Springer

3rd International Conference on Image Processing and Capsule Networks
(ICIPCN-2022)

May 20-21, 2022 | Bangkok, Thailand

ICIPCN
Proceedings

Technical Sponsors

45	Medchain for Securing Data in Decentralized Healthcare System using Dynamic Smart Contracts *R Priyadarshini, Mukil Alagirisamy, N Rajendran*
46	Multipurpose Linux tool for Wi-Fi based attack, Information gathering and Web vulnerability scanning automations *Ceronmani Sharmila, Gopalakrishnan J, Shanmuga Prasath P, Daniel Y*
47	Customer Engagement through Social Media and Big Data Pipeline *Rubeena Rustum, J Kavitha, PVRD Prasada Rao, Jajjara Bhargav, G Charles Babu*
48	Performance Analysis of CNN models using MR images of Pituitary Tumour *Ashwitha K*
49	Detection of Facebook Addiction Using Machine Learning *Md. Zahirul Islam, Ziniatul Jannat, Md. Tarek Habib, Md. Sadekur Rahman, Gazi Zahirul Islam*
50	Mask R-CNN based Object Detection in Overhead Transmission Line from UAV Images *D Satheeswari, Leninisha Shanmugam, N M Jothi Swaroopan*
51	Insights into Fundus Images to Identify Glaucoma using Convolutional Neural Network *Digvijay J Pawar, Yuvraj K Kanse, Suhas S Patil*
52	An Implementation Perspective on Electronic Invoice Presentment and Payments *B Barath Kumar, C N S Vinoth Kumar, Suguna R, M Vasim Babu, M Madhusudhan Reddy*
53	Multi-model DeepFake Detection using Deep and Temporal Features *Jerry John, Bismin V Sherif*
54	Real-Time Video Processing for Ship Detection using Transfer Learning *V Ganesh, Johnson Kolluri, Amith Reddy Maada, Mohammed Hamid Ali, Rakesh Thota, Shashidhar Nyalakonda*
55	Smart Shopping using Embedded based Autocart and Android app *V Sherlin Solomi, C Srujana Reddy, S Naga Tripura*
56	Gastric Cancer Diagnosis Using MIFNet Algorithm And Deep Learning Technique *Mawa Chouhan, Corinne Veril D, Prerana P, A Kumaresan*
57	Cold Chain Logistics Method using to Identify Optimal Path in Secured Network Model with Machine Learning *Vijaykumar Janga, Desalegn Awoke, Assefa Senbato Genale, B Barani Sundaram, Amit Pandey, P Karthika*
58	Brain Tumor Image Enhancement Using Blending of Contrast Enhancement Techniques *Deepa Abin, Sudeep Thepade, Yash Vibhute, Sphurti Pargaonkar, Vaishnavi Kolase, Priya Chougule*
59	A Systematic and Novel Ensemble Construction Method for Handling Data Stream Challenges *Rucha Chetan Samant, Suhas H Patil*
60	A Smart Garbage System for Smart Cities using Digital Image Processing *Sivaranjani P, Gowri P, Bharathi Mani Rajah Murugan, Ezhilarasan Suresh, Arun Janarthanan*

Customer Engagement through Social Media and Big Data Pipeline

Rubeena Rustum[1], J.Kavitha[2],
PVRD Prasada Rao[3], Jajjara Bhargav[4]
and G.Charles Babu[5]

[1,5]Gokaraju Rangaraju Institute of Engineering and Technology(Autonomous),
Bachupally, Hyderabad,Telangana
[2]BVRIT HYDERABAD College of Engineering for Women, Hyderabad,
Telangana
[3]Koneru Lakshmaiah Education Foundation, Vaddeswaram,Guntur, A.P
[4]Chalapathi Institute of Engineering and Technology (Autonomous),Lam,Guntur
rubeena.rustum@gmail.com

Abstract. Engagement of customers through social media has gained considerable popularity in recent years in the field of digital marketing. Especially with the rise of technological revolution in business operations, utilizing sophisticated technology for strategic development of businesses has been seen. In this regard, data pipeline can be considered as an efficient, automated and sophisticated technology that uses a systematic data management process for voluminous data. The paper thus aims to investigate the beneficial scope of aligning social media with data pipeline technology for enhancing customer engagement. Through an empirical analysis of existing secondary resources as a viable and beneficial method for research, the study has developed a comprehensive understanding of data pipelines and social media platforms contributing to the enhancement of customer engagement.

