International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(REVIEW ARTICLE)

Check for updates

Emotion analysis based on belief of targeted individual supporting insider threat detection

Jason Slaughter *, Carole E. Chaski, and Kellep Charles

Capitol Technology University, 11301 Springfield Rd, Laurel, MD 20708, United States.

International Journal of Science and Research Archive, 2024, 11(02), 226–237

Publication history: Received on 21 January 2024; revised on 03 March 2024; accepted on 06 March 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.11.2.0393

Abstract

Unintentional Insider Threat is the concept that an insider threat event may occur unintentionally versus maliciously. Individuals who believe they are being targeted may be at increased risk of being insider threats. Based on a previous survey titled A Survey of Unintentional Medical Insider Threat Category, it was found that both medical and psychological problems may lead to feeling targeted. It was further found that Insider Threat programs should be updated to include trained personnel in both medicine and psychology in addition to cybersecurity to address the risk.

Keywords: Insider; Threat; Detection; Targeted; Medical

1. Introduction

This article builds off the original research in the peer-reviewed article A Survey of Unintentional Insider Threat Category, published at https://ijcsit.com/ijcsit-v14issue2.php.

The background is individuals who believe they are being targeted and may become insider threats.

The focus is on individuals who may have medical problems over psychological problems or detecting actual real-world targeting occurring.

However, this study could consider either of those two cases.

To expand on the first author's background and experience, it was initially thought that hypervigilance caused actions and reactions in their behavior.

However, after further investigation and many medical tests and brain scans, it was determined to be a combination of two sleep disorders resulting in various known and documented symptoms. The life experience ultimately led to performing this series of studies.

The first author found that many insider threat programs do not appear to be staffed by a multidisciplinary group capable of detecting an unintentional insider threat incident that medical symptoms may cause.

This led to the thought that several areas of detection and response could be improved within insider threat programs.

During the onset of the initial symptoms, the first author believed they were being targeted or that a family member was being targeted due to initially hearing an individual in Target state a family member's name. Then, “We have her

* Corresponding author: Jason Slaughter

phone." There were several other apparent indicators. Some indicators were spoken out loud near the first author, and others appeared in cyberspace. This led to the first author posting the indicators on social media to get eyes on whatever was occurring. This activity persisted for several days until everyone convinced the first author that it wasn't real and only occurring in the author's head. For years, the first author believed this to be correct until, out of the blue, it happened again.

The second occasion occurred in June of 2023, with various words and phrases appearing or being spoken in context to the first author. The first one was the word rock. The word didn't initially have any context or meaning for the first author. Until it was passed around across multiple networks and ended with the rock-on emoji symbol. Similar words and phrases were used over the months, and then individual names the first author was familiar with were used. In this instance, the first author reacted differently. However, everything was documented, and a bag of words was created. A report was generated and passed on to be reviewed to determine if there was a cyber actor.

Also, the first author's LinkedIn page appeared to be hacked, and the hacker wrote, "You coder now," on the account. Also of interest was what appeared to be cell phone hacking, which the first author also noted and reported.

Of those reports, nothing has occurred or been briefed back to the first author regarding the incident.

However, once the report was sent, the activity stopped across social media platforms, and people speaking near the first author also stopped.

Since then, the first author has moved across the country. On the second day in the new house, the router was factory reset by a rogue Apple device that was ethernet-connected according to the router. However, on inspection of the router, no device was physically connected. The first author performed an additional factory reset on the device and then locked down all the security settings.

In January 2024, the first author underwent additional psychological testing with Veterans Affairs. It was determined that the events were not psychological, and the likelihood of a cyber actor being in play increased significantly.

To date, it is undetermined who the cyber actors or the physical actors are. Various scenarios are possible, but the first author won't delve into those currently.

1.1. Purpose of the Study

The study aims to further the knowledge of the detection of insider threats and to open a new area of study where individuals with medical problems may also be considered in insider threat detection versus the current approaches.

1.2. Theoretical Framework

The theoretical framework of this paper is based on the detection of personality and behavior, which is widely accepted for measuring a person's reaction. The authors used this model to develop profiles for insider threats. They also employed natural language processing to analyze personal content and detect potential insider threats.

1.3. Research Questions

- RQ1: What is the individuals' response when they feel targeted?
- RQ2: Does the response to RQ1 fit into a known behavior for insider threat detection?
- RQ3: Does the data analysis provide sufficient information to determine if the individual is suffering from a medical or psychological condition, or is this an actual real-world event?

