

# A Data Analytic Approach in the Thematic Classification of the Reasons and Perspectives of Adolescents' Social Media Engagement



Elaine Marie Aranda<sup>1</sup>, Roy Francis Navea<sup>2</sup>

<sup>1</sup>De La Salle University, Philippines, [elaine.aranda@dlsu.edu.ph](mailto:elaine.aranda@dlsu.edu.ph)

<sup>2</sup>De La Salle University, Philippines, [roy.navea@dlsu.edu.ph](mailto:roy.navea@dlsu.edu.ph)

## ABSTRACT

Social media is one of the leading platforms where people and organizations meet. With the technological advancements in the world wide web, smart mobile devices and Internet connectivity, an increase in social media engagement is highly observable. People of all ages, one way or the other, has a social media account and their perspective, reasons, and content-preferences vary. In this study, the experience of social media engagement from the perspective of young people is analyzed and thematically classified using a data analytic approach which focuses on natural language processing (NLP). Results show that the social media engagement experience of the respondents reflects what social media is to them and for them expressed by their reasons and perspectives, respectively. The reasons of their engagement are basically connected to the contents and features of social media platforms that suit their purpose, intention, and goals of engagement expressed by textual response and analyzed by a learning algorithm that fits multiclass models for support vector machines (SVM). The profound social media engagement of the youth leads to a wide spectrum of responses and behaviors that would affect their mental health as defined by their reasons and perspectives.

**Key words:** Data analytics, engagement, perspectives, reasons, social media

## 1. INTRODUCTION

Social media has been a part of the 21<sup>st</sup> century lifestyle and has been an avenue for great opportunities in terms of interaction with people from different parts of the world. Social media platforms are interactive computer interface technologies that facilitate and allow users to create and share information, ideas, career or business interests, and other forms of expression thru virtual communities and

networks. By design, social media is based on the Internet, electronic in nature, and is accessible by fixed and mobile devices connected online [1]. People engage in this platform using their computer units, smartphones, or tablets using a web application to interact with friends, family, and later adopted for business, leisure, religion, politics, and government. The 2020 social media statistics shows an estimate of 3.6 billion social media users around the world [2] and adolescents are found to be the avid users of online social networking sites taking a larger percentage as compared to adult users [3], [4].

Social media affects the well-being of a person with its rapid development in recent years together with the advances in Internet connectivity and availability of smart mobile devices. The engagement of people in social media matters as it brings potential benefits and risks depending on the perspective of the user and the reason of engagement on a general or specific social media content. The frequent use of social media and having the urge to check notifications are signs of constant curiosity and have been linked to anxiety and/or depressive symptoms [5]. Seeing the frequent updates of social media friends and the possibility of encountering negative comments from personal posts may contribute to the development of negative emotions [6]. From whatever device a social media account is accessed, the amount of time spent with it was found to be positively correlated to higher level of anxiety, belittled well-being, low self-esteem, and increased feelings of depression, cognitive overload, stress, and physique dissatisfaction [7]–[10]. In effect, other important tasks are not given enough priority as time for them were allotted browsing the social media. The longer time users spend on social media platforms reflects a larger instance that the individuals' feeling of not having completed anything meaningful becomes pronounced leading them to depressive symptoms such as anxiety, general discontent, or lack of motivation [11]. It is possible that frequent social media use fulfills adolescents' need for connectedness but alternatively social activity could also signify feelings of loneliness and anxiety [12]. Internet addiction in adolescents is more likely related to anxiety symptoms than substance dependence or obsessive-

compulsive symptoms which can lead to indulgence on the Internet. Excessive use of internet and social networking platforms could weaken the connection between individuals. As a result of that, individuals may feel lonelier and depressed [13]. The number of social media accounts can be a reason for anxiety. It is correlated with the level of anxiety due to the overwhelming demand for feedback from those accounts [14]. The link between social media and mental health problems is based on lack of social support from significant relationships and they tend to look for someone online to approve their behavior and fulfill their needs for belongingness [15].

Adolescence is a transition time between childhood and adulthood. It is a period of desire for connectedness with peers. With this, they tend to seek out social media to look for new friends, as they might have difficulty in expressing themselves in face-to-face encounters with their peers, or they may feel that they lack significant social adeptness. On one hand, social media is a place they feel more comfortable to express themselves and allow them to be anonymous by disguising their identity or creating a false impression on creating a completely false persona [12]. Moreover, the more they are wasting time in social media without a clear purpose, the larger less they are spending time with their family members, leading to tension with their parents and family members or loss of significant relationships because of the Internet.

