Information Management and Business Review (ISSN 2220-3796)
Vol. 16, No. 3S(a), pp. 1161-1169, Oct 2024

Factors Influencing Data Partiality in Artificial Intelligence

Faten Elina Kamaruddin[1], *Nur Hanisah Mohamad Razali[2], Ahmad Fuzi Md Ajis[1], Nur Rifhan A Rahim[2], Siti Noorhaslina Abd Halim[2], Ainol Mardhiyah Rahmat[2]
[1]Faculty of Computing, Informatics and Mathematics, UiTM Johor Branch, Malaysia
[2]Faculty of Business and Management, UiTM Johor Branch, Malaysia
fatenelina@uitm.edu.my, *hanisahmr@uitm.edu.my, ahmadfuzi@uitm.edu.my, rifhan@uitm.edu.my, snhaslina@uitm.edu.my, dhiyarahmat@uitm.edu.my
Corresponding Author: Nur Hanisah Mohamad Razali

Abstract: This study proposes a conceptual framework to investigate factors influencing the data partiality in Artificial Intelligence (AI). The popularity of AI is world-renowned machines that can represent human intelligence under the program of computer systems. However, the academic research on data partiality focusing on AI is limited across the bibliographic database sources. This study aims to address the gaps by proposing a developed framework that integrates three factors: the AI algorithm, black data, and user revise terminology highlighted in the past literature. The AI algorithm refers to the issues on the training data as a dataset used in the tools, which stimulates the data partiality as the outcome retrieved by the user. The black data is influencing data partiality on the existence of unknown data. The user revise terminology represented on the keywords used by the user to search for information, which incorrect keywords with not specify will lead to the AI to give all related information as an output without filter. The framework asserts that these three elements directly affect the partiality of data in AI. A quantitative methodology will be used in this study to cover the collection of survey data from the community under the MDEC program called Global Online Workforce (GLOW). The framework contributes a theoretical understanding of AI algorithms, black data, and user-revised terminology that influence data partiality in AI. In future research, the framework can be extended to test the data partiality in AI tools used in information agencies, as these bodies govern the safeguards of the accuracy of the information.

Keywords: *Data partiality, artificial intelligence, black data, algorithm, user revise terminology*

1. Introduction

Artificial intelligence (AI) algorithms are widely employed by corporations, political entities, and other groups to make decisions that profoundly impact individuals and society. They might affect everyone, everywhere, at any moment, with their decisions. While artificial intelligence (AI) can solve problems in various domains of daily life, it also carries risks, such as the potential of partial data from the information it receives. Data partiality is neither novel nor exclusive to AI, therefore it is unattainable to eliminate the danger of bias in an AI system (Schwartz et al., 2023). The presence of partial data in artificial intelligence (AI) is a crucial matter that has recently received much focus. AI systems that contain prejudice can result in unintentional ethical, social, and legal problems (Agarwal et al., 2022). Within machine learning, partial data are frequently acquired, resulting in the well-known saying "garbage in, garbage out" (Tempke & Musho, 2022). Research has demonstrated that AI models can possess inherent partial data, which mirrors the underlying inequities in society (Zhou et al., 2024). Addressing partial data in AI is an intricate task because data partiality can be deeply embedded in the data utilized to train AI models (Caliskan et al., 2017). Partial datasets can contribute to the continuation of inequalities.

Moreover, as shown in the news, data partiality is a significant concern from a commercial standpoint, with approximately 22% of firms lacking systematic methods to identify data partiality in AI technologies. (Business Times, 5 January 2024). Even though AI can help a business grow, communicate with customers in real-time, manage operations, and find new ways to innovate and expand, its efficiency is still not fully utilized. In this case, since Malaysia recognizes the significance of AI governance, there is still a need to ensure data partiality in AI training data may be used for good (Saeidnia, 2023). Nowadays, the number of internet users worldwide is more than five billion users. Despite a tenfold rise in the number of websites over the past decade, the quantity of pages indexed by Google has neared 50 billion (Techjury, 2023). Information is rapidly growing in the virtual environment with the internet's widespread use. This is aligned with the statement reported by a commissioned study by Malaysia Digital Economy Corporation (MDEC) Malaysia's big data analytics market is anticipated to grow to US$ 1.9 billion (RM8.7 billion) by 2025 (Online, 2023). The growing number of users on the internet will then lead to an increased amount of new information being produced, which will cause higher data partiality problems in AI as well. For instance, based on a study made by Lorenz et al. 2024, artificial intelligence ChatGPT, the results about drug allergies aligned with the prevailing allergy and immunology guidelines; however, the chatbot proved to be an untrustworthy information source due to inaccuracies in all the referenced materials. The errors raised are due to data partiality in ChatGPT (Lorenz et al., 2024).

