SAS® GLOBAL FORUM 2021

Paper SAS 1202-2021

Making Text Analytics More Approachable

Sundaresh Sankaran, SAS Institute

ABSTRACT

Text Analytics proves most effective when it is approachable to business users, packaged within easily deployable and accessible applications, and helps you make better and effective decisions at scale. Core text analytics capabilities, including those offered by SAS , help analysts and data scientists categorize documents, extract information, discover topics and identify sentiment. In this paper, we offer a selection of applications built over these core capabilities, each of which help you obtain answers to questions that motivate you as a business users of AI, and enable you to receive recommendations, gain insight, and make decisions. Our hope is that these applications serve to expand your understanding of the applicability of text analytics capabilities, and the value it brings to your organization.

INTRODUCTION

Text Analytics needs to be more approachable to business users. Observing the variety of text analytics applications available today, including the SAS Visual Text Analytics® offering, we perceive that many are targeted towards the developer of analytics, such as a text analyst, taxonomy developer, or a data scientist. SAS Visual Text Analytics®, for example, offers an easy to use, GUI-based solution in order to categorize documents, extract information from documents, discover topics and identify sentiment.

While this is a powerful application which offers a hybrid approach comprising Machine Learning and rules-based intelligence, it's oriented towards developers of Artificial Intelligence (viz. the data scientist or text analyst), and viewed along the analytical lifecycle, offers capabilities that stop at the stage of the analysts visualizing results. A recent change, which will be referred to in this paper, extends the analytics lifecycle by making the output and results of analytics accessible to business-focused and operations-focused stakeholders, and helps them obtain answers to their questions.

In addition, SAS has been continuously challenged (and welcomes such challenges) by customers such as yourselves to come up with solutions that solve more specific business problems, rather just offer up the direct end-results of analytics. The stage has now shifted to where we would like to help the consumer of Artificial Intelligence (AI). Examples of AI consumers are CxOs, operational decision-makers & line managers. Providing more extensible, accessible, and business-specific applications, together, helps to make Text Analytics more approachable.

WHO IS THE CONSUMER OF AI?

A consumer of AI differs from the data scientist or text analyst in terms of the type of questions they ask of the data. For example, a Customer Support Manager may like to use text analytics not to just ask ' What is the sentiment?', or they may not stop with 'what are the topics that customers called about', but ask questions more in terms of "What can I do to stop customers from getting angry?" In short, a business user looks for recommendations, areas of further analysis, and additional insights. The type of questions that business users

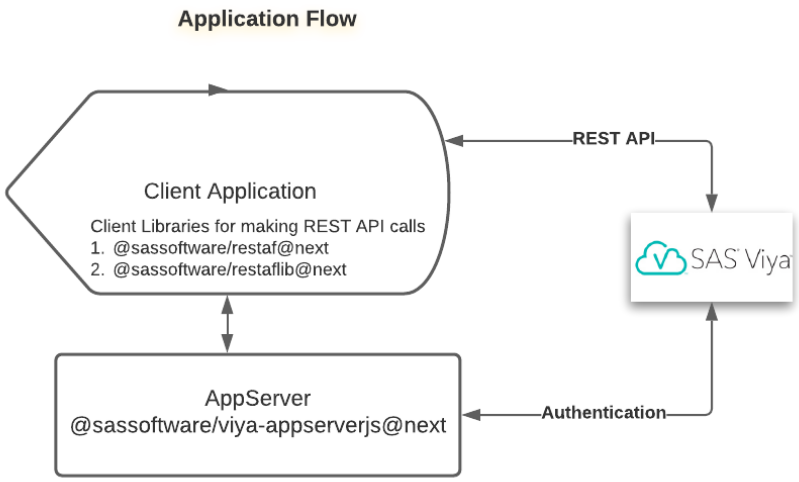
ask are fed by different motivations. These consumers of AI concentrate more on outcomes, rather than process or analytical methods.

We therefore take up the question of how to make text analytics more approachable to the business user. We adopt a fairly broad definition of the term approachable and look at all ways to reduce the mystery around text analytics and make it easier to consume , easier to deploy and easier to understand. While this paper focuses on text analytics and Natural Language Processing (NLP) applications, one may like to make similar arguments for other branches of analytics as well.

