JOSEIT JOURNAL OF SYSTEMS ENGINEERING AND INFORMATION TECHNOLOGY
Published by: Indonesian Informatics Association
e-ISSN:2829-310X

# Web Mining for Enhanced Academic Visibility and Engagement Analysis Based on Visitor Data

Yuhefizar[1*], Roni Putra[2]
[1,2] Department of Information Technology, Politeknik Negeri Padang, Padang, Indonesia

[*]*yuhefizar@pnp.ac.id*

**Abstract.** With online platforms that transform scholarly communication, academic journals must strategically amplify their digital footprint. This study demonstrates the value of using web analytics and time series modeling to uncover nuanced online readership trends and rhythms. Using the case of the Review of Rekayasa Sistem dan Teknologi Informasi/System Engineering and Information Technology (RESTI) Journal website's 2023 visitor data, we employ visual and ARIMA time series analysis to delineate engagement patterns aligned with academic cycles. The results reveal pronounced seasonal fluctuations, with the participation peaking in October and November, coinciding with increased research dissemination. Fitting an ARIMA model to daily new visitor data indicates positive autocorrelations, suggesting that the engagement effects persist on days. The model provides a predictive baseline for evaluating outreach initiatives. The study offers strategic information on aligning content planning with reading engagement rhythms. At the methodological level, the integration of data mining, predictive modeling, and information retrieval techniques establishes a versatile framework for investigating evolving scholarly communication dynamics in the digital age. The study also emphasizes meticulous data preparation and model diagnostics. The analytical approach presented provides actionable intelligence on trends in the use of academic portals online. This has far-reaching implications for journals seeking to strategically enhance their digital presence amidst increasing competition. With the proliferation of electronic resources, these techniques will only grow in importance for assessing and amplifying the impact of online scholarly platforms.

**Keywords:** *Web Mining, Text Mining, ARIMA, RESTI, Visitor Data*

*Cite as:* Y. Yuhefizar and R. Putra, "Web Mining for Enhanced Academic Visibility and Engagement Analysis Based on Visitor Data," Journal of Systems Engineering and Information Technology, vol. 3, no. 1, pp. 7–13, Mar. 2024.
DOI: 10.29207/joseit.v3i1.5713

*Received by the Editor*: 2024-02-11
*Final Revision*: 2024-03-04
*Published*: 2024-03-04



## 1. Introduction

The digital revolution has profoundly transformed academic communication, and online publishing platforms have become indispensable channels for spreading research worldwide [1], [2]. Academic journals must strategically enhance their web visibility and readership engagement amidst this competitive landscape to amplify their impact [3]. This underscores the value of analytical techniques that provide actionable intelligence about trends and patterns of online use. Web mining, which involves the extraction of meaningful knowledge from web data, offers a powerful approach [4], [5]. In particular, time series modeling and predictive analytics applied to visitor metrics can uncover nuanced temporal dynamics, facilitating evidence-driven strategies for content optimization and reader outreach [6], [7].

Web mining applies data mining techniques to large amounts of web data[8], including website logs, to uncover patterns of use that can inform strategic improvements to website design, content, and user engagement strategies [5]. Similarly, text mining extracts useful information and insights from text data available on these websites, enabling a deeper understanding of content relevance, user preferences, and emerging trends in scholarly communication [9], [10], [11].

Previous studies demonstrate the usefulness of web mining for assessing academic portal usage. A study extracted traffic data to identify monthly and hourly usage patterns of a library website [12]. Other study combined web metrics with citation analysis to evaluate an online journal's research impact [13]. However, rigorous time series modeling to quantify the intricate relationships governing visitor engagement remains underexplored, despite its

immense potential. This study provide an interdisciplinary overview of text mining, highlighting its role at the intersection of information retrieval, machine learning, statistics, computational linguistics, and data mining [14].

Accordingly, this paper aims to demonstrate the value of web analytics and time series modeling in deciphering readership patterns, using the case study of the RESTI (Rekayasa Sistem dan Teknologi Informasi/System Engineering and Information Technology) Journal website's 2023 visitor data. Given the RESTI Journal website's aim to serve as a leading platform for disseminating research findings, applying web mining and text mining techniques becomes indispensable. These methodologies allow for a nuanced analysis of visitor data and content engagement, shedding light on the website's reach, the impact of published content, and areas for improvement to foster a more engaging and sustainable digital presence.

