Business Excellence and Management
Volume 14 Issue 1 / March 2024

USING SOCIAL MEDIA ANALYTICS AND PARSING PROGRAMS TO IDENTIFY AND ANALYZE THE TARGET AUDIENCE

Vasily KASHKIN
American University of Phnom Penh, Cambodia
vasily@kashkin.com.cn

Yuriy ANDROSIK
Belarusian State Technological University, Minsk, Belarus
androsik@belstu.by

Abstract
Social media analytics is the backbone of successful marketing. It allows you to accurately find and understand the target audience of the company. However, this requires analyzing a huge amount of publicly available data. This is where software tools come to the rescue, allowing such analysis to be carried out in real time. Nevertheless, there are a large number of software products on the market, each of which has a different set of functions - from complex products to highly specialized ones. A wide or vice versa narrow range of functions makes it difficult to choose a specific software solution for a specific task. For this reason, the main purpose of the article is to make a comparative analysis of the functions of the most popular social media analytics software tools designed to identify and analyze the target audience.
Keywords: Social network, Social media, Social media analytics, Social media analytics software tools, Text analytics, Image analytics, Audio analytics, Video analytics, Target audience.
DOI: https://doi.org/10.24818/beman/2024.14.1-04

1. INTRODUCTION

In the modern world, social networks are becoming increasingly popular and, in fact, are the driving force for the dissemination of information in various fields, including business. From an economic point of view, the reason for such popularity lies in the low costs of creating, receiving and disseminating information, and everywhere, that is, with a fairly wide coverage of an audience that goes beyond any one country.

2. LITERATURE REVIEW

In international literature, there is still no single and clear definition of social networks, with which most scientists and practitioners would agree. Basically, the term “social media” is used to describe various online channels and online platforms where various kinds of information are created and shared. This may

include forums, bulletin boards, review sites, opinion exchange sites, blogs and microblogs, sites that distribute media content (Ruhi, 2014).

But it is necessary to understand that the term "social network" is based on two concepts. The first is that there should be activity in a social network, or rather, activity streams, in simple words, a set of certain actions (activity streams). Such flows include communication, content creation and sharing, user collaboration and interaction. They reflect precisely the social aspect of the network. The second is that the totality of such actions are carried out in the context of connections and relationships between people. Such connections in structure are nothing more than a graph of social connections. This idea is borrowed from the mathematical theory of graphs, reflecting the fact that people are the nodes of a graph, and links between nodes, or edges, indicate relationships between people. It is these two concepts that reflect the content and form of a social network and distinguish a social network from other forms of computer communication (Ruhi, 2014).

