WORKING PAPER SERIES

234, 2025

FROM SCREENS TO GRIDS: A WEB SCRAPPING ANALYSIS OF ELECTRICITY ACCESS IN SUB-SAHARAN AFRICA

Pius Gamette
Department of Applied Economics, School of Economics,
University of Cape Coast, Ghana
E-mails: pius.gamette@ucc.edu.gh / gamettepius@gmail.com

Simplice A. Asongu
The Institute of Convergence Science, Korea University,
Seoul, South Korea
E-mails: asongusimplice@yahoo.com / asongus@afridev.org

Abstract

In view of the lagging position of sub-Saharan Africa (SSA) compared to other regional blocs in accessing electricity coupled with the wide disparity in rural-urban access to electricity, this study adopts the scrapping method to efficiently gather and analyse social discourse in real time. From the YouTube platform, this study examined 887 comments from January 2014 to July 2024. The result revealed that "ANC," "Africa," "south," "power," and "will" were the most frequently discussed topics in the discourse of electricity access. Also, more positive-tone comments were recorded than negative-tone remarks within the period. In terms of subjective polarity analysis, comments were positively inclined. The top ten topics that emerged were "ANC", "Energy", "People", "Country", "Government", "Electricity", "Eskom", "Africa", "Power" and "South". The heatmap showed that there were a few topics with high correlation values. For cluster analysis, "South", "Country", "Eskom", "Power" and "Energy" had positive sentiments and contents. On the contrary, "Government", "Africa" and "People" were topics of negative sentiments and contents. Only "ANC" connected neutral sentiments and content. "ANC", "Africa" and "country" had an 81.77% degree of centrality in the public digital discourse on access to electricity in SSA countries. This study has important ramifications for SSA, an energy-poor region towards achieving Sustainable Development Goal (SDG) 7.

Keywords: sentiment analysis, electricity, discourse, Africa, web scraping
JEL Classification: C8, L94, Q4

1. Introduction

Electricity access is essential for social services (including health and education) and economic development in both developed and emerging economies (Asongu & Le Roux, 2024; Banerjee, Mishra & Maruta, 2021). Emphasis is placed on this by the Sustainable Development Goals (SDGs), which aim to guarantee that by 2030, all people have access to clean, affordable and dependable energy. African homes, businesses, industries and institutions should all have access to modern, reliable, efficient and reasonably priced energy by 2063, according to a proposal made by the African Union Commission (Munezero Buzingo, 2021). These objectives all seem to suggest that a nation's ability to acquire electricity is essential to its ability to prosper economically. Albeit, sub-Saharan Africa (SSA) is home to more than two-thirds of the world's population without access to electricity (Byaro & Mmbaga, 2022). Estimates from the International Energy Agency show that from 2019 to 2021, there were 4% more individuals in SSA without access to modern electricity. Furthermore, according to Valickova and Elms (2021), 1.1 billion individuals who lived in rural areas did not have access to electricity in the region. As per the provided data, the electrification rate in SSA stands at 45%, while it approaches 94% in Asian nations (Byaro & Mmbaga, 2022). One way that SSA countries can switch to renewable energies is through the use of electricity. Making the switch to renewable energy looks more than just like a choice; it is a sign of hope that energy production and environmental preservation will co-exist peacefully in the future and promise an endless supply of energy (Ovchinnikova et al., 2021; Razi & Dincer, 2022).

Social media platforms have grown significantly and quickly as a result of the growing importance of renewable energy in international talks. This has had a significant impact on our modern social structure (Saud et al., 2020). Since these platforms were initially imagined as digital spaces for forming personal connections and engaging in casual interactions, they have undergone a substantial metamorphosis. From their early beginnings as digital meet-up locations, they have expanded dramatically to become persuasive engines of conversation, information conversation, and base-level mobilisation (Hargreaves & Middlemiss, 2020). Social networking architecture facilitates the ad hoc emergence of diverse groups, from hyperlocal forums debating solar panel fittings in the neighbourhood to vast international mobilisations uniting around the more important goal of mitigating climate variability via sustainable energy activities (Boulianne, 2020). The digital insurgency has lowered information access blockades and greatly increased the voices of people who might have been marginalised in influential debates in pre-digital times. Disenfranchised groups in the past, campaigners without major platforms and regular people who are passionate about sustainability now have the tools to transmit, inform and motivate on a global level. Furthermore, the interactive and real-time features of these platforms

contribute to their unmatched effectiveness. They have the power to organise large crowds, mould popular perceptions and occasionally have a direct or indirect impact on legislative choices (Boulianne, 2020). The shared voice emanating from these digital echo perspectives often influences the strategies and stances of legislators, corporations and decision-makers, who are becoming more and more attuned to the pulse of online emotion (Gilardi et al., 2022).

The diverse communities in SSA have become increasingly vocal on social media platforms regarding their want to transit and access electricity. The increasing popularity of conversations surrounding the attainment of a carbon-neutral SSA region and the promotion of environmental protection has made social media sites like Facebook, Instagram and X as indispensable tools for sentiment analysis and information sharing. They work as catalysts, dispelling misconceptions, educating the public about the many advantages of sustainable energy practices and cultivating an informed support system (Qazi et al., 2019). The resources and techniques we have at our disposal develop even additional important as we strive to understand the dominant values and general consciousness of the digital dispensation. This is particularly crucial given that digital content is available from a wide range of sources. The core of this research is text analysis and sentiment analysis. These two interconnected computer techniques offer an advanced means of comprehending and deciphering the intricate web of digital communication (Nandwani & Verma, 2021). A bright light in this area of study is sentiment analysis, which makes use of sophisticated linguistic algorithms and machine learning models. This method, at its core, explores the nuanced nuances of emotion that are present in written content. Sentiment analysis goes beyond superficial reading to uncover the nuanced emotional undertones, which may signal happiness, ambivalence, discontent, or open hostility. Essentially, it aims to chart the emotional terrain of a conversation, identifying not only the explicit topics but also the implicit emotions that give such messages context and personality (Wankhade, Rao & Kulkarni, 2022). Sentiment analysis is supported by the similarly potent field of text analysis. Text analysis employs a distinct methodology, even though the former is delicate to the emotional incidences of the content. It meticulously sifts through the massive volumes of data, looking for patterns, recurring themes, and semantic relationships. This method builds a thorough understanding of the subject or problem at hand by dissecting narratives, connections, and contextual cues (Macanovic, 2022).

Sentiment analysis can be applied to assess the level of access to electricity given its regional relevance. There are a lot of different opinions, statements, conversations and anecdotes about getting power online. These come from a variety of sources, including both seasoned industry experts and everyday consumers. This large corpus of material may yield invaluable insights into the various points of view that influence the conversation by utilising text analysis and

emotion analysis (Jain et al., 2021). This study provides a thorough understanding of the attitudes, concerns, hopes and doubts held by a diverse digital populace, which improves understanding of accessing electricity in an energy-poor region like SSA in the age of digitalisation. According to Isoaho, Gritsenko and Mäkelä (2021), combining reliable sentiment and text analysis tools with the vibrant and ever-expanding online discourse about renewable energy can yield insightful results. Given the urgency and the significant potential of renewable energy solutions in SSA countries, including access to electricity, the digital age can promote connectedness, spread knowledge and enable preventative actions. Social media platforms have evolved greatly from being merely online forums to powerful discussion spaces with the power to influence policy and set global trends. Digital discourse and electricity accessibility together offer a specific perspective and a well-thought-out implementation strategy. This extensive study uses sentiment analysis to investigate the complex nature of internet conversations via YouTube. Compared to other social platforms, YouTube provides a combination of video, audio and text data (video descriptions, titles, comments, transcripts, and subtitles) (Senevirathne, Kumara & Banujan, 2024). Furthermore, YouTube is one of the largest repositories of user-generated content, with billions of videos and a vast amount of user interaction through likes, dislikes, and comments providing a basis for a more robust sentiment analysis (Cooper, 2023). Its goal is to address the ambitions, fears, and goals of the people in SSA countries who patronise social media platforms while providing a complex and multifaceted understanding of accessing and using electricity in SSA countries.

This study offers three distinct contributions. (i) By using sophisticated natural language processing techniques and web scraping to examine public opinion on electricity accessibility in SSA, this study presents a fresh methodological approach. Comparing this methodological breakthrough to standard survey methods, it provides a more dynamic and comprehensive picture of public discourse by enabling the collection and analysis of large-scale, real-time data from a variety of internet sources. (ii) By studying electricity availability in SSA, the study adds to our understanding of its socio-political elements. The research reveals how public perceptions and narratives regarding electricity accessibility are constructed, highlighting the socio-economic, political, and regional aspects that influence these beliefs. It does this by analysing attitudes expressed in online debates. (iii) By fusing socioeconomic and demographic data with sentiment analysis findings, this study will offer insightful policy information. Patterns and correlations that might guide focused interventions and policy choices will be made visible through this integration. Through the analysis of public opinion in various sub-regions and demographic groupings, this research provides evidence-based suggestions to policymakers to help them better handle challenges related to energy availability.

Apart from the first section, the study is structured as follows: Section 2 provides pertinent studies on social media analytics and renewable energy and public opinion. Section 3 outlines the methodology, followed by sections 4 and 5 with the experimental data and discussions of the findings and conclusions and policy implications, respectively.