Online platforms of social interaction only allow seeing and hearing posts, comments, and any other media contents. Hence, the full functionality of the human senses to understand and comprehend these social media contents is subject to misinterpretation and misunderstanding [16]. This social media constraint can induce anxiety and expose vulnerable users to negative online experiences [17]. Many adolescents are experiencing this when exposed to social media contents and engaged in its usage on a regular basis. Adolescents who are not properly guided may eventually feel empty, lonely and may gradually turn to the use of social media as an escape route from their problems.

Given these circumstances, this paper aims to explore the social media engagement experience of young people in terms of their perspective and reasons of engagement in a qualitative and quantitative approach using data analytics and machine learning algorithms. These are known methods in quantifying qualitative data for analysis and knowledge building [18], [19]. This could provide a better understanding of their situation and capture their young insights about social media expressed through their responses in texts, analyzed, and given a learned classification and interpretation.

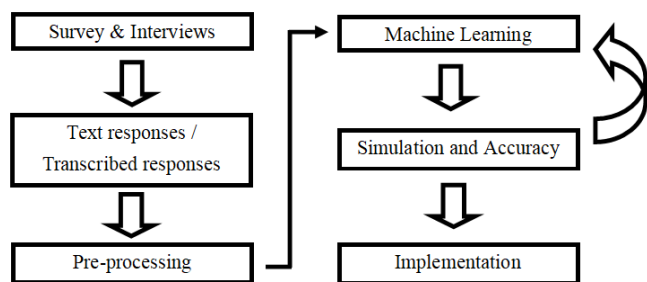
Social media is one of the most common platforms of communication, entertainment, and leisure. Browsing social media contents consumes time without even knowing it because of its interest-stirring mechanism that captures people specially the young. With the availability of mobile devices and Internet connectivity, the young people are hooked-up with their devices and browsing social media sites with a different set perspective, reasons why they engage less or a lot depending on the contents that captures their attention, satisfies their needs, or fills their emptiness. In this study, the experience of social media engagement from the perspective of young people is explored with the objective of determining their reasons and perspectives based on their textual response to survey questionnaires and interviews.

Qualified participants are 18 to 21 years of age with a minimum social media engagement of 6 hours in a day. A minimum of 6 participants was interviewed with a balanced count of male and female respondents. Google forms were utilized to gather data and interview details were recorded and transcribed. Themes were formed from the responses and were used to attain the objectives as well as provide other relevant information about the social media engagement experience of young people.

The young people may have different perspectives on how they look at social media. Some may look at it as a way to reach-out, meet new friends, and acquaint themselves with others. Some may see at it as a learning portal where wide range of contents are available for directed self-learning and intellectual investment. Some may perceive at it as an opportunity to market, grow their business, promote their goods, and expand their network. But for some, social media serves as an avenue in releasing their burst of emotions or a channel to get help and support. The perspectives of young people vary, and they contribute to their level of engagement that provides a sense of purpose why they hook themselves in a short or long duration of time. The wide range of social media contents contribute to the level of their engagement as interest is being stirred up by posts coming from different purpose-driven origins. Guided surveys and interviews were used to gather information from the respondents. through a

## 2. METHODOLOGY

To describe the social media experience of the youth, an online questionnaire was created, and a virtual interview was conducted. Figure 1 shows the data analytic approach for thematic classification which quantifies qualitative information from surveys and interviews.



**Figure 1:** Data Analytic Approach for Thematic Classification

**2.1 Survey Questionnaire and Interview**

The survey questionnaires were prepared using Google forms and distributed online by sharing the link to the target respondents. To qualify, the respondent should be 18 to 21 years of age who is using social media apps for a minimum of 6 hours. A minimum of 6 participants was interviewed with a balanced count of male and female respondents. An introductory message is initially presented in the questionnaire that explains the rationale of the study and the importance of the voluntary participation. Consent was obtained as the respondents continue since communication information will be asked for the virtual interview. The demographics of the respondents were asked including their name, age, nationality, gender, and whether they are studying or not.

