

# An Artificial Intelligence Framework on SAS® Viya® to Counter International Human Trafficking

Tom Sabo, SAS Institute Inc

## ABSTRACT

Efforts to counter human trafficking internationally must assess data from a variety of sources to determine where best to devote limited resources. These varied sources include the US Department of State's Trafficking in Persons (TIP) reports, verified armed conflict events, migration patterns, and social media. How can analysts effectively tap all the relevant data to best inform decisions to counter human trafficking?

This paper builds on two previous SAS Global Forum submissions, which apply SAS text analytics to the TIP reports as well as the Armed Conflict Location and Event Data (ACLED) Project. We propose to show a framework supporting Artificial Intelligence on SAS® Viya® for integrating and exploring all data related to counter human trafficking initiatives internationally, incorporating the TIP, and ACLED sources as a starting point. The framework will include SAS rule-based and supervised machine learning text analytics results not available in the original data sets, providing a depth of computer-generated insight for analysts to explore. We will ultimately show how this extensible framework provides decision makers with capabilities for countering human trafficking internationally, and how it can be expanded as new techniques and sources of information become available.

## INTRODUCTION

Human trafficking, also known as modern day slavery, is a global phenomenon, and no country is exempt. Victims of modern slavery are exploited in every region of the world, in forced labor or commercial sex in the real world of industry and on the pages of the Internet. The enormity of the problem necessitates the development of a unified, comprehensive response from world leaders to collectively address a crime that defies all borders<sup>1</sup>.

There is good news in terms of countries and organizations identifying and collecting human trafficking data. In 2009, only twenty-six countries had an institution which collected and disseminated data on trafficking, while by 2018, this number had risen to sixty-five<sup>2</sup>. Furthermore, over the years, the development of standards in data collection has been considered by the international community as a key activity to enhance national responses in the field of trafficking in persons<sup>3</sup>. Multi-stakeholder partnerships are critical to collect and use this data to combat human trafficking. They must exist vertically between national, regional, and local governments, and horizontally between law enforcement, service providers, and other key actors such as humanitarian organizations and other NGOs within and across communities<sup>1</sup>.

The evolving landscape surrounding the international human trafficking problem and the evolution of sources of data related to addressing the problem has necessitated a new analytics strategy. These sources of data have become too extensive and diverse for manual search and analysis alone. In this paper, we propose an enterprise framework to counter international human trafficking, which leverages evolving sources of data, and enables analytics and artificial intelligence (AI) methods for these counter trafficking purposes. Using these methods, data can be assessed and used to help set up and inform these multi-stakeholder partnerships to provide direction for counter human trafficking action. We will explore the specifics of these AI and analytics capabilities within this paper as well.

How would an AI-driven framework benefit counter human trafficking at an international level? We offer the following goals:

1. **Strategic** - Improve the accuracy and value of strategic data and depict this through interactive visualizations. This will provide stakeholders with a more complete picture of patterns and trends in what is happening, improving strategic decisions and providing impact.
2. **Operational/Tactical** - Improved the specific responses and provide analytics to inform these specific responses. For example, using analytics to identify that a militant group attacks villages regularly in the northeast region of a country gives organizations information that can target a response.
3. **Data model** - As data collection improves, organizing, and categorizing this data becomes paramount. Our proposed counter trafficking framework can take existing data models from sources and enhance them or create data models using previously unstructured data, such as reports, news sources, or notes.
4. **Integration** - As new sources of data become available, these can be incorporated alongside other sources of data assisting with the strategic, operational, and tactical goals.

For the purposes of this paper, we will use the definition of Artificial Intelligence (AI) as the simulation of human intelligence by machines. AI is particularly adept at processing and analyzing large amounts of data to provide targeted courses of action for human consideration. It applies machine learning, deep learning, and natural language processing (NLP) to solve actual problems, including counter human trafficking<sup>4</sup>. We define machine learning as a branch of AI based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention. Deep learning extends machine learning by using multiple network layers, and is used in recognizing speech, identifying objects, and extracting patterns within text. We will show how these capabilities fit into the Counter Trafficking Framework.

In the following sections, we build upon two previous global forum submissions to showcase a combination of SAS technology that supports the four goals listed above. It is our aim to show what stakeholders can do more effectively using these capabilities. As such, we will focus on three different areas of improvement for strategic, operational, and tactical operations to counter human trafficking.

1. How can stakeholders incorporate multiple data sources in one visualization environment, and best analyze information from those sources of data to make strategic decisions? This section will use *SAS Visual Analytics* and *SAS Visual Text Analytics*.
2. How can stakeholders better target indicators of human trafficking conditions in freeform text? In this section we will also leverage *SAS Visual Analytics* and *SAS Visual Text Analytics*.
3. Can we better target the tactical where, what, and how of human trafficking conditions and events? This section focuses on the use of *SAS Visual Investigator*.

By the end of this paper, the reader will become familiar with how to apply different aspects of the SAS analytics life cycle to the problem. We will focus on natural language processing techniques, machine learning, and point to deep learning capabilities that will help provide targeted courses of action for human consideration. The reader will gain an understanding of how key capabilities in *SAS Visual Analytics*, *SAS Visual Text Analytics*, and *SAS Visual Investigator* all fit together to improve counter trafficking efforts in a single solution environment via the set of goals above.

## ANALYZING DATA FROM MULTIPLE SOURCES IN A COMMON ENVIRONMENT

Human trafficking is a crime that is committed with the intention to hurt people. In conflict situations, traffickers can operate with even greater impunity<sup>2</sup>. Armed conflicts can drive vulnerabilities to trafficking in persons in different ways. Thus, we propose in our counter trafficking framework capabilities that can link patterns of trafficking in persons with armed conflict event data to better understand how armed conflict creates and exacerbates human trafficking.

Areas with weak rule of law and lack of resources to respond to crime provide traffickers with a fertile terrain to carry out their operations. This is exacerbated by more people in a desperate situation, lacking access to basic needs. Some armed groups might exploit civilian victims of attacks or kidnappings, including children, for sexual exploitation, sexual slavery, forced marriage, armed combat and various forms of forced labor<sup>2</sup>. Furthermore, refugees or those who leave a country under duress are vulnerable to exploitation by traffickers when fleeing violence and while residing in refugee camps<sup>1</sup>. In Figure 1 below from the United Nations Office of Drugs and Crime's 2018 report on Human Trafficking, they depict the various ways in which victims are exploited in conflict areas and while fleeing conflict areas.

