S I L E S I A N U N I V E R S I T Y O F T E C H N O L O G Y P U B L I S H I N G H O U S E

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2023
ORGANIZATION AND MANAGEMENT SERIES NO. 185

SENTIMENT ANALYSIS OF COMMENTS POSTED ON YOUTUBE VIDEOS RELATED TO PHOTOVOLTAICS

Marcin WYSKWARSKI

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

Purpose: Based on sentiment analysis of comments posted on YouTube, determining people's thoughts, feelings and opinions on photovoltaics.
Design/methodology/approach: Comments posted on videos were downloaded automatically. The comments' content has undergone preprocessing. All characters other than letters, URLs, hashtags, emojis, and words used to search for videos taken out of their text. The comment's sentiment value was determined. To display the proportion of favourable, negative, and neutral comments, visualisations were created. The word cloud was employed to display the comments' most popular words.
Findings: For comments posted on videos related to photovoltaics, proportions of positive, negative and neutral comments were determined. The information about the number of published videos, the view count of videos, the length of videos, the number of published comments, and the length of comments has been obtained.
Research limitations/implications: Only comments posted on videos which contained the word "photovoltaic" were downloaded, Only Polish-language comments' content was examined. Without author oversight, sentiment analysis was carried out automatically by the "ccl emo" service. Only viewpoints expressed by YouTube users were analysed. It was assumed that if the title of the video contains the word photovoltaic, its comments content is related to photovoltaic.
Practical implications: Automated assessment of people's opinions regarding photovoltaics.
Originality/value: Opinions on photovoltaics were collected. Based on the growing number of videos and comments, it was found that interest in photovoltaics in Poland is steadily growing.

Keywords: sentiment analysis, YouTube, photovoltaics, text mining.

Category of the paper: research paper, case study.

1. Introduction

As the impacts of global warming become increasingly evident, there is a rising concern regarding the adverse effects of the conventional energy industry on the environment. Communities are actively engaging in measures aimed at decreasing the emission of greenhouse gases (Decuypere et al., 2022; Peng et al., 2013; Pestana et al., 2018). On a global scale, the adoption of renewable energy sources is expanding to contribute towards the mitigation of air pollution and the reduction of carbon emissions (Dincer et al., 2000; Moriarty, Honnery, 2011). Numerous nations recognize the advantages of green energy, resulting in policy shifts concerning energy procurement (Bórawski et al., 2019; Eyl-Mazzega, Mathieu, 2020; Omri et al., 2015; Pellerin-Carlin et al., n.d.; Salim, Rafiq, 2012).

Public acceptance and support for renewable energy play a crucial role in facilitating the transition towards a low-carbon energy system (Kim et al., 2021). Public sentiment and opinions regarding renewable energy have been conducted (Hamilton et al., 2019; Kim et al., 2021; Lee, 2022; Noblet et al., 2015; Peñaloza et al., 2022; Qazi et al., 2019; Stokes, Warshaw, 2017).

The worldwide market for renewable energy sources (RES), particularly in the solar and wind sectors, is experiencing a consistent growth trajectory, unaffected even by the challenges posed by the coronavirus pandemic (Bhuiyan et al., 2021; Bilgili, Ozturk, 2015; Eroğlu, 2021; Quitzow et al., 2021). Out of all renewable energy sources, photovoltaic technology holds the highest potential due to its affordability and straightforward installation process (Alves dos Santos et al., 2021; C.B. et al., 2021; Castilho et al., 2021; Mota et al., 2020).

The photovoltaic sector in Poland is characterized by a decentralized structure, predominantly relying on micro installations. By the conclusion of 2019, micro-installations represented over 70% of Poland's overall installed photovoltaic capacity. Residential participation in photovoltaics was promoted through solar energy support programs such as the governmental initiative "My Electricity", and long-term European Union support under the Regional Operational Programs (Grębosz-Krawczyk et al., 2021).

In recent years, Poland's households, industries, and service sectors have grappled with escalating electricity bills (Chomać-Pierzecka et al., 2022). Rising electricity prices have increased interest in photovoltaics as an alternative, but choosing the right solution can be complex. In Poland, several factors contribute to this situation, including but not limited to:

- optimal installation size (avoiding excessive electricity production that may not be economically viable) (Zrównoważonego et al., 2015),
- terms and conditions for accounting for excess electricity production with the distribution system operator (Zator, Lambert-Torres, 2021),
- evaluating the cost-effectiveness of investing in an electricity storage system (Zator, Lambert-Torres, 2021),

– limited familiarity with technical criteria for selecting the suitable energy solution, often leading to purchase decisions primarily driven by factors such as installation cost, lifespan of photovoltaic panels, availability of implementation options, and aesthetic considerations (Chomać-Pierzecka et al., 2022).

In the era of digital advancements, individuals often share their thoughts and opinions on social media platforms, expressing their ideas to a wide audience. Sentiment analysis can be utilized as a means to analyse and understand people's thoughts, emotions, and opinions. Sentiment analysis can be an alternative approach to traditional surveys and interviews. It provides an automated approach to analyse sentiment, emotions, and opinions expressed in written language. It is a process of analysing, processing, generalizing and making sense of emotionally charged subjective texts (Deng et al., 2022). It offers a method for extracting valuable insights from textual data efficiently and effectively (Xu et al., 2022). It can be performed to assess an individual's perspective or inclination towards a subject or issue, determining whether it leans towards a positive or negative viewpoint (Pang et al., 2002).

A big source of data with people's opinions can be the comments posted on YouTube videos. YouTube is an online video platform that is quickly expanding and receives nearly two billion views daily (Aydın, Yılmaz, 2021; Snelson, 2011). According to data as of March 14, 2023, 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, 2021). As the world's biggest video platform, YouTube showcases a diverse range of media content produced by either companies or individuals. This content encompasses music videos, promotional videos for products, vlogs, review videos, and educational content (Muhammad et al., 2019).

A variety of tools can be used for analysing data retrieved from the Internet. Due to a significant volume of data, methods such as text mining, data mining, machine learning, topic modelling, sentiment analysis and similar approaches are employed. The exploration of data collected from social media constitutes a new field. Its popularity is growing due to cost-effectiveness, easy access, and the element of anonymity (Das et al., 2015, 2019; Evans-Cowley, Griffin, 2012). There are many studies in the literature about sentiment analysis on data extracted from the Internet (Ağrali, AYDIN, 2021; Pang, Lee, 2004, 2008; Read, 2005). The use of sentiment analysis of text to find out people's 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).

Sentiment analysis is a tool to understand how society perceives photovoltaics. It is a valuable tool for policymakers, investors and businesses alike, helping to shape a positive image and support the development of this sustainable form of energy. Sentiment analysis provides valuable feedback on societal attitudes towards photovoltaics, which, in turn, can impact decision-making, investments, and actions in the field.

2. Research Methodology

On March 22, 2023, in service, YouTube 2.960 videos related to photovoltaics were found. This was accomplished using the Python "scrapetube" (Twersky, n.d.) library. This library allows search for videos without the official YouTube API (application programming interface). The title of the video had to include one or more of the following nouns and/or adjectives in Polish:

- nouns: "fotowoltaika", "fotowoltaice", "fotowoltaiką", "fotowoltaikę", "fotowoltaiki", "fotowoltaiko", "fotowoltaik", "fotowoltaikach", "fotowoltaikami", "fotowoltaikom",
- adjectives: "fotowoltaiczna", "fotowoltaiczną", "fotowoltaicznego", "fotowoltaicznej", "fotowoltaicznemu", "fotowoltaiczni", "fotowoltaicznych", "fotowoltaicznym", "fotowoltaicznymi", "fotowoltaiczne", "fotowoltaiczny".

