



Artificial Intelligence for the Analysis of the Security Situation in Burkina Faso ★

Abdoul Fataoh Kaboré, Maïmouna Ouattara, Rodrique Kafando, Abdoul Kader Kaboré, Tegawendé F. Bissyandé

► To cite this version:

Abdoul Fataoh Kaboré, Maïmouna Ouattara, Rodrique Kafando, Abdoul Kader Kaboré, Tegawendé F. Bissyandé. Artificial Intelligence for the Analysis of the Security Situation in Burkina Faso ★. 2024. hal-04425416

HAL Id: hal-04425416

<https://hal.science/hal-04425416>

Preprint submitted on 30 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Artificial Intelligence for the Analysis of the Security Situation in Burkina Faso^{*}

Abdoul Fataoh Kaboré¹, Maïmouna Ouattara², Rodrique Kafando³, Aminata Sabané⁴, Abdoul Kader Kaboré⁴, and Tegawendé F. Bissyandé⁵

Centre d'Excellence Interdisciplinaire en Intelligence Artificielle pour le
Développement (CITADEL)
<https://citadel.bf>

Abstract. In the face of the insecurity caused by terrorism that Burkina Faso has been experiencing since 2015, the population doesn't hesitate to express their feelings. The various reactions of the population are expressed through comments on different social platforms, thereby creating a significant amount of data. Analyzing these opinions can provide assistance in decision-making related to security. This analysis can be accomplished through techniques and methods offered by artificial intelligence (AI). In this article, we introduce a web scraping tool to gather data for our research. Subsequently, we employ sentiment analysis and hate speech detection models based on transformers [1]. Through this research, our contributions are as follows: establishing a textual database or corpus related to the security situation, proposing a methodological approach based on AI for analyzing this data, and suggesting a dashboard for visualizing the analysis results.

Keywords: security situation · web scraping · artificial intelligence · transformers.

1 Introduction

Since August 2015, Burkina Faso has been the target of terrorist attacks, affecting both the Defense and Security Forces (FDS) and civilian populations. This situation of insecurity has widespread repercussions across the country. Socially, the situation report from OCHA [2] counts, as of April 30, 2022, 1,520,012 internally displaced persons, of which 59.13% are children, and 4,258 closed schools. The population expresses various opinions and sentiments about the country's security situation. Nowadays, with the proliferation of Information and Communication Technologies (ICT), a significant portion of these reactions is channeled through social media. Analyzing these opinions could aid decision-making related to security.

Artificial intelligence offers numerous methods and techniques for data collection and analysis. It enables sentiment analysis, hate speech detection, named

^{*} CITADEL-UVBF

entity extraction, and more. Hence, artificial intelligence could contribute to addressing the security challenge in Burkina Faso. However, applying AI methods requires a substantial amount of data. To address this, we focus our research questions on the following points: how to acquire a sufficient quantity of data for AI methods application? What relevant analyses can be conducted based on the data to enhance decision-making?

The objective of our research is to gather data on Burkina Faso’s security situation and analyze it using artificial intelligence’s methods and techniques. In this work, we make the following contributions: establishing a textual database or corpus on the security situation, proposing an AI-based methodological approach for data analysis, and suggesting a dashboard for visualizing analysis results.

This article comprises five sections, with this introduction being the first. The second section presents our methodological approach. Section 3 demonstrates the application of our approach. In section 4, we present the various obtained results and their interpretation. Section 5 encompasses a conclusion and the prospects of our work.

2 Methodological Approach

In this work, we propose an approach divided into two main parts: the construction of a dataset and the analysis of this data using artificial intelligence models. The first part involves collecting, preprocessing, and storing textual data related to Burkina Faso’s security situation. As for the second part, it entails applying AI models to this dataset to conduct various analyses. These two parts are interconnected to ensure a reliable analysis based on up-to-date information. Figure 1 provides an overview of the architecture of our approach. Our work is available on GitHub¹.

3 Methodological Application

3.1 Construction of the Dataset

For our initial work, we are using the website *lefaso.net*² as our data source. On this platform, we collect publications (or articles) related to the current security situation in Burkina Faso. These articles are gathered and displayed in a paginated manner under a section titled “*Terrorist Attacks*”. To achieve this, we have implemented a web scraping module to retrieve these articles as well as the comments made by users on these articles. Consequently, we have acquired textual data that we preprocess to retain important information and the appropriate format.

Data preprocessing involves converting all text to lowercase, removing non-alphanumeric characters, and eliminating accents from accented letters. The preprocessed data is stored in *json* format within a list. The essential details are listed in Table 1.

