 for ordering and then you had to do a reorder if you wanted another 20. That is time	0.818
6956	I have checked out other printing sites, and of course have found costs for some items such as 500 to 1000 postcards to be less, but I found it easier for me to make the card and more choices in	0.813

Display 3. Phrase Cloud Visualization

The **Documents** pane helps you understand how these topics are related to your data. You can review the results by investigating the **Relevance** score for each document in the **Documents** pane in various levels of detail, from viewing a quick concordance match, as shown in Display 5, to reading the entire document. Also, the **Terms** table and the phrase cloud visualization add more cues to help you decide which topics to use and which ones to ignore.

You can improve the accuracy of the topic discovery process by adjusting the individual term weights or by using "stop term" lists. A "stop term" list indicates which words or phrases to ignore. These are words that do not add value to the analysis; examples are the company name, agents' names, and common phrases. Display 4 illustrates how to drop the term "able" in the **Concepts** pane.



Kept Terms

Terms and Synonyms	# Docs	Concept
address	511	
satisfy	510	
cost	506	
delivery	505	
last	498	
always	493	
recommend	487	
picture	480	
excellent	474	
expect	460	
size	448	
photo	446	
charge	444	
wish	439	
check	434	
able	430	
thank you	405	
keep	399	

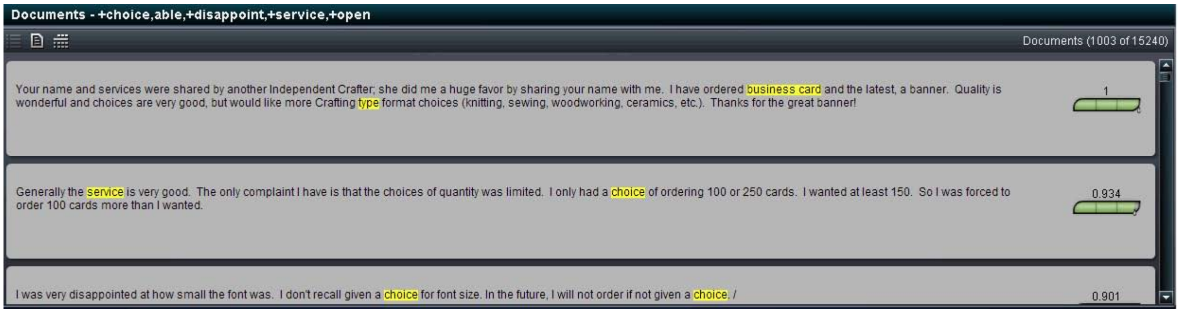
Dropped Terms

Terms and Synonyms	# Docs	Concept
able	430	

Display 4. Keep or Drop Terms

After exploring the matches in your documents, you can decide whether to combine topics or move to rule generation. A quick reading of the matching documents in Display 4 shows the following issues (in addition to some very positive comments about the service of the company):

- choice of printing quantities
- choice of font size
- blurry image
- disappointed with small font



Documents - +choice,able,+disappoint,+service,+open

Documents (1003 of 15240)

Your name and services were shared by another Independent Crafter; she did me a huge favor by sharing your name with me. I have ordered business card and the latest, a banner. Quality is wonderful and choices are very good, but would like more Crafting type format choices (knitting, sewing, woodworking, ceramics, etc.). Thanks for the great banner!

1

Generally the service is very good. The only complaint I have is that the choices of quantity was limited. I only had a choice of ordering 100 or 250 cards. I wanted at least 150. So I was forced to order 100 cards more than I wanted.

0.934

I was very disappointed at how small the font was. I don't recall given a choice for font size. In the future, I will not order if not given a choice. /

