THE ACCURACY LEVEL OF CHATGPT IN INFORMATION SEARCH BY SOCIAL SCIENCE STUDENTS

NURUL NAJIHAH BINTI MAZLI

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

APPROVAL

"I hereby admit that I have read this thesis and in my opinion, this thesis meets the scope and quality for the purpose of awarding Bachelor Degree of Technology Management (Innovation Technology) with Honors"

Signature :
Name Of Supervisor : DR. KAMARUDIN BIN ABU BAKAR
Date : 25 JANUARY 2025

Signature :
Name Of Panel : DR. NORUN NAJJAH BINTI AHMAT
Date : 25 JANUARY 2025

THE ACCURACY LEVEL OF CHATGPT IN INFORMATION SEARCH BY SOCIAL SCIENCE STUDENTS

NURUL NAJIHAH BINTI MAZLI

This report is submitted in partial fulfillment of the requirement for the award
Bachelor Degree of Technology Management (Innovation Technology)
with Honors

Faculty of Technology Management and Technopreneurship
Universiti Teknikal Malaysia Melaka

FEBRUARY 2025

DECLARATION OF ORIGINAL WORK

I hereby declare that this final year project with the title

"THE ACCURACY LEVEL OF CHATGPT IN INFORMATION SEARCH BY SOCIAL SCIENCE STUDENTS"

This is the result of my research except as cited in references.

SIGNATURE	:
NAME	: NURUL NAJIHAH BINTI MAZLI
DATE	: 3 FEBRUARY 2025

DEDICATION

To my dearest family and friends,

This final year project marks the culmination of years of dedication and hard work. As I stand at this significant milestone, I cannot express enough gratitude for your unwavering support and encouragement throughout my academic journey. There were moments of doubt and frustration, but your belief in me never wavered. Your constant motivation pushed me to strive for excellence, and your understanding during late nights and demanding deadlines made it all manageable.

A special thankyou to my supervisor Dr Kamarudin Bin Abu Bakar and my panel Dr. Norun Najjah Binti Ahmat, whose invaluable guidance and mentorship played a crucial role in shaping my approach to this project. Your challenging questions and insightful feedback helped me refine my ideas and develop critical thinking skills that will undoubtedly benefit me in the future.

ACKNOWLEDGEMENT

Bismillahirrahmanirrahim,

In the name of Allah, The Most Gracious, The Most Merciful. Praise to Allah S.W.T. who blessed me with the strength and dedication to complete this thesis. Peace and prayers be upon His Final Prophet and Messenger Muhammad.

Firstly, I would like to express my sincere gratitude to the following individuals and organizations for their invaluable support throughout the completion of this final-year project:

My supervisor, Dr. Kamarudin Bin Abu Bakar, for your guidance, encouragement, and insightful feedback throughout the research and writing process. Your expertise in artificial intelligence was crucial in shaping this project, especially when I was struggling to finish this final-year project.

Finally, I am grateful to my family and friends for their unwavering support and encouragement during this challenging but rewarding journey.

ABSTRACT

Nowadays, artificial intelligence is widely used in every sector including in education. Most students especially university students, are very efficient in using AI tools like ChatGPT to help them do their work like writing assignments and essays and finding pieces of information. They already made the AI tool one of their reference materials. It is good as they have a lot of sources that students can refer to while doing their work. However, the students should not be too dependent on the information provided by the ChatGPT. This is because it is still a technology and the accuracy level also is not convincing. Therefore, this research aims to study the accuracy level of information retrieved by social science students using ChatGPT and determine the relationship between three independent variables (validity, reliability, and relevance) and dependent variables (information search by social science students). This study focused on the Universiti Teknikal Malaysia Melaka (UTeM) social science students. The data was collected from 215 respondents through the questionnaire survey. Therefore, the result from the Multiple Regression Analysis and Pearson's Correlation Coefficient showed that accuracy levels which are validity, reliability, and relevance had a significant relationship and strong relationship to information search by social science students. To conclude, through this research, hope that students will be more careful with the information provided by ChatGPT and ensure that it is valid before using it.

Keywords: ChatGPT, Artificial intelligence (AI), Social science students, Accuracy level, Validity, Reliability, Relevance, Information search

ABSTRAK

Pada masa kini, kecerdasan buatan digunakan secara meluas dalam setiap sektor termasuk dalam pendidikan. Kebanyakan pelajar terutamanya pelajar universiti sangat cekap menggunakan alatan AI seperti ChatGPT untuk membantu mereka melakukan kerja mereka seperti menulis tugasan dan esei dan mencari cebisan maklumat. Mereka telah menjadikan alat AI sebagai salah satu bahan rujukan mereka. Ia bagus kerana mereka mempunyai banyak sumber yang boleh dirujuk oleh pelajar semasa membuat kerja mereka. Walau bagaimanapun, pelajar tidak boleh terlalu bergantung kepada maklumat yang diberikan oleh ChatGPT. Ini kerana ia masih merupakan teknologi dan tahap ketepatannya juga tidak meyakinkan. Oleh itu, penyelidikan ini bertujuan untuk mengkaji tahap ketepatan maklumat yang diperoleh oleh pelajar sains sosial menggunakan ChatGPT dan bertujuan untuk menentukan hubungan antara tiga pembolehubah tidak bersandar (kesahan, kebolehpercayaan, dan perkaitan) dan pembolehubah bersandar (pencarian maklumat oleh pelajar sains sosial). Kajian ini tertumpu kepada pelajar sains sosial di Universiti Teknikal Malaysia Melaka (UTeM). Data dikumpul daripada 215 responden melalui tinjauan soal selidik. Oleh itu, hasil daripada Analisis Regresi Berganda dan Pekali Korelasi Pearson menunjukkan tahap ketepatan iaitu kesahan, kebolehpercayaan, dan perkaitan mempunyai hubungan yang signifikan dan hubungan yang kuat terhadap pencarian maklumat oleh pelajar sains sosial. Kesimpulannya, melalui penyelidikan ini, diharapkan pelajar lebih berhati-hati dengan maklumat yang diberikan oleh ChatGPT dan memastikan ianya adalah satu maklumat yang sahih sebelum menggunakannya.

Kata Kunci: ChatGPT, Kecerdasan buatan, Pelajar sosial sains, Tahap ketepatan, Kesahan, Kebolehpercayaan, Perkaitan, Pencarian maklumat

TABLE OF CONTENT

CHAPTER 2 LITERATURE REVIEW

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX

CHAPTER 1

INTRODUCTION

1.1 Background of Study

As technology rapidly evolves, education undergoes a shift and transformation. Education and technology are interconnected and influence each other. Technology has facilitated and influenced educational advancement (Potasheva et al., 2019; Mhlanga, 2023; Baidoo-Anu & Owusu Ansah, 2023). Technology advancements are causing traditional practices to struggle to keep pace with society, and they demand a change in the mindset of the people to benefit from it (Tubachi,2018). As people constantly need ways to resolve their information problems relatively quickly, information retrieval (IR) has always been a topic in cutting-edge research (Kadir et al., 2018). This is because the improvements in the fields over time have made the possibility of finding information efficient, due to the development of various IR mechanisms like ChatGPT, Gemini, and Perplexity that ensure information is easily found or obtained (Sambe,2017). However, ChatGPT is the most popular AI used in machine learning among these students.

ChatGPT, or Chat Generative Pre-Trained Transformer, is an artificial intelligence (AI) chatbot that uses natural language processing to create human-like conversational dialogue. It can respond to questions and compose various written content, including articles, social media posts, essays, code, and emails (Hetler, 2023). Furthermore, it is also defined as a conversational AI model that uses a machine-learning framework to communicate and generate intuitive responses to human inputs (Kanade, 2023). Moreover, as ChatGPT is trained on larger volumes of text, including books, articles, and web pages, it can generate accurate responses on diverse topics, from science and technology

to sports and politics. This chatbot is a perfect AI tool for students to use in searching for information efficiently.

Likewise, most higher education students, especially university students, use the ChatGPT tool efficiently while doing assignments or studying. Nowadays young generations are more familiar with technology, they already made the AI tool one of their reference material. However, the students should not be too dependent on the information provided by the ChatGPT as the accuracy level is not convincing. Therefore, the research aims to determine the accuracy of information retrieved by social sciences students using ChatGPT and analyze which measure of accuracy level has the most significant influence when using ChatGPT for social science research. Based on the findings, the researcher also provides evaluation strategies that can enhance the accuracy and credibility of information retrieved by social sciences students using ChatGPT.

1.2 Operational Definition

This study will explain in detail the technical terms of independent variables and dependent variables that the researcher has used.