2. Review of Literature

2.1 Social Media Analytics

The process of converting social media data into useful knowledge is known as social media analytics (SMA). Research on text mining in the renewable energy field has exploded in the last few years. Direct experimental procedures and those that make use of indirect datasets are the two main analytical approaches. The former includes modeling trials, such as Butkowski et al. (2020), and on-site questionnaires, such as Liu et al. (2021). On the other hand, the latter involves gathering data from particular forums for public discussion. Leading text analysis platforms are Facebook (Nuortimo, Karvonen & Härkönen, 2020), Instagram (Nuortimo, Karvonen & Härkönen, 2020; Vespa et al., 2022), TikTok (Söğüt & Öngel, 2022), X (previously Twitter) (Corbett & Savarimuthu, 2022; Kim et al., 2022) and news media sources (Dehler-Holland, Okoh & Keles, 2022). Examining public opinion and the most talked-about subjects in text analysis are two of the most popular goals (Dehler-Holland, Okoh & Keles, 2022). The other goal is to determine how new technology is seen to be legitimate. As per Zafar et al., (2021), public sentiment refers to a concept, idea, or opinion that stems from a mindset or an emotion regarding a particular circumstance. Text-mining analytics is used to analyse data that has been produced and text connected to the issues of renewable energy. To identify important themes discussed concerning wind power technology, for instance, topic modeling has been utilised (Dehler-Holland, Okoh & Keles, 2022).

A corpus of texts derived from official papers, internet forums and media sources is used for this. An application that is linked to a social media application programming interface (API) was used to extract data. In addition, a list of keywords associated with the environment is used to filter the chosen remarks and categorise the subjects (Corbett & Savarimuthu, 2022). When web crawling, a few primary keywords are usually employed as a baseline for a particular context (Jeong, Ha & Lee, 2021). To find the most common words, frequency analysis is another tool that may be used in addition to topic modeling. This method differs from survey-based analysis in that it uses a respondent-provided questionnaire with pre-specified items. According to Zafar et al., (2021), these components clarify the relationship between independent and dependent variables in a population. The adoption of renewable energy technology will be modeled using multinomial logistic regression with outcome variables for accepting, not accepting and doubtful. In social media analytics, topics and possibilities, significant conversations and challenges mentioned in user-generated material are identified using topic modeling and sentiment analysis (Chen & Xie, 2020). In order to perform sentiment analysis, the natural language toolbox needs to be modified

for various research topics. Software data, like that from M-adaptive media monitoring, can be utilised by sentiment analysis algorithms (Nuortimo, Karvonen & Härkönen, 2020).

2.2 Sentiment Analysis and Energy Transition: Empirical Perspective

Research showed that sentiment analysis could be done with "R" and that it may be used in a variety of industries, such as energy and banking (Klinkhammer, 2022). Sentiment analysis has found great riches in the enormous reaches of social media, particularly in industries such as energy and finance. The prospect of sentiment analysis in forecasting market movements was highlighted by a study that examined the usability of several NLP models to analyse social media sentiments with a focus on the cryptocurrency space. The study found correlations between sentiment metrics and changes in the price of Bitcoin. Another fascinating work investigated how to use transformer-based models that have already been trained to curate social media corpora to remove irrelevant tweets. The study's remarkable F1 scores point to the technique's potential for use as a social media dataset pre-processing phase (Arnold et al., 2023). To gauge public opinion of electric vehicles, sentiment analysis applications have also been used in the field of electric vehicles (Şenyapar, 2023). These systems use both token-wise and document-wise sentiment analysis.

According to recent scholarly research, public opinion toward renewable energy supplies is becoming more and more important. The Northeastern part of the United States has more positive ideas about solar energy than the Southern region, according to a study that looked at over 71 thousand messages throughout January and July 2020. ideas on solar energy varied significantly between the states. Targets for Renewable Portfolio Standards (R.P.S.) and net metering legislation, as well as market conditions pertaining to renewable energy, were found to be correlated with public opinion by Kim et al. (2020). Another study that looked at the emotions of people in the UK and India regarding tweets about power found that people were against growing electricity prices even with government subsidy programs (Kaur & Edalati, 2022). A sentiment classification model based on machine learning was used to analyse public attitude towards wind energy in Norway, and the results showed an upsurge in anti-wind tweets throughout the 2018–19 season (Vågerö & Zeyringer, 2023). This research underscores the significance of opinion and text analysis in assessing public opinion and developing policies and strategies related to renewable energy.

3. Methodology

3.1 Site Selection for Countries

The SSA region is the focal area for this sentiment analysis as it continues to trail in access to electricity compared to other regions across the globe hence is an area of concern for energy access discourse (Blimpo, Postepska & Xu, 2020; Zogo et al., 2024). With the onset of COVID-19 pandemic, it was estimated that 580 million people worldwide were without access to electricity in 2019—of which 759 million were from SSA—will have even more of a problem in 2030 (Valickova & Elms, 2021). According to comparable statistics, the situation will deteriorate as 75% of people in 2025 in comparison to 48% and 67% in 2010 and 2017, respectively (de Abajo Llamero 2023). Although 46% of people in SSA have access to electricity, up from 33% in 2010, some households are close to the national grid but are not connected due to high cost, lack of information, unreliable grid and administrative barriers in using electricity (Sievert & Steinbuks 2020). Several SSA countries have access to YouTube for video sharing and content consumption. However, the availability of YouTube servers specifically depends on the region's internet infrastructure, government regulations and regional partnerships with global companies like Google (which owns YouTube) (Stocker et al., 2017). Notable SSA countries with robust internet infrastructure that ensure seamless streaming of YouTube include South Africa, Kenya, Ghana, Uganda, Ethiopia, Senegal, Cote D'Ivoire, Tanzania, Zambia, Botswana, Namibia, Malawi, Togo, Mozambique, Rwanda and Burkina Faso. While YouTube itself doesn't have country-specific servers in every nation, the platform's content is typically accessible through the global network of Google data centers, which are located in various regions around the world, including SSA (Noonan & Piatt, 2015). Specifically, Google has invested in infrastructure to improve internet speed and accessibility across the continent which indirectly affects YouTube's performance (Phokeer, 2021). YouTube attracts a significant and diverse user base across various demographics, particularly urban, tech-savvy and younger populations. This therefore makes it a valuable proxy for understanding the sentiments of digitally active individuals who often drive online discourse and influence public opinion.

3.2 Data collection

Using social networking statistics from YouTube, we gathered posts and comments from YouTube users from January 1, 2014 to July 31, 2024, based on accessibility of electricity in the SSA. Access to the YouTube platform was facilitated through the global network of Google data centers which are located in various SSA countries with support from the various governments in the installation of internet infrastructure. Data was collected over 10 years and 7 months period to cater for the

changes in comments expressed from SSA countries as these countries have initiated varied electrification programmes, integrated technology in their energy provision and are vulnerable to the influence of climate change.

3.3 Data preprocessing

To increase the accuracy of the data as well as capture data over the study, we generated Application Programming Interfaces (APIs) from Google Colab and further used to get access to public comments on electricity access in SSA from YouTube. This was carried out before starting the sentiment analysis (Palomino & Aider, 2022; Dang, Moreno-García & De la Prieta, 2020). This required a number of actions including data cleaning to capture geolocation filters, normalization, addressing outliers, tokenization and removal of stop words.

3.4 Handling noisy data text cleaning

We start our preparation with a thorough text cleaning procedure. This entails eliminating comments outside of SSA countries, non-alphanumeric letters, HTML elements, and URLs— common sources of noise in text data. To maintain consistency, we also transform all text to lowercase and standardise the case.

3.5 Handling noisy data normalization

We use text normalisation techniques beyond simple cleaning, such as eliminating stop-words (words that are frequently used but unlikely to have any significance, like "and," "the," etc.) and using lemmatisation to condense words to their dictionary or base form. This stage aids in narrowing the analysis's emphasis to significant words.

3.6 Addressing outliers' statistical analysis for anomaly detection

To find word frequency and emotion score outliers, we perform a statistical study. Words that are flagged for examination and, if determined to be outliers, eliminated, are those that occur very frequently yet do not add to sentiment or thematic relevance.

3.7 Addressing outliers' sentiment score distribution analysis

In order to locate and evaluate possible outliers, we examine the sentiment score distribution throughout the sample. We ascertained whether articles with extreme emotion ratings

that exhibit a large deviation from the mean are the consequence of irregularities in the data collection or preprocessing, they undergo a manual examination

3.8 Tokenization

The text was divided up into discrete words, or "tokens." We can examine the frequency and distribution of words in the text using this common text analysis method.

3.9 Stop word removal

We eliminated terms like "the," "is," and "and" that are frequently used but have no real significance. In text analysis, these terms are frequently omitted in order to concentrate on the words that provide more information.

3.10 Lemmatization

Words were stripped down to their most basic or root form (e.g., "accessing" to "access"). This lessens the complexity of the analysis by combining many usages of the same word. We preprocessed and analysed comments about the availability and usage of electricity in SSA countries. We chose rows in the dataset where the term "electricity" appears in the "comments" column to narrow the focus to renewable energy-related topics. This enables us to selectively limit the dataset to remarks about electricity.

3.11 Sentiment analysis

We used sentiment analysis to examine people's views on power access by examining the pre-processed postings and comments. To look into sentiment analysis, we employed linguistic investigation and word count to develop a text analysis package for determining the sentiment value of words in different categories (e.g., affective, intellectual, social, and positive/negative assertions (Li, Goh & Jin, 2020). The degree of positive and negative emotions shown for each post or comment was another way we looked at SSA's access to electricity (Blimpo, Postepska & Xu, 2020).