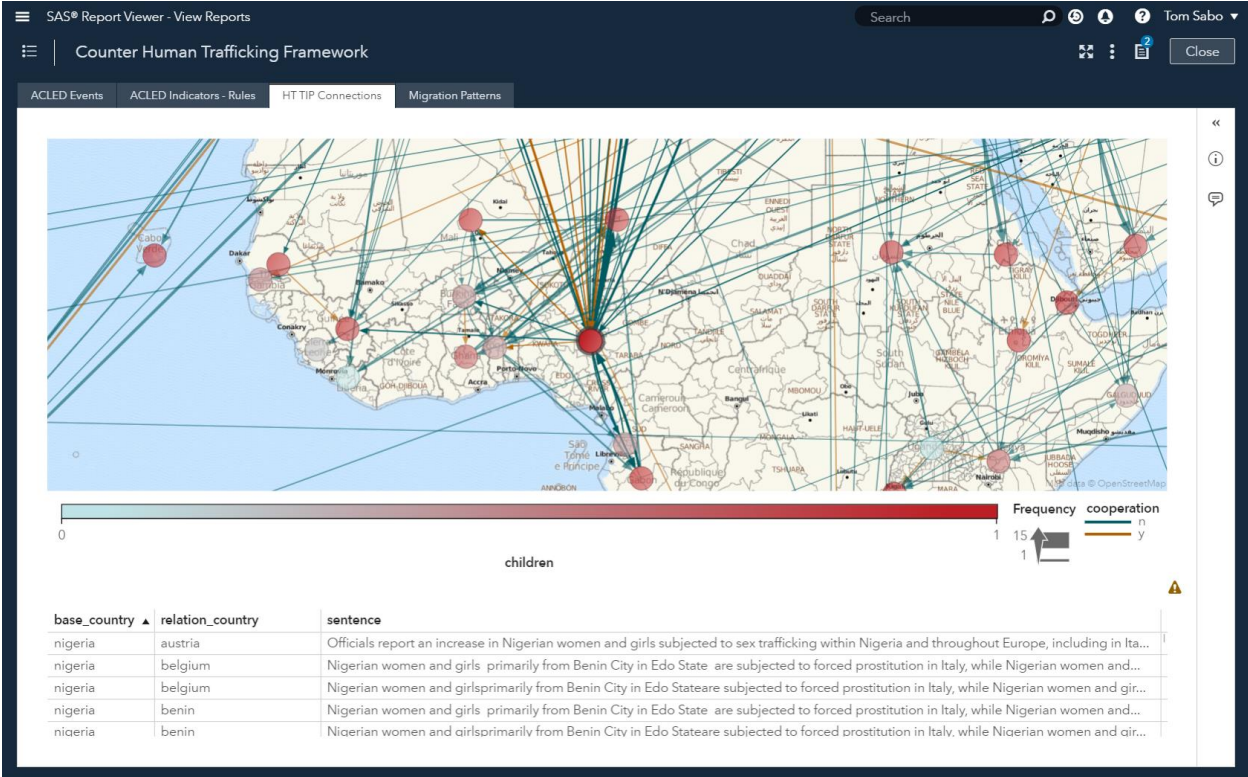


**Figure 1: Reported forms of trafficking in persons directly and indirectly related to armed conflict**

In 2016, we, at SAS, performed a study on extending the Armed Conflict Location and Event Data project<sup>5</sup> (ACLED) through text analytics<sup>6</sup>. The key innovation of the project enabled machine learning to model the text associated with each conflict type, including violence against civilians, to assess a deeper level of themes for each type. For this paper, we replicated the 2016 work on the ACLED data in SAS Visual Analytics for 2013-2017 armed conflict events. Dashboard visualizations of the analysis were subsequently presented in SAS Visual Analytics. This helped answer the question of what types of violence was being committed to civilians in the reports. A number of these sub-themes could be considered directly related to generating human trafficking conditions. Kidnaping and abduction could be considered directly in line with exploiting victims in conflict areas, as shown in Figure 1. Incidents where villages were bombed could cause displacements of many families and individuals, resulting in exploitation while fleeing conflict areas. In this process, we enhanced an existing source of data, to give it more relevance to the counter human trafficking problem.

In 2018, we underwent an investigation into the US Department of State Trafficking in Persons (TIP) reports with the goal to assess international human trafficking patterns<sup>7</sup>. The key innovation of that project enabled AI methods through SAS Visual Text Analytics to identify and geospatially visualize patterns in trafficking across countries. Using these methods, we were able to identify from unstructured data who was being trafficked (men, women, children), what type of trafficking is occurring (labor or sex trafficking), and whether the countries in question are in cooperation to address the problem. In the process, we generated a new structured source of data that captured these patterns, where formerly there was only free-form text.

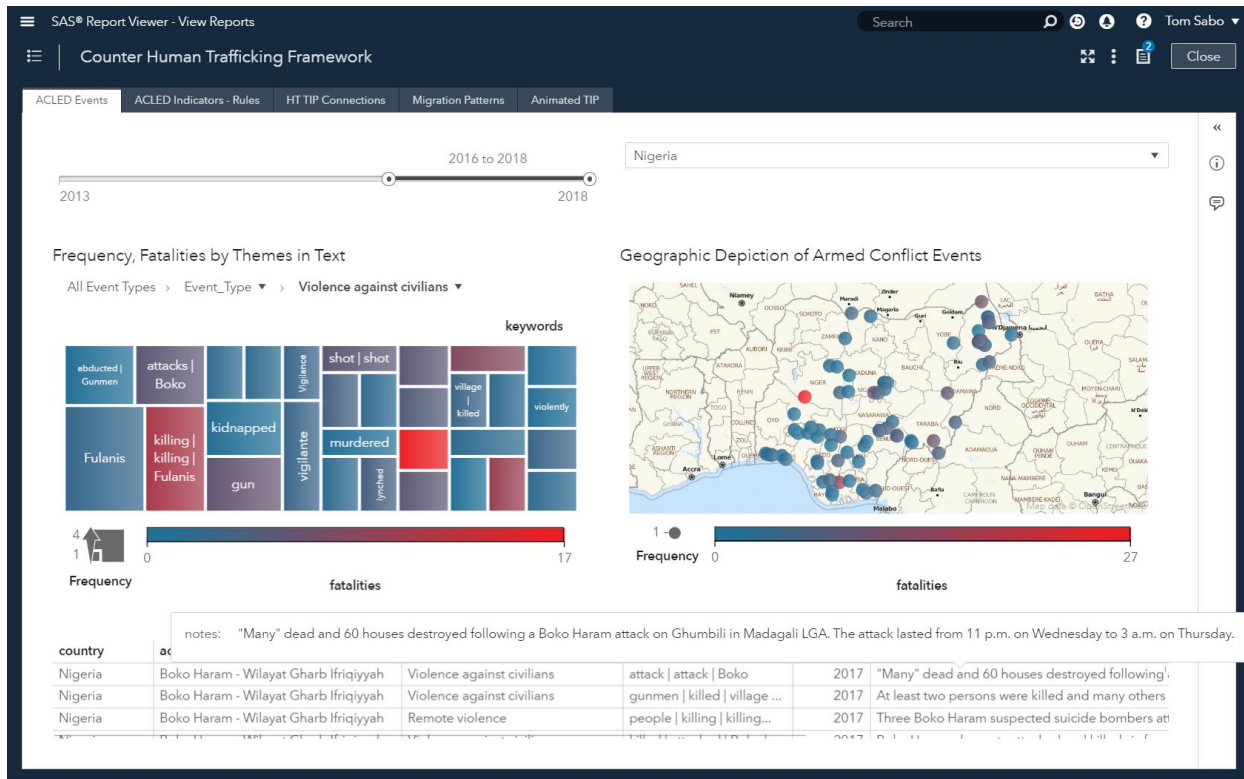
These sources work extremely well together to help analyze the causes of trafficking in geographic locations, particularly where they related to armed conflicts. Figure 2 depicts patterns in trafficking in west and central Africa, derived from TIP reports spanning 2013-2017. An analyst assessing the reports can be questioned regarding the conditions surrounding trafficking in the region during this time span or might want to understand the armed conflict events that contribute to these trafficking patterns. Using this framework, they can answer these questions.