1.4. Nature of the Study

The nature of the study is to use the ALIAS system of machine learning solutions to analyze Twitter data for individuals who believe they are being targeted and may become insider threats.

1.5. Significance of the Study

This study furthers the research into the detection of insider threats. It focuses on potential insider threats with a medical problem, causing them to believe they are being targeted. This study does not analyze psychological events or real-world targeting events that may be occurring. However, the methodology used may be further refined for those areas of study.

2. Literature Review

The concept of unintentional insider threats refers to individuals who inadvertently pose a risk to an organization's security or sensitive information. These individuals may not have malicious intent, but their actions or behavior can still lead to unintended consequences [1]. Unintentional insider threats are one of the three categories of insider threats, with the other two being traitors and masqueraders [1].

According to a foundational study on unintentional insider threats, they are defined as current or former employees who unintentionally cause harm to an organization's security [2]. Understanding the nature of unintentional insider threats is essential to detect and mitigate their potential risks effectively.

Sentiment analysis, also known as opinion mining, is a technique used to analyze and determine the emotional tone behind a piece of text. It involves using natural language processing and machine learning algorithms to identify and classify sentiments expressed in the text, such as positive, negative, or neutral [4].

In detecting insider threats, sentiment analysis can be applied to analyze the sentiment of publicly available information, known as Open-Source Intelligence (OSINT), to identify potential indicators of insider threats [5].

By analyzing the sentiment of OSINT data, organizations can gain insights into individuals' emotional states and detect any signs of disgruntlement, dissatisfaction, or potential malicious intent [3].

Table 1 describes some known medical problems that may cause individuals to feel targeted and lead to an unintentional insider threat event.

Table 1 Table of Medical Problems of Unintentional Medical Insider Threats

Medical Condition	Description
Epilepsy	A neurological disorder characterized by recurring seizures.
Encephalitis	Inflammation of the brain can cause seizures and other neurological symptoms.
Meningitis	Inflammation of the membranes surrounding the brain and spinal cord, which can cause seizures and other neurological symptoms
Stroke	Disrupting blood flow to the brain can cause seizures and other neurological symptoms.
Traumatic Brain Injury	An injury to the brain caused by an external force can cause seizures and other neurological symptoms.
Alcohol Withdrawal	Abrupt cessation of alcohol consumption can cause seizures and other neurological symptoms.
Brain Tumors	Abnormal growths in the brain can cause seizures and other neurological symptoms.
Sleep Disorders	A broad category that encompasses several symptoms related to sleep and wakefulness. It is known to cause anxiety, paranoia, and hallucinations across all five senses when left undiagnosed and untreated.

3. Research Methodology

The research method will be natural language processing using the ALIAS [6] suite of tools to analyze Twitter data for individuals who believe they are being targeted. Then, further analysis will determine if there is potential for them to be an Insider Threat and if this technique can be used for Insider Threat detection.

3.1. Population

The participants of the research will be anonymous posts of Twitter data. All identifying information will be removed.

3.2. Sample

The sample data comprises approximately four hundred and fifty thousand tweets collected from Open-Source Intelligence (OSINT) using standard keyword searches to determine whether the sample data was correct for the study. Specifically, the word targeted was selected, and the data was retrieved and paid for on Tweet Binder.

3.3. Materials/Instruments

The instrument will be ALIAS's [6] machine learning library of capabilities that will be used to analyze and determine how the research questions will be answered. ALIAS [6] includes a list of machine learning capabilities to be leveraged during the research.

Custom Python code will also be written as needed to clean and transform the dataset.

3.4. Data Collection and Data Analysis

Data was collected initially from Twitter. There are 450,000 Tweets in various languages. The English

Tweets have been split out for the initial analysis using a custom Python script. The Non-English Tweets will be used in future studies to continue this analysis.

The sentiment analysis data control set was selected from Twitter using "watched" to choose the 10,000 Tweets control set.

Additionally, a future study will analyze handwritten writing samples from an individual who thought they were being targeted.

To begin the analysis, the 450,000 tweets were converted to JSON format from an original Excel source file. This allowed the data to be loaded into the ALIAS system much faster than the initial load times found with the Excel file.

The following images capture how the data is loaded into the ALIAS system and analyzed.