3.12 Topic modelling

We applied Latent Dirichlet Allocation (LDA), a topic modeling technique, to find recurring themes or subjects in news stories. LDA is a probabilistic model that assumes each document

consists of a mixture of a predefined number of themes, and that each word in the paper can be related to one of the subjects (Kim et al.,2021). Topic modeling is a statistical modeling technique that is used to determine the abstract "topics" that occur in a collection of documents (Dang, Moreno-García & De la Prieta, 2020). One method that is frequently employed in topic modeling is Latent Dirichlet Allocation. (LDA). Every document is a mixture of a predetermined number of subjects, and each word in the written content can be connected to one of those subjects, according to the evolutionary probabilistic model called LDA. The definition of a subject in LDA is a distribution over a specified vocabulary. The word "electricity" may contain phrases like "electric," "hydro," and "electrical," each of which has a distinct likelihood.

3.13 Natural language processing (NLP) techniques: suitability

NLP approaches are especially well-suited for our study because of their capacity to handle and examine substantial amounts of unstructured text data from social media platforms. NLP's robust and adaptable analytical capabilities are required due to the various and dynamic nature of language used on social media sites to highlight problems obtaining electricity. Through the application of these methods, we are able to methodically classify opinions shared on social media, revealing public attitudes and stories on the availability of power in SSA nations.

3.14 Latent Dirichlet Allocation (LDA): suitability

Our work is especially well-suited to the LDA model for a number of reasons. In order to get insight into the thematic landscape of the debate surrounding access to electricity, it first enables us to identify hidden themes within our large collection of social media comments on getting electricity. Second, LDA's versatility and capacity to reveal latent theme structures in vast amounts of data without the need for pre-labeling make it perfect for analysing a wide range of complex and extensive news data. Last but not least, the probabilistic character of LDA provides a sophisticated comprehension of subject composition and prevalence, which is essential for deciphering the intricate interaction of themes in the accessing electricity process. Thus, the LDA model in conjunction with NLP approaches provides a strong framework for topic and sentiment analysis in huge textual datasets (Palomino & Aider, 2022). These approaches are theoretically sound and useful in the real world for breaking down the complex ways in which language is employed on social media to communicate ideas and feelings in order to obtain power. With the use of this strategy, we can accomplish our research goal of learning how the general public in SSA nations perceives access to electricity, which will provide significant knowledge to the area.

3.15 Correlation and Clustering Analysis

In order to investigate the connections between the remarks made on SSA's access to power, we performed a correlation study. Moreover, a clustering analysis was done to put related comments in one group. This gave information about the recurring themes or subjects in the comments as well as how the sentiment scores relate to them. We have clustered and conducted a correlation study based on the sentiment value of social media comments. Cosine similarity, which calculates the cosine of the angle between two vectors, was used to do the correlation study. This represents the degree of similarity between the documents, regardless of size. K-means clustering was employed for the clustering; it is a kind of unsupervised machine-learning approach for classifying objects or clusters (Li, Goh & Jin, 2020). Based on their TF-IDF vectors, which indicate the relative relevance of various remarks on social media sites, the social media comments in this instance were clustered.

3.16 Software and tools

Python was used for all aspects of data collection, preprocessing, analysis, and visualisation. NLTK and pandas were used for text preprocessing, VaderSentiment was used for sentiment analysis, sklearn was used for LDA and correlation analysis, and matplotlib and seaborn were used for data visualisation.

3.17 Primary algorithms and models

We utilise the VADER (Valence Aware Dictionary and Sentiment Reasoner) model for sentiment analysis, which is well-known for its precision in interpreting social media texts and complex language. It is generally known that VADER is successful in a number of domains (Marrapu, Senn & Prybutok, 2024). VADER: A Parsimonious Rule-based Model for Social Media Text Sentiment Analysis Eighth International Weblog and Social Media Conference. Based on (Negara & Triadi, 2021), the Latent Dirichlet Allocation (LDA) model is used for topic modeling. Our strategy for identifying dormant subjects related to obtaining power on social media platforms is based on this basic study.

4. Results and Discussion

The results section presents the findings of the sentiment analysis and the polarity of the sentiments expressed. We would also discuss the various topics modeled from the sentiments, correlation and clustering analysis results, which provide insights into the relationships between social media comments and the common themes or topics among them. Lastly, the section highlights the degree of centrality of the comments on access to electricity in SSA countries.

4.1 Word Frequency

Word clouds for every topic were also created using a module. For every topic, this will present a word cloud with the frequency of each word in the topic indicated by its size (Figure 1).

Figure 1: key words linked to access to electricity in SSA countries

As shown in Table 1, the visual representations highlight the most frequent and significant themes associated with access to electricity in SSA countries. The major highlighted words are "Anc", "Africa", "Power", "South" and "Will". These word clouds show that there is an active and multifaceted discussion around electricity access in SSA countries. The word "Anc" is bold in the word could due to the heightened discourse on how the current government in South Africa is tackling electricity challenges in South Africa. Moreover, "Africa" is highly mentioned in the discourse on accessing electricity due to the sub-region's contribution to augmenting Africa's

14

access to electricity over the years. The highlighted words reflect the sub-region's public's perspective on the urgency of these issues and the collective desire to enhance access to electricity and drive change through dialog, advocacy and action.

4.2 Sentiment Analysis

Every YouTube remark about the accessibility of electricity in Sub-Saharan Africa (SSA) was given a sentiment score by us using a sentiment analysis model. Higher sentiment scores indicate more positive emotion while more negative sentiment is indicated by lower sentiment scores. The sentiment score is a numerical value that captures the overall sentiment of the text. Our sentiment analysis tool of choice was the VADER (Valence Aware Dictionary and Sentiment Reasoner) model, a rule-based lexicon designed to be sensitive to the sentiments expressed on social media (Hutto et al., 2014).
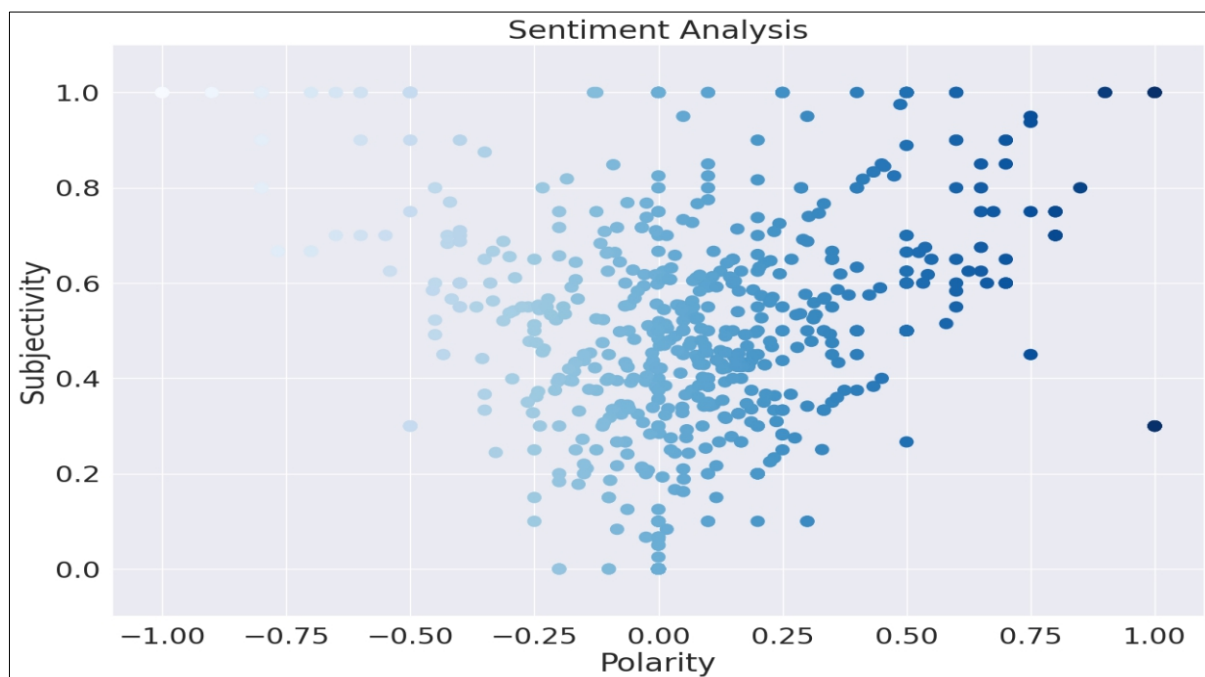

Figure 2: sentiment analysis of comments from YouTube on access to electricity in SSA countries

The sentiment analysis shows that many of the comments expressed on the YouTube platform indicate more positive sentiments. From the scatter diagram indicated in Figure 2, many of the scatter plots are dense and concentrated in the positive quadrant of Figure 2.

