

From Query to Communication: A Data-Driven Study of Indonesians Google Search Behavior

Assyari Abdullah¹, Yefni², Yasril Yazid³, Dwi Sutri Astuti⁴, Edison⁵,
Muhammad Soim⁶, Akmal Khairi⁷

Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru,
Indonesia^{1,2,3,4,5,6}

Sekolah Tinggi Agama Islam Imam Asy Syafii Pekanbaru, Indonesia⁵

Correspondence Author : assyariabdullah@uin-suska.ac.id

Abstract

The purpose of this article is to examine how people in Indonesian society use the Google Search Engine to find information as part of the digital public communication process. Exploratory Data Analysis (EDA) is the foundation of the Big Data Analytics strategy used in this study. Using data from the Kaggle Machine Learning and Data Science Community Portal, the study examines the "Google Searches of Indonesian People 2019-2024" dataset. Data collection, data cleaning (managing missing values and outliers), descriptive statistical analysis (mean, median, mode), data visualisation (histogram, scatter plot), finding patterns and correlations between variables, and interpretation of results to make preliminary conclusions are the main steps in exploratory data analysis (EDA), which aims to comprehend the structure and features of the data prior to further modelling. The results of this study offer scholars important new information: sports, entertainment, politics, and economics are the most searched topics in Indonesia. Bar charts, pie charts, and work cloud are examples of data visualisations that show how information searches reflect not just cognitive requirements but also social anxieties, cultural goals, and collective concerns. The results of this study undoubtedly support the notion that Google Search functions as a silent communication tool where people actively participate in the process of negotiating social orientation and meaning on an individual basis but with a foundation in collective dynamics. This study also demonstrates the importance of a data-driven approach in the study of digital communication and offers an interactive framework for reading Google Search as a hidden yet significant public discursive space. This research provides a theoretical contribution in expanding the study of big data-based communication that is practically useful for policymakers, media practitioners, academics who want to understand the dynamics of digital public discourse specifically.

Keywords: Query, Digital Communication, Google Search, Data-Driven Communication, Exploratory Data Analysis (EDA).

INTRODUCTION

In the digital era, searching for information through search engines has become the main activity of the global community in accessing, understanding, and responding to the social environment (Efthimiadis et al., 2009; Riyaz, 2017; White et al., 2015). Google Search, as the most dominant search engine, has transformed the way individuals interact with information and shape meaning on social, political, cultural, and economic issues (Abdullah, Jayus, et al., 2024; Dijk, 2020; Halavais, 2013, 2018; Lewandowski, 2015; Morozov, 2011; Rowlands et al., 2008; Zimmer, 2008). Search is no longer just the process of finding information, but has become part of the process of public communication and identity negotiation in the digital space (Hargittai & Walejko, 2008; Huang & Efthimiadis, 2009; Masimengo, 2025; Na & Park, 2012; Taylor, 2005; Treem & Leonardi, 2013). In Indonesia, the use of Google Search reflects not only the thirst for information, but also the response to events that drive public opinion online (Krismayani & Mafar, 2024; Jayus et al., 2024)

The increasing dependence of people on Google Search during the COVID-19 pandemic, elections, natural disasters, the economy and viral phenomena confirm that online search activities now have a significant communicative weight (Abdullah et al., 2024; Coiro & Dobler, 2007; Donadelli, 2015; Götz & Knetsch, 2019; White & Horvitz, 2009; Jansen et al., 2008; Eysenbach, 2002; Sumaiyah et al., 2024). When people are looking for information about the "best vaccine", "2024 presidential candidate", or "fuel price increase", they not only want to know the factual answer, but are also navigating the broader social discourse (Cornell University, 2020; Fallows, 2005; Niu & Kelly, 2014; Rochmawati & Nurmandi, 2020). Google Search has become part of a "cyber communication system" that connects individuals to global narratives, opinions, and discourse despite the nature of search being individual and hidden (Hassan, 2008; Morozov, 2011; Rowlands et al., 2008).

Research related to search engines has recently experienced significant developments across various disciplines, including communication science and information science, and digital sociology (Showkat & Gull, 2020; Acharya, 2024). Jansen & Spink (2006) explains that information search behavior reflects highly complex and contextual patterns of information needs. Halavais (2013) see Google as a cultural actor that shapes public perception. Meanwhile, the results of research by Höchstötter & Lewandowski (2009) and Granka (2010) Highlight how search results affect opinion formation and information selection. In the health sector, Eysenbach (2002) reveals that Google Search is the primary reference for health information, although it is not always accurate. However, in communication studies, the study of Google Search as a form of public communication is still limited and has not been explored, especially in the context of developing countries (Astra, 2025).

Most previous research has placed Google Search in the context of informational efficiency and search effectiveness, rather than as a reflective and narrative public communication channel (Halavais, 2013, 2018; Jansen et al., 2008; Jansen & Spink, 2006; Treem & Leonardi, 2013; White & Horvitz,

2009; Zimmer, 2008). There is still little research that uses the dataset of public search results as a basis for analyzing the dynamics of digital communication quantitatively and interpretatively (Jayus et al., 2025) . In addition, there has been no comprehensive research that uses a data-driven exploratory analysis approach to Google Search data among Indonesian people to understand how search behavior shapes social discourse. This opens up opportunities for interdisciplinary studies that combine data analysis with digital communication theory (Abdullah, Yazid, et al., 2024; Krismayani & Mafar, 2024).