These nouns and adjectives are in all possible grammatical cases for the Polish language and are translations of the terms "photovoltaics".

In the next step for each video, comments posted by users were downloaded. The "youtube-comment-downloader" package was used for this. It is a simple script for downloading YouTube comments without using the YouTube API (*Youtube-Comment-Downloader*, n.d.).

In the next step, the author removed:

- comments were written in languages other than Polish,
- the comments whose content was the same as the content of other comments and were posted by the same user (it was frequently an advertisement for a company's services, products, or jobs); the content was treated as a string of characters and compared using the comparison operator "==".

Then the comments' content was pre-processed. URLs, hashtags, emojis, user names, terms used to search for videos, and all characters other than letters were removed from the comments. Next, the number of words in the cleaned content of each comment was checked. Comments with less than 4 words have been removed. After these operations, the number of comments was 136.416. These comments were posted on one of the 1565 videos. The remaining 1395 videos had no comments, or those comments were removed during the pre-processed stage.

In the next step the ***ccl_emo***[1] service, created by CLARIN-PL[2], was used. In Polish, this service is also known as "Wydźwięk" and "Sentiment" (in English). It is a service for statistically analysing texts' overtones and emotions (Grubljesic et al., 2019; Janz et al., n.d.). Also, others CLARIN-PL's services were used. These were:

[1] https://wiki.clarin-pl.eu/pl/nlpws/services/ccl_emo; https://clarin-pl.eu/index.php/wydzwiek/
[2] CLARIN-PL is a Polish scientific consortium, part of the European Research Infrastructure CLARIN (Common Language Resources and Technology Infrastructure) (*CLARIN-PL*, n.d.)

- Any2txt - a service that transforms text files (e.g. doc, docx, xlsx) into text.
- Speller2 - a service that verifies the text's spelling. It uses a tool called Autocorrect[3] for this.
- Wcrft2 - is a basic morpho-syntactic tagger for Polish.
- WSD - a service for word sense disambiguation, which works for Polish texts. As a source of possible senses, it uses plWordNet, which consists of lexical units grouped into synsets that are linked by lexico-semantic relations. A lexical unit represents a lexical meaning and is a triple: lemma, part of speech and sense identifier (Janz et al., n.d.).

The selected lexical units stored in plWordNet were added emotive annotation. Lexical units were described by (Janz et al., n.d.):

- sentiment polarity: it was assessed on a 5-point scale: strong and weak vs. negative and positive, plus neutral;
- basic emotions: gladness, trust, enjoyment of something expected, sadness, anger, fear, disgust, and surprise with something unpredictable - these emotions were derived from the 8 basic emotions identified by Plutchik and his Wheel of Emotions (Plutchik, 1980; Wierzbicka, 1992a, 1992b);
- fundamental human values: utility, good of another man, truth, knowledge, beauty, happiness, uselessness, harm, ignorance, error, ugliness, unhappiness - basic human values indicated by (Puzynina, 1992) were used.

Table 1.
Example of calculating the sentiment of a comment

Sample comment	Niech ***sprawdzi*** [1] lodówkę. U mnie przez ***uszkodzoną*** [-1] uszczelkę rachunki ***mocno*** [1] latem wzrosły.
Sentiment calculation	***sprawdzi*** [1] + ***mocno*** [1] = 2 ***uszkodzoną*** [-1] = -1
	The number of positive words (2) > The number of negative words (1)
	The sentiment of the comment = positive

Sources: original research.

In the next step, each comment was examined to determine the number of words it contained with annotated basic emotions and fundamental human values. According to Table 2, the example comment consisted of 3 words expressing "gladness", 1 word indicating " enjoying something expected ", and 1 word conveying "trust".

[3] https://languagetool.org/pl/

Table 2.
The number of words with annotated basic emotions for sample comment

Comment Id	gladness	enjoying something expected	trust	disgust	fear	anger	surprise with something unpredictable	sadness
Ugw3f5BC2Nq6nDyXWA14AaABAg	1	1	1	0	0	0	0	0

Sources: original research.

3. Results

Using a ribbon chart shown in Figure 1, the number of retrieved videos has been presented with a division into videos that received comments and videos without comments. More detailed information on this subject has also been presented in Table 3. As can be seen, the number of movies began to significantly increase in the year 2021 (with 547 published videos), reaching its peak in the year 2022 (with 1805 videos). In the year 2023, the number of videos is smaller. However, it's important to remember that the movies were searched on March 22, 2023. Therefore, it's uncertain how many more videos will be published in the upcoming months of 2023. Attempting to estimate their number in the year 2023 is difficult because the month of publication is unknown (only the year was known).

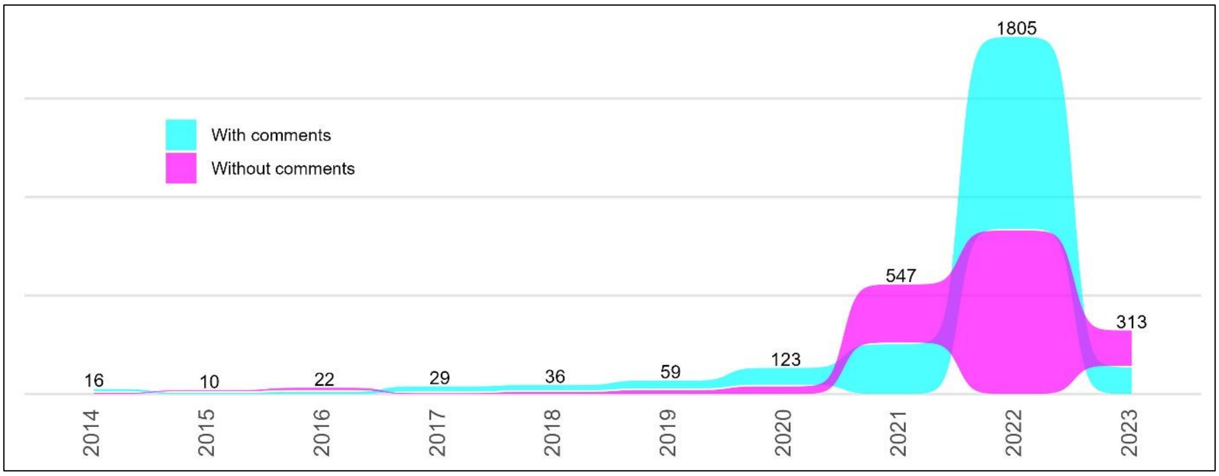

Figure 1. Distribution of the retrieved videos over the years. Note: videos for 2023 only gathered until March 22, 2023.

Source: original research.

Table 3.
Distribution of the retrieved videos over the years. Note: videos for 2023 only gathered until March 22, 2023

	2014	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**Total**
With comments	11 (69%)	5 (50%)	11 (50%)	25 (86%)	26 (72%)	40 (68%)	85 (69%)	252 (46%)	975 (54%)	135 (43%)	**1565** (53%)
Without comments	5 (31%)	5 (50%)	11 (50%)	4 (14%)	10 (28%)	19 (32%)	38 (31%)	295 (54%)	830 (46%)	178 (57%)	**1395** (47%)
Total	**16**	**10**	**22**	**29**	**36**	**59**	**123**	**547**	**1805**	**313**	**2960**

Sources: original research.