4.3 Polarity Sentiment Analysis

To ascertain the sentiment conveyed in a text, polarity sentiment analysis is a method used in natural language processing (NLP) (Dang, Moreno-García & De la Prieta, 2020). This sentiment can be categorized into different levels of polarity, usually ranging from positive, neutral, to negative.
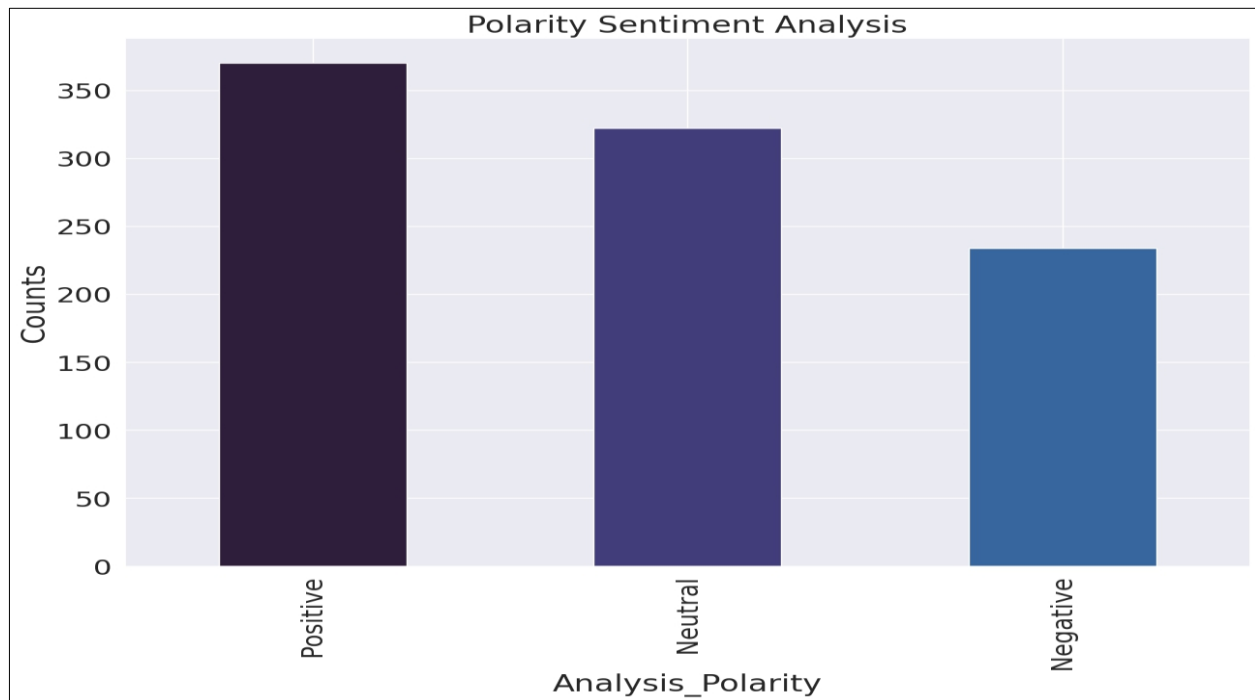

Figure 3: polarity of sentiments analysis from YouTube comments on access to electricity in SSA countries

From the 883 comments retrieved from the YouTube platform, Figure 3 indicated that the sentiments were polarized towards positive sentiments followed by neutral comments and lastly negative comments. The implication is that based on the videos retrieved from the YouTube platform, the public expressed more positive sentiments compared to negative sentiments indicating the prospects of the various households in accessing electricity in SSA countries.

4.4 Robustness Check for Sentiment Analysis

For confirmation of whether the comments expressed on YouTube are sentiments or facts, the study conducted a robustness check. As shown in Figure 4, the robustness check confirmed

that the comments retrieved from YouTube were more opinions as compared to facts. Within the context of checking the robustness of the comments expressed, the subjective nature of the comments was mostly not verifiable with few comments being objectives and verifiable through evidence or reliable sources.
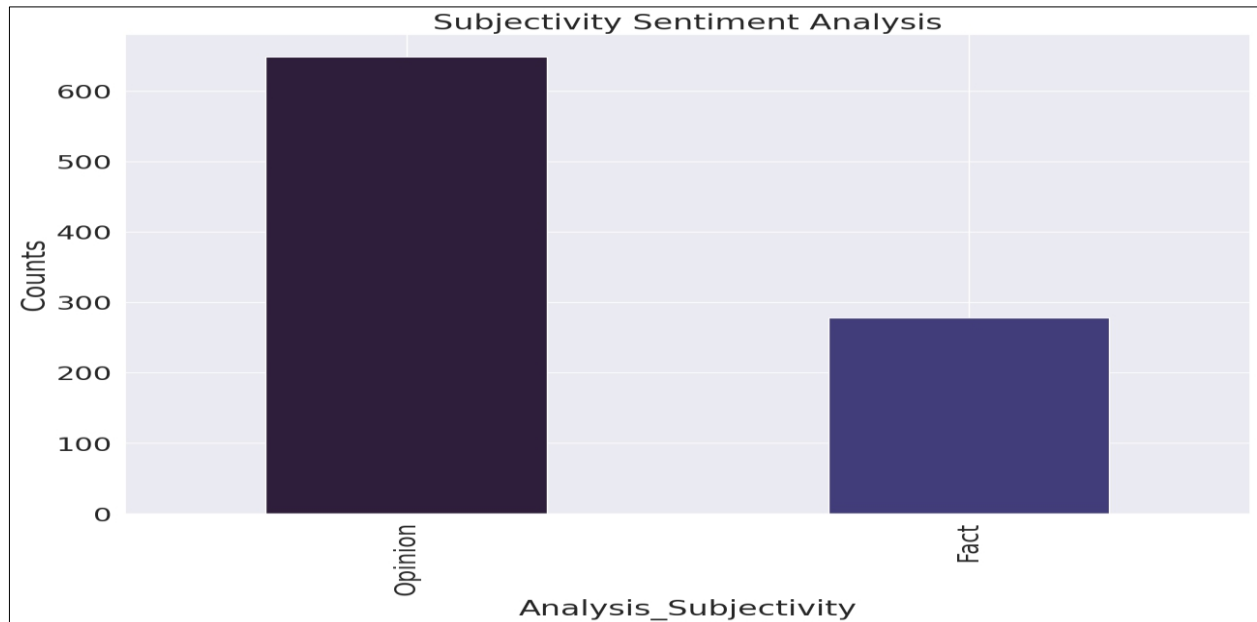

Figure 4: Robustness check of comments from YouTube

4.5 Topic Modelling

Using a topic modelling method like Latent Dirichlet Allocation (LDA), we were able to determine which subjects were most common. LDA is a type of probabilistic model in which, on the assumption that each document consists of a mixture of a certain number of topics, every expression in the document can be associated to one of the subjects (Negara & Triadi, 2021). To find the most popular subjects, we ran Latent Dirichlet Allocation (LDA) on the YouTube comments.
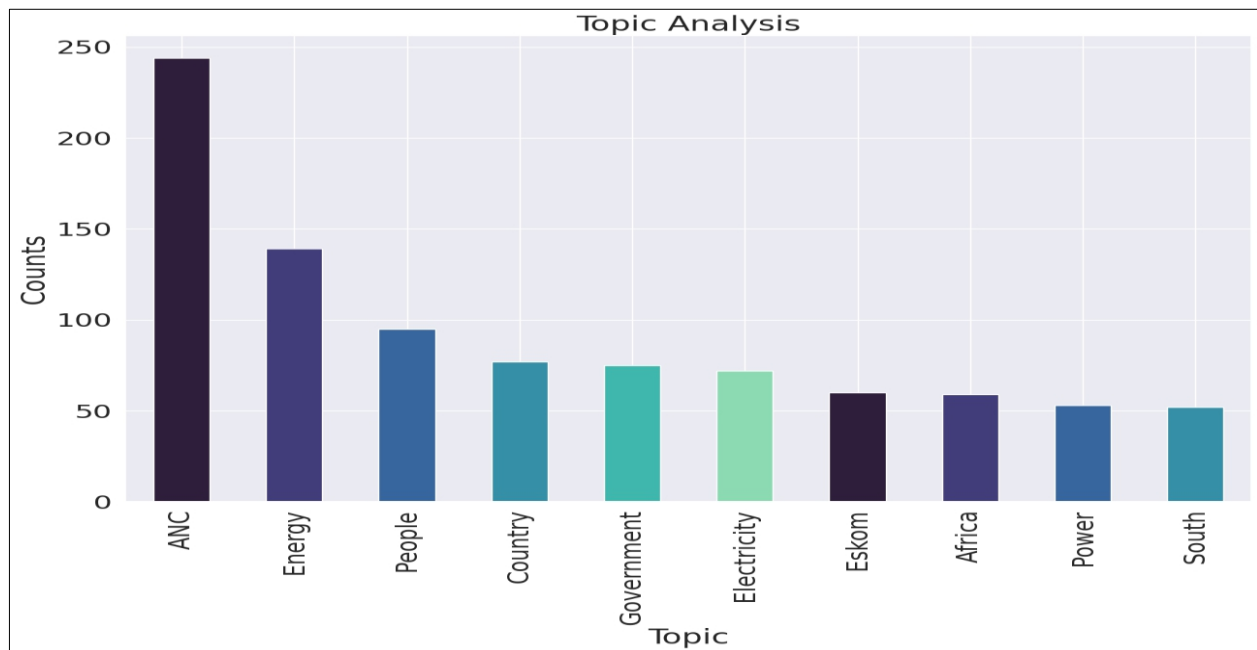
Figure 5: Topic Analysis of Comments from YouTube

From Figure 5, the top ten topics generated from the comments expressed on the YouTube platform in relation to access to electricity were "Anc", "Energy", "People", "Country", "Government", "Electricity", "Eskom", "Africa", "Power" and "South". Intuitively, the digital discourse on electricity access in SSA countries is heavily full of comments on the ANC party in South Africa suggesting that the major political party in South Africa is a dominant topic in the digital discourse of electricity access in SSA countries. Other notable topics include "Energy", "People", "Country", "Government" and "Electricity". On the other hand, the less domineering topic in the digital discourse of access to electricity are "Eskom", a power distribution company in South Africa followed by "Africa", "Power" and "South". The last four topic though are frequent in the digital discourse of electricity access in SSA countries, they are not compared to "Anc", "Energy", "People", "Country", "Government" and "Electricity".

