Comments Analysis on Social Media: A Review

Brian Alafwan\textsuperscript{1*}, Manahan Siallagan\textsuperscript{1}, Utomo Sarjono Putro\textsuperscript{1}

\textsuperscript{1}Institut Teknologi Bandung, Indonesia

Abstract

As the number of people using and participating in social media grows, academics become interest in studying this new media, specifically comment analysis, in order to comprehend public opinion and user behavior. However, there are no studies that map the development of comment analysis domain, which would be valuable for future research. To address the issue, we examine prior publications using PRISMA approach, and offer suggestions for further research. An investigation was conducted to locate pertinent publications published in databases between 2010 and 2022. On the basis of our examination of 115 relevant articles, we found that, within the scope of methodology, prior researches employ two methods (sentiment and content analysis) and three tools (human, software, and mixed coders), and the majority of them concentrate on gathering data from western countries, covering numerous platforms and topics. Based on these findings, we recommend that future research in comment analysis should synthesize methods and instruments. In addition, examine areas that have not been fully explore in terms of platforms (e.g., Instagram and Tiktok), topic (e.g., local government), and regions (e.g., eastern countries) that would be valuable in order to enhance the body of knowledge in this domain.

Keywords: Online Comment; Social Media Communication; Social Media; Systematic Literature Review.

Received on 18 January 2023, accepted on 20 August 2023, published on 6 September 2023

Copyright © 2023 Alafwan \textit{et al.}, licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetis.3843

\*Corresponding author. Email: brian.alafwan@sbm-itb.ac.id

1. INTRODUCTION

Social media has gained popularity as a source of information (Zaleski et al., 2016), and a place for expressing opinions on a number of topics (Humprecht et al., 2020). Furthermore, the easier and more adaptable the device for accessing social media, the stronger and more powerful the connection will be (Lingam & Aripin, 2017). This connection is exemplified by the fact that every citizen can leave comments on various social media platforms (Ziegele et al., 2014), and liberated to express their opinions and attitudes regarding both public and sensitive issues due to the nature of online social spaces (see, Al-Zaman, 2021; Brochado et al., 2019; De Brün et al., 2014; King & McCashin, 2022; Towner, 2012). This phenomenon drives scholars to investigate the content, behavior, and patterns of citizens' comments on social media using comment analysis method (Chumwatana, 2018). Additionally, the analysis of comments constitutes a potentially interesting data source to mine for obtaining implicit knowledge about users, post, categories, community interests (Siersdorfer et al., 2010) and to comprehend the generative deliberation potential of the emerging technology and its capacity to create a virtual public sphere (Ksiazek, 2015).
Since more than a decade ago, scholars have been interested in analyzing public opinion generated from comment sections on a variety of topics, including food topics (Danner & Menapace, 2020; Pantelidis, 2010), enterprise (Dekay, 2012), celebrity gossip (Eronen, 2014), jihadists in Syria (R. da Silva & Crilley, 2013), fashion (Jung & Kim, 2016), flaming behavior (Lingam & Aripin, 2017), television show (Lacalle & Simelio, 2016), news (Kalogeropoulos et al., 2017), product and service (Chumwatana, 2018), news platform (Ben-David & Soffer, 2018), depression (Gaus et al., 2021; Tao & Jacobs, 2019), immigrant (Andersen, 2019; Utami; Zahra et al., 2023), sexual harassment (Colliver & Coyle, 2020), election (Rossini et al., 2020), policy (E. O. Silva & Flynn, 2020), scientific discovery (Walker & Malson, 2020; Supardam et al., 2023; Utami, 2023), Covid-19 (Al-Motlaq, 2021; Suter et al., 2022), and various other topics. It demonstrates that comment analysis is an expanding field of study.

To the best of our knowledge, no previous research has attempted to comprehensively summarize the existing body of knowledge in the field of comment analysis. As a result, there is a limited understanding of the strengths and limitations of the research that has been conducted in this field, including methodologies, regions, tools, research objects, and platforms used. To address this gap, we propose a systematic literature review (SLR) on comment analysis literatures in order to advance the current state of knowledge (Hamid et al., 2016; Roza et al., 2023; Saefudin et al., 2023). Such a review would be useful for informing future research that aims to further explore and expand the existing body of knowledge on comment analysis in new media, which based on research gaps (Ojong et al., 2021; Tandon et al., 2020; Tremmel et al., 2017; Malta, 2023).

To achieve the objectives of this study, a SLR was conducted using PRISMA approach. Our research material consists of 115 scholarly articles published between 2010 and 2022. Our research findings have academic implications, namely mapping how past studies have addressed comment analysis and providing recommendations for future research. More specifically this study aims to: 1) conduct a comprehensive literature review on comment analysis; 2) Identify the various methodologies, regions, tools, research objects, and platforms explored by prior studies; and 3) provide recommendations for future research.

This paper is organized as follows. This article, as indicated, presents the studies conducted regarding comment analysis. The subsequent section describes in detail our research methodology, and then how previous research on comment analysis was carried out. The final section summarizes the primary contributions of the study, mainly recommendations for future research.