For the social media information proper, it is important to determine if they are patrons of social media apps and what are those apps that they are using. If more than one, the list was asked in order of priority in terms of usage. The period of use is a determinant for them to qualify in the virtual interview, hence, a minimum of 6 hours per day was set. The frequency of social media usage was asked on a week before the actual day when they answered the survey. The usage frequency was blocked as not once last week, once last week, twice last week, up to everyday last week. Instances considered are between waking-up and going to bed.

The personal experiences of the respondents were explored by asking them to describe their social media use in a day as well as their eagerness to log-in or check their accounts. The most common platform was asked, and they shared their reasons why they are were opted to check updates and look at its contents. The respondents identified their preferred contents to look-at as well as the benefits that they can get from them.

The virtual interview was made using either Google meet, Zoom, Microsoft Teams, or through a mobile call. The conversations were recorded, and transcripts of the conversation was written down for analysis.

Data obtained from the survey and interview were kept in the Google drive of the proponents subject to the data privacy guidelines of the University.

**2.2 Data Analysis**

A total of 28 answered the survey and 6 of them, 3 males and 3 females, were found to be highly immersed in social media activities. A thematic analysis approach using Microsoft Excel was used following the procedures as proposed in [20]. Thematic analysis is a well-known method used in qualitative research applied to a set of text or interview transcripts. Common words, phrases, ideas, patterns, or topics were grouped together to develop themes. Thematic approach includes familiarization, coding, generation of themes, review of themes, defining and naming themes, and doing the write-up [20]. Familiarization was performed by repeatedly reading and understanding the transcripts of the interview proper. It is important to understand the context of the words and phrases based on how the conversation went through during the interview. Coding was performed by segmenting sentences based on the meanings of the words or phrases uttered by the respondents in response to the interview questions. The codes were obtained by simplifying the segments that captures its essence. The codes were grouped according to similar thoughts and color-coded for better visual identification. With the similar codes obtained, themes were generated to describe the common denominator of the codes. The themes were revised several times to come up with better words to capture the codes.

**2.3 Thematic Classification and Machine Learning**

The thematic classification and machine learning is performed in MATLAB using the Text Analytics Toolbox. The pre-processed and coded text inputs from the respondents were saved in a .csv file together with their text description and categorial labels. These labels were converted in a category column and their histograms were obtained. Data partition for training was performed at a 30:70 ratio with 30% of the data used for training and 70% for testing. Text tokens were created as well as the bag of words. Documents in the bag of words containing no words were removed as well as their entry labels. A support vector machines (SVM)-based supervised classification model was used using the word frequency counts from the bag-of-words model and the labels [21], [22]. A multiclass linear classifier specifies the counts of the bag-of-words model to be the predictor, and the event type labels to be the response. Learners are linear and support sparse data input.

The labels of the test data were predicted using the trained model and the classification accuracy was calculated. Recurrent training was performed to attain an accuracy of

more than 90%. The model was tested with new sets of data to demonstrate its implementation.

### 3. RESULTS AND DISCUSSION

Accounts from 6 individuals who are from age 18-21 years old with more than 8 hours of social media exposure a day were selected. They are those who are using social media every day for the past week and skipping meals and excessively exposed to social media in between meals. The accounts were selected from 28 respondents of the structured Google form survey. The names of the respondents were changed to ensure confidentiality and protect their personal information.

The researchers read and re-read the transcripts several times, segments were made and coded. Initial descriptions of significant and consistent statements were noted, these were translated to themes. The purpose is to capture both the description and how the respondents experience social media engagement, where the core essence of the experience is.

The analysis presents themes with which the codes are grouped. The experiential themes represent the concept of young people of what social media engagement is and how young people make sense of and give reasons of their social media engagement. These includes the behavioral attachments in the use of social media. The themes also provide an array of healthy and unhealthy behaviors in the use of social media, and the social media content that makes young people engaged. The derived themes on the perspective of social media engagement are related to (1) cognitive process, (2) affective response, and (3) behavioral attachment.