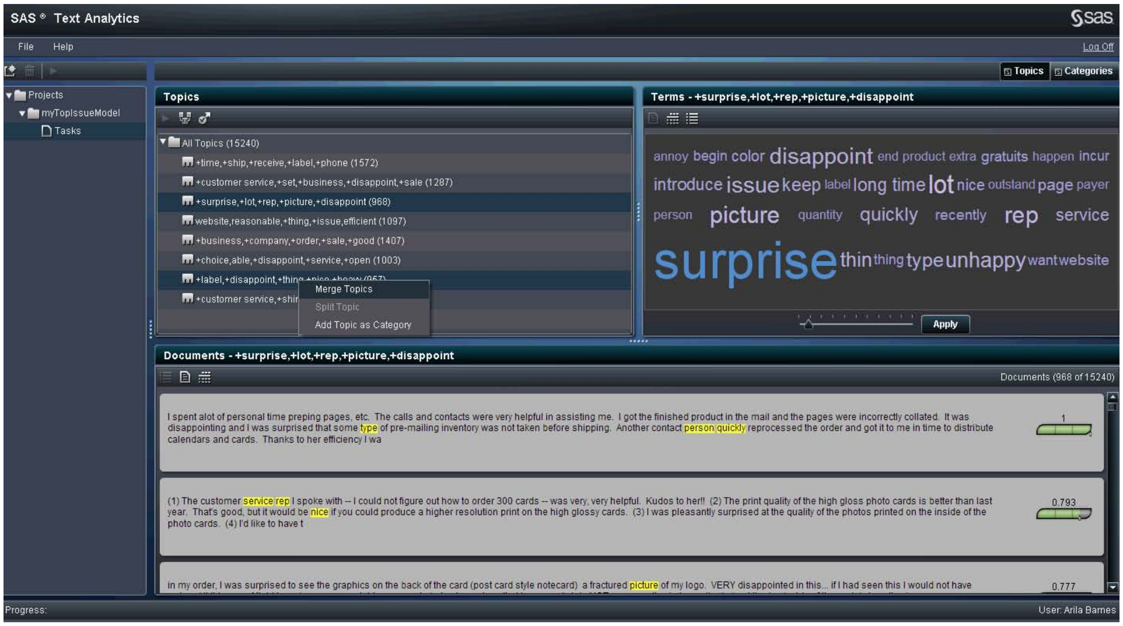
0.901

Display 5. Matching Documents

Notice that "blurry image" was not highlighted by the automatic topic discovery, so you can add a rule manually to catch that issue in the model. For more information, see the section "Editing or Adding Rules."

MERGING OR SPLITTING TOPICS

You can merge topics to simplify results when terms are similar. In this example, it makes sense to merge the topics that contain negative terms such as "disappoint." To merge topics, select two or more topics and either click the ⊞ icon or right-click and select **Merge Topics** from the context menu (as shown in Display 5). Alternatively, you can split a topic. To split a topic, select it, and then either click the ⊞ icon or right-click and select **Split Topic** from the context menu.



Topics

All Topics (15240)
+time,+ship,+receive,+label,+phone (1572)
+customer service,+set,+business,+disappoint,+sale (1287)
+surprise,+lot,+rep,+picture,+disappoint (968)
website,reasonable,+thing,+issue,efficient (1097)
+business,+company,+order,+sale,+good (1407)
+choice,able,+disappoint,+service,+open (1003)
+label,+disappoint,+thin
+customer service,+shir

Terms - +surprise,+lot,+rep,+picture,+disappoint

annoy begin color disappoint end product extra gratuits happen incur
introduce issue keep label long time lot nice outstand page payer
person picture quantity quickly recently rep service
surprise thin thing type unhappy want website

Documents - +surprise,+lot,+rep,+picture,+disappoint

Documents (968 of 15240)

I spent alot of personal time preping pages, etc. The calls and contacts were very helpful in assisting me. I got the finished product in the mail and the pages were incorrectly collated. It was disappointing and I was surprised that some type of pre-mailing inventory was not taken before shipping. Another contact person quickly reprocessed the order and got it to me in time to distribute calendars and cards. Thanks to her efficiency I wa — 1

(1) The customer service rep I spoke with -- I could not figure out how to order 300 cards -- was very, very helpful. Kudos to her!! (2) The print quality of the high gloss photo cards is better than last year. That's good, but it would be nice if you could produce a higher resolution print on the high glossy cards. (3) I was pleasantly surprised at the quality of the photos printed on the inside of the photo cards. (4) I'd like to have t — 0.793