The main findings of the study indicates that through the ETL pipeline (Extraction- Transformation-Loading), large columns data is managed sequentially and swiftly. The automated system can be used with both flexibility and control to manage the data flow. The businesses are able to control the data flow to their advantage and increase visibility and interaction. Simultaneously the analytics process of the big data aids the decision-making process that ensures customer behavior and market demands are being considered accurately. The study also considers certain challenges such as lack of increased storage capacity, high volume data, and consistency on that can be addressed through further development of advanced architecture.

Keywords: data pipeline, ETL pipeline, Big Data management, Big Data analytics, social media, social media marketing, digital marketing, customer engagement.

1. Introduction

Social media platforms have been recognized as one of the most effective and engaging platforms through which various industries promote products and services or simply establish relationships with their customers. Recent inclination to utilize sophisticated technology for enhancing marketing policies among target demographic has been seen throughout various global industries. Organizations are seen to utilize various data analytics technologies for processing vast quantities of customer data and perform tasks such as online engagement rates, response to new and old products, sharing feedback and so on. These are only surface level operations that are aided through sophisticated data analytics technology such as data pipeline.

Data pipeline is an advanced data processing and analytics technology that is able to retrieve and process data and store it in a data warehouse. The three prime elements in data pipelines consist of a source, processing of data and destination [1]. The apparent simplicity of its process is far more critical in application as it is used by organizations to organize real-time data, process it and store it. Organizations are highly interested in processing large-scale digital information and analyze it for assessing market demands, leading to creation of personalized digital marketing policies. On a global scale, the expenditure for social media marketing is expected to increase 10.93% by 2025 [2]. Especially, with the increased rate of using various mobile applications, it has become a necessity to process and store data of each customer. Hence, the paper will address the issues of large-scale data processing in general and aim at investigating the beneficial use of data pipelines in increasing the rate of customer engagement through social media.
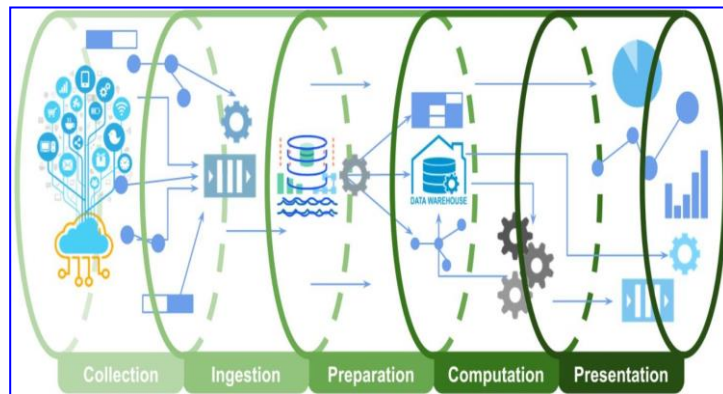

Figure 1: Big Data Pipeline Architecture [1]

1.1 Analysis of data pipeline technology

The overall usage of digital information gathered from various social media platforms have grown in variety, volume and velocity. Big data or the enormous volume of data gathered through social media interaction can be processed and stored through dada pipelines [3]. Such a sophisticated technology is able to support big data and has evolved to ensure that data processing is conducted in a clear and logical sequence. The interaction of customers throughout various social media platforms are processed and assessed through predictive analytics [4]. It leads to the enhancement of engagement through proper marketing. In other words, as the organizations monitor the rate of engagement, they are able to understand the direction of market demands and cater towards the personal desires of customers.
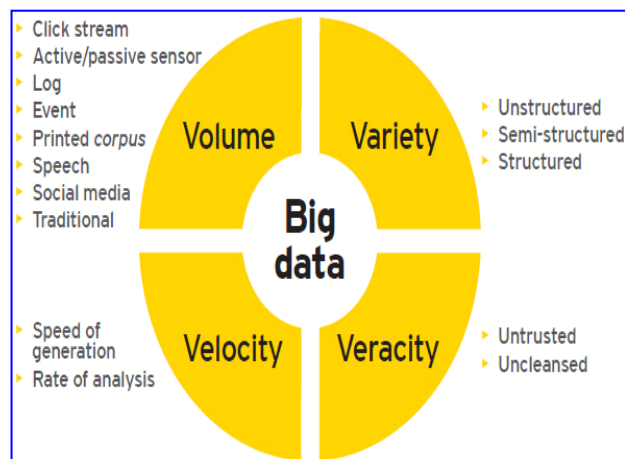
Figure 2: Three Vs of Big Data [3]