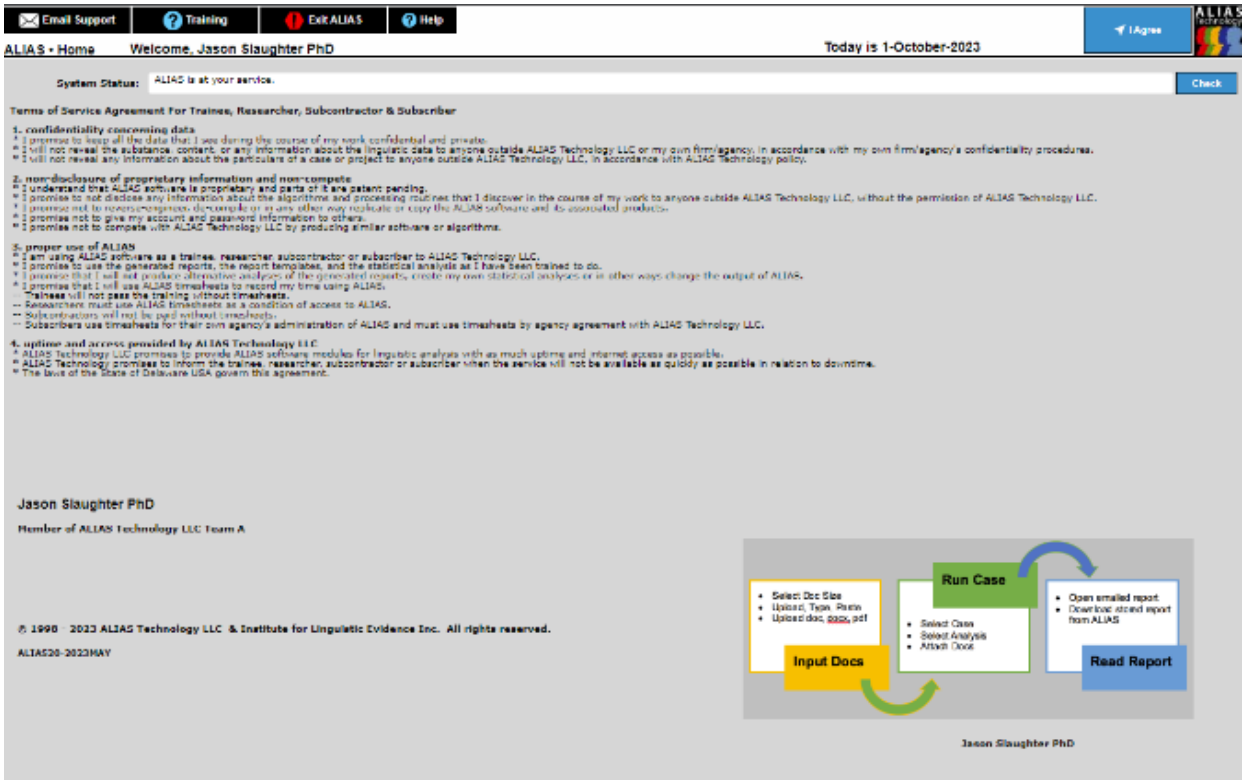

Figure 1 ALIAS Homepage

In Figure 1, the login to the ALIAS system was completed. From here, author one navigated to the main ALIAS functionality shown in Figure 2.