**Figure 2: SAS® Visual Analytics depicts patterns in human trafficking surrounding Nigeria and West Africa**

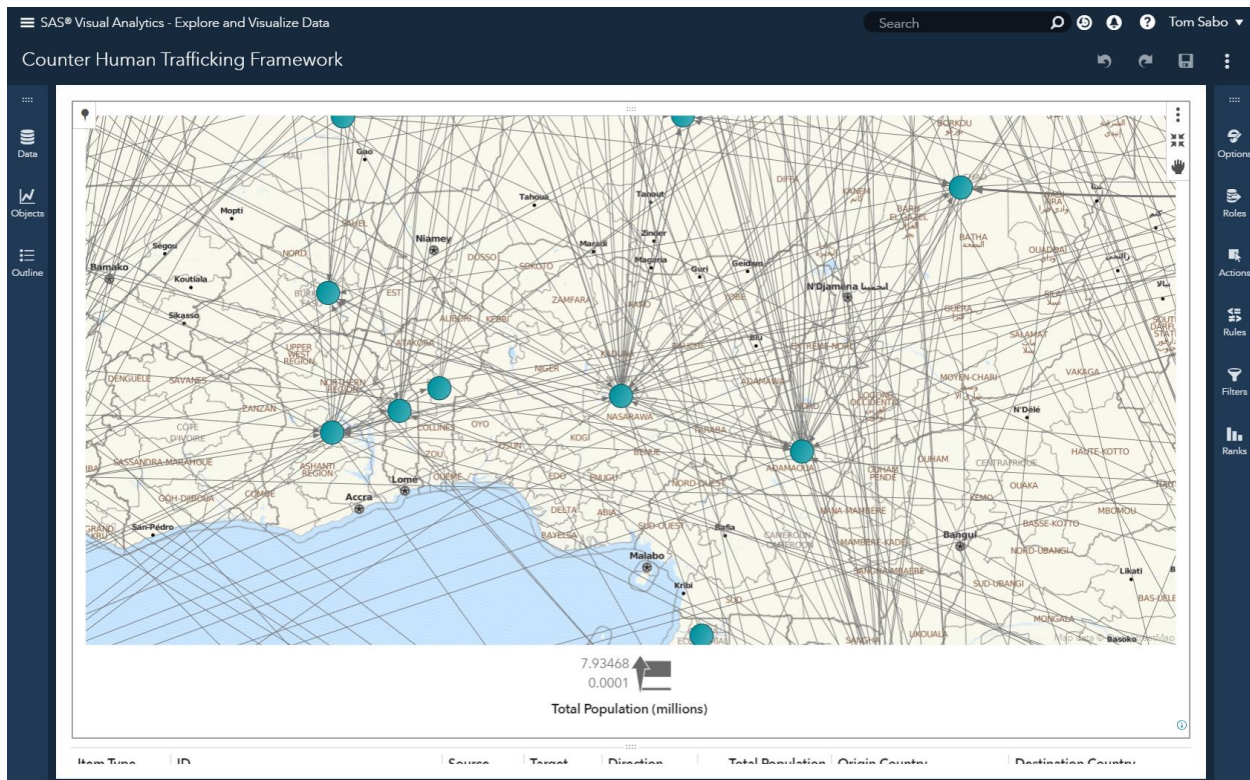
SAS Visual Analytics enables drilldown between data sources. For example, an analyst interested in the human trafficking situation in Nigeria could look specifically at patterns of violence against civilians. Figure 3 below shows events related to violence against civilians from 2016-2018 in Nigeria against the ACLED data. The depiction shows how we enhanced the data with SAS Visual Text Analytics modeling capabilities to include subthemes of violence, including village attacks and kidnaping. This can be used to assess the effectiveness of campaigns in reducing the impact of military groups. In this example, the data shows that violence against civilians continues in Nigeria in recent years, including but not limited to the militant group Boko Haram. Such attacks result in death and

displacement, generating conditions where families could be vulnerable to subsequent human trafficking either in the conflict area, or when fleeing the conflict area.



**Figure 3: Assessment of ACLED events in SAS® Visual Analytics for human trafficking vulnerabilities**

This system could be used to compare relative amounts of violence year over year and assess whether this correlates with migration patterns. The UNHCR, the UN Refugee Agency, hosts a database that contains data about populations of concern from 1951 to 2017, and this can be used to investigate different aspects of these populations, including general composition by location of residence or origin, status (refugees, internally or externally displaced persons, and so on.)<sup>8</sup>. We obtained migration data from and assessed in the same solution environment whether the patterns of migration also match up with the reported patterns of human trafficking from the TIP reports. Such comparisons help validate the data, and subsequently, any strategic and tactical actions that result from use of the data. Figure 4 below depicts in SAS Visual Analytics the migration patterns surrounding Nigeria and West Africa from 2013-2017. These patterns of migration from UNHCR are very similar to the trafficking patterns extracted from the TIP reports. Trafficking and migration patterns overlap surrounding Nigeria and Niger, as well as in the West African countries.



**Figure 4: Assessment of HHNCR migration data to verify human trafficking vulnerabilities**

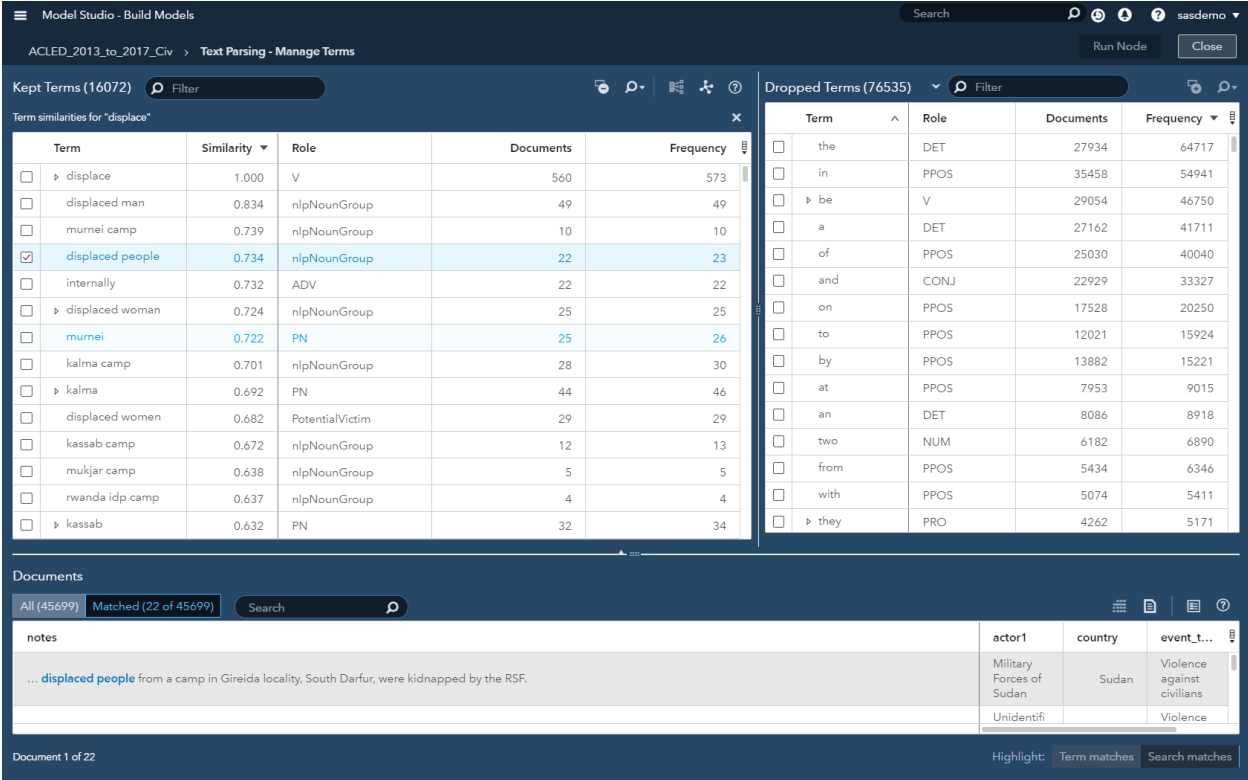
In summary, we showcased how data from multiple sources can be analyzed using AI and machine learning methods, and subsequently visualized and related to each other. In the process, we showed how such methods assist **strategic** operations, partly through improved access to relevant data and connections, and partly through verifying assertions in one source against others. Furthermore, we indicated how the use of such a system can help augment existing **data models** and generate new ones from purely unstructured data sources (in this case, the TIP reports). Finally, this process highlighted the **integration** of these data sources, and we will explore environments suitable for investigation further in this paper as well when we review the role of SAS Visual Investigator in this process.