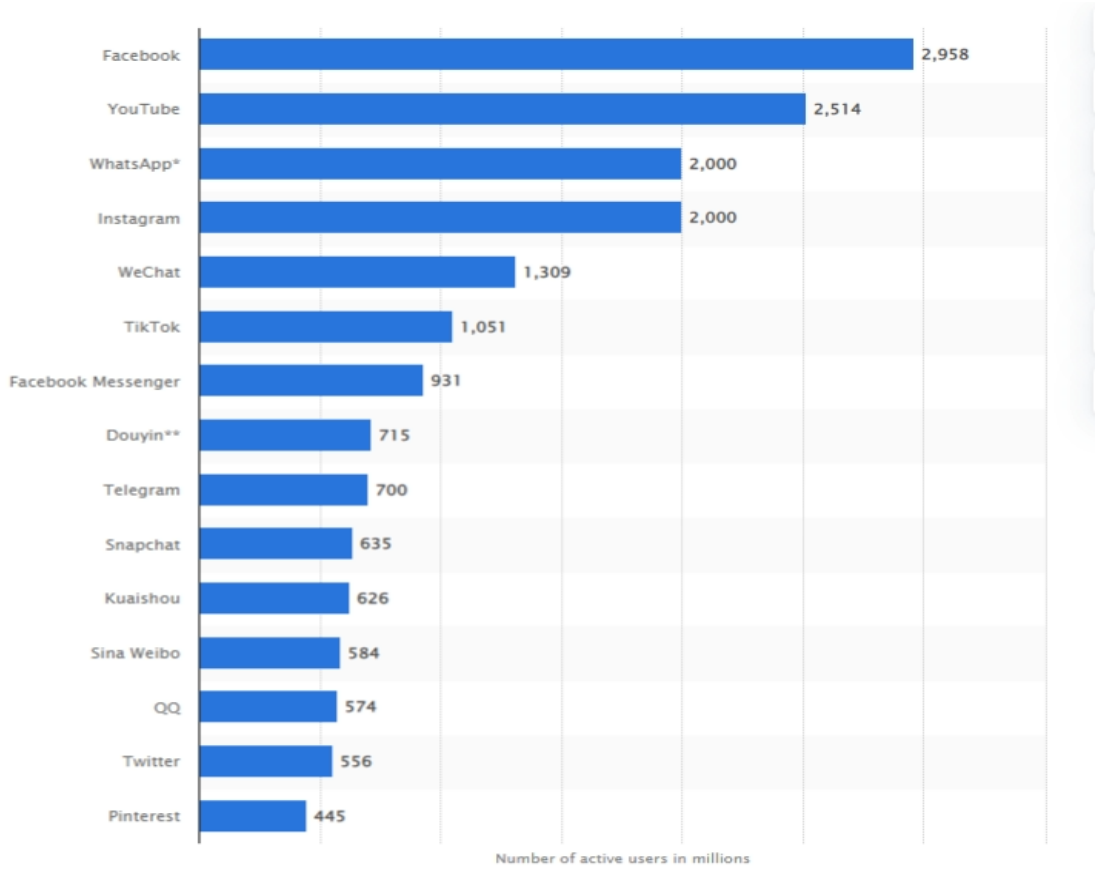

FIGURE 1. NUMBER OF ACTIVE USERS OF THE LEADING SOCIAL MEDIA, MLN
Source: statista.com

Businesses are more interested in networks with a wide coverage of active users. The top 5 social networks in the global context include Facebook, YouTube, WhatsApp, Instagram and WeChat. The order of other social networks and the number of active users can be seen in the diagram below (Statista, 2023). The main reason for the interest on the part of business is that the social network aggregates a huge amount of data about users, and therefore about potential customers, their behavior, interests, habits, consumption patterns, reactions. Such information makes it possible to distinguish one network user from another. Which in turn, especially from a marketing point of view, carries valuable information for identifying consumer segments and analyzing the target audience. Thus, the actions of people and their connectedness in social networks serve as the basis for developing models of interaction between business and consumer. At the heart of these developments is the analysis of users on the network and their digital footprints and activities.

Social networks accumulate digital footprints of various formats. This is mainly text data, including url addresses, images, video, audio, location data, social media and user metadata (Singh & Verma, 2022).

All of these types of information can be used by businesses to identify users and refine their profiles. Various methods of analysis are applied to these types of data. The study of these methods, their development and adaptation to social networks lies in the field called social media analytics (SMA).

It should be noted here that the term “social network analysis” (SNA) exists separately. From the point of view of the English language, the term “social media” is defined as forms of electronic communication (for example, websites for social networks and microblogging), with through which users create online communities for the exchange of information, ideas, personal messages and other content (for example, videos) (Merriam-Webster.com Dictionary, 2023). That is, this definition includes the term “social networks”. However, in academic literature, the term “social network analysis” is defined as the study of social structures using graph theory, a branch of discrete mathematics, being a key method of sociology. And this method appeared at the beginning of the 20th century, long before the advent of the first computer, for the study of social groups (Freeman, 2004). Later, the method was borrowed, adapted and SNA became one of the research methods in the field of SMA. Thus, if in the future we talk about the analysis of social networks, then we should understand here SMA, not SNA.

In a broader sense, social media (network) analytics is understood as the activity associated with the development and evaluation of information tools and environments for collecting, monitoring, analyzing, summarizing and visualizing social network data, usually due to the specific requirements of the target application (Zeng et al., 2010). In a narrow sense, as a rule, social network analytics is understood directly as the process of analyzing social network data.