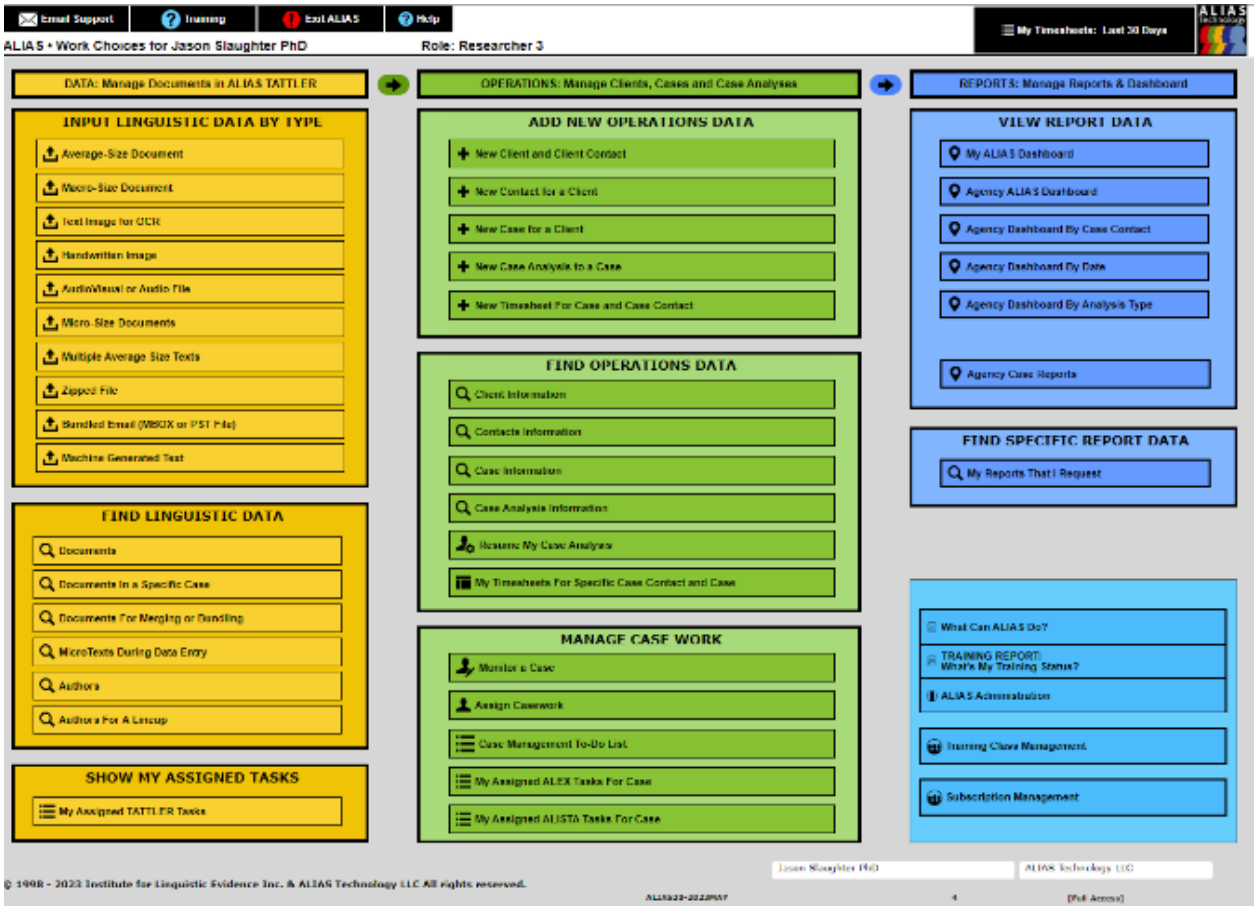

Figure 2 ALIAS Selections

From the selections in Figure 2 the micro text selection was chosen due to the tweets being considered micro texts by the ALIAS system.