## TARGETING VICTIMS OF HUMAN TRAFFICKING IN FREE-FORM TEXT

Text analytics, machine learning, and deep learning capabilities can be used to generate models that flag victims of human trafficking in freeform text. Such models can be trained to look for appropriate context, and subsequently flag conditions that cause human trafficking vulnerabilities in regions. This would enable analysts to better target conditions surrounding human trafficking victims, make recommendations, and act. In this section, we will explore how to apply AI capabilities against a set of data to build up a rule-based approach to flagging potential human trafficking vulnerabilities.

A rules-based approach to flagging human trafficking victims can be based on lists of classifier terms that meet criteria. For example, I might know that when the term civilians, women, or village is present in the ACLED data, that is usually indicative of some form of violence against civilians, creating a human trafficking vulnerability in the affected region. SAS Visual Text Analytics provides AI capabilities to assess term similarity, that is, terms used in a similar context. As shown in Figure 5 below, I can use terms representing known targets to derive other terms. In this case, I show term similarities for “displace” to identify displaced individuals, nationalities affected by displacement, and violence affecting certain

displaced camps. In this semi-automated manner, we can grow our list of relevant terms from my initial set of terms, to a much more complete set in a short amount of time.



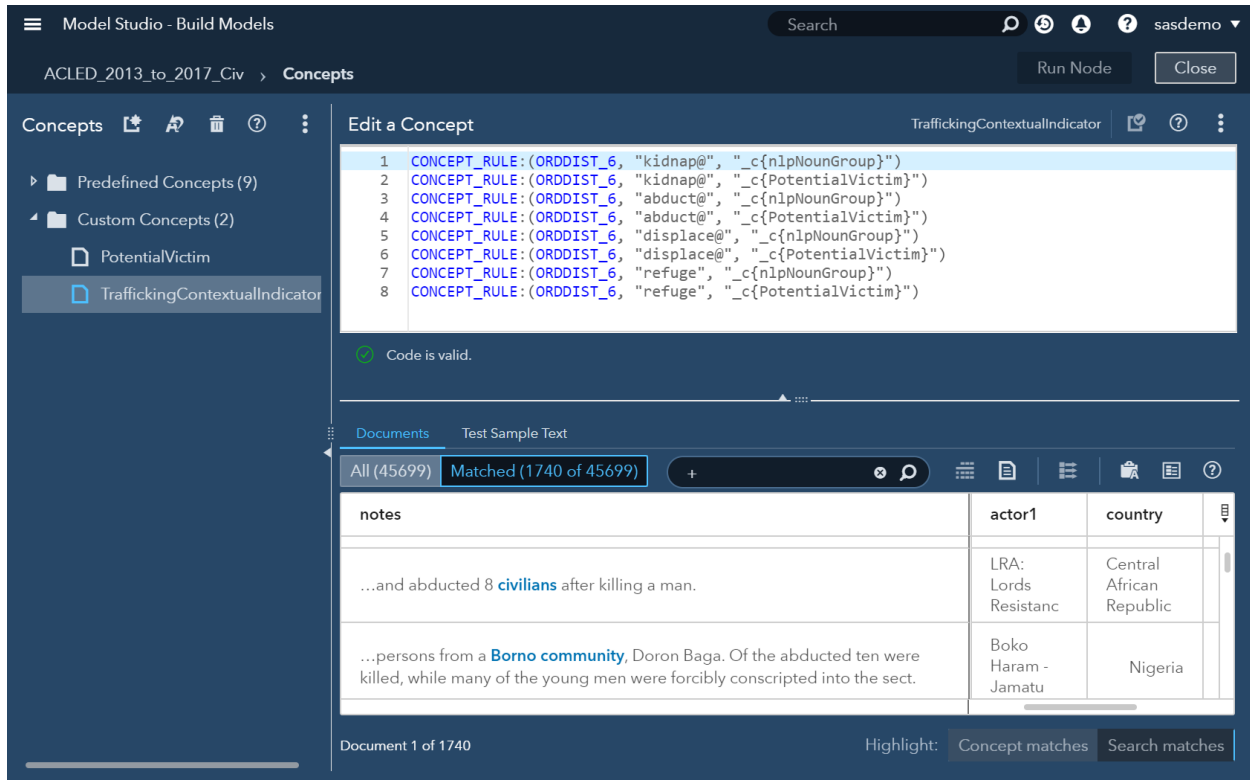
**Figure 5: Using term similarities in SAS® Visual Text Analytics to grow a classifier list of terms representing human trafficking vulnerabilities**

Following is a more complete set of classifier terms we developed using these methods, which can be used in to assess any source of armed conflict data for potential trafficking vulnerabilities.

- CLASSIFIER: civilians
- CLASSIFIER: several civilians
- CLASSIFIER: civilian home
- CLASSIFIER: civilian convoy
- CLASSIFIER: women
- CLASSIFIER: civilian family
- CLASSIFIER: displaced civilians
- CLASSIFIER: family
- CLASSIFIER: displaced family
- CLASSIFIER: village
- CLASSIFIER: villagers
- CLASSIFIER: entire village
- CLASSIFIER: refugee camp
- CLASSIFIER: refugee shelter
- CLASSIFIER: refugees
- CLASSIFIER: children
- CLASSIFIER: women
- CLASSIFIER: young women
- CLASSIFIER: local women
- CLASSIFIER: displaced woman
- CLASSIFIER: displaced women

Furthermore, identifying noun groups and terminology related to potential victims in the context of trafficking-related actions (kidnaping, abduction, displacement, seeking refuge) is indicative of trafficking vulnerabilities as well. Figure 6 below shows such rules in SAS Visual Text Analytics, flagging the classifier terms listed above or any noun groups within 6 terms of a trafficking-related action. The rule depicted in Figure 6 utilizes an operator called an ORDDIST to accomplish this, which looks for terms or entities in the specified order within a number of terms (in this case, 6). Such sophisticated rules help analysts move past