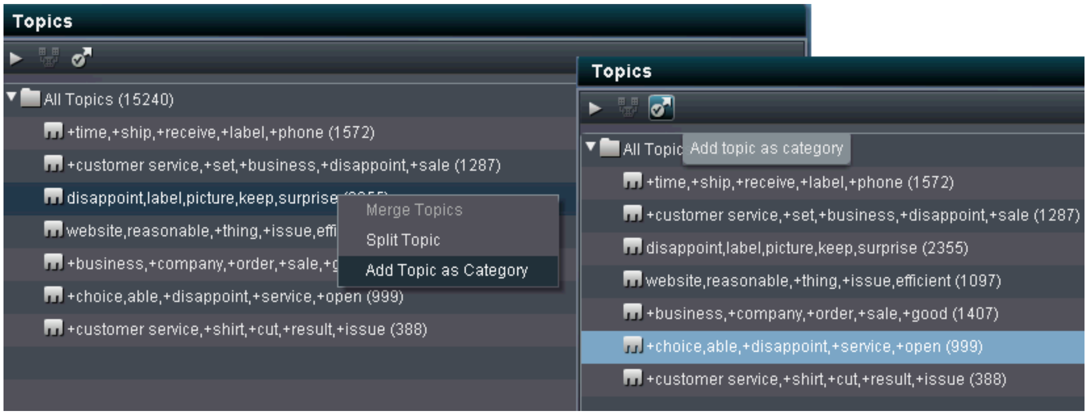
in my order, I was surprised to see the graphics on the back of the card (post card style notecard) a fractured picture of my logo. VERY disappointed in this... if I had seen this I would not have — 0.777

Display 6. Splitting and Merging Topics

CATEGORIZING CONTENT

Now that you have examined topics to understand what is going on in the data, you can decide to promote some or all of those topics as categories to begin building a taxonomy. A *taxonomy* is a hierarchical organization of categories that are useful for classifying (and organizing) unstructured data according to your business needs. You can promote topics by right-clicking and selecting **Add Topic as Category** from the context menu or by clicking the ♂ icon, as shown in Display 6. In the **Categories** pane, you can rename and add new categories as you build the taxonomy. The taxonomy captures your business domain and organizes your rules for future categorization.

The research community has approached taxonomic classification through a variety of techniques from the areas of text mining, natural language processing, and computational linguistics. With the explosion of unstructured data, organizations are eager to find automated methods of taxonomic classification in order to improve scalability, speed, consistency, and the ability to consistently reprocess data while ensuring the desired accuracy levels within the subjective context of their business. The UTAI combines the power of automatic text mining with the rule-based approach of SAS Content Categorization to provide you with finer-grained control of accuracy. When you promote your topics as categories, the UTAI uses the rule-builder functionality of SAS Text Analytics to automatically generate initial rules in Boolean syntax.

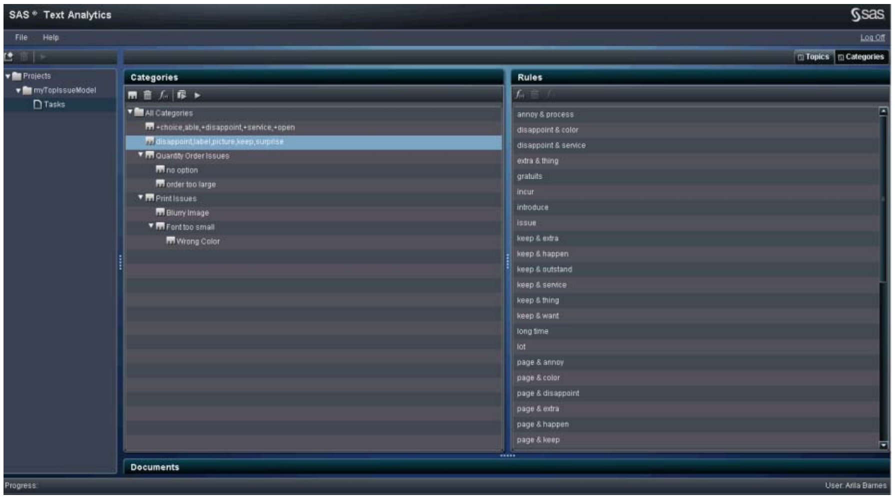


All Topics (15240)
+time,+ship,+receive,+label,+phone (1572)
+customer service,+set,+business,+disappoint,+sale (1287)
disappoint,label,picture,keep,surprise
website,reasonable,+thing,+issue,effi
+business,+company,+order,+sale,+g
+choice,able,+disappoint,+service,+open (999)
+customer service,+shirt,+cut,+result,+issue (388)

All Topic
+time,+ship,+receive,+label,+phone (1572)
+customer service,+set,+business,+disappoint,+sale (1287)
disappoint,label,picture,keep,surprise (2355)
website,reasonable,+thing,+issue,efficient (1097)
+business,+company,+order,+sale,+good (1407)
+choice,able,+disappoint,+service,+open (999)
+customer service,+shirt,+cut,+result,+issue (388)

Display 7. Generate Boolean Rules

The generated rules can be further edited and reviewed in the **Categories** view, as shown in Display 8.