This research offers an important contribution in expanding the horizons of digital communication studies by placing Google Search activities as communicative practices. Theoretically, this research enriches the framework of thinking about information-seeking as communication and brings together the communication systems approach with individual search behavior (Abdullah, 2020; Castells, 2009, 2011; Hargittai & Walejko, 2008; McQuail, 2000, 2010; Sundar & Limperos, 2013; Taylor, 2005; Treem & Leonardi, 2013). Practically, the findings of this study can be used by journalists, policymakers, and media researchers to read the dynamics of public attention in real-time. This information is important in the post-truth era, when stories in the news and people's own searches work together to shape how people see things.

This article analyzes search trends for important issues such as politics, health, entertainment, and the economy using the "Google Searches of Indonesian People 2019-2024" dataset. This study will carry out Exploratory Data Analysis (EDA) of the most searched phrases by the people that illustrates the digital communication process that takes place in a personal context but affects the entire population (Coiro & Dobler, 2007; Lewandowski, 2015; Napoli, 2003, 2011, 2014; Van Dijck, 2013; White & Horvitz, 2009). The primary objective of this research is to determine how the concerns, goals and orientations of the public which change according to the social dynamics reflect in the information that is sought.

The study of data-driven communication studies advances by an integration of big data analysis and interpretive methods. This study found that the responses placed by the informants constitute a discourse that is not merely a reaction to the problem of ignorance, but also an affective and strategic communicative act. This study can also contribute as a useful reference for the advancement of research on digital communication in Indonesia that is more empirically grounded and reflects the actual communicative habits of the Indonesians.

The current study aims to use an interpretative framework that may help to improve a field of data-driven communication studies through big data analytic. In addition to its analytical and conceptual contributions to the study of communication, knowledge is not merely an escape from ignorance, it is also an emotional reaction and a strategic form of communication (Abdullah et al., 2025; Castells, 2011; Eggensperger & Redcross, 2018; Grivy, 2025; Luke et al., 2024; Martin et al., 2024; McQuail, 2010; Sundar & Limperos, 2013; Treem & Leonardi, 2013; Zimmer, 2008). The study can also be a useful tool for

improving digital communication research in Indonesia that is based on real data and shows how Indonesians really communicate.

METHODOLOGY

This research utilizes a Big Data Analytics framework with Exploratory Data Analysis (EDA) technique to analyse the information searches pattern on the Google Search tools or google.co.id owned by the Indonesian people in the year range of 2019 – 2024. It helps the researchers to discover, scrutinise and analyse a data set which is large enough to learn something significant about digital communicative search behaviour (Kitchin, 2014; Lohr, 2012; Mayer-Schönberger, 2013; Provost & Fawcett, 2013; Ward & Barker, 2013).

The data set used in this study was downloaded from Kaggle with the title of “Google Searches of Indonesian People 2019-2024” by Nugrah Nurrohman (Nurrohman, 2024). This dataset consists of hundreds of records that show the most common search terms by Indonesians for the last five years. The information has been divided into categories, namely politics, economy, entertainment, and health. This set of data provides search intensity per word and category (no date associated with the individual entries) giving the impression of evolution over time in digital public conversation.

The research used Google Colaboratory (Google Colab) to ensure effective processing of large-scale data (Google, 2025). Python is used as a programming language which uses popular analytical libraries like pandas, matplotlib, seaborn, wordcloud (Matplotlib Development, 2025; McKinney, 2010; Mueller, 2020; Washbasin, 2021). To clean and transform text data, the pre-processing stage in any analytical procedure. The process of turning the search volumes into numbers is termed transformation. There is grouping and aggregation of the data based on themes and frequencies. Visualization like bar charts, pie charts and word clouds etc... is another factor to showcase important trends in searches (Rahm & Do, 2000; Wickham, 2014).

The first steps in Data Analysis are to collect the data, clean it (remove missing values and outliers), calculate descriptive statistics (mean, median and mode), visualize it (using histograms and scatter plots), detect patterns and correlations in the variables and analyze the outcome to draw some preliminary conclusions; all this is by understanding the structure and properties of the data before building any models. The approach will focus on the most popular searches, themes and interpretations with an emphasis on digital public communication. All this will take place by comprehensive EDA (AC, 2019; Das, 2025, 2025; Tukey, 2005; Weng, 2021). The research categorizes search phrases into thematic classifications that represent social, political, economic, and cultural dimensions of society. The classification is carried out manually based on the meaning of each term which is then visualized to create a digital discourse map of society (Alauddin et al., 2025; Lazer et al., 2014; Michel et al., 2011; Mohr & Bogdanov, 2013; Shen et al., 2025; St. Amant & Cohen, 1998).