4.6 Topic Polarity Distribution

Within each of the top ten topics, each topic is made up of three polars (positive, neutral and negative). Figure 6 shows the extent of topic polarity as generated from the comments expressed on the YouTube channel.
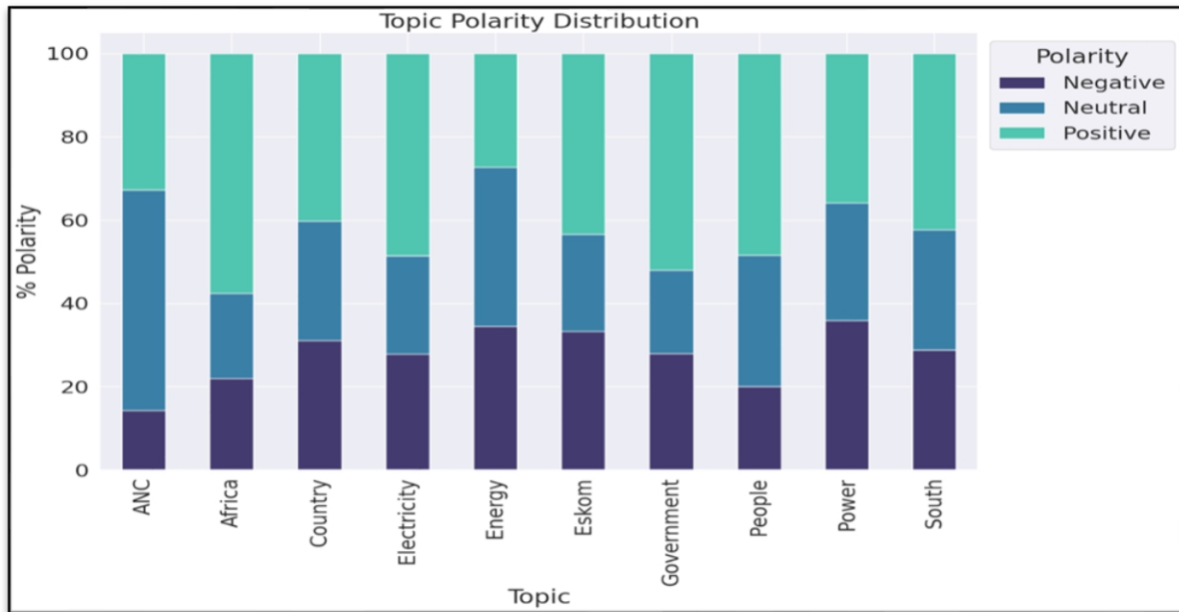
Figure 6: Topic polarity distribution based on the top ten topics

As indicated in Figure 6, neutral sentiments dominated both positive and negative comments when "Anc" topics were raised. On the contrary, positive sentiments dominated when topics concerning "Africa", "Country", "Electricity", "Eskom", "Government", "People", "Power" and "South" were raised. In another context, positive, neutral and negative sentiments were equally shared when the "Anc" topic are raised on YouTube platform.

4.7 Correlation

The correlation matrix displays each pair of top subjects from the YouTube platform along with their cosine similarity. The similarity between two values is high when it is near one and low when it is near zero. Whereas a negative correlation shows that one topic's sentiment scores tend to rise when the other topic's sentiment scores fall, and vice versa, a positive correlation shows that the sentiment scores of the two topics tend to rise or fall together (Nolasco & Oliveira, 2020).

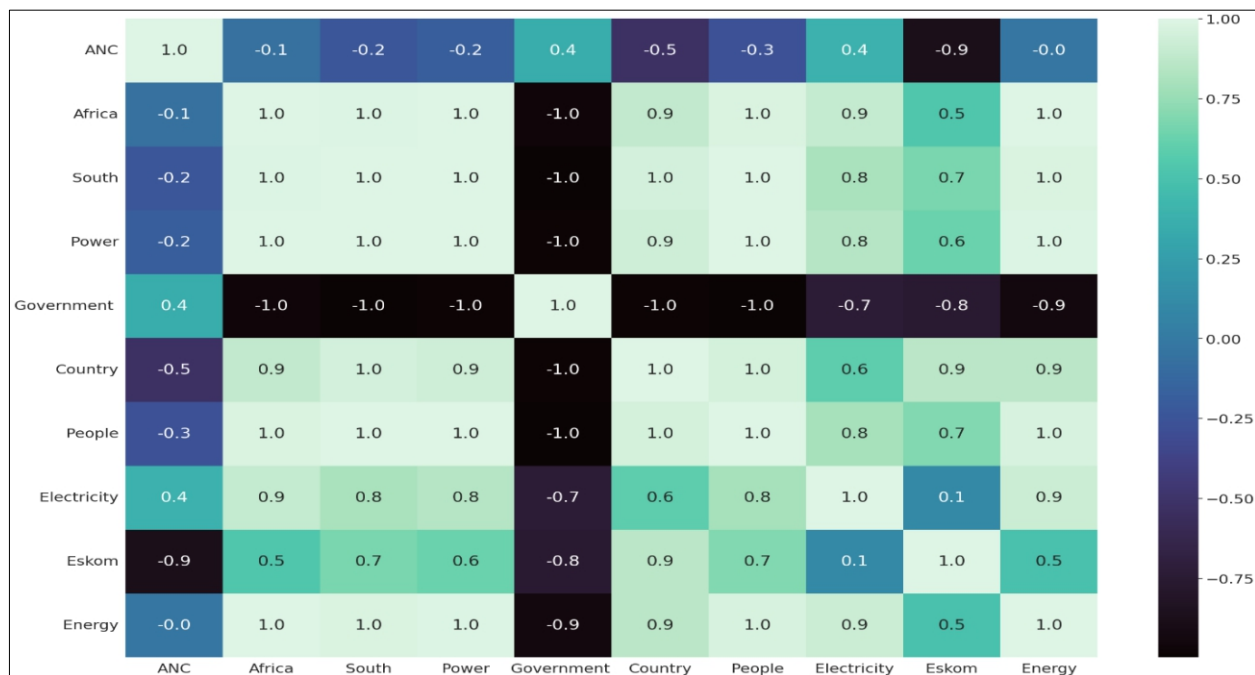|	ANC	Africa	South	Power	Government	Country	People	Electricity	Eskom	Energy
ANC	1.0	-0.1	-0.2	-0.2	0.4	-0.5	-0.3	0.4	-0.9	-0.0
Africa	-0.1	1.0	1.0	1.0	-1.0	0.9	1.0	0.9	0.5	1.0
South	-0.2	1.0	1.0	1.0	-1.0	1.0	1.0	0.8	0.7	1.0
Power	-0.2	1.0	1.0	1.0	-1.0	0.9	1.0	0.8	0.6	1.0
Government	0.4	-1.0	-1.0	-1.0	1.0	-1.0	-1.0	-0.7	-0.8	-0.9
Country	-0.5	0.9	1.0	0.9	-1.0	1.0	1.0	0.6	0.9	0.9
People	-0.3	1.0	1.0	1.0	-1.0	1.0	1.0	0.8	0.7	1.0
Electricity	0.4	0.9	0.8	0.8	-0.7	0.6	0.8	1.0	0.1	0.9
Eskom	-0.9	0.5	0.7	0.6	-0.8	0.9	0.7	0.1	1.0	0.5
Energy	-0.0	1.0	1.0	1.0	-0.9	0.9	1.0	0.9	0.5	1.0

Figure 7: Results of correlation analysis showing

The heatmap in Figure 7 shows the connection among the top ten topics when issues of electricity access are raised on YouTube. Higher resemblance is shown by a deeper colour, and lower similarity is indicated by a lighter colour. Several of the top 10 topics have high similarity (black squares) on the heatmap, indicating that their content and emotion are comparable. Additionally, several pairs of the most popular topics show low similarity (light squares), indicating that their emotion and content diverge. The relationships between the top 10 topics that are similar to one another are made clearer to us by this visualisation.

4.8 Clustering

We performed a clustering analysis in addition to the correlation analysis to group comparable comments based on their sentiment scores. Figure 8 shows how K-means clustering, a machine learning technique, splits the data into K distinct, non-overlapping clusters. Each subject is allocated to the cluster whose centre, also called the centroid, was closest to the topic's sentiment score.
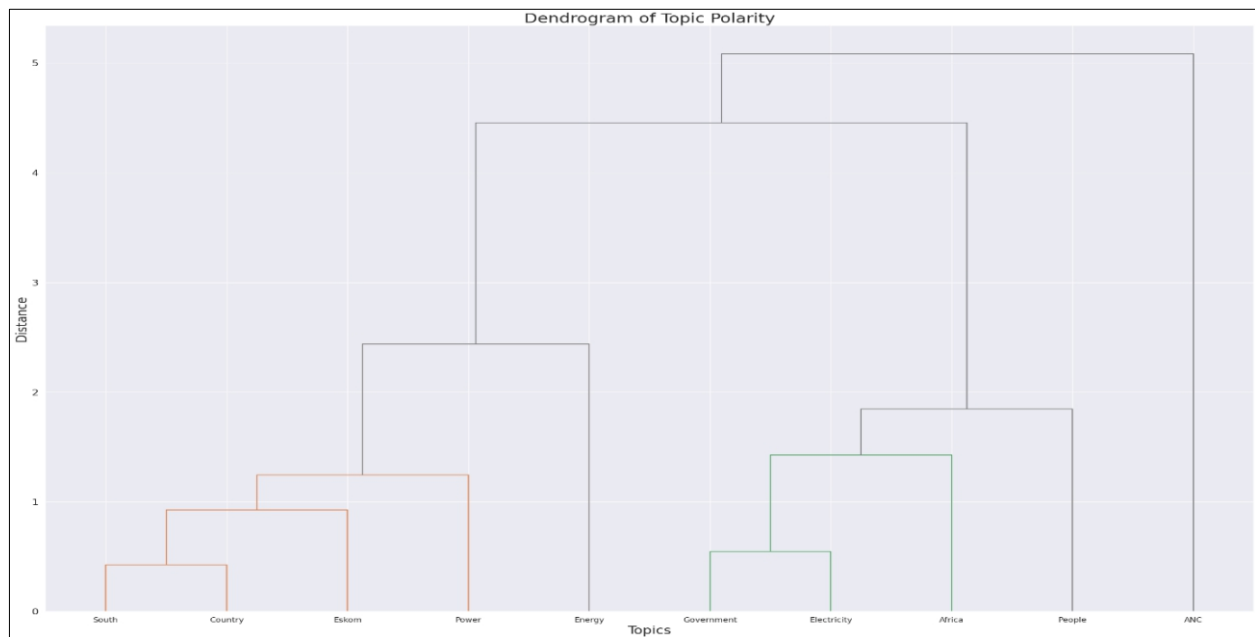
Figure 8: Dendrogram of Topic Polarity

The overall number of topics in each cluster was displayed in the clustering findings from Figure 8. The findings shed light on the connections between various subjects and the recurring themes or subjects in them. Due to their similarity in both content and sentiment, topics with comparable sentiments were grouped together. From Figure 8, "South", "Country", "Eskom", "Power" and "Energy" had positive sentiments and contents. On the contrary, "Government", "Africa" and "People" were topics with negative sentiments and contents. Only ANC connected neutral sentiments and content.