Extraction of big data from its current source location, processing or transforming it into a reliable format. A secured loading or storage process is conducted at the third stage of the data pipeline where the stored data is analyzed through sophisticated machine learning technology [5]. The most beneficial aspect of the data pipeline that is used by business is its ability to

extract data at any point of data processing [6]. Hence, the data pipeline technology can be used to provide organizations with the required data for analytics. It not only contributes to saving time of conducting data extraction from the source point alone but also enables companies to save the extracted data. Huge volume of data is then analyzed automatically in order to perceive the market demands and customer attitudes towards a certain product or service.

Data ingestion pipeline is used for the purpose of big data analysis. AI has been used inrecent times for the improvement of ingestion practices [7]. It also helps in avoiding redundant loading of the processed data that uses an automated system. In essence, data pipeline technologies are able to provide optimal scope for extracting and processing big datain a systematic manner.

1.2   Benefits of data pipeline for customer engagement

Data pipelines can be used to optimize the process of big data extraction, transformation and loading. Customer data can be extracted from various applications used by them [8]. As companies grow more and more aware about the behavior and demands of customers, they are able to employ policies for raising social media engagement through effective marketing. It also helps in providing the organizations with ample knowledge regarding their retail sales rates and post engagement, aiding business-related decision-making processes. Thisinformation is highly beneficial to the decision-making process as it assists in generating strategies for attaining substantial competitive advantage.

The elements of data pipeline technology help in managing the e-commerce segment of various businesses and increases customer engagement through the accurate data analytics. Using social media analytics tools for analyzing customer behavior helps in the decision- making process [9]. Customers express their feedback and attitudes towards a certain product or service through social media posting, sharing, commenting or liking. Such data is then analyzed to generate personalized, suggestive and attractive marketing policies to further enhance customer engagement. On the other hand, through social media platforms, various organizations enhance their interaction with individual customers [10]. It not only establishes a sustainable and loyal relationship among a brand's targeted customer base but it also ensures that through such direct interactions, market demands are being recognized. In essence, the big data extracted from social media platforms through data pipelines contributes to critical business operations that have significant influence on an organization's manufacturing or production policies and marketing policies. Thus, the use of data pipeline technology for big data processing leading to proper analytics systems helps organizations to engage

with their customers in a personalized manner that naturally leads to competitive advantage.

2 Methods and Materials

The paper has conducted an empirical analysis on the topic of social media and data pipeline optimization in customer engagement. Construction of systematic and sequential methodology for the conduction of research enhances scope for logical and valid outcomes [11]. Various secondary resources on the topic had been gathered to analyze in order to meet the aim of the paper. In this regard, a descriptive design to the analysis of gathered resources had been conducted. Descriptive research design helps in the development of a research that is based on the development of knowledge through analyzing existing data or information [12]. Furthermore, a deductive research design had been adopted in order to ensure in-depth analysis. The overall qualitative approach undertaken for this particular paper had helped it to generate new ideas and hypotheses based on multidimensional existing literature on the topic.