Social media analytics involves the passage of 4 stages: data discovery, data collection, data preparation and analysis (Stieglitz et al., 2018). In modern academic literature, this is called the Social Media Analytics

Methodological Framework or SMA MF. At its core, it is a scheme by which it is necessary to analyze social network data. This approach is the result of a refinement of the social network data analysis framework, called CUP SMA MF (Fan & Gordon, 2014). In this case, C - caption (capturing data), U - understand (understanding data), P - present (providing data). In the case of automated data analysis, we are talking about the implementation of these stages in the form of a software product.

Capture involves the collection of information from various sources, data preprocessing, and extraction of relevant data. Understanding data involves applying various analytical methods and techniques to the captured data in order to subsequently draw conclusions relevant to the task. Presentation of data is the final procedure during which the results of the analytical work done are summarized and evaluated, conclusions are drawn up.

From a marketing point of view, the most important stage of social media analytics is the stage of understanding, that is, understanding the data through the methods and techniques of analysis. At the initial stage of analytical work, it is the knowledge of the main methods of data processing that allows us to form expectations regarding possible results. In other words, knowing the method and methodology of analysis, we understand which of the spectrum of possible results will be obtained. In the same way, knowing what method or technique we need, we determine the function of the program, which we will use later.

The situation is more complicated when, starting from a business task, we are trying to determine which functions should be used. Here the problem appears at the very first stage of the analysis - detection. This means that the complexity lies in the identification of relevant information and its subsequent selection from the entire data array. The problem is compounded by the fact that there has already been a transition from the “simple data” of social networks to the “big data” of social networks (Abkenar et al., 2021). This means that all the information that is generated within social networks is big data. Therefore, it is impossible to analyze the entire array of available data, since almost all companies have limited computing power.

In a full-fledged software, modules should be implemented that are responsible for detecting, collecting, preparing, analyzing data and presenting the results of the analysis. There are software solutions that contain some specific functionality. For example, the processes of collecting and structuring data (optionally partial pre-processing of information) are usually called parsing. The parser program collects information from web pages and stores it in some structured form.

There are software products that analyze individual types of content. Since there are 4 universal data types that are fundamentally different from each other, the corresponding types of analytics are distinguished (Rahman & Reza, 2022):

Business Excellence and Management
Volume 14 Issue 1 / March 2024

1. Text analytics
2. Image analytics
3. Audio analytics
4. Video analytics

Many researchers are used to working with numbers. However, they should be categorized as text analytics because the numbers in this case are text (but in numeric format). Although purely technically any digitized content is reduced to a number.

A separate software product can be developed for each type of content. This situation is due to the fact that different content requires the implementation of different methods, algorithms and analysis techniques. According to scientific literature research, the most popular are sentiment analysis (54.54%), predictive analysis (9.09%), theme modeling (9.09%), cluster analysis (6.81%), natural language processing (6.81%), network analysis (6.81%), text analysis (4.54%), trend analysis (2.27%), spatial analysis (2.27%), comparative analysis (2.27%), event detection (2.27%) (Madila et al., 2021). As you can see, most of the popular algorithms relate to text analytics. It can be assumed that this is due to the fact that the text carries the largest semantic load in relation to the size that it occupies during storage. Therefore, text analysis takes less time and resources in general, including computational ones. For example, a text file contains a much larger semantic load than a video file or an audio file of comparable size.

This is how we designate a view of social network analytics from the side of an analyst - a user of a software product who collects and analyzes data. However, social media analytics can be perceived not only as a concept of data collection and analysis, focusing on online discourses, but also as a concept of searching for emerging trends (Geissinger et al., 2023). At the same time, effective analytics is always associated with the company's business metrics, so the functionality of software products should be as close as possible to the set of such metrics. This is how we designate a view of social network analytics from the side of the end customer, which is business.