From a previous study, the impact of data partiality in AI results in the exclusion of marginalized library communities can be observed in the context of library and information fields (Kim et al., 2021). The study conducted by Kim et al., 2021, found AI tends to select a range of people by limiting the selection of the material upon the request made by library users. This situation caused partial data obtained from AI, which the AI does not fulfill the library function, which libraries main for a wide range of people to get reference materials. Users depend on libraries to offer impartial and objective information; any evidence of partiality in AI algorithms might erode people's trust. Based on Ferrara (2023), the issue related to partial data in AI led to reinforce preconceptions and discrimination among AI users. This happened when AI was built with various datasets, thus the output is given in search results, procedures, recommendations, and decision making then became partial. For instance, a study conducted by Saeidnia (2023) found that data partiality in AI tools leads to partial datasets in biometric technology, such as facial identification was set to be fair skin faces, the AI having a problem recognizing or verifying users with darker skin faces. Data already set in AI is not reflected as a true output in most situations, and the difference in the result disproportionately impacts specific individuals or groups (Akter et al, 2021).

On top of the issues, previous research found factors that influence data partiality in AI, which were algorithm in AI (Daneshjou et al., 2021), black data (AFM Ajis et al., 2022), and user-revised terminology (Atman Uslu & Yildiz Durak, 2022). According to Daneshjou et al. (2021), the study demonstrated that a lack of clarity in AI algorithms resulted in data partiality inside AI systems. In most AI tools, the algorithms are crafted using training data (Saeidnia, 2023). The algorithm using training data did not mirror the variety of types of data entered by the user, which led AI to select the trained data, which is partial. A suggestion from Chen (2023) stated that it is important to study the algorithm used in AI to identify the presence of partial data. Furthermore, another factor that causes data partiality in AI is black data. Black data is derived from the dark data situation, which refers to the uncontrollable data, which occurs mostly when managing a large amount of data. In information technology fields, the vast amount of data is increasing rapidly, and this has caused problems with the technology tools to manage the data. Those unmanaged data, however, have meaningful meaning based on the situation. The AI tools will give the information based on the user searching action, thus those unmanaged data will be called as well as a result (De La Peña & Granados, 2023). The dissemination of inaccurate information during a commercial transaction can negatively impact the profitability and overall performance of the user of an AI (Ntoutsi et al., 2020). Moreover, every information in the Internet environment has a limited lifespan (Marwala et al., 2015). Inaccurate data containing partial data elements given by AI tools may affect an individual's decision-making. This study will identify the elements of black data that emerged in AI tools that could lead to partial data and disrupt the user decision-making process. On the other hand, the third factor emphasized in this paper that causes data partiality in AI is user-revised terminology. With the development of new online tools, various terms and sometimes combined terms can formulate the terminology captured and processed by AI tools (Jumabek et al., 2024). Consequently, this gives rise to the development of terminology that amalgamates issues from several disciplines and affects the terminology process by an AI. Furthermore, traditional terminology is occasionally modified or augmented with new AI elements to accurately represent modern perspectives and needs (Jumabek et al., 2024). The present scenario gives rise to a lack of consistency and standardization in the terminology, impeding the accessibility of mathematical knowledge for different user groups.

Therefore, this study is intended to discover the factors influencing data partiality in AI. Throughout this paper, three factors were identified: algorithms in AI, black data, and user review terminology.

2. Literature Review

Artificial Intelligence (AI)

Artificial intelligence has become the driving force behind all breakthrough technology in the 20th century. A burgeoning area of technology called artificial intelligence uses models of human brain networks to infer patterns from certain information. Artificial intelligence enables machines to execute tasks that otherwise necessitate human intelligence (Preckel et al., 2020; Schoser, 2023). Artificial intelligence is becoming more prevalent in every field due to its human-like intelligence and capacity for problem-solving. Artificial intelligence is becoming increasingly necessary in the current environment, as computers are used for theoretical and practical purposes. In the context of information management, AI tools are characterized as intelligent systems capable of adapting to their environment while operating with limited knowledge and resources (Wang, 2022). Computational algorithms in artificial intelligence can vary from basic, rule-based directives to intricate methodologies such as machine learning and deep learning (Raj, 2019). For example, these algorithms may be used in pattern recognition, data processing, decision-making, or data-driven learning (Gignac & Szodorai, 2024). Thus, artificial intelligence can be defined operationally as the maximum ability of an artificial system to do a novel standardized task with accurate scoring through computer algorithms.