2. MATERIALS AND METHODS

This systematic review was conducted in accordance with the PRISMA principles, which provide a structured approach for synthesizing existing evidence and providing transparent and thorough reports that can assist decision making based on evidence (Page et al., 2021). The Scopus database and Scopus-indexed publishers such as Taylor & Francis, Springer, Science Direct, and Emerald were utilized to obtain publications relevant to our study. In addition, a snowball search step was used to identify further publications that were published between the period 2010 until 2022, this process followed the pattern that was established by Macke & Genari (2019) and Tandon et al. (2020). Furthermore, we included in academic publications studies on comment analysis of social media platforms that were published in English (e.g. books, book chapters, and peer-reviewed articles) and exclude studies that were primarily concerned with developing analytical tools or analyzing the effect of comments on other factors.

We began our search with the predefined keywords ‘social media’, ‘comment’, and ‘analysis’, as well as ‘social media OR comment’ OR ‘analysis’ This preliminary search produced over 1,965 results. The authors limited the number of papers included in the review by restricting the search to the title, abstract, and keywords generated by an algorithm from the citations or bibliographies of the records. After deleting duplicates and articles that did not meet the requirements, our initial search revealed 340 references, which served as the basis for our research. After reviewing the titles and abstracts of these papers, the authors discarded 11 that did not retrieved, leaving 329 to be retrieved.

We eliminated 214 papers as unrelated or accepted in our inclusion area to our examination because they focused insufficiently on the study of comments analysis, for example, by examining the effect of other variables or focusing on the comment distribution pattern. Based on this reason, we reserve 115 for evaluation. By analyzing platforms, and then provide recommendations for future study based on this data and utilizing the remaining 115 references, we construct a mapping of approaches, methods, contexts, and. The procedure depicted in Figure 1.

3. RESULTS

Year of Paper Distribution
Based on 115 eligible papers, the distribution of these papers began in 2010 and is ongoing to this day. According to the data, the most publications (17 papers) occurred in 2020, followed by 14 papers in 2018 and 12 papers in 2021. Although there has been some fluctuation from year to year, figure 2 indicate that the number of scholars exploring comment analysis on social media has increased over time, supporting the trend pointed by Madden et al. (2012).
To evaluate the influence and interest of previous studies, using a reliable bibliometric tool, specifically Google Scholar (Harzing & Alakangas, 2016), we track how many times each paper has been cited. The ten articles that received most citations are listed in Table 1.

As shown in the table below, the most influential paper is *Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments* (Coe et al., 2014), which received 845 citations, followed by paper that discuss similar topic (Santana, 2014) which has 530 citations. However, this paper analyzed the user-comment sections of newspapers that allow anonymity and those that do not; then, compares the two in terms of civility. And the third most cited paper is *Electronic Meal Experience: A Content Analysis of Online Restaurant Comments* (Pantelidis, 2010), which has 497 citations.

<table>
<thead>
<tr>
<th>Titles</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments</td>
<td>845</td>
</tr>
<tr>
<td>Virtuous or Vitriolic: The effect of anonymity on civility in online newspaper reader comment boards</td>
<td>530</td>
</tr>
<tr>
<td>Electronic Meal Experience: A Content Analysis of Online Restaurant Comments</td>
<td>497</td>
</tr>
<tr>
<td>How Useful are Your Comments? Analyzing and Predicting YouTube Comments and Comment Ratings</td>
<td>403</td>
</tr>
<tr>
<td>Civility 2.0: a comparative analysis of incivility in online political discussion</td>
<td>391</td>
</tr>
<tr>
<td>Commenting on YouTube Videos: From Guatemalan Rock to El Big Bang</td>
<td>271</td>
</tr>
<tr>
<td>What Creates Interactivity in Online News Discussions? An Exploratory Analysis of Discussion Factors in User Comments on News Items</td>
<td>225</td>
</tr>
<tr>
<td>How large companies react to negative Facebook comments</td>
<td>225</td>
</tr>
<tr>
<td>Engaging the Social News User Comments on news sites and Facebook</td>
<td>201</td>
</tr>
<tr>
<td>Sentiment classification of consumer generated online reviews using topic modeling</td>
<td>194</td>
</tr>
</tbody>
</table>

**Number of Paper's Citation**

Following Dabi et al. (2020), in order to obtain further insight, we also collect citations from Scopus as another credible source as mentioned by Harzing & Alakangas (2016), in order to identify the most influential authors. The ten authors that received most citations are listed in Table 2. According to the data we present, the most influential author based on citations is Sanjay Sharma (2019), who has received a total of 22,503 citations. Followed by Mike Thelwall (2018) with a total of 20,518 citations, and Stefan Gossling (2017), with a total of 14,023 citations.
### Distribution of Regions

In the context of region distribution, we mapped two types of papers, namely first, papers that focus on comments in general, in the sense that they do not focus on specific regions, for example papers written by Siersdorfer et al. (2010) and Thelwall (2011). The second is the type of papers that limits the object of their research in a particular region, for instance papers written by Rowe (2015a) and Ziegele et al. (2014). To illustrate our collected data, we give Figure 3 shown below:

![Figure 3. Distribution of Regions](image)

According to our data, we discovered publications that limit their research of social media comments to particular areas of the region (several studies do not specify the region examined). There are 42 articles that analyze social media comments in the North America region, including United States, Canada and Mexico. Then there are 22 publications that cover the Western Europe, including Ireland, France, Netherland, Germany and United Kingdom; and 7 papers that analyze social media comments in Southern Europe, including Italy, Turkey and Spain. This indicates that the researchers concentrated in Western region. Other regions, such as in Asia region have not yet been extensively investigated by researchers.