¹ <https://github.com/abdoufataoh/security-situation-analysis>

² www.lefaso.net

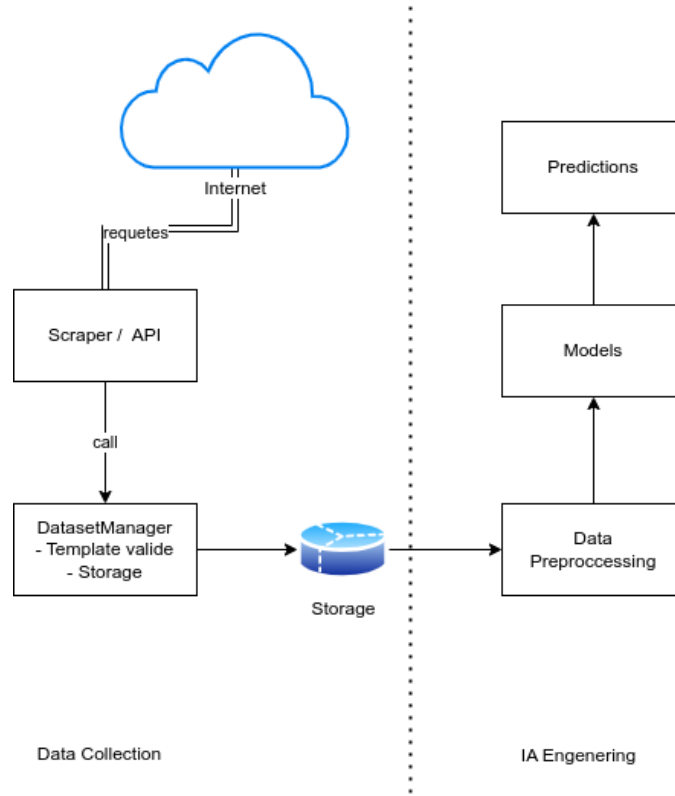


Fig. 1. Architecture

We have collected 2120 articles with a total of 29560 comments. The distribution of the number of comments per article is illustrated in Figure 2.

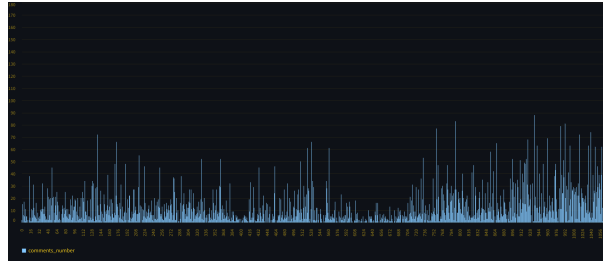
3.2 Sentiment Analysis

We utilize *4-sentiments-with-flaubert* [3], a pre-trained model on French language data, designed for sentiment analysis. Given input text, it classifies the text into four categories:

- **POSITIVE:** A positive sentiment is a favorable or pleasant feeling. For instance, an internet user believing that the Burkinabe authorities are on the right path to defeat terrorism.
- **NEGATIVE:** A negative sentiment is an unfavorable or unpleasant feeling. For example, fear following a terrorist attack; a citizen might experience negative sentiments due to fear for their own safety or that of their loved ones.

Table 1. List of Article Keys

Key	Description
<i>article_type</i>	Article type (report, article, etc.)
<i>article_title</i>	Title of the article
<i>published_date</i>	Publication date of the article
<i>origin</i>	Source of the article
<i>url</i>	Article access URL
<i>content</i>	Article content
<i>comments</i>	List of comments on the article
<i>comments_number</i>	Number of comments

**Fig. 2.** Distribution of the Number of Comments per Article

- MIXED: A mixed sentiment combines both positive and negative feelings. A citizen might feel both sadness for the attack victims and anger towards the perpetrators, or both fear for personal safety and hope for the country’s recovery.
- OBJECTIVE: A sentiment can be considered objective if it’s based on facts and evidence rather than personal opinions or beliefs. Objective sentiments might stem from understanding the reasons behind an attack, relying on facts like conflict history, attackers’ motivations, etc., rather than personal preferences or beliefs. However, it’s essential to note that even in this case, individuals might still hold subjective feelings about the situation.