- All Categories
  - +choice,able,+disappoint,+service,+open
  - disappoint,label,picture,keep,surprise
  - Quantity Order Issues
    - no option
    - order too large
  - Print Issues
    - Blurry Image
    - Font too small
      - Wrong Color

annoy & process
disappoint & color
disappoint & service
extra & thing
gratuits
incur
introduce
issue
keep & extra
keep & happen
keep & outstand
keep & service
keep & thing
keep & want
long time
lot
page & annoy
page & color
page & disappoint
page & extra
page & happen
page & keep

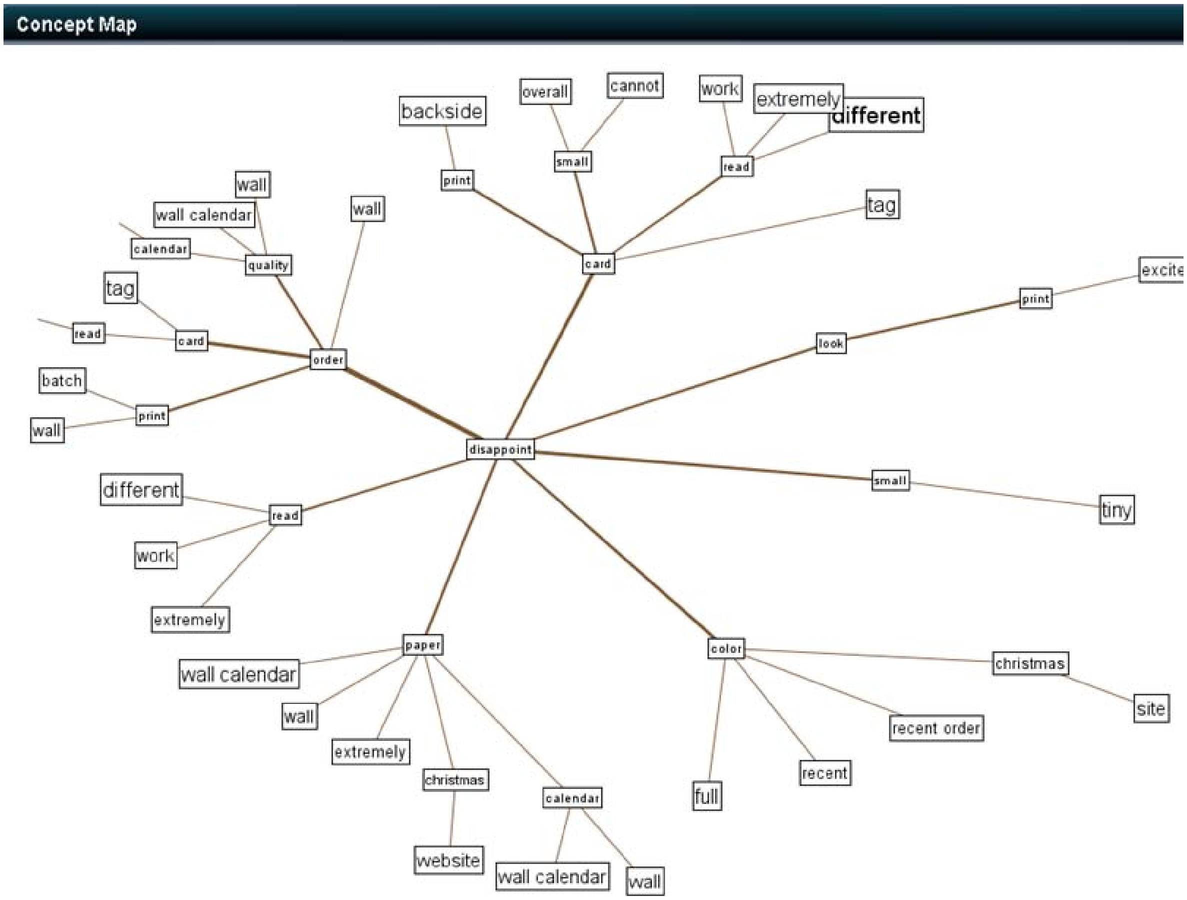
Display 8. Build the Taxonomy

By convention, the generated rule names use “&” for AND Boolean rules, “~” for NOT rules, and spaces for OR rules.

You can add new categories by clicking the ▥ icon, and you can add new rules by clicking the ƒ icon. You can rename both categories and rules, and you can arrange them hierarchically to build the taxonomy. A list of Boolean rules defines each category, and matches for the rules are displayed in the **Documents** pane.