### Typology of Comment Analysis

Research on the analysis of comments in social media, in our results, involves two types of research. The first form of research is analysis that classifies or codes comments from social media users, as seen in the research of R. da Silva & Crilley (2016) and Lingam & Aripin (2017b). And second, is research that analyzes the sentiments of comments from social media users, an example of which can be found in the research of Teh et al. (2018) and Fortunati et al. (2021). Our result are shown in Figure 4 below.
First, Content analysis is a method for qualitatively or quantitatively (Rowe, 2015b) determining the message content by the use of consistent and objective criteria (Aharony, 2012; Hsieh et al., 2019; Madden et al., 2013) typically accompanied by coding techniques (Fu et al., 2016). We discovered 85 studies that analyzed comments using a content analysis methodology. The researchers’ content analysis resulted in a variety of classifications of users comments.

Second, Sentiment refers to positive or negative feelings about something (Io & Lee, 2020; Veletsianos et al., 2018), which are typically shared through a variety of mediums including social media. Sentiment analysis is a technique for assessing the opinions and feelings represented in texts (Calheiros et al., 2017; Chang & Chen, 2017; Fortunati et al., 2021; Tur-Viñes & Castelló-Martínez, 2019; Xu et al., 2016) which then yield polarizations on particular themes including positive, negative and neutral sentiment (Cummins & Nambudiri, 2022; Piccinelli et al., 2021; Tur-Viñes & Castelló-Martínez, 2019). In our research, we discovered 30 studies examining the sentiment of social media comments made by users. Similar to content analysis, sentiment analysis studies encompass a vast array of topics and platforms.

Difference of Analyzing Tools

In addition to the two categories of research regarding comment analysis, there are also a variety of methods used by researchers to analyze user comments. According to Teh et al. (2018), tools are used by researchers to identify meaning as well as patterns in comments. We found that researchers was divided into three type when using tools to analyze comment analysis, namely human-coder, software-coder and mix between them and shown in Appendix 1.

As implicitly explained by Graham & Wright (2015) and Tao & Jacobs (2019), human-coder refers to the manual analysis conducted by the researcher without software analysis. We discovered 80 papers of comments analysis either content analysis or sentiment analysis which used this technique. In contrast to the human-coder, the software-coder refers to the analysis carried out using the comment analysis software tool (Oksanen et al., 2015). Examples of this software are NVivo (Liew & Hassan, 2021), SentiStrengh (Oksanen et al., 2015), MAXQDA (Danner & Menapace, 2020), Vader (Cummins & Nambudiri, 2022) and various other software. We found 23 papers which conducting this technique. Mixed-coder refers to the use of human- and software-coder in conducting analysis. For instance, this can be seen in research conducted by Boon-Long & Wongsurawat (2015). Our result are shown in Figure 5 below:
Social media is a platform that allows users to comment on many issues (Lange, 2014), ranging from technology (Fortunati et al., 2021), actor & actress (Parvaresh & Tayebi, 2018) to breastfeeding (Grant, 2016). The scattered comments then encouraged the researchers to undertake an analysis (Zheng et al., 2021). We found that researchers of social media user comments have researched a wide array of objects (see, Appendix 2). As our data shown, numerous researchers have investigated users comments in relation to News, consider the studies undertaken by Pinto-Coelho et al. (2019) as an example. Followed with comment analysis in covid-19 context receive the second-most research. An example of research on this topic can be seen in Al-Motlaq’s (2021) research. Despite this findings, we shall discus in the following chapter why these two objects are in such popular among scholars for study.

Difference of Context

Difference of Platforms

Typically, while conducting study on social media comments, researchers focus on one or a few social media platforms. This can be seen, for instance, in Ben-David & Soffer, (2019) research which is limited to Facebook. Another example is Bolsover (2016), who conducted research on two platforms for his analysis: Facebook and Weibo. The numerous platforms that have been investigated by researchers are included in Appendix 3. According on the data we have gathered, it appears that Youtube (36 papers), Facebook (28 papers), and news websites (21 papers) are the most frequently researched platforms, as depicted in Figure 6. According to R. da Silva & Crilley (2016), social media platforms such as YouTube, Facebook, Twitter, and others permit users to comment and express their opinion, one possible reason that encourages researchers to investigate these platforms.
4. DISCUSSION

In this research, a literature review of social media comment analysis has been conducted. We showed the distribution of articles by year, most influential papers and authors, region, method, tools, and contexts. On the basis of these data, we conclude that the analysis of social media comments conducted by researchers has a broad reach, encompassing a variety of context and methodologies. In the following section, we will examine and provide recommendations based on the findings reported in the previous section.

Urgency for more research on comments in the future

Our data indicate a decrease in social media comment analysis papers, specifically between 2020 and 2022. According to Global Active User Figure (2022a), the number of social media users expanded from 2.18 million at the beginning of 2012 to 4.95 million at the beginning of 2022, which is related to the activities of users who post comments on social media. The increasing number of social media users from year to year should also lead to an increase in study on social media comments. Because, as Zheng et al. (2021) said, social media comments have supplied a rich data for research. This indicates that research on comment analysis is becoming increasingly crucial. This is, in our opinion, a recommendation for social media academics to perform additional research.