1.2.1 ChatGPT

ChatGPT in this research refers to the involvement of ChatGPT in the information search process for social science students. In this research, the researcher aims to understand how ChatGPT impacts the information search accuracy of social science students. ChatGPT can be defined as an NLP model created by OpenAI based on the Generative Pretrained Transformer (GPT-3) architecture, which was initially developed for language creation tasks, such as machine translation (OpenAI, 2022; Qadir, J., Taha. 2020). It is designed to create human-like text based on a specific request or dialogue that enables

natural and open-ended conversations (OpenAI,2022). Unlike previous AI language models, ChatGPT can create new content and ideas through enhanced learning from human feedback and express them in real-time conversations. Moreover, the new development approach has enabled ChatGPT to respond to follow-up questions, acknowledge mistakes, refuse false assumptions, and reject inappropriate queries (Ngo, T.T.A, 2023).

Furthermore, ChatGPT has several techniques for collecting and analyzing information by using a large corpus of text data as input during the pre-training phase. It comes from a variety of sources like books, articles, websites, and other text-based sources. For instance, during pre-training, the model utilizes unsupervised learning to discover patterns and correlations in data. After pre-training, GPT can be tailored to specific tasks or domains, like text categorization or language translation. Fine-tuning requires training the model on a small dataset of labeled samples relevant to the job at hand. Next, it is through the input processing where GPT analyzes and interprets incoming inputs, including sentences and paragraphs, using pre-trained and fine-tuned models. This includes breaking down information into smaller parts, such as words or phrases, and examining their connections. GPT uses “conceptual analysis” to interpret information within its context. When analyzing input, it is important to evaluate individual words and their interactions with the surrounding context. After going through all the steps, GPT can analyze input and generate several outputs, including text answers, summaries, and translations. The model develops output depending on its knowledge of the input and patterns acquired during pre-training and fine-tuning. As a result, GPT is an extremely effective technique for gathering and evaluating natural language data. GPT uses AI and NLP to evaluate huge quantities of text data rapidly and accurately, making it a useful tool for various applications (Shubhrajyotsna Aithal & Aithal, 2023).

Therefore, due to the varied tools of ChatGPT, social science students are very active in involving the use of ChatGPT in their information-searching process. This is because it can help them with their assignments, provide feedback and revision guidelines, and provide writing assistance. Consequently, it can be used to help them develop their skills. and enhance the

learning process (Abdullayeva & Musayeva, 2023; Ausat et al., 2023). ChatGPT can be used as a quasi-virtual library for students because it provides users with access to a vast array of information and resources (Verma, M. 2023). It is facilitated for students as they can search for variations of information anywhere and anytime without the need to go to the library. It also helps to save time in searching for resources and the accuracy of information can be trusted to some extent after knowing the information collection process that ChatGPT goes through.

1.2.2 Social Science Students

Social science students in this research are defined as undergraduate students currently enrolled in a social science program at Universiti Teknikal Malaysia Melaka (UTeM). These programs encompass a variety of fields that study human behavior and social phenomena, including sociology, psychology, anthropology, political science, economics, and geography.

Therefore, in this study, the research targets social science students at the undergraduate level as they will be actively engaged in information search activities related to their research topic. This includes tasks such as analyzing how students use ChatGPT to locate datasets, statistics, and other quantitative data related to their social science research topic. This could involve students working on research papers, theses, dissertations, or other projects requiring in-depth exploration of a social science subject. From this, the researcher can examine how students leverage ChatGPT to identify relevant academic articles, journals, and research studies within the information needed. Furthermore, it is important to take social science students as participants because they have a foundational understanding of research methodologies relevant to the course. This ensures they are familiar with the process of formulating research questions, searching for credible sources, and critically evaluating information.

This research is suitable for social science students as they utilize ChatGPT as a tool to retrieve information for their research projects it can help students better understand concepts they are struggling with by providing customized, interactive explanations. Moreover, ChatGPT can be used to develop innovative projects and resources that are needed by the students (Kalla, Kuraku, et al., 2023). The researcher can see how the social science students will acknowledge the accuracy level of information from ChatGPT as they are good at analyzing information attentively.

1.3 Research Question

1. What kind of accuracy level of ChatGPT in information search by social sciences students?
2. How effective is the accuracy level of ChatGPT in information search by social science students?
3. Which accuracy level of ChatGPT has the most influence on information search by social science students?

1.4 Research Objective

1. To determine the accuracy level of ChatGPT in information search by social sciences students.
2. To evaluate the impact of the accuracy level of ChatGPT in information search by social science students.
3. To analyze which measure of accuracy level of ChatGPT has the most significant influence on information search by social science students.

1.5 Problem Statement

Artificial intelligence advancements have led to the invention of ChatGPT, an innovative system that responds to natural language instructions like a human and is increasingly being used in a variety of academic fields, including social sciences (Hetler, 2023). Research has shown that ChatGPT can be a beneficial tool for students in social sciences, delivering benefits such as its natural language generation capability and scalability making it ideal for applications like customer service chatbots and language translation (Ngo, T.T.A, 2023). It provides human-like, coherent responses, which improves user experience and satisfaction. ChatGPT's capacity to manage large volumes of conversations at once eliminates human involvement, enhancing efficiency and students' social satisfaction. Furthermore, ChatGPT's scalability provides quick responses and simultaneous management of large conversations, making it perfect for automated customer support or language translation services that reduce human participation while increasing user satisfaction. ChatGPT is also efficient as it generates responses quickly and handles several conversations, making it appropriate for jobs such as customer service and language translation, saving time and money through process automation (Kalla, Smith, et al., 2023).

However, concerns have been raised regarding the accuracy level of using ChatGPT in information search, particularly in the field of humanities and social studies. Issues such as authorship, information reliability, validity, relevance, and depth of information have created a sense of distrust in the information provided. This is because ChatGPT's large datasets may result in biased replies, spreading stereotypes or prejudice. Moreover, ChatGPT's limited knowledge base, derived from training data, might result in inaccurate or unhelpful responses to user inquiries, thus leading to irritation and an unpleasant user experience. While ChatGPT can assist students in tasks like generating responses to frequently asked questions, writing essays, and analyzing data, it is crucial to ensure that ChatGPT correctly replicates what it is designed to assess (Kalla, Smith, et al., 2023).

Based on the issues faced by the ChatGPT, a researcher found that more focus should be placed on ensuring all the information provided is relevant and authentic to the students. In this situation, the accuracy level of ChatGPT in information search by social sciences students is influenced by factors like task complexity, training data quality, and the need for human oversight to ensure unbiased and reliable results. Further research and guidelines are essential to maximize the benefits of ChatGPT while addressing the concerns associated with its usage in information search. Hence, a quantitative study is proposed to determine the accuracy of information retrieved by social sciences students using ChatGPT in terms of validity, reliability, and relevance. Furthermore, this research also wants to evaluate strategies that can enhance the accuracy and credibility of information retrieved by social sciences students using ChatGPT and to analyze which measure of accuracy level has the most significant influence when using ChatGPT for social science research. For this study, the respondents selected are social sciences students at the Universiti Teknikal Malaysia, Melaka. Thus, this topic is very interesting to research as it aims to provide students with a clearer understanding of ChatGPT's accuracy for information search, allowing them to leverage its benefits while maintaining critical thinking skills.

1.6 Significance of the study

The study fills a gap in the existing knowledge about the accuracy level of ChatGPT usage in information search by social science students, through quantitative analysis, within the context of Universiti Teknikal Malaysia Melaka, UTeM. While research on ChatGPT has grown significantly, studies addressing the level of accuracy are still limited. The recent review highlights the scarcity of research on the accuracy level of ChatGPT information search. This research fills this gap by focusing on the accuracy level as students rely heavily on accurate and verifiable information. This study holds significant value for several reasons.

Firstly, understanding the accuracy level of ChatGPT in information searches by social science students holds significant value in ensuring academic integrity. By evaluating the reliability and correctness of information retrieved through ChatGPT, this research aims to contribute to maintaining high academic standards. There are several strategies such as reduced plagiarism risk where identifying the accuracy level of ChatGPT- sourced information helps students avoid unintentional plagiarism. Furthermore, this study helps students to understand the limitations of ChatGPT that encourage them to distinguish between credible sources and potentially biased or inaccurate information generated by the tool. This research also encourages students to have ethical research practices as it can highlight the importance of critical evaluation of all information sources, including those retrieved through ChatGPT. This fosters ethical research practices by encouraging students to verify facts and acknowledge the limitations of AI tools.