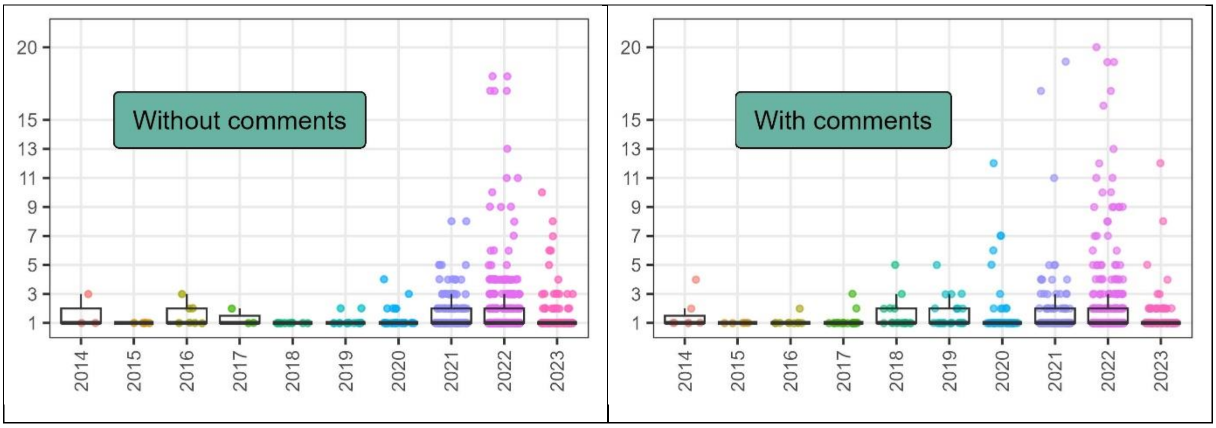

Figure 2. The number of videos published by users.

Sources: original research.

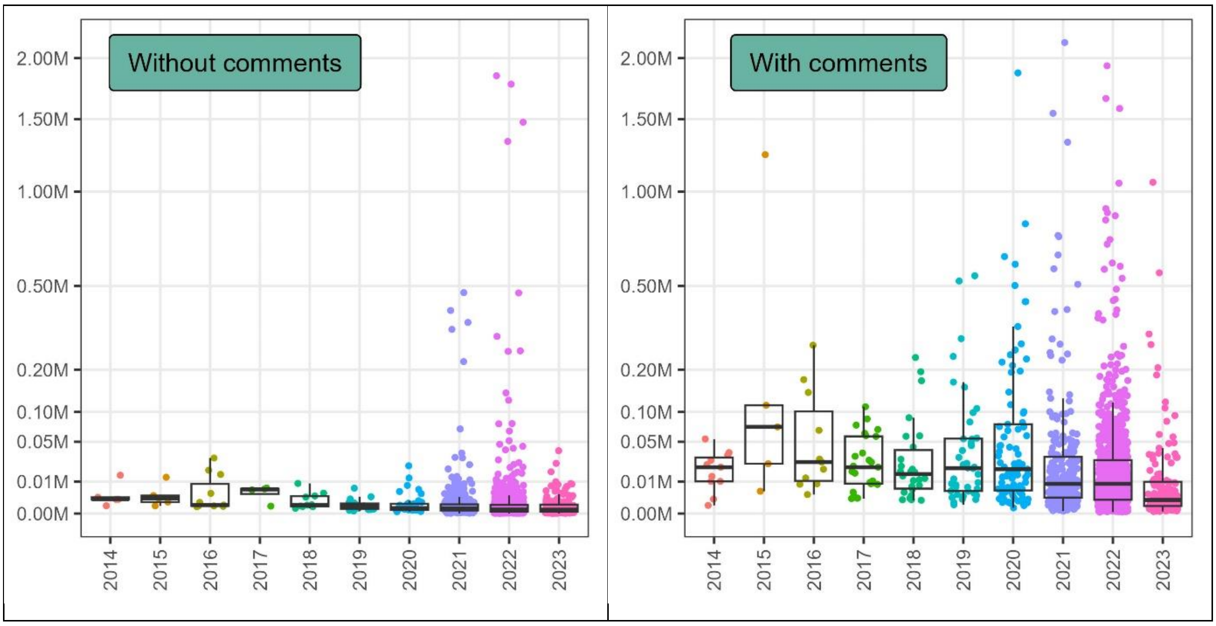

Figure 3. Views count of videos.

Sources: original research.

In Figure 3, the views count of videos is depicted, divided into videos with comments and without comments. To improve the clarity of the figure, a square root transformation was applied to the values presented on the y-axis. This has compressed high values while making low values more spread out. The successive values on the y-axis were determined by the author. Each point on the chart represents the number of views for a single video. The box plots depicted in the figure provide insights into the distribution of view counts for the videos. It can be observed that videos with comments were more frequently viewed compared to those without comments. Longer box plots for videos with comments indicate that the view counts exhibited more diverse values. From the figure, it's also evident that one of the videos from 2021 has been viewed more than 2 million times.

Table 4 presents the number of comments obtained by videos categorized by years. It's evident, for instance, that the total number of comments received by videos from 2021 is 21,909. Some of these (12,991 comments) were replies to other comments. The presence of comments that are replies to other comments indicates an exchange of information between users.

Table 4.
Total number of comments by year

	2014	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Comments	3 (100%)	20 (50%)	52 (44%)	139 (55%)	184 (32%)	714 (40%)	4438 (39%)	8918 (41%)	40990 (47%)	6700 (52%)
Comments as replies	0 (0%)	20 (50%)	65 (56%)	113 (45%)	384 (68%)	1050 (60%)	7046 (61%)	12991 (59%)	46463 (53%)	6121 (48%)
Total	**3**	**40**	**117**	**252**	**568**	**1764**	**11484**	**21909**	**87453**	**12821**

Sources: original research.

The left side of Figure 4 depicts the number of comments received by videos in each year. Each video is represented as a single point on the chart. For example, one of the videos in 2016 received 51 comments. The box plots in the figures allow us to observe the distribution of the number of comments received. For instance, the third quartile for the year 2021 is 52, indicating that 75% of the films in that year received 52 or fewer comments. To improve the clarity of the figure, some numbers have been modified. If the number of comments exceeded 500, it was randomly replaced with a whole number ranging from 501 to 600. The number of videos with comments above 500 was 37 (for the year 2019: 660 comments; for the year 2020: 820, 959 and 2963 comments; for the year 2021: 508, 516, 548, 550, 587, 724, 958, 973, 1025 and 1206 comments; for the year 2022: 502, 517, 519, 530, 543, 546, 587, 620, 629, 678, 710, 888, 1106, 1155, 1370, 1408, 1533, 1788, 2370 and 2470 comments; for the year 2023: 689, 867, and 1264 comments). These numbers are now represented on the charts in the range of 500 to 600. The maximum number of comments received by one of the films in the year 2020, was 2963 comments. The right side of the figure represents the distribution of the number of comments received without categorization by year. In addition to the box plot, a density plot is also included. We can deduct from it, that the most frequently received number of comments is 1.

