

## Google trends and online media data for supply and demand information in waste management evaluation in Jakarta

Robert Kurniawan<sup>1,3,6</sup>, Agung Purwanto<sup>6</sup>, Anugerah Karta Monika<sup>2,3</sup>, Krismanti Tri Wahyuni<sup>2,3</sup>,  
Muhammad Yunus Hendrawan<sup>3</sup>, Mohamad Andrian Isnaeni<sup>4,5</sup>

<sup>1</sup>Department of Statistical Computing, Faculty of Statistics, Politeknik Statistika STIS, Jakarta, Indonesia

<sup>2</sup>Department of Statistics, Faculty of Statistics, Politeknik Statistika STIS, Jakarta, Indonesia

<sup>3</sup>Statistics Indonesia (BPS), Jakarta, Indonesia

<sup>4</sup>Department of Economic, Faculty of Economics and Business, Sebelas Maret University, Surakarta, Indonesia

<sup>5</sup>Local Development Planning, Research and Development Agency (Bappeda Litbang) Sragen Regency, Sragen, Indonesia

<sup>6</sup>Department of Population and Environmental Education, Faculty of Post-Graduate, State University of Jakarta, Jakarta, Indonesia

### Article Info

#### Article history:

Received Nov 28, 2022

Revised Jan 15, 2023

Accepted Jan 17, 2023

#### Keywords:

Demand of information

Google trends

Impulse response function

Latent dirichlet allocation

Supply of informastion

Waste management

### ABSTRACT

Demand and supply of information in online media, particularly regarding waste management, remain hampered by a number of obstacles. Consequently, the objective of this study is to determine the public's interest in waste management knowledge based on demand data obtained from Google trends and to determine the most recent events in waste management by analyzing online news content. As a result, vector autoregressive (VAR) with impulse response function (IRF) and latent dirichlet allocation (LDA) are utilized as the analysis method. An important finding of this study is that it takes at least four weeks for individuals to absorb waste management information. Therefore, it is necessary for the government and the pentahelix component to sit together in order to reduce the community's information acquisition delay. Waste management, which is the subject of the shared information, should guide the selection of keywords by information providers.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



### Corresponding Author:

Robert Kurniawan

Department of Statistical Computing, Faculty of Statistics, Politeknik Statistika STIS

Jatinegara, 13330, East Jakarta, DKI Jakarta, Indonesia

Email: robertk@stis.ac.id

## 1. INTRODUCTION

Waste is a big problem faced by big cities in developing countries such as Indonesia, not least the city of Jakarta [1]–[3]. Jakarta is a megapolitan city and the capital of Indonesia, with an area of 664.01 km<sup>2</sup>, with 10,576 million people. With a population density of 15,927.47 inhabitants/km<sup>2</sup>, Jakarta is one of the megapolitan cities in Indonesia with population density, and waste management is problematic [4]. The main urban problems are waste management [5] and poor sanitation, thus adversely affecting health. It is because the rate of urbanization and population density is increasing [6].

Some studies mention that problems in waste management are caused by education. Collaborative education and policy enforcement are necessary for controlling and managing waste [7]. Formal education plays a vital role in waste management [4]. Lower education tends to be more littering [8]. Though waste management cannot be entirely handed over to state institutions, there needs to be the active involvement of the entire community through awareness and participation carried out by the government [9]. In addition to formal education, informal education is also needed and proven to affect public awareness and concern for community management [4]. Public concern and awareness can be seen from behavior seeking information on managing waste. Some people still lack knowledge about the adverse effects of poorly managed waste piles

[6]. The waste management depends on the community's social behavior [10]. Social media platforms are targeted by stakeholders in publicity about waste management and education on the harmful effects of solid waste [6].

Along with social media as a channel for presenting information to the public, Google has developed into one platform that feeds information to search engines. With the presence of propaganda and publicity through online media [10], Google can also be helpful to see the behavior of the public in searching for information [11]. Google is the most commonly used tool to search for specific information [12]. In terms of information coverage and accessibility, Google is excellent [13].

Google trends analyze internet users obtained from Google's search engine database as a multi-sided one that allows two different interest groups to meet benefits [14]–[16]. Groups of people who provide information and groups that seek information. We used the Google Trends approach to see the public's interest in finding waste management information in this study. While our online news media uses a group approach that provides information supply about waste. Some researchers use Google trends to see people's interest in certain information. For example, people's interest in seeking health information [17]–[22], fashion [23], socioeconomic [24], [25]. However, no research has been conducted on public interest in waste management information. So this research takes up space as an exciting thing to research. Especially the interest of the people of Jakarta in seeking information related to waste. The intended management includes reuse, reduce, and recycle (3R), as 3R is considered a significant element in waste management [26].

Waste management analysis with a big data approach has been done. One source of big data is weChat, which is used to see how households manage waste [10]. The United nations (UN) considers the calculation of SDG's indicators by using google trends [25]. In addition to google trends, which is used to measure people's interest in information, online news is present to provide services to supply the information needed. In other studies, online news was used to determine the construction of media coverage of forest fires [27], handling Coronavirus disease (COVID-19) [28], climate change [29], mental health during COVID-19 [30], healthcare stocks forecasting [31], classification and trends of vaccines and vaccinations [32], and restrictions on the use of plastic waste [33]. While online news is related to waste management information, especially in Indonesia, there has never been any research that explains it. So it is the opportunity to research how the grouping of waste topics that become the supply of information for the community is taken from online news.

Based on the context, the primary research question is whether the community effectively receives waste management information. Second, how do related parties create information for it to be accepted by the larger community? using Google trends data, this study aims to determine the public's interest in waste management information and the amount of time required for the public to receive online information. Another objective is to investigate the grouping of waste-related news topics in online news sources in Jakarta. It is to determine whether the results align with Google trends searches. So this research can provide a picture of the community's waste management information supply and demand. The results obtained later will provide stakeholders with policy recommendations to improve waste management education in the community.