Data partiality in Artificial Intelligence (AI)

According to Ferrer et al., 2021, data partiality in artificial intelligence (AI) refers to individuals or groups' unfair or unjust treatment because of partiality embedded within AI algorithms or systems. Equitable access to information may be hampered by the employment of AI systems in the information sector, which could reinforce prejudice and inequality. Inaccurate data, skewed decision-making procedures, and opaque AI algorithms are all potential sources of partial data (Saeidnia, 2023). Addressing data partiality and discrimination in AI is crucial for ensuring the completeness and accuracy of information (Ferrer et al., 2021; Akter et al., 2021). This entails analyzing and addressing incomplete data in training datasets, enhancing algorithmic openness and interpretability, engaging various stakeholders in the creation and assessment process, and enforcing ethical standards and rules (Malek, 2022; Gerards & Zuiderveen, 2018). Data partiality and discrimination in AI are not deliberate but stem from the foundational data and algorithms employed (Chen, 2023). If these issues are not resolved, AI systems have the potential to exacerbate already existing inequities and expand gaps in the services provided by the information industry. Advocating for ethical AI practices and emphasizing education and awareness is crucial to mitigate data partiality and guarantee fairness, accountability, and transparency in AI technology within the information economy. Consequently, it is imperative to perpetually assess and enhance AI systems to mitigate the effects of incomplete data and promote justice and equality.

Algorithm in AI

Artificial intelligence algorithms are engineered to render conclusions or forecasts based on patterns and data (Malek, 2022). Nonetheless, these algorithms may unintentionally mirror the data partiality and prejudices inherent in the training data or their programming (Ferrer et al., 2021; Malek, 2022*)*. Algorithms are configured so that a computer can process instructions and solve problems*.* According to Akter et al., 2021, Algorithmic bias refers to the phenomenon wherein AI systems generate discriminatory outcomes or decisions that disproportionately impact specific individuals or groups. This may occur when the training data utilized to construct the AI system is incomplete or lacks variety, resulting in biased predictions or recommendations. Furthermore, discrimination in AI may arise when algorithms are employed to make decisions in domains such as recruitment, lending, or criminal justice. These algorithms can sustain or exacerbate discriminatory practices if trained on incomplete data. An AI hiring tool may unintentionally discriminate against specific groups if educated on historical data that embodies biased hiring practices (Ferrer et al., 2021; Akter et al., 2021). These data partialities in AI systems were inherent to the datasets and algorithmic procedures utilized to create AI applications. These issues frequently occur when algorithms are trained on certain data and cannot extrapolate beyond it. The inaccuracy could result from various sources, including heterogeneous data, incorrect data, complicated data represented in simpler mathematical representations, algorithmic partial like over- and under-fitting, handling outliers, data cleaning, and imputation issues.

Black data

Black data in information management refers to data beyond the control of custody, searchability, and awareness, rendering its usage banned, deceptive, or impracticable. This is an entirely opaque data scenario called black data. Black data comprises the following characteristics: unknown and misled data. In the realm of unidentified data, utilizing any artificial intelligence platform to seek information may exacerbate the influence of dark data on business performance and operations. The scenario in which undiscovered data remains unregulated or is in possession yet unrecognized or is beyond the comprehension of the AI provider despite identifiable sources, constituting the origin of the unknown data. In addition, unknown data were derived from uncertain data, defined as data with inconsistent context or unclear meaning, making it impossible to identify and forecast the data's significance or impact. Thus, data partiality came from an information provider such as an AI platform that had a relation with unknown data whereby the receiver or user of information with partial data tends to accept partial information if utilized without critical review. On the other hand, misled data given by the information provider, such as AI tools, could greatly impair the user. Inaccurate information supplied by any party may disrupt corporate operations and incur additional obligations for the organization (Ajis et al., 2022). Misleading information is typically the result of communication errors and arises from asymmetrical data, where the recipient interprets the contextual meaning of the information differently. Therefore, data partiality is related to misled information, as some data is asymmetrical for instance, from the output obtained through various AI tools, the information is inconsistent and contradicts multiple AI tools and others.

