SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2024
ORGANIZATION AND MANAGEMENT SERIES NO. 196

EXPLORING PUBLIC OPINION: A SENTIMENT ANALYSIS OF YOUTUBE COMMENTS ON HEAT PUMP VIDEOS

Marcin WYSKWARSKI

Institute of Economy and Informatics, Faculty of Organization and Management, Silesian University of Technology; marcin.wyskwarski@polsl.pl, ORCID: 0000-0003-2004-330X

Purpose: This study aims to ascertain public sentiments, thoughts, and opinions about "heat pumps" through sentiment analysis of comments posted on YouTube videos.
Design/methodology/approach: Video comments were automatically downloaded and pre-processed. This preprocessing involved removing all characters except for letters, as well as eliminating URLs, hashtags, emojis, and search-related words from the text. The sentiment value of each comment was then assessed. Visual representations were employed to depict the distribution of positive, negative, and neutral sentiments across the comments. Additionally, a word cloud was used to highlight the most frequently occurring words within these comments.
Findings: For the comments on videos about heat pumps, the proportions of positive, negative, and neutral sentiments were calculated. Additionally, data concerning the number of videos published, the total views received by these videos, their duration, the number of comments made, and the length of these comments were collected and analysed.
Research limitations/implications: The study focused exclusively on comments from videos containing the words "heat pump" in their titles, assuming relevance to the subject of heat pumps. Only comments in Polish were selected for analysis. The sentiment analysis was conducted autonomously using the "ccl emo" service, without manual oversight. This analysis was limited to assessing the sentiments of YouTube users, based on the assumption that the comments reflect discussions pertinent to the video content.
Practical implications: Automated analysis of public opinions on photovoltaics.
Originality/value: Opinions about heat pumps were gathered and analysed. The increasing number of videos and comments indicates a steady rise in interest in heat pumps within Poland.
Keywords: sentiment analysis, YouTube, heat pump, text mining.
Category of the paper: research paper, case study.

1. Introduction

As global warming becomes increasingly evident in the environment, society is confronted with numerous climate-related challenges aimed at substantially reducing greenhouse gas emissions from the heating and cooling of buildings (Decuypere et al., 2022). The heating and cooling of buildings account for one-tenth of anthropogenic greenhouse gas emissions. (Edenhofer, 2015). These emissions are expected to rise sharply in the coming decades (Ürge-Vorsatz et al., 2015). Heating systems require a swift transition to low-carbon alternatives to meet global climate targets (Martiskainen et al., 2021).

Thanks to technological advancements, heat pumps now have the potential to halve or more the emissions from heating and cooling in various settings (Billimoria et al., 2021). Heat pumps are versatile, utilizing renewable energy sourced from the air, water, ground, or exhaust from buildings to provide heating and cooling (Nowak, 2018). Two prevalent types of heat pumps are air-source, which transfers heat to and from the outdoor air, and ground-source, which transfers heat to and from the ground (Kircher, Zhang, 2021). Heat pumps can serve multiple functions, including heating and cooling buildings, generating electricity, and supplying hot water (Soltani et al., 2019). In recent years, heat pump technology has advanced significantly, improving both in efficiency and in heating performance at low temperatures (Chua et al., 2010). Despite their benefits, heat pumps encounter several barriers to widespread adoption, including the fact that their lifetime costs are not always competitive with existing technologies, such as natural gas furnaces (Billimoria et al., 2021). Even when the operating costs of heat pumps are competitive, the initial purchase and installation expenses can be prohibitively high (Bergman, 2013). Additional barriers to the adoption of heat pumps include locating installers who are knowledgeable about modern heat pump technology, selecting the appropriate models and sizes of heat pumps, and navigating the process of finding and applying for rebates, tax credits, and other incentives (Snape et al., 2015).

In the digital era, it has become common for people to express their opinions on social media. These expressions of thoughts, feelings, and judgments can be analyzed using a technique known as sentiment analysis. Sentiment analysis offers an automated approach to examining sentiment, emotion, and opinion expressed in written language (Xu et al., 2022). It involves the analysis, processing, summarizing, and interpretation of subjective texts with emotional nuances, such as valuable feedback on individuals, events, products, etc., shared by users online (Deng et al., 2022). It can be conducted to evaluate an individual's viewpoint or inclination towards a subject or issue, determining whether it leans towards a positive or negative perspective (Pang et al., 2002).

A significant data source containing people's opinions can be the comments posted on YouTube videos. YouTube is an online video platform that is rapidly expanding and receives nearly two billion views daily (Aydın, Yılmaz, 2021; Snelson, 2011). According to data as of May 09, 2024, more than 5 billion YouTube videos are viewed each day, there are 2.5 billion monthly active YouTube users, and more than 500 hours of YouTube videos are uploaded per minute (Omnicore, n.d.). As the world's largest video platform, YouTube presents a diverse array of media content created by both companies and individuals. This content includes music videos, promotional videos for products, vlogs, review videos, and educational content. (Muhammad et al., 2019).

A variety of tools can be utilised for analysing data retrieved from the internet. Due to the substantial volume of data, techniques such as text mining, data mining, machine learning, topic modelling, sentiment analysis, and similar approaches are commonly employed. The examination of data gathered from social media represents a burgeoning field. Its popularity is increasing due to its cost-effectiveness, easy accessibility, and the element of anonymity (Das et al., 2015, 2019; Evans-Cowley, Griffin, 2012). The literature contains numerous studies on sentiment analysis conducted on data extracted from the Internet (Ağrali, Aydin, 2021; Pang, Lee, 2004, 2008; Read, 2005). The utilization of sentiment analysis of text to ascertain public opinions on renewables was presented in (Corbett, Savarimuthu, 2022; Ibar-Alonso et al., 2022; Jain, Jain, 2019a, 2019b; Kim et al., 2021; Loureiro, Alló, 2020; Zarrabeitia-Bilbao et al., 2022).

2. Research Methodology

On May 5, 2024, a total of 4025 videos related to heat pumps were found on YouTube. Python 'scrapetube' library was utilized for this purpose (Twersky, n.d.). This library enables video searches on YouTube without the need for the official YouTube API. For conducting video searches, the author employed 57 phrases associated with heat pumps. These phrases were various combinations of the following three 2-grams: “pompa ciepła” (heat pump), “pompa gruntowa” (ground source heat pump), “pompa powietrzna” (air source heat pump). Examples of phrases included, among others, the following:

- “pompa ciepła”, “pompie ciepła”, “pomp ciepła” etc.
- “gruntowa pompa”, “pompa gruntowa”, “pompy gruntowej” etc.
- “pompa powietrzna”, “powietrzna pompa”, “powietrznej pompie” etc.

In the next step, comments posted by users under each video were downloaded using the "youtube-comment-downloader" package. This package provides a straightforward script for downloading YouTube comments without the need to utilise the YouTube API (*Youtube-Comment-Downloader*, n.d.).

The next phase involved the author removing:

- comments written in languages other than Polish,
- comments that were identical to others posted by the same user (often advertising services, products, or job opportunities from a company); the content was treated as a character string and compared using the "==" operator.