## 2. METHOD

The supply of information about the environment is needed even though it is unknown who the demand for information is [34]. There is research related to the supply-demand of drug information in Thailand [35]. Therefore, the knowledge of supply and demand of waste management information is essential to research, which in the end, this study gained an idea of how public interest in waste management. This study used big data derived from the internet, Google trends, and scraping from national online news: *detik.com*.

### 2.1. Google trends

Google trends is a tool provided by Google that allows users to view and analyze data on search trends. It provides information on how often specific terms or phrases are searched for on Google over time. This data is presented in the form of a graph, with search volume plotted on the y-axis and time plotted on the x-axis. The data can be filtered by location, language, and category, and users can compare the search volume of multiple terms or phrases. Google trends can be used for a variety of purposes, including keyword research, market research, and identifying trends in public opinion. There is research that describes the characteristics of Google trends as part of big data [36]. The term 'big data' in this context does not refer to the number of samples, but rather to observations made at a global level.

The sample period for Google trends is typically high frequency and up-to-date, making it useful for nowcasting. Google trends data is available for specified periods, such as daily, weekly, or monthly, as chosen by the user. It is presented in the form of an index derived from search volume index (SVI) [37]. The data used in this study is sourced from Google trends, covering the period from October 1, 2017 to October 1, 2021. The keywords used in Google trends are listed in Table 1 and treated as variables in the model.

Table 1. Search keywords in Google trends

Keyword	Description
<i>Sampah</i>	<i>Sampah</i> is an Indonesian vocabulary that describes waste. It refers to how people search about waste on the internet.
Waste	Waste refers to keyword that describe how people search about how managing waste
Reuse	Reuse refers to a keyword that describes a way of managing waste (from Reuse, Reduce, Recycle)
Reduce	Reduce refers to a keyword that describes a way of managing waste (from Reuse, Reduce, Recycle)
Recycle	Recycle refers to a keyword that describes a way of managing waste (from Reuse, Reduce, Recycle)

## 2.2. Online news from detik.com

The data used in this study is secondary data in the form of news with the keyword "*Sampah di Jakarta*" (waste in Jakarta) sourced from *detik.com*. *detik.com* was a reference source for news gatherings because it is Indonesia's most popular news site [38]. In addition, *detik.com* is also included in the top five sites with the highest amount of traffic in Indonesia. The news collected in this study has a publishing time from October 20, 2017, to October 26, 2021.

Data preprocessing is the initial stage of text mining that aims to prepare data before further processing or analysis can be done. Data preprocessing is required to convert text from human language into a machine-managed format, structure unstructured text, and maintain keywords useful for representing topics [39]. Before preprocessing the news text data, the title and the article's content are merged into one context, and then the filtering process is carried out to filter the location of the news. News locations not from the DKI Jakarta area were not used in this study. Two thousand four hundred twenty-one articles originate from DKI Jakarta and will be analyzed further.

The preprocessing stage of this research data begins with text normalization, which includes converting text into non-capital letters or case *folding*, eliminating special characters and punctuation, and eliminating white space. After the process is complete, the input document will go through the process of converting non-standard words (slang words) into common words, removal of words that often appear (stopwords), and the process of converting words into essential words (stemming). The results of text cleaning can be viewed through the word cloud in Figure 1.

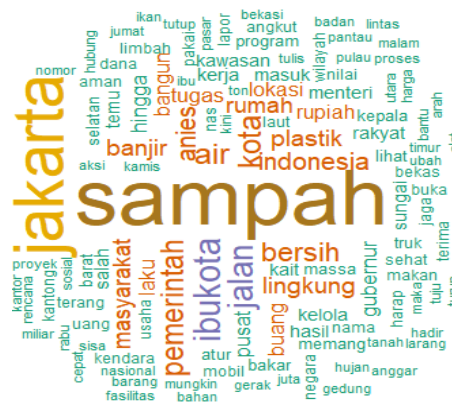


Figure 1. Wordcloud data text cleaning results (in Indonesian language), collected by Scrapy in Python

The number of news in the period October 2017 to October 2021 is seen in Figure 2, which shows that the most news related to waste in Jakarta was in February 2020. It can happen because of a flood event throughout January 2020 in Jakarta. So there is much news about the waste associated with the flood incident in Jakarta in February 2020. Then, much news about waste in Jakarta occurred in March 2018. Much waste was found in Jakarta Bay during that period, so many online mass media reported it.

Furthermore, news about the waste that occurs a lot can be seen in October 2018. The problem of waste in DKI Jakarta and Bekasi city was revealed. The three incidents triggered the media to report it more often. It can be seen that every month during the period October 2017 to October 2021, there is news about waste. It means that, from the supply side, information about waste is always available. Not to mention if it is added with other information obtained from other online media or websites from official government agencies or NGOs. This information will be more and more if it is added with other information obtained from other online media or websites from official government agencies or NGOs.