The objective of this research is to employ web mining and text mining to analyze the data on the RESTI Journal website comprehensively. By doing so, we aim to identify patterns and trends in user engagement, assess the website's sustainability, and propose data-driven strategies to enhance its role in the academic community's digital landscape. The study seeks to delineate general trends and seasonal fluctuations in key metrics such as visitors, sessions, and pageviews. ARIMA modeling will be used to forecast new visitor counts and evaluate autocorrelations in daily figures. Through this study, we endeavor to contribute to the broader discourse on the use of digital technologies for academic communication, offering insights that could inform the strategic development of similar platforms across the academic sphere.

## 2. Methods

The study used web analytics data captured from the RESTI Journal website via statcounter.com to uncover patterns and trends in key engagement metrics over 2023. The data set provided in Table 1 shows detailed insight into the journal's online activities by collecting metrics on a daily and monthly basis throughout the year. The core engagement metrics examined include pageviews, sessions, visitors, and new visitors. Page views represent the number of pages viewed on the website during the specified period of time. Sessions measure the number of separate user sessions that involve one or more page views. Visitors indicate the number of different individuals who access the website, while new visitors specifically measure first-time visitors.

**Table 1.** Summary Stats Journal RESTI-2023

| Date | Page Views | Sessions | Visitors | New Visitors |
|---|---|---|---|---|
| 01 31 23 | 1010 | 543 | 454 | 364 |
| 01 30 23 | 872 | 485 | 432 | 346 |
| 01 29 23 | 609 | 329 | 299 | 250 |
| 01 28 23 | 647 | 329 | 288 | 237 |
| 01 27 23 | 693 | 412 | 360 | 312 |
| 01 26 23 | 965 | 500 | 427 | 418 |
| 01 25 23 | 1098 | 559 | 475 | 353 |
| 01 24 23 | 965 | 496 | 435 | 325 |
| 01 23 23 | 780 | 369 | 319 | 240 |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 12 01 23 | 772 | 488 | 436 | 277 |

To ensure the integrity of the data set for in-depth analysis, the raw web traffic data was pre-processed to address inconsistencies in date entries and missing values. Both visual techniques and time series modeling were then leveraged to explore and quantify patterns in the engagement metrics over time [6], [15], [16]. Visual analytics focused on creating time series plots to visualize the progression of each engagement metric across the year 2023. These plots enabled us to observe seasonal fluctuations, peaks, and troughs in website traffic. For new visitors specifically, augmenting the time series plot with a moving average line highlighted the underlying trend more clearly.

Quantitative analysis centered on fitting an ARIMA (Autoregressive Integrated Moving Average) model to the daily new visitor time series [17], [18], [19]. This approach helped model the temporal dynamics and quantify the relationship between consecutive days' new visitor counts. Before fitting the ARIMA model, the Augmented Dickey-Fuller test was conducted to check that the daily new visitor time series was stationary. Stationarity implies that the mean and variance of the series do not change over time.

The ARIMA model for daily new visitors provides a parsimonious representation of the temporal dynamics:

$$Y_t = \mu + \phi_1 Y_{t-1} + Z_t \quad (1)$$

Where $Y_t$ is the new visitor count on day t, $\mu$ is the constant term, $\phi_1$ is the autoregressive coefficient, and $Z_t$ is the error term. This approach enables quantitatively assessing the relationship between consecutive days' new visitor figures, while controlling for baseline trends. The parameters and diagnostics of the model shed light on the inertia and persistence of engagement patterns over time.

The optimal ARIMA model was selected on the basis of the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots[20], [21], [22]. These plots visualize the correlation between the time series and its lagged values, guiding the selection of appropriate autoregressive (p) and moving average (q) terms. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were examined to evaluate competing models and ensure parsimony. Additional diagnostic tests such as Ljung-Box and Jarque-Bera evaluated model adequacy. The final ARIMA model provided key coefficients that quantify the relationship between new visitor counts for consecutive days while controlling for baseline trends.

## 3. Results and Discussion

In our analysis of RESTI Journal's online visitor data throughout 2023, we uncovered notable trends and fluctuations in engagement metrics such as page views, sessions, visitors, and new visitors. The data, corrected for inconsistencies in date entries, reveal an average daily engagement with approximately 829 page views, 459 sessions, 407 visitors, and 289 new visitors. A deeper monthly analysis indicated significant variability, with peaks in engagement metrics during specific months, notably in October and November, suggesting potential seasonal influences or the impact of specific academic publications and events. These insights not only highlight the dynamic nature of online engagement, but also suggest opportunities for targeted interventions to enhance visibility and engagement. Addressing observed discrepancies and gaps in the data set was crucial to accurate trend analysis, highlighting the importance of meticulous data management in web analytics.