Subsequently, the content of the comments was pre-processed. This included the removal of URLs, hashtags, emojis, user names, search terms, and any characters that were not letters. The word count of each cleaned comment was then verified, and comments containing fewer than four words were excluded. After these processes, there were 185874 comments remaining.

These comments were associated with one of 2857 videos. The other 1168 videos either had no comments or their comments were eliminated during the pre-processing phase.

In the following step, the ccl_emo (https://wiki.clarin-pl.eu/pl/nlpws/services/ccl_emo; https://clarin-pl.eu/index.php/wydzwiek) service, developed by CLARIN-PL[1], was utilized. Known as "Wydźwięk" in Polish and "Sentiment" in English, this service is designed for the statistical analysis of texts' tones and emotions. (Grubljesic et al., 2019; Janz et al., n.d.).

Additionally, other services provided by CLARIN-PL were employed. These included:

- Any2txt - a service that converts text files (e.g. doc, docx, xlsx) into plain text.
- Speller2 - a service that checks the spelling of the text, utilizing a tool known as Autocorrect (https://languagetool.org/pl) for this purpose.
- Wcrft2 - is a basic morpho-syntactic tagger for the Polish language.
- WSD - a service designed for word sense disambiguation, specifically tailored for Polish texts. It employs plWordNet as a source of potential meanings, which organizes lexical units into synsets connected through lexico-semantic relationships. Each lexical unit encapsulates a lexical meaning and is defined by three elements: a lemma, a part of speech, and a sense identifier (Janz et al., n.d.).

The selected lexical units stored in plWordNet were enhanced with emotive annotations. Lexical units were characterized by (Janz et al., n.d.):

- sentiment polarity: it was evaluated using a 5-point scale, ranging from strong and weak through negative and positive, to neutral;
- basic emotions: gladness, trust, enjoyment of something expected, sadness, anger, fear, disgust, and surprise at something unpredictable. These emotions correspond to the eight basic emotions identified by Plutchik in his Wheel of Emotions (Plutchik, 1980; Wierzbicka, 1992a, 1992b);
- fundamental human values: utility, the good of another, truth, knowledge, beauty, happiness, uselessness, harm, ignorance, error, ugliness, unhappiness. These are the basic human values as indicated by (Puzynina, 1992).

In the next step, the sentiment of each comment was determined based on the polarity of the words it contained. A comment is considered to have a negative sentiment if it contains more negative words than positive ones. On the other hand, a comment is classified as having a positive sentiment when the number of negative words is fewer than the positive ones. A neutral sentiment is noted when the ratio of positive to negative words is equal. Additionally, each comment was analysed to identify the number of words annotated with basic emotions and fundamental human values.

[1] CLARIN-PL is a Polish scientific consortium, part of the European Research Infrastructure CLARIN (Common Language Resources and Technology Infrastructure) (*CLARIN-PL*, n.d.)

3. Results

A ribbon chart shown in Figure 1 illustrates the total number of videos retrieved, categorized into those with comments and those without. Additional detailed information can be found in Table 1. Importantly, there was a notable rise in video publication starting in 2019, with 129 videos published, and reaching a peak in 2023 with 1867 videos. Although the production of videos in 2024 has declined, it should be noted that the data was collected up until May 09, 2024. Therefore, the total count of videos for 2024 might increase as more videos could be published in the latter months of the year. As the chart shows, in the analysed dataset, the number of videos with comments was lower than those without comments only in the year 2012.

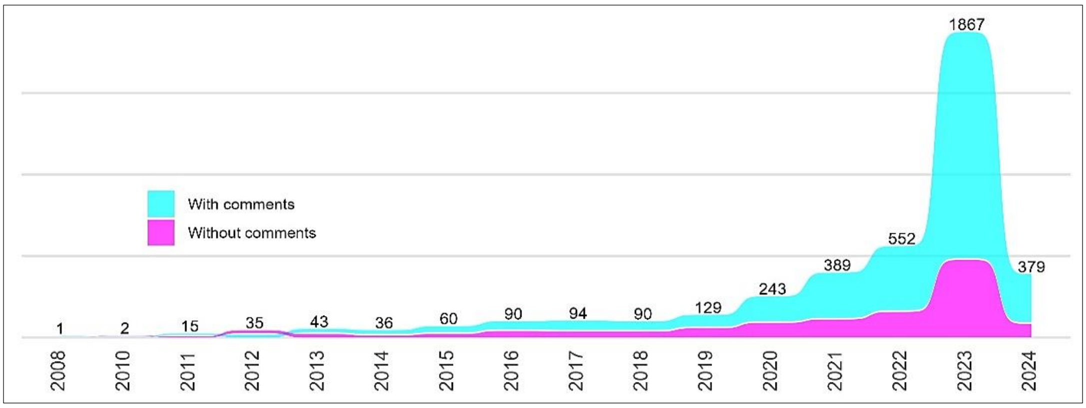

Figure 1. Distribution of the retrieved videos over the years. Note: videos for 2024 only gathered until May 9, 2024.

Sources: original research.

Table 1.
Distribution of the retrieved videos over the years. Note: videos for 2024 only gathered until May 9, 2024

Year	With comments	Without comments	Total
2008	1 (100%)	(0%)	1
2010	1 (50%)	1 (50%)	2
2011	9 (60%)	6 (40%)	15
2012	17 (48,57%)	18 (51,43%)	35
2013	22 (51,16%)	21 (48,84%)	43
2014	24 (66,67%)	12 (33,33%)	36
2015	38 (63,33%)	22 (36,67%)	60
2016	50 (55,56%)	40 (44,44%)	90
2017	58 (61,7%)	36 (38,3%)	94
2018	54 (60%)	36 (40%)	90
2019	72 (55,81%)	57 (44,19%)	129
2020	153 (62,96%)	90 (37,04%)	243
2021	278 (71,47%)	111 (28,53%)	389
2022	395 (71,56%)	157 (28,44%)	552
2023	1390 (74,45%)	477 (25,55%)	1867
2024	295 (77,84%)	84 (22,16%)	379
Total	2857 (70,98%)	1168 (29,02%)	**4025**

Sources: original research.

Figure 2 illustrates the number of films published by individual users. For example, in 2011, videos without comments were published by fifth users. One of them published two videos, while the remaining four each published one film. The box plots shown on the diagrams allow us to observe the distribution of the number of published films. For instance, the third quartile for films with comments in 2023 is 3, indicating that 75% of users published three or fewer films (in this case, meaning two or one film). To enhance the clarity of the chart for films with comments, 15 values were modified. If the number of published films exceeded 15, it was randomly replaced with a whole number between 16 and 20. The modified counts of published films ranged from 16 to 45. These numbers are now represented on the chart in the range from 16 to 20.