3. MATERIALS AND METHODS

At the first stage we studied 106 social media analytics and parsing programs. At the second stage we have selected only those programs, that actually perform the stated functions and are most well-known in the market. These 30 social media analytics and parsing programs were used for detailed research.

The list of analyzed programs includes such popular products as Buffer, Sprout Social, Khoros Marketing, Brandwatch, Audiense, Talkwalker, Synthesio Social Intelligence Suite, Comscore Media Metrix Multi-Platform, Equifax Data-Driven Marketing (Datalicious OptimaHub), Pulsar CORE, Helixa, 1plusX, Zoomph, NOW Affinio, Gravy Audiences, Socialeads, MVPindex, Tribyl, Persona by Delve AI, Dstillery, KPI6,

Demographics Pro, People Pattern, Linkfluence Radarly (from Meltwater), YouGov Audience Intelligence, Appier AIXON, Tubular, SparkToro, GWI , Infegy Atlas. The full list of analyzed programs and their main functionality are presented in the Appendix 1, Table 2.

Most programs are standardized and applicable to a wide range of companies. They measure the main indicators of social networks, such as user reach, engagement, influence, sentiment, and activity. Some, on the contrary, are quite applied and highly specialized, which cannot be used in any company. Therefore, there are difficulties in choosing suitable software, between its functions and business objectives. And ultimately, between finding plausible links between user activity or a company's social media initiative and the company's sales and profits.

Our purpose was to analyze the most popular software products for social media analytics and text analysis, including those that can find and analyze the target audience. We have aimed to determine which functions of these programs help to identify and analyze target audience.

From the standpoint of their functionality, they can only be aimed at analyzing the target audience and related aspects such as demographic and psychographic characteristics of the audience, which in the literature is called Target Audience Analytics (target audience analytics). Or they can have a wider and generally complex functionality, where the analysis of the target audience is only part of the program.

It should be noted that most programs work on the basis of machine learning models and artificial intelligence technologies.

4. RESULTS AND DISCUSSION

Within the framework of these products, 4 enlarged functions can be distinguished - this is the management of publications / content, interaction with users, monitoring and listening to social networks and analytics. First of all, we are interested in the analytical capabilities of these products, especially in relation to the target audience. They are contained in the last two enlarged functions.

In the course of testing the programs, we identified 27 functions. Typical functions include the following 24 functions:

own channel analysis	hashtag analysis	demographic analysis	trend analysis
user portrait analysis	competitor analysis	psychographic analysis	influencer analysis
access to historical data	preferences and interests analysis	human personality analysis	audience analysis and segmentation
post analysis	campaign analysis	sentiment analysis	data classification and clustering
conversation and topic	image recognition	reaction analysis	analysis of behavior

Business Excellence and Management
Volume 14 Issue 1 / March 2024

analysis			
automated collection of customer data	video analysis	brand mention analysis	location analysis

Such functions as audio analysis, automated market analysis, analysis of indicators affecting business goals are not common. Perhaps this unpopularity is due solely to the choice of programs. On the other hand, it is possible that audio, market, and metrics analysis is more difficult to use for target audience analysis.

The frequency of functions is shown in the diagram below.