User revise terminology

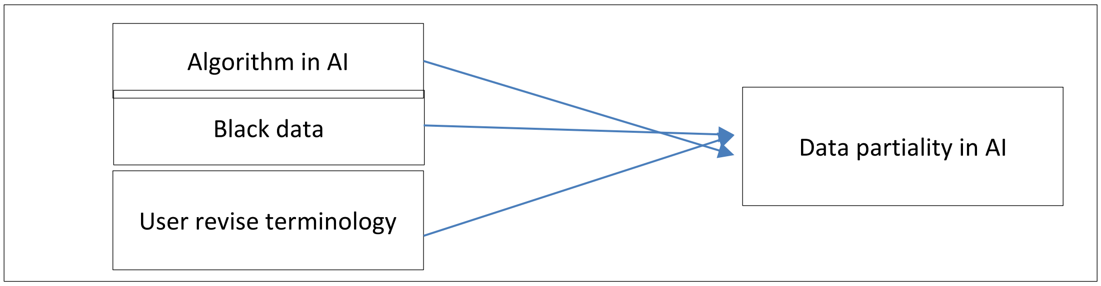
According to Adegbilero-Iwari, Oluwadare and Adegbilero-Iwari (2023), in today lives, finding information and knowledge is a major time-consuming activity for people. Nowadays, the Internet serves as the primary source for information acquisition, as opposed to the conventional printed materials-filled spaces like libraries where information searchers once conducted their business (Sharit, Hern´andez, Czaja, & Pirolli, 2008; Tsai, 2022). Today, individuals will primarily opt for the Internet or AI-integrated search engines, as these tools facilitate access to extensive knowledge with minimal effort (Wellings & Casselden, 2017). Finding information is a complex cognitive activity involving cognitive, metacognitive, problem-solving, and decision-making techniques (Atman Uslu & Yildiz Durak, 2022; Monchaux, Amadieu, Chevalier & Marin´e, 2015). People look up information on any topic on the internet (Canan Güng¨oren, Gür Erdo˘gan, & Kaya Uyanık, 2019).

However, increasing knowledge in online spaces necessitates more adept search techniques (Atman Uslu & Yildiz Durak, 2022). According to Da et al., 2011, searching for information online can direct the user's attention and encourage their behavior toward accepting partial data without evaluating the result. By conducting internet research, one can lessen the likelihood of an information crash by mitigating the partiality present in the information provider ecosystem. This study will investigate the impact of internet searches on the data partiality produced by AI systems. The technology platforms that govern the presentation and retrieval of information online significantly influence user search behaviors and the performance of AI technologies. For instance, Google is a popular search engine that frequently modifies its algorithms and business structures in generally unclear ways to individual users (Mager et al., 2023). Thus, the variety of datasets used by Google search engines could be inaccurate and lead to data partiality. Typical users lack awareness of the mechanisms that determine the data they access, the prioritization of search results, and the selection of related links (Hoeyer et al., 2024). The difficulties multiply when dealing with big language models such as ChatGPT, Bard, or Bing Chat. In this situation, so-called generative artificial intelligence (AI) provides consumers with synthesized information based on the statistical likelihood that a given sentence will answer their query. While AI can quickly generate highly understandable responses, it will also be prone to errors and may be incorrect and complete (Hoeyer et al., 2024). Users might use the same or various keywords, however, the sources of information given by AI tools remain unknown to the user and might cover outdated information as well that contains inherent partial data or skewed information (Bender et al., 2021). The generated reactions may result in diverse repercussions, including perpetuating or amplifying existing prejudices, misattribution, and disinformation.

Proposed Framework

Figure 1 illustrates the proposed conceptual framework that was created for this study. The framework is built upon a comprehensive literature analysis to investigate the elements that impact data biases in artificial intelligence. Accordingly, a systematic literature review proposed a model with three core constructs, namely black data, AI algorithm, and user revise terminology as the Independent Variable (IV) with data partiality in AI as the Dependent Variable (DV).