Next, this study can promote critical thinking skills in students. This is because understanding the limitations of ChatGPT might assist social science students build the critical thinking skills required for evaluating information reliability and avoiding potential biases in AI-generated content. This will prepare students to be more discriminating researchers and consumers of information in the digital era. Moreover, it can enhance analytical abilities, and this research can contribute to a deeper understanding of research methodologies. Students learn to critically analyze information from ChatGPT, leading to stronger overall analytical abilities.

Finally, the research can inform educational practices. By identifying how to maximize accuracy and minimize pitfalls associated with ChatGPT, educators, and librarians can develop better strategies to guide social science students in effectively utilizing this tool. This will ultimately optimize educational practices and enhance students' research skills.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter focuses on defining and operationalizing the independent and dependent variables that will allow an author to assess the accuracy level of ChatGPT in information searches by social science students. The author conducted a comprehensive literature review, examining the relevant articles published in academic journals, reports, and books exploring ChatGPT's information retrieval capabilities. This varied set of resources allowed the author to capture theoretical frameworks and real-world implementations. To ensure the most relevant arguments for the study have been selected, the researcher extensively investigated present theories and models. Furthermore, following the same effective approach, the author defined the specific variables that would form the foundation of the research survey questionnaires. To strengthen the discussion, the chapter incorporates critical reviews, where the author adds valuable insights and opinions to enrich the discussion. Drawing upon all the gathered information, a robust theoretical framework will be constructed in Chapter 3. This framework will systematically trace the relationship between the independent variables and the dependent variables. Finally, chapter 3 will also convert these relationships into specific hypotheses that represent the ideas about the ways variables interact with each other.

2.2 The Accuracy Level of ChatGPT

According to Laskowski, N., & Tucci, L. (2022), the concept of artificial intelligence (AI) revolves around creating intelligent machines, particularly computer systems that can imitate human cognitive functions. This has led to breakthroughs in areas like expert systems, where machines can offer specialized knowledge, natural language processing for machines to understand and respond to human language, and speech recognition, enabling machines to interpret spoken language and machine vision and extract information from visual data. AI leads to various applications in different disciplines, such as education. In educational institutions, AI will make it possible to comprehend and manage a data-collecting procedure to incorporate it into an educational efficiency plan (Marino-Romero et al., 2022). However, a critical aspect of any educational tool is its accuracy. Inaccurate information or misleading responses can hinder learning outcomes.

Leveraging the power of GPT language model technology, this chatbot surpasses the limitations of a typical program. It tackles a wide range of text-based requests with remarkable skill. Whether users need clear answers to basic questions or assistance with more demanding tasks like essay writing or navigating conversations about productivity issues, GPT's sophisticated capabilities offer an exceptional level of support and understanding (Lund & Wang, 2023; Tlili *et al.*, 2023). While its versatility is promising for educational applications like question answering, summarizing factual topics, generating practice questions, and giving ideas for report writing, students need to be aware of potential limitations related to accuracy. For this study, the 3 independent variables for the accuracy level of ChatGPT namely validity, reliability (Meida Rabia Sihite *et al.*, 2023), and relevance (Zhang *et al.*, 2024).

2.2.1 Validity

Validity refers to whether the information is accurate and reflects the truth (Middleton, 2023)**.** In the case of ChatGPT, validity concerns the accuracy and truthfulness of the information it produces when responding to the student's prompts. It is important for students they use this tool to search for information about their assignments, check essay writing, ask questions, and more. They need valid information for references. With the validity of the ChatGPT's information, it can save students time in searching data physically at the library. However, AI tools like ChatGPT have limitations in giving the truth info. Several factors can compromise the validity of information generated by ChatGPT. As such, validity comprises four variables: data biases, contextual understanding limitations (McCoy, 2024), focus on fluency over accuracy, and lack of source attribution (Marr, 2023).

2.2.1.1 Data Biases

Data biases are systemic errors or inconsistencies within a dataset. These biases can impact the accuracy and fairness of analyses, machine learning models, and decision-making processes. The same goes for ChatGPT, the data produced from this AI may be inaccurate or biased, which users should be aware of to avoid replicating these errors in their work (Baidoo-Anu & Owusu Ansah, 2023). This is because ChatGPT is trained using enormous datasets gathered from the internet. This data may contain inherent biases due to the sources and opinions it represents. These biases can be apparent in ChatGPT's responses, resulting inaccurate or misleading information (Kalla, Smith, *et al.*, 2023). According to Hammoud (2023), the training data of ChatGPT, sourced from Websites, articles, and online forums reflects the biases of their creators, leading to discriminatory language patterns, and unbalanced perspectives that can influence ChatGPT's responses. These biases can have significant consequences. For instance, if ChatGPT analyzes historical data with biases against a specific social group,

it may reinforce mistakes and hinder social justice efforts. In education, ChatGPTs can provide different responses to the same question based on the wording of the prompt. This variability in responses can indicate bias in the model's understanding or interpretation of the input, leading to inconsistent or potentially skewed outputs (Fujimoto S and Takemoto K, 2023). Furthermore, ChatGPT's training data biases can influence creative outputs, such as poems or code. A lack of variety in authorship or perspective within the data may result in outcomes that reinforce stereotypes or fail to capture every aspect of the human experience. To minimize bias, selecting and curating the training data carefully and continually monitoring ChatGPT's responses to identify and correct potential biases is essential (Kalla, Smith, *et al.*, 2023).

2.2.1.2 Contextual Understanding Limitations

ChatGPT, a powerful language model, faces limitations in understanding complex contexts and language which can lead to misinterpretations of user prompts and inaccurate or irrelevant responses (Kalla, Smith, *et al.*, 2023; Ray, 2023). Despite its advanced capabilities, the model struggles with long-term context, ambiguity, and topic shifting, making it challenging to maintain consistency and coherence in conversations (JACOBY, 2023). For example, when students have a conversation, they refer back to something mentioned earlier. ChatGPT might struggle to maintain a coherent thread across extended dialogues. Its focus on the immediate prompt can lead it to overlook previously established context, resulting in disconnected or irrelevant responses. Moreover, language is rife with ambiguity. Sarcasm, double meanings, and implicit references can easily trip up ChatGPT. The model might interpret a sarcastic statement literally or miss the underlying sentiment entirely. This can lead to awkward or misleading responses that fail to capture the true intent of the user. Natural conversations often flow organically, shifting topics as the dialogue

progresses. ChatGPT, however, might struggle to keep up with these shifts. It might cling to the original topic or abruptly jump to unrelated areas, creating a disjointed and confusing experience for the user (JACOBY, 2023). The consequences of these limitations can be frustrating for users. For example, students might be providing detailed instructions, only to have ChatGPT misinterpret a key step due to its limited grasp of the overall context. To improve the contextual understanding of large language models could involve supplementing training data with factual knowledge bases that can provide ChatGPT with a broader understanding of the world and how concepts interrelate and developing more sophisticated attention models that allow ChatGPT to focus not just on the immediate prompt but also on relevant elements of the conversation history. By addressing these limitations, large language models like ChatGPT can evolve into even more powerful tools for communication and understanding. However, it is important to remember that achieving a true human-level understanding of language remains a complex challenge. As research progresses, we can expect ChatGPT and similar models to become more adept at navigating the complexities of human conversation, ultimately leading to richer and more meaningful interactions.