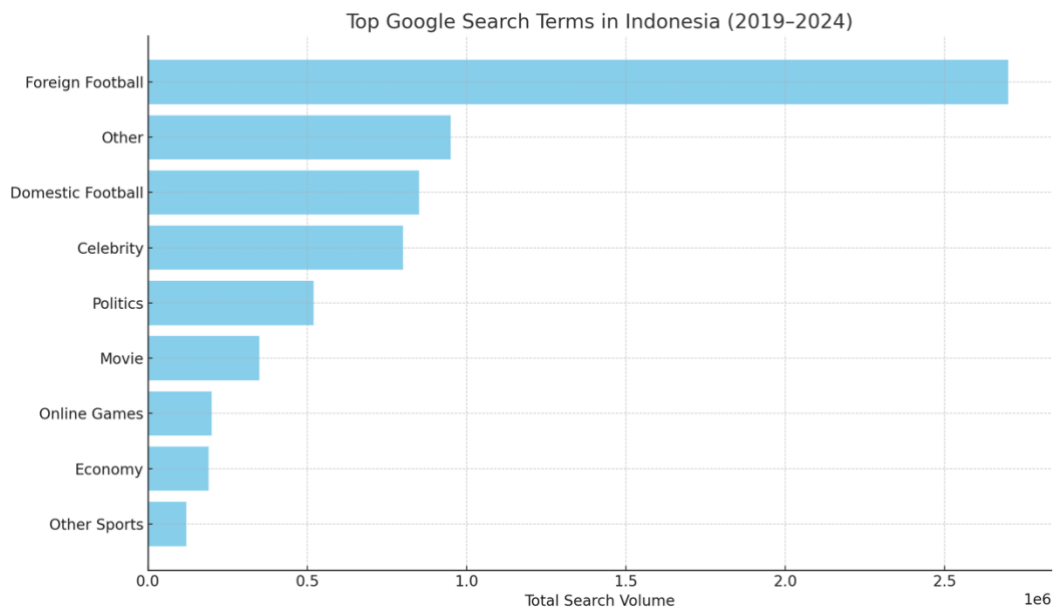
This allows the researchers to relate numerical data from digital searches to meaning in human communication. Outcomes-based result-oriented methodologies relevant to communication studies driven by big data, requiring data analysis for the understanding of production, circulation and consumption of information in contemporary digital society (Boyd & Crawford, 2012; Bruns, 2019; Napoli, 2014).

RESULTS AND DISCUSSION

Overview of Google Search Data (2019-2024)

The preliminary analysis of the "Google Searches of Indonesian People 2019-2024" dataset indicates that the search patterns of Indonesians over the past five years have been predominantly influenced by prevalent themes related to sports, entertainment, and political matters. According to the image below, terms like "Overseas Ball", "Domestic Ball" and "Artist" rank well. There are around more than 800011 total monthly searches for the first, and about 27 million for the last one. The public digital search behavior in Indonesia is prominently influenced by the interplay between sportainment and fame as digital mass media trends show in this period.

Fig. 1 Top 15 Google Search Terms in Indonesia (2019—2024)



The community further demonstrated their investment in understanding the pressing affairs of the day and in chasing alternative forms of entertainment with categories like "Politics," "Movies," "Online Games. The data not only displays informative and recreational search habits but also a significant finding is Google Search, which has become a major tool for solving current societal needs. As the net played a part in counteracting censorship by

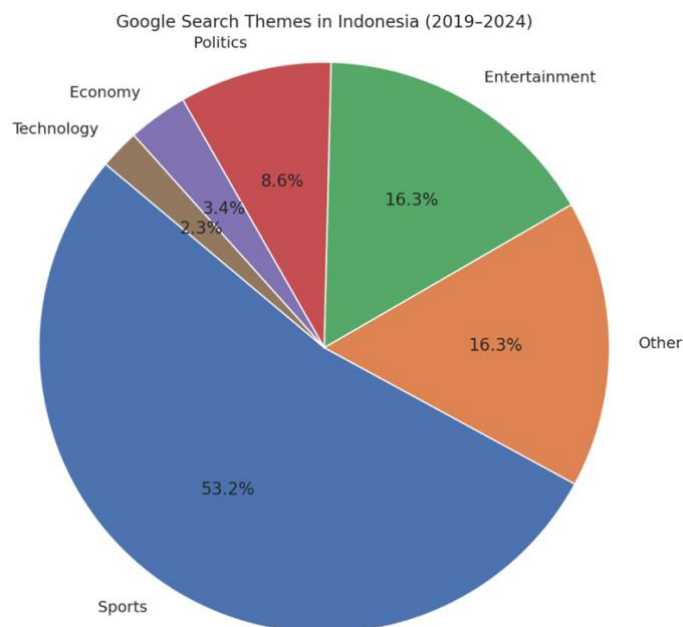
not providing internet identities to groups, Web search engines increasingly, and surely, function alike and search tools for individuals.

Increased Search volume of a term is an algorithmic phenomenon. It takes place when many people search for it again and again. Consequently, a search mapping as a part of EDA is descriptive and gives background study for later communicative analysis. The study shows that understanding search as a form of communication implies that category mapping and story exploration should be theme based.

Top Search Categories and Their Communicative Significance

Thematic categorization of a large number of search phrases indicates that categories related to sports, were the ones with the highest search volume; followed through entertainment, politics and business. More than 30%, overseas ball, domestic ball and such types of search have share in the sports sector. Football and other sport do not only serve entertainment esmercial but also has become a part of a digital culture that shapes the identity and discourse of the Indonesian people.

Fig. 2 Google Search Themes in Indonesia (2019–2024)



The entertainment category, featuring the likes of Artist and Film, is a measure of how the public feels about celebrities and stories. The "artist" or "film" who is sought in the constitution of communication is but a symbolic consumption; it expresses goal, social identity and digital fame. At the same

time, political classifications like “Politics” and “Presidential Candidates” suggest that people are going to Google Search to look for material related to government, policy, and electoral competition, especially before major events such as elections.

The technology and online gaming sectors are increasingly becoming part of the digital search landscape and suggest that digital media is not only an influential communication medium but also a platform for socializing. As such, the vital section of this classification is a reflection of the collective interests, the affected experiences, and the public communication practice that gets distributed through the personal search. Within the frame of digital communicating, searching is not simply the act of discovering information but rather a way of participating that takes place within discursive publics which together create social meaning.