Other objects that need to be explored

Based on the number of citations displayed in Table 1, it is clear that the majority of citations pertain to studies that analyze comments in the context of news, more precisely in the study on incivility and civility in users comments. For example, Santana (2014) and Rowe (2015a) discovered that user anonymity can lead to incivility. Due to the fact that civility is the essential tone and practice of democracy (Coe et al. 2014), research on how this practice can occur becomes urgently necessary in order to improve the pattern of civil comments. We are also on the same line as the researchers, but there is an absence of research examining civility and incivility in the context of government, particularly citizens' interaction with their own local municipalities. With the growing adoption of social media by local governments over the past decade (Mergel, 2013), online medium has become a tool for citizens to directly communicate with the government (Fushimi, 2022) and promote democratic values (Lihong He et al., 2019). Engagement, participation, and public trust in local government’s social media (Jaiyen et al., 2020), as well as the influence on policy-making (Mergel, 2019) and monitoring of government performance (Mergel, 2013a, 2014, 2019; W. N. Wu & Jung, 2016), indicate the significance of this topic and the need for additional research.

Asian region as the highest source of social media users

We find that research concentrate on evaluating social media comments from certain region. The North America is the most researched region, the second is Western Europe and then followed by Southern Europe. One possible explanation for this tendency is that Western regions have the most social media users. Humprecht et al. (2020) and Thelwall (2011) demonstrated that Facebook and Youtube are the most popular social media network in the United States. However, as shown in the Global Active User Figure (2021) report, region outside of the Western also have a large number of social media users. This is demonstrated by their everyday use of social media. The Philippines is ranked first among regions where the average user spends approximately 4 hours and 15 minutes every day on social media. Indonesia spends approximately 3 hours and eight minutes per day, while Malaysia spends 3 hours and 1 minute per day. We argue, based on these facts, that social media participation in Asia is also very high. This information should also serve as a guideline for future research, with the Asian region serving as the study's point of focus.

Differences in approach and new recommendations

As stated in the preceding section, research on social media comments is divided into two categories: content analysis and sentiment analysis. Content analysis is a research method that involves making the content of messages manifest by identifying as objectively as feasible their properties. While sentiment analysis focuses on automatically assessing whether a text has an opinion, recognizing whether the polarity or sentiment represented is favorable, negative, or neutral, and extracting an author's assessment of particular features of a topic.

Content analysis is utilized by researchers because it can identify latent meanings in comments objectively and systematically. It also utilized since it may collect data without a theoretical basis, can contribute to generalization due to the size of the dataset, and can provide as evidence for specific hypotheses (Bolsover, 2016; Yi et al., 2016). Meanwhile, sentiment analysis is applied when the researchers want to examine the attitudes and
opinions expressed by users in the comments (Jelodar et al., 2021; Oksanen et al., 2015). Positive, negative, or neutral sentiments are usually used to mapped during sentiment analysis (Lan He et al., 2020).

Based on these findings, in order to comprehend the pattern and substance of user comments (Walker & Malson, 2020) and (Hsieh et al., 2018), we propose combining content analysis with sentiment analysis. Since content analysis alone cannot reveal the sentiment of comments, and sentiment analysis cannot categorize comments. Consequently, combining them might complement each approach’s deficiencies.

Advantages and Disadvantages of Tools

There are three tools that researchers use to analyze social media comments. As previously stated, the three tools are the human coder, the software coder, and the mixed-coder. Researchers utilize software-coder because it eliminates bias (Siersdorfer et al., 2010) and simplifies analysis (Koltsova & Nagornyy, 2019). For instance, Thelwall & Mas-Bleda (2018) and Veletsianos et al., (2018) utilize SentiStrength because it provides them with high accuracy. In contrast, researchers prefer human-coders because they can identify subtle context and hidden meaning in comments, whereas software-coders cannot (Teh et al., 2018; Walker & Malson, 2020).

Other researchers utilize these two tools concurrently. The simultaneous use of both of these technologies to analyze social media comments is referred to as mixed-coding. These researchers utilize these two tools because they are more efficient, provide a high level of reliability, and can compensate for each tool’s limitations (Koltsova & Nagornyy, 2019; Oksanen et al., 2015; Su et al., 2018). We concur with the researchers’ use of a mixed-coder for the analysis, however we have some suggestions. Our argument is that, despite the fact that both are utilized in combination, the human coder should be viewed as the final judge of the analysis, because software-coders are incapable of analyzing the complex context of user comments. We follow the research of Ernst et al. (2017) and Thelwall (2018), which states that automated sentiment tools are complementary. This means that the human coder retains the primary focus of the study, with software providing assistance.

TikTok and Instagram as platforms that need to be research

In the context of the platform, we discovered that YouTube was the most examined platform, followed by Facebook and the News Website. Because these platforms enable user’s engagement and interaction (R. da Silva & Crilley, 2013), and are widely used, popular, and influential on their users’ behavior, they have been investigated intensively by researchers Al-Zaman (2021) and Liebig et al. (2017) Arancibia & Montecino, (2017). In besides these three platforms, we discovered that Instagram (1440 million users) and TikTok (1023 million users) also have a very significant user base (Figure, 2022b). Therefore, we argue that future researchers must also explore the two platforms.