CONCEPT MAP VISUALIZATION

The concept map is a powerful visual tool that enables you to further tune the process of creating rules. The map helps you apply your domain knowledge by exploring the topics and their associated terms to generate rules that are based only on the relationships that apply to your domain. Display 9 shows a concept map of the term “disappoint.” With one glance, you see strong connections with “small,” “card,” “color,” and “order.” If you explore “card,” you see terms such as “small,” “print,” and “different.” There is definitely an issue around the color of the printed cards and the size of the print.



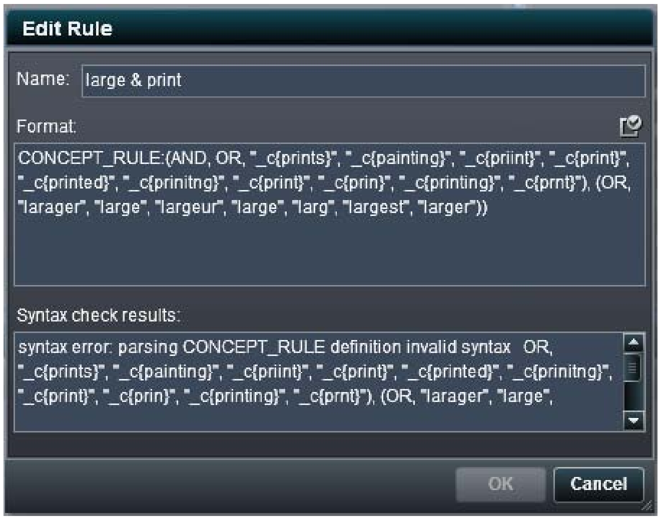
Display 9. Concept Map Visualization

EDITING OR ADDING RULES

The UTAI uses the advanced linguistic technologies in SAS® Enterprise Content Categorization to provide the following sets of Boolean operators that offer the most flexibility:

- Boolean operators: **AND**, **OR**, **NOT**
- counting operators: number of occurrences and count of distinct terms
- proximity operators: operators that are based on word distance and the scope of the sentence or paragraph
- contextual operators: operators that are based on order, XML fields, position within the document, alignment or overlap, and so on

In the **Rules** pane, you can add new rules or edit the ones that are generated by the application. By interacting with the concept map, you can adjust the rules and categories to arrive at the final classification scheme. You can check syntax by clicking the ☑ icon in the rule editor pop-up window, as shown in Display 9. After all the rules are validated, you can build the model again.



Name: large & print

Format:

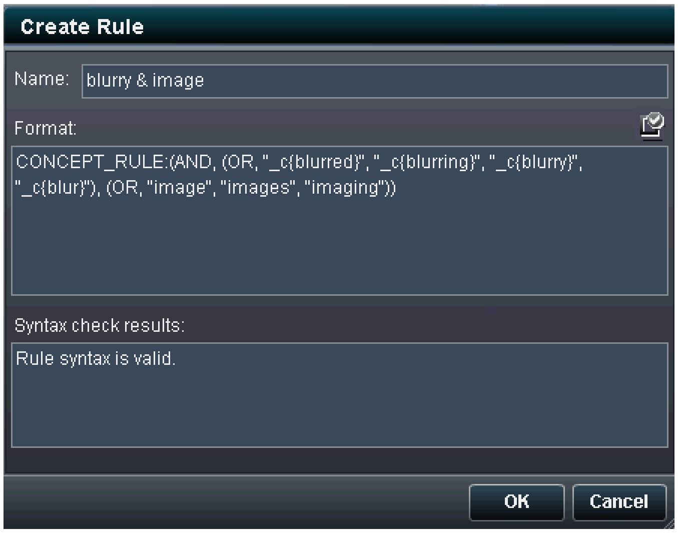
```
CONCEPT_RULE:(AND, OR, "_c{prints}", "_c{painting}", "_c{priint}", "_c{print}",
"_c{printed}", "_c{prinitng}", "_c{print}", "_c{prin}", "_c{printing}", "_c{prnt}"), (OR,
"larager", "large", "largeur", "large", "larg", "largest", "larger"))
```

Syntax check results:

```
syntax error: parsing CONCEPT_RULE definition invalid syntax   OR,
"_c{prints}", "_c{painting}", "_c{priint}", "_c{print}", "_c{printed}", "_c{prinitng}",
"_c{print}", "_c{prin}", "_c{printing}", "_c{prnt}"), (OR, "larager", "large",
```

Display 10. Syntax Validation

Earlier you noticed that issues such as “blurry image” and “small print” were not highlighted by the automatic topic discovery process in Display 3. Certainly, these are issues that should not persist for the printing company customers. You can create simple Boolean rules for these issues, as shown in Display 10 and Display 11.