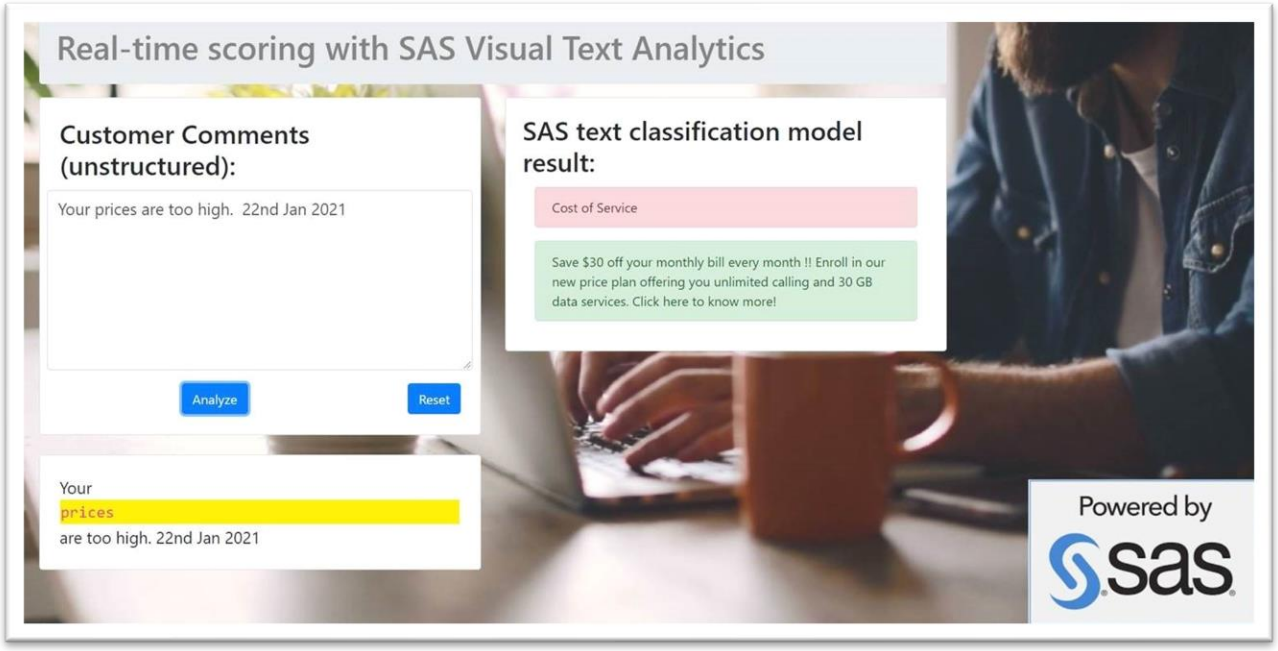
PROVIDE INSIGHTS AND RECOMMENDATIONS

AI consumers require applications that analyze text or a document in real time, preferably calling capabilities from the same application where the text was generated in, and as part of a defined workflow that has downstream applications (what to do with the text analytics result). To illustrate this point, we created an example that combined core text analytics capabilities, reporting capabilities, and decisioning and model management solutions from SAS. Add in a healthy dose of standard, easy to deploy web development frameworks such as NodeJS, and we have an application that can be easily deployed and interact with analytics provided by a SAS Viya environment. You may consider these as suggestions through which you can use SAS in friendly integration with other technologies to execute business decisions, thus leading to better analytics Return on Investment (ROI).

One aspect of approachability is to provide the business user an application which is easy to use and answers their questions. Enterprises may choose to embed text analytics capabilities within their business applications . To illustrate this, we demonstrate an application, which uses SAS Viya REST APIs and open source libraries such as restaf (https://github.com/sassoftware/restaf). The restaf library simplifies making REST API calls to a SAS Viya® environment. The following diagram illustrates how this application works.