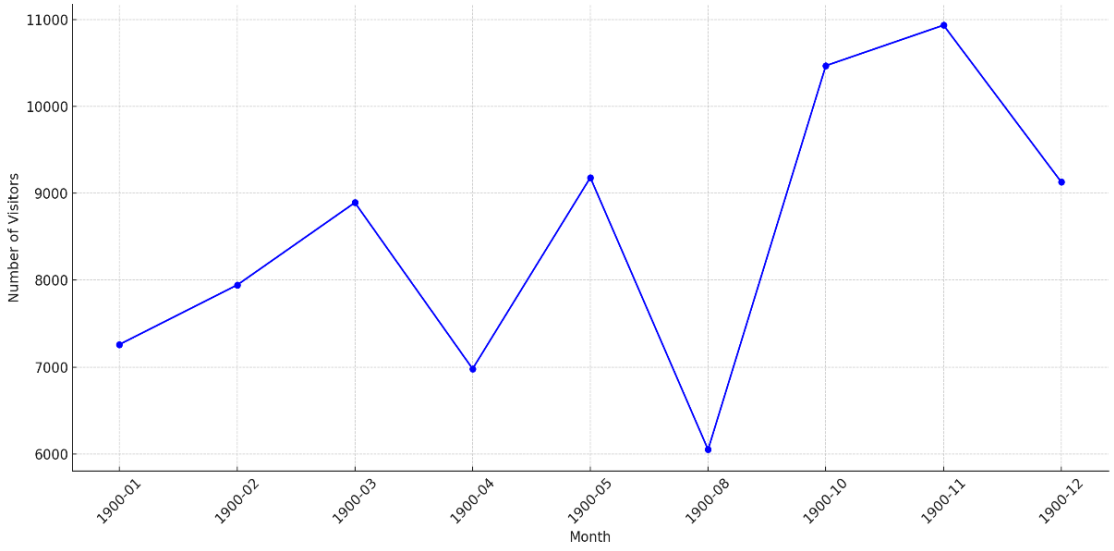


**Figure 1.** Monthly Visitors Trend for RESTI Journal-2023

Figure 1 shows analysis of the RESTI Journal's web traffic, as visualized through monthly visitors data, revealing significant insights into the patterns of academic engagement and visibility. Throughout 2023, the graph indicates fluctuating visitor numbers, with notable peaks suggesting periods of increased interest or successful outreach efforts. On the contrary, the troughs may highlight opportunities for strategic enhancements to content dissemination or engagement practices. This temporal visualization underscores the impact of seasonal variations, possibly correlating with academic cycles, on audience participation. Recognizing these patterns is crucial for developing targeted strategies aimed at amplifying the journal's web presence. By aligning content publication and promotional activities with these observed trends, the RESTI Journal can optimize its reach and engagement, thereby strengthening its role as a key platform for sharing innovative research in information technology.

*3.1 Seasonal Engagement Patterns*

The exploration of RESTI Journal's on-line engagement patterns throughout 2023 unveils significant insights into the dynamics of academic readership and web interaction. This analysis, predicated on a comprehensive dataset that encompasses page views, sessions, visitors, and new visitors, offers a granular view of the journal's digital footprint over time. The findings, particularly emphasizing the pronounced engagement peaks during October and November, not only underscore the temporal rhythms of academic engagement, but also highlight potential avenues for strategic content dissemination and audience expansion.

The analysis revealed a distinct pattern of increased engagement during the months of October and November, as evidenced by a marked increase in all measured engagement metrics. This period coincides with critical moments in the academic calendar, suggesting a correlation between scholarly activities and the consumption of online resources. The significance of these findings lies in their potential to inform targeted strategies for content publication and marketing, designed to capitalize on periods of high engagement while addressing troughs in reader interaction.