Figure 1: Proposed conceptual framework



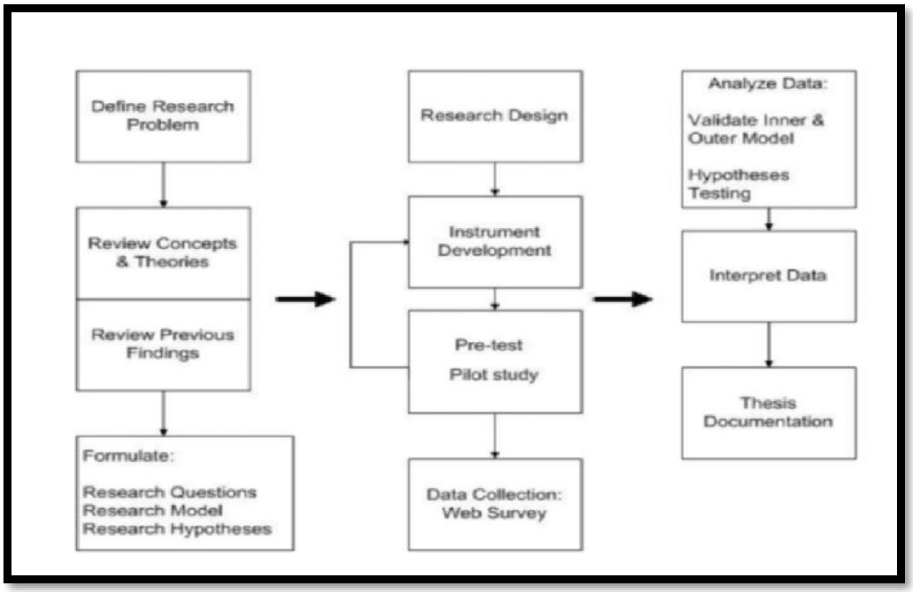
3. Research Methodology

To assess the proposed model, this study will be using a quantitative research design involving statistical analysis of data sets and surveys to evaluate the partiality and accuracy of data used in AI systems within the information industry, as illustrated in Figure 2 (Sekaran, 2003). Therefore, samples of respondents will be selected using purposive sampling, which is among the community practices of using AI applications or systems that only cover one type of AI engine known as the AI chatbot. The scope of the AI chatbot programs will be narrowed down to ChatGPT and Copilot to initiate the investigation before expanding the coverage of more types of chatbot programs. The proposed selected samples are considered expert samples as they are involved in using the AI chatbot programs and teaching it to the community under the MDEC program called Global Online Workforce (GLOW). There will be approximately 646 respondents available for this study.

Survey questionnaires will be used to gather the data sets required for the study. Reliability and validity for the survey will be executed by executing content validity and face validity. This would involve expert review from the field and pilot testing to obtain fine instruments for the study. Further, a few samples from the targeted population will be asked to review the instrument's usability. Afterward, the instrument's reliability will depend on the internal consistency of Cronbach Alpha calculation on the pilot testing datasets.

The data collected will then be analyzed using a few data analysis processes. This study will use descriptive analysis to explore and measure the level of data partiality in AI systems within the information industry. It is also suggested that the data dispersion among findings be explored to investigate variance for each variable and understand the data variability. Further, inferential analysis will also be executed to examine the impact of independent variables on the data partiality in AI systems within the information industry. Besides, correlation analysis is also planned to be executed to understand the relationship among variables and their strength towards each other. However, these would be depending on the normality of the data collected. The software selected to perform the analysis will be Statistical Package for the Social Sciences (SPSS).

Figure 2: Research Process (Sekaran, 2003)



4. Discussion

Data partiality is crucial as artificial intelligence (AI) systems are developed and implemented. These elements significantly influence the efficacy, equity, and dependability of AI applications in various fields. Data partiality is essential to the efficacy and equity of AI systems. To tackle these problems, the data must be carefully chosen, the algorithms must be designed, and the keywords used and other viewpoints must be included in the development process. Various artificial intelligence systems are highly effective but also equitable and reliable by putting strong safeguards in place to identify and reduce prejudice and guarantee data veracity.

It is crucial to study the factors influencing data partiality in artificial intelligence to enhance the data accuracy and completeness obtained from any AI tools available nowadays. Prior research primarily concentrated on using and implementing artificial intelligence across fields, with relatively less focus on highlighting factors causing the data partiality.