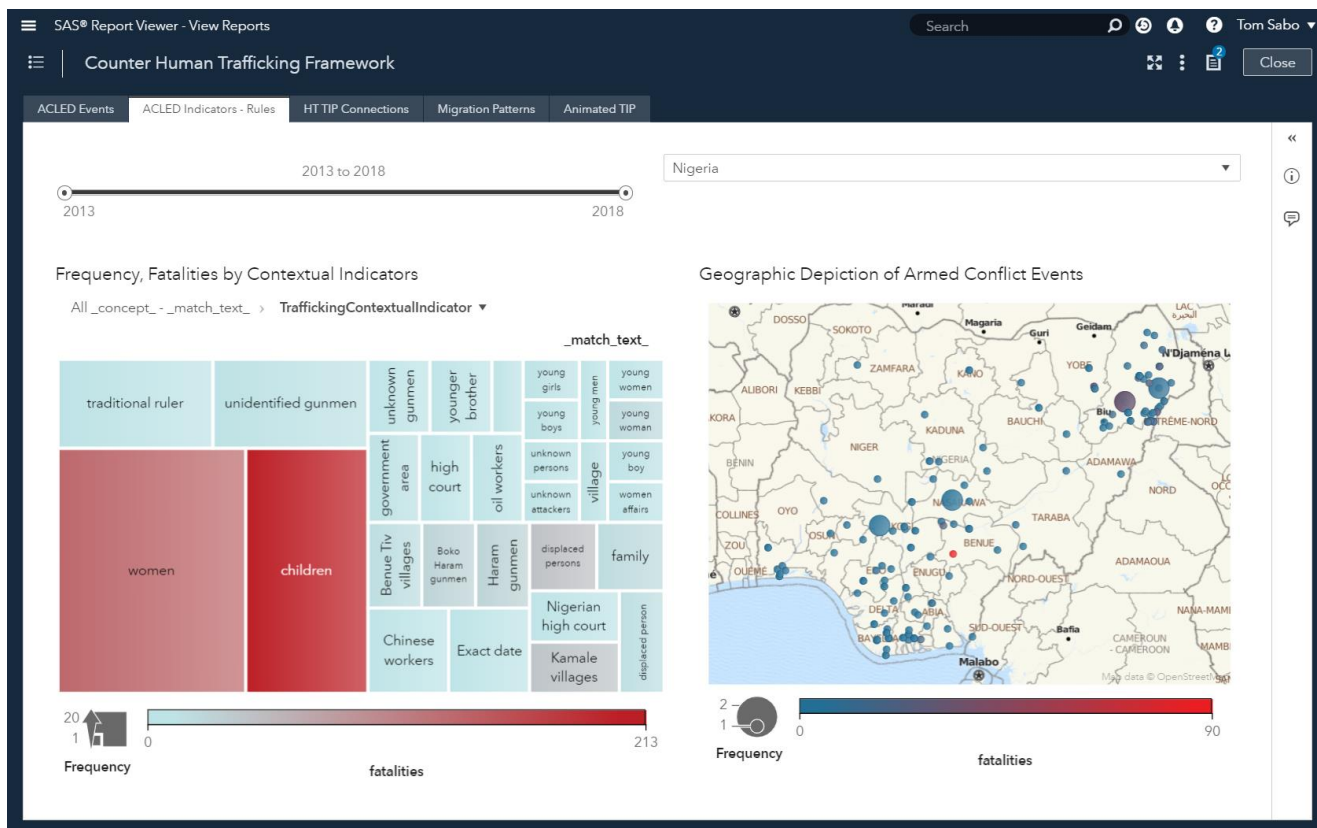
keyword search to only return matches in the correct context. These patterns can be also feed into a deep learning approach as we will discuss at the conclusion of this paper.



**Figure 6: SAS® Visual Text Analytics identifies victims in the context of trafficking-related actions**

We use SAS to score the larger data set with these victim-related concept matches, and they are subsequently available to review in the same visualization environment as we discussed in the previous section. Figure 7 below shows an example of this, focusing on events creating trafficking vulnerabilities in Nigeria. This visualization differs from Figure 3 in that it focuses exclusively on the various types of victims. The rule identifies some expected victims, including women and children, but also identifies Chinese workers, Kamale villages, displaced persons, and oil workers as indicators to trafficking vulnerabilities or potential targets of trafficking. It also highlights a high number of fatalities related to attacks involving children in this region.





**Figure 7: Geospatial Visualization of contextual trafficking vulnerabilities in SAS® Visual Analytics**

This exercise gives the analyst new insight into the ways in which individuals, families, and communities are victimized by violent circumstances, as well as answer where and how these actions take place. For instance, while armed conflict events are generally dispersed across Nigeria, most of the armed conflict events involving children take place in the northeast region of Nigeria. Such analysis improves the **strategic** value of the data, as well as provides answers to **operational** and **tactical** questions, such as assessing the impact of a militant group in creating trafficking vulnerabilities. Finally, this assessment of trafficking vulnerabilities adds to the **data model** of information surrounding the trafficking problem. In the conclusion section at the end of this paper, we will overview how to use this rules-based approach to generate training data for machine learning and deep learning models.

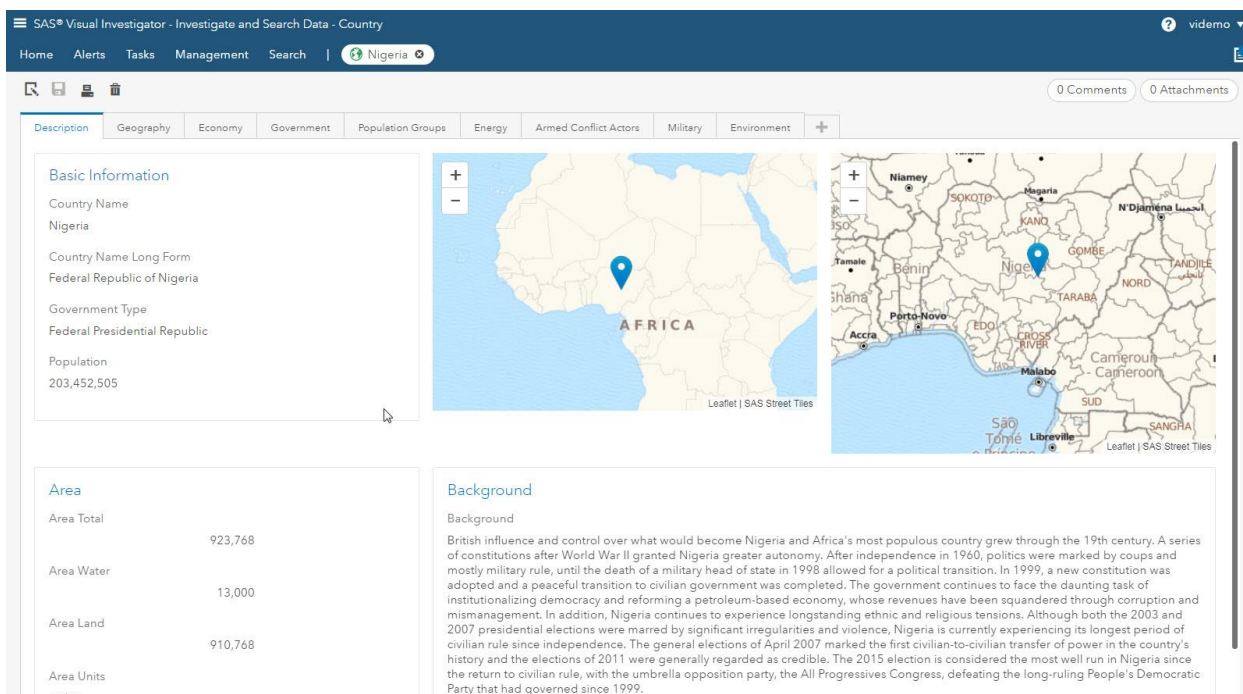
## VISUAL INVESTIGATOR: TARGETING THE WHERE, WHAT AND HOW

The previous sections focused on the strategic value of the data. Counter trafficking organizations and coalitions also need to better assess how armed groups are acting in a region so they can predict future events and take appropriate countermeasures. There are many questions an investigator could ask of this data to better make an operational or tactical decision. These include:

- Who are the armed groups that are operating in a country?
- What are the different methods of attack that an armed group uses?
- How do these methods of attack contribute directly or indirectly to human trafficking?