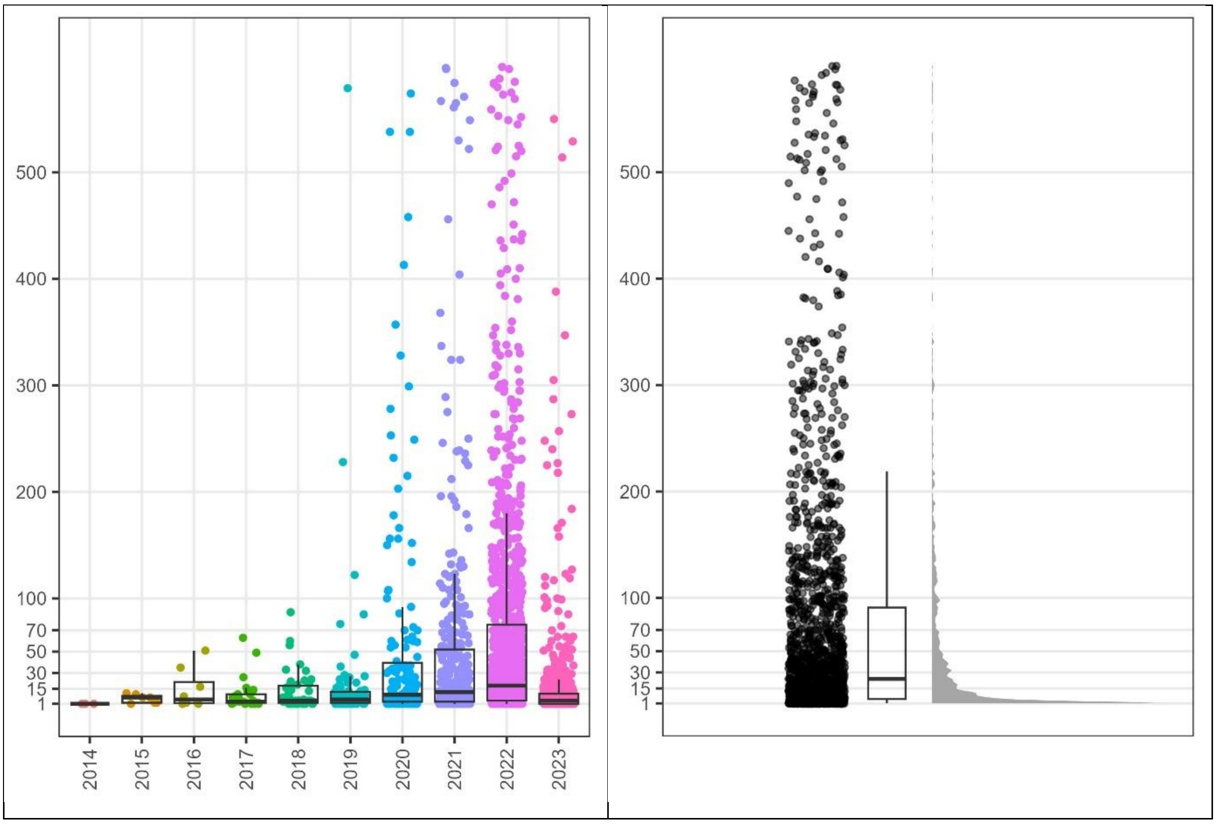

Figure 4. The number of comments received by videos.

Sources: original research.

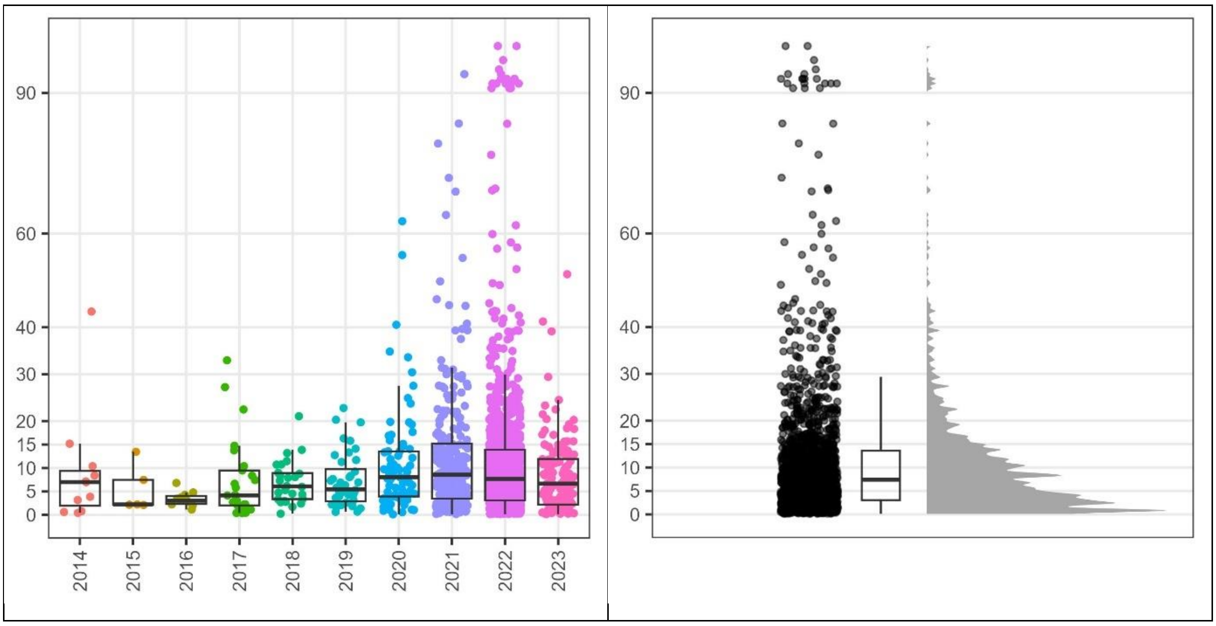

Figure 5. The duration of videos with comments in minutes.

Sources: original research.

Figure 5 presents in minutes the duration of films with comments. The left side represents the length of films divided by year. For example, the longest film from 2018 lasted approximately 21 minutes. The box plots shown in the figure allow us to see the distribution of film durations. For instance, the third quartile for the year 2021 is around 15.2 minutes, meaning that 75% of films in that year lasted 15.2 minutes or less. To enhance the readability of the drawing, some durations have been modified. If the duration exceeded 90 minutes, it was randomly changed to a number between 90 and 100. The number of films with a duration over 90 minutes was 17 (the durations of these films were as follows: 91, 101, 104, 104, 109, 115, 130, 132, 142, 150, 153, 169, 185, 186, 187, 191, 202, 438, 607 minutes). These durations are represented on the plots in the range from 90 to 100. The maximum film duration was 607 minutes. The right side of the drawing presents the distribution of film durations without division by years. In addition to the box plot, a density plot is included. From it, we can infer that films most commonly lasted around 1 minute.

Figure 6 and Table 5 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 95% of all comments. The largest group consisted of comments composed of 5 words. The largest group of comments consisted of 5 words. There were 6191 such comments.

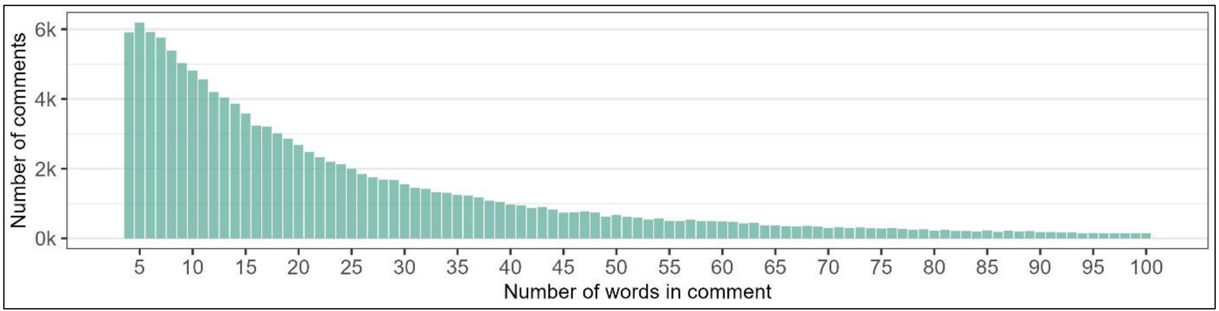

Figure 6. Number of words in comments.