The materials used for the study had been gathered from a targeted and systematic search for materials in certain electronic databases such as ProQuest and Google Scholar. Peer-reviewed and published journals, newspaper articles and official reports had been gathered. A systematic sampling method had been adopted as well which consisted of inclusion- exclusion criteria. Setting inclusion-exclusion arteritis assists in discovering only those materials which are beneficial for the course of a given research [13]. The criteria set for this particular study pertaining to the selection of beneficial and informative journal articles had included, peer-reviewed articles, publications within the last 5 years, publications in English language and publications containing certain keywords such as data pipeline, big data, big data analytics, marketing, digital marketing, customer engagement and social media. Thus, adoption of these specific methods has aided the process of data collection and data analysis, leading to the findings of the research that is valid, relevant and evidence-based

3. Results

Advantages of data pipeline technology for enhancing social media engagement

There are various advantages that are attained by a business through a data pipeline. Businesses use data pipelines to enhance their capacity for predictive analysis and measure the rate of activities or engagement in social media. Social media is a thriving platform that allows individuals to interact freely with a brand [14]. Businesses use this opportunity to retrieve the data of users to get accurate information regarding the brand's ability to meet the key

performance indicators (KPI) [15]. The activities of a potential customer on social media are assessed to analyze their demands, timing of purchase, and inclination for purchasing within a price range and so on. In essence, as a key advantage of such data streaming and analysis helps businesses to perceive the buying behavior of customers and offer products and services accordingly.

As businesses continue to align their social media platforms with data pipelines, they are essentially in control of regulating the flow of data. This helps businesses to gain relevancy and thus further engagement. Social media is being used as a significant platform for promoting and interacting with customers [16]. The current demographic is more prone to using social media than engaging in any other forms of engagement platforms. With the use of data pipeline technology, a brand is able to manage the data in a flexible manner [17]. Such flexibility leads to the beneficial alignment of data to the customers, leading to enhancement in social media interaction among customers.

Figure 3: Social media engagement strategies [16]

Management of such vast amounts of data and using it to meet the specific needs of a business is an advantageous side of data pipeline technology. The ETL pipeline (Extraction- Transformation-Loading) helps in the formation of an effective network [18]. Utilization of Thai networks in social media aids the process of customer engagement.
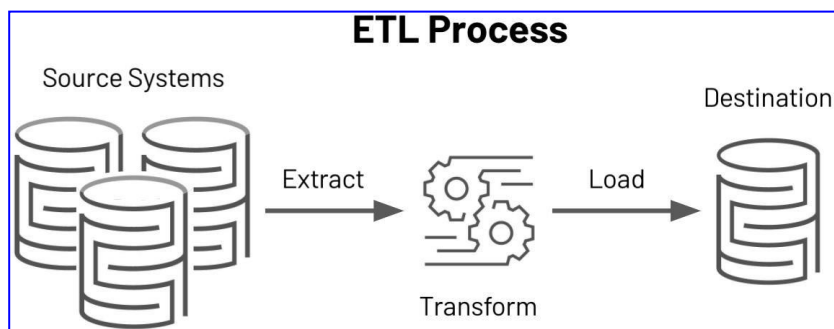
Figure 4: ETL pipeline process [18]

Challenges in big data analytics with data pipeline

There are several challenges that are faced in big data management without proper data pipeline architecture. The large volume of data remains at risk of being lost in the processing phase or may not reach the data warehouse [19]. Additionally, the major challenges identified in the field of big data analytics is proper management and integration process. In many cases, such processes are reliable in human understanding [20]. However, in recent times, development of AI-based systems for such operations has reduced this particular challenge. Regardless, lack of proper knowledge regarding the system limits the ability of a business to use it and optimize information to its own benefits.