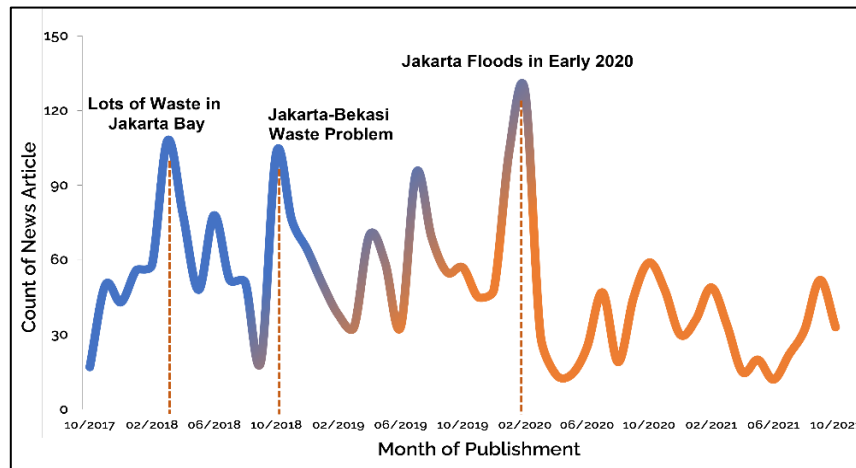


Figure 2. Graph of the amount of waste-related news in Jakarta

### 2.3. Vector autoregressive

Vector autoregressive (VAR) models are a class of statistical models used to analyze multivariate time series data. In a VAR model, each variable is modeled as a linear combination of its own past values and the past values of all other variables in the model. This makes VAR models a generalization of autoregressive (AR) models, which model each variable as a linear combination of its own past values only. To estimate a VAR model, the first step is to specify the lag structure, or the number of past values of each variable to include as predictors. The model is then fit to the data using ordinary least squares (OLS) regression. Once the model is fit, it can be used to make forecasts, test hypotheses about relationships between variables, and analyze the impact of shocks or interventions on the system [40].

VAR model does not distinguish between exogenous and endogenous variables. All variables are treated as exogenous. With the characteristics of such equations, it can be said that VAR is a statistical method that describes the relationship between variables simultaneously [41]. VAR models are suitable for analyzing the dynamic and causal relationships between economic variables, whereas OLS regression models cannot perform such an analysis. VAR models can also handle autocorrelation issues that arise from time series data, which OLS regression models cannot [42]. The VAR model used in this study is a VAR model lag 1 or written as VAR (1) that includes only one lag of each variable as a predictor. The shape of the model is as (1):

$$Y_t = A_0 + A_1 Y_{t-1} + e_t \quad (1)$$

$Y_t$  is a vector consisting of google trends variables used, namely *sampah*, reuse, reduce, recycle and waste.  $Y_{t-1}$  is lag 1 of vectors  $Y_t$ ,  $A_0$ , and  $A_1$  contain regression coefficients, and  $e_t$  is the term error of the VAR (1) model.

### 2.4. Impulse response function (IRF)

In VAR model, the IRF shows how the variables in the model respond over time to a shock in a single variable. To compute the IRF, a VAR model is first estimated using time series data. Once the model is fit, the IRF is calculated by setting the shock or intervention in one variable to a large value at a specific time point, and then simulating the model forward in time to see how the variables respond. The resulting IRF shows the dynamic response of each variable to the shock over time. The IRF shows how each variable in the model responds to the shock over time by looking for patterns in the responses of different variables, such as whether some variables respond strongly than others, or whether there are delays or lags in the responses. Impulse responses are most often interpreted through grid graphs of the individual responses of each variable to an implemented shock over a specified time horizon [40].

### 2.5. Topic modeling

The topic modeling is done after all the data goes through the preprocessing. Implementation of topic modeling in this study was carried out by applying latent dirichlet allocation (LDA) modeling. LDA is a probabilistic generative model topic modeling technique for corpora that is used to extract themes from a document's content [43]. The concept of the LDA method is based on a document represented as a random mixture of hidden topics, in which each topic is characterized by a set of probability sets of words representing the words included in a topic. The LDA assumes the following generative processes for each document  $w$  in the corpus [44]: i) Choose Poisson  $(\xi).N \sim$ , ii) Select  $Dir(\alpha).\theta \sim$ , and iii) For each word:  $Nw_n$ .

- Select the topic  $z_n \sim \text{Multinomial}(\theta)$ .
- Choose a word from a multinomial, probability conditioned on the topic.  $w_n p(w_n | z_n, \beta) z_n$

Before using LDA to implement topic modelling, each method is used to determine the best parameters that are applied to modeling through tuning parameters. After the topic modeling process is done, the resulting topic model needs to be evaluated to see effectiveness in grouping topics through a coherence score. Coherence score is a standard measure used to evaluate topic models by measuring the semantic similarity score of words in a topic [45].

### 3. RESULTS AND DISCUSSION

#### 3.1. VAR

The study used the VAR (1) model estimated with ordinary least square. This model uses five variables that index keywords in google trends, namely *SAMPAH*, REUSE, REDUCE, RECYCLE, WASTE. The VAR (1) equation system produces five equations in which each variable is a dependent variable with an independent variable lag 1 of all variables. The estimated VAR (1) model is done with stationary data at the level using the EViews 12 student version presented in Table 2.

Table 2 shows  $SAMPAH_{t-1}$  influence *SAMPAH*, REUSE, REDUCE, and WASTE.  $REUSE_{t-1}$  influence REUSE and WASTE.  $REDUCE_{t-1}$  influence REDUCE and RECYCLE.  $RECYCLE_{t-1}$  influence *SAMPAH* and REDUCE.  $WASTE_{t-1}$  influence RECYCLE and WASTE. The IRF is calculated using an estimated equation to understand the dynamic impact of word search on Google when a shock or change of one standard deviation occurs. The IRF for each variable indicates how that variable responds to a shock in another variable. A shock of one standard deviation is used. This is also known as an innovation [40]. This response is plotted on a graph, with the horizontal axis representing the time following the shock and the vertical axis showing the value of the response for each variable in the event of a shock to a particular variable, including the variable itself.

In Figure 3, the IRF for the variable *SAMPAH* is shown in response to a shock of one standard deviation in the variables *SAMPAH*, REUSE, REDUCE, RECYCLE, and WASTE. The horizontal axis represents the time following the shock, and the vertical axis shows the response of the *SAMPAH* variable. The IRF can be used to determine the duration of the dynamic impact of the shock by looking for a flat or leveling off curve. In this case, it appears that the impact of the shock on the search for the word "waste" will be stable for up to four periods, or one month, in the future. The same pattern is observed for the variables REUSE, REDUCE, RECYCLE, and WASTE in response to shocks in the same variables. This suggests that when the keywords *SAMPAH*, REUSE, REDUCE, RECYCLE, and WASTE are searched on Google, the resulting impact will last for up to one month. This could be interpreted as the potential impact of publicizing information related to these keywords or topics.