Word Frequency and Dominant Public Concerns

The word cloud visualization below provides a visual representation of the terms most frequently searched by Indonesian people through Google Search during the period 2019 to 2024. The size of each word reflects the volume of searches, the larger the word displayed, the higher the search frequency. In this word cloud, words such as "Bola", "Artist", "Politics", "Film", and "Online Game" appear as the dominant elements, confirming the results of previous classifications that sports and entertainment are the main focus of the digital search of Indonesian people.

Fig. 3 Word Cloud of Top Google Search Terms in Indonesia



The dominance of certain terms in the word cloud also shows a collective concern for issues of a creative and political nature. People are not only looking for information for entertainment, but also to understand the ongoing social, political, and economic conditions. This shows that people's search behavior is a form of response to external dynamics both local and global. In the framework of communication, each term in the word cloud reflects the topics that make up the agenda setting in the digital space.

When you read terms such as “Other” (and similar), they refer to open-category data and indicate large undeclared searches. The wording in this statement is the first indicator of a new discursive space, something traditional media analysis may not have done. The frequency pattern in digital public communication can be perceived as an indicator of the cognitive, affective and social needs of society to respond to rapidly changing conditions that often occur in modern society.

Thematic Classification of Search Behavior

Subsequent analysis of the datasets reveal that each theme area had a collection of dominant phrases suggestive of different systems of public attention. “Football Abroad” and “Domestic Football” get the most search traffic across the different datasets. This tendency indicates that the Indonesian public has high interest in football, both domestic and foreign. This affirms the position of sports as an element of digital culture and a meeting point of collective communication through sports content and sportainment.

On the other hand, the entertainment category is dominated by terms such as "Artist", "Film", and "Drama", which indicate that public figures and popular media remain the main source of search in the context of relaxation, fandom, and social identification. The search for artists reflects the form of symbolic consumption and the role of the media as a link between society and the world of celebrities. It is also a reflection of "aspirational communication" in which internet users project interest and curiosity through the search for public figures.

Table 1. Top 9 Search Volume and Themes

Search Term	Total Volume	Theme
Foreign Football	2,794,902	Sports
Other	1,157,700	Other
Domestic Football	881,314	Sports
Celebrity	890,290	Entertainment
Politics	614,482	Politics
Movie	267,036	Entertainment
Online Games	160,310	Technology
Economy	320,448	Economy
Other Sports	106,760	Sports

The only political theme, then "Politics" and its derivatives are the main focus. This shows that the Indonesian people are quite active in seeking information about the dynamics of power, elections, and national political events, especially during periods of political contestation, both national, such

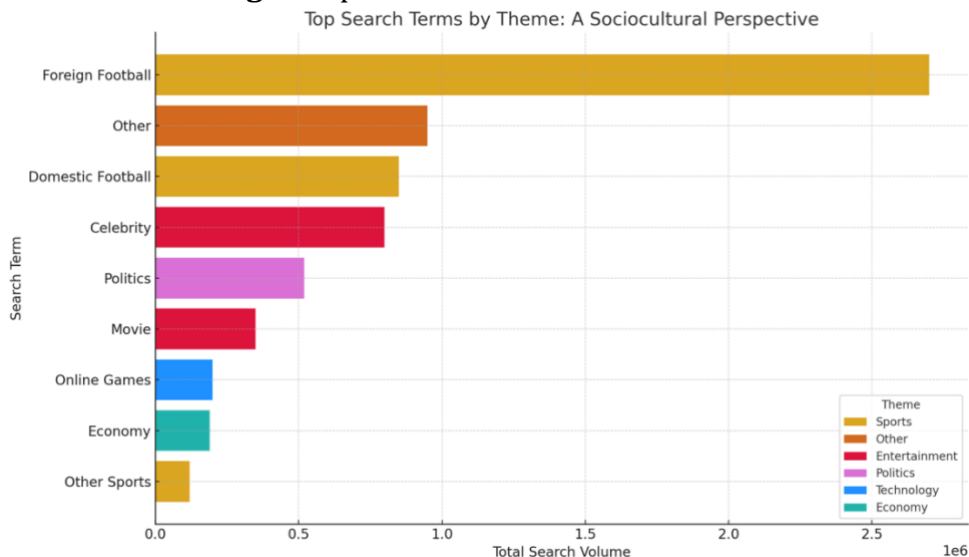
as the Presidential and Member of Parliament of the Republic of Indonesia and the election of regional heads. This phenomenon reflects that Google Search has become a major source in shaping citizens' political understandings, replacing the role of traditional mass media in some contexts.

Meanwhile, the "Other" and "Technology" classifications keep terms that cannot be directly incorporated into the main theme, but still show high relevance. Keywords in this category require some more analysis as they are largely indicative of emerging interests, trending topics or those which have not been classified systematically. The classification of Googling activities carried out by Indonesians, apparently, is not just one thing like the intention. However, it turns out to be varied and everything.

Google Search and Sociocultural Interpretation of Information

Searching the information through Google is a source of cognitive, social and cultural dimension of civilization. In the context of Indonesia, this dataset showed that people were in search of information not only in order to solve their problems, but also in order to adjust to social change, strengthen group identity, and influence perception of reality. We see large search queries for "overseas ball" and "artist", which prove that sports and celebs are not merely entertainment media, but manifest forms of social bonding and aspirational cultural signifiers.

Fig. 4 Top Search Terms a Sociocultural



Queries for terms like "Politics," "Presidential Candidate," and "Fuel Prices" indicate that the public utilizes Google Search as a resource to navigate social and political ambiguity. In a culture characterized by diminished faith in formal institutions, independent information retrieval serves as a mode of self-communication that supplants traditional information authorities (Castells, 2009; Lim, 2013). Google Search serves as both a repository of information

and a discursive platform where citizens collaboratively construct their comprehension of the nation, economy, and future.