Name: blurry & image

Format:

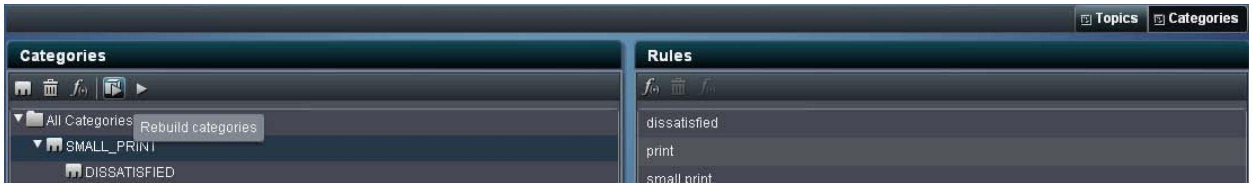
```
CONCEPT_RULE:(AND, (OR, "_c{blurred}", "_c{blurring}", "_c{blurry}",
"_c{blur}"), (OR, "image", "images", "imaging"))
```

Syntax check results:

```
Rule syntax is valid.
```

Display 11. Create a Simple Boolean Rule

When you are satisfied with all the rules, you can build the complete taxonomy model by clicking the **Rebuild categories** icon as shown in Display 12.



Categories:

- All Categories
  - SMALL_PRINT
    - DISSATISFIED

Rules:

- dissatisfied
- print
- small print

Display 12. Rebuild Categories

You can find the generated model (`rules.li`) in the SAS folder for the project. You can import this model to SAS® Content Categorization Server in order to score new content within its existing SAS Content Categorization deployment.

DEPLOYING A MODEL

When you created the project, you specified a SAS library. When you add new documents to this library, you can score them interactively by rerunning the project for the model that you created. You can deploy this model in more automated environments by using a DS2 program and the DS2 procedure, or you can deploy the model on the grid for big data scenarios. The following section shows how you can use PROC DS2.

USING PROC DS2 FOR DEPLOYMENT

DS2 is a new SAS programming language that uses packages and methods to provide data abstraction. DS2 either executes by using PROC DS2 within a SAS session or executes directly within selected databases where SAS® Embedded Process is installed. The DS2 packages included in UTAI are TKCAT and TKTXTANIO. TKTXTANIO is a utility package that is required by TKCAT.