#### 3.1 Reasons for Social Media Engagement

There are compelling reasons why people engage and immerse themselves in social media activities. Though social media is a virtual thing, it has the capability to integrate the online and offline experiences of people and organizations [23]. Figure 2 shows the histogram of the reasons which includes 1. As a distraction, 2. As a habit or an addiction, 3. As a learning experience, 4. As a positive outlet to reduce or release stress, 5. As a way to express opinion, 6. As a way to kill time, 7. A way to share or manage personal updates, 8. As a way to stay connected with or help people, 9. For fun and entertainment, 10. For updates and information, 11. Not equivalent to in-person interaction or experience, 12. Social media engagement as a less authentic experience, 13. Social media increases insecurities, 14. Increases motivation. The highest in occurrence in reasons are “as a habit or an addiction” and “as a way to stay connected with or help people”.

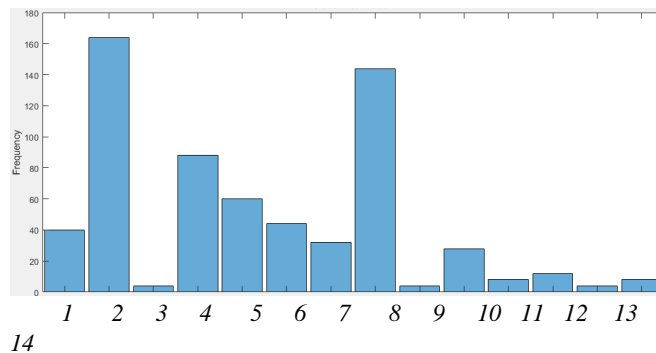


Figure 2: Histogram of Reasons for Social Media Engagement

#### 3.2 Social Media User Perspective

Adolescents actively engage in social media with cognitive, emotional, and behavioral perspectives [24]. These perspectives are validated by satisfaction, positive emotions, and for some, problematic and/or addictive social media use [24], [25]. Social media engagement refers to the extent of an individual immersing into social media activities. Social media engagement refers to the state of cognitive and emotional absorption in the use of social media tools [26]. The concept of engagement comprises cognitive, attitudinal, and behavioral attachment [27]. The segmented responses were coded and grouped with common themes and for this instance, provide answers to what is social media experience to young people. Figure 3 shows the themes of perspectives. Perspective themes are more on behavioral attachment to social media.

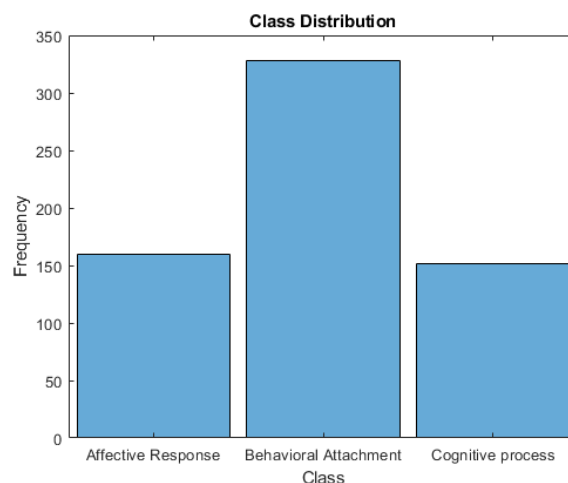


Figure 3: Histogram of Thematic Perspectives in Social Media Engagement

*Cognitive Response:* Social media engagement is seen as a source of information, a means by which one gets connected and updated as well as a utility to express opinion, interact and learn.

**Affective Response:** Social media engagement is seen a source of entertainment, fun and excitement. A channel that gratifies drives, wants, needs, emotions, interests, motivation as well as insecurities. Social media engagement as an alternative to face-to-face interaction is seen to be less authentic and is not equivalent to in-person interaction.

**Behavioral Attachment:** Social media engagement is a means by which one gets distracted, a habit-forming activity that is addictive, and a pattern by which engagement becomes part of daily living and lifestyle.

### 3.3 Classification Results

With the iterative training process, the classifier accuracies obtained are 91.76% for the reasons class while 93.75% for the perspective class as shown in Table 1. The classifier is a full, trained, multiclass, error-correcting output codes model that uses  $K(K-1)/2$  binary support vector machines (SVM) models using the one versus-one coding design, where K is the number of unique class labels.