Sources: original research.

Table 5.
Number of words in comments

Number of words in the comment	Number of comments
from 04 to 10	39016 (28,6%)
from 11 to 20	35268 (25,85%)
from 21 to 40	31900 (23,39%)
from 41 to 100	23467 (17,2%)
from 101 to 250	5996 (4,4%)
from 251 to 500	668 (0,49%)
from 501 to 1000	82 (0,06%)
from 1001 to 1388	14 (0,01%)
Total	**136411 (100%)**

Source: original research.

Table 5 shows the percentage distribution of comments based on the number of words. The largest group consisted of comments composed of 4 to 10 words, making up 28.6 per cent of all comments. 14 comments had between 1001 and 1388 words.

Figure 7 presents the relationship among three film parameters: duration (measured in minutes), number of received comments, and number of views. The relationship is visualized through three scatter plots. No adjustments were made to the outliers. The publication year of each film is represented by colour.

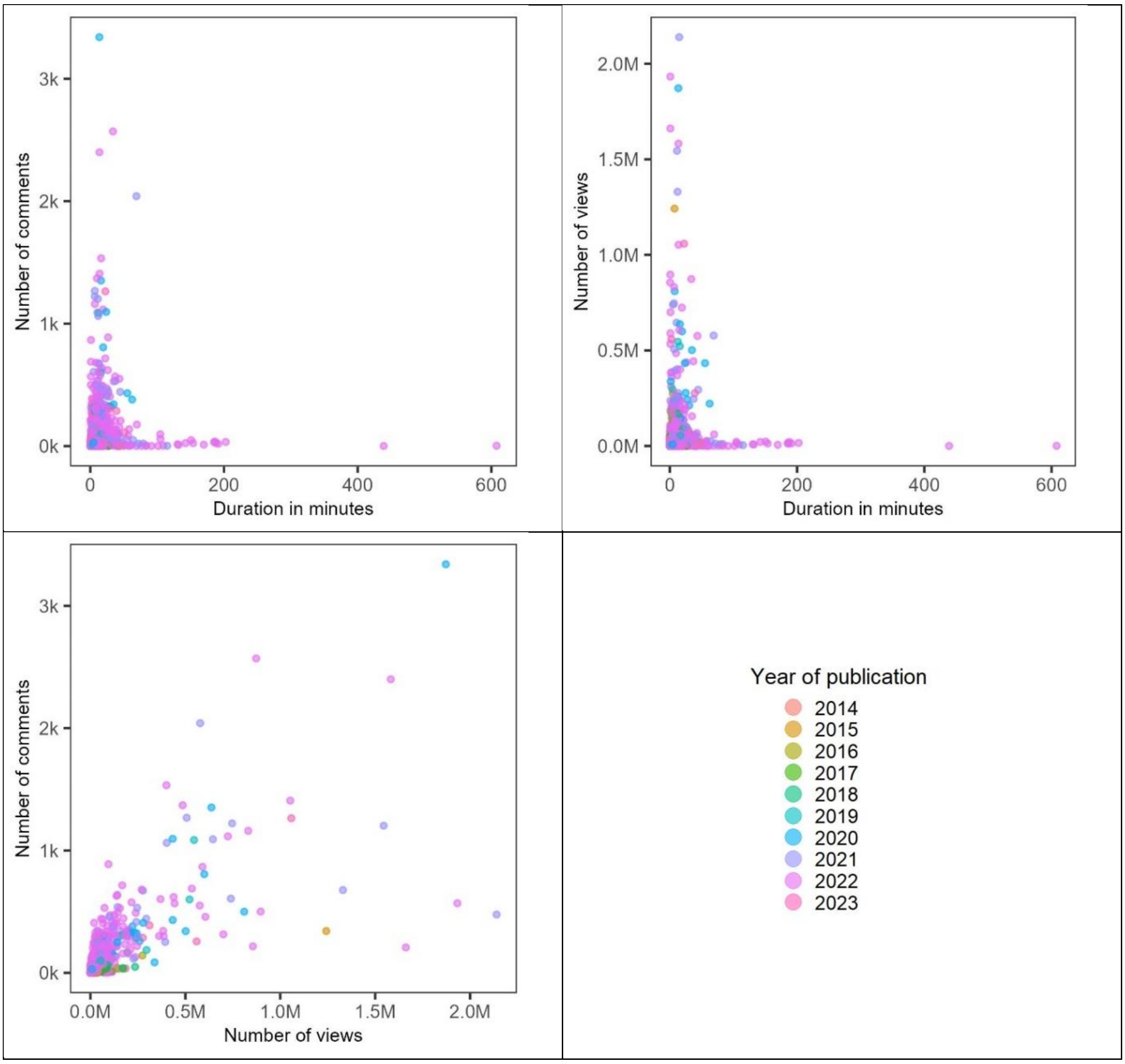

Figure 7. Relationship between duration, number of comments and views of films.

Sources: original research.

Figure 8 shows the percentage of positive, negative and neutral comments by year. We can see from it that in the year 2021 17.7% of comments had negative, 42.9% neutral and 39.4% positive sentiment. Starting in 2019, the number of positive comments decreased, while the number of neutral comments increased. The count of negative comments from 2017 to 2023 ranged between 17.7 and 21.3 per cent.

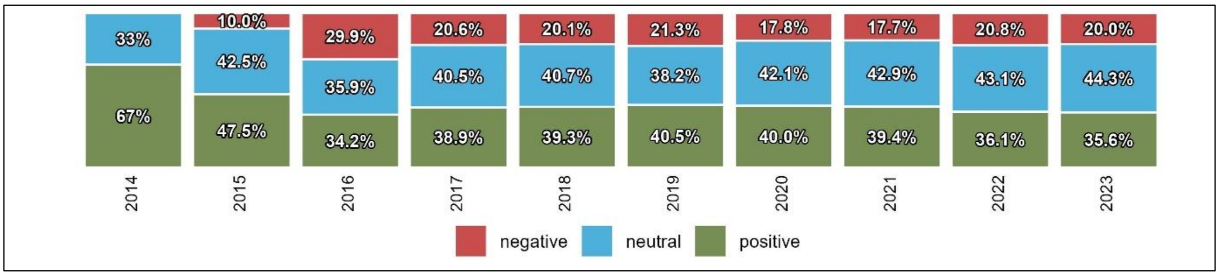

Figure 8. Percentage of positive, negative, and neutral comments.

Sources: original research.

Figure 9 presents the distribution of positive, negative, and neutral comments received by individual videos. The data is segmented by years. The distribution of comments is illustrated using box plots. Individual data points at a value of 100% for negative and neutral comments (received by videos for the year 2017) suggest that among these videos, some only received negative or neutral comments. The third quartile at approximately 25% for negative comments in 2022 indicates that 75% of the videos had a maximum of 25% negative comments.