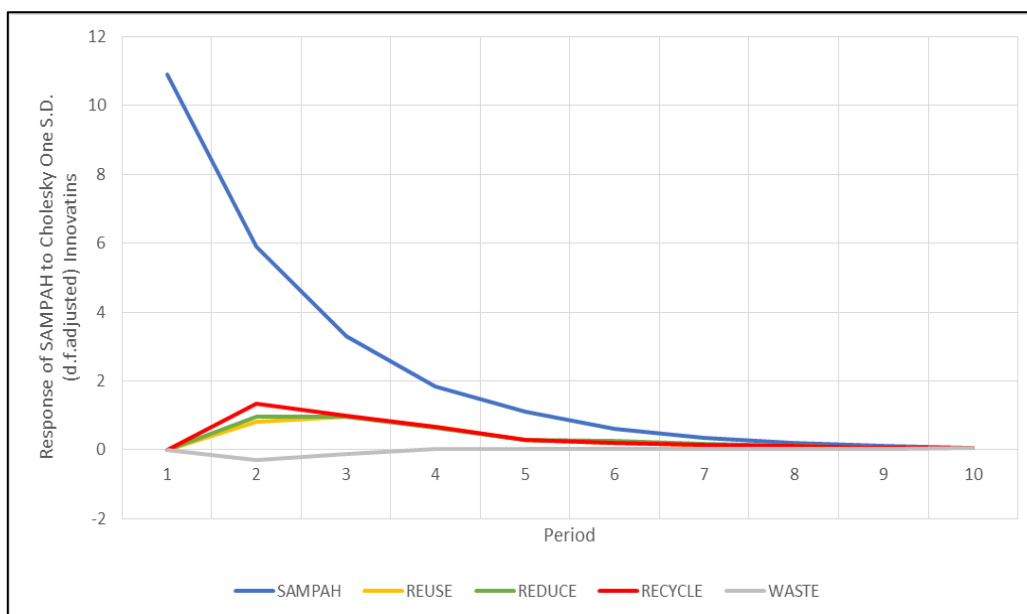


Figure 3. The impulse response function (IRF) of the variables *SAMPAH*, REUSE, REDUCE, RECYCLE, and WASTE in response to a shock of one standard deviation in the *SAMPAH* variable

Table 2. Estimated vector autoregressive model lag (1)

Independent Variables	Equation (dependent variables)				
	$SAMPAH_t$	$REUSE_t$	$REDUCE_t$	$RECYCLE_t$	$WASTE_t$
Trash <sub>t-1</sub>	0.507*	0.022*	0.042*	0.013	0.04*
REUSE <sub>t-1</sub>	-0.109	0.228*	-0.037	0.204	-0.389*
REDUCE <sub>t-1</sub>	0.474	0.045	0.227*	0.15*	0.032
RECYCLE <sub>t-1</sub>	0.9*	-0.013	0.378*	0.03	0.14
WASTE <sub>t-1</sub>	0.096	-0.016	0.005	0.103*	0.143*
C	9.302*	0.293	1.544*	1.259*	2.46*
R-squared	0.399	0.188	0.299	0.183	0.084

\*significant at 5 percent level

### 3.2. Topic modeling

Topic modeling is done on the data set of waste problems in Jakarta completed through text preprocessing. It is done first to determine the parameters of the number of best topics produced in modeling, based on the size of the coherence score. This coherence evaluation measures how high-scoring words in a topic are similar to each other in terms of semantic similarity. It distinguishes semantically understandable topics from statistical inference artifacts. At the beginning of the determination, a set of 12 topic models was set to see the k-topics that provided maximum coherence score.

After the iteration is complete, the model with ten topics has the highest coherence value, as shown in Figure 4. The distribution of words that make up each topic can be known from the ten best topics while generated, as shown in Table 3. It can be known that waste-related news in Jakarta can be grouped into ten topics based on the LDA method. From the grouping of topics, waste news in Jakarta published by *detik.com* contains quite diverse topics. Among them is Topic 10, which is related to discovering waste goods in Jakarta. The group of topics 1 and Topic 9 reported the role of the DKI Jakarta government in handling waste is quite often reported. It can be seen with the many words of government and government officials loaded along with words related to waste.

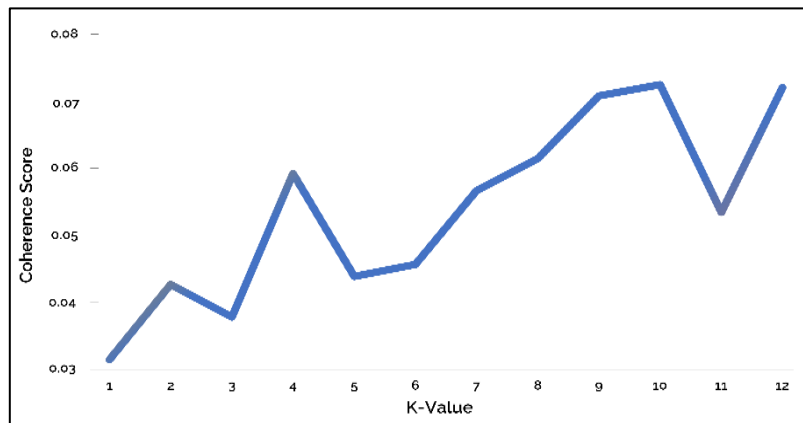


Figure 4. Coherence score results tuning parameters k-topic latent dirichlet allocation

Keywords related to the problem of waste in general and specification of the waste problem in Jakarta can be seen in Topics 2, 3, 4, 6, 7, and 8. Some specific issues, such as related to the problem of plastic waste in Jakarta, are also contained in news articles published by *detik.com*. In addition, the problem of waste in restaurants and the discovery of waste on the street due to the mass gathering can highlight how the source of waste problems in Jakarta has sprung up. The existence of waste that is not managed correctly triggers the emergence of news in group Topic 5, which impacts the problem of flooding in Jakarta. From the news that has been done processing and modeling, it can also be known that news related to waste handling programs and waste cleaning activities in Jakarta have also been reported. Education through this news is crucial because it can add insight and information to the community that waste in Jakarta is worth regretting and needs special attention from every level of society.