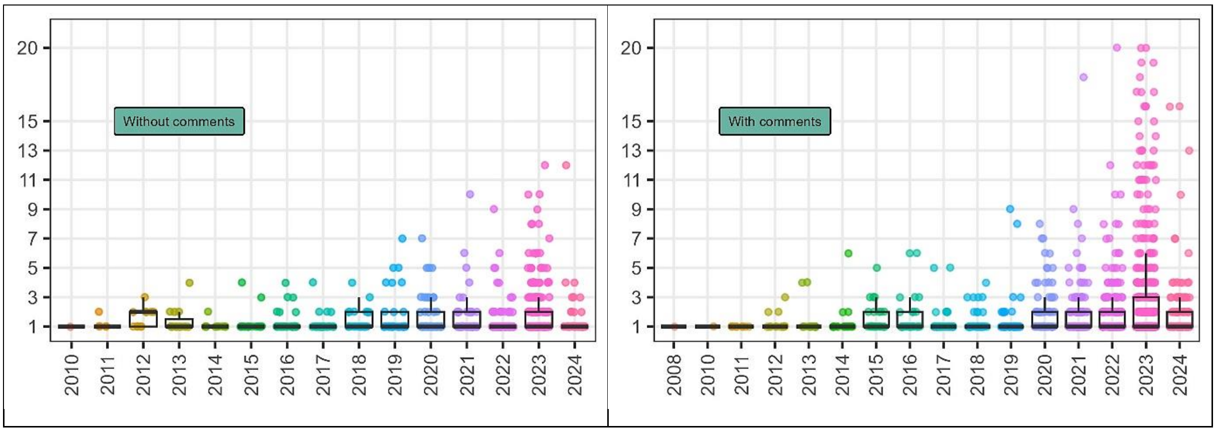

Figure 2. The number of videos published by users.

Sources: original research.

Figure 3 shows the view counts for videos, segmented into those with comments and those without. Each data point on the graph represents the view count for an individual video. To enhance the figure's clarity, a square root transformation was applied to the y-axis values, compressing higher values and expanding lower ones for better visibility. Further clarity was achieved by modifying view counts exceeding 0.8M, assigning them a fixed value between 1.04M and 1.05M. Among the videos without comments, one, published in 2024, has accumulated approximately 3.79M views to date. In contrast, among the videos with comments, there are 14 videos; the most viewed has attracted about 32.3M views, while the view counts of the remaining 13 videos range from 0.9M to 4M. The box plots shown in the figure illustrate the distribution of view counts across the videos. It can be observed that there were significantly more videos with comments, and they were viewed more frequently compared to those without comments. In the analysed dataset, there were no videos with comments from 2009, and none without comments from 2008 and 2009.

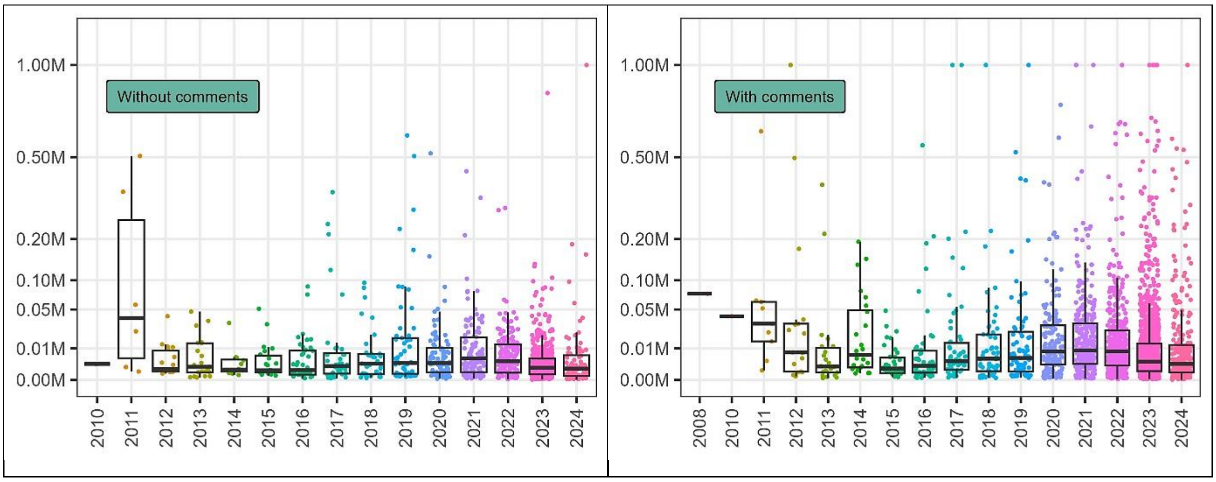

Figure 3. View counts of videos.

Sources: original research.

Table 2 presents the annual distribution of comments received by videos. For instance, videos from 2023 accumulated a total of 29,008 comments, of which 14,560 were replies to other comments. Comments in the analysed dataset were only received by videos from the year 2012 onwards. The high proportion of reply comments underscores the active exchange of information among users.

Table 2.

Total number of comments by year

Year	Comments	Comments as replies	Total
2012	20 (100%)	0 (0%)	20
2013	34 (100%)	0 (0%)	34
2014	39 (84%)	7 (16%)	46
2015	32 (68%)	15 (32%)	47
2016	54 (59%)	37 (41%)	91
2017	1003 (68%)	465 (32%)	1468
2018	318 (51%)	300 (49%)	618
2019	595 (46%)	697 (54%)	1292
2020	2637 (39%)	4011 (61%)	6648
2021	10327 (45%)	12417 (55%)	22744
2022	14161 (38%)	22491 (62%)	36652
2023	33882 (41%)	48478 (59%)	82360
2024	14006 (41%)	19848 (59%)	33854
		Total	**185874**

Sources: original research.

The left side of Figure 4 illustrates the annual count of comments received by videos, with each video depicted as a single point on the chart. For instance, a video from 2013 garnered 14 comments. The box plots included in the figure demonstrate the distribution of comment counts. For example, the third quartile for 2022 is 36, suggesting that 75% of the videos that year received no more than 36 comments. To enhance the clarity of the figure, certain values have been adjusted. The number of comments exceeding 500 were randomly replaced with whole numbers between 501 and 600. There were 40 films for which the number of comments in a given year exceeded 500 (the number of received comments ranged from 511 to 3535).

The right side of the figure shows the distribution of comments received, not categorized by year, and includes both a box plot and a density plot, which indicates that the most common number of comments received is 1. On this chart, the number of comments exceeding 500 was also modified. There were 50 such films. Throughout the study period, they received between 502 and 5020 comments in total.