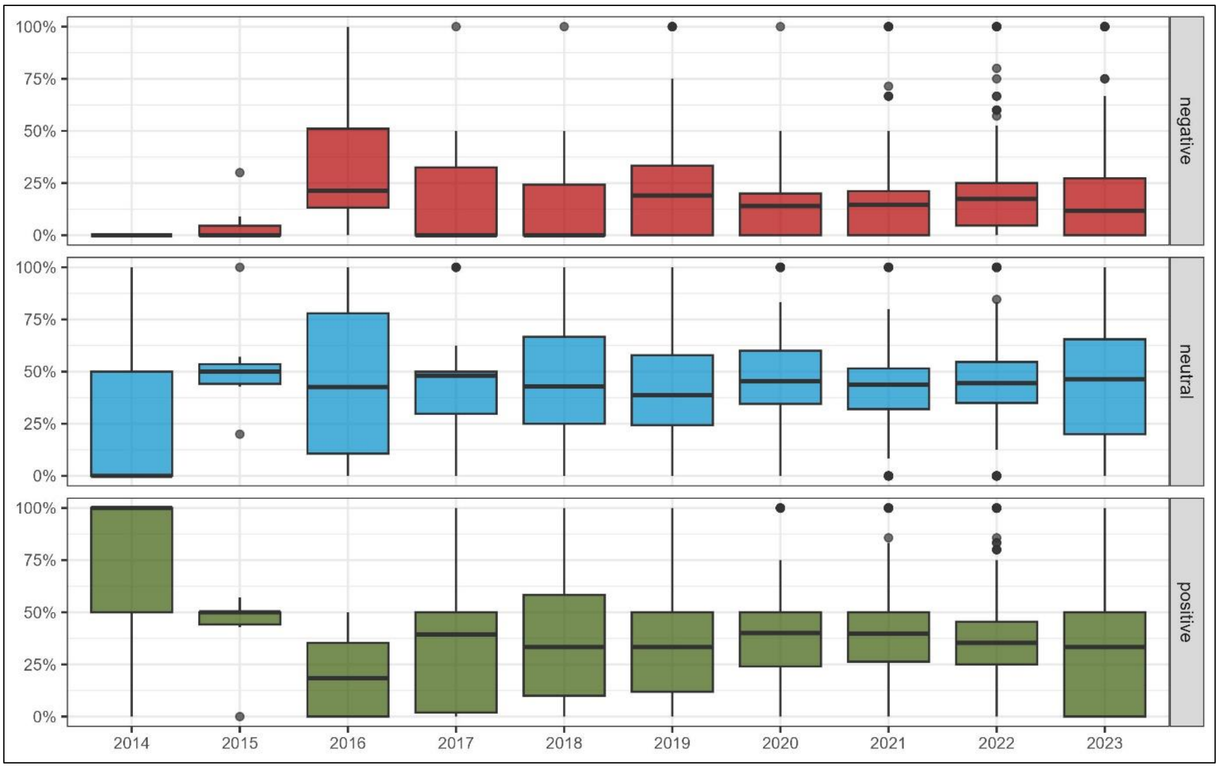

Figure 9. Distribution of the percentage of comments received by individual videos.

Sources: original research.

In Figure10, the annual distribution of words with annotated basic emotions is depicted. Positive emotions are denoted by the colour green, representing emotions like gladness, enjoying something expected, and trust. Negative emotions are indicated in red, representing emotions like disgust, fear, anger, surprise with something unpredictable, and sadness. It is evident that words associated with positive emotions had the following proportions in 2020: gladness accounted for 26.0%, enjoyment of something expected was 8.8%, and trust was 13.8%.

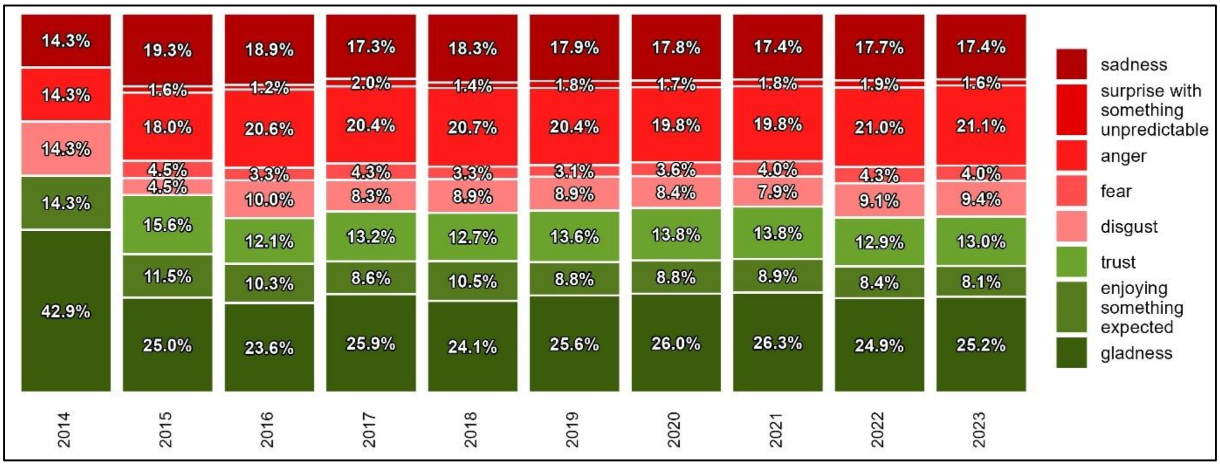

Figure 10. Percentage of words with annotated basic emotions by year.

Sources: original research.

Figure 11 presents the distribution of words with annotated fundamental human values. Positive human values are represented by the colour green (beauty, happiness, good of another man, utility, knowledge). Among the positive human values, not once did the "truth" occur. The negative human values are marked in red (unhappiness, error, harm, ignorance, uselessness, ugliness).

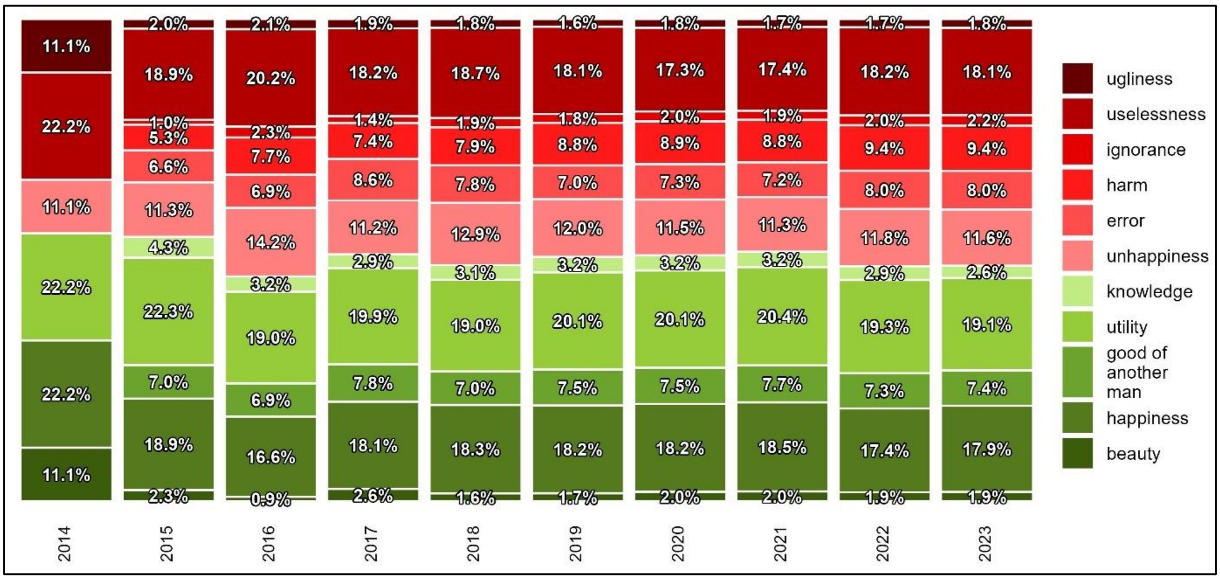

Figure 11. Percentage of words with annotated fundamental human values by year.