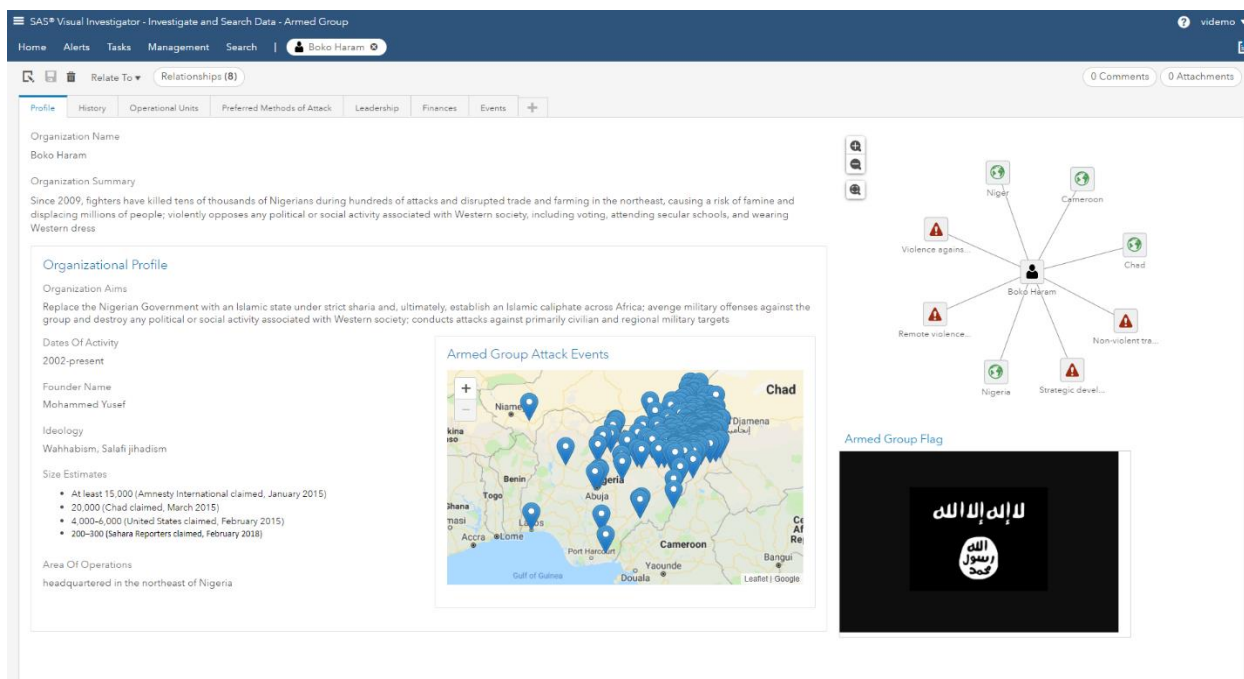
Answering tactical questions of the data relies on multiple sources and analysis on those sources. In the previous sections we showcased how to incorporate multiple data sources related to counter human trafficking into one visualization environment, and how to apply text analytics to target indicators of human trafficking conditions in freeform text. The next step is to present this in a framework which can logically assist analysts in an investigation. To do so, we offer *SAS Visual Investigator*, and will walk through an example workflow that helps to address the above three questions, as well as help assess veracity of cited sources.

A landing page for investigation would enable the analyst to assess the overall situation in a country. For the purposes of example here, we focus on Nigeria, using information drawn from the CIA World Factbook<sup>9</sup> and Wikipedia<sup>10</sup> as well as data from the ACLED project. The landing page would provide orientation information including location, population, area, and provide the ability to drill down in areas such as economy, government, population groups, and armed conflict actors, as shown in below in Figure 8 below.



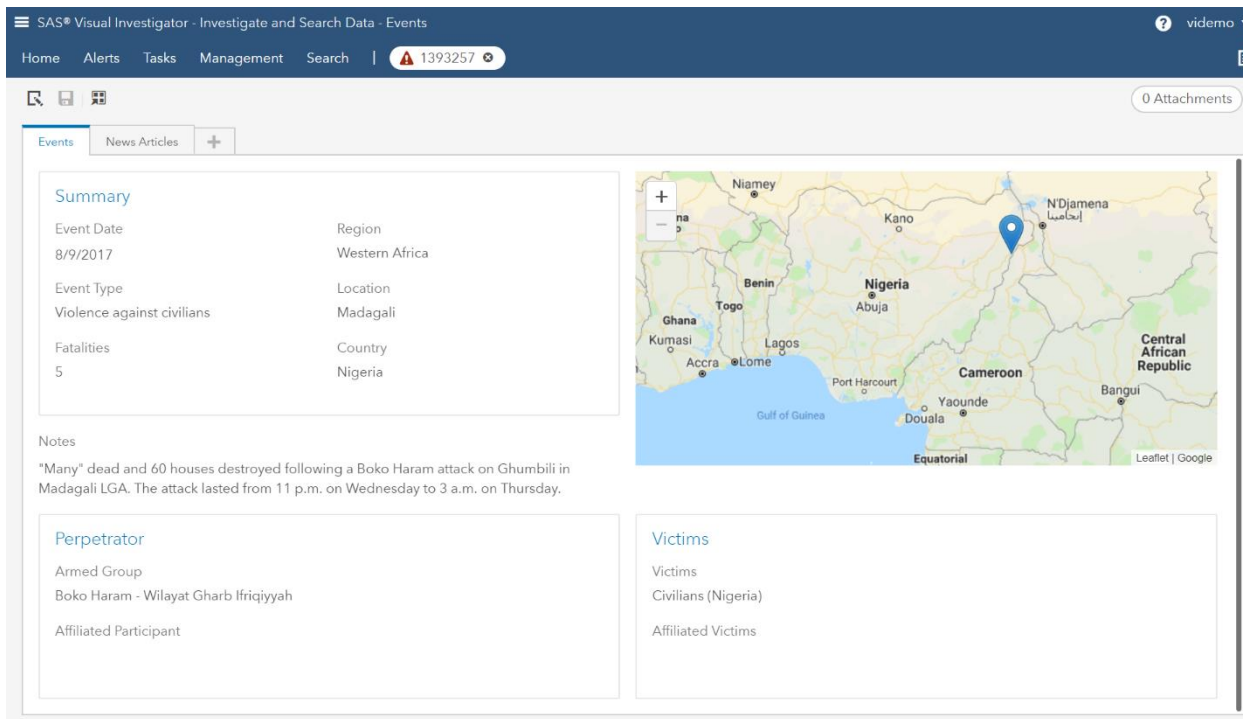
**Figure 8: Investigation Landing Page for Nigeria in SAS® Visual Investigator**

A user could choose to drill down from here into any of the armed conflict groups, which are impacting the stability of Nigeria and influencing the trafficking situation through the Armed Groups Tab in Figure 8 above. For example, this enables exploration of Boko Haram as depicted in Figure 9 below. At this level, we draw structured and unstructured information from all available sources to show how Boko Haram operates across national borders, influencing Chad, Cameroon and Niger in addition to Nigeria via a basic link chart diagram, including the types of violence that Boko Haram is associated with. We also display pertinent information including the current flag, area of operation, and size estimates. Finally, all associated armed conflict events are marked on a map and are available for further drilldown. As an example, use of this investigative capability, the geographic events could be compared to economic data to answer the question of whether more recruits for Boko Haram are coming from the north/remote areas of Nigeria. It could also answer whether counter trafficking efforts are more successful in wealthier areas, and whether counter trafficking measures should now be focused on the northeast region of Nigeria.