2.2.1.3 Focus on Fluency over Accuracy

A large language model like ChatGPT has great potential for communication and information access. However, important questions regarding how users ensure the validity of the information they generate have arisen. This study explores the possible pitfall of prioritizing fluency over accuracy in ChatGPT responses and the fine dance required to maintain its trustworthiness. One major concern is the generation of reasonable but imprecise responses. ChatGPT, in its effort to sound convincing, may create a narrative that deviates from reality. Imagine a student asking about a historical event. ChatGPT can make up a compelling story, but it is riddled with factual inconsistencies (Wankhede, 2024). This can be particularly

challenging because fluency may influence users and overlook inaccuracies. Furthermore, fluency-driven responses can unintentionally spread misinformation and misunderstandings. Moreover, ChatGPT which cannot distinguish between fact and fiction, can confidently assert fantastical claims. While these may be expressed poetically, they have no basis in fact and can mislead consumers who accept them at face value. Another challenge is contextual awareness. ChatGPT cannot distinguish between scientific inquiries, fictional narratives, and folklore. Depending on the context, this can lead to inappropriate responses (Marr, 2023). For example, a user seeking medical advice may receive a beautifully crafted but medically unsound recommendation. This potential danger highlights the importance of balancing fluency and accuracy. Achieving this balance is not easy. A thorough focus on accuracy can result in robotic and uninteresting language, defeating the purpose of user interaction. However, sacrificing accuracy completely undermines ChatGPT's credibility. Users rely on it for factual information, and consistent fluency without accuracy can erode their trust (Haven, 2023). User expectations also play a role. The fluency of ChatGPT's responses can lead users to assume a depth of knowledge that may not exist. They might not critically evaluate the information received, potentially accepting inaccuracies as truth. The ethical implications of misleading information are significant. Imagine ChatGPT, in its fluency way, advising someone to ingest harmful substances. This underscores the need for an ethical approach that prioritizes both fluency, which keeps users engaged, and accuracy, which ensures they receive reliable information. Fortunately, ChatGPT can learn iteratively through user feedback. As users correct inaccuracies, they can adapt and improve their responses. This continuous learning process allows it to refine the delicate dance between fluency and accuracy. In conclusion, ChatGPT's validity rests on its ability to maintain a balance between captivating fluency and unwavering truth. It must navigate a tightrope, neither sacrificing its ability to engage users nor forsaking its responsibility to provide accurate information. As users, we need to be smart when using ChatGPT. Even though it can write well, it is important to check the facts it gives us. By working together, we can help ChatGPT get better at giving us the correct information.

2.2.1.4 Lack of Source Attribution

ChatGPT's ability to generate human-quality text holds immense potential. However, its lack of source attribution presents a significant challenge to its validity and ethical use. This essay explores the importance of source attribution and proposes strategies to address this issue. One major concern is the risk of plagiarism (Hosam Alamleh *et al.*, 2023). ChatGPT, trained on a massive dataset of internet text, might inadvertently reproduce existing material without acknowledging the source. Imagine a student using ChatGPT-generated text for an essay without proper citations. This could lead to accusations of plagiarism, highlighting the critical need for source attribution. Furthermore, the lack of attribution can contribute to the spread of misinformation (Harrison, 2022). When ChatGPT presents information without citing sources, users have no way to verify its accuracy. This can be particularly dangerous when dealing with sensitive topics. Proper source attribution allows users to evaluate the credibility of the information presented. Source attribution also enhances trust and authenticity. When users know where information comes from, they are more likely to trust it. Without clear attribution, ChatGPT's responses may appear less reliable, hindering user confidence. This is particularly important in academic settings, where proper citation is essential for building strong arguments and avoiding plagiarism. Unintentionally submitting unattributed content can have serious consequences for students. Another practical concern is detection by text-matching software (Sain & Negi, 2024). Educational institutions often use plagiarism detection tools. Encouraging users to explicitly cite sources helps responses avoid detection and ensures academic integrity. Several approaches can mitigate the risks associated with a lack of source attribution in ChatGPT. Implementing honor codes that emphasize originality and proper citation can create a culture of academic integrity among users. Acknowledging ChatGPT's limitations is also crucial. Educators should emphasize that it lacks real-time access to information, and users should verify their responses with credible sources (Koblyakov, 2024). Designing assignments that require data extraction from images or visual sources can also incentivize original research and analysis. Integrating oral

discussions into assessments allows educators to evaluate students' understanding beyond text-based responses generated by ChatGPT. While fluency and coherence are key strengths of ChatGPT, its validity and ethical use hinge on proper source attribution. By implementing the strategies outlined above, we can ensure that ChatGPT is a valuable tool for communication and learning, promoting academic integrity and responsible information consumption.

2.2.2 Reliability

According to the Cambridge Dictionary, reliability refers to the quality of being reliable or providing consistently accurate information. This is important when using AI tools like ChatGPT, especially in the academic field to prevent users from getting inaccurate information. Unreliable information can have serious consequences, leading to plagiarism, wrong conclusions, and wasted time (Shah *et al.*, 2022). For instance, unreliable information can lead to poor decision-making. In research, incorrect data or inaccurate information can skew results and lead to flawed conclusions. These flawed conclusions can then misinform policy decisions or hinder scientific progress. In conclusion, reliability is the foundation of trustworthy information. By prioritizing reliable information, we ensure the quality of research, foster sound decision-making, and prevent misinformation. This is especially important in academic settings where the pursuit of truth and knowledge is paramount. however, several factors affect the reliability of ChatGPT such as training data, user input, and outdated data (Saravanan, 2022).

2.2.2.1 Training Data

ChatGPT's capabilities are based on its training data. This enormous text and code dataset underlies the system's ability to generate human-quality writing and reply to user questions (AI,2024). However, the quality and diversity of this data are essential in establishing the accuracy of the information it generates. One major concern is the possibility of bias and inaccuracies within the training data (Sutaria, 2022). Training ChatGPT on a data set heavily skewed toward a particular political ideology will result in the information generated possibly reflecting that bias, gradually encouraging people to a particular opinion. Furthermore, if the data includes factual errors or outdated information, ChatGPT may unintentionally spread these shortcomings in its responses. This emphasizes the significance of using high-quality, thoroughly validated data during the training process. In conclusion, ensuring the quality of training data is paramount for fostering trust in ChatGPT. Developers can make ChatGPT a credible and dynamic tool for users seeking knowledge and engaging in creative inquiry by painstakingly screening data for biases, inaccuracies, and obsolete material, as well as including a diverse range of sources and opinions. However, responsibility does not stop there. Users must also be attentive, keeping a critical eye on the data supplied by ChatGPT and cross-referencing it with trustworthy sources as needed.

2.2.2.2 User Input

While ChatGPT's ability to process information and generate human-like text is impressive, its effectiveness relies on a crucial factor, especially the users like students. The way users interact with ChatGPT has a considerable impact on the reliability of the information it offers. This human touch, through the questions they ask, can be a double-edged blade. On one hand, precise and well-defined questions can help ChatGPT reach its full potential, whereas confusing or misleading questions can send it off track.

One major challenge lies in the realm of ambiguous queries. According to SAP Help Portal (2023), it is defined as one where there is no specific query, it could have more than one meaning, asking for several responses, or not clearly defining the subject or object. Imagine a student asking about the way they can get good grades. This lacks the specific details about the subject or the student's learning style. ChatGPT might provide generic advice that is not particularly helpful. Thus, it will highlight the necessity of students asking precise, well-defined inquiries. Users may help ChatGPT generate more accurate and relevant information by clarifying goals, providing details, and avoiding too general terminology (Jones, 2024). To conclude, the reliability of information obtained via ChatGPT is not simply based on its internal workings. The way students or users interact with it, specifically, the clarity and precision of users' prompts, is crucial. Recognizing unclear queries is key to user ChatGPT collaboration for reliable, and informative findings.

2.2.2.3 Outdated Data

According to Koenders & Konning (2023), outdated data hinders ChatGPT's ability to understand newer information. This limitation probably will be short-lived, as developments are rapid. However, the constant flow of information tests ChatGPT's reliability. Consider asking about a recent education breakthrough. If the training data has not been updated in a while, ChatGPT may be uninformed of the latest development. Outdated information can mislead users and hinder their ability to make informed decisions, especially on rapidly evolving topics (Loh, 2022). Thus, regularly updating the training data with the latest recent research, news, and discoveries is essential for ensuring ChatGPT gives accurate and relevant information.

2.2.3 Relevance

Relevance is one of the most important points that need to be taken care of when looking for information using ChatGPT. This is because it determines how well the AI model responds to user inquiries. ChatGPT's ability to assess relevance has an impact on the quality of the information provided, ensuring that responses are closely aligned with the user's expectations and requirements. Besides, maintaining a high level of relevance increases user satisfaction, improves productivity, and facilitates better decision-making processes, making ChatGPT a more effective tool for various tasks, including answering inquiries and generating content. However, several factors can impact the irrelevant information provided by ChatGPT such as lack of keyword appearance, content quality, user interaction, and link structure (Carlson, 2023).

2.2.3.1 Keyword Appearance

Keywords play a crucial role in guiding AI models like ChatGPT to understand the context and generate accurate responses. Keyword appearance is defined as the function of obtaining information about users and their behavior (bomba, 2015). If the keywords are unclear, missing, or misused, it can have a major impact on the relevance of the information produced by ChatGPT. Keywords are essential in helping ChatGPT understand the context and generate proper responses. When keywords are missing or not used correctly, ChatGPT may fail to recognize the user's purpose, producing irrelevant or incorrect information. This can lead to ChatGPT returning results that do not match the user's inquiry or expectations, reducing the overall quality and utility of the interaction (somlith,2023; Team, 2023). As a result, ensuring the relevant keywords are successfully incorporated into prompts is critical for improving the efficiency and relevance of ChatGPT in generating meaningful responses.