Indonesia is a heterogeneous country, thus the variation in geographies, social economy, and culture also varies the search themes. Expressions like "Online Gaming" and "Korean Drama" are mostly associated with the young and urban middle class who are part of the digital global culture. On the other hand, "costs of essential goods" or "social support" related phrases may be more of interest to financially strained lower-middle-class citizens. Thus, the search data manifest social fragmentation and interaction of sectors of digital cultures in one shared search space.

The conclusion that can be drawn is that it indicates that Indonesians create their own social knowledge through web searches rather than just consume traditional media. According to Pink et al., 2016 Google search can be regarded as a form of what may be termed 'silent discourse' communication in the framework of digital ethnography. That is a type of engaged communication which is not taking place in a public forum but which expresses a shared sentiment on what the public considers is important, urgent or interesting. As such, it correctly represents a small portion of the broader social reality being constructed digitally.

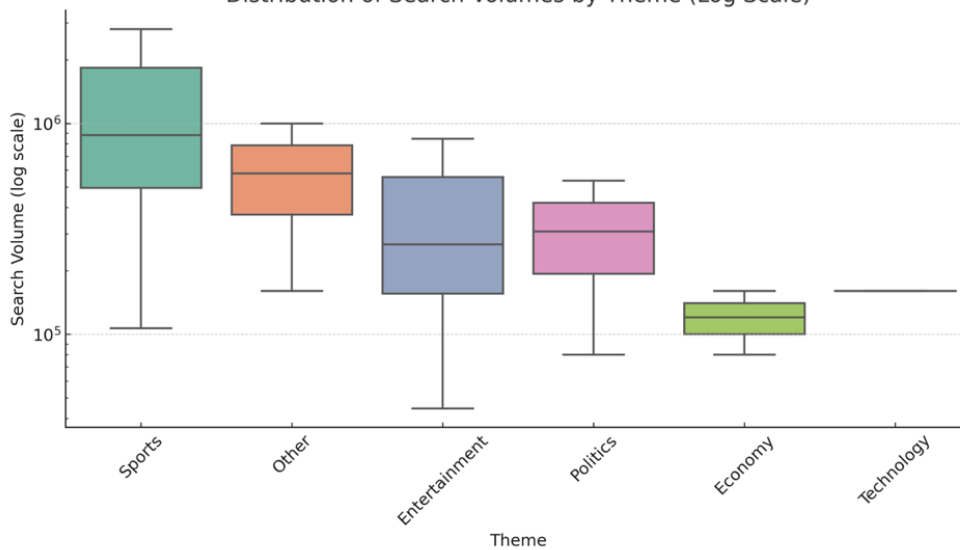
All of this data proves that Indonesian people's information-seeking behaviour is not random, but rather shaped by their own value systems, socio-cultural practices, and the economic-political pressures that they experience. Google now serves as a symbolic representation that captures the culture of society and acts as a medium of silent communication. It articulates the identity, curiosity, and collective worries of Indonesians.

Searching Information as a Communicative Practice in the Digital Realm.

The examination of the 2019-2024 Google Searches of Indonesian People shows that information search is not a neutral matter, but a coding act in the digital realm. Looking at communication in the present day, googling information is an individual meaning-making process, one that is essentially embedded in broader social processes (Boyd & Crawford, 2012; Castells, 2011; Jayus et al., 2024; McQuail, 2010; Sundar & Limperos, 2013; Treem & Leonardi, 2013). This means that users select information, interpret meanings and structure their own knowledge without assistance, the outcomes are shaped by the media experiences of users, and by their own values and backgrounds.

Fig. 5 Distributions of search volume by theme

Distribution of Search Volumes by Theme (Log Scale)



Google Search does not only operate as tool for meeting information need in this case but also works as a “communication medium” through which users actively construct personal stories of social reality. For example, when a person types “presidential candidate 2024” or “fuel prices today”, they enter into a public discussion and try to place themselves within the wider social discourse. The nature of self-mass communication is highlighted by this, as it refers to the potential of digital media to allow an individual to communicate in a self-directed manner and impact others (Castells, 2009)

Moreover, searching using Google is personal, but not entirely private. Every time there’s a search, it impacts the algorithms behind the platform as well as its trends and rationale. Search practice refers to the two-way communication system associated with the technology designed to shape search outcomes and may be vital in doing so(Halavais, 2018; Van Dijck, 2013; Zimmer, 2008). It will be vital to the ways that employed in the process of searching practice. As such, search behavior should be seen as dependent upon user agency and technological mediation.

This technique widens the focus of communication studies from the obvious – (overt, verbal, dialogical) communication – to implicit communication (search, click, attention). Users are not talking to anyone when executing a Google search. Not directly, at least. Yet there is some form of communication taking place. In fact, the search itself communicates something. It expresses an intention, a worry, or an expectation. It is a contribution to this social realm, after all. A silent dialogue, as it were. (Schroeder, 2016; Vlassenroot et al., 2021). People engage in communication behaviours that are hidden, yet steeped in social and cultural meaning using Google Search.

Google Indonesia user queries can be seen as a digital communicative artifact originating from the dynamics of collective action. The history of ideas has traditionally relied on explicit articulations and conscious beliefs. However, we are now facing a new challenge: the quest for indigenous ideas that have the power to shape change. This task proves to be not only difficult but also significant in its own right.