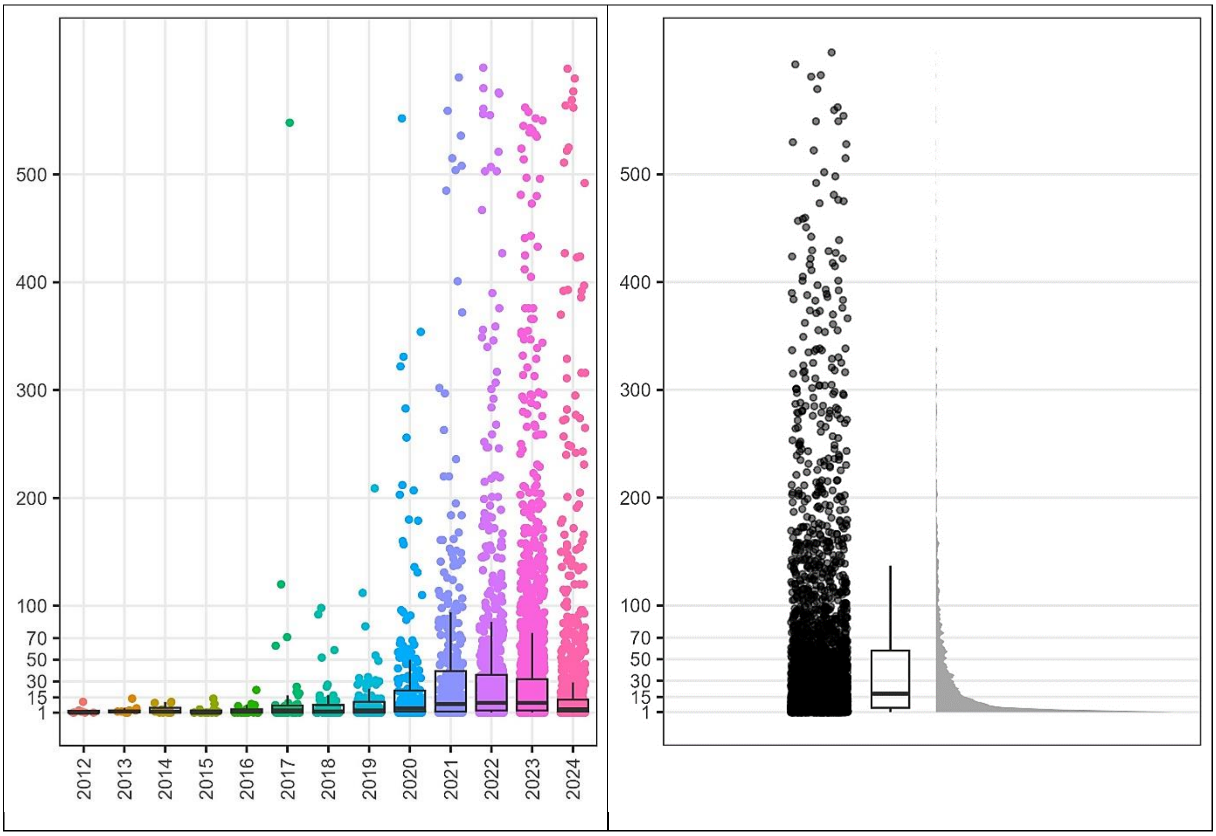

Figure 4. The number of comments received by videos.

Sources: original research.

Figure 5 shows the duration of films with comments in minutes. On the left side, the lengths are segmented by year. For instance, the longest film in 2011 was approximately 10 minutes. The box plots in the figure illustrate the distribution of film durations. For example, for 2022, the third quartile is about 16 minutes, indicating that 75% of the films from that year were 16 minutes or shorter. o improve the legibility of the diagram, some durations have been adjusted. If a film's duration exceeded 90 minutes, it was randomly adjusted to a number between 90 and 100. A total of 32 films had durations over 90 minutes, with specific durations ranging from 93 to 196 minutes. These adjusted durations are depicted on the plots within the 90 to 100-minute range. On the right side, the diagram shows the distribution of film durations without annual segmentation. Alongside the box plot, a density plot is also included, from which we can deduce that the most common film duration is approximately 1 minute.

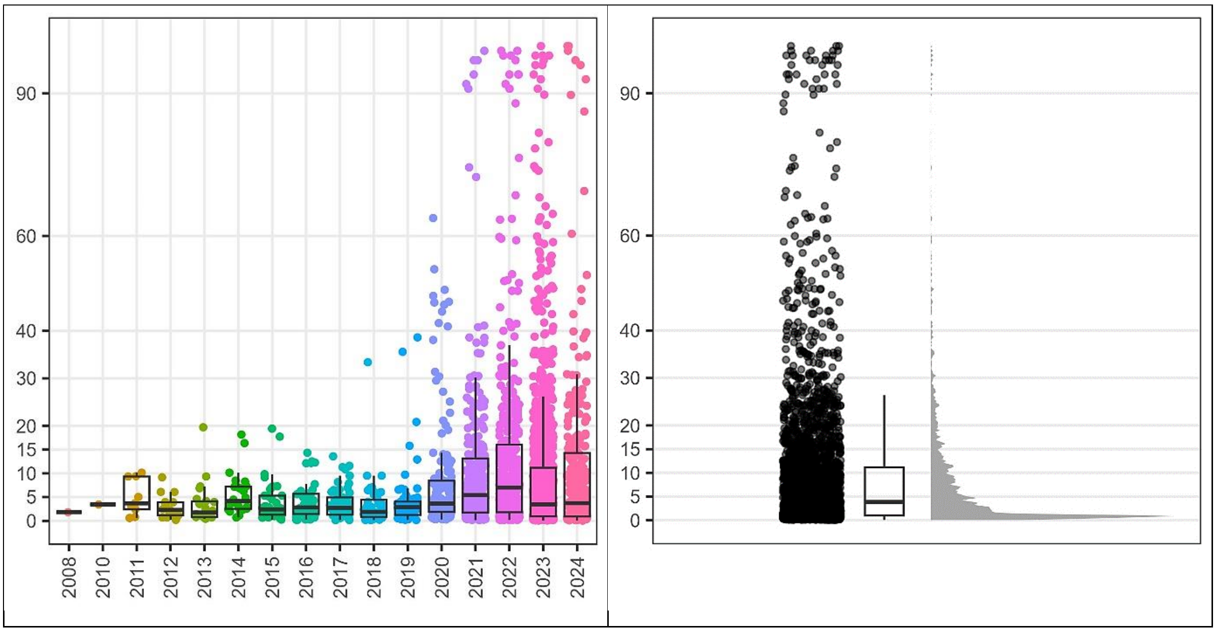

Figure 5. The duration of videos with comments in minutes.

Sources: original research.

Figure 6 and Table 3 present information about the number of words in comments. Figure 6 illustrates the number of comments containing from 4 to 100 words. These comments constituted approximately 94,8% of all comments. The largest group consisted of comments composed of 6 words. The largest group of comments consisted of 5 words. There were 9085 such comments.

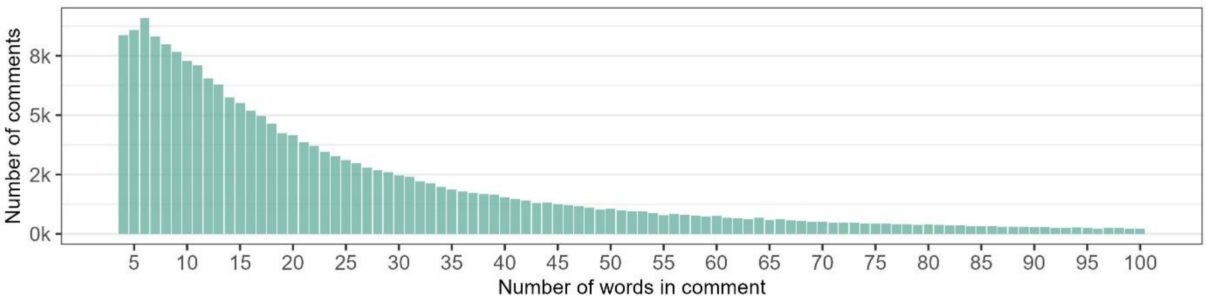

Figure 6. Number of words in comments.

Sources: original research.