**Figure 9: SAS® Visual Investigator Drilldown into Armed Conflict Actor - Boko Haram**

Finally, we can drill down into individual events, as shown in Figure 10 below. This enables the investigator to assess what happened in the event, including attack vectors such as village attacks where houses are destroyed, and assess the impact of such actions on making an area vulnerable to human trafficking. This can answer questions such as whether or not the displacement of villagers from such attacks is widespread or isolated.



**Figure 10: SAS® Visual Investigator Drilldown into a Specific Armed Conflict Event in Nigeria**

Boko Haram is not the only factor affecting instability in Nigeria. Using such an interface, one can assess whether Boko Haram is the primary or a secondary factor influencing human trafficking in and from Nigeria. Further, it can help assess whether sources of intelligence data are reliable. While using this framework to assess events, we turned up information that is inaccurate. We identified a photo in a recent news article depicting an older version of the Boko Haram flag than is currently in use. This is another area where deep learning AI capabilities in computer vision, object detection, and recognition could provide automated assistance. This could serve to not only detect the flag but identify a mismatch in the date of the report and the flag depicted in the photograph, effectively identifying this as a stock photo.

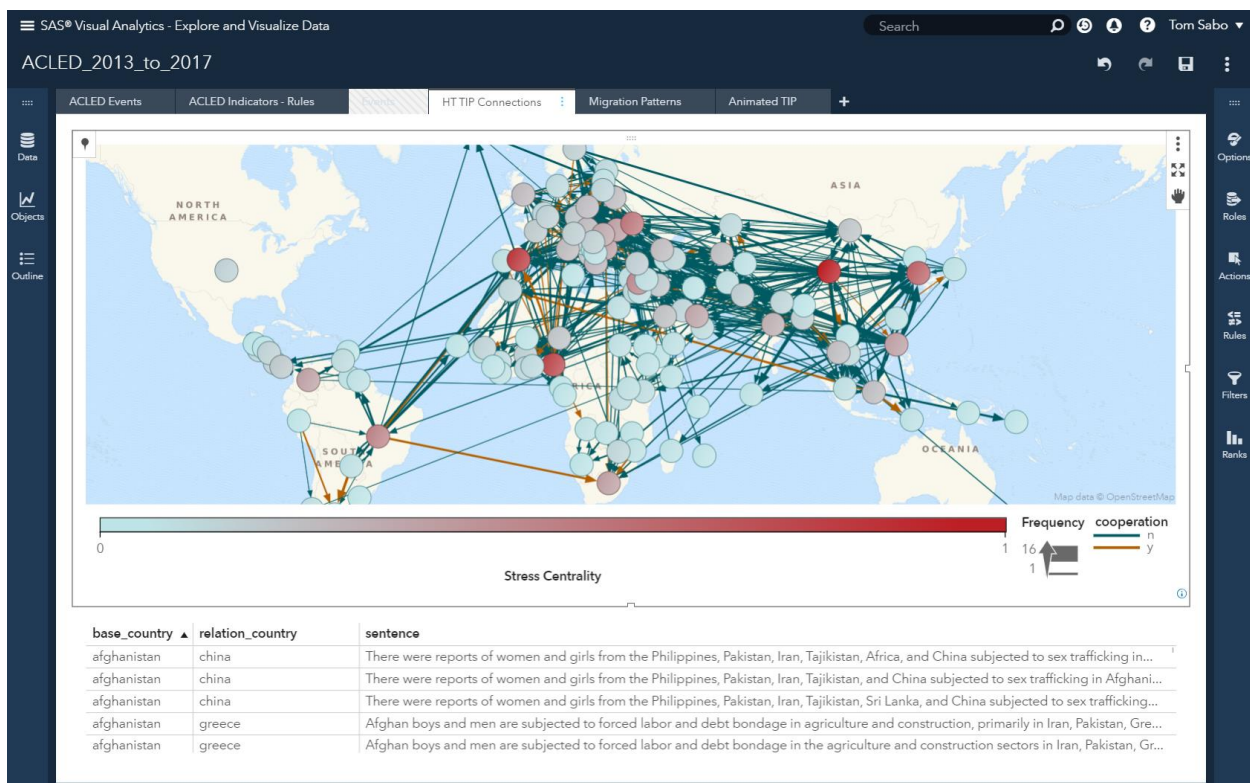
In summary, such capabilities improve the **data model**, enabling organization and categorization in a way that assists investigations, as well as providing an environment for **integrating** new sources of data to assist with the **operational/tactical** responses to international human trafficking.

## CONCLUSION

This paper explored how to incorporate multiple data sources into a single visualization environment. This included how to apply analytics to those sources of data, as well as how to better target indicators of human trafficking conditions in freeform text. The visualizations offered both an analytical and investigative oriented perspective on the data. By answering these questions, we demonstrated how such a solution could be leveraged **strategically** to improve situational awareness of international human trafficking. We also showed how to apply it **tactically/operationally** to influence specific response in trafficking areas such as Nigeria. Finally, such a solution supports **integrating** multiple data sources into **data models** suitable for investigative and strategic work, as well as extending the value of the data through AI, structured analytics, and text analytics.

As stated in the introduction, data collection to support these counter trafficking efforts is on the rise<sup>2</sup>. Furthermore, nearly every country now has legislation in place criminalizing human trafficking. Counter trafficking partnerships are in development as well. For example, in 2017, the Governor of Edo State in Nigeria declared human trafficking to be one of his top priorities and created the Edo State Task Force to combat trafficking in persons. This includes NGO participants, national agencies that prohibit Trafficking in persons, immigration services, and numerous other stakeholders<sup>1</sup>. This is good news. However, as the international community accelerates progress to build capabilities and cooperation<sup>2</sup>, applying analytical capabilities to support these efforts is paramount. As any country implements counter trafficking laws, analysts can quantitatively measure how effective those laws are through a system like what is proposed in this paper. Such capabilities will support the task-forces and interagency partnerships as well by giving a strategic and tactical view into the data supporting counter trafficking initiatives.

Regarding analytics and AI capabilities that can be directed to these efforts, this paper is just scratching the surface. For example, a variety of network metrics comes out-of-the-box with the network diagram visualization capability available in SAS Visual Analytics. Metrics include "Stress Centrality" that indicates how frequently a node would be crossed when taking the shortest path between nodes<sup>11</sup>. As shown in Figure 11 below, this highlights areas in red as more central to the international community for human trafficking, including Nigeria, Spain, China, and South Korea.