**Table 1:** Classification Accuracy Results

Classification Task	Accuracy
Reasons	91.76%
Perspective	93.75%

### 3.4 Synthesis of Themes and Key Points

In this study, the social media engagement experience of young people is studied in terms of their perspective and reasons of engagement. The cognitive process, affective response, and behavioral attachment reflect the perspective of the respondents. The cognitive process refers to the perspective wherein social media is seen as a catalyst of the mind as it provides information where people learn by interaction. Engagement is the progression from interacting with the interface physically to becoming cognitively immersed in the content offered by it and then onto proactively spreading the outcomes of this involvement [28]. The social engagement of the respondents is affective in the sense that they are affected, changed, or renewed accordingly as per what content they are engaged. Engagement as a multidimensional concept comprises not only of behavioral (actions) but also cognitive (thoughts), and emotional (feelings) aspects that are molded by the information coming from external sources like social media. As such, the behavior of a person engaged in social media has the potential to be altered at a faster rate as compared to those who are less engaged and exposed to it contents. Behavioral satisfaction leads to attachment. People who lack social support from significant relationships to look for someone online to approve their behavior and fulfill their needs for belongingness [15]. On one hand, social media is a place they feel more comfortable to express themselves

and allow them to be anonymous by disguising their identity or creating a false impression on creating a completely false persona [12]. Moreover, viewing videos and reading comments is included as a form of engagement. Users may choose to remain passive by simply consuming content or play an active role by participating in various interactions, and even repurpose content to fit their needs [17].

## 4. CONCLUSION

This study has been conceptualized because of the increasing number of social media platforms followed and used by people, specifically the young ones, the adolescents. Their immersive exposure to the use of social media may have impact on their personalities as well as in their relationships. Data for this study were taken from a range of social media users where young people shared their experiences of using social media including their concept of social media, the content that increases their engagement as well as their reasons for using social media.

The social media engagement experience of the respondents reflects their perspective as to what social media is to them and for them. The reasons of their engagement are basically connected to the contents and features of social media platforms as expressed in their textual response.

Natural language processing was used by creating a learning algorithm that can classify the themes of the reasons and perspective of the adolescents as they engage in social media platforms. Histograms were presented as well as the classification accuracy of an SVM-based classifier which yields more than 90%.

## REFERENCES

- [1] M. Dollarhide, **Social Media Definition**, *Dotdash*, 2020. <https://www.investopedia.com/terms/s/social-media.asp>.
- [2] J. Clement, **Number of global social network users 2017-2025**, *Statista*, 2020. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>.
- [3] A. Dobrea and C.-R. Păsărelu, **Impact of Social Media on Social Anxiety: A Systematic Review**, 2016.
- [4] J. Chen, **Social media demographics to inform your brand’s strategy in 2020**, *SproutSocial*, 2020. <https://sproutsocial.com/insights/new-social-media-demographics/>.
- [5] L. Wit, A. Stratan, F. Lamers, P. Cuijpers, and B. W. Penninx, **Are sedentary television watching and computer use behaviors associated with anxiety and depressive disorders?**, *Psychiatry Res.*, vol. 186, pp. 239–243, 2011.