There is a critical challenge that is faced by business in the utilization of data pipeline architecture as SMEs fail to integrate it due to financial lack. Both technological and financial infrastructure is required for the development and integration of sophisticated technology [21]. The technology itself may be faced with certain challenges as there is a possibility of a gap between real-time data entry and ETL process (Extraction-Transformation-Loading) [22]. The recent integration of AI-based systems in data processing and analytics has however decreased the time lag. Another challenge faced in recent years is the rising velocity, variety and most importantly volume of data. Each stream of data is to be processed through a data pipeline after which proper analytics can be conducted [23]. Hence, gap in processing capacity due lack of proper digital architecture leads to challenges in implementation as well as lack of consistency, security and so on.
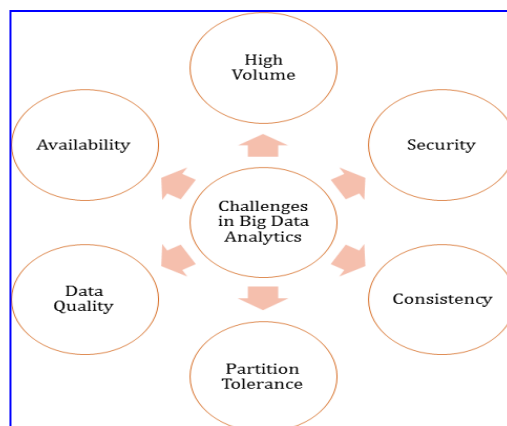
Figure 5: Challenges in Big Data analytics [23]

Optimization of data pipeline for enhancing social media engagement and organizational benefits

The organizations are able to benefit exponentially through data processing and analysis optimization that is aided by data pipeline technology. The organizations are able to gain profit and obtain substantial competitive advantage [24]. Enhancing customer engagement through social media is attained through analysis and predictive assessment [25]. As KPIs are also maintained with the use of data pipelines, it brings about significant opportunities for profitability.

On the other hand, the flexibility provided to the organizations as data pipelines are being used assists in their processes of active engagement. Social media analytics has been considered as an emerging tool for aiding the progress of business [26]. Data pipelines are also able to provide a clear view of the data flow based on which decision-making process is conducted. In essence, it can be stated that linkage of social media and data pipelines helps in increasing customer engagement and help businesses to formulate strategies for further marketing growth. The data flow of social media is tracked and analyzed in order to provide the organizations with ample insight into the needs of necessary modifications.

Social media marketing has been acknowledged as an effective segment for promotingproducts and services effectively. As the worldwide wealth of digital information rises, businesses use this opportunity to place their products or services in a strategic manner [27]. It essentially helps increase product visibility and ensure customers are effectively engaged with the brand through social media platforms. Hence, the most significant benefit that data pipelines are able to provide the organizations is the ability to analyze voluminous data and use it to retain relevance and interaction. Keeping track of the customer's opinions and latest trends are also analyze to strategically place product suggestions through various social media platforms, enhancing scope for sustained profitability.

4. Discussion

Analysis of the relevant literature has indicated the positive scope of using data pipeline in connection with social media to ensure a positive growth in customer interaction. Most optimal method of data processing and analysis has

been identified through the implementation of a data pipeline system [15]. As a data pipeline provides the optimal scope for arranging data in a sequential and synchronized format, it aids the development of a reproducible system in which businesses are able to gain insight into the gathered data anduse it to aid their decision-making process. Additionally, it has also been indicated that the automated ETL pipelines enable a timely analysis of valuable data.

The flexibility offered in data pipelines can be used for social media platforms to ensure that the data flow of raw data travels uninterruptedly to its destination. The point of data flow within a pipeline system ensures sequential flow which can be controlled by a business [22]. The main benefit of aligning data pipelines and social media is the enhancement in capacityto process large volumes of data, generated every second of the day. Being in control of the data processing procedures, it has been indicated that the development of an effective and interactive relationship between the brands and customers is possible by virtue of relevant data flow.

Data pipelines are also able to extract big data and distribute them in various relevant sets. Such a simplified structure for voluminous data helps in backing up data and redistributing in case of a crashed server [23]. Analysis of the data gathered from various social media platforms, a business is also able to assess the behavior and attitudes of customers. Gaining the trust of customers is an important part of effective marketing. However, customer engagement enhancement is not merely limited to the marketing procedures. It is extended towards a controlled flow of data to ensure that product visibility is optimal.
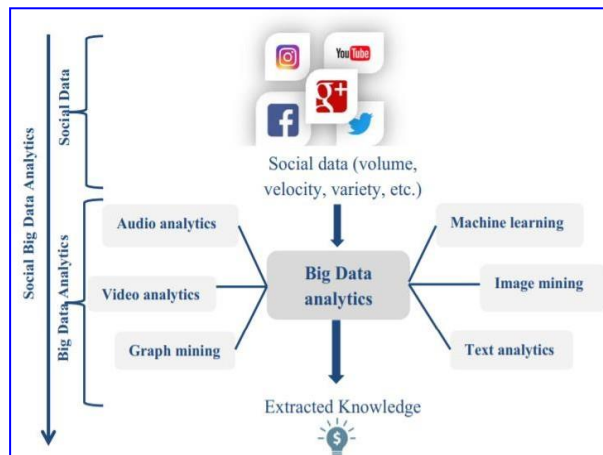
Figure 6: Big data analytics for social media [4]