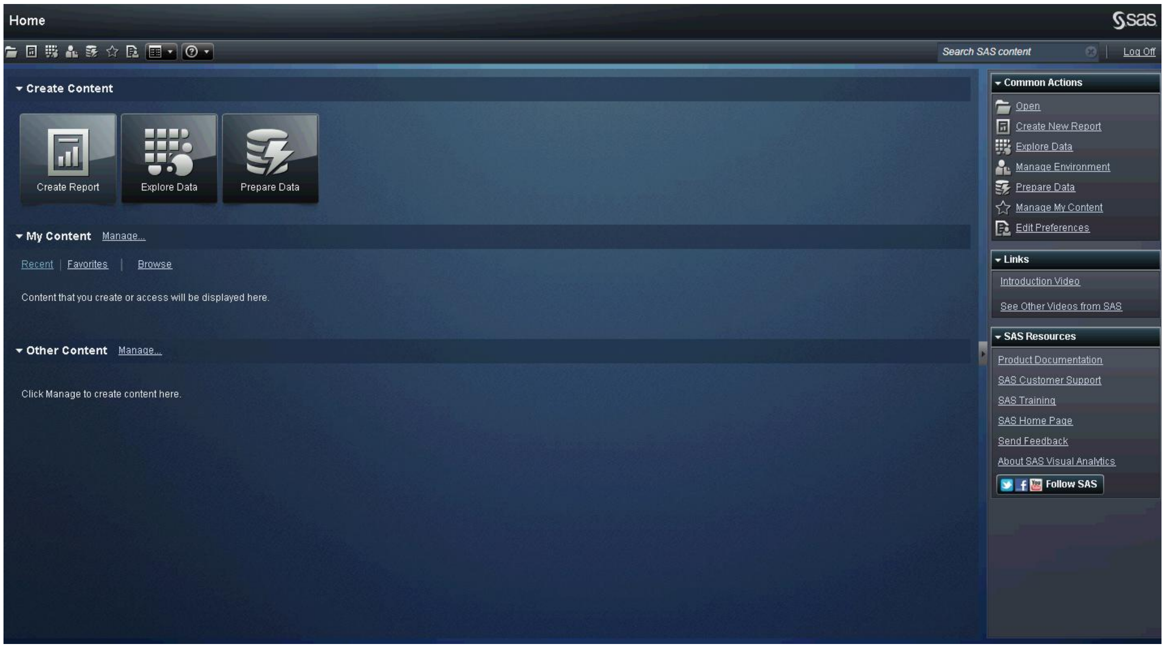
TKCAT Source Code Sample

The following code illustrates how to use the TKCAT package to apply the model that you build in the SAS UTAI:

```
libname mydata 'C:\SAS Data Sets';

proc ds2;
 require package tkcat; run;
 require package tktxtanio; run;
 /* oe_comment is the column that contains text in your data set */
 table result(drop=(oe_comment status current_concept total_concepts transact document
settings model));

 dcl package tkcat cat(); /*TKCAT categorization engine package */
 dcl package tktxtanio txtanio(); /*Utility package required by TKCAT*/
 dcl binary(8) transact;
 dcl binary(8) document;
 dcl binary(8) settings;
 dcl binary(8) model;
 retain transact;
 retain settings;
 retain model;

 method init();
 /* Create a transaction to use for scoring documents */
 transact = cat.new_transaction();
 /* Create the default settings for the transaction */
 settings = cat.new_apply_settings();
 /* Set the model created in UTAI to apply to new documents */
    model = txtanio.new_local_file('..\mytopissuesmodel\conf\rules.li');
 status = cat.set_model(settings, model);
 if status NE 0 then put 'ERROR: set_binary fails';
 /*    Initialize the categorization engine with your model */
 status = cat.initialize_concepts(settings);
 if status NE 0 then put 'ERROR: initialize_concepts fails';
 end;

method run();
 set mydata.opreporting; /* Data set that contains text to analyze */
 /*Specify document to score */
 document = txtanio.new_document_from_string(oe_comment);
 status = cat.set_document(transact, document);
 if status NE 0 then put 'ERROR: set_document fails';
 /* Apply the model to the document transaction*/
 status = cat.apply_concepts(settings, transact);
 if status NE 0 then put 'ERROR: apply_concepts fails';
/* Iterate for each concept to get details */
 total_concepts = cat.get_number_of_concepts(transact);
 current_concept = 0;
 do while (current_concept LT total_concepts);
 myterm = cat.get_concept(transact, current_concept);
 tag = cat.get_concept_name(transact, current_concept);
 parent = cat.get_parent(transact, current_concept);
 sentence = cat.get_sentence(transact, current_concept);
 output;
 current_concept = current_concept+1;
 end;
 /* Reset document transaction for each document */
 cat.clean_concepts(settings, transact);
 txtanio.free_object(document);
end;

method term();
 /* Clean up variables*/
 cat.free_transaction(transact);
 cat.free_apply_settings(settings);
 txtanio.free_object(model);
end;
run;
```