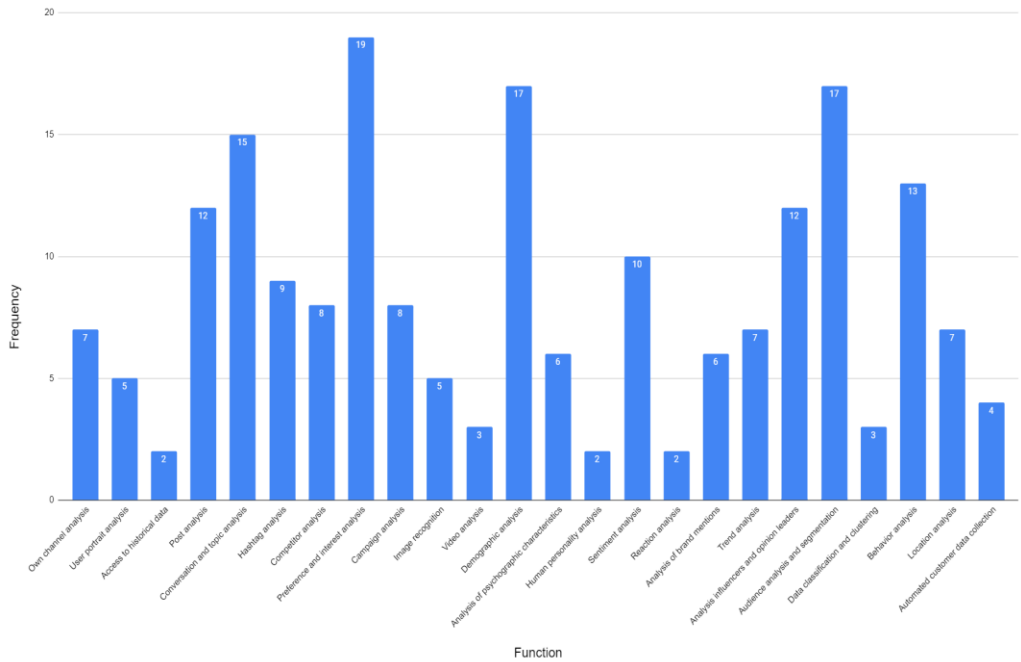

FIGURE 2. FREQUENCY OF FUNCTIONS
Source: Developed by the authors

The most popular features are preference and interest analysis, demographic analysis, audience analysis and segmentation, conversation and topic analysis, behavioral analysis. They are contained in the functions on average of every second program. And in general, they allow you to conduct a primary analysis of the target audience. The functions of analyzing psychographic and personality characteristics, analyzing the location and portrait of the user have a relatively small number of programs. These functions allow for a more detailed analysis of the target audience and already with reference to a specific territory. As a rule, specialized programs have such functions.

Separately, we note that part of the functionality can be replaced by the ChatGPT neural network. It can be used for text and sentiment analysis. However, at the moment, the neural network cannot make a full analysis of the target audience and its segmentation.

5. CONCLUSION

Social networks are a valuable source of information. They store a huge array of data, which, with appropriate software processing, allows you to form not only a portrait of the target audience, but also a portrait of a particular social network user. Different programs have different functions.

Every second program in our set (see Appendix 1) is suitable for the primary analysis of the target audience. The main features it should include are preference and interest analysis, conversation and topic analysis, and behavioral analysis. For an in-depth study of the target audience, it is necessary to select a highly specialized program that has the functions of analyzing psychographic and personal characteristics, analyzing the user's portrait in conjunction with his location. In the long term, a significant function is access to historical data, as well as exporting data to your own database in the appropriate format.

REFERENCES

Abkenar, S.B., Kashani, M.H., Mahdipour, E. & Jameii, S.M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. Telematics and Informatics, 57: 101-517.

Fan, W. & Gordon, M. D. (2014). The power of social media analytics. Communications of the ACM, 57(6): 74-81.

Freeman, L.C. (2004). The development of social network analysis. A study in the sociology of science. ΣP Empirical Press, Vancouver, BC Canada. 217 pages.

Geissinger, A., Laurell, C., Öberg, C. & Sandström, C. (2023). Social media analytics for innovation management research: A systematic literature review and future research agenda, Technovation, 123: 10271.

Madila, S., Dida, M. & Kaijage, S. (2021). A review of usage and applications of social media analytics. Journal of Information Systems Engineering and Management, 6(3), em0141. https://doi.org/10.21601/jisem/10958.

Merriam-Webster.com Dictionary (2023). Social Media. Merriam-Webster, Retrieved 2023 from https://www.merriam-webster.com/dictionary/social%20media.

Rahman, M.S. & Reza, H. (2022). A systematic review towards big data analytics in social media. Big Data Mining and Analytics, 5(3): 228-244.

Ruhi, U. (2014). Social media analytics as a business intelligence practice: current landscape & future prospects, Journal of Internet Social Networking & Virtual Communities, Vol. 2014 (2014), Article ID 920553, DOI: 10.5171/2014.920553.