It has also been gathered that certain challenges are faced by organizations in the process of data pipeline integration for management and processing. A comprehensive awareness of the entire system hinders the process of proper utilization [19]. Along with that, the lack of financial support has also been identified as one of the challenging factors pertaining to the implementation of sophisticated technology. Moreover, as the wealth of digital data on a global scale continues to increase, feasible options for adequate management, processing and analysis become highly dependable on the capacity of a system's data storage. Real-time data gathered from social media contributes to the generation of voluminous data every day. Hence, the overall capacity of the data pipeline to stream sequential data poses a challenge. However, despite these challenges, constant modification and innovative discoveries made in the field enhances its overall capacity and enables it to be used efficiently for managing big data and providing scope to the businesses to use it as a digital tool for increasing customer engagement.

5. Conclusion

Social media is used as a platform for active interaction between various businesses and customers. Increased rate of online interaction between a business and its target customers leads to the establishment of loyal and beneficial relationships. Increase in social media engagement is thus considered to be a significant part of increasing profitability and competitive advantage. Fulfillment of such business agenda may be possible with the use of data pipeline for social media engagement enhancement.

The data pipeline uses a systematic process of data extraction, transformation and loading (ETL) for sequential management of big data. With the growth of digital data, processing such voluminous data through a controlled yet flexible technology is required which is provided by a data pipeline. However, the processing and storage capacity is required to be in accordance with the growing volume, velocity and variety of data. It has been gathered that with the alignment of social media and data pipelines, businesses are able to control the data flow and direct the raw data to a beneficial destination that ultimately leads to the enhancement of customer engagement. In conclusion, it can be used as a potent tool for strategic placement, increased interaction and decision-making process.

References

1. del Rio Astorga, D., Dolz, M.F., Fernández, J. and García, J.D., 2017. A generic parallel pattern interface for stream and data processing. *Concurrency and Computation: Practice and Experience*, *29*(24), p.e4175.

2. Statista.com, 2021. *Social Media Advertising*. Available at: https://www.statista.com/outlook/dmo/digital-advertising/social-media-advertising/worldwide [Accessed on: 2nd November, 2021]

3. Prim, J., Uhlemann, T., Gumpfer, N., Gruen, D., Wegener, S., Krug, S., Hannig, J., Keller, T. and Guckert, M., 2021. A data-pipeline processing electrocardiogram recordings for use in artificial intelligence algorithms. *European Heart Journal*, *42*(Supplement_1), pp.ehab724-3041.

4. Sebei, H., Taieb, M.A.H. and Aouicha, M.B., 2018. Review of social media analytics process and big data pipeline. *Social Network Analysis and Mining*, *8*(1), pp.1-28.

5. Helu, M., Sprock, T., Hartenstine, D., Venketesh, R. and Sobel, W., 2020. Scalable data pipeline architecture to support the industrial internet of things. *CIRP Annals*, *69*(1), pp.385-388.

6. Therrien, J.D., Nicolaï, N. and Vanrolleghem, P.A., 2020. A critical review of the data pipeline: how wastewater system operation flows from data to intelligence. *Water Science and Technology*, *82*(12), pp.2613-2634.

7. Akanbi, A. and Masinde, M., 2020. A Distributed Stream Processing Middleware Framework for Real-Time Analysis of Heterogeneous Data on Big Data Platform: Case of Environmental Monitoring. *Sensors*, *20*(11), p.3166.