INTEGRATING WITH SAS VISUAL ANALYTICS

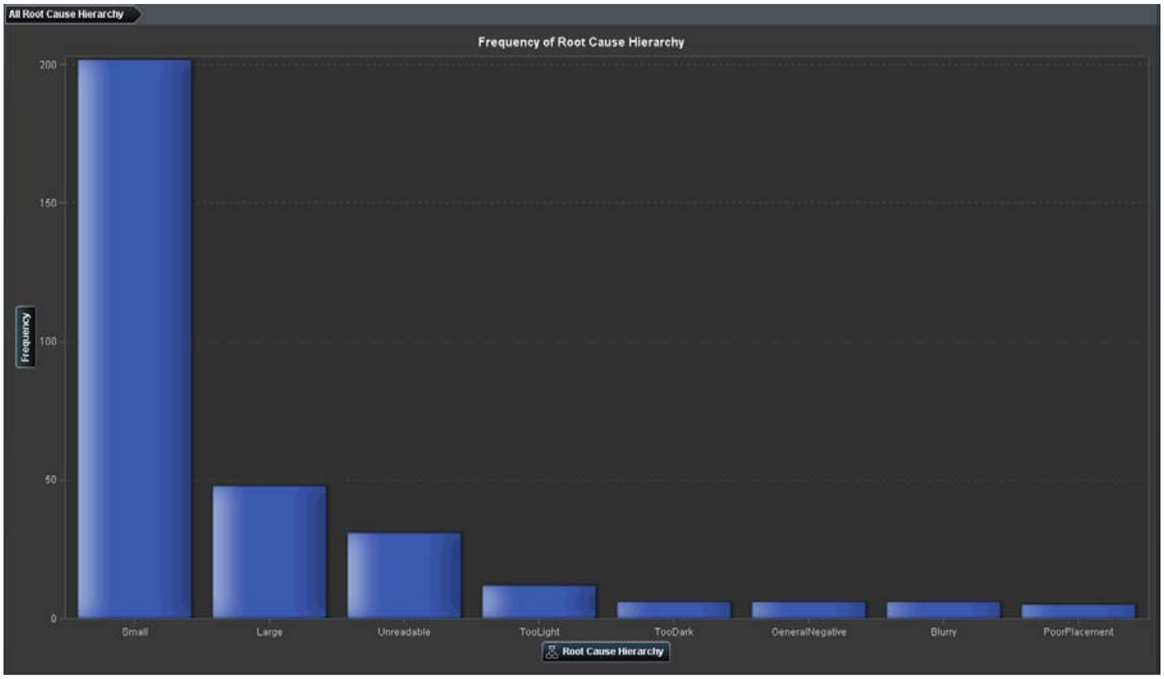
The UTAI web application has several built-in interactive reports, including the concept map and the phrase cloud. In SAS Visual Analytics you can easily create additional reports from the data that are scored by the UTAI. From the SAS Visual Analytics hub, you can select **Create Report** in the SAS® Visual Analytics Designer. Display 12 shows how to access SAS Visual Analytics Designer to build a custom report. You can build a simple histogram as shown in Display 14 to display the main categories in which new documents are placed when the model is applied to new data, and you can compare that histogram to the similar report on previous data. This type of report is useful for monitoring whether certain issues persist.



Content that you create or access will be displayed here.

Click Manage to create content here.

Display 13. SAS Visual Analytics Hub



Display 14. Custom Report Frequency of Root Cause

CONCLUSION

Text is a largely unused asset in many organizations. Firms need to interpret, summarize, and report on information that is contained in documents. This paper demonstrates the convenience of a single user interface in a familiar call center scenario. It also helps you understand the benefits of a single web application that provides a framework for interactively discovering and building a content categorization model. In a single installation, the SAS Unified Text Analytics Interface enables multiple users to access the combined power of the algorithms in SAS Text Analytics and SAS Content Categorization. It is a complementary solution in the SAS Text Analytics product suite that is also tightly integrated with SAS® High-Performance Analytics technologies. SAS Unified Text Analytics Interface is formally planned to be released under the name SAS® Contextual Analysis in Q3 2013.

Although this paper focuses on call center data, the same principles apply to the collection of unstructured data in any organization. Furthermore, the rich analytical tools that SAS offers can augment the analysis of unstructured text to help organizations understand the virtues of moving beyond reporting to proactive, forward-looking business analytics that reduce uncertainty, predict with precision, optimize performance, and minimize risks in their business.

NOTES

1. “HP Turns 2.5 Billion Customer Transactions into Customer Intimacy,” http://www.sas.com/success/hp-big-data.html?utm_medium=RSS&utm_source=SASCustomerStory. Accessed Feb. 14, 2013.
2. “Humana Reduces Call Center Volume with Text Analytics,” http://www.sas.com/success/humana.html?utm_medium=RSS&utm_source=SASCustomerStory. Accessed Feb. 14, 2013.
3. Beth Schultz, “SAS Boosts Text Mining & Extends Model Support,” *All Analytics*, Oct. 9, 2012, http://www.allanalytics.com/author.asp?section_id=1411&doc_id=252067.

RECOMMENDED READING

- *SAS® Contextual Extraction Studio User Guide*
- *SAS® DS2 Language Reference*

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Arila Barnes
SAS Institute Inc.
10 Fawcett Street
Cambridge, MA 02138-1175
Phone: (617) 576-6800, Ext. 54213
Fax: (617) 576-6888
E-mail: arila.barnes@sas.com
Web: www.sas.com

Jared Peterson
SAS Institute Inc.
100 SAS Campus Drive
Cary, NC 27513
Phone: (919) 531-4274
Fax: (919) 677-8000
E-mail: jared.peterson@sas.com
Web: www.sas.com

Saratendu Sethi
SAS Institute Inc.
10 Fawcett Street
Cambridge, MA 02138-1175
Phone: (617) 576-6800, Ext. 54246
Fax: (617) 576-6888
E-mail: saratendu.sethi@sas.com
Web: www.sas.com