5. CONCLUSIONS

The aim of this study is to identify existing studies on comment analysis and to suggest future research based on the existing research gaps. There are seven findings from our study, including, first, the number of publications on comment analysis is decreasing in 2022, according to the distribution of research years, we discover in the first place. We demonstrate the need for further publications on this subject since, as social media usage rises annually, more academics are able to access the data for analysis, making research potential becoming increasingly essential. Second, our results reveal that research on civility and incivility in the news realm is highly cited, reflecting the interest in this topic among academics. We urge further exploration of civility and incivility in the context of citizens interacting with their local government, an issue that has been largely forgotten but seems to be important given the increased use of social media by local governments. There is a considerable quantity of data available from the comments posted by residents on local government social media, which gives an opportunity for scholars to investigate this topic.

Third, researchers concentrated more on Western regions as a result of our findings. In addition, as we previously demonstrated, there are a lot of social media users in Asian regions as well, thus we draw the conclusion that future studies should include study on Eastern region.

Fourth, we demonstrate that sentiment analysis and content analysis are the two main approaches used in commentary analysis research. Additionally, we discussed the advantages and disadvantages of these two techniques in terms of comprehending user patterns and behavior while leaving comments on social media. Based on these conclusions, we subsequently offer suggestions for combining methods for researching user comments on social media, with the goal of creating a more comprehensive and in-depth approach for examining user comments on social media.
Fifth, we demonstrate the three tools that researchers employ to examine comments including human-coder, software-coder, and mixed-coder. We also examine the advantages and limitations of each tool used for analysis in previous section. Following the researchers who combined the two techniques, we also recommend that future researchers should mix the two tools. We do point out that the two tools combined must place a strong emphasis on the human coder as the primary assessor. Because human-coders is better able to discern subtle messages than software-coders. However, the assistance of a software-coder is also necessary, since it can make the analysis process run more efficiently.

Finally, prior studies has primarily focused analyzing comments on specific social media platforms, including YouTube, Facebook, and news websites. Therefore, we suggested that future research include examination of comments on Instagram and TikTok, as these platforms have a vast user base and deserve investigation. This will contribute to a more thorough comprehension of comments on social media platforms.

6. REFERENCES


[55] Figure, G. A. U. (2022a). Internet Users Over Time. Dare Portal.

[56] Figure, G. A. U. (2022b). The World’s Most Used Social Media Platforms. Dare Portal.

[57] Fortunati, L., Sorrentino, A., Fiorini, L., &


https://doi.org/10.1002/asi
APPENDICES

Appendix 1:

<table>
<thead>
<tr>
<th>Tool</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human-Coder</td>
<td>(Abdalla et al., 2021; Aharony, 2012; Al-Motlaq, 2021; Al-Rawi, 2014; Al-Zaman, 2021; Andersen, 2019; Arancibia &amp; Montecino, 2017; Artime, 2016; Bolsover, 2016; Burton &amp; Schlieman, 2021; Canter, 2013; Canty et al., 2019; Carpenter &amp; Stautd Willet, 2021; Chang &amp; Chen, 2017; Coche &amp; Le Blond, 2018; Coe et al., 2014; Cohen et al., 2017; Colliver &amp; Coyle, 2020; Craft et al., 2015; D’Errico &amp; Paciello, 2018; R. da Silva &amp; Crilley, 2013; Dekay, 2012; Diakopoulos &amp; Naaman, 2011; Dördeivi, 2020; Döring &amp; Mohseni, 2019; Dubovi &amp; Tabak, 2020; Eronen, 2014; Ferrucci &amp; David Wolfgang, 2021; Foruntani et al., 2021; Gaus et al., 2021; Graham &amp; Wright, 2015; Grant, 2016; J. He, 2021; Lan He et al., 2020; Hsieh et al., 2019; Humprech et al., 2020; Jelodar et al., 2021; Juhary, 2012; King &amp; McCashin, 2022; Kopacz &amp; Lawton, 2013; Ksiazek, 2015; Lacalle &amp; Simelio, 2016; Lange, 2014; Lie &amp; Sandel, 2020; Lingam &amp; Aripin, 2017; Madden et al., 2013; McCluskey &amp; Hmielowski, 2012; Miller, 2015, 2019; Moreau et al., 2020; Murthy &amp; Sharma, 2019; Neumayer, 2012; Nguyen &amp; Romaniuk, 2014; Pantelidis, 2010; Parvaresh &amp; Tayebi, 2018; Pavalanathan &amp; Choudhury, 2015; Piccinnelli et al., 2021; Pinto-Coelho et al., 2019; Rabab’ah &amp; Alali, 2020; Rossini, 2022; Rossini et al., 2020; Rowe, 2015b, 2015a; Santana, 2014; Seely, 2017; M. T. da Silva, 2013; E. O. Silva &amp; Flynn, 2020; Strandberg &amp; Berg, 2013; Su et al., 2018; Suter et al., 2022; Tao &amp; Jacobs, 2019; Tsou et al., 2014; Walker &amp; Malson, 2020; Whiting et al., 2019; Yi et al., 2016; Ziegele et al., 2014).</td>
</tr>
<tr>
<td>Software-Coder</td>
<td>(Amarasekara &amp; Grant, 2019; Brochado et al., 2019; Calheiros et al., 2017; Chumwatana, 2018; Cummins &amp; Nambudiri, 2022; Danner &amp; Menapace, 2020; Deschamps, 2014; Foruntani et al., 2021; Hille &amp; Bakker, 2014; Io &amp; Lee, 2020; Koltsova &amp; Nagorny, 2019; Loures et al., 2018; Mishra &amp; Sharma, 2019; Oksanen et al., 2015; Paasch-Colberg &amp; Strippel, 2022; Rocca et al., 2020; Siersdorfer et al., 2010; Szabó et al., 2021; Thelwall &amp; Mas-Bleda, 2018; Tur-Viñes &amp; Casteló-Martínez, 2019; J. Wu &amp; Hong, 2022; Xu et al., 2016; Zaleski et al., 2016).</td>
</tr>
<tr>
<td>Mixed-Coder</td>
<td>(Bloch, 2016; Boon-Long &amp; Wongsurawat, 2015; Chu et al., 2020; De Brûn et al., 2014; Edgerly et al., 2013; Ernst et al., 2017; Fu et al., 2016; McCambridge, 2022; Teh et al., 2018; Thelwall, 2018, 2011; Zheng et al., 2021).</td>
</tr>
</tbody>
</table>