CONCLUSION

According to the study, Google Search used for gathering information is not simply a transaction of data in a technical sense but a digital communication process that is complicated and reflective. According to an Exploratory Data Analysis (EDA) study received by Tempo on the dataset “Google Searches of Indonesian People 2019-2024”, the Indonesian people use search engines as a means of expressing, exploring and orienting themselves on social, political, cultural and economic issues.

The findings showed that the leading search categories were sports, entertainment, politics and economics– which indicates that a great deal of discussion on digital media takes place on the lives of the rich and famous, both local and abroad. The bar chart, pie chart, and word cloud showed that the people largely focused on Foreign Ball, Artist, and Politics. These are major indications of the society. Thematic categories showcase different styles of online search behavior and are closely associated with the social dynamics of the environment.

From a communicational point of view, people’s Google Search is a non-verbal communication. It asserts that searching is part of engagement in the public sphere, which does not involve self-expression. Although it seems like a personal search effort, this is really part of a larger social web that creates discursive publics that generate shared meanings in communities that are linked algorithmically and have similar information needs.

The strategies adopted by communication researchers to study extensive data driven research need hardly a mention. They help in the observance of elaborate and covert information forms. Search data is a significant alternative to more immediate, contextual and empirically driven ways of conducting public communication analysis especially in the context of the profoundly fragmented and algorithmic world of digital media.

REFERENCES

- Abdullah, A. (2020). Public Relations in The Era of Artificial Intelligence: Peluang atau Ancaman? *ARISTO*, 8(2), 406. <https://doi.org/10.24269/ars.v8i2.2629>
- Abdullah, A., Jayus, J., Rozita, R., & Sumaiyah, S. (2025). Pemanfaatan Artificial Intelligence dalam Aktivitas Public Relations: Humas Masa Depan dan Masa Depan Humas. In D. Kartikawangi & E. Setyaningtyas (Eds.), *Perspektif Kecerdasan Artificial dalam Kehumasan*. LSPR Publishing. <https://omp.lspr.edu/index.php/omp/catalog/book/43>

- Abdullah, A., Jayus, J., Yazid, Y., Sumaiyah, Sumaiyah, Khairi, A., Edison, E., Astuti, D. S., & Mutasir, M. (2024). Navigating Society 5.0: Unraveling the Dynamics of a People-Centric Super-Smart Society. *ARISTO*, 12(2), 642–659. <https://doi.org/10.24269/ars.v12i2.9038>
- Abdullah, A., Yazid, Y., Jayus, J., Sumaiyah, S., Khairi, A., Edison, E., & Astuti, D. S. (2024). Google Trends and Indonesia Presidential Elections 2024: Predictor of Popularity Candidate in Digital Age. *POLITICON: Jurnal Ilmu Politik*, 6(2), 273–300. <https://doi.org/10.15575/politicon.v6i2.34636>
- AC. (2019, February 9). Memahami Data Dengan Exploratory Data Analysis. *Data Folks Indonesia*. <https://medium.com/data-folks-indonesia/memahami-data-dengan-exploratory-data-analysis-a53b230cce84>
- Acharya, K. (2024). *How Google Search Works*. <https://doi.org/10.22541/au.172961967.70431246/v1>
- Alauddin, M., Schrader, A., Pegg, M. J., & Amyotte, P. (2025). Investigating dust explosibility using exploratory data analysis. *Process Safety and Environmental Protection*, 108241. <https://doi.org/10.1016/j.psep.2025.108241>
- Astra, S. (2025). *How Google's 2025 Search Changes Will Reshape Communications in IoT*. AI Business. <https://aibusiness.com/iot/how-google-s-2025-search-changes-will-reshape-communications-in-iot>
- Boyd, D., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Castells, M. (2009). *Communication Power*. Oxford University Press.
- Castells, M. (2011). *The Rise of the Network Society*. John Wiley & Sons.
- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, 42(2), 214–257. <https://doi.org/10.1598/RRQ.42.2.2>
- Cornell University. (2020). Search Engines as a Mirror for Society. *Networks Course Blog*. <https://blogs.cornell.edu/info2040/2020/11/11/search-engines-as-a-mirror-for-society/>
- Das, S. B. (2025, January 15). Exploratory Data Analysis: Techniques and Best Practices. *Medium*. <https://medium.com/@dasskbibekananda/exploratory-data-analysis-techniques-and-best-practices-e465def27b6d>
- Dijk, J. van. (2020). *The Network Society* (4th ed.). SAGE Publications.
- Donadelli, M. (2015). Google search-based metrics, policy-related uncertainty and macroeconomic conditions. *Applied Economics Letters*, 22(10), 801–807. <https://doi.org/10.1080/13504851.2014.978070>