- [6] L. J. Labrague, **Facebook use and adolescents' emotional states of depression, anxiety, and stress**, *Heal. Sci. J.*, vol. 8, no. 1, pp. 80–89, 2014.
- [7] W. Chen and K. H. Lee, **Sharing, liking, commenting, and distressed? the pathway between Facebook interaction and psychological distress**, *Cyberpsychology, Behav. Soc. Netw.*, vol. 16, no. 10, pp. 728–734, 2013.
- [8] H. T. G. Chou and N. Edge, **They are happier and having better lives than I am': The impact of using facebook on perceptions of others' lives**, *Cyberpsychology, Behav. Soc. Netw.*, vol. 15, no. 2, pp. 117–121, 2012.
- [9] D. A. de Vries, J. Peter, H. de Graaf, and P. Nikken, **Adolescents' Social Network Site Use, Peer Appearance-Related Feedback, and Body Dissatisfaction: Testing a Mediation Model**, *J. Youth Adolesc.*, vol. 45, no. 1, pp. 211–224, 2016.
- [10] E. Kross *et al.*, **Facebook Use Predicts Declines in Subjective Well-Being in Young Adults**, *PLoS One*, vol. 8, no. 8, pp. 1–6, 2013.
- [11] M. Wright, **Cyberbullying Victimization through Social Networking Sites and Adjustment Difficulties: The Role of Parental Mediation**, *J. Assoc. Inf. Syst.*, vol. 19, no. 2, 2018.
- [12] C. Barry, C. Sidoti, S. Briggs, S. Reiter, and R. Lindsey, **Adolescent social media use and mental health from adolescent and parent perspectives**, *J. Adolesc.*, vol. 61, pp. 1–11, 2017, doi: 10.1016/j.adolescence.2017.08.005.
- [13] K. Kircaburun, **Self-Esteem, Daily Internet Use and Social Media Addiction as Predictors of Depression among Turkish Adolescents**, *Journal of Education and Practice*, vol. 7, 2016.
- [14] G. Escobar-viera *et al.*, **Use of multiple social media platforms and symptoms of depression and anxiety: A nationally-representative study among U.S. young adults**, *Comput. Human Behav.*, vol. 69, pp. 1–9, 2017, doi: 10.1016/j.chb.2016.11.013.
- [15] B. Keles, N. McCrae, and A. Grealish, **A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents**, *Int. J. Adolesc. Youth*, vol. 25, no. 1, pp. 79–93, 2020, doi: 10.1080/02673843.2019.1590851.
- [16] P. S. N. Lee, L. Leung, V. Lo, C. Xiong, and T. Wu, **Internet Communication Versus Face-to-face Interaction in Quality of Life**, *Soc. Indic. Res.*, vol. 100, no. 3, pp. 375–389, 2011.
- [17] O. Calancie, L. Ewing, L. D. Narducci, S. Horgan, and S. Khalid-Khan, **Exploring how social networking sites impact youth with anxiety: A qualitative study of facebook stressors among adolescents with an anxiety disorder diagnosis**, *Cyberpsychology*, vol. 11, no. 4, 2017.
- [18] P. Kadu and A. V. Zadgaonkar, **Knowledge extraction from text document using open information extraction technique**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 2280–2283, 2020, doi: 10.30534/ijatcse/2020/208922020.
- [19] R. Kherde and J. Sanghavi, **Text detection and recognition techniques from images using react native**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 1562–1567, 2020, doi: 10.30534/ijatcse/2020/100922020.
- [20] V. Braun and V. Clarke, **Using thematic analysis in psychology**, *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006, doi: 10.1191/1478088706qp063oa.
- [21] H. T. Sueno, B. D. Gerardo, and R. P. Medina, **Dimensionality reduction for classification of filipino text documents based on improved bayesian vectorization technique**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 8037–8042, 2020, doi: 10.30534/ijatcse/2020/162952020.
- [22] A. El Kah and I. Zeroual, **The effects of Pre-Processing Techniques on Arabic Text Classification**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 10, no. 1, pp. 41–48, 2021, doi: 10.30534/ijatcse/2021/061012021.
- [23] E. Mosconi and M. C. Perreault, **Social Media Engagement: Content Strategy and Metrics Research Opportunities**, in *51st Hawaii International Conference on System Sciences*, 2018, pp. 3568–3577, doi: 10.24251/HICSS.2018.451.
- [24] M. A. Throuvala, M. D. Griffiths, M. Rennoldson, and D. J. Kuss, **A 'control model' of social media engagement in adolescence: A grounded theory analysis**, *Int. J. Environ. Res. Public Health*, vol. 16, no. 23, 2019, doi: 10.3390/ijerph16234696.
- [25] F. D. O. Santini, W. J. Ladeira, D. C. Pinto, M. M. Herter, C. H. Sampaio, and B. Babin, **Customer engagement in social media: a framework and meta-analysis**, *J. Acad. Mark. Sci.*, 2020, doi: 10.1007/s11747-020-00731-5.
- [26] A. Keating and G. Melis, **Social media and youth political engagement: Preaching to the converted or providing a new voice for youth?**, *Br. J. Polit. Int. Relations*, vol. 19, no. 4, 2017.
- [27] S. Yang and M. Kang, **Measuring blog engagement: Testing a four-dimensional scale**, *Public Relat. Rev.*, vol. 35, pp. 323–324, 2009.
- [28] B. G. Smith and T. D. Gallicano, **Terms of engagement: Analyzing public engagement with organizations through social media**, *Comput. Human Behav.*, vol. 53, pp. 82–90, 2015.