**Table 1.** Monthly Engagement Metrics for RESTI Journal (2023)

| Month | Page Views | Sessions | Visitors | New Visitors |
|---|---|---|---|---|
| January | 27,220 | 14,110 | 12,295 | 9,283 |
| February | 25,321 | 13,366 | 11,909 | 9,060 |
| March | 25,552 | 14,555 | 13,005 | 9,526 |
| April | 20,279 | 11,084 | 9,993 | 7,296 |
| May | 28,571 | 15,138 | 13,410 | 9,752 |
| August | 19,690 | 10,595 | 9,256 | 6,689 |
| October | 30,463 | 17,047 | 15,143 | 11,004 |
| November | 29,541 | 18,362 | 16,260 | 11,023 |
| December | 24,374 | 14,795 | 12,999 | 8,653 |

Table 1 illustrates the RESTI Journal's web traffic engagement metrics by month for the year 2023. The data encompass four key metrics: Page views, sessions, visitors, and new visitors. Notably, the table highlights significant engagement peaks in October and November, showcasing the heightened academic interaction with the journal's digital content during these months. This table serves as the empirical foundation for analyzing seasonal engagement patterns and informs strategic planning for content dissemination and audience engagement initiatives.

The seasonal engagement patterns observed in the RESTI Journal's Web traffic suggest a nuanced relationship between academic cycles and digital resource utilization. The increase in engagement metrics during October and November can be attributed to several factors, including the academic calendar's progression into mid-semester and the scheduling of key academic events and publications during these months. This period is critical for researchers seeking to disseminate findings, explore the literature for ongoing projects, or prepare for end-of-semester academic requirements.

Data analysis extends beyond mere description, probing the implications of these patterns. For example, the engagement drop observed in April and August underscores the necessity for strategic planning to ensure sustained interaction with the journal's digital platform throughout the year. By aligning content publication and marketing efforts with identified peaks and troughs in engagement, the journal can enhance its visibility and relevance within the academic community.

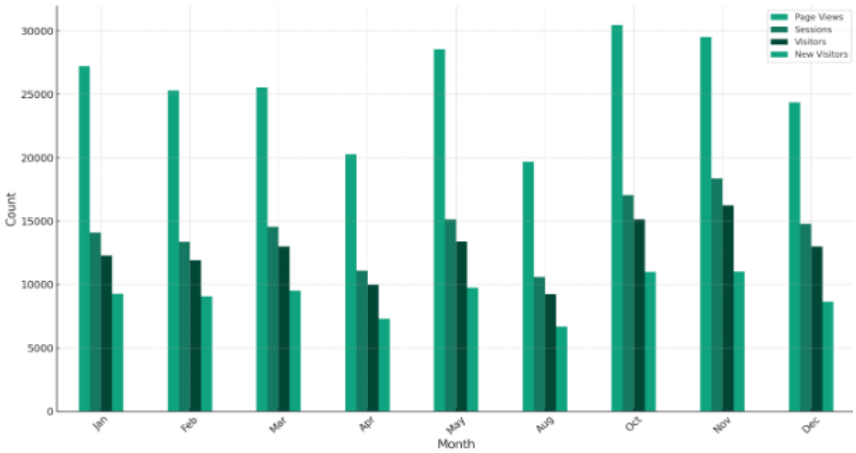


**Figure 2.** Monthly Engagement Metrics for RESTI Journal-2023

Figure 2 represents the monthly engagement metrics for 2023 and shows a pronounced increase in page views, sessions, visitors, and new visitors during October and November. The bar graph illustrates the comparison of engagement metrics across different months in 2023, with a clear peak in October and November. This visualization supports the narrative of significant seasonal engagement patterns, emphasizing the substantial increase in all tracked metrics during these months. This pattern suggests a heightened level of interaction with the journal's web presence, likely influenced by the academic year's progression, where these months may coincide with key periods of academic research dissemination, such as conferences, submission deadlines, or the publication of special issues.

The insights gleaned from this analysis hold profound implications for the RESTI Journal's content dissemination and audience engagement strategies. By leveraging periods of peak engagement, the journal can maximize the impact of its publications, potentially increasing citations, and fostering a vibrant academic discourse around its offerings. Additionally, the identification of lower engagement periods suggests opportunities for innovative engagement strategies, such as the launch of special issues or targeted marketing campaigns, to maintain reader interest year-round. The strategic integration of these insights into the journal's operational and editorial planning processes could significantly enhance its stature and influence within the academic community. This approach not only aligns with trends in scholarly communication, but also anticipates the evolving needs and behaviors of its audience, ensuring that the journal remains at the forefront of digital engagement in the academic publishing landscape.