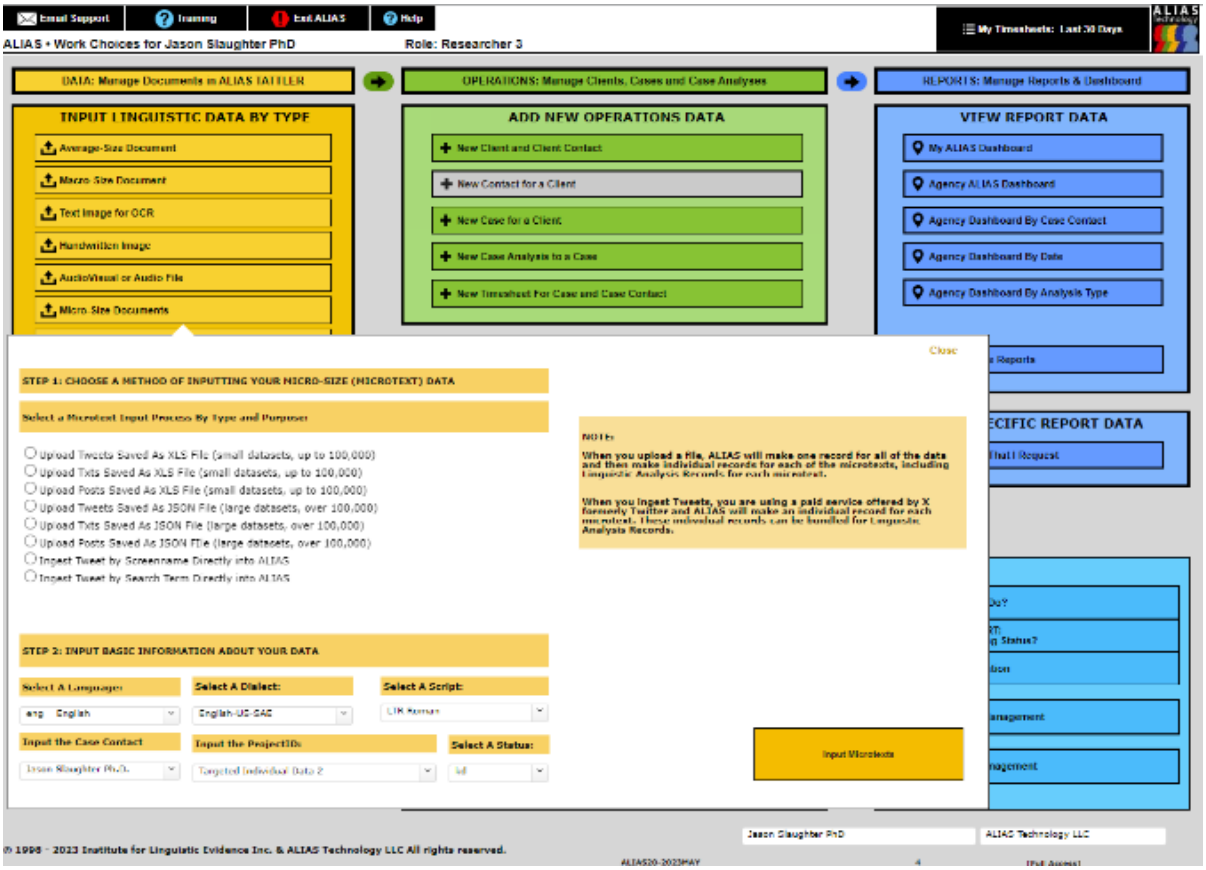

Figure 3 Choose Upload Type

In Figure 3 the upload tweets as JSON were selected due to the format of the tweets to be uploaded into ALIAS.