Using the VAR, the google trends keywords used in research (*sampah*, reuse, reduce, recycle and waste) after being given shock (in the form of information/campaign about waste) will have an impact on the community in finding the information after the next four weeks. It means that the information provided by online media (websites or online news) is slow to be received by the public. The impact of the news about the

3R given is delayed. For example, in terms of online media, waste as the cause of Jakarta's flooding in early January 2020 was reported by many online media in February 2020. Whereas as Hagi and Wright [16] mentioned, the Google platform can be a multi-side platform between information users and information providers meet, which is expected to occur a break in the time of acceptance.

Table 3. List of topics and distribution of latent dirichlet allocation topic modeling results

Topic	Top 10 word distributions	Topic discussion
T1	capital city, city, government, jakarta, waste, rupiah, anies, jakarta_capital city, wake up, funds	Related to Jakarta keywords in general
T2	waste, jakarta, clean, assignment, nas, location, transport, throw away, road, truck	Related to waste cleaning activities in Jakarta
T3	house, eat, healthy, Jakarta, isolation, sick, family, rupiah, mother, until	Related to the problem of waste in the restaurant
T4	plastic, waste, environment, indonesia, waste, bags, plastic_waste, arrange, community, plastic_bags	Related to the problem of plastic waste
T5	flood, water, jakarta, river, anies, rain, waste, clean, house, capital city	Related to the impact of waste on floods in Jakarta
T6	road, mass, burn, jakarta, car, driving, location, action, waste, assignment	Related to waste from the masses on the street
T7	indonesia, program, community, effort, environment, price, work, economics, develop, government	Related to environmental programs in Indonesian society
T8	indonesia, waste, jakarta, hook, people, real, islam, country, money, road	Related to waste in Jakarta in general
T9	jakarta, capital city, anies, governor, waste, city, jakarta_capital city, sandiaga, people, governor_capital city	Related to the role of the Jakarta government in handling waste
T10	jakarta, meet, safe, waste, saleable, west, ban, goods, report, result	Related to waste in Jakarta

The internet users in households in Indonesia in the last five years have increased rapidly, reaching 78.18%, and the growth of internet users is in line with the growth of mobile phone users [46]. In the period 2016 to 2020, it was noted that the percentage of the population accessing the internet increased from 25.37% to 53.73%. The increase does not necessarily encourage the public to find waste management information quickly. People need time to get that information. So that the increase in internet users cannot be used as a reference in the receipt of information, especially waste management information.

Google's search engine algorithm that influences certain content will appear. It means that competition in determining keywords from the content used in the news also impacts user searches related to these keywords. This study is only limited to 5 words used as keywords. It is a research opportunity in the future to research essential keywords related to waste management. While the topics generated from online news use the LDA method, no one uses keywords reduce, reuse and recycle (3R). The 3R behavior promotion or campaign is one of the main ways in reducing waste related to policies increasing public participation and contribution to reduce waste [47]. Campaigns should improve waste management habits that more emphasize 3R behavior. Furthermore, LDA results only classsify online news into common problem topics. It does not lead to the topic of waste management specifically.

This results in delays in information reaching users (the public), in addition to the three critical waste management keywords (reuse, reduce, and recycle) being absent from online news. The delay is allegedly due to a lack of user awareness of the consequences of improper waste management and less widespread publicity about waste management. Laziness and recklessness become the dominant factor that is not good in waste management [48]. The knowledge gap between young people and the elderly about the environment impacts problems in waste management [49]. The methods of delivery of waste management programs can be through online mass media, social media, training, websites, education, newspapers, government policies, and NGOs [50]. Online media platforms are becoming critical with the rapid delivery of information in the digital age. So that public interest in seeking information about waste management can increase significantly. It requires a strong synergy between the government, academia, society, NGOs, private sectors, and media called Penta helix collaboration [51], [52].

It can be known from the results of LDA on topics 2 and 6 that the critical role of the government in addressing waste is a policymaker in waste management. Based on Law of Republic of Indonesia No. 18 of 2008 on waste management and government regulation of Republic Indonesia No. 81 of 2012 on the management of household waste and another waste akin to household waste, the regulations should have a good impact, but the reality shows that the regulations are not apply maximally yet. Cooperation is needed between policyholders (central, regional, private, and community governments) so that the community gets policies related to waste management better than before [53].

#### 4. CONCLUSION

The government takes four weeks to communicate its policies to the public via online media. Because of the delay in receiving information, it is hoped that news through online media will pay more attention to the supply and demand for waste management information. Thus, if the keywords used in creating news content are adequate, the public, as Google platform users, will find it easier to find information on waste management. A flood, for example, will generate extensive media coverage of waste-related issues. There is a need for further research into the selection of dominating or important terms in the description of waste management, particularly in the context of 3R behavior. Solving problems with unstructured data, such as that found in online news articles or other data sources related to waste management, is also possible using the deep learning approach. In addition, a survey of five collaborators, known as Penta helix, is needed to combine field data with the results of the Penta helix. Surveying each of these groups will help us figure out how best to raise awareness, what policies should be put in place, what kind of information should be disseminated to the public, and how the private sector can help support waste reduction campaigns and management. It will be interesting to watch how waste management is linked to the spatial layout of the community in Indonesia, given the country's size and diversity of people.