2.2.3.2 Content Quality

ChatGPT prioritizes high-quality, unique content and will prefer to rank pages with well-written, informative material over those with low-quality or duplicate content (Carlson, 2023). This is because content quality plays an important role in determining the relevance and accuracy of information provided by ChatGPT. High-quality content improves ChatGPT's ability to provide meaningful and appropriate responses that correspond closely to user requests and expectations. On the other hand, poor content quality might cause ChatGPT to provide irrelevant or erroneous information, negatively impacting the overall user experience and reducing ChatGPT's effectiveness (community, 2023; Rodgers, 2023). For instance, if a social science student asks ChatGPT about the correlation between social media usage and depression rates among teenagers, the presence of quality content on this topic would enable ChatGPT to provide informed responses that align with the student's query. But, if the content available is sparse, outdated, or lacks depth and credibility, ChatGPT may struggle to offer relevant and accurate information, leading to potentially misleading or incomplete responses for the student's research inquiries. In conclusion, users like students can dramatically increase the quality of ChatGPT responses by ensuring that the content is well-structured, useful, and contextually relevant. This reduces the possibility of irrelevant information being shown.

2.2.3.3 User Interaction

User interaction involves individuals engaging with artificial intelligence technology by providing prompts or queries, and ChatGPT responds by generating conversational interactions (Fergus et al., 2023). According to Carlson (2023), ChatGPT considers how users interact with a website, such as how long they stay on the site, if they click on links, and whether they return. Popular websites with high levels of user involvement are more likely to rank well in ChatGPT. When users interact with ChatGPT,

their queries, communication style, and question clarity contribute to helping the AI model produce relevance responses. Users play a crucial role as they have a substantial impact on the relevance of the information supplied by ChatGPT. Effective user interaction, defined by clear and specific queries, can assist ChatGPT have a better understanding of the user's intent and providing more exact information. While, giving incomplete or confusing interactions might lead to erroneous responses from ChatGPT, as the AI model relies on human input to generate acceptable outputs (Bustamante, 2023). For example, when a student asks a vague question about the ways society impacts individuals without specifying the context or variables of interest, the AI may struggle to generate a relevant and accurate response. As a result, the quality and effectiveness of user interactions with ChatGPT have a direct impact on the relevance and accuracy of the data generated in response to user inquiries.

2.2.3.4 Link Structure

In this context, ChatGPT analyses the structure of links within and between websites to identify the value and relevance of pages. A page with numerous high-quality, relevant inbound links, for example, is more likely to rank higher than one with fewer or lower-quality links (Carlson,2023). Nevertheless, a disorganized or spammy link structure may lead ChatGPT to irrelevant or low-quality sources, causing inaccurate or misleading responses. Therefore, the link structure plays a crucial role in shaping the information ChatGPT processes and subsequently influences the relevance and accuracy of the responses it generates (Templeton, 2024).

2.3 Information Search by Social Science Students

Information searching, also known as querying, is a well-defined process of seeking targeted information based on a clear understanding of the user's needs (Fulton & McGuinness, 2016). Understanding how students look for and analyze information is critical to improving information literacy and research abilities in this academic area. Sociology, psychology, anthropology, economics, political science, and other disciplines are all part of the social sciences (Rindawati et al., 2021). Each discipline has its methods, theories, and sources of information. Consequently, social science students' information-seeking behavior may differ depending on the specific criteria and conventions of their field of study. However, a key aspect of information search behavior is understanding students' information needs and goals.

Studies by Johnson (2019) and Wang et al. (2020) have highlighted that social science students frequently seek information to support their research projects, coursework, literature reviews, and thesis work. These information requirements can vary from empirical data and statistical analysis to theoretical frameworks and scholarly literature. Next, social science students conduct research using various kinds of information sources. Library databases, academic journals, books, government documents, and reliable websites are commonly used sources. However, technological advancements have transformed the information search landscape for social science students. Digital libraries, online databases, search engines, citation management systems, and AI-powered platforms like ChatGPT have all become essential components of the information retrieval process. These technological tools provide accessibility, and efficiency but also raise concerns about information relevance and validity. The introduction of AI tools like ChatGPT introduces new variables into information search behavior, potentially influencing aspects like frequency of information seeking, the channel of awareness, the purpose of use, the goodness of the result, and self-assessment of ChatGPT competence in utilizing the tool (Karunaratne & Adesina, 2023).

2.3.1 Frequency of Information Seeking using ChatGPT

The frequency of information-seeking refers to how often students actively search for information to meet their academic needs (Tubachi,2018). According to Darrel et al. (2023), the study on ChatGPT usage among programming students indicates that they explore the tool primarily for project tasks, assignments, and personal curiosity which matches their reasons for seeking information. The study, shows students choose ChatGPT above other AI tools for information retrieval due to its perceived effectiveness. The frequency of using ChatGPT by students is also the highest and get some good comments while using ChatGPT (Firaina & Sulisworo, 2023). Even though ChatGPT has some challenges, like network issues and difficulty in formulating queries, students usually trust the tool and find it useful in giving relevant information, encouraging more frequent use. The increased usage and trust in the tool indicate that ChatGPT has the potential to influence the frequency of information-seeking among social science students.

2.3.2 The Channel of Awareness

Although the AI tool is commonplace nowadays, the popularity among the population studied is not as high as typically forecasted (Rudolph et al., 2023). Channel of awareness is defined as the familiarity consumers have with a particular product or service (Kopp, 2022). According to a study conducted by the Pew Research Center, two-thirds of U.S. teens, particularly those in the 12-17 age group, are already aware of the existence of ChatGPT. Otherwise, in the survey results in existing studies, students have become aware of the tool via different channels, mainly friends and social media (Kanade, 2023). This aligns with the general trend of young adults using social media to discover new technologies and tools. However, in educational institutions, Universities or colleges may introduce students to ChatGPT through workshops, seminars, or by incorporating it into their teaching

techniques. As the tool's popularity develops, this channel is likely to gain relevance. In conclusion, social sciences students might be familiar with the ChatGPT as they have many friends and are already exposed to technological advances.

2.3.3 The Purpose of Use

The "purpose of use" refers to the specific reasons or tasks for which individuals employ a particular tool or service. In the context of information search, understanding the purpose of use for ChatGPT among social science students is crucial to optimizing its role in their research practices. ChatGPT can be used to identify relevant academic literature by summarizing research topics, identifying keywords, and creating preliminary literature review drafts (OpenAI, 2023). This can save students time and effort during the initial stages of research. Based on the survey of this study, users mostly used the tool when carrying out their project tasks or writing their assignments as a part of their academic purposes. On top of that, some users utilize it based on pure personal interest (Kanade, 2023). This trend aligns with the broader reasons individuals seek information, as identified by Tubachi (2018), which can include completing academic work. In conclusion, this purpose aligns with social science students' proposition as they always need to search for information about academic work.

2.3.4 The Goodness of Result

In the artificial intelligence (AI) tools, ChatGPT is one of the most popular compared to other competitive tools. This is because the result of findings from ChatGPT is more effective compared to other competitive tools. Apart from that, ChatGPT has been improved much better from the

existing format. This makes sense for students to trust the tools compared to other conventional AI applications and find them effective in providing relevant information for their academic needs (Kanade, 2023; Jowarder, 2023). Moreover, the ChatGPT has inserted a lot of data, it can help students generate ideas, and divergent thinking, improve problem-solving, and enhance collaboration (OpenAI, 2023). ChatGPT improves the academic achievement of social science students by helping them understand challenging concepts and giving relevant study materials. The tool's ability to support students' information demands and give valuable information may contribute to improved academic outcomes (Jowarder, 2023).

2.3.5 Self-Assessment of ChatGPT Competence

In the context of social science students' information searches, self-assessment of ChatGPT competency refers to a student's perception of their ability to use ChatGPT effectively for research purposes. Understanding this perception is important to promoting responsible and critical use of the tool. This is to help students to not overconfidence and misinformation when taking information from the ChatGPT. Other than that, the self-assessment of students' competency is a strong measure of their information literacy, leading to reduced anxiety and improved confidence in using the tool in the future for those who are underconfidence and missed opportunities of the ChatGPT tools (Kanade, 2023). To conclude, self-assessment competence helps the researcher to identify the way to balance the confidence of social science students to harness ChatGPT effectively in their research endeavors.