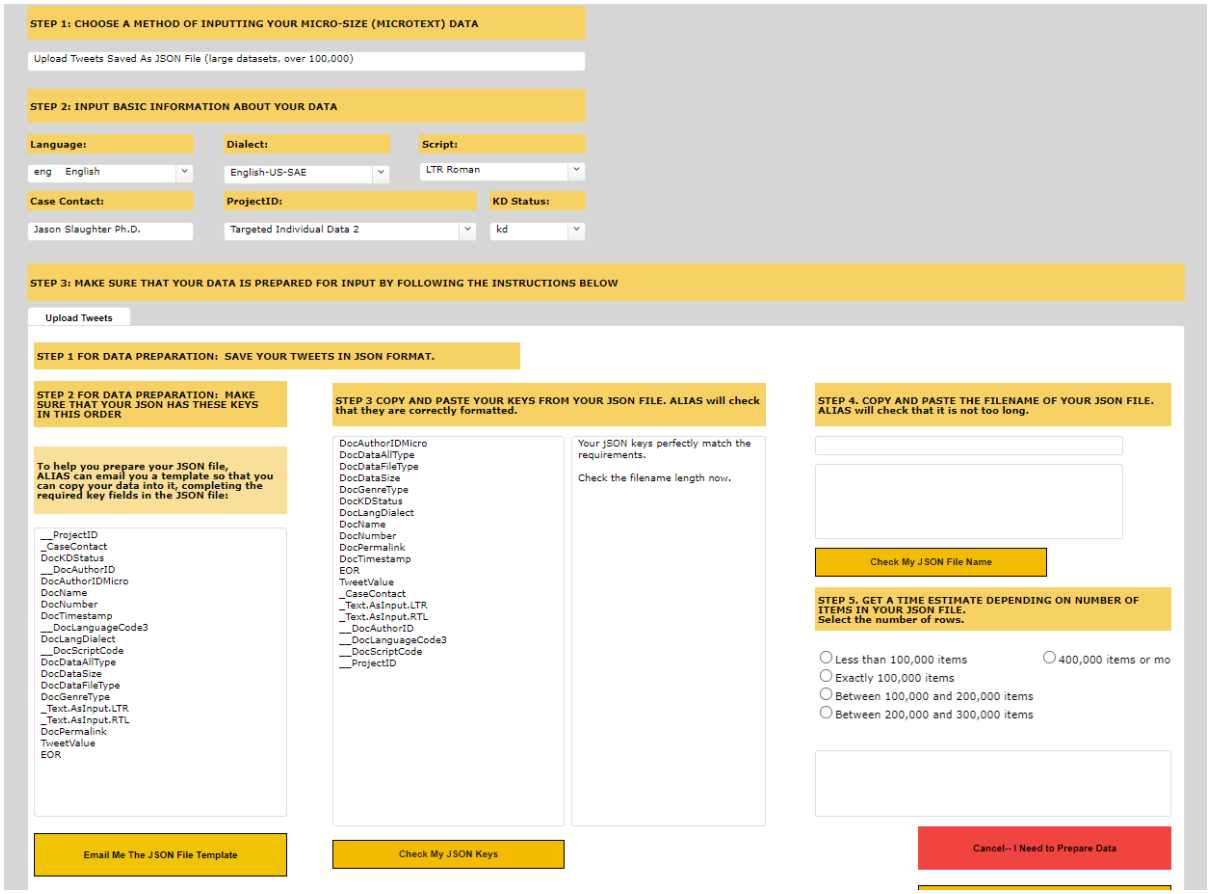

Figure 4 Check JSON Keys

In Figure 4 a check is performed to make sure the JSON keys match what ALIAS is expecting for the ingest of the data. Provided the keys all match during this step, the filename length check can be performed, and then the upload can continue. However, if this doesn't pass, the end user must update the JSON until it matches what ALIAS expects for the upload.

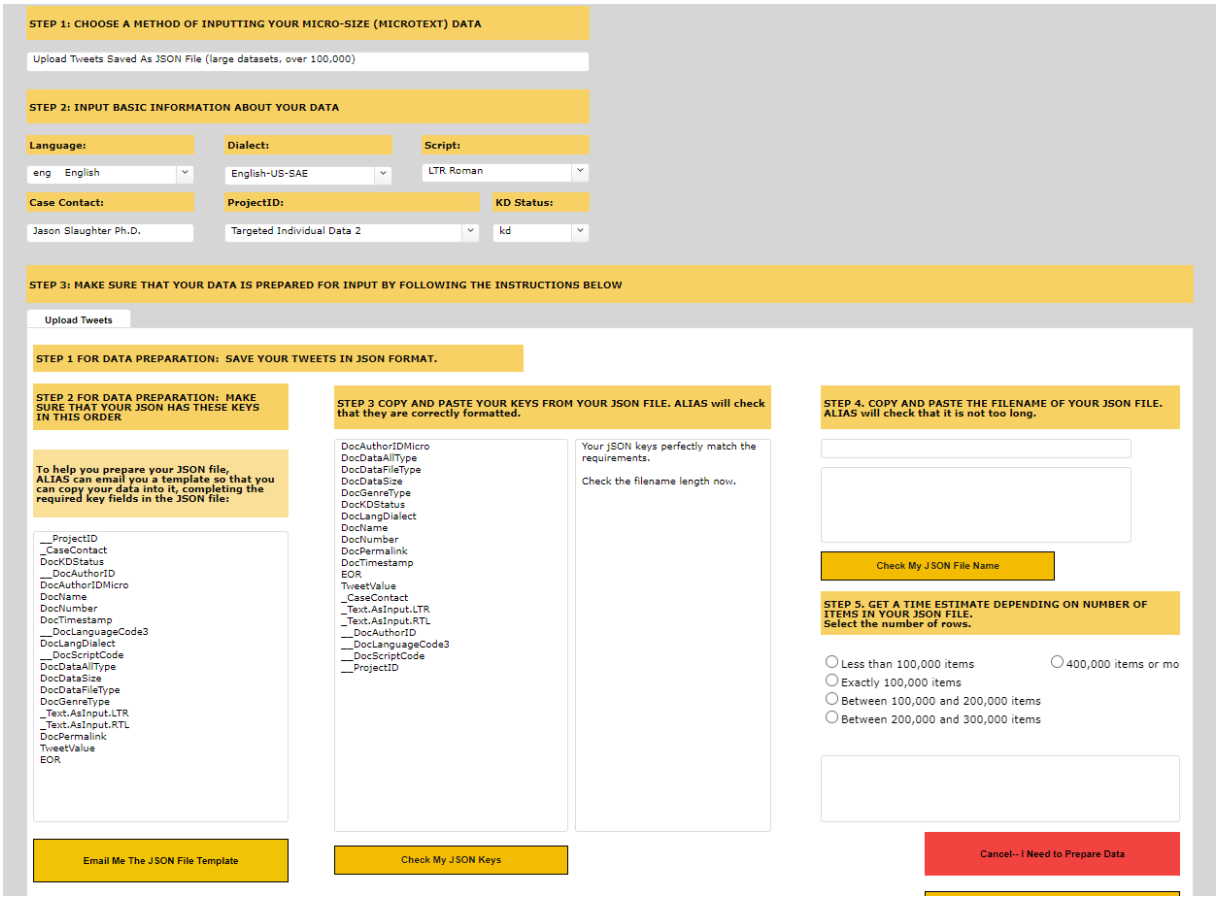

Figure 5 Check Filename Length

In Figure 5, the final check is performed to determine if the filename length is short enough for the system. If the file name is too long, the file must be renamed until the length is correct for ALIAS to allow the upload to continue.