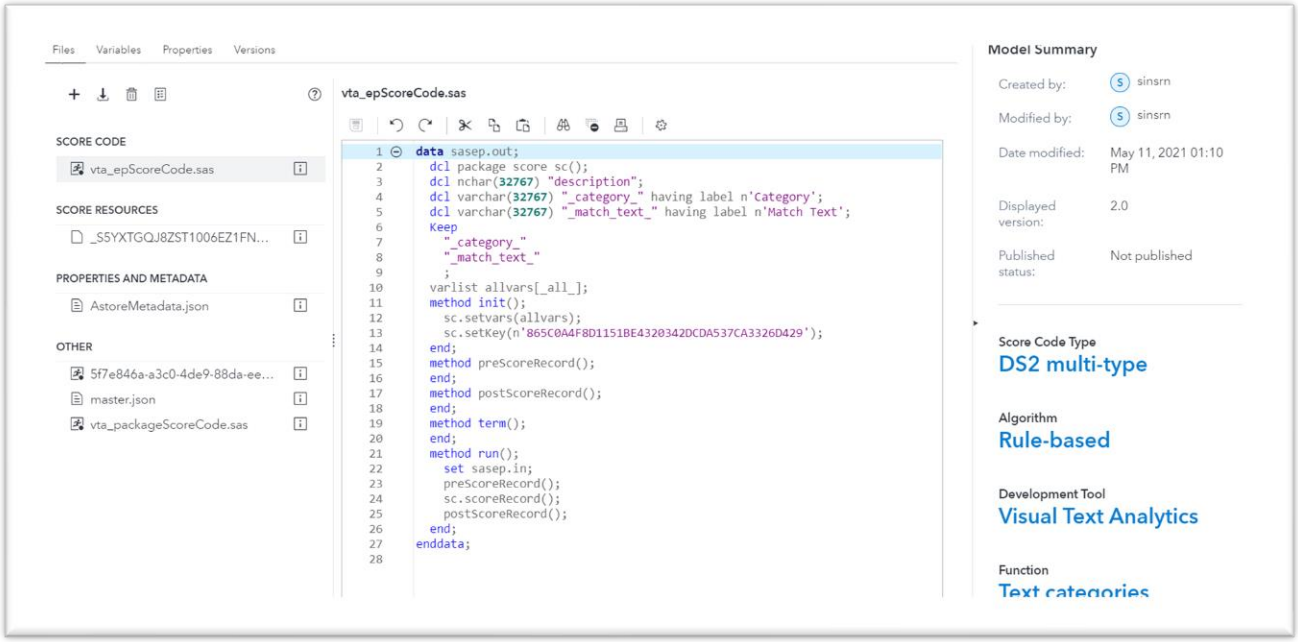
This application can be deployed in many types of environments including on-premise and the cloud, making use of Kubernetes. Users can build similar applications and can refer https://github.com/sassoftware/restaf-uidemos/ for examples.



For this specific use-case, we look at the business use-case of large volumes of feedback and complaints that are sent to customer support centers, a typical situation for service industries with large customer bases. One needs to balance handling complaints efficiently as per proper priority, with a need to make sure your customer doesn't get angry and churn. There are also support to sales opportunities you may like to take advantage of. Using the application described earlier, we demonstrate how, upon providing feedback, a text analytics model categorizes such feedback and also provides a recommendation which is targeted towards the customer and looks to either mitigate negative impact or increase customer engagement. It could also be used for providing a recommendation to the agent.

How is this possible ? Recent features introduced to SAS Visual Text Analytics® now allow for categories, concepts, topics, and sentiment models to get published to Microanalytic services (MAS), which is SAS's powerful mechanism and API meant for executing models and decision flows. Upon registration of a text analytics model belonging to these types from Model Studio (the application which is used to create text analytics models), a package is created which contains the binary file and necessary score code, in astore format, which enables a single observation of input to get analyzed, leading to a single output object, which contains the results in a JSON structure. The format used for delivering output is called the DATAGRID format.

Here is an example of the DS2 wrapper code used to execute the model. Note the other necessary model files, including the binary file (astore) which is included in the package.



The required model is then published to Micro Analytics Services (MAS) through the Publish button, which can be accessed through either SAS Model Manager® or SAS Intelligent Decisioning®.