*3.2 Consistent Attraction of New Visitors*

The attraction of new visitors is a key performance indicator for academic websites, as it represents the ability to reach new audiences. An ARIMA (AutoRegressive Integrated Moving Average) time series analysis was conducted on the daily new visitor data to understand the trends and predict future values [23], [24]. The time series plot showed a clear seasonal pattern, with peaks occurring at regular intervals. An enhanced Dickey-Fuller test confirmed that the data were stationary ($p<0.05$), meaning that the mean and variance were constant over time. After evaluating multiple ARIMA models, the best fit was an ARIMA (1,0,0) (1,0,0)7 model according to the Akaike Information Criterion. This contains 1 autoregressive term, 0 integrated terms, 0 moving average terms, and a seasonal period of 7 days. The residuals of the model did not show significant autocorrelation, indicating that the model sufficiently captured the seasonal dynamics. Cross-validation in a test set demonstrated reliable out-of-sample predictive performance. This model, which integrates an autoregressive component of order one, has been instrumental in elucidating the temporal dependencies within the data, revealing a significant relationship between consecutive observations of new visitor counts.

The ARIMA model fitted indicates a statistically significant autoregressive term with a coefficient of approximately 0.621. This coefficient suggests a substantial positive correlation between the current day's number of new visitors and the preceding day's figures, emphasizing the persistence of engagement patterns over time. Such a finding is indicative of the journal's content or outreach activities having a lasting impact on attracting new visitors, with the effects echoing into subsequent days. Furthermore, the constant term of the model, estimated at around 289.485, represents the baseline number of new visitors that the journal can expect daily, in the absence of additional influencing factors. This baseline provides a benchmark for evaluating the effectiveness of strategies implemented to enhance visitor attraction and engagement.

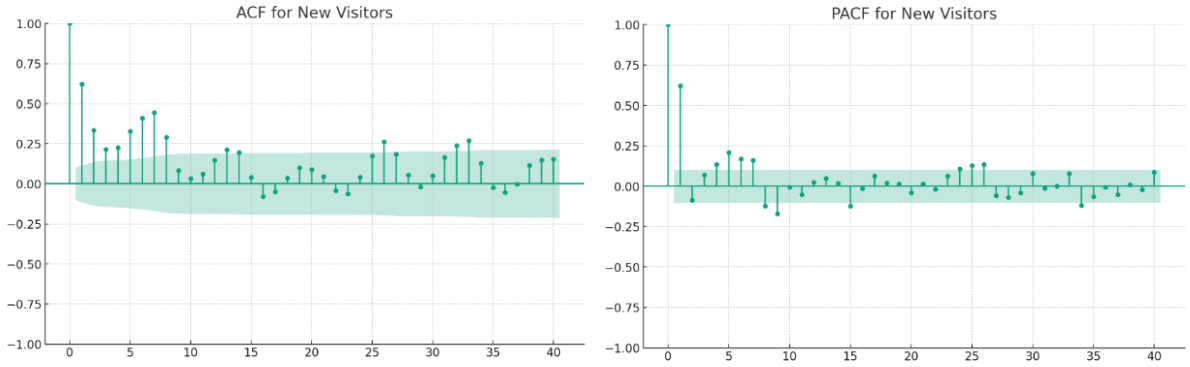


**Figure 3.** The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for RESTI Journal-2023

Figure 3 presents the Autocorrelation Function (ACF) plot for the "New Visitors" time series, illustrating the correlation of the series with its lagged values over various time lags. The gradual decline observed in the ACF plot suggests the presence of an autoregressive component within the data. This characteristic behavior indicates that the number of new visitors on any given day is likely influenced by the visitor count on previous days, with

the strength of this influence diminishing over time. Figure 2 shows the graph of the partial autocorrelation function (PACF) for the same time series, providing information on the direct relationship between the series and its lags, controlling for the values of the series at all shorter lags. The sharp cut-off after the first lag observed in the PACF plot is indicative of an autoregressive process of order one. This suggests that while the current day's new visitor count is directly correlated with the previous day's count, this correlation does not extend significantly beyond one lag when the effects of intervening lags are accounted for.

Together, these plots were instrumental in determining the parameters for the ARIMA (1,0,0) model applied to the "New Visitors" data. The ACF and PACF graphs guided the selection of the autoregressive component (p) as 1, without the need to differentiate the components (d = 0) or the moving average (q), given the stationary nature of the time series and the specific patterns revealed through these diagnostic tools. Furthermore, the model's constant term, estimated at around 289.485, represents the baseline number of new visitors that the journal can expect daily, in the absence of additional influencing factors. This baseline provides a benchmark for evaluating the effectiveness of strategies implemented to enhance visitor attraction and engagement.