**Figure 11: SAS® Visual Analytics Network Diagram of TIP Connections Highlighting Stress Centrality**

There is more to discuss regarding the insertion of AI capability into this framework. In this paper, we explored targeting victims of human trafficking in freeform text. In the process, we developed a rules-based approach to assess victims in particular contexts using SAS methods in Visual Text Analytics. The rules developed from this approach can feed machine learning or deep learning models for entity extraction purposes. The challenge we are

addressing is that labeled training data for the purposes of named entity extraction (NER) modeling is laborious to manually generate, particularly for recurrent neural network algorithms, which require extensive training data to optimize. Using a rules-based approach, we can score thousands of examples sentences to flag entities, such as victims of human trafficking, in context, and in seconds. By feeding these example instances into a RNN (Recurrent Neural Network) deep learning model or a CRF (Conditional Random Fields) machine learning model, it is likely that these models would identify additional instances of trafficking victims not covered by the original rules-based approach, improving the breadth of what the models are able to extract and flag as victims of human trafficking<sup>12</sup>. Such an approach would extend the AI-oriented capabilities of the counter human trafficking framework.

The addition of text analytics capabilities to the Visual Investigator interface would assist with the tactical response. This would involve flagging human trafficking victims in context or being able to review the articles and events that are specifically associated with potential human trafficking. A risk score could be generated for each event based on structured data available and the narrative, with emphasis on events that result in kidnapping or displacement.

This counter trafficking framework can be generalized to contraband movement, drug trafficking, counter terrorism, and defense purposes. At the national, state, and local level, it can be generalized to law enforcement for victim identification and criminal prosecution. This is an area of future work, demonstrating in a concrete manner how text analytics and artificial intelligence can help flag indicators of potential human trafficking in law enforcement incident data. This can help identify situations where individuals have been trafficked and need resources to help them escape a life of trafficking. Automated models run against textual data can provide a huge lift in terms of time-to-value versus manual review alone.

In summary, this paper shows how a variety of data sources that inform human trafficking can be enhanced, integrated and assessed in an analytical and investigative way. First, this framework evolves the assessment of patterns to assist the strategic response. In addition, these capabilities enhance organizations' abilities to take tactical and operational actions. With these analytical applications, we seek to move the ball just a bit further in countering human trafficking at all levels. It is our aim that organizations that fight human trafficking, be it local or international, will leverage these capabilities to rescue those who have been trafficked, and counter those who would engage in modern day slavery.

## REFERENCES

1. U.S. Department of State. 2018. "Trafficking in Persons Report." Accessed January 7, 2018. <https://www.state.gov/j/tip/rls/tiprpt/2018/>.
2. United Nations Office on Drugs and Crime. 2018. "Global Report on Trafficking in Persons." Accessed January 7, 2019. <https://www.unodc.org/unodc/en/data-and-analysis/glotip.html>
3. Institute for International Research on Criminal Policy, 2004, The Siamsect Files: Standardised Templates and Blueprint for EU-wide Collection of Statistical Information and Analysis on Missing and Sexually Exploited Children and Trafficking in Human Being; International Organization for Migration and Austrian Federal Ministry of Interior, 2009, Guidelines for the collection of data on trafficking in human beings, including comparable indicators; International Labour Office and the European Commission, 2009, Operational indicators of trafficking in human beings; Results from a Delphi survey; International Centre for Migration Policy Development,

2009, Handbook on Anti-Trafficking Data Collection in South-Eastern Europe: Developing Regional Criteria.

4. Gupta, Saurabh. 2018. *Artificial Intelligence with SAS; Special Collection*. Cary, NC : SAS Institute Inc. <https://support.sas.com/content/dam/SAS/support/en/books/free-books/artificial-intelligence-with-sas.pdf>
5. ACLED Data; Bringing Clarity to Crisis. 2019. "About". Accessed January 7, 2019. <http://www.acleddata.com/>.
6. Sabo, Tom. 2016. "Extending the Armed Conflict Location and Event Data Project with SAS® Text Analytics." Proceedings of the SAS Global Forum 2016 Conference. Cary NC: SAS Institute Inc. Available <https://support.sas.com/resources/papers/proceedings16/SAS6380-2016.pdf>.
7. Sabo, Tom. 2018. "Using SAS® Text Analytics to Assess International Human Trafficking Patterns." Proceedings of the SAS Global Forum 2018 Conference. Cary NC: SAS Institute Ince. Available <https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/1986-2018.pdf>.
8. The UN Refugee Agency. 2019. "Persons of Concern; Population Statistics". Accessed January 7, 2019. [http://popstats.unhcr.org/en/persons\\_of\\_concern](http://popstats.unhcr.org/en/persons_of_concern).
9. Central Intelligence Agency; The World Factbook. "Africa::Nigeria". Accessed February 22, 2019. <https://www.cia.gov/library/publications/the-world-factbook/geos/ni.html>
10. Wikipedia; The Free Encyclopedia. "Boko Haram". Accessed February 22, 2019. [https://en.wikipedia.org/wiki/Boko\\_Haram](https://en.wikipedia.org/wiki/Boko_Haram)
11. SAS Visual Analytics 8.3; Working with Report Content . "Working with Network Analysis Objects". Accessed February 25, 2019. <https://go.documentation.sas.com/?cdcId=vacdc&cdcVersion=8.3&docsetId=vaobj&docsetTarget=n0kcvb9vm0kd1sn1b199uu8hgq1w.htm&locale=en>
12. SAS Blogs. "Reduce the cost-barrier of generating labeled text data for machine learning algorithms." Accessed February 25, 2019. <https://blogs.sas.com/content/sgf/2019/01/16/generating-labeled-text-data-for-machine-learning-algorithms/>

## ACKNOWLEDGMENTS

Much collaboration went into the development of this paper. Thanks to Bryan Goodliffe for driving the effort of generating the data model and visualizations for the investigative framework, and for helping with honing the message of what a counter human trafficking platform can deliver. Thank to Christina Hsiao and Adam Pilz for their pioneering work on using rule-based text extraction to feed machine learning CRF and deep learning RNN models. Thanks to Emily McRae for ongoing text analytics oriented review of methods which went into the machine learning models and the rule-based approach. Thanks also to Kelsey Young for recommendations on network diagram metrics available in SAS Visual Analytics, which provide insight into the TIP country connections.

## RECOMMENDED READING

- Jade, Teresa. Wilsey, Biljana Belamaric. Wallis, Michael. 2019. *SAS Text Analytics for Business*

*Applications: Concept Rules for Information Extraction Models.* Cary, NC: SAS Institute Inc.

- Sabo, Tom. 2014. SAS Institute white paper. “*Text Analytics in Government: Using Automated Analysis to Unlock the Hidden Secrets of Unstructured Data.*” Available [http://www.sas.com/en\\_us/whitepapers/text-analytics-in-government-106931.html](http://www.sas.com/en_us/whitepapers/text-analytics-in-government-106931.html).
- Chakraborty, G., M. Pagolu, S. Garla. 2013. *Text Mining and Analysis; Practical Methods, Examples, and Case Studies Using SAS®.* Cary, NC: SAS Institute Inc.
- Sabo, Tom. 2017. “*Applying Text Analytics and Machine Learning to Assess Consumer Financial Complaints.*” Proceedings of the SAS Global Forum 2017 Conference. Cary, NC: SAS Institute Inc. Available <http://support.sas.com/resources/papers/proceedings17/SAS0282-2017.pdf>
- Reamy, Tom. 2016. “*Deep Text; Using Text Analytics to Conquer Information Overload, Get Real Value from Social Media, and Add Big(ger) Text to Big Data.*” Medford NJ: Information Today, Inc.

## CONTACT INFORMATION

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

Tom Sabo, Principal Solutions Architect  
1530 Wilson Blvd.  
Arlington, VA 22209  
SAS Federal LLC  
+1 (703) 310-5717  
[tom.sabo@sas.com](mailto:tom.sabo@sas.com)  
@mrTomSab

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.