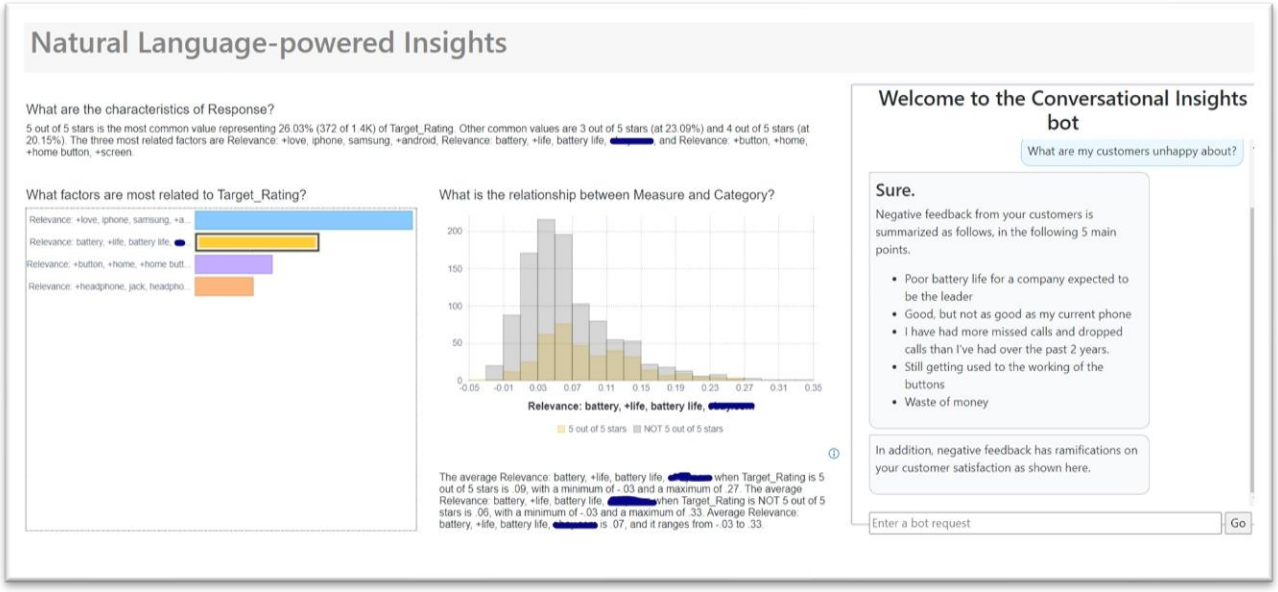
It is this resultant endpoint that gets called from within a web application, which determines that the initial category, to which the feedback belongs to, refers to "Cost of Service". This "Cost of Service" category is simultaneously presented as a result, but also used as an input variable to another Decision Flow, which determines an appropriate next step, in this case, recommending a lower priced plan to the user. Using MAS enables us to quickly arrive at and execute decisions at scale.

ASK DIFFERENT QUESTIONS OF YOUR DATA

Another aspect of approachable text analytics is that it helps us answer the questions that business users ask of their data in an easier to understand fashion. Business users like to obtain answers to their questions through a conversational medium and appreciate the faster time to value they obtain were they to retrieve these answers through the use of a conversational medium. We enable these conversational insights through many approaches and SAS technologies – Natural Language Understanding, Natural Language Conversations & Natural Language Generation, all provided by SAS Conversation Designer®, and SAS Visual Analytics®, surfaced through the Visual Analytics Software Development Kit (SDK).

Natural Language Understanding deals with how software can take up an utterance made by humans, and apply a combination of rules and machine-learning techniques in order to understand the correct intent, and derive required parameters in order to execute a query or run a program. SAS Conversation Designer®, a chatbot design application provided by SAS, helps chatbot designer access a Natural Language Understanding service in order to frame intents and attach them to the correct conversational flows. Natural Language Conversations (NLC) is a service within SAS Conversation Designer®, which determines the flow of many components of a conversation, known as dialogues, and executes desired programs within these dialogues in order to satisfy or fulfill the intent. Finally, it is up to the third Natural Language-related capability, Natural Language Generation, to frame an easy to understand

answer and communicates the same back to the user. Learn more about the many capabilities of SAS Conversation Designer® at www.sas.com/conversation-designer.



A major component of approachable text analytics lies in how the final results are communicated to the user. As indicated by the saying 'a picture is worth a thousand words', the results provided by the approachable application should get the message across as succinctly as possible. Visualization provides us with an additional tool to facilitate this. You use the SAS Visual Analytics Software Development Kit (SDK) as an additional element in your web application to render the appropriate visualization, as shown in the above example. Upon the business user asking for insights on the recent feedback he has received, you witness the following actions.

1. The conversational agent triggers a text analytics procedure to summarize the feedback for the given time period, resulting in an extractive text summary getting generated, highlighting the main points for specific negative feedback.

2. Using the same parameters (the feedback data item, and a date range, if provided), the conversational agent also passes a query to a SAS Visual Analytics® report

3. Based on the query parameters, the SAS Visual Analytics® report provides a dashboard relating negative feedback to overall customer satisfaction which is rendered on screen through a Visual Analytics SDK element.

This way, you obtain not just the data or information that you asked for, but also additional insights which seek to provide contextual information that helps you make better decisions.