Singh, R.K. & Verma, H.K. (2022). Effective parallel processing social media analytics framework, Journal of King Saud University - Computer and Information Sciences, 34(6, Part A): 2860-2870.

Statista (2023). Global social networks ranked by number of users. Retrieved January 2023 from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

Stieglitz, S., Mirbabaie, M., Ross, B. & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. International Journal of Information Management, 39: 156-168.

Zeng, D., Chen, H., Lusch, R. & Li, S.-H. (2010). Social media analytics and intelligence. IEEE Intelligent Systems, 25(6): 13-16.

APPENDIX 1

TABLE 1. ANALYZED PROGRAMS AND THEIR FUNCTIONS. PART 1

Program	Buffer	Sprout Social	Khoros	Brandwatch	Audiense	Talkwalker	Synthesio	Comscore Media Metrix	Equifax	Pulsar CORE	Helixa	1plusX	Zoomph	NOW Affinio	Gravy Audiences
Content management	1	1	1	1											
User interaction	1	1	1	1					1	1					
Listening to social networks		1	1	1		1	1			1					
Publication of posts	1	1													
Profile/account performance analysis	1	1	1	1		1	1			1					
Own channel analysis	1	1		1		1	1			1			1		
User portrait analysis				1	1										
Access to historical data	1			1											
Post analysis	1			1	1								1	1	
Conversation and topic analysis	1	1		1		1				1	1				
Hashtag analysis	1	1		1		1	1					1			
Competitor analysis		1	1	1		1	1							1	
Preference and interest analysis		1			1			1		1	1		1	1	1
Campaign analysis		1	1	1			1			1		1			
Image recognition						1							1		
Video analysis						1							1		
Demographic analysis	1	1			1			1		1	1				
Analysis of psychographic characteristics					1						1				
Human personality analysis					1						1				
Sentiment analysis			1	1		1	1			1			1		
Reaction analysis		1													
Analysis of brand mentions				1		1	1			1					
Trend analysis			1			1				1				1	
Analysis influencers and opinion leaders		1	1		1	1	1				1			1	

Audience analysis and segmentation			1	1	1		1	1	1	1	1	1		1	
Data classification and clustering						1						1		1	
Behavior analysis					1			1		1					1
Location analysis	1	1													1
Automated customer data collection			1	1		1			1						
Atypical function	Post generation, automated creation of landing pages	Building chatbots	Analysis of indicators that affect business goals			Audio analysis	Market analysis								

TABLE 2. LIST OF ANALYZED PROGRAMS. PART 2

Program	Socialeads	MVPindex	Tribyl	Persona	Dstillery	KPI6	Demographics Pro	People Pattern	Linkfluence Radarly	YouGov Audience Intelligence	Appier AIXON	Tubular	SparkToro	GWI	Infegy Atlas
Content management															
User interaction															
Listening to social networks															
Publication of posts															
Profile/account performance analysis	1		1												
Own channel analysis															
User portrait analysis					1			1		1					
Access to historical data															
Post analysis	1	1							1		1	1	1		1
Conversation and topic analysis	1		1	1		1			1		1	1	1		1
Hashtag analysis				1								1	1		
Competitor analysis								1							1
Preference and interest analysis	1		1	1		1		1		1	1	1	1	1	1
Campaign analysis		1													1
Image recognition						1			1		1				
Video analysis												1			
Demographic analysis		1		1	1	1	1	1		1		1	1	1	1
Analysis of psychographic characteristics				1			1	1		1					

Business Excellence and Management
Volume 14 Issue 1 / March 2024

Human personality analysis															
Sentiment analysis	1					1			1						1
Reaction analysis															1
Analysis of brand mentions									1			1			
Trend analysis									1			1			1
Analysis influencers and opinion leaders							1	1				1		1	1
Audience analysis and segmentation		1		1				1	1	1	1			1	
Data classification and clustering															
Behavior analysis		1	1	1	1			1		1	1	1	1		
Location analysis		1		1			1								1
Automated customer data collection															