Appendix 2:

<table>
<thead>
<tr>
<th>Context</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>GoPro 360</td>
<td>(Burton &amp; Schlieman, 2021)</td>
</tr>
<tr>
<td>Photography</td>
<td>(Barry, 2014)</td>
</tr>
<tr>
<td>K-Artist</td>
<td>(Xu et al., 2016)</td>
</tr>
<tr>
<td>Discrimination</td>
<td>(Abdalla et al., 2021)</td>
</tr>
<tr>
<td>Cancer</td>
<td>(Fortunati et al., 2021)</td>
</tr>
<tr>
<td>Topic</td>
<td>References</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hotel Review</td>
<td>(Calheiros et al., 2017; Chang &amp; Chen, 2017; Yi et al., 2016)</td>
</tr>
<tr>
<td>Oscar Movie Nomination</td>
<td>(Jelodar et al., 2021)</td>
</tr>
<tr>
<td>Roboid</td>
<td>(Fortunati et al., 2021)</td>
</tr>
<tr>
<td>Television Series</td>
<td>(Loures et al., 2018)</td>
</tr>
<tr>
<td>Selective Dorsal Rhizotomy (SDR)</td>
<td>(Canty et al., 2019)</td>
</tr>
<tr>
<td>Various Context in Youtube</td>
<td>(Madden et al., 2013; Siersdorfer et al., 2010; Thelwall, 2011)</td>
</tr>
<tr>
<td>Pro-Anorexia</td>
<td>(Oksanen et al., 2015)</td>
</tr>
<tr>
<td>TedTalks</td>
<td>(Tsou et al., 2014)</td>
</tr>
<tr>
<td>Hate speech on Islam</td>
<td>(Dordević, 2020; Ernst et al., 2017)</td>
</tr>
<tr>
<td>TedX</td>
<td>(Veletsianos et al., 2018)</td>
</tr>
<tr>
<td>STEM</td>
<td>(Amarasekara &amp; Grant, 2019)</td>
</tr>
<tr>
<td>Sexism</td>
<td>(Döring &amp; Mohseni, 2019)</td>
</tr>
<tr>
<td>Sarcasm</td>
<td>(Teh et al., 2018)</td>
</tr>
<tr>
<td>Wine</td>
<td>(Hsieh et al., 2019)</td>
</tr>
<tr>
<td>Sentiment Towards Males and Females</td>
<td>(Thelwall &amp; Mas-Bleda, 2018)</td>
</tr>
<tr>
<td>Gender Ratio</td>
<td>(Thelwall, 2018)</td>
</tr>
<tr>
<td>Local Government Social Media</td>
<td>(Rocca et al., 2020)</td>
</tr>
<tr>
<td>Rant Comment</td>
<td>(Lange, 2014)</td>
</tr>
<tr>
<td>Actress</td>
<td>(Parvaresh &amp; Tayebi, 2018)</td>
</tr>
<tr>
<td>Travel</td>
<td>(Cohen et al., 2017)</td>
</tr>
<tr>
<td>Obesity</td>
<td>(De Brún et al., 2014)</td>
</tr>
<tr>
<td>Online Hacking Group</td>
<td>(Al-Rawi, 2014)</td>
</tr>
<tr>
<td>Breastfeeding</td>
<td>(Grant, 2016)</td>
</tr>
<tr>
<td>Rape Culture</td>
<td>(Zaleski et al., 2016)</td>
</tr>
<tr>
<td>Word of Mouth</td>
<td>(Nguyen &amp; Romaniuk, 2014)</td>
</tr>
<tr>
<td>Male and Female Athlete</td>
<td>(Coche &amp; Le Blond, 2018)</td>
</tr>
<tr>
<td>Covid-19</td>
<td>(Al-Motlaq, 2021; Al-Zaman, 2021; Suter et al., 2022; Zheng et al., 2021)</td>
</tr>
<tr>
<td>Educational Professional</td>
<td>(Carpenter &amp; Staudt Willet, 2021)</td>
</tr>
<tr>
<td>Air Travel</td>
<td>(Piccinelli et al., 2021)</td>
</tr>
<tr>
<td>Personality Disorder</td>
<td>(King &amp; McCashin, 2022)</td>
</tr>
<tr>
<td>Skincare</td>
<td>(Cummins &amp; Nambudiri, 2022)</td>
</tr>
<tr>
<td>Airline Service Quality</td>
<td>(Brochado et al., 2019)</td>
</tr>
<tr>
<td>News</td>
<td>(Aharony, 2012; Bolsover, 2016; Canter, 2013; Coe et al., 2014; J. He, 2021; Hille &amp; Bakker, 2014; Humprech et al., 2020; McCluskey &amp; Hmielowski, 2012; Pinto-Coelho et al., 2019; Rabab’ah &amp; Alali, 2020; Rowe, 2015b; Santana, 2014; Seely, 2017; Ziegele et al., 2014)</td>