BUILD AND EXTEND ACCELERATORS

While considering options to make text analytics more approachable, you should also take advantage of pre-built models, rules, or definitions that make it easier for your analytics teams to develop such applications and provide you better time to value. Applications based on text analytics utilize models built by analysts with linguistic and machine learning expertise. But you may find it tough to find such resources easily nowadays due to the huge

gap between demand and supply of data science practitioners. The shortage of skills that is currently experienced in the marketplace give data scientists, text analysts included among them, the status of unicorns. One way you can still make progress is to use accelerators and pre-built components so that you don't need to start every project from scratch.

You have the option of taking advantage of pre-defined concepts that SAS Visual Text Analytics® provides, which reduces the burden on your text analytics team to come up with rules to identify commonly used entities. You benefit through quickly identifying common entities such as references to people, locations, and organizations, among other definitions, to understand your text data better. Here is a brief description of the nine pre-defined concepts that SAS Visual Text Analytics® provides.

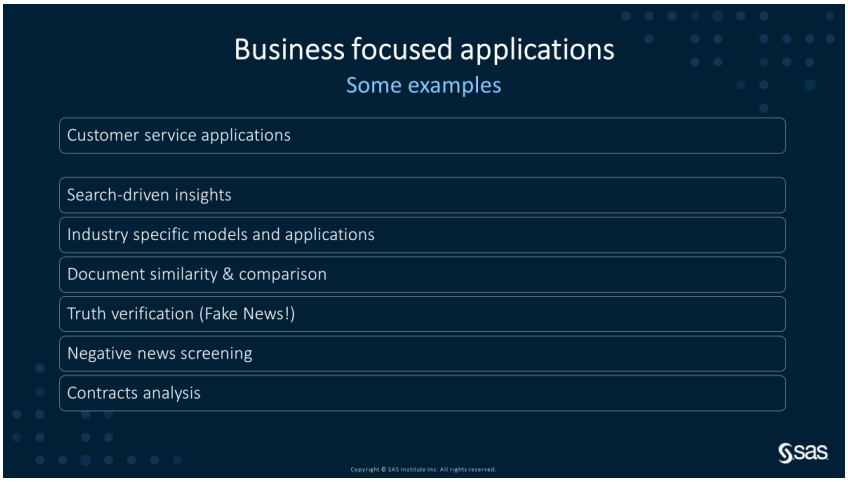
- nlpDate – references to mentions of dates within a text document
- nlpMeasure – references to measures in a document – usually some numeric value accompanied by a unit of measure – for example, six weeks, 15^0 Celsius
- nlpMoney – references to a monetary value, or a unit of currency.
- nlpNounGroup – commonly occurring phrases or group of terms
- nlpOrganization – references to organizations mentioned in the text. Example, the United Nations
- nlpPercent – references to percentages
- nlpPerson – references to mentions of names , or people in the text document
- nlpPlace – references to locations or other places in the text document
- nlpTime – references to mentions of time within a text document

This feature provides you a jumpstart in solving your business problem. At the same time, it is important to also customize the definition of these entities to your specific business problems. For example, what if you wanted a rule to be very specific in identifying Tacoma, the car vs. Tacoma the city in Washington state?

For this reason, SAS provides you the capabilities to customize these concept definitions further, using the subject matter expertise of your text analytics team. Working with pre-defined entities, the amount of effort that your team may have to engage in is much lesser than what it would have been, had they designed the same from scratch. The net benefit is that your organization, overall, is more productive in harnessing the value of text analytics models and can concentrate more on building business focused applications to access them. Some examples of business focused applications are provided in the next section.

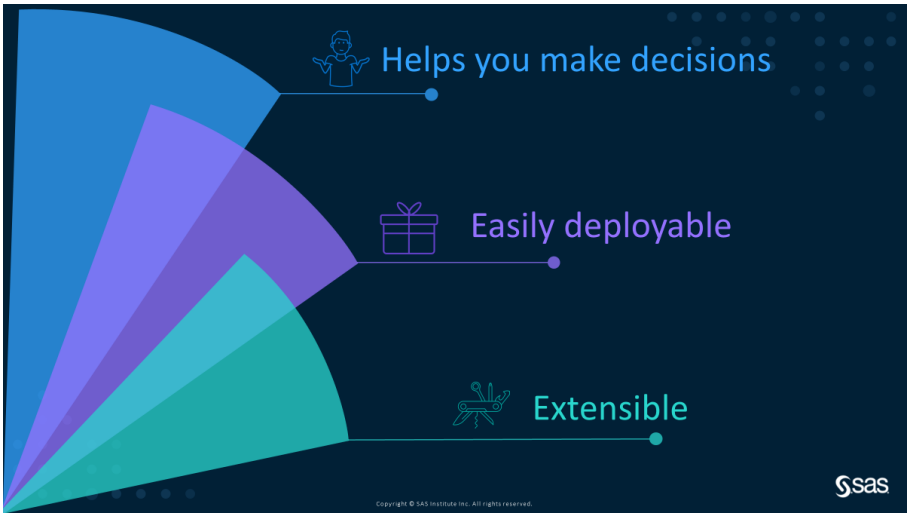
OFFERING A CHOICE OF APPROACHABLE APPLICATIONS