- Efthimiadis, E. N., Huang, J., Spink, A., & Jansen, J. (2009). Query formulation in web search. *Proceedings of the American Society for Information Science and Technology*, 46(1), 1–3. <https://doi.org/10.1002/meet.2009.1450460131>
- Eggensperger, J., & Redcross, N. (2018). *Data-Driven Public Relations Research: 21st Century Practices and Applications* (1st ed.). Routledge. <https://doi.org/10.4324/9781315196688>
- Eysenbach, G. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ*, 324(7337), 573–577. <https://doi.org/10.1136/bmj.324.7337.573>
- Fallows, D. (2005). *Search Engine Users Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve*. Pew Research Center. https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2005/PIP_Searchengine_users.pdf
- Google. (2025). *Google Colaboratory* [Computer software]. <https://colab.research.google.com/>
- Götz, T. B., & Knetsch, T. A. (2019). Google data in bridge equation models for German GDP. *International Journal of Forecasting*, 35(1), 45–66. <https://doi.org/10.1016/j.ijforecast.2018.08.001>
- Granka, L. A. (2010). The Politics of Search: A Decade Retrospective. *The Information Society*, 26(5), 364–374. <https://doi.org/10.1080/01972243.2010.511560>
- Grivy. (2025). *The Art of Creating—Data-driven Communication Strategy*. <https://business.grivy.com/blog/the-art-of-creating-data-driven-communication-strategy>
- Halavais, A. (2013). *Search engine society* (First edition). Polity.
- Halavais, A. (2018). *Search engine society* (Second edition). Polity.
- Hargittai, E., & Walejko, G. (2008). THE PARTICIPATION DIVIDE: Content creation and sharing in the digital age1. *Information, Communication & Society*, 11(2), 239–256. <https://doi.org/10.1080/13691180801946150>
- Hassan, R. (2008). *The Information Society: Cyber Dreams and Digital Nightmares*. Polity.
- Höchstötter, N., & Lewandowski, D. (2009). What users see – Structures in search engine results pages. *Information Sciences*, 179(12), 1796–1812. <https://doi.org/10.1016/j.ins.2009.01.028>
- Huang, J., & Efthimiadis, E. (2009). Studying query reformulation strategies in search logs. *Proceedings of the American Society for Information Science and Technology*, 46(1), 1–3. <https://doi.org/10.1002/meet.2009.14504603117>
- Jansen, B. J., Booth, D. L., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*, 44(3), 1251–1266. <https://doi.org/10.1016/j.ipm.2007.07.015>

- Jansen, B. J., & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, 42(1), 248–263. <https://doi.org/10.1016/j.ipm.2004.10.007>
- Jayus, J., Abdullah, A., Mustafa, M., & Sumaiyah, S. (2025). YouTube, Public Discourse, and the ‘Makan Siang Gratis’ Program: An Analysis of Toxicity Comments on the Liputan6 Channel. *Proceeding Jogjakarta Communication Conference*, 3(1). <https://jcc-indonesia.id/proceeding/index.php/jcc/article/view/383>
- Jayus, J., Sumaiyah, S., Mairita, D., & Abdullah, A. (2024). Media Sosial sebagai Media Kampanye Politik Menjelang Pemilu 2024. *JURNAL SIMBOLIKA Research and Learning in Communication Study*, 10(1), 72–81. <https://doi.org/10.31289/simbolika.v10i1.11468>
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. SAGE Publications Ltd. <https://doi.org/10.4135/9781473909472>
- Krismayani, I., & Mafar, F. (2024). Tren Pencarian Informasi Masyarakat Indonesia Menggunakan Google Search Engine. *Lentera Pustaka: Jurnal Kajian Ilmu Perpustakaan, Informasi Dan Kearsipan*, 10(1), 53–60. <https://doi.org/10.14710/lenpust.v10i1.55156>
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), 1203–1205. <https://doi.org/10.1126/science.1248506>
- Lewandowski, D. (2015). Evaluating the retrieval effectiveness of web search engines using a representative query sample. *Journal of the Association for Information Science and Technology*, 66(9), 1763–1775. <https://doi.org/10.1002/asi.23304>
- Lim, M. (2013). The Internet and Everyday Life in Indonesia: A New Moral Panic? *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, 169(1), 133–147. <https://doi.org/10.1163/22134379-12340008>
- Lohr. (2012). *Big Data’s Impact in the World*. The New York Times. <https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>
- Luke, S., Bon, E., Dommett, K., Gibson, R., Lecheler, S., & Kruikemeier, S. (2024). Editorial: Data-Driven Campaigning in a Comparative Context—Toward a 4th Era of Political Communication? *Media and Communication*, 12, 9227. <https://doi.org/10.17645/mac.9227>
- Martin, J. A., Camaj, L., & Lanosga, G. (2024). Audience engagement in data-driven journalism: Patterns in participatory practices across 34 countries. *Journalism*, 25(7), 1578–1596. <https://doi.org/10.1177/14648849241230414>
- Masimengo, T. (2025). Google search volume index and banks’ capital adequacy. *Cogent Economics & Finance*, 13(1), 2494163. <https://doi.org/10.1080/23322039.2025.2494163>

- Matplotlib Development. (2025). *Matplotlib: Visualization with Python* (Version v3.10.8) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.592536>
- Mayer-Schönberger, V. (with Cukier, K.). (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (1st ed). HarperCollins Publishers.
- McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>
- McQuail, D. (2000). Some reflections on the western bias of media theory. *Asian Journal of Communication*, 10(2), 1–13. <https://doi.org/10.1080/01292980009364781>
- McQuail, D. (2010). *Mcquail's mass communication theory*. SAGE Publications.
- Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, 331(6014), 176–182. <https://doi.org/10.1126/science.1199644>
- Mohr, J. W., & Bogdanov, P. (2013). Introduction—Topic models: What they are and why they matter. *Poetics*, 41(6), 545–569. <https://doi.org/10.1016/j.poetic.2013.10.001>
- Morozov, E. (2011). *The Net Delusion: The Dark side of internet freedom*. PublicAffairs.
- Mueller, A. (2020). *WordCloud for Python documentation—Wordcloud 1.8.1 documentation*. Github.Io. https://amueller.github.io/word_cloud/
- Na, K., & Park, M. S. (2012). Searcher's perceptions for query reformulation behavior on the web. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–4. <https://doi.org/10.1002/meet.14504901333>
- Napoli, P. M. (2003). *Foundations of communication policy: Principles and process in the regulation of electronic media* (2. print). Hampton Press.
- Napoli, P. M. (2011). *Audience Evolution: New Technologies and the Transformation of Media Audiences*. Columbia University Press.
- Napoli, P. M. (2014). Digital intermediaries and the public interest. *Telecommunications Policy*, 38(11), 1088–1100. <https://doi.org/10.1016/j.telpol.2014.03.001>
- Niu, X., & Kelly, D. (2014). The use of query suggestions during information search. *Information Processing & Management*, 50(1), 218–234. <https://doi.org/10.1016/j.ipm.2013.09.002>
- Nurrohman, N. (2024). *Google Searches of Indonesian People 2019-2024* [Dataset]. <https://www.kaggle.com/datasets/nugrahmaindonesia/google-searches-of-indonesian-people-2019-2024>
- Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking* (First edition). O'Reilly.