Table 3.
Number of words in comments

Number of words in the comment		Number of comments
from 4 to 10		57302 (27,40%)
from 11 to 20		54386 (26,00%)
from 21 to 40		49882 (23,85%)
from 41 to 100		36757 (17,58%)
from 101 to 250		9693 (4,63%)
from 251 to 500		1018 (0,49%)
from 501 to 1388		102 (0,05%)
	Total	**209140**

Source: original research.

Table 3 displays the percentage distribution of comments by word count. The predominant group comprised comments of 4 to 10 words, accounting for 27.4 percent of all comments. Additionally, there were 102 comments containing between 501 and 1388 words.

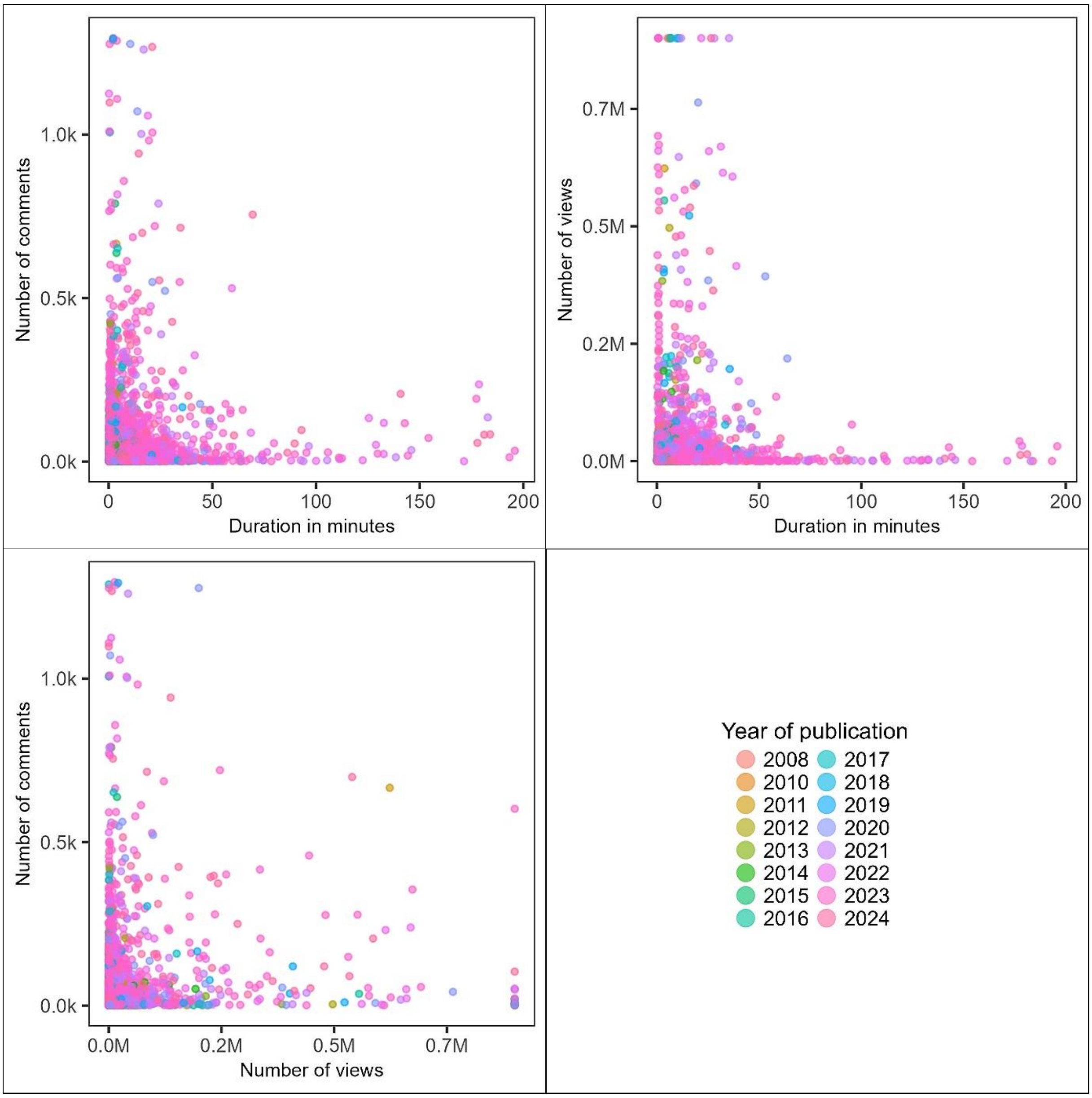

Figure 7. Relationship between duration, number of comments and views of films.

Sources: original research.

Figure 7 illustrates the correlations between three film attributes: duration (in minutes), the number of comments received, and the number of views. These relationships are depicted using three scatter plots, with the publication year of each film indicated by different colours. To enhance the clarity of the charts, selected values were modified. If the Number of views exceeded 800k, it was replaced with a random number ranging from 900.1k to 900.5k (there were 14 such instances). If the Number of comments exceeded 1.2k, it was replaced with a random number between 1.25k and 1.3k (there were 12 such instances). The variable Duration in minutes was not modified.

Figure 8 displays the annual distribution of positive, negative, and neutral comments. From this, it is evident that in 2022, 4.5% of comments were negative, 85.8% were neutral, and 9.5% were positive.

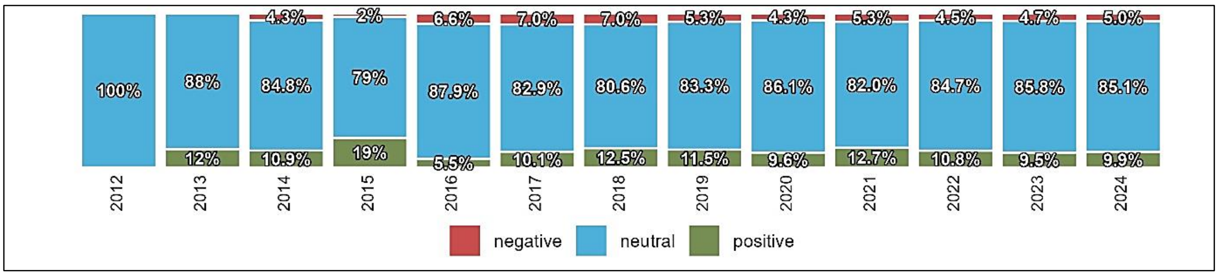

Figure 8. Percentage of positive, negative, and neutral comments.

Sources: original research.