4.9 Degree of Centrality

The degree of centrality quantifies the quantity of links a node possesses with other nodes within the network (Chiranjeevi et al., 2024). A node with a high degree of centrality in the topic network has several connections to other subjects, suggesting that it is a significant and prominent issue in the dataset. From Figure 9, "Anc", "Africa" and "Country" were at the centre of centrality of the sentiments expressed on the YouTube channel. Thus, digital discourse on access to electricity cannot happen without the mention of the ANC party, Africa as a region and the particular country in view. It is therefore advisable that policymakers keep in tune with all the discourses that happen on access to electricity in SSA without reference to these three topics.
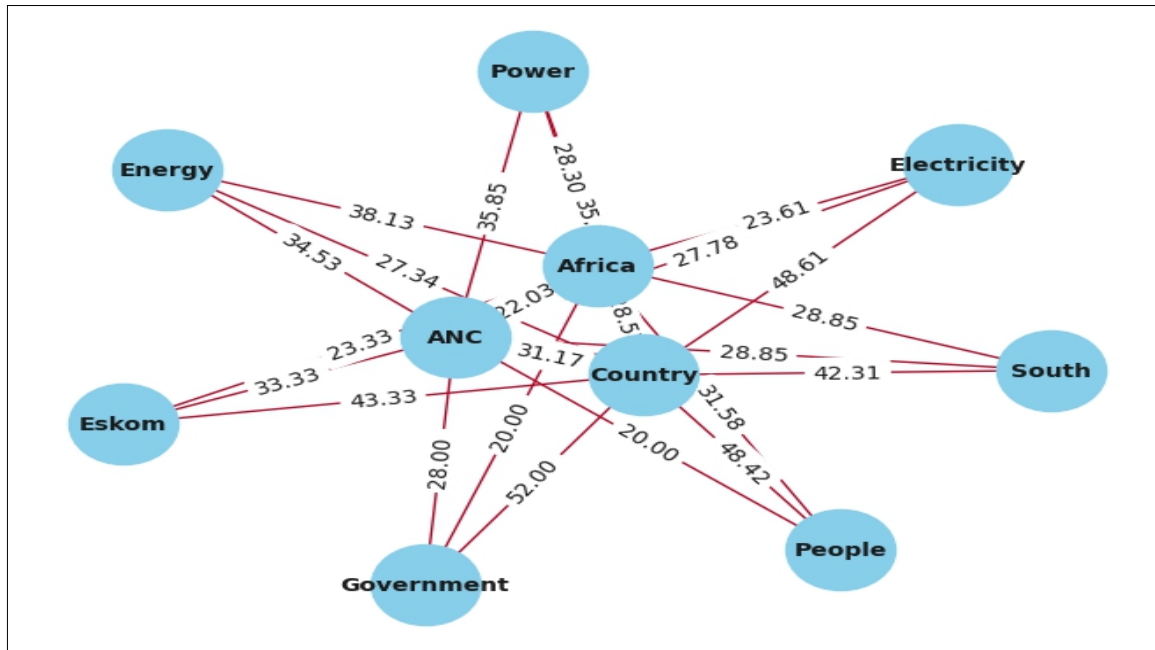
Figure 9: Network output of top ten topic

Among the three topics of centrality, 81.77% was derived from their total sum on "Anc", "Africa" and "Country". Thus, the three topics dominate most digital discussions on access to electricity in SSA countries. In the case of the subsidiary nodes, the degree of connection between one node and another fell below 50 % except for the connection between government and country where the connection was 52% however decreased to 48.61% when extended to electricity issues in SSA countries.

4.6 Discussion

Sentiment analysis is the process of looking at and evaluating people's ideas, feelings, assessments, attitudes and emotions as they relate to a certain topic in written documents. Politicians, investors, and consumers can make more strategic decisions if they are aware of the trends in areas that are important to society, like access to energy. In this study, we used comments gathered from YouTube to analyse how society responded to the availability of electricity in Sub-Saharan African nations. Following an analysis of the data gathered from YouTube, the following points were found:

- Electricity generation is core to the growth and development of SSA countries against the backdrop of growing comments on social media platforms, particularly YouTube.

- Public digital sentiments on access to electricity vary among YouTube users thereby reflecting the level of heterogeneity among the population in SSA countries.

- Though public sentiments were skewed in the positive sense, the public was silent on the source of electricity generation (renewable and non-renewable sources). This is very pertinent considering the Sustainable Development Goal (SDG) 7 and African Goal 63 which emphasizes providing affordable, efficient and reliable energy. In many SSA countries, there is an abundance of solar energy which can serve as a reliable source of electricity generation however public sentiments on YouTube did not factor in the source of generation. In light of the goal of pursuing a regional goal of a carbon-neutral region, public sentiments on the YouTube platform were oblivious to the electricity generation contributes to carbon emission in the SSA region or not. Other countries within the region rely on hydro as the mainstay of generating electricity for household consumption. With the growing interest in exploring alternative ways of generating electricity in SSA countries, some have also begun to explore wind as a potential source of generating electricity however the public sentiments did not extend to emerging sources of electricity generation.

- Positive sentiments dominate negative sentiments on issues of electricity access in SSA countries suggesting that the public foresees that the region has the prospective of enhancing the region's access to electricity. However, it became clear that the public sentiments in SSA countries are mostly emanating from South Africa with the discovery of "Anc", "Africa" and "South" as topical issues in the discourse of electricity.

- Together "Anc", "Africa" and "Country" have approximately 81.77% connections on access to electricity. This suggests that the discourse of access to electricity cannot take place in SSA countries forgoing the African continent and countries within the SSA region. The presence of "Anc" gives the impression that political parties play critical roles in the provision of electricity to households in SSA countries.

Public opinions on access to electricity in SSA countries can also be found through an examination of comments expressed by YouTube users. With these evaluations at their disposal, policymakers may be able to influence public opinion regarding electricity accessibility and possibly even accelerate the process. Through organised educational and informational initiatives, people's perceptions of access to electricity can be changed. Within the context of this study, data gathered between 2014 and 2024 generated an 81.77% degree of centrality that was

obtained validates that public discourse on the YouTube platform is impossible to have on its own without consulting the individual SSA countries, each of which faces unique difficulties in providing electricity. The results imply that throughout time, social perceptions may alter as a result of education, technology development and information transmission.

Unlike previous research (Zarrabeitia-Bilbao et al., 2023; Lammers et al., 2023; Kolosok, Chygryn & Onyshchenko, 2022), we did not use social media platforms like Twitter data for our analysis; instead, we used data from YouTube because it included many technical details, no character limit and was easy to acquire historical data. It has been determined that YouTube data are more useful for analysing significant aspects of electricity accessibility, such as cost, risk and environment, given the wealth of technical data available on the subject. As is the case with data from all social media networks, it is not possible to say that the data we used accurately depicts all people. Analysing the data gathered from YouTube, it can be shown that there is a dearth of digital information about the sources used in the SSA countries to generate electricity. Data from various social media platforms can be gathered to enhance the study and provide a more thorough examination of access to electricity. Although prior research has frequently employed single-labeled ternary classification models (Nandwani & Verma, 2021; Ruz, Henríquez & Mascareño, 2020), our investigation employed a triple classification model (positive, neutral, and negative) to analyse digital feelings on the YouTube platform.

5.0 Conclusion and Policy Implications

This study investigated public perspective on electricity access in SSA countries. To address this issue, we gathered posts and comments from YouTube on electricity access issues and topics. The study collected dataset comprised 883 comments from January 2014 to July, 2024. We conducted five (5) major analysis – sentiment analysis, topic modelling, correlation and cluster analysis and network anlysis. We conducted a word cloud, sentiment polarity, subjectivity analysis and robustness check as part of the sentiment analysis. Regarding the word cloud, dominance of "Anc" and "Africa" suggest heightened discussions as well as reflect not only national concerns but also regional efforts to improve infrastructure and energy policies towards access to electricity. In policy terms, the high prominence of governance-related themes "ANC" and regional groupings "Africa" underscores political parties and regional cooperation in tackling energy deficits in the sub-region.