</tr>
<tr>
<td>Teaching</td>
<td>(Juhary, 2012)</td>
</tr>
<tr>
<td>Celebrity</td>
<td>(Eronen, 2014; Whiting et al., 2019)</td>
</tr>
<tr>
<td>Policy</td>
<td>(Deschamps, 2014; Ksiazek, 2015; E. O. Silva &amp; Flynn, 2020)</td>
</tr>
<tr>
<td>Presidential</td>
<td>(Rossini et al., 2020; M. T. da Silva, 2013)</td>
</tr>
<tr>
<td>Topic</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Politic</td>
<td>(Artine, 2016; Rossini, 2022; Rowe, 2015a)</td>
</tr>
<tr>
<td>Music Group</td>
<td>(Murthy &amp; Sharma, 2019)</td>
</tr>
<tr>
<td>Journalism</td>
<td>(Craft et al., 2015; Graham &amp; Wright, 2015)</td>
</tr>
<tr>
<td>Mental Health</td>
<td>(Pavalanathan &amp; Choudhury, 2015; J. Wu &amp; Hong, 2022)</td>
</tr>
<tr>
<td>Anti-Immigrant</td>
<td>(Bloch, 2016)</td>
</tr>
<tr>
<td>Discourse Pattern</td>
<td>(Fu et al., 2016)</td>
</tr>
<tr>
<td>Immigrant</td>
<td>(Andersen, 2019; D’Errico &amp; Paciello, 2018)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>(Pantelidis, 2010)</td>
</tr>
<tr>
<td>Companies</td>
<td>(Dekay, 2012)</td>
</tr>
<tr>
<td>Anti-Fasist Protest</td>
<td>(Neumayer, 2012)</td>
</tr>
<tr>
<td>Proposition Campaign</td>
<td>(Edgerly et al., 2013)</td>
</tr>
<tr>
<td>Native America</td>
<td>(Kopacz &amp; Lawton, 2013)</td>
</tr>
<tr>
<td>Mass Shooting</td>
<td>(Miller, 2015)</td>
</tr>
<tr>
<td>Handphone</td>
<td>(Boon-Long &amp; Wongsurawat, 2015)</td>
</tr>
<tr>
<td>British Jihadist</td>
<td>(R. da Silva &amp; Crilley, 2013)</td>
</tr>
<tr>
<td>Online Flaming</td>
<td>(Lingam &amp; Aripin, 2017)</td>
</tr>
<tr>
<td>Movie</td>
<td>(Lacalle &amp; Simelio, 2016)</td>
</tr>
<tr>
<td>Online Anger</td>
<td>(Arancibia &amp; Montecino, 2017)</td>
</tr>
<tr>
<td>Consumer</td>
<td>(Mishra &amp; Sharma, 2019)</td>
</tr>
<tr>
<td>Service Brand</td>
<td>(Chumwatana, 2018)</td>
</tr>
<tr>
<td>Online Incivility</td>
<td>(Su et al., 2018; Szabó et al., 2021)</td>
</tr>
<tr>
<td>Depression</td>
<td>(Gaus et al., 2021; Tao &amp; Jacobs, 2019)</td>
</tr>
<tr>
<td>Holocaust</td>
<td>(Miller, 2019)</td>
</tr>
<tr>
<td>YouTubers</td>
<td>(Tur-Viñes &amp; Castelló-Martínez, 2019)</td>
</tr>
<tr>
<td>Media</td>
<td>(Koltsova &amp; Nagornyy, 2019)</td>
</tr>
<tr>
<td>Sexual Violence</td>
<td>(Colliver &amp; Coyle, 2020)</td>
</tr>
<tr>
<td>Football</td>
<td>(Moreau et al., 2020)</td>
</tr>
<tr>
<td>Science</td>
<td>(Duboví &amp; Tabak, 2020; Walker &amp; Malson, 2020)</td>
</tr>
<tr>
<td>Wester Opinion on China</td>
<td>(Lan He et al., 2020)</td>
</tr>
<tr>
<td>Violence</td>
<td>(Chu et al., 2020; McCambridge, 2022)</td>
</tr>
<tr>
<td>Hotel Robot</td>
<td>(Io &amp; Lee, 2020)</td>
</tr>
<tr>
<td>Ethnic Chinese</td>
<td>(Lie &amp; Sandel, 2020)</td>
</tr>
<tr>
<td>Organic Food</td>
<td>(Danner &amp; Menapace, 2020)</td>
</tr>
<tr>
<td>Journalism</td>
<td>(Ferrucci &amp; David Wolfgang, 2021)</td>
</tr>
<tr>
<td>Hate Comment</td>
<td>(Paasch-Colberg &amp; Strippel, 2022)</td>
</tr>
</tbody>
</table>