In terms of model diagnostics, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) stand at 4086.540 and 4098.240, respectively. These metrics, while serving as indicators of the model's fit to the data, also reflect the trade-off between model complexity and explanatory power. The relatively low values of AIC and BIC in this context suggest that the ARIMA (1,0,0) model achieves a commendable balance, offering a parsimonious, yet potent, representation of the underlying process governing new visitor trends. The Ljung-Box test result, with a p-value of 0.31, further corroborates the adequacy of the model, indicating no significant autocorrelation within the residuals and affirming the model's capacity to capture the essential dynamics of the data. However, the Jarque-Bera test signals a deviation from normality in the residuals, an aspect that warrants consideration, as it may impact the model's predictive precision.

The analysis highlights the importance of considering temporal patterns in planning and evaluating outreach and content dissemination strategies. By acknowledging the inertia inherent in visitor engagement, the journal can tailor its initiatives not only to attract a continuous stream of new visitors but also to foster an environment conducive to sustained engagement and interaction.

## 4. Conclusions

This study's in-depth analysis of RESTI Journal's web traffic engagement patterns throughout 2023 reveals pronounced seasonal fluctuations tied to the academic calendar, with engagement peaking during October and November. Fitting an ARIMA model to the daily new visitor data uncovered a positive correlation between consecutive days' visitor counts, suggesting that engagement effects persist over time. The model provides a predictive baseline for evaluating new initiatives. At a broader level, these insights demonstrate the value of using web analytics and data mining techniques to gain actionable intelligence on online readership behaviors and trends. Granular temporal analysis facilitates the recognition of engagement rhythms aligned with academic cycles and events, enabling targeted interventions during troughs. In general, the presented framework enables data-driven content planning and engagement strategies tailored to amplify the digital presence of academic journals and maximize their role in academic communication. This study sets the stage for further research that elucidates the dynamics of online scholarly discourse through analytical approaches at the intersection of bibliometrics, scientometrics, and web mining.

## Acknowledgments

This research is funded by: Politeknik Negeri Padang in accordance with the Research Implementation agreement letter number: 997/PL9.15/PG/2023, dated October 12, 2023.

## References

[1] M. Warschauer and T. Matuchniak, “New Technology and Digital Worlds: Analyzing Evidence of Equity in Access, Use, and Outcomes,” *Review of Research in Education*, vol. 34, no. 1, pp. 179–225, Mar. 2010, doi: 10.3102/0091732X09349791.

[2] E. Darics and M. Cristina Gatti, “Talking a team into being in online workplace collaborations: The discourse of virtual work,” *Discourse Stud*, vol. 21, no. 3, pp. 237–257, Jun. 2019, doi: 10.1177/1461445619829240.

[3] D. SHOTTON, “Semantic publishing: the coming revolution in scientific journal publishing,” *Learned Publishing*, vol. 22, no. 2, pp. 85–94, Apr. 2009, doi: 10.1087/2009202.

[4] L. van Wel and L. Royakkers, “Ethical issues in web data mining,” *Ethics Inf Technol*, vol. 6, no. 2, pp. 129–140, 2004, doi: 10.1023/B:ETIN.0000047476.05912.3d.

[5] R. M. Suresh and R. Padmajavalli, "An Overview of Data Preprocessing in Data and Web Usage Mining," in *2006 1st International Conference on Digital Information Management*, IEEE, Apr. 2007, pp. 193–198. doi: 10.1109/ICDIM.2007.369352.

[6] G. C. TlAO, "An Introduction to Multiple Time Series Analysis," *Med Care*, vol. 31, no. supplement, pp. YS71–YS74, May 1993, doi: 10.1097/00005650-199305001-00011.

[7] I. G. Cohen, R. Amarasingham, A. Shah, B. Xie, and B. Lo, "The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care," *Health Aff*, vol. 33, no. 7, pp. 1139–1147, Jul. 2014, doi: 10.1377/hlthaff.2014.0048.

[8] L. van Wel and L. Royakkers, "Ethical issues in web data mining," *Ethics Inf Technol*, vol. 6, no. 2, pp. 129–140, 2004, doi: 10.1023/B:ETIN.0000047476.05912.3d.