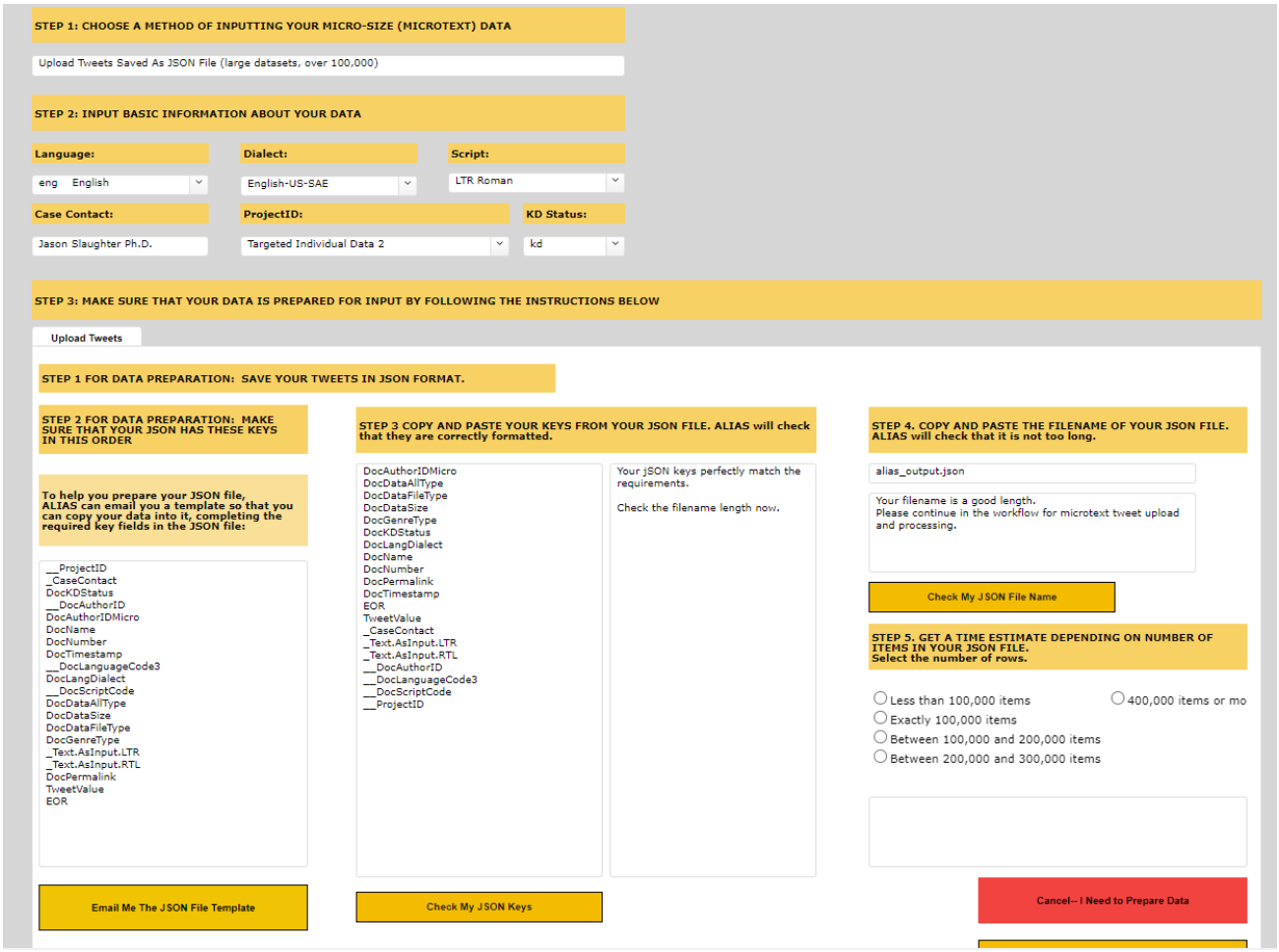

Figure 6 Prior to JSON Upload

In Figure 6, the user selects the JSON file that was previously validated to be uploaded into the system. Each of the darker yellow "Click" buttons is used to upload and then process the incoming data. Once the final step is complete on this page, ALIAS navigates the user to the main ALIAS functionality.

Table 1 Bag of Words for Targeted

Word	Frequency
targeted	456293
people	50506
new	29755
attack	28436

In Table 1, a Bag of Words analysis was performed using NLTK to generate a simple analysis of the word frequencies across the Tweet dataset. As expected, "targeted" has the highest count since it was the selector word for the dataset.