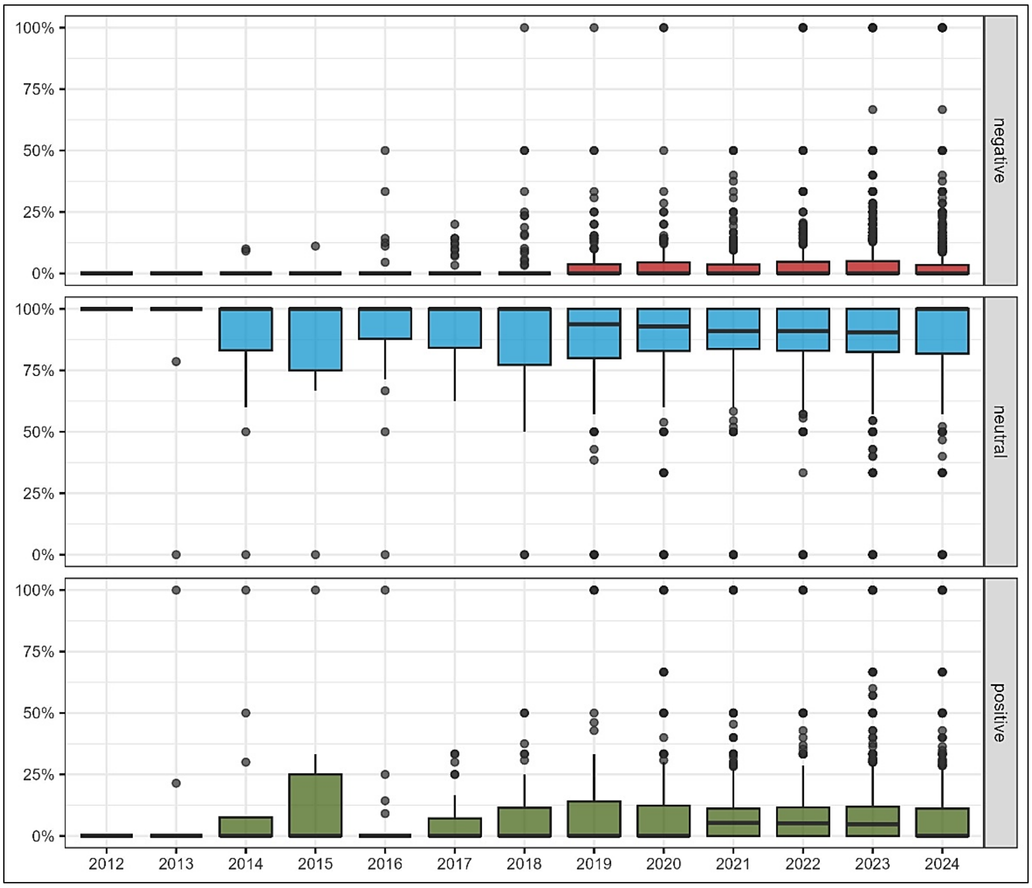

Figure 9. Distribution of the percentage of comments received by individual videos.

Sources: original research.

Figure 9 illustrates the distribution of positive, negative, and neutral comments received by individual videos, segmented by year. This distribution is represented using box plots. Individual data points showing a value of 100% for negative, positive, or neutral comments indicate that some videos received exclusively negative, positive, or neutral comment. Additionally, From the chart, we can see, for instance, that the third quartile for positive comments in 2015 being approximately 25% suggests that up to 75% of the videos received no more than 25% positive comments.

Figure 10 displays the yearly distribution of words annotated with fundamental emotions. The color green signifies positive emotions, which include feelings such as gladness, enjoyment of something expected, and trust. Conversely, negative emotions are marked in red, encompassing emotions like disgust, fear, anger, surprise with something unpredictable, and sadness.

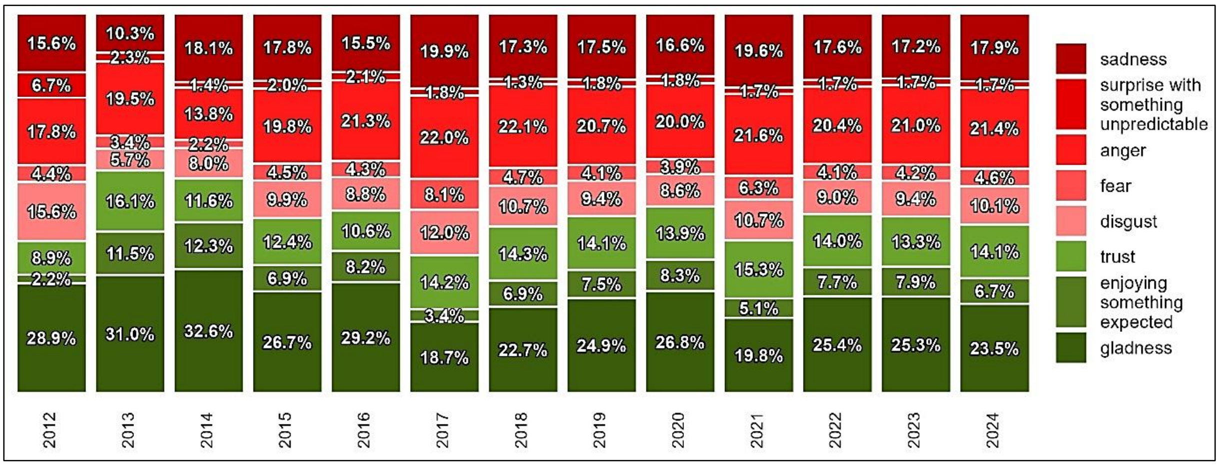

Figure 10. Percentage of words with annotated basic emotions by year.

Sources: original research.

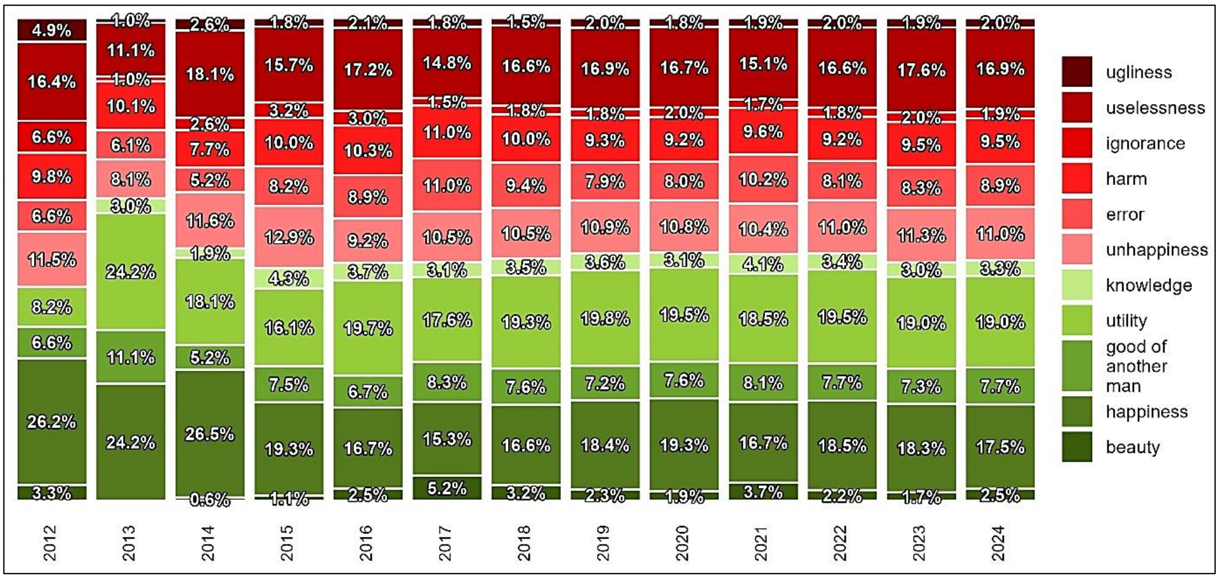

Figure 11. Percentage of words with annotated fundamental human values by year.