- Rahm, E., & Do, H. H. (2000). Data Cleaning: Problems and Current Approaches. *IEEE Data(Base) Engineering Bulletin*, 23, 3–13. <https://api.semanticscholar.org/CorpusID:260972099>
- Riyaz, A. (2017). An Investigation into the 'I can Google it' Information Seeking Behaviour of the Academic Community and the Implications for the Delivery of Academic Library Services for Developing Countries. *Journal of the Australian Library and Information Association*, 66(2), 180–182. <https://doi.org/10.1080/24750158.2017.1331523>
- Rochmawati, E., & Nurmandi, A. (2020). Public precaution awareness: A case study from Google search trend during Covid-19 outbreak in Indonesia. *Authorea*. <https://doi.org/10.22541/au.158766297.74766951>
- Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., Withey, R., Jamali, H. R., Dobrowolski, T., & Tenopir, C. (2008). The Google generation: The information behaviour of the researcher of the future. *Aslib Proceedings*, 60(4), 290–310. <https://doi.org/10.1108/00012530810887953>
- Schroeder, R. (2016). Big Data and Communication Research. In R. Schroeder, *Oxford Research Encyclopedia of Communication*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.276>
- Shen, Y., Benke, B., Ashtiani, M., Huang, M., & Simonen, K. (2025). Exploratory Data Analysis of a North American Whole Building Life Cycle Assessment datasets. *Building and Environment*, 286, 113655. <https://doi.org/10.1016/j.buildenv.2025.113655>
- Showkat, N., & Gull, M. (2020). The Study of Google Search Trends for an Effective Communication during COVID-19 Pandemic. *ADVANCE*. <https://doi.org/10.31124/advance.12138732.v1>
- St. Amant, R., & Cohen, P. R. (1998). Intelligent Support for Exploratory Data Analysis. *Journal of Computational and Graphical Statistics*, 7(4), 545–558. <https://doi.org/10.1080/10618600.1998.10474794>
- Sumaiyah, S., Abdullah, A., Jayus, J., Yazid, Y., Astuti, D., Khairi, A., & Edison, E. (2024). Building Hope in Crisis: Global Public Service Broadcaster Innovation During the Covid-19 Pandemic. *ARISTO*, 13(1), 128–154. <https://doi.org/10.24269/ars.v13i1.9004>
- Sundar, S. S., & Limperos, A. M. (2013). Uses and Grats 2.0: New Gratifications for New Media. *Journal of Broadcasting & Electronic Media*, 57(4), 504–525. <https://doi.org/10.1080/08838151.2013.845827>
- Taylor, M. (2005). Using the Google Search Appliance for Federated Searching: A Case Study. *Internet Reference Services Quarterly*, 10(3–4), 45–55. https://doi.org/10.1300/J136v10n03_06
- Treem, J. W., & Leonardi, P. M. (2013). Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association. *Annals of the International Communication Association*, 36(1), 143–189. <https://doi.org/10.1080/23808985.2013.11679130>
- Tukey, J. W. (2005). *Exploratory data analysis* (Repr.). Addison-Wesley.
- Van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media* (1st ed.). Oxford University Press New York. <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>

- Vlassenroot, E., Chambers, S., Lieber, S., Michel, A., Geeraert, F., Pranger, J., Birkholz, J., & Mechant, P. (2021). Web-archiving and social media: An exploratory analysis: Call for papers digital humanities and web archives – A special issue of international journal of digital humanities. *International Journal of Digital Humanities*, 2(1-3), 107-128. <https://doi.org/10.1007/s42803-021-00036-1>
- Ward, J. S., & Barker, A. (2013). Undefined By Data: A Survey of Big Data Definitions. *Computer Science*.
- Waskom, M. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- Weng, J. (2021, April 18). *Exploratory Data Analysis (EDA): A Practical Guide and Template for Structured Data*. Medium. <https://towardsdatascience.com/exploratory-data-analysis-eda-a-practical-guide-and-template-for-structured-data-abfbf3ee3bd9>
- White, R. W., & Horvitz, E. (2009). Cyberchondria: Studies of the escalation of medical concerns in Web search. *ACM Transactions on Information Systems*, 27(4), 1-37. <https://doi.org/10.1145/1629096.1629101>
- White, R. W., Richardson, M., & Yih, W. (2015). Questions vs. Queries in Informational Search Tasks. *Proceedings of the 24th International Conference on World Wide Web*, 135-136. <https://doi.org/10.1145/2740908.2742769>
- Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59(10). <https://doi.org/10.18637/jss.v059.i10>
- Zimmer, M. (2008). The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance. In A. Spink & M. Zimmer (Eds.), *Web Search* (Vol. 14, pp. 77-99). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75829-7_6