#### ACKNOWLEDGEMENTS

The authors would like to thank Politeknik Statistika STIS and Universitas Negeri Jakarta for contributing to the publication of this article. Thank you to the reviewers for giving very helpful feedback that made it possible for this paper to be published.




#### REFERENCES

- [1] E. Friedberg and M. E. Hilderbrand, *Observing policy-making in Indonesia*. Observing Policy-Making in Indonesia, Springer Singapore, pp. 1-232, 2016, doi: 10.1007/978-981-10-2242-5.
- [2] A. Brotosusilo and D. Handayani, "Dataset on waste management behaviors of urban citizens in large cities of Indonesia," *Data in Brief*, vol. 32, 2020, doi: 10.1016/j.dib.2020.106053.
- [3] S. Nanda and F. Berruti, "Municipal solid waste management and landfilling technologies: a review," *Environmental Chemistry Letters*, vol. 19, no. 2, pp. 1433–1456, 2021, doi: 10.1007/s10311-020-01100-y.
- [4] A. Brotosusilo, D. Utari, H. A. Negoro, A. Firdaus, and R. A. Velentina, "Community empowerment of waste management in the urban environment: More attention on waste issues through formal and informal educations," *Global Journal of Environmental Science and Management*, vol. 8, no. 2, pp. 209–224, 2022, doi: 10.22034/GJESM.2022.02.05.
- [5] A. Maghmoumi, F. Marashi, and E. Houshfar, "Environmental and economic assessment of sustainable municipal solid waste management strategies in Iran," *Sustainable Cities and Society*, vol. 59, no. April, p. 102161, 2020, doi: 10.1016/j.scs.2020.102161.
- [6] S. De and B. Debnath, "Prevalence of Health hazards associated with solid waste disposal- a case study of Kolkata, India," *Procedia Environmental Sciences*, vol. 35, pp. 201–208, 2016, doi: 10.1016/j.proenv.2016.07.081.
- [7] P. Ten Brink *et al.*, *Guidelines on the Use of Market-based Instruments to Address the Problem of Marine Litter*. 2009.
- [8] C. Alexander, C. Smaje, R. Timlett, and I. Williams, "Improving social technologies for recycling," *Proceedings of Institution of Civil Engineers: Waste and Resource Management*, vol. 162, no. 1, pp. 15–28, 2009, doi: 10.1680/warm.2009.162.1.15.
- [9] S. Das, S. H. Lee, P. Kumar, K. H. Kim, S. S. Lee, and S. S. Bhattacharya, "Solid waste management: Scope and the challenge of sustainability," *Journal of Cleaner Production*, vol. 228, pp. 658–678, 2019, doi: 10.1016/j.jclepro.2019.04.323.
- [10] P. Jiang, Y. V. Fan, and J. J. Klemeš, "Data analytics of social media publicity to enhance household waste management," *Resources, Conservation and Recycling*, vol. 164, no. June 2020, p. 105146, 2021, doi: 10.1016/j.resconrec.2020.105146.
- [11] A. Brodeur, A. E. Clark, S. Fleche, and N. Powdthavee, "COVID-19, lockdowns and well-being: evidence from Google trends," *Journal of Public Economics*, vol. 193, p. 104346, 2021, doi: 10.1016/j.jpubeco.2020.104346.
- [12] H. R. Jamali and S. Asadi, "Google and the scholar: The role of Google in scientists' information-seeking behaviour," *Online Information Review*, vol. 34, no. 2, pp. 282–294, 2010, doi: 10.1108/14684521011036990.
- [13] J. Brophy and D. Bawden, "Is Google enough? Comparison of an internet search engine with academic library resources," *Aslib Proceedings: New Information Perspectives*, vol. 57, no. 6, pp. 498–512, 2005, doi: 10.1108/00012530510634235.
- [14] A. Hagi, "Two-sided platforms: Product variety and pricing structures," *Journal of Economics and Management Strategy*, vol. 18, no. 4, pp. 1011–1043, 2009, doi: 10.1111/j.1530-9134.2009.00236.x.
- [15] A. Hagi and D. Spulber, "First-party content and coordination in two-sided markets," *Management Science*, vol. 59, no. 4, pp. 933–949, 2013, doi: 10.1287/mnsc.1120.1577.
- [16] A. Hagi and J. Wright, "Multi-sided platforms," *International Journal of Industrial Organization*, vol. 43, pp. 162–174, 2015, doi: 10.1016/j.ijindorg.2015.03.003.
- [17] S. C. Wilson, D. A. Daar, S. Sinno, and S. M. Levine, "Public interest in breast augmentation: analysis and implications of Google trends data," *Aesthetic Plastic Surgery*, vol. 42, no. 3, pp. 648–655, 2018, doi: 10.1007/s00266-017-1038-1.
- [18] N. Mahroum *et al.*, "Capturing public interest toward new tools for controlling human immunodeficiency virus (HIV) infection exploiting data from Google Trends," *Health Informatics Journal*, vol. 25, no. 4, pp. 1383–1397, 2019, doi: 10.1177/1460458218766573.
- [19] J. D. Tijerina, S. D. Morrison, D. G. Vail, G. K. Lee, and R. Nazerali, "The utility of Google trends data for analyzing public interest in breast procedures," *Annals of plastic surgery*, vol. 82, no. 5S Suppl 4, pp. S325–S331, 2019, doi: 10.1097/SAP.0000000000001806.
- [20] G. C. Wu *et al.*, "Leveraging Google trends to investigate the global public interest in rheumatoid arthritis," *Rheumatology international*, vol. 39, no. 8, pp. 1439–1444, 2019, doi: 10.1007/s00296-019-04297-6.
- [21] D. C. Landy, B. P. Chalmers, T. J. Utset-Ward, and M. P. Ast, "Public interest in knee replacement fell during the onset of the COVID-19 pandemic: a Google trends analysis," *HSS Journal*, vol. 16, pp. 24–28, 2020, doi: 10.1007/s11420-020-09794-0.