### Appendix 3:

<table>
<thead>
<tr>
<th>Platforms</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>YouTube</td>
<td>(Al-Rawi, 2014; Amarasekara &amp; Grant, 2019; Arancibia &amp; Montecino, 2017; Boon-Long &amp; Wongsurawat, 2015; Burton &amp; Schlieman, 2021; Canty et al., 2019; Colliver &amp; Coyle, 2020;</td>
</tr>
<tr>
<td>Platform</td>
<td>References</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Flickr</td>
<td>Barry, 2014; McCambridge, 2022; Suter et al., 2022; Zheng et al., 2021</td>
</tr>
<tr>
<td>Twitter</td>
<td>Abdalla et al., 2021; Canty et al., 2019; Chumwatanana, 2018; R. da Silva &amp; Crilley, 2013; Lacalle &amp; Simelio, 2016; Moreau et al., 2020; Nguyen &amp; Romaniuk, 2014</td>
</tr>
<tr>
<td>Facebook</td>
<td>Abdalla et al., 2021; Al-Motlaq, 2021; Al-Zaman, 2021; Andersen, 2019; Bolsover, 2016; Canty et al., 2019; Chumwatanana, 2018; Cohen et al., 2017; D’Errico &amp; Paciello, 2018; R. da Silva &amp; Crilley, 2013; Dekay, 2012; Hille &amp; Bakker, 2014; Humprecht et al., 2020; Lacalle &amp; Simelio, 2016; Lie &amp; Sandel, 2020; Mishra &amp; Sharma, 2019; Moreau et al., 2020; Nguyen &amp; Romaniuk, 2014; Parvaresh &amp; Tayebi, 2018; Rocca et al., 2020; Rossini, 2022; Rossini et al., 2020; Su et al., 2018; Szabó et al., 2021; Tao &amp; Jacobs, 2019; Teh et al., 2018; Walker &amp; Malson, 2020; Ziegele et al., 2014</td>
</tr>
<tr>
<td>News Website</td>
<td>Canter, 2013; Coe et al., 2014; Craft et al., 2015; Danner &amp; Menapace, 2020; Đorđević, 2020; Eronen, 2014; Ferrucci &amp; David Wolfgang, 2021; Graham &amp; Wright, 2015; Koltsova &amp; Nagornyy, 2019; Ksiazek, 2015; McCluskey &amp; Hmielowski, 2012; Paasch-Colberg &amp; Strippel, 2022; Pinto-Coelho et al., 2019; Rabab’ah &amp; Alali, 2020; Rowe, 2015a, 2015b; Santana, 2014; Seely, 2017; M. T. da Silva, 2013; Strandberg &amp; Berg, 2013; Zaleski et al., 2016</td>
</tr>
<tr>
<td>TripAdvisor</td>
<td>Brochado et al., 2019; Calheiros et al., 2017; Chang &amp; Chen, 2017; Hsieh et al., 2019; Yi et al., 2016</td>
</tr>
<tr>
<td>MyAnimelist</td>
<td>Loures et al., 2018</td>
</tr>
<tr>
<td>Ted Website</td>
<td>Tsou et al., 2014</td>
</tr>
<tr>
<td>Reddit</td>
<td>Carpenter &amp; Staudt Willet, 2021; Cummins &amp; Nambudiri, 2022; Lan He et al., 2020; Pavalanathan &amp; Choudhury, 2015; Teh et al., 2018</td>
</tr>
<tr>
<td>Website Magazine</td>
<td>Whiting et al., 2019</td>
</tr>
<tr>
<td>Internet Forum</td>
<td>Danner &amp; Menapace, 2020; De Brün et al., 2014</td>
</tr>
<tr>
<td>Dailymail.uk</td>
<td>Grant, 2016</td>
</tr>
<tr>
<td>MySpace</td>
<td>Nguyen &amp; Romaniuk, 2014</td>
</tr>
<tr>
<td>Yahoo!</td>
<td>Coche &amp; Le Blond, 2018</td>
</tr>
<tr>
<td>Italian National Consumer Union website</td>
<td>Piccinelli et al., 2021</td>
</tr>
<tr>
<td>Zomato</td>
<td>Calheiros et al., 2017</td>
</tr>
<tr>
<td>Sacbee.com</td>
<td>Diakopoulos &amp; Naaman, 2011</td>
</tr>
<tr>
<td>Wikileaks</td>
<td>Aharony, 2012</td>
</tr>
<tr>
<td>Various Digital Platforms</td>
<td>Eronen, 2014</td>
</tr>
<tr>
<td>Imigrant Website</td>
<td>Bloch, 2016</td>
</tr>
<tr>
<td>Discussion Space</td>
<td>(Fu et al., 2016)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>eating.co.uk</td>
<td>(Pantelidis, 2010)</td>
</tr>
<tr>
<td>from Pew Research Centre</td>
<td>(Artme, 2016)</td>
</tr>
<tr>
<td>Weibo</td>
<td>(J. He, 2021; Io &amp; Lee, 2020)</td>
</tr>
<tr>
<td>Weaboo</td>
<td>(Bolsover, 2016)</td>
</tr>
<tr>
<td>Baidu Tieba</td>
<td>(Chu et al., 2020)</td>
</tr>
<tr>
<td>Instagram</td>
<td>(J. Wu &amp; Hong, 2022)</td>
</tr>
</tbody>
</table>