2.4 Theoretical Framework

The theoretical framework is defined as a structure that guides research by relying on a formal theory, constructed by using an established, coherent explanation of certain phenomena and relationships (Grant & Osanloo, 2014). In this research, the framework shows the independent variables and dependent variables that will be examined. The independent variable is the accuracy level of ChatGPT, which is validity, reliability, and relevance. The dependent variable is information search by social science students. It is important to clearly understand the independent and dependent variables.

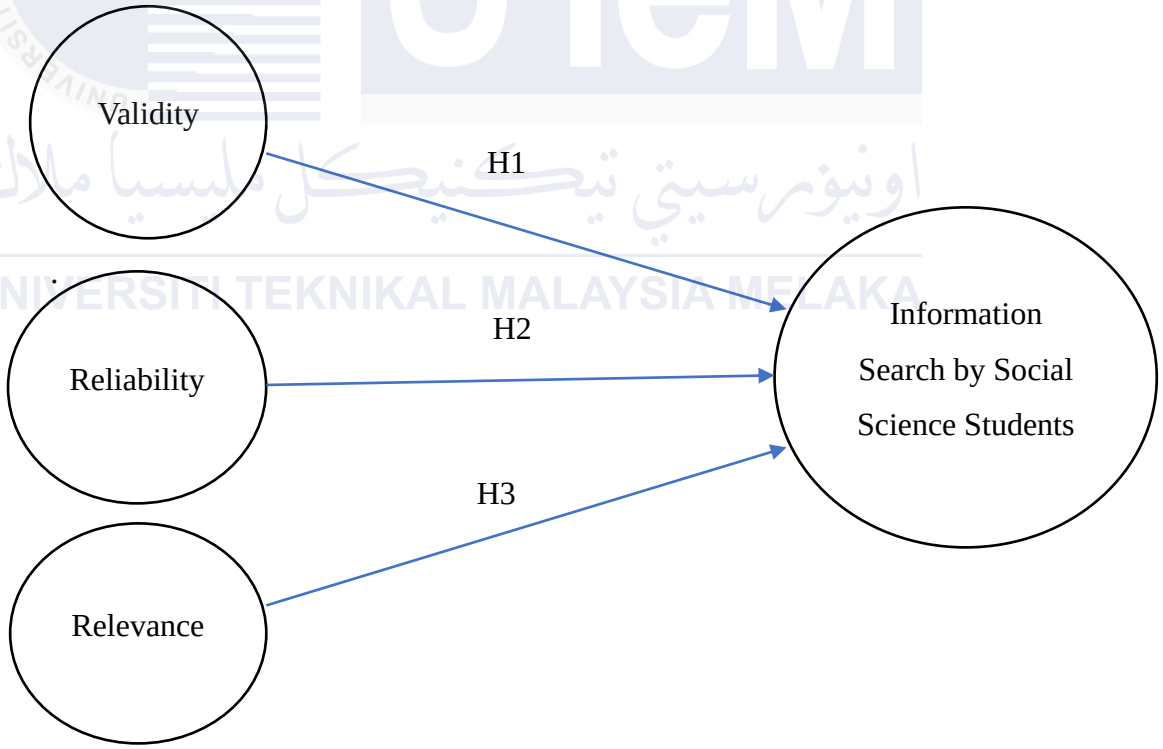

Figure 2.1: Proposed research model by this researcher

2.5 Research Hypothesis

The hypothesis of the study is based on the literature review done by this researcher.

Hypothesis 1

Ho: There is no significant relationship between validity and information search by social science students.

Ha: There is a positive significant relationship between validity and information search by social science students.

Hypothesis 2

Ho: There is no significant relationship between reliability and information search by social science students.

Ha: There is a positive significant relationship between reliability and information search by social science students.

Hypothesis 3

Ho: There is no significant relationship between relevance and information search by social science students.

Ha: There is a positive significant relationship between relevance and information search by social science students.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

In this chapter, the researcher will explain the method adopted by this research. Research methodology is the specific procedures or strategies used to gather, select, process, and analyze information about a topic. In a research paper, the methodology section allows the reader to examine the overall validity and reliability of the study critically (Heever,2022). This research uses quantitative study, some steps must be followed. Thus, the procedures and methods used also need to be explained. Furthermore, several components should be included in this research methodology such as research design, research method, research instruments, research location, time horizon, and data analysis method. It is a sequential method process that should be followed to conduct research.

3.2 Research Design

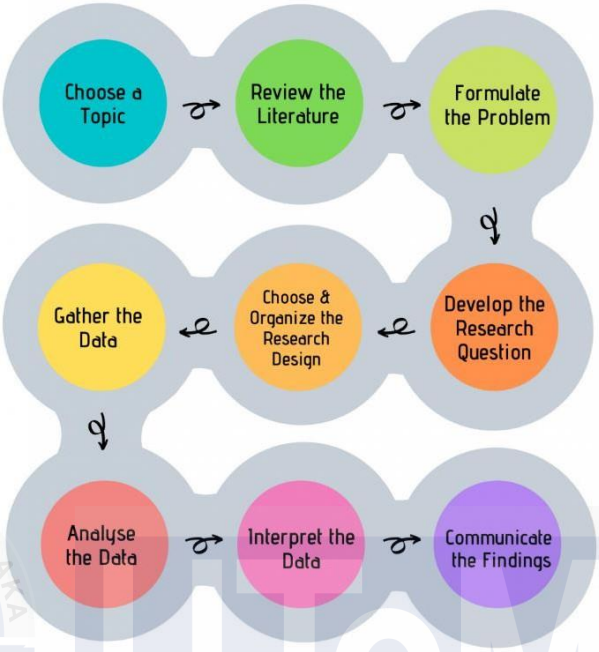

Figure 3.1: Stages of Research Process (Source: Sheppard, 2020)

According to Akhtar (2016), a research design is the arrangement of conditions for data collection and analysis in such a way as to combine relevance to the research purpose with economy and method. Furthermore, it is also a plan, structure, method, and research concaved to acquire a guaranteed search question and control variance. Research design is the overall plan or structure guiding the research process (Alam,2023). The research design aims to create a suitable framework for a study. The selection of a research approach is a critical design in the research process as it defines how relevant information for a study will be gathered (Abu-Taieh *et al.*,2020)

3.3 Research Method

According to the University of Newcastle Library (2023), research methods are the strategies, processes, or techniques used to collect data or evidence for analysis to discover new knowledge or gain a better understanding of an area of study. This study will test the hypothesis and describe the link between independent variables and dependent variables. To the test, three methods can be applied in this research, which are qualitative, quantitative, and mixed methods. However, this researcher chooses to use quantitative studies as it is more suitable for accomplishing this research.

Quantitative research is a form of research that uses natural science methodologies to generate numerical data and actual truths. Its goal is to establish a cause-and-effect link between two variables using mathematical, computational, and statistical methodologies. Moreover, the research is also referred to as empirical research because it can be classified, ranked, or measured using units of measurement. Thus, graphs and tables of raw data can be created using quantitative research, making it easier for the researcher to analyze the results (Ahmad *et al.*, 2019).

3.4 Location of Research

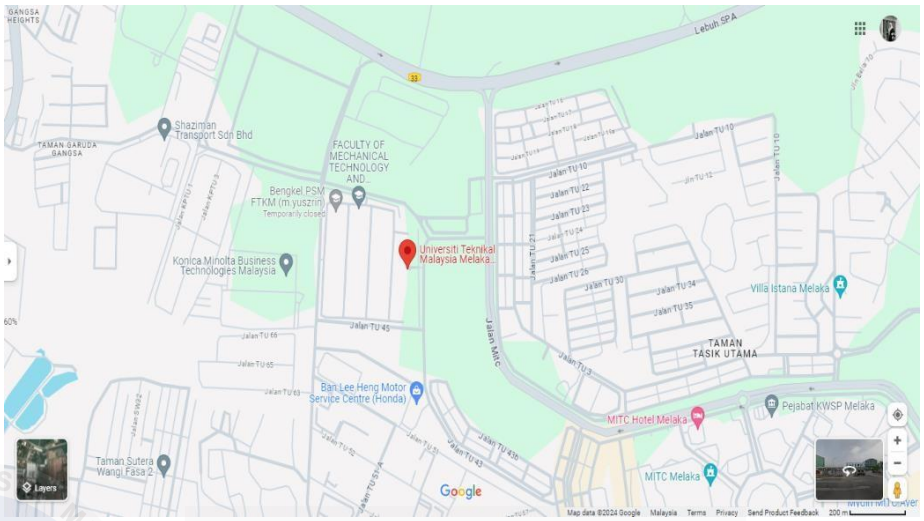

Figure 3.2: The location of Universiti Teknikal Malaysia Melaka (UTeM)

Source: Google Maps

Figure 3.2 above shows the location of the research that has been conducted at Universiti Teknikal Malaysia Melaka (UTeM). This location had been selected as UTeM has a growing number of students enrolled in social science-related programs that can be the respondent and it is also near the researcher places. Furthermore, UTeM has a strong emphasis on research and innovation. Conducting this study at UTeM aligns with the institution's objectives to explore and integrate new technologies into education and research. Lastly, students at UTeM are diverse in terms of backgrounds, which allows for varied perspectives and testing scenarios. This diversity can contribute to a more comprehensive assessment of ChatGPT's accuracy in catering to different information needs.