- [22] S. Kardeş, "Public interest in spa therapy during the COVID-19 pandemic: analysis of Google Trends data among Turkey," *International Journal of Biometeorology*, vol. 65, no. 6, pp. 945–950, 2021, doi: 10.1007/s00484-021-02077-1.
- [23] E. S. Silva, H. Hassani, D. Ø. Madsen, and L. Gee, "Googling fashion: Forecasting fashion consumer behaviour using Google Trends," *Social Sciences*, vol. 8, no. 4, 2019, doi: 10.3390/socsci8040111.
- [24] D. Seth, H. Gittleman, J. Barnholtz-Sloan, and J. S. Bordeaux, "Association of socioeconomic and geographic factors with google trends for tanning and sunscreen," *Dermatologic Surgery*, vol. 44, no. 2, pp. 236–240, 2018, doi: 10.1097/DSS.0000000000001324.
- [25] H. Hassani, X. Huang, S. Macfeely, and M. R. Entezarian, "Big data and the united nations sustainable development goals (UN SDGs) at a glance," *Big Data and Cognitive Computing*, vol. 5, no. 3, 2021, doi: 10.3390/bdcc5030028.
- [26] K. Kabirifar, M. Mojtahedi, C. Wang, and V. W. Y. Tam, "Construction and demolition waste management contributing factors coupled with reduce, reuse, and recycle strategies for effective waste management: A review," *Journal of Cleaner Production*, vol. 263, p. 121265, 2020, doi: 10.1016/j.jclepro.2020.121265.
- [27] I. Prawira, R. E. Irawan, Muslikhin, L. W. Evelina, and M. Rizkiyansyah, "The media agenda-setting on Indonesian forest fires: Content analysis study on three Indonesia online news media," *IOP Conference Series: Earth and Environmental Science*, vol. 729, no. 1, 2021, doi: 10.1088/1755-1315/729/1/012082.
- [28] Prayudi, S. Fatolah, and J. Shafira, "Covid-19 on online news media: A discourse analysis of Indonesian government crisis management on kompas.com," *Jurnal Komunikasi: Malaysian Journal of Communication*, vol. 37, no. 3, pp. 38–53, 2021, doi: 10.17576/JKMJC-2021-3703-03.
- [29] R. Djalante, J. Jupesta, and E. Aldrian, *Correction to: Climate change research, policy and actions in Indonesia (springer climate, 10.1007/978-3-030-55536-8)*, 2021.
- [30] S. Widati and A. S. A. Nordin, "Mental health during COVID-19: An overview in online mass media Indonesia," *Asia-Pacific Journal of Public Health*, vol. 34, no. 2–3, pp. 300–301, 2022, doi: 10.1177/10105395211063169.
- [31] P. Jariyapan, J. Singvejsakul, and C. Chaiboonsri, "A machine learning model for healthcare stocks forecasting in the US stock market during COVID-19 period," in *Journal of Physics: Conference Series*, 2022, vol. 2287, no. 1, doi: 10.1088/1742-6596/2287/1/012018.
- [32] M. K. Das, D. Singh, and S. Sharma, "Media news on vaccines and vaccination: The content profile, sentiment and trend of the online mass media during 2015–2020 in India," *Clinical Epidemiology and Global Health*, vol. 10, no. October 2020, p. 100691, 2021, doi: 10.1016/j.cegh.2020.100691.
- [33] A. Kang, L. Ren, C. Hua, M. Dong, Z. Fang, and M. Zhu, "Stakeholders' views towards plastic restriction policy in China: Based on text mining of media text," *Waste Management*, vol. 136, no. September, pp. 36–46, 2021, doi: 10.1016/j.wasman.2021.09.038.
- [34] K. Arts *et al.*, "Supply of online environmental information to unknown demand: the importance of interpretation and liability related to a national network of river level data," *Scottish Geographical Journal*, vol. 131, no. 3–4, pp. 245–252, 2015, doi: 10.1080/14702541.2014.978809.
- [35] K. Pinyopornpanish, W. Jiraporncharoen, K. Thaikla, K. Yoonut, and C. Angkurawaranon, "Sedative and analgesic drugs online: a content analysis of the supply and demand information available in Thailand," *Substance Use and Misuse*, vol. 53, no. 4, pp. 641–647, 2018, doi: 10.1080/10826084.2017.1355386.
- [36] N. Askitas, "Google search activity data and breaking trends," *IZA World of Labor*, no. November, pp. 1–10, 2015, doi: 10.15185/izawol.206.
- [37] N. Woloszko, "Tracking activity in real time with Google Trends," *OECD iLibrary*, no. 1634, p. 54, 2020. doi: <https://doi.org/10.1787/18151973>.
- [38] R. Morgan, "detik.com, KOMPAS.com and LIPUTAN6 are Indonesia's leading news websites," *Roy Morgan*, 2018. <http://www.roymorgan.com/findings/7622-top-10-indonesian-news-websites-march-2018-201806172227%0Ahttp://files/489/7622-top-10-indonesian-news-websites-march-2018-201806172227.html>.
- [39] T. Gonçalves and P. Quaresma, "Evaluating preprocessing techniques in a text classification problem," *Unisinos*, pp. 841–850, 2005, [Online]. Available: <http://www.research.att.com/>.
- [40] E. R. Ziegel and W. Enders, *Applied Econometric Time Series*, Technometrics, vol. 37, no. 4, 1995.
- [41] L. Charfeddine and M. Kahia, "Impact of renewable energy consumption and financial development on CO2 emissions and economic growth in the MENA region: A panel vector autoregressive (PVAR) analysis," *Renewable Energy*, vol. 139, pp. 198–213, 2019, doi: 10.1016/j.renene.2019.01.010.
- [42] D. Van Dinh, "Impulse response of inflation to economic growth dynamics: VAR model analysis," *Journal of Asian Finance, Economics and Business*, vol. 7, no. 9, pp. 219–228, 2020, doi: 10.13106/JAFEB.2020.VOL7.NO9.219.
- [43] C. Bird, T. Menzies, and T. Zimmermann, *The Art and Science of Analyzing Software Data*. Elsevier, 2015.
- [44] D. M. Blei, A. Y. Ng, and M. T. Jordan, "Latent dirichlet allocation," *Advances in Neural Information Processing Systems*, vol. 3, pp. 993–1022, 2002.
- [45] S. H. Mohammed and S. Al-Augby, "LSA & LDA topic modeling classification: Comparison study on E-books," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, pp. 353–362, 2020, doi: 10.11591/ijeecs.v19.i1.pp353-362.
- [46] A. D. R. Adawiah and T. Sutarsih, *Indonesian Telecommunication Statistics 2014*. Jakarta, Indonesia: Central Bureau of Statistics - Statistics Indonesia, 2014.
- [47] A. Minelgaitė and G. Liobikienė, "Waste problem in European Union and its influence on waste management behaviours," *Science of the Total Environment*, vol. 667, pp. 86–93, 2019, doi: 10.1016/j.scitotenv.2019.02.313.
- [48] A. Sewak, S. Deshpande, S. Rundle-Thiele, F. Zhao, and R. Anibaldi, "Community perspectives and engagement in sustainable solid waste management (SWM) in Fiji: A socioecological thematic analysis," *Journal of Environmental Management*, vol. 298, no. August, p. 113455, 2021, doi: 10.1016/j.jenvman.2021.113455.
- [49] J. K. Debrah, D. G. Vidal, and M. A. P. Dinis, "Raising awareness on solid waste management through formal education for sustainability: A developing countries evidence review," *Recycling*, vol. 6, no. 1, pp. 1–21, 2021, doi: 10.3390/recycling6010006.
- [50] G. B. Zamri *et al.*, "Delivery, impact and approach of household food waste reduction campaigns," *Journal of Cleaner Production*, vol. 246, 2020, doi: 10.1016/j.jclepro.2019.118969.
- [51] R. E. Caraka, M. Noh, R. C. Chen, Y. Lee, P. U. Gio, and B. Pardamean, "Connecting climate and communicable disease to penta helix using hierarchical likelihood structural equation modelling," *Symmetry*, vol. 13, no. 4, 2021, doi: 10.3390/sym13040657.
- [52] E. P. Purnomo, F. Aditya, A. T. Fathani, L. Salsabila, and D. Rachmawatie, "Penta-Helix approach as a strategy to recovery tourism in Bali due to COVID-19 pandemic," *ACM International Conference Proceeding Series*, no. C, pp. 122–127, 2021, doi: 10.1145/3484399.3484417.
- [53] A. S. Suryani, "Community perception and analysis willingness to Pay towards the public paid plastic bag policy perception and willingness to study in Jakarta and Bandung," *Kajian*, vol. 21, no. April 2012, pp. 359–376, 2017, doi: <https://jurnal.dpr.go.id/index.php/kajian/article/view/784>.