Building upon the principles expressed above, we suggest more business-focused applications that organizations can look to build, which best utilize the results of text analytics and provide insights to business users. The primary challenge for organizations, when considering building such applications, has been with regards to customization, taking care of industry-specific nuances, and integrating these applications within an organizational workflow. The examples provided in this paper are an indication of our attempt to address these challenges by providing approaches, tools and examples, which organizations can take up and modify for their business purposes.



CONCLUSION

The approaches, tools and examples provided through this paper lead us to specify some broad principles which we believe help in making text analytics more approachable.



First, look at text analytics as a means, and not an end in itself. The ultimate objective should be to look at using text analytics, along with a combination of other objectives, in order to make business decisions at scale. All the examples illustrated in this paper, and those created when preparing this paper, depended not just on text analytics, but other useful approaches such as Decision management, Business Rules, and Visualization.

Next, adopt deployment approaches that help you quickly build accessible applications for your business needs. A common approach is to embed these capabilities directly into your existing websites and business applications and in this context, note the recent advances from SAS to make this possible. Watch the SAS Viya® release video for May 2021 to learn more : https://www.youtube.com/watch?v=EFmiQgcT3kQ&list=PLVV6eZFA22Qy3Phh79SwcT-4tB84vnQ48

Finally, consider your approachable application to be extensible and incorporate other approaches. There are continuous advances both within the field of text analytics as well as

other analytical approaches which work on the same data or the results of text analytics. For example, upon extracting common entities such as people, organizations, and locations latent in a set of documents, you get additional value if you were to look at the relationships of these entities through a graph-based structure, also known as a network diagram. Further techniques based on network analytics also help you identify closely connected entities, all of which further the cause of obtaining more insights.

REFERENCES

1. Website Kumar, Deva. "restaf - An Easy Button To Access SAS Viya Using SAS Viya REST APIs", github.com/sassoftware, 2nd Jan 2020. Available at https://github.com/sassoftware/restaf/wiki.
2. Website SAS Institute. "SAS Visual Analytics SDK for developers", https://developer.sas.com/guides/visual-analytics-sdk.html.
3. Website SAS Institute. "SAS Conversation Designer sample connector", https://github.com/sassoftware/conversation-designer-sample-connector.
4. Website Furbee, Joe. "Authentication to SAS Viya: a couple of approaches" blogs.sas.com, 25th Jan 2019. Available at https://blogs.sas.com/content/sgf/2019/01/25/authentication-to-sas-viya/.
5. Website SAS Viya Release Highlights, May 2021, available at the SAS YouTube channel, https://www.youtube.com/watch?v=EFmiQgcT3kQ&list=PLVV6eZFA22Qy3Phh79SwcT-4tB84vnQ48

ACKNOWLEDGMENTS

I acknowledge and thank Deva Kumar, SAS institute, creator of the restaf framework (https://github.com/sassoftware/restaf-uidemos/tree/k8s) which enabled creation of different assets supporting this paper.

Thanks also to my colleagues on the SAS Conversation Designer® Research and Development team, especially Chris Barefoot and Ben Tomlinson, for their help on various aspects relating to the SAS Conversation Designer® custom connector, which enabled one of the example applications in this paper.

RECOMMENDED READING

- https://github.com/sassoftware/restaf-uidemos/tree/k8s
- https://communities.sas.com/t5/SAS-Communities-Library/Building-a-simple-web-interface-for-your-chatbot/ta-p/714197
- https://blogs.sas.com/content/subconsciousmusings/2021/04/29/utilizing-text-models-with-ease/

CONTACT INFORMATION

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

Sundaresh Sankaran
SAS Institute
Sundaresh.sankaran@sas.com

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

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