3.5 Research Strategy

A research strategy is an organized strategy of action that a researcher follows throughout the study process. It establishes a framework for conducting research efficiently and effectively. It also helps to choose the right data collection and analysis procedure (Walia & Chetty, 2020). Besides, it helps a researcher to decide the instrument that will be used in this study. Under the quantitative strategy, a researcher can collect the data by using survey questionnaires, experiments, secondary data analysis, and structured observations (Hassan, 2022). In this study, a quantitative study questionnaire was used as the main source of data. A random sample will be drawn and the survey administered, containing a Likert Scale designed to address the research questions.

3.5.1 Survey Method

In this research, the researcher used the survey method to collect the data as human behavior can be illustrated and implemented by the result of a survey method (Ponto,2015). According to Formplus Blog (2021), a survey method is a procedure, tool, or approach for gathering information in research that involves asking questions to a designated set of individuals. It typically allows the transmission of information between research participants and the person or organization conducting the study. A survey technique can help the researcher collect quantitative data, analyze descriptive and inferential statistics, and identify potential relationships between independent and dependent variables (Saunders *et al.*,2016). Surveys come in 2 major formats which are paper forms or online forms. The study will use survey online forms as they are easy to administer because they can be sent to respondents via email or social media.

Besides, the questionnaire will be used for the survey strategy to study quantitatively after collecting standardized data. This is because questionnaires are the best instrument for collecting large amounts of data and information with a huge population of respondents in a cost-effective way. A questionnaire is a collection of questions or objects that seek information from respondents about their attitudes, experiences, or opinions. It can be used to obtain quantitative and qualitative data (Bhandari, 2021). Thus, the survey method was appropriate for data gathering and analyzing the relationship between the accuracy of information retrieved by social sciences students using ChatGPT.

As is known, the survey method was usually associated with a deductive approach through a web questionnaire, which is a quantitative method. The web questionnaire uses Google Forms to represent the survey questions, answered, and submitted online by every respondent. Google Form has been chosen as it is free to use and has integration with Google services which makes the process of collecting data and analysis easier.

3.5.2 Questionnaire Method

In this study, the questionnaires will be distributed to a systematically selected sample of social science UTeM students to collect primary data for this research. The researcher designed the questionnaire to examine the accuracy level of ChatGPT in information searches by social science students. In the questionnaire, there will be three sections designed for survey research.

The first section would be contained to analyze the demographic of respondents. Closed-ended multiple-choice questions were used to assess respondents' demographic attributes. In the second section, the researcher will focus on the independent variable which is the accuracy level of ChatGPT in information searches by social science students. Lastly, the third section determines the information searchers by social science

students. The questionnaire was evaluated on the Likert Scala format. Likert Scala is a type of rating scale, commonly found on survey forms or questionnaires, that assesses the way people feel and degrees of agreement and can be beneficial in many various scenarios (solmaz,2023). It has a five-point rating scale where 1 represents "strongly disagree", 2 represents "disagree", 3 points represent "natural", 4 represents "agree", and 5 represents "strongly agree". According to research (Joshi *et al.*, 2015), the Likert scale was developed to measure 'attitude' in a scientifically acceptable and proven manner. An attitude can be characterized as favored methods of behaving or reacting in a certain context founded in the generally long-lasting organization of beliefs and ideas acquired through social interactions. The three sections of the questionnaires are shown in Table 3.1.

Table 3.1: Questionnaire Method

Section	Content
A	Respondent Background: • Gender • Age • Course • Year
B	Assessment of Independent Variable • The Accuracy Level of ChatGPT (Validity, Reliability, and Relevance)
C	Assessment of Dependent Variable • Information Searchers by Social Science Students

Table 3.2: Likert-Scale Survey

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

3.5.3 Population and Sample

The target population for this research is Universiti Teknikal Malaysia Melaka (UTeM) students. The researcher aims to pass out the question to social science students from Year 1 until Year 4. This population is chosen because it directly relates to the research questions about the kind of accuracy level of ChatGPT in information search by social science students, the impact of the accuracy level of ChatGPT in information search by social science students, and the accuracy level of ChatGPT has the most influence on information search by social science students.

The sampling method is the act of examining a population by acquiring and evaluating data. It is the basis of the data where the sample space is huge (BYJU'S, 2023). Besides, sampling techniques have been categorized into two major types which are probability sampling methods and non-probability sampling methods. In this study, the researcher chose probability sampling as the sampling method. This is because it involves randomly selecting individuals from a population with an equal chance of being chosen.

The researcher used a simple random sampling method for the selection process of sampling. Alvi (2016) states that simple random sampling has an equal chance of selecting each element of the population. The simple sampling method ensures that each unit in the sample has an equal probability of being included. If the population is homogeneous, this technique provides a neutral and more accurate parameter estimate (Singh and Masuku, 2014). Thus, the researcher has identified that there are 2220 social science students at UTeM. According to Table 3.3 from Krejcie and Morgan (1970), the researcher needs to distribute the questionnaire to 327 social science students at UTeM through Google Forms.

Table 3.3: Krejcie and Morgan (1970) sample determination table.

Source: Kakumbi & Phiri (2022).

Table for Determining Sample Size from a Given Population					
N	**S**	**N**	**S**	**N**	**S**
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
N is population size.	**S is sample size**				

3.5.4 Pilot Testing

Pilot testing is a small-scale study undertaken before an actual experiment to test and refine processes. According to Tan (2024), a pilot test is a small-scale evaluation of the whole research plan, including instruments, procedures, sampling, and data analysis. The primary goal of a pilot test is not to answer specific research questions but to prevent researchers from launching a large-scale study without adequate knowledge of the method proposed, in essence, a pilot test is carried out to prevent the possibility of an unfortunate law in a study that costly in terms of time and money (Pilot & Beck, 2017). Furthermore, pilot tests also are to improve the questionnaire's reliability, validity, and practicability. Piloting is administering the questionnaire to a sample of respondents who represent the target study sample and using statistical analysis and feedback to limit the number of items in the questionnaire to a manageable quantity (Cohen et al.,2013). In the pilot test, respondents' recommendations and information were gathered to create the final survey questionnaire. There are at least 30 participants who will be chosen for the pilot test (Yahaya, Onipe. 2024).

3.6 Time Horizon

The time horizon refers to the duration taken to conduct the research (Saunders *et al.*,2016). There are two types of time horizons which are cross- sectional studies and longitudinal studies. In this study, the researcher chose to conduct the research using a cross-sectional study because of the limited to the short amount of time for obtaining an analysis of data and concluding within a period of this research. Wang & Cheng (2020) defined cross-sectional studies are observational studies that look a data from a population at a certain point in time. They are frequently used to evaluate the prevalence of health outcomes, analyze health factors, and describe the characteristics of a population. Moreover, cross-sectional is

also known as studying a phenomenon at a specific time (Saunders *et al.,*2016). The researcher conducted this study in a short period from March 2024 until February 2025. The researcher distributed the questionnaire to the respondents from October 2024 until December 2024.

3.7 Validity

Surucu & Maslakci (2020) state that validity is defined as whether the measuring instrument measures the behavior or quality it is intended to assess. It is also a measure of how well the measuring instrument accomplishes its function. In this study, validity tests are used to ensure that scale expressions accurately measure the intended goal of the research. Therefore, testing the validity of the measurement instrument is more challenging, but more important than determining its reliability. For the research to be effective, the measuring equipment must accurately reflect the intended results. Thus, using an approved measurement instrument ensures accurate analysis results. The researcher chooses to use internal validity which refers to the degree to which a research study develops a reliable cause-and-effect relationship (Cuncic, 2022). It is suitable for this research as the researcher wants to know if there is any relationship between the two variables. A survey questionnaire's internal validity refers to the statistical correlation of a group of questions with an analytical factor or result.

3.8 Data Analysis

In this chapter, data analysis has been made to measure the regression, Pearson's Correlation Coefficient, and demographic data of the respondents.