**BIOGRAPHIES OF AUTHORS**

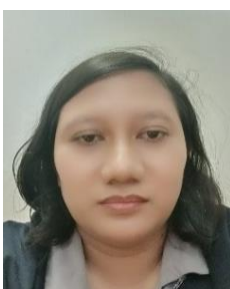
**Robert Kurniawan**    is a researcher at the Statistics Polytechnic STIS Jakarta who focuses on social science, disaster management and the environment, and big data. He is currently pursuing a doctorate in population and environmental education at the State University of Jakarta. he is actively conducting research and writing reference books with ISBNs such as Easy Understanding of Nonparametric Statistics in the Health Sector and Regression Analysis using R. He hopes that big data can provide new insights into waste management. He can be contacted at email: robertk@stis.ac.id.






**Prof. Dr. Agung Purwanto**    received his doctorate in population and environmental education from Jakarta State University in 2012 and was inaugurated as an environmental education professor in 2022. Currently, he serves as the Director of the Population and Environmental Education Study Program at Jakarta State University. ICT and big data are two of the subjects taught. He can be contacted at email: agungpurwanto@unj.ac.id.






**Anugerah Karta Monika**    graduated from Padjadjaran University in Applied Statistics for under graduate degree and Applied Economics for master degree. Well trained econometrician, specialization in official statistics, with more than 20 years working experience in Statistics Indonesia and currently working as Secretary of Academic Senate of Politeknik Statistika STIS in Indonesia. She can be contacted at email: ak.monika@stis.ac.id.






**Krismanti Tri Wahyuni**    is currently a lecturer at Politeknik Statistika STIS. She holds a M.Si. degree in economics from the IPB University. She has authored or coauthored more than 15 publications. Her research interests include economics, econometrics, and official statistics. He can be contacted at email: krismanti@stis.ac.id.



**Muhammad Yunus Hendrawan**    received the S.Tr.Stat (Bachelor of Applied Statistics) in Computational Statistics from Politeknik Statistika STIS, Jakarta. He is currently working as statistician at Directorate of Census and Survey Methodology Development in Statistics Indonesia (Badan Pusat Statistik). His research interests include statistics, big data, text mining, and official statistics. He can be contacted at email: yunus.hendrawan@bps.go.id.



**Mohamad Andrian Isnaeni**    is currently young planner at local planning, research and development agency of Sragen Regency, Central Java Province. He graduated from Pantheon Sorbonne University and Diponegoro University in urban planning for master degree. His research interests include local economic development, poverty, and statistics. He can be contacted at email: andriansrg@gmail.com.