Algorithmic partiality is a serious issue that arises when AI systems generate incomplete or unjust results. AI algorithmic partiality is a complex problem that calls for all-encompassing approaches to identify, quantify, and lessen its impacts. Data partiality can be introduced into AI algorithms through their design, which includes the selection of optimization criteria and objective functions. Algorithms optimized for total accuracy could overlook the variations in performance among other subgroups. To address this issue, it is important to create fairness-aware learning algorithms that can intervene at various points in the decision-making process to reduce partial data (Redi & Alameda-Pineda, 2019). Although it may pose difficulties to entirely eradicate partial data from AI systems, ongoing endeavors exist to provide uniform procedures for recognizing and controlling data partiality in AI (Schwartz et al., 2022).

Black data describes information unsuitable for AI systems because it is erroneous, out-of-date, incomplete, or has other flaws. Black data has a substantial and complex impact on the accuracy of artificial intelligence. Additionally, incomplete or inaccurate data can cause models to overfit the flawed data or underfit the actual patterns, which degrade performance. This study is needed to address the issues raised by black data, which involves maintaining data quality, implementing bias detection and mitigation techniques, encouraging transparency and explainability, and abiding by legal and ethical requirements. Prejudices or historical inaccuracies in black data can be learned and reinforced by AI systems, perpetuating data partiality.

Uncertain or poorly defined terminology can cause AI systems to interpret partial data. When disparate data sources employ diverse wording to explain comparable ideas, the lack of standardized terminology might introduce data partiality. AI systems that cannot successfully reconcile the disparities may make partial interpretations and choices due to this contradiction. These technologies' user-revised nomenclature may unintentionally induce biases, producing unfair and discriminating results. This study investigates how terminology contributes to prejudice in AI tools, the consequences of this bias, and methods to lessen its effects.

5. Conclusion and Recommendations

As for the conclusion, this study's findings might help address the factors leading to data partiality in AI systems. Data partiality can hinder the effective implementation of AI tools by providing inaccurate and incomplete data. When considering partiality and fairness, it is crucial to include those who may be directly affected by AI systems in the concerns of the emergence of AI algorithms, black data, and user-revised terminology as for now. It is crucial to examine the concept of accuracy that could result in increased surveillance and data collection on disadvantaged communities.

Additionally, it is essential to recognize the framework developed from this study can be commercialized by providing awareness and training on leveraging data partiality of AI tools hoarded by any information industry. The concept built is relevant to highlight the issues of data partiality in AI tools in the context of data accuracy, which can be implemented to specific domain knowledge. Hence, the untested proposed theory can be tested in future research by focusing on specific information fields such as library AI tools and lead to the emergence of new theory in the context of data partiality in AI from the information management field throughout the development of algorithm or formula to assist in reducing the partiality percentage of data obtain from AI tools.

On the other hand, this paper sheds light on how AI understands various professions, which is important knowledge about the data partiality inherent in AI systems. Given the extent of data partiality in generative AI tools, the AI community needs to take rapid stock of its practices and take corrective action. As our reliance on AI grows, ensuring these technologies are designed and implemented fairly and equally is critical. Creating generative AI systems that are not just technologically sophisticated but also ethically molded by a dedication to equity and inclusivity must be prioritized. By doing this, data partiality could be reduced by not undermining the AI user's objectives with the technology as the information generator.

References

Adegbilero-Iwari, O. E., Oluwadare, T., & Adegbilero-Iwari, I. (2023). Predictors of online health information-seeking behavior of non-medical undergraduate students might vary. *Health Information & Libraries Journal*. https://doi.org/10.1111/hir.12479

Agarwal, A., Agarwal, H., & Agarwal, N. (2022). Fairness Score and process standardization: a framework for fairness certification in artificial intelligence systems. *AI And Ethics*, *3*(1), 267–279. https://doi.org/10.1007/s43681-022-00147-7

Ahmed, W., & Ameen, K. (2017). Defining big data and measuring its associated trends in the field of information and library management. *Library Hi Tech News*, *34*(9), 21–24. https://doi.org/10.1108/lhtn-05-2017-0035

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, *60*, 102387. https://doi.org/10.1016/j.ijinfomgt.2021.102387

Atman Uslu, N., & Yildiz Durak, H. (2022). The relationships between university students' information-seeking strategies, social-media-specific epistemological beliefs, information literacy, and personality traits. *Library & Information Science Research*, *44*(2), 101155. https://doi.org/10.1016/j.lisr.2022.101155

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623)