For the polarity of public sentiments on access to electricity, the majority of SSA sentiments were positive, next neutral and finally negative. There are less unfavourable opinions about power access in the bulk of public conversation. The greater percentage of favourable sentiments suggests that people are aware of possible improvements in the supply of electricity, whilst the neutral sentiments point to continuing conversations devoid of strong viewpoints. Even though they are the least common, the unfavourable sentiments draw attention to the difficulties and annoyances that people currently face when trying to access electricity. It can be concluded that there is a general feeling of hope in spite of the ongoing low access to electricity in SSA. This could be attributed to frantic effrots of SSA government's new energy infrastructural investments and legislative initiatives meant to enhance the availability of electricity. Negative sentiments serve as a reminder that issues like outages, pricing and unequal distribution still exist, even when some conversations are factual rather than opinionated. The encouraging sentiments is a cue for policymakers to know that current initiatives might be working well and laying the groundwork for future advancements. However, the presence of unfavourable sentiments draws attention to issues that need immediate solution, like unstable supply and affordability issues. The ambivalent opinions highlight the necessity of ongoing public participation and information sharing to guarantee that conversations around power access stay productive and focused on finding solutions.

Regarding power access in SSA, the vast majority of the YouTube comments that were collected were opinions rather than factual assertions. Rather from being unbiased evaluations of the state of electricity in SSA nations, the majority of comments are opinion-based and may

represent personal optimism or biases. Although factual debates are not as prevalent, the existence of certain verifiable, evidence-based remarks suggests that they do occur. The necessity of properly assessing social media conversation is highlighted by this, as opinions offer important insights into public sentiment but should not be confused with factual data. Since public opinions impact policy discussions and investment choices, it is crucial to balance popular opinion with factual information when recommending policies. The preponderance of opinion-based criticism indicates that more verified facts about energy availability has to be disseminated in order to ensure that discussions are informed.

The correlation analysis heatmap reveals that high positive correlation between "Electricity" and words such as "Africa," "South," "Power," "People," and "Energy" suggests that discussions around electricity are deeply intertwined with broader issues of energy distribution, regional considerations and public engagement. This reflects a general public concern about access and energy security in SSA, as electricity is often associated with conversations about infrastructure, development, and its social impact. However, the low positive association with "ANC" and "Eskom" indicates that, although these terms are partly related to electricity, they do not show up as prominent themes in this context. However, compared to more general energy and societal issues, political and institutional discourses are implied to be less strongly linked to broader talks about electricity. Public scepticism about the government's role in addressing electricity challenges is reflected in the high negative correlation between "Electricity" and "Government," which suggests an inverse relationship and indicates that discussions about electricity access frequently deviate from government action. This highlights a perceived disconnect between policy efforts and actual service delivery.

For policymakers interested in finding solutions to households with prolonged electricity shortages, these results highlight the necessity of focused interventions in specific regions of SSA nations. Negative feelings indicate annoyance with continuous outages and affordability concerns, according to sentiment research, even though public opinion is still generally positive about increases in energy access. In response, policymakers should prioritise investments in energy infrastructure in high energy straved areas thereby ensuring that energy is distributed fairly as well as increase subsidies for low-income households. Furthermore, the strong inverse relationship between "Electricity" and "Government" suggests a gap between what the public expects from the government and what it really does, underscoring the need for open, accountable policies and communication tactics to restore confidence.

In the area of policy response, the robustness check confirms the high percentage of opinion-based conversations on electricity access, indicating that personal experiences rather than empirical data drive public opinion. This implies that governments and other energy players have a critical role to play in providing accurate, fact-based information regarding power programs, problems and policy objectives. Public opinion analysis should be used by policymakers to create interventions that specifically address issues with pricing, governance transparency, and service dependability.

The correlation heatmap shows that some topics differ in how they are framed emotionally and thematically. Some debates on access to electricity deal with different facets of the problem, while others offer unique perspectives and feelings. Public opinion on electricity access is complex; some debates highlight common issues, while others offer fresh viewpoints. The groups of very similar subjects show recurrent themes that guide discussions and represent common issues and lobbying initiatives. In the meantime, the existence of low-similarity topics indicates that a wide variety of communities' and stakeholders' experiences, worries, and priorities are covered in public discourse. It is possible to detect prevailing public concerns and implement focused actions and policy responses by identifying which themes often match. On the other hand, knowledge of low-similarity themes sheds light on neglected or marginalised problems that need more research. This analysis also emphasises the significance of a nuanced approach to public participation, making sure that policies pertaining to electricity access take into account both commonly held concerns and important but little-discussed issues.

This study has important ramifications for expanding access to electricity in SSA nations since it gives decision-makers data-driven insights into public opinion. Examining public opinion around electricity access is essential to reaching Sustainable Development Goal (SDG) 7, which calls for universal access to modern, affordable, and dependable energy. In light of the growing focus on regional and global energy objectives, policymakers should develop and execute focused energy policies that not only increase access to electricity but also alleviate the inequalities between impoverished urban and rural households. Real-time feedback from public mood analysis, especially via social media, helps decision-makers match energy investments to the demands of the community. Additionally, by combining fundamental analysis techniques, the study identifies key themes and trends in electricity discourse, enabling policymakers to prioritise investment in off-grid renewable solutions, energy infrastructure improvements and affordability mechanisms like subsidies or pay-as-you-go systems. The ability to distinguish between objective and subjective discussions also ensures that policies are informed by both empirical data and

public opinion, strengthening governance and public trust. Finally, this research provides a framework for evaluating immediate public reactions to energy policies, assisting policymakers in adjusting strategies accordingly. The results also show how governance and electricity access are related, highlighting the necessity of regional cooperation, institutional accountability, and coherent policies in tackling power outages. The clustering and correlation analysis also highlights important energy-related issues, helping policymakers decide which issues need more research and prompt action. By utilising these insights, policymakers should proactively endeavour to supply underprivileged communities with inexpensive, dependable and efficient power, making sure that policies are not only technically sound but also in line with the public's goals and lived experiences.

5.1 Limitations and Future Recommendations

This study has some significant limitations even if it has various implications, as previously indicated. Firstly, the analysis did not take demographic data into account. The age group that utilises social media most frequently may be favoured by the fact that we used data from YouTube, a topic-oriented social media platform. Second, geographical and sociodemographic data were not taken into account in this study. We were unable to include the sociodemographic profiles and geographic locations of the users on the chosen social media platforms in the analysis since it was not possible to determine them. Future studies should incorporate socio-demographic data because prior research has demonstrated that regional variations play a significant role in shaping public opinions toward obtaining electricity (Blimpo, Postepska & Xu, 2020). Future researchers could benefit from taking into account users' regional features and socio-demographic data. Utilising pertinent hashtags like "region" and "age" to extract user demographic data from other social media platforms can offer greater insights into how various groups see and discuss access to and use of electricity in SSA nations. Second, even if our study's findings might have offered a reliable interpretation of public opinions, more research is still required so that policymakers can make informed decisions. Nonetheless, the suggested approach might be appropriate for examining current and emerging public opinion patterns via social media platforms. The results of this study also demonstrate how different social media datasets can be used to give semantic interpretations. We discovered that most conversations on access to power are likely to mention certain terms regularly. In light of these findings, future studies ought to employ more intricate and multiplicative analyses in order to provide a clearer picture of the public's attitudes regarding the availability of power in SSA nations and other areas with low levels of energy (Chirwa & Odhiambo, 2020).

References

Arnold, M. V., Dodds, P. S., & Danforth, C. M. (2023). Curating corpora with classifiers: A case study of clean energy sentiment online (arXiv:2305.03092). arXiv. https://doi.org/10.48550/arXiv.2305.03092

Asongu, S. A., & Le Roux, S. (2024). Governance, debt service, information technology and access to electricity in Africa. *International Journal of Finance & Economics*.

Banerjee, R., Mishra, V., & Maruta, A. A. (2021). Energy poverty, health and education outcomes: evidence from the developing world. *Energy economics*, *101*, 105447.

Blimpo, M. P., Postepska, A., & Xu, Y. (2020). Why is household electricity uptake low in Sub-Saharan Africa?. *World Development*, *133*, 105002.

Blimpo, M. P., Postepska, A., & Xu, Y. (2020). Why is household electricity uptake low in Sub-Saharan Africa?. *World Development*, *133*, 105002.

Boulianne, S. (2020). Twenty years of digital media effects on civic and political participation. *Communication research*, *47*(7), 947-966.

Butkowski, C. P., Dixon, T. L., Weeks, K. R., & Smith, M. A. (2020). Quantifying the feminine self (ie): Gender display and social media feedback in young women's Instagram selfies. *New Media & Society*, *22*(5), 817-837.

Byaro, M., & Mmbaga, N. F. (2022). What's new in the drivers of electricity access in sub-Saharan Africa?. *Scientific African*, *18*, e01414.

Chen, X., & Xie, H. (2020). A structural topic modeling-based bibliometric study of sentiment analysis literature. *Cognitive Computation*, *12*, 1097-1129.

Chiranjeevi, M., Dhuli, V. S., Enduri, M. K., Hajarathaiah, K., & Cenkeramaddi, L. R. (2024). Quantifying Node Influence in Networks: Isolating-Betweenness Centrality for Improved Ranking. *IEEE Access*.

Chirwa, T. G., & Odhiambo, N. M. (2020). Electricity consumption and economic growth: New evidence from advanced, emerging and developing economies. *International Journal of Energy Sector Management, 14*(1), 1-19.

Cooper, C. R. (2023). The identification of YouTube videos that feature the linguistic features of English informal speech. *Applied Corpus Linguistics*, *3*(3), 100068.

Corbett, J., & Savarimuthu, B. T. R. (2022). From tweets to insights: A social media analysis of the emotion discourse of sustainable energy in the United States. *Energy Research & Social Science*, *89*, 102515.

Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. *Electronics*, *9*(3), 483.

de Abajo Llamero, C. (2023). Adding a new analysis framework to support large scale electrification financing plans in countries lacking universal electricity access.

Dehler-Holland, J., Okoh, M., & Keles, D. (2022). Assessing technology legitimacy with topic models and sentiment analysis–The case of wind power in Germany. *Technological Forecasting and Social Change*, *175*, 121354.

Gilardi, F., Gessler, T., Kubli, M., & Müller, S. (2022). Social media and political agenda setting. *Political communication*, *39*(1), 39-60.

Hargreaves, T., & Middlemiss, L. (2020). The importance of social relations in shaping energy demand. *Nature Energy*, *5*(3), 195-201.

Isoaho, K., Gritsenko, D., & Mäkelä, E. (2021). Topic modeling and text analysis for qualitative policy research. *Policy Studies Journal*, *49*(1), 300-324.

Jain, R., Gupta, M., Taneja, S., & Hemanth, D. J. (2021). Deep learning based detection and analysis of COVID-19 on chest X-ray images. *Applied Intelligence*, *51*, 1690-1700.

Jeong, S. W., Ha, S., & Lee, K. H. (2021). How to measure social capital in an online brand community? A comparison of three social capital scales. *Journal of Business Research*, *131*, 652-663.

Kaur, P., & Edalati, M. (2022). Sentiment analysis on electricity Twitter posts. *arXiv preprint arXiv:2206.05042.*

Kim, S. Y., Ganesan, K., Dickens, P., & Panda, S. (2021). Public sentiment toward solar energy—opinion mining of twitter using a transformer-based language model. *Sustainability, 13*(5), 2673.

Klinkhammer, D. (2022). Sentiment Analysis with R: Natural Language Processing for Semi-Automated Assessments of Qualitative Data. *arXiv preprint arXiv:2206.12649.*

Kolosok, S., Chygryn, O., & Onyshchenko, S. (2022, June). Leadership and Global Energy Security: Sentiment Analysis of Twitter Using a Named Entity Recognition. In *PRIZK International Conference-Novel Insights in the Leadership in Business and Economics After the COVID-19 Pandemic* (pp. 361-375). Cham: Springer Nature Switzerland.

Lammers, R., Jansma, S. R., Veldkamp, B. P., Machens, A. K., de Visser, M., & Gosselt, J. F. (2023). 'I Tweet about Our# GreenEnergy'—Automated Classification of Social Identity and Opinion Mining of the Dutch Twitter Discourse on Green-Energy Technologies. *Sustainability, 15*(22), 16106.

Li, L., Goh, T. T., & Jin, D. (2020). How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis. *Neural Computing and Applications, 32*, 4387-4415.

Liu, W. J., Liu, P., Lei, W., Jia, Z., He, X., Shi, W., ... & Wu, G. (2024). Surveillance of SARS-CoV-2 at the Huanan seafood market. *Nature, 631*(8020), 402-408.

Loureiro, M. L., & Alló, M. (2020). Sensing climate change and energy issues: Sentiment and emotion analysis with social media in the UK and Spain. *Energy Policy, 143*, 111490.

Macanovic, A. (2022). Text mining for social science–The state and the future of computational text analysis in sociology. *Social Science Research, 108*, 102784.

Marrapu, S., Senn, W., & Prybutok, V. (2024). Sentiment Analysis of Twitter Discourse on Omicron Vaccination in the USA Using VADER and BERT. *Journal of Data Science and Intelligent Systems.*

Munezero Buzingo, S. (2021). Challenges and opportunities in developing a common legal and regulatory framework on energy in East African Community.

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social network analysis and mining, 11*(1), 81.

Negara, E. S., & Triadi, D. (2021). Topic modeling using latent dirichlet allocation (LDA) on twitter data with Indonesia keyword. *Bulletin of Social Informatics Theory and Application, 5*(2), 124-132.

Nolasco, D., & Oliveira, J. (2020). Mining social influence in science and vice-versa: A topic correlation approach. *International Journal of Information Management, 51*, 102017.

Noonan, C. F., & Piatt, A. W. (2015). *Global Social Media Directory. A Resource Guide* (No. PNNL-23805 Rev. 1). Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Nuortimo, K., Karvonen, E., & Härkönen, J. (2020). Establishing social media firestorm scale via large dataset media analytics. *Journal of Marketing Analytics, 8*(4), 224-233.

Ovchinnikova, S., Schneider, E., Borovkov, A., & Lyamina, A. (2021, December). Energy potential of modern wind turbines in the Krasnodar Krai. In *IOP Conference Series: Earth and Environmental Science* (Vol. 937, No. 4, p. 042049). IOP Publishing.

Palomino, M. A., & Aider, F. (2022). Evaluating the effectiveness of text pre-processing in sentiment analysis. *Applied Sciences, 12*(17), 8765.

Phokeer, A. D. (2021). Improving content delivery in low-resource networks: a case study of the African Internet Ecosystem.

Qazi, A., Hussain, F., Rahim, N. A., Hardaker, G., Alghazzawi, D., Shaban, K., & Haruna, K. (2019). Towards sustainable energy: a systematic review of renewable energy sources, technologies, and public opinions. *IEEE access, 7*, 63837-63851.

Razi, F., & Dincer, I. (2022). Renewable energy development and hydrogen economy in MENA region: A review. *Renewable and Sustainable Energy Reviews, 168*, 112763.

Renewable Energy Discourse on Social Media: A Cross-Platform Sentiment Analysis. *Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi, 26*(2), 466-480.

Ruz, G. A., Henríquez, P. A., & Mascareño, A. (2020). Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. *Future Generation Computer Systems, 106*, 92-104.

Saud, M., Ida, R., Abbas, A., Ashfaq, A., & Ahmad, A. R. (2020). The social media and digitalization of political participation in youths: An Indonesian perspective. *Society, 8*(1), 83-93.

Senevirathne, H. U., Kumara, B. T. G. S., & Banujan, K. (2024, February). YouTube Video Categorization Based on Closed Captions. In *2024 4th International Conference on Advanced Research in Computing (ICARC)* (pp. 212-217). IEEE.

Sievert, M., & Steinbuks, J. (2020). Willingness to pay for electricity access in extreme poverty: Evidence from sub-Saharan Africa. *World Development, 128*, 104859.

Söğüt, Y., & Öngel, A. (2022). Examination of TikTok application in the axis of uses and gratifications approach, user experience and entertainment culture. *Erciyes İletişim Dergisi, 9*(2), 667-690.

Stocker, V., Smaragdakis, G., Lehr, W., & Bauer, S. (2017). The growing complexity of content delivery networks: Challenges and implications for the Internet ecosystem. *Telecommunications Policy, 41*(10), 1003-1016.

Vågerö, O., & Zeyringer, M. (2023). Can we optimise for justice? Reviewing the inclusion of energy justice in energy system optimisation models. *Energy Research & Social Science, 95*, 102913.

Valickova, P., & Elms, N. (2021). The costs of providing access to electricity in selected countries in Sub-Saharan Africa and policy implications. *Energy Policy, 148*, 111935.

Valickova, P., & Elms, N. (2021). The costs of providing access to electricity in selected countries in Sub-Saharan Africa and policy implications. *Energy Policy, 148*, 111935.

Vespa, M., Kortsch, T., Hildebrand, J., Schweizer-Ries, P., & Volkmer, S. A. (2022). Not all places are equal: using instagram to understand cognitions and affect towards renewable energy infrastructures. *Sustainability, 14*(7), 4071.

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review, 55*(7), 5731-5780.

Zafar, A. U., Shen, J., Ashfaq, M., & Shahzad, M. (2021). Social media and sustainable purchasing attitude: Role of trust in social media and environmental effectiveness. *Journal of Retailing and Consumer Services, 63*, 102751.

Zarrabeitia-Bilbao, E., Jaca-Madariaga, M., Rio-Belver, R. M., & Álvarez-Meaza, I. (2023). Nuclear energy: Twitter data mining for social listening analysis. *Social Network Analysis and Mining, 13*(1), 29.

Zogo, T. E., Mbassi, C. M., & Asongu, S. A. (2024). Coup d'état and access to electricity in sub-Saharan Africa. *The Electricity Journal, 37*(4-5), 107406.

Profile of Authors

Mr Pius Gamette is a PhD Candidate at the University of South Africa in Energy Economics and holds Master of Philosophy (Economics) from the University of Cape Coast, Ghana. He is currently with the Scholarly Communication Unit at the University of Cape Coast. His research interests are Energy Economics, Climate Change and Health Economics. He is an active member of the African Economic Research Consortium (AERC), Kenya. Mr Pius Gamette can be contacted at: pius.gamette@ucc.edu.gh

Prof. Simplice Asongu holds a PhD from Oxford Brookes University and is currently the Lead Economist and Director of the African Governance and Development Institute (Yaounde, Cameroon), the Lead Economist and Director of the European Xtramile Centre of African Studies (Liege, Belgium) and Lead Economist and Co-director of "Centre de Recherche pour le Developpement Economique" (Bangui, Central African Republic). He is also a Senior Research Fellow at the Africa Growth Institute (Cape Town, South Africa); PhD Supervisor for several universities in Nigeria, Ghana, Botswana, Zimbabwe and South Africa. Simplice A. Asongu is Visiting Professor in The Institute of Convergence Science, Korea University, Seoul, South Korea. He is also a Visiting Professor in the Tashkent State University of Economics, Tashkent, Uzbekistan. Simplice Asongu is the corresponding author and can be contacted at: asongusimplice@yahoo.com