Sources: original research.

Figure 12 shows the most frequent words in comments. They are presented in the form of a word cloud. By analysing these words, it is possible to determine what the comments were about. The more frequently a word occurred in the comments, the bigger it is in the cloud.

Figure 12. Most frequently used words.

Sources: original research.

4. Conclusion

The analysis of data concerning videos related to photovoltaics (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 photovoltaics is continuously growing, particularly after the year 2018. Only 12 users (0.79 per cent) published between 17 and 112 videos;
- 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;
- in recent years, the number of comments with a neutral sentiment has been increasing, while the number of comments with a positive sentiment has been decreasing;
- 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;
- in 2022 and 2023, there was a slight increase in the percentage of words with annotated negative basic emotions and with annotated negative fundamental human values;
- analysing the most frequently used words, it can be assumed that comments addressed the following issues related to photovoltaics:
  - households as one of the main user groups of photovoltaics – words: “dom” (eng. house), “domowy” (eng. domestic), “budynek” (eng. building), “prosument” (eng. prosumer), “własny” (eng. own),
  - the main components of a photovoltaic installation – words: “słoneczny” (eng. solar), “ogniwo” (eng. cell), “panel”, “falownik” (eng. inverter), “moduł” (eng. module), “akumulator” (eng. battery), “bateria” (eng. battery), “magazyn” (eng. storage),
  - amount of energy produced by the photovoltaic installation over a given period – words and abbreviations like: “kwh” (kWh, eng. Kilowatt-hour), “kilowatogdzina” (eng. kilowatt-hour), “produkować” (eng. to produce), “wyprodukować” (eng. to produce), “produkcja” (eng. production), prąd (eng. “electricity”), “energia (eng. energy), “elektryczny” (eng. electric), “rocznie” (eng. annually), “roczny” (eng. annual), “rok” (eng. year), “miesiąc” (eng. month), “wynik” (“result”),
  - photovoltaic installation capacity and the factors affecting it - words: “moc” (eng. power), “kw” (eng. kW), “kilowatt” (eng. kilowatt), “k” (eng. kilo), “dach” (eng. roof), “kierunek” (eng. direction), “metr” (eng. metre), “powierzchnia” (eng. area),

- considering the purchase of an electric or plug-in hybrid car – words: “auto” (eng. car), “samochód” (eng. car), ładować (eng. to charge),
- complaints about intensive persuasion to buy photovoltaic – words: “dzwonić” (eng. to call), “telefon” (eng. phone), “sprzedać” (eng. to sell), “sprzedawać” (eng. to sell), “sprzedaż” (eng. sales),
- financial support for the purchase of photovoltaic – words: „dopłata” (eng. subsidy), “dotacja” (eng. subvention), “dostać” (eng. to get), “rząd” (eng. government), “przepis” (eng. law), “ustawa” (eng. law),
- the profitability of investment in photovoltaics – words: “koszt” (eng. cost), “kosztować” (eng. to cost), “opłacalny” (eng. worthwhile), “opłacać” (eng. to be worth), “zwrot” (eng. return on investment), “zwrócić” (eng. return on investment), “kredyt” (eng. credit),
- analysing the most frequently used words can determine that comments did not only concern photovoltaic but also in general electricity and heat production from various energy sources – “elektrownia” (eng. power plant), “eneriga” (eng. energy), “gaz” (eng. gas), “gazowy” (eng. gas), „prąd” (eng. electricity), “farma” (eng. farm), “węgiel” (eng. coal), “wiatrak” (eng. wind turbine), “wiatr” (eng. wind), “woda” (eng. water), “źródło” (eng. source), “grzać” (eng. to heat), “grzałka” (eng heater),“ciepło” (eng. heat), “pompa” (eng. pomp).

Analysis of data related to videos about photovoltaics on YouTube showed growing interest in this topic, especially after 2018, with an increase in the number of videos and comments. It is worth noting that videos with comments were viewed more often, and comments with neutral sentiment increased while those with positive sentiment decreased. Analysis of the most frequently used words shows a variety of topics, from household use to investment profitability. The conducted research confirms that comments on YouTube videos can be a source of data that can be used to understand people's thoughts, feelings, and opinions on "photovoltaics". It's worth noting that in this study, only the opinions of Polish-speaking YouTube users were identified.

References

1. Ağrali, Ö., Aydin, Ö. (2021). Tweet Classification and Sentiment Analysis on Metaverse Related Messages. *Journal of Metaverse*, *1*(1), 25–30.
2. Alves dos Santos, S.A., João, J.P., Carlos, C.A., Marques Lameirinhas, R.A. (2021). The impact of aging of solar cells on the performance of photovoltaic panels. *Energy Conversion and Management: X*, *10*, 100082. https://doi.org/10.1016/J.ECMX.2021.100082

3. 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
4. Bhuiyan, M.A., An, J., Mikhaylov, A., Moiseev, N., Danish, M.S.S. (2021). Renewable Energy Deployment and COVID-19 Measures for Sustainable Development. *Sustainability, Vol. 13*(8), 4418. https://doi.org/10.3390/SU13084418
5. Bilgili, F., Ozturk, I. (2015). Biomass energy and economic growth nexus in G7 countries: Evidence from dynamic panel data. *Renewable and Sustainable Energy Reviews*, *49*, 132–138. https://doi.org/10.1016/J.RSER.2015.04.098
6. Bórawski, P., Yashchenko, T., Sviderskyi, A., Dunn, J.W. (2019). Development of renewable energy market in the EU with particular regard to solar energy. *Conference Proceedings. Determinants Of Regional Development*, *1*, 12–13. https://doi.org/10.14595/CP/01/003
7. Marques Lameirinhas, R.A., João, J.P., Fernandes, C.A.F. (2021). Comparative study of the copper indium gallium selenide (CIGS) solar cell with other solar technologies. *Sustainable Energy & Fuels*, *5*(8), 2273–2283. https://doi.org/10.1039/D0SE01717E
8. Castilho, C.D.S., Torres, J.P.N., Fernandes, C.A.F., Lameirinhas, R.A.M. (2021). Study on the Implementation of a Solar Photovoltaic System with Self-Consumption in an Educational Building. *Energies, Vol. 14*(8), 2214. https://doi.org/10.3390/EN14082214
9. Chomać-Pierzecka, E., Kokiel, A., Rogozińska-Mitrut, J., Sobczak, A., Soboń, D., Stasiak, J. (2022). Analysis and Evaluation of the Photovoltaic Market in Poland and the Baltic States. *Energies, Vol. 15*(2), 669. https://doi.org/10.3390/EN15020669
10. *CLARIN-PL* (n.d.). Retrieved from: http://clarin-pl.eu/, June 5, 2022
11. 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
12. 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
13. 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
14. 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.
15. 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.
16. Dincer, I. (2000). Renewable energy and sustainable development: a crucial review. *Renewable and Sustainable Energy Reviews*, *4*(2), 157–175.