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Canan Gungoren, Z., Gur Erdogan, D., & Kaya Uyanik, G. (2019). Examination of Preservice Teachers' Lifelong Learning Trends by Online Information Searching Strategies. *Malaysian Online Journal of Educational Technology*, *7*(4), 60–80. https://doi.org/10.17220/mojet.2019.04.005

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, *10*(1). https://doi.org/10.1057/s41599-023-02079-x

Corallo, A., Crespino, A. M., Vecchio, V. D., Lazoi, M., & Marra, M. (2023). Understanding and Defining Dark Data for the Manufacturing Industry. *IEEE Transactions on Engineering Management*, *70*(2), 700–712. https://doi.org/10.1109/tem.2021.3051981

DA, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *The Journal of Finance*, *66*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Daneshjou, R., Smith, M., Sun, M. & Rotemberg, V. & Zou, J. (2021). Lack of Transparency and Potential Bias in Artificial Intelligence Data Sets and Algorithms: A Scoping Review. JAMA Dermatology. 157. 10.1001/jamadermatol.2021.3129.

De La Peña, N. & Granados, OM. (2023). Artificial intelligence solutions to reduce information asymmetry for Colombian cocoa small-scale farmers. Information Processing in Agriculture [Internet]. Available from: https://doi.org/10.1016/j.inpa.2023.03.001

Ferrara, E. (2023). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. Sci [Internet]. 6(1), 3. Available from: https://doi.org/10.3390/sci6010003

Ferrer, X., Nuenen, T. V., Such, J. M., Cote, M., & Criado, N. (2021). Bias and Discrimination in AI: A Cross-Disciplinary Perspective. *IEEE Technology and Society Magazine*, *40*(2), 72–80. https://doi.org/10.1109/mts.2021.3056293

Gerards, J., & Zuiderveen, B. F. (2020). Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3723873

Gignac, G. E., & Szodorai, E. T. (2024). Defining intelligence: Bridging the gap between human and artificial perspectives. *Intelligence*, *104*, 101832. https://doi.org/10.1016/j.intell.2024.101832

Hoeyer, K., Couturier, A., Barawi, K., Drew, C., Grundtvig, A., Lane, E., Munk, A. K., Whiteley, L. E., & Munsie, M. (2024). Searching for information about stem cells online in an age of artificial intelligence: How should the stem cell community respond? *Stem Cell Reports*, *19*(2), 159–162. https://doi.org/10.1016/j.stemcr.2023.12.009

How Many Blogs are Published Per Day in 2023? (2023, July 28). Techjury. https://techjury.net/blog/blogs-published-per-day/

Jumabek, S., Tuwelbay, K., Allabay, A. & Geldibayev, B. (2024). Current Trends in The Formation of Mathematical Terminology on An Online Platform. International Journal of Advanced Science Computing and Engineering [Internet]. 6(1), 7–12. Available from: https://doi.org/10.62527/ijasce.6.1.188

Kim, J, Giroux M, Lee JC. When do you trust AI? The effect of number presentation detail on consumer trust and acceptance of AI recommendations. Psychology and Marketing [Internet]. 2021 May 8;38(7):1140–55. Available from: https://doi.org/10.1002/mar.21498

Lorenz, K., Raviv, O., DIMOV, G., Dimova, M., Graneiro, A. L., & Dimov, V. (2024). Quality of Drug Allergy Information Obtained Via An Artificial Intelligence Bot Versus Clinical Guidelines. *Journal of Allergy and Clinical Immunology*, *153*(2), AB154. https://doi.org/10.1016/j.jaci.2023.11.507

Mager, A., Norocel, O. C., & Rogers, R. (2023). Advancing search engine studies: The evolution of Google critique and intervention. *Big Data & Society*, *10*(2). https://doi.org/10.1177/20539517231191528

Malaysia Digital Economy Corporation. (n.d.). MDEC. https://mdec.my/

Malek, M. A. (2022). Criminal courts' artificial intelligence: the way it reinforces bias and discrimination. *AI And Ethics*, *2*(1), 233–245. https://doi.org/10.1007/s43681-022-00137-9

Marwala, T. & Hurwitz, E. (2015). Artificial Intelligence and Asymmetric Information Theory. arXiv (Cornell University) [Internet]. 2015 Jan 1; Available from: https://arxiv.org/abs/1510.02867

Md Ajis, A. F., Zakaria, S., & Ahmad, A. R. (2022). Modelling Dark Data Lifecycle Management: A Malaysian Big Data Experience. *International Journal of Academic Research in Business and Social Sciences*, *12*(3). https://doi.org/10.6007/ijarbss/v12-i3/12363