Sources: original research.

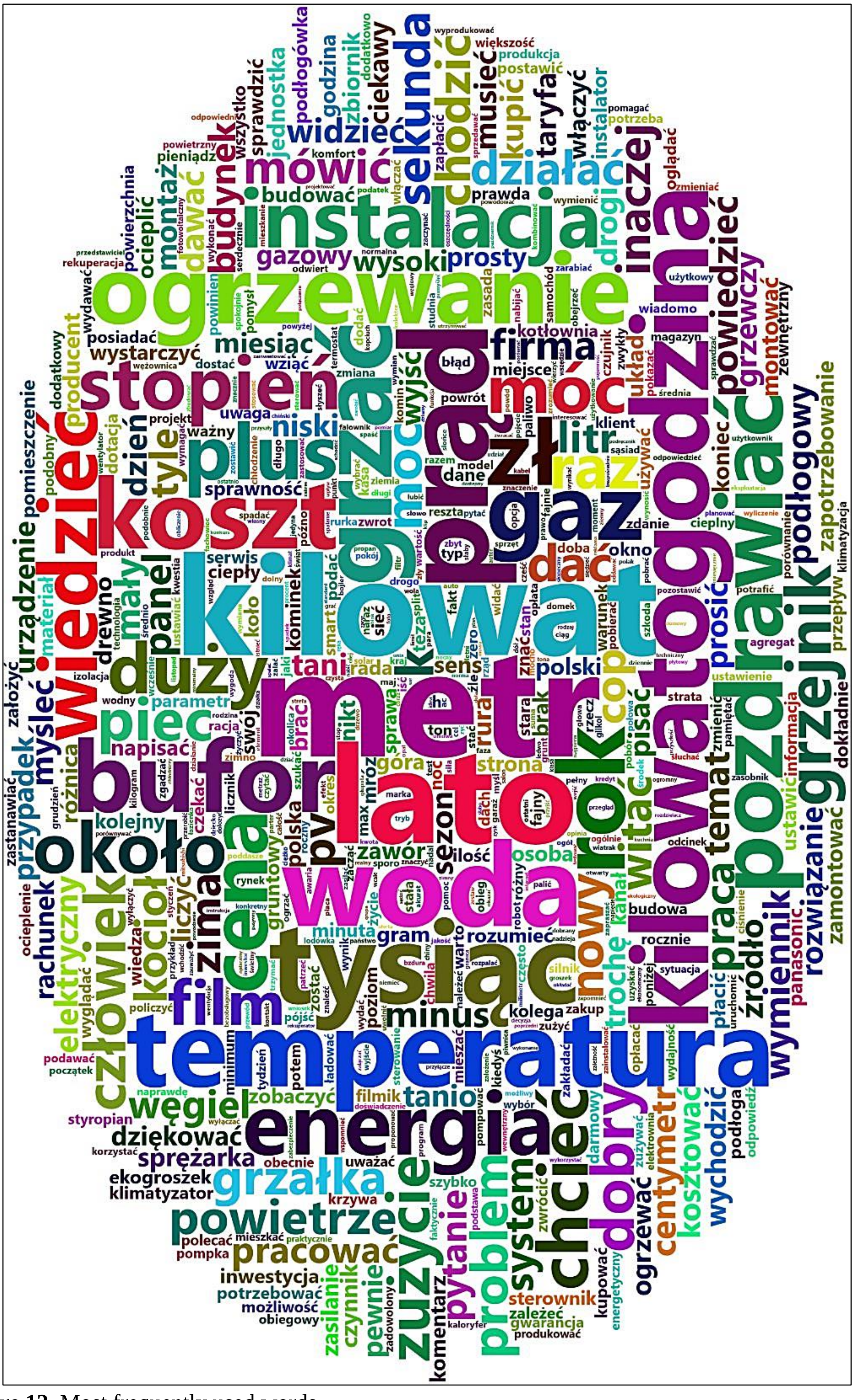

Figure 12. Most frequently used words.

Sources: original research.

Figure 11 illustrates the distribution of words annotated with fundamental human values. Positive human values are depicted in green, including beauty, happiness, the good of another man, utility, and knowledge. Notably, the value "truth" does not appear among the positive human values. Negative human values are highlighted in red and include unhappiness, error, harm, ignorance, uselessness, and ugliness.

Figure 12 displays a word cloud of the most commonly used words in comments. This visualization allows for an analysis of the themes present in the comments. The size of each word in the cloud is proportional to its frequency of occurrence, with more common words appearing larger. The two most frequently occurring words, "pompa" (ang. pump) - 69,288 occurrences and "ciepło" (ang. heat) - 64,207 occurrences, were removed from the word cloud.

4. Conclusion

The analysis of data concerning videos related to heat pump (including the number of published videos, their view counts, the number of comments, and their sentiment) allowed for drawing the following conclusions:

- corresponding to an increasing number of videos and comments, it can be observed that interest in heat pump is continuously growing, particularly after the year 2019;
- it's evident that there is variation in the views count of videos; videos with comments were more frequently viewed;
- retrieved videos most commonly had one comment and a duration of one minute;
- among the comments, those with a neutral tone are predominant;
- starting from 2019, there are generally twice as many positive comments as negative ones;
- the videos exhibit variation in the distribution of the number of positive, negative, and neutral comments, among them were those that received only negative, positive, or neutral comments;
- analysing the most commonly used words reveals that comments were not solely focused on heat pumps but also covered general topics:
  - heating systems for houses and buildings, using different equipment and energy carriers,
  - system for heating water and air,
  - the type of heat pump,
  - the air quality,
  - systems for producing electricity from solar energy,
  - the information about solutions encouraging the replacement of non-ecological heat sources with ecological ones.

The research conducted confirms that comments on YouTube videos can serve as a valuable data source for understanding people's thoughts, feelings, and opinions about heat pumps. It should be noted that this study specifically analysed the views of Polish-speaking YouTube users.