https://econpapers.repec.org/RePEc:eee:rensus:v:4:y:2000:i:2:p:157-175
17. Eroğlu, H. (2021). Effects of Covid-19 outbreak on environment and renewable energy sector. *Environment, Development and Sustainability, 23*(4), 4782–4790. https://doi.org/10.1007/S10668-020-00837-4/FIGURES/5
18. 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
19. Eyl-Mazzega, M.A., Mathieu, C. (2020). The European Union and the energy transition. *Lecture Notes in Energy, 73,* 27–46. https://doi.org/10.1007/978-3-030-39066-2_2/FIGURES/1
20. Grębosz-Krawczyk, M., Zakrzewska-Bielawska, A., Glinka, B., Glińska-Neweś, A. (2021). Why Do Consumers Choose Photovoltaic Panels? Identification of the Factors Influencing Consumers' Choice Behavior regarding Photovoltaic Panel Installations. *Energies, Vol. 14*(9), 2674. https://doi.org/10.3390/EN14092674
21. 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
22. Hamilton, L.C., Hartter, J., Bell, E. (2019). Generation gaps in US public opinion on renewable energy and climate change. *PLOS ONE, 14*(7), e0217608. https://doi.org/10.1371/JOURNAL.PONE.0217608
23. Ibar-Alonso, R., Quiroga-García, R., Arenas-Parra, M. (2022). Opinion Mining of Green Energy Sentiment: A Russia-Ukraine Conflict Analysis. *Mathematics, Vol. 10*(14), 2532. https://doi.org/10.3390/MATH10142532
24. 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
25. 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
26. 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.
27. 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, Vol. 13*(5), 2673. https://doi.org/10.3390/SU13052673
28. Lee, J. (2022). A Governance Structure Based on an Opinion Analysis of Local Stakeholders of Saemangeum Floating Photovoltaic Power Plants Project: Using Text

Mining for Each Subject. *Journal of People, Plants, and Environment*, *25*(6), 595–606. https://doi.org/10.11628/KSPPE.2022.25.6.595

29. 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
30. Moriarty, P., Honnery, D. (2011). What is the global potential for renewable energy? *Renewable and Sustainable Energy Reviews*, *16*, 244–252. https://doi.org/10.1016/j.rser.2011.07.151
31. Mota, F., Neto Torres, J.P., Ferreira Fernandes, C.A., Marques Lameirinhas, R.A. (2020). Influence of an aluminium concentrator corrosion on the output characteristic of a photovoltaic system. *Scientific Reports*, *10*(1), 1–16. https://doi.org/10.1038/s41598-020-78548-z
32. 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
33. Noblet, C.L., Teisl, M.F., Evans, K., Anderson, M.W., McCoy, S., Cervone, E. (2015). Public preferences for investments in renewable energy production and energy efficiency. *Energy Policy, 87*, 177–186. https://doi.org/10.1016/J.ENPOL.2015.09.003
34. Omnicore (2021). YouTube by the Numbers: Stats, Demographics & Fun Facts. *Omnicoreagency.Com*. https://www.omnicoreagency.com/youtube-statistics/
35. Omri, A., Daly, S., Nguyen, D.K. (2015). A robust analysis of the relationship between renewable energy consumption and its main drivers. *Http://Dx.Doi.Org/10.1080/00036846.2015.1011312*, *47*(28), 2913–2923. https://doi.org/10.1080/00036846.2015.1011312
36. Pang, B., Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. *ArXiv Preprint Cs/0409058*.
37. Pang, B., Lee, L. (2008). *Opinion mining and sentiment analysis. Found Trends Inf Retr, 2(1-2)*, 1-135.
38. 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
39. Pellerin-Carlin, T., Vinois, J.-A., Rubio, E., Fernandes, S., Delors, J., Letta, E. (n.d.). *Making the energy transition a european success tackling the democratic, innovation, financing and social challenges of the energy union*.
40. Peñaloza, D., Mata, É., Fransson, N., Fridén, H., Samperio, Á., Quijano, A., Cuneo, A. (2022). Social and market acceptance of photovoltaic panels and heat pumps in Europe:

A literature review and survey. *Renewable and Sustainable Energy Reviews*, *155*, 111867. https://doi.org/10.1016/J.RSER.2021.111867

41. Peng, J., Lu, L., Yang, H. (2013). Review on life cycle assessment of energy payback and greenhouse gas emission of solar photovoltaic systems. *Renewable and Sustainable Energy Reviews*, *19*, 255–274. https://doi.org/10.1016/J.RSER.2012.11.035
42. Pestana, D.G., Rodrigues, S., Morgado-Dias, F. (2018). Environmental and economic analysis of solar systems in Madeira, Portugal. *Utilities Policy*, *55*, 31–40. https://doi.org/10.1016/J.JUP.2018.09.001
43. Plutchik, R. (1980). Emotion: A Psychoevolutionary Synthesis. *The American Journal of Psychology*, *93*(4), 751. https://doi.org/10.2307/1422394
44. Puzynina, J. (1992). *Język wartości*. 264.
45. Qazi, A., Hussain, F., Rahim, N.A.B.D., Hardaker, G., Alghazzawi, D., Shaban, K., Haruna, K. (2019). Towards Sustainable Energy: A Systematic Review of Renewable Energy Sources, Technologies, and Public Opinions. *IEEE Access*, *7*, 63837–63851. https://doi.org/10.1109/ACCESS.2019.2906402
46. Quitzow, R., Bersalli, G., Eicke, L., Jahn, J., Lilliestam, J., Lira, F., Marian, A., Süsser, D., Thapar, S., Weko, S., Williams, S., Xue, B. (2021). The COVID-19 crisis deepens the gulf between leaders and laggards in the global energy transition. *Energy Research & Social Science*, *74*, 101981. https://doi.org/10.1016/J.ERSS.2021.101981
47. Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. *Proceedings of the ACL Student Research Workshop*, 43–48.
48. Salim, R.A., Rafiq, S. (2012). Why do some emerging economies proactively accelerate the adoption of renewable energy? *Energy Economics*, *34*(4), 1051–1057. https://doi.org/10.1016/j.eneco.2011.08.015
49. 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.
50. Stokes, L.C., Warshaw, C. (2017). Renewable energy policy design and framing influence public support in the United States. *Nature Energy 2017 2:8*, *2*(8), 1–6. https://doi.org/10.1038/nenergy.2017.107
51. Twersky, C. (n.d.). *scrapetube*. Retrieved from: https://pypi.org/project/scrapetube/, June 13, 2023.
52. Wierzbicka, A. (1992a). *Semantics, culture, and cognition : universal human concepts in culture-specific configurations*. 487.
53. Wierzbicka, A. (1992b). Defining Emotion Concepts. *Cognitive Science*, *16*(4), 539–581. https://doi.org/10.1207/S15516709COG1604_4
54. 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
55. *youtube-comment-downloader* (n.d.). Retrieved from: https://pypi.org/project/youtube-

comment-downloader/, June 13, 2023.

56. 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
57. Zator, S., Lambert-Torres, G. (2021). Power Scheduling Scheme for DSM in Smart Homes with Photovoltaic and Energy Storage. *Energies*, *Vol. 14*(24), 8571. https://doi.org/10.3390/EN14248571
58. Zrównoważonego, M.I.-K., Energetycznego, R., Energetyki, W., Paliw, A., Mirowski, T., Sornek, K. (2015). Potential of prosumer power engineering in Poland by exampleof micro PV installation in private construction. *Polityka Energetyczna – Energy Policy Journal*, *18*(2), 73–84. https://epj.min-pan.krakow.pl/Potential-of-prosumer-power-engineering-in-Poland-by-example-nof-micro-PV-installation,96084,0,2.html