Minsky, M. (1961). Steps toward Artificial Intelligence. *Proceedings of the IRE*, *49*(1), 8–30. https://doi.org/10.1109/jrproc.1961.287775

Monchaux, S., Amadieu, F., Chevalier, A., & Mariné, C. (2015). Query strategies during information searching: Effects of prior domain knowledge and complexity of the information problems to be solved. *Information Processing & Management*, *51*(5), 557–569. https://doi.org/10.1016/j.ipm.2015.05.004

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C. 7 Staab, S. (2020). Bias in data-driven artificial intelligence systems-An introductory survey. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, *10*(3). https://doi.org/10.1002/widm.1356

Online, T. S. (2023, July 24). *Changing the world for the better with IOT and Big Data*. https://www.thestar.com.my/starpicks/2023/07/24/changing-the-world-for-the-better-with-iot-and-big-data

Outlook 2024 Latest News & Headlines-THE BUSINESS TIMES. (n.d.). The Business Times. https://www.businesstimes.com.sg/keywords/outlook-2024

Preckel, F., Golle, J., Grabner, R., Jarvin, L., Kozbelt, A., Müllensiefen, D., Olszewski-Kubilius, P., Schneider, W., Subotnik, R., Vock, M., & Worrell, F. C. (2020). Talent Development in Achievement Domains: A Psychological Framework for Within- and Cross-domain Research. *Perspectives on Psychological Science*, *15*(3), 691–722. https://doi.org/10.1177/1745691619895030

Redi, M., & Alameda-Pineda, X. (2019). Opinion column: Fairness, accountability and transparency (in multimedia). *ACM SIGMultimedia Records*, *11*(3), 1–1. https://doi.org/10.1145/3524460.3524462

Raj, J. (2019). A Comprehensive Survey on the Computational Intelligence Techniques and its Applications. *Journal of ISMAC*, *01*(03), 147–159. https://doi.org/10.36548/jismac.2019.3.002

Saeidnia, H. (2023). Using ChatGPT as Digital/ Smart Reference Robot: How may ChatGPT Impact Digital Reference Services? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4441874

Saeidnia, H. R. (2023). Ethical artificial intelligence (AI): confronting bias and discrimination in the library and information industry. *Library Hi Tech News*. https://doi.org/10.1108/lhtn-10-2023-0182

Schoser, B. (2023). Editorial: Framing artificial intelligence to neuromuscular disorders. *Current Opinion in Neurology*, *36*(5), 424–426. https://doi.org/10.1097/wco.0000000000001190

Schwartz, R., Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). Towards a standard for identifying and managing bias in artificial intelligence (Vol. 3, p. 00). US Department of Commerce, National Institute of Standards and Technology.

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). *Towards a standard for identifying and managing bias in artificial intelligence*. https://doi.org/10.6028/nist.sp.1270

Sekaran, U., & Bougie, R. (2003). Research methods for business: A skill building approach. john Wiley & sons.

Sharit, J., Hernández, M. A., Czaja, S. J., & Pirolli, P. (2008). Investigating the Roles of Knowledge and Cognitive Abilities in Older Adult Information Seeking on the Web. *ACM Transactions on Computer-Human Interaction*, *15*(1), 1–25. https://doi.org/10.1145/1352782.1352785

Tempke, R., & Musho, T. (2022). Autonomous design of new chemical reactions using a variational autoencoder. *Communications Chemistry*, *5*(1). https://doi.org/10.1038/s42004-022-00647-x

Tsai, P. S. (2022). Research on information searching strategies in high school students' quality of argumentative essay writing. *Interactive Learning Environments*, *31*(10), 6799–6817. https://doi.org/10.1080/10494820.2022.2046108

Wang, P. (2022). Intelligence: From definition to design. In International workshop on self-supervised learning (pp. 35–47). PMLR.

Wellings, S., & Casselden, B. (2017). An exploration into the information-seeking behaviors of engineers and scientists. *Journal of Librarianship and Information Science*, *51*(3), 789–800. https://doi.org/10.1177/0961000617742466

Zhou, M., Abhishek, V., Derdenger, T., Kim, J., & Srinivasan, K. (2024, March 5). *Bias in Generative AI*. arXiv.org. https://doi.org/10.48550/arXiv.2403.02726