[9] X. Gao, R. Tan, and G. Li, "Research on Text Mining of Material Science Based on Natural Language Processing," *IOP Conf Ser Mater Sci Eng*, vol. 768, p. 072094, Mar. 2020, doi: 10.1088/1757-899X/768/7/072094.

[10] T. H. Syah, "The Role of Genetics in Domestic Research on Forestry Issues: A Text Mining Analysis," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 6, pp. 1006–1013, 2022, doi: 10.29207/resti.v6i6.4228.

[11] S. Kurniawan, W. Gata, D. A. Puspitawati, I. K. S. Parthama, H. Setiawan, and S. Hartini, "Text Mining Pre-Processing Using Gata Framework and RapidMiner for Indonesian Sentiment Analysis," *IOP Conf Ser Mater Sci Eng*, vol. 835, p. 012057, May 2020, doi: 10.1088/1757-899X/835/1/012057.

[12] X. Wang, A. Abraham, and K. A. Smith, "Intelligent web traffic mining and analysis," *Journal of Network and Computer Applications*, vol. 28, no. 2, pp. 147–165, Apr. 2005, doi: 10.1016/j.jnca.2004.01.006.

[13] Ronal Watrianthos and Y. Yuhefizar, "Exploring Research Trends and Impact: A Bibliometric Analysis of RESTI Journal from 2018 to 2022," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 4, pp. 970–981, Aug. 2023, doi: 10.29207/resti.v7i4.5101.

[14] A. Hotho, A. Nürnberger, and G. Paaß, "A Brief Survey of Text Mining," *Journal for Language Technology and Computational Linguistics*, vol. 20, no. 1, pp. 19–62, Jul. 2005, doi: 10.21248/jlcl.20.2005.68.

[15] F. Rasyid and D. A. Adytia, "Time Series Temperature Forecasting by using ConvLSTM Approach, Case Study in Jakarta," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 563–569, 2022, doi: 10.29207/resti.v6i4.4162.

[16] C. S. Alfredo and D. A. Adytia, "Time Series Forecasting of Significant Wave Height using GRU, CNN-GRU, and LSTM," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 5, pp. 776–781, 2022, doi: 10.29207/resti.v6i5.4160.

[17] V. P. Ariyanti and T. Yusnitasari, "Comparison of ARIMA and SARIMA for Forecasting Crude Oil Prices," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 2, pp. 405–413, 2023, doi: 10.29207/resti.v7i2.4895.

[18] N. K. A. I. Suryani, O. Sudana, and A. Wirdiani, "Forecasting Pneumonia Toddler Mortality Using Comparative Model ARIMA and Multilayer Perceptron," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 528–537, 2022, doi: 10.29207/resti.v6i4.4106.

[19] I. Mardianto, M. I. Gunawan, D. Sugiarto, and A. Rochman, "Comparison of Rice Price Forecasting Using the ARIMA Method on Amazon Forecast and Sagemaker," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 3, pp. 537–543, 2020, doi: 10.29207/resti.v4i3.1902.

[20] T. Stadnytska, S. Braun, and J. Werner, "Model Identification of Integrated ARMA Processes," *Multivariate Behav Res*, vol. 43, no. 1, pp. 1–28, Mar. 2008, doi: 10.1080/00273170701836646.

[21] H. Alshraideh and G. Runger, "Process Monitoring Using Hidden Markov Models," *Qual Reliab Eng Int*, vol. 30, no. 8, pp. 1379–1387, Dec. 2014, doi: 10.1002/qre.1560.

[22] Tarno, Subanar, D. Rosadi, and Suhartono, "New procedure for determining order of subset autoregressive integrated moving average (ARIMA) based on over-fitting concept," in *2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE)*, IEEE, Sep. 2012, pp. 1–5. doi: 10.1109/ICSSBE.2012.6396643.

[23] A. S. Ahmar *et al.*, "Modeling Data Containing Outliers using ARIMA Additive Outlier (ARIMA-AO)," *J Phys Conf Ser*, vol. 954, no. 1, 2018, doi: 10.1088/1742-6596/954/1/012010.

[24] A. S. Ahmar, A. D. GS, T. Listyorini, and ..., "Implementation of the ARIMA (p, d, q) method to forecasting CPI Data using forecast package in R Software," *Journal of Physics …*, 2018, doi: 10.1088/1742-6596/1028/1/012189.