3.3 Hate Speech Detection

In this analysis, our goal is to detect hate speech in comments. For this purpose, we employ *hubert-mono-french*, a model specialized in this task. It classifies text as either hateful or not hateful, using the labels:

- Hate
- No_Hate

4 Results

4.1 Sentiment Analysis

By using the *4-sentiments-with-flaubert* [3] model, we have successfully determined the sentiments of internet users through their comments on the security situation in Burkina Faso. Figure 3 depicts the evolution of different sentiments on a monthly basis. The curve labeled *negative_sentiment* (in red) generally stands above the other curves. Notably, this curve exhibits significant spikes indicating a very high negative sentiment, which usually arises after significant attacks: the first in December 2016 with the attack on Nassoumbou, in November 2019 when the country suffered high human losses due to the attack on Semafo de Boungou, in June 2019 with the attack on the village of Solhan [5], and in August 2022 with the attack on Nohao near the city of Bittou. We also observe a reduction in negative comments starting from September 2022.

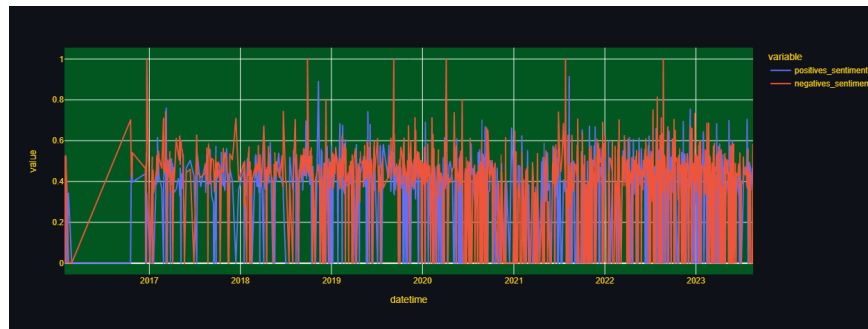


Fig. 3. Sentiment Evolution Curve Based on Monthly Granularity

4.2 Detection of Hate Speech

Analyzing comments using the *hubert-mono-french* [4] model enabled us to create Figure 4. This figure presents the evolution of the number of comments containing hate speech on a monthly basis. It is evident that hate speech is present in user comments, but overall, non-hateful messages dominate.

5 Conclusion

The objective of this study was to gather and analyze data on the security situation in Burkina Faso. We collected data using web scraping techniques, which were subsequently employed to train transformer models for sentiment analysis and hate speech detection. The results from our visualizations reveal an increase

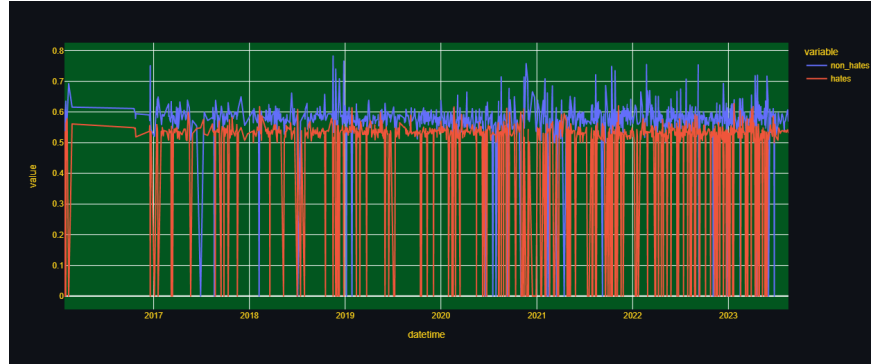


Fig. 4. Evolution of Hate Speech Comments Based on Monthly Granularity

in negative sentiments and hate speech during periods marked by terrorist attacks. These efforts lay the groundwork for a system that could contribute to decision-making regarding Burkina Faso's security situation.

Future directions for our work include collecting data from heterogeneous sources such as other social media platforms. Additionally, we plan to identify spatial named entities within the collected data to visualize the relationship between these named entities and the various analyses derived from the data we've gathered.

References

1. LIN, Tianyang, WANG, Yuxin, LIU, Xiangyang, et al. A survey of transformers. arXiv preprint arXiv:2106.04554, 2021.
2. <https://reports.unocha.org/fr/country/burkina-faso>
3. J  r  my Demange, Four sentiments with FlauBERT, (2021), Hugging Face repository, <https://huggingface.co/DemangeJeremy/4-sentiments-with-flaubert>
4. Aluru, Sai Saket and Mathew, Binny and Saha, Punyajoy and Mukherjee, Animesh Deep Learning Models for Multilingual Hate Speech Detection, 2020
5. https://fr.wikipedia.org/wiki/Massacre_de_Solhan