References

1. Ağrali, Ö., Aydin, Ö. (2021). Tweet Classification and Sentiment Analysis on Metaverse Related Messages. *Journal of Metaverse*, *1*(1), 25–30.
2. Aydın, E., Yılmaz, E. (2021). YouTube as a Source of Information on Echocardiography: Content and Quality Analysis. *Acta Cardiologica Sinica*, *37*(5), 534–541. https://doi.org/10.6515/ACS.202109_37(5).20210514A
3. Bergman, N. (2013). Why is renewable heat in the UK underperforming? A socio-technical perspective. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, *227*(1), 124–131.
4. Billimoria, S., Guccione, L., Henchen, M., Louis-Prescott, L. (2021). The Economics of Electrifying Buildings: How Electric Space and Water Heating Supports Decarbonization of Residential Buildings. In: *World Scientific Encyclopedia of Climate Change: Case Studies of Climate Risk, Action, and Opportunity, Vol. 3* (pp. 297–304). World Scientific.
5. Chua, K.J., Chou, S.K., Yang, W.M. (2010). Advances in heat pump systems: A review. *Applied Energy*, *87*(12), 3611–3624.
6. *CLARIN-PL* (n.d.). Retrieved from: http://clarin-pl.eu/, June 5, 2022.
7. 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. https://doi.org/10.1016/J.ERSS.2022.102515
8. Das, S., Dutta, A., Medina, G., Minjares-Kyle, L., Elgart, Z. (2019). Extracting patterns from Twitter to promote biking. *IATSS Research*, *43*(1), 51–59. https://doi.org/10.1016/j.iatssr.2018.09.002
9. Das, S., Sun, X., Dutta, A. (2015). Investigating User Ridership Sentiments for Bike Sharing Programs. *Journal of Transportation Technologies*, *5*(2), 69–75. https://doi.org/10.4236/jtts.2015.52007
10. Decuypere, R., Robaeyst, B., Hudders, L., Baccarne, B., de Sompel, D. (2022). Transitioning to energy efficient housing: Drivers and barriers of intermediaries in heat pump technology. *Energy Policy*, *161*, 112709.
11. Deng, H., Ergu, D., Liu, F., Cai, Y., Ma, B. (2022). Text sentiment analysis of fusion model based on attention mechanism. *Procedia Computer Science*, *199*, 741–748.
12. Edenhofer, O. (2015). *Climate change 2014: mitigation of climate change, Vol. 3*. Cambridge University Press.

13. Evans-Cowley, J.S., Griffin, G. (2012). Microparticipation with Social Media for Community Engagement in Transportation Planning. *Transportation Research Record: Journal of the Transportation Research Board*, *2307*(1), 90–98. https://doi.org/10.3141/2307-10
14. Grubljesic, T., Coelho, P.S., Jaklic, J. (2019). The Shift to Socio-Organizational Drivers of Business Intelligence and Analytics Acceptance. *Journal of organizational and end user computing*, *31*(2), 37–64. https://doi.org/10.4018/JOEUC.2019040103
15. Ibar-Alonso, R., Quiroga-García, R., Arenas-Parra, M. (2022). Opinion Mining of Green Energy Sentiment: A Russia-Ukraine Conflict Analysis. *Mathematics, Vol. 10*, *10*(14), 2532. https://doi.org/10.3390/MATH10142532
16. Jain, A., Jain, V. (2019a). Renewable Energy Sources for Clean Environment: Opinion Mining. *Asian Journal of Water, Environment and Pollution*, *16*(2), 9–14. https://doi.org/10.3233/AJW190013
17. Jain, A., Jain, V. (2019b). Sentiment classification of twitter data belonging to renewable energy using machine learning. Https://Doi.Org/10.1080/02522667.2019.1582873, *40*(2), 521–533. https://doi.org/10.1080/02522667.2019.1582873
18. Janz, A., Kocoń, J., Piasecki, M., Zaśko-Zielińska, M. (n.d.). *plWordNet as a Basis for Large Emotive Lexicons of Polish CLARIN-PL View project Liner2-A Customizable Framework for Automatic Text Annotation (NER, TimeX, Events) View project*. Retrieved from: https://www.researchgate.net/publication/322684200, June 5, 2022.
19. 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. https://doi.org/10.3390/SU13052673
20. Kircher, K.J., Zhang, K.M. (2021). Heat purchase agreements could lower barriers to heat pump adoption. *Applied Energy*, *286*, 116489.
21. Loureiro, M.L., Alló, M. (2020). Sensing climate change and energy issues: Sentiment and emotion analysis with social media in the U.K. and Spain. *Energy Policy*, *143*, 111490. https://doi.org/10.1016/J.ENPOL.2020.111490
22. Martiskainen, M., Schot, J., Sovacool, B.K. (2021). User innovation, niche construction and regime destabilization in heat pump transitions. *Environmental Innovation and Societal Transitions*, *39*, 119–140.
23. Muhammad, A.N., Bukhori, S., Pandunata, P. (2019). Sentiment Analysis of Positive and Negative of YouTube Comments Using Naïve Bayes – Support Vector Machine (NBSVM) Classifier. *2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 199–205. https://doi.org/10.1109/ICOMITEE.2019.8920923
24. Nowak, T. (2018). Heat Pumps: Integrating technologies to decarbonise heating and cooling. *European Copper Institute*.
25. Omnicore (n.d.). *YouTube by the Numbers: Stats, Demographics & Fun Facts*. Omnicoreagency.Com. Retrieved from: https://www.omnicoreagency.com/youtube-

statistics/, May 9, 2024.
26. Pang, B., Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. *ArXiv Preprint Cs/0409058*.
27. Pang, B., Lee, L. (2008). *Opinion mining and sentiment analysis. Found Trends Inf. Retr., 2(1-2),* 1-135.
28. Pang, B., Lee, L., Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing ({EMNLP} 2002)*, 79–86. https://doi.org/10.3115/1118693.1118704
29. Plutchik, R. (1980). Emotion: A Psychoevolutionary Synthesis. *The American Journal of Psychology*, *93*(4), 751. https://doi.org/10.2307/1422394
30. Puzynina, J. (1992). *Język wartości,* 264.
31. Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. *Proceedings of the ACL Student Research Workshop*, 43–48.
32. Snape, J.R., Boait, P.J., Rylatt, R.M. (2015). Will domestic consumers take up the renewable heat incentive? An analysis of the barriers to heat pump adoption using agent-based modelling. *Energy Policy*, *85*, 32–38.
33. Snelson, C. (2011). YouTube across the Disciplines : A Review of the Literature. *Journal of Online Learning and Teaching*, *7*(1), 159–169. http://www.watchknow.org.
34. Soltani, M., Kashkooli, F.M., Dehghani-Sanij, A.R., Kazemi, A.R., Bordbar, N., Farshchi, M.J., Elmi, M., Gharali, K., Dusseault, M.B. (2019). A comprehensive study of geothermal heating and cooling systems. *Sustainable Cities and Society*, *44*, 793–818.
35. Twersky, C. (n.d.). *scrapetube*. Retrieved from: https://pypi.org/project/scrapetube/, June 13, 2023.
36. Ürge-Vorsatz, D., Cabeza, L.F., Serrano, S., Barreneche, C., Petrichenko, K. (2015). Heating and cooling energy trends and drivers in buildings. *Renewable and Sustainable Energy Reviews*, *41*, 85–98.
37. Wierzbicka, A. (1992a). *Semantics, culture, and cognition : universal human concepts in culture-specific configurations*, 487.
38. Wierzbicka, A. (1992b). Defining Emotion Concepts. *Cognitive Science*, *16*(4), 539–581. https://doi.org/10.1207/S15516709COG1604_4
39. Xu, Q.A., Chang, V., Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, *3*, 100073. https://doi.org/10.1016/J.DAJOUR.2022.100073
40. *youtube-comment-downloader* (n.d.). Retrieved from: https://pypi.org/project/youtube-comment-downloader/, June 13, 2023.
41. Zarrabeitia-Bilbao, E., Morales-I-gras, J., Rio-Belver, R.M., Garechana-Anacabe, G. (2022). Green energy: identifying development trends in society using Twitter data mining to make strategic decisions. *Profesional de La Información*, *31*(1). https://doi.org/10.3145/EPI.2022.ENE.14