8. de Oliveira Santini, F., Ladeira, W.J., Pinto, D.C., Herter, M.M., Sampaio, C.H. and Babin, B.J., 2020. Customer engagement in social media: a framework and meta- analysis. *Journal of the Academy of Marketing Science*, *48*, pp.1211-1228.

9. Pääkkönen, P. and Jokitulppo, J., 2017. Quality management architecture for social media data. *Journal of Big Data*, *4*(1), pp.1-26.

10. Li, M.W., Teng, H.Y. and Chen, C.Y., 2020. Unlocking the customer engagement- brand loyalty relationship in tourism social media: The roles of brand attachment and customer trust. *Journal of Hospitality and Tourism Management*, *44*, pp.184-192.

11. Xanthopoulou, D., 2017. Capturing within-person changes in flow at work: Theoretical importance and research methodologies. In *Flow at Work* (pp. 50-65). Routledge.

12. Bloomfield, J. and Fisher, M.J., 2019. Quantitative research design. *Journal of the Australasian Rehabilitation Nurses Association*, *22*(2), pp.27-30.

13. Patino, C.M. and Ferreira, J.C., 2018. Inclusion and exclusion criteria in research studies: definitions and why they matter. *JornalBrasileiro de Pneumologia*, *44*, pp.84-84.

14. Lin, H.C., Swarna, H. and Bruning, P.F., 2017. Taking a global view on brand post popularity: Six social media brand post practices for global markets. *Business Horizons*, *60*(5), pp.621-633.

15. Hunt, K. and Gruszczynski, M., 2021. The influence of new and traditional media coverage on public attention to social movements: the case of the Dakota Access Pipeline protests. *Information, Communication & Society*, *24*(7), pp.1024-1040.

16. Wang, X., Baesens, B. and Zhu, Z., 2019. On the optimal marketing

aggressiveness level of C2C sellers in social media: Evidence from china. *Omega*, *85*, pp.83-93.

17. Baljak, V., Ljubovic, A., Michel, J., Montgomery, M. and Salaway, R., 2018. A scalable realtime analytics pipeline and storage architecture for physiological monitoring big data. *Smart Health*, *9*, pp.275-286.

18. Bala, M., Boussaid, O. and Alimazighi, Z., 2017. A Fine-Grained Distribution Approach for ETL Processes in Big Data Environments. *Data & Knowledge Engineering*, *111*, pp.114-136.

19. Elragal, A. and Klischewski, R., 2017. Theory-driven or process-driven prediction? Epistemological challenges of big data analytics. *Journal of Big Data*, *4*(1), pp.1-20.

20. Arunachalam, D., Kumar, N. and Kawalek, J.P., 2018. Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. *Transportation Research Part E: Logistics and Transportation Review*, *114*, pp.416-436.

21. Wang, L. and Alexander, C.A., 2020. Big data analytics in medical engineering and healthcare: methods, advances and challenges. *Journal of medical engineering &technology*, *44*(6), pp.267-283.

22. Moly, M., Roy, O. and Hossain, A., 2019. An advanced ETL technique for error-free data in data warehousing environment. *Int. J. Sci. Res. Eng. Trends*, pp.554-558.

23. Ardagna, C.A., Bellandi, V., Bezzi, M., Ceravolo, P., Damiani, E. and Hebert, C., 2018. Model-based big data analytics-as-a-service: take big data to the next level. *IEEE Transactions on Services Computing*, *14*(2), pp.516-529.

24. Liu, Y., Jiang, C. and Zhao, H., 2019. Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media. *Decision Support Systems*, *123*, p.113079.

25. Li, F., Larimo, J. and Leonidou, L.C., 2021. Social media marketing strategy: definition, conceptualization, taxonomy, validation, and future agenda. *Journal of the Academy of Marketing Science*, *49*(1), pp.51-70.

26. Lee, I., 2018. Social media analytics for enterprises: Typology, methods, and processes. *Business Horizons*, *61*(2), pp.199-210.

27. Hajirahimova, M.S. and Aliyeva, A.S., 2017. About big data measurement methodologies and indicators. *International Journal of Modern Education and Computer Science*, *9*(10), p.1.