Table 2 Bag of Words for Watched

Word	Frequency
watched	10127
like	1065
one	818
movie	803
last	710
never	709
time	694
first	638

For the "watched" dataset, the same Bag of Words analysis was run to compare the two datasets. Based on the initial results, no words were commonly found when running the Bag of Words individually on the datasets.

Table 3 Emotion Analysis – Watched

Emotion	Count	Average Score
anger	3230	0.792957
fear	657	0.737411
joy	4714	0.868765
love	297	0.845870
sadness	862	0.869382
surprise	240	0.813566

In Table 3 an emotion analysis was performed on the full dataset for "watched" of ten thousand Tweets and then the average score per emotion was used along with the count of occurrences. Interestingly "joy" had the highest count for the "watched" dataset.

Table 4 Emotion Analysis – Targeted

Emotion	Count	Average Score
anger	6352	0.891600
fear	637	0.762074
joy	1659	0.875495
love	220	0.976948
sadness	1020	0.869003
surprise	112	0.956244

In Table 4, an emotion analysis was performed on ten thousand Tweets from the "targeted" dataset. Anger was the highest-occurring emotion in the "targeted' dataset.

Comparing the two results, those who talk about being watched appear to find joy when talking about it. However, individuals who talk about being targeted appear to be angry based on the initial results.

The research questions asked the following questions:

- RQ1: What is the individuals' response when they feel targeted?
- RQ2: Does the response for RQ1 fit into a known behavior for insider threat detection?

In the case of RQ1, the initial answer based on the emotion analysis of the data appears to be that the individuals feel anger when talking about the word targeted.

In the case of RQ2, emotion detection could be tied into insider threat programs.

3.5. Assumptions

The analysis assumes that RQ1 and RQ2 will be answered. RQ3 is more of an unknown, and the data may not be sufficient, or the instrument may not answer it.

Limitations

Sample data and data cleanliness limit the study. The sample only uses a subset of the four hundred and fifty thousand tweets. A larger sample size or alternative samples from other OSINT sources may improve the results.

3.6. Ethical Assurances

The data did not collect any identifying information about the respondents. This article has not captured or produced publicly identifiable information (PII) beyond the names in the acknowledgments section.

4. Conclusion

The ALIAS [6] analysis answered RQ1 and RQ2 sufficiently to continue the research using handwriting with alternative data samples. It was found that individuals who discuss the word "targeted" appear to exhibit anger in their writing based on the emotional analysis of the data. This can be detected and added to an insider threat program to support organizations in protecting themselves from insider threats.

RQ3 was not sufficiently answered with the available data. RQ3 will be moved to the following journal article in this series, where handwriting and additional Tweets will be used to determine if handwriting can be used to build a behavior model. Then the same behavior can be detected in the Twitter data.

This future work is planned to be published in an IEEE journal shortly after this article's publication.

Compliance with ethical standards

Acknowledgments

Jason Slaughter wishes to acknowledge the following individuals:

- Dr. Kellep Charles Ph.D.
- Dr. Carole E. Chaski, Ph.D.
- Supporting family.
- Supporting team members at MITRE.
- Brian Seborg
- Jay Fultz

The author's affiliation with The MITRE Corporation is provided for identification purposes only and is not intended to convey or imply MITRE's concurrence with or support for, the positions, opinions, or viewpoints expressed by the author.

The author's affiliation with The MITRE Corporation is provided for identification purposes only and is not intended to convey or imply MITRE's concurrence with or support for the positions, opinions, or viewpoints expressed by the author.

MITRE Case: 23-3877 (Independent Effort)

Disclosure of conflict of interest

No conflict of interest is to be disclosed.

References

[1] A Review of Insider Threat Detection: Classification, ... (n.d.) Retrieved August 5, 2023, from www.mdpi.com/2076-3417/10/15/5208

[2] Protect Against Unintentional Insider Threats: The risk of an .... (n.d.) Retrieved August 5, 2023, from arxiv.org/pdf/2103.04744

[3] -Analysis of Unintentional Insider Threats Deriving from .... (n.d.) Retrieved August 5, 2023, from www.ieee-security.org/TC/SPW2014/papers/5103a236.PDF

[4] Detecting Potential Insider Threat: Analyzing .... (n.d.) Retrieved August 5, 2023, from www.hindawi.com/journals/scn/2018/7243296/

[5] https://ijcsit.com/ijcsit-v14issue2.php

[6] https://aliastechnology.com/