3.8.1 Reliability

To collect the right data and information, this research must be valid and reliable to test. Reliability refers to the ability to measure components to produce consistent results when used at different times (Surucu & Maslakci, 2020). It is also described as the level to which measures are accurate when various persons measure on different occasions, under different situations, supposedly using separate devices that assess the construct or skill (Edwin,2019). To maximize reliability, statistical tests such as Cronbach's alpha evaluate the internal consistency of major survey variables, providing a quantifiable estimate of response consistency. It consists of an alpha coefficient between 0 to 1 in Cronbach's Alpha.

Table 3.4: Cronbach's Alpha Coefficient Range and Strength of Association

Source: (Zahreen Mohd Arof *et al.*, 2018)

No	Coefficient of Cronbach's Alpha	Reliability Level
1	More than 0.90	Excellent
2	0.80-0.89	Good
3	0.70-0.79	Acceptable
4	0.6-0.69	Questionable
5	0.5-0.59	Poor
6	Less than 0.59	Unacceptable

Table 3.4 based on Zahreen Mohd Arof *et. al* (2018) shows Cronbach's Alpha Coefficient Range and Strength of Association. According to the table above, Cronbach's Alpha values equal to or greater than 0.70 were considered acceptable. However, Cronbach's Alpha values

of 0.80 are considered good, while 0.90 or higher are regarded as excellent. It is poor if Cronbach's Alpha is less than 0.60 and unacceptable if Cronbach's Alpha is less than 0.59.

3.8.2 Linear Correlation

Correlation determines the degree of relationship between an independent variable and the dependent variable in question. According to Senthilnathan (2019), the correlation coefficient is a statistic that quantifies the degree of link between the two variables. Applications have two types of correlation coefficients such as the Pearson Correlation Coefficient and Spearman's Rank Correlation Coefficient. This paper focuses on the applications of the Pearson Correlation Coefficient in determining the relationship between independent and dependent variables.

Pearson's Correlation Coefficient (r) is a measurement quantifying the strength of the association between two variables. In this case, Pearson's Correlation Coefficient was used to determine the strength of the connection between the accuracy level of ChatGPT and the information searchers by social science students, to see whether the correlation is significant or not. Based on Gogtay and Thatte (2017), the correlation coefficient can be recognized using the value presented in Figure 3.3.

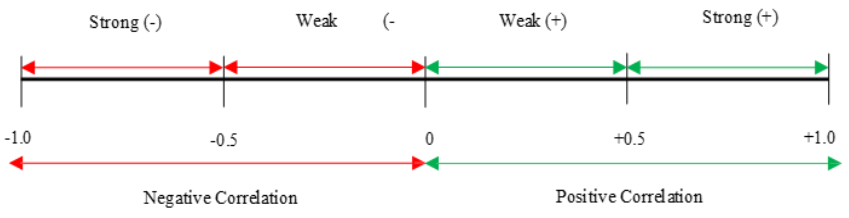

Figure 3.3: Basic spectrum of interpreting correlation coefficient
(Source: Gogtay and Thatte, 2017, p. 79)

Based on Figure 3.3 (Gogtay and Thatte, 2017), if the trend of a variable is positive and almost comparable to another variable, there may be a probability of a positive relationship between them, which can yield a positive correlation coefficient. Suppose one variable's trend is positive and practically negative to another. In that case, there is a probability of a negative relationship between the two variables, which might result in a negative correlation coefficient. The coefficient of correlation R ranges from -1 to +1, with -1 < R <+1. There is no method to interpret the correlation coefficient.

3.8.3 Linear Regression Analysis

Linear regression analysis is used to predict the value of a dependent variable based on the value of the independent variable. In this case, the researcher used multiple linear regression as it is relevant to this study. Multiple linear regression analysis estimates the association between two or more independent variables and one dependent variable. The researcher used multiple linear regression analysis to determine the strength of the relationship between the accuracy level of ChatGPT and information searchers by social science students. Furthermore, this analysis aims to examine the value of the dependent variable at a certain value of the independent variable. Moreover, the equation of multiple linear regression analysis is shown below:

Equation: $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_n X_n + €$

Where:

Table 3.5: Equation of Simple Linear Regression Analysis

Source: (San, 2021)

y	pendent variable (Information Searchers by Social Science Students)
B_0	**tercept**, the predicted value of **y** when all other parameters are set to 0)
$B_1 X_1$	**gression coefficient (B_1)** of the first independent variable has on the predicted **y** value (Validity)
B_2X_2	e second independent variable (Reliability)
B_3X_3	e third independent variable (Relevance)
$B_n X_n$	e regression coefficient of the last independent variable
e	e **error** of the estimate, or how much variation there is in the researcher's estimate of the regression coefficient

According to San (2021), Table 3.5 is the equation that shows how the dependent variable (information search by social science students) is influenced by the independent variables (validity, reliability, and relevance). Furthermore, each regression coefficient (β) measures the contribution of its respective independent variables to the prediction of *y*, while the intercept (β_0) provides the baseline prediction. Lastly, the error term (€) highlights how much variation there is in the researcher's estimate of the regression coefficient.

3.8.3.1 R Square

The coefficient of determination (R^2) measures a model's goodness of fit. In the regression context, it is a statistical measure of how well the regression line fits the real data. It is therefore significant when a statistical model is employed to forecast future outcomes or to evaluate hypotheses. Specifically, it estimates the amount of variance in the dependent variable that the model can explain. The one presented here is widely used

$$R^2 = 1 - \frac{sum\ squared\ regression\ (SSR)}{total\ sum\ of\ squares\ (SST)}$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

Figure 3.4: Equation of R-Square

Source: (Newcastle University, 2024)

According to Newcastle University (2024), figure 3.4 shows the equation of R-square. The sum squared regression is the sum of the residuals squared, and the total sum of squares is the sum of the data's distance from the mean, squared. The percentage number ranges from 0 to 1. Simply said, the greater a model's prediction accuracy, the closer its R-squared value is to one. If R-squared is zero, the model cannot predict the result any better than estimating the dependent variables' average value. However, R-squared values between 0 and 1 indicate that the model predicts the result partially. Its estimates are more accurate than the average value, but they are not perfect. If R-squared equals one, the model fully predicts the outcome.

3.8.3.2 F Value

The F value in regression is the result of a test in which the null hypothesis is that all regression coefficients are zero. In other words, the model lacks predictive capability. Furthermore, the F test determines the statistical significance of the regression equation as a whole. It is calculated by dividing the explained variance by the unexplained variance. An F-value of more than 4.0 is usually statistically significant but should check the table to be sure. If the F value is significant, the regression equation will help to comprehend the relationship between X and Y. Mathematically, the F-statistics are calculated as:

$$F = \frac{TSS - RSS}{P - 1} \div \frac{RSS}{n - p}$$

Where:

Table 3.6: Equation of F value

Source: (Prove that F-statistic follows F-distribution, 2017)

TSS	(Total Sum of Square) represents the total variability in the dependent variable.
RSS	(Residual Sum of Square) represents the unexplained variability after fitting the model.
(p)	The number of predictors (including the constant)
(n)	The number of observations.

Table 3.6 from Prove that F-statistic follows F-distribution (2017) shows the equation of F-value. From the formula, the numerator (TSS−RSS)/(p−1) calculates the average variability explained by the predictors. Furthermore, the dominator RSS/(n−p) calculates the average unexplained variability after fitting the model. By using this equation, researchers can test whether the predictors as a group significantly impact the dependent variable.

3.8.3.3 T-Value

T-values are used in linear regression to determine whether a given independent variable (or feature) is statistically significant in the model. A statistically considerable variable strongly correlates with the dependent variable and improves the model's accuracy. In other words, the t-value refers to how these hypothesis tests evaluate sample data using t-values. T-values are a type of test statistic. Hypothesis tests employ the test statistic derived from the sample to compare it to the null hypothesis. If the test statistics are sufficiently extreme, it suggests that the data are incompatible with the null hypothesis, allowing the researcher to reject it (Frost, 2019).

CHAPTER 4

RESULT AND DISCUSSION

4.1 Introduction

This chapter presents the results and discussion of the questionnaire from respondents who answered about the accuracy level of ChatGPT in information searches by social science students. SPSS software version 27 will be used to code and enter all results received from the questionnaire to perform data analysis. The analyses include pilot testing, reliability testing, mean, correlation, regression analysis, and descriptive statistics. In this research, a total of 327 UTeM students were invited to answer the survey questions. However, only 215 respondents completed and returned the survey questions through online the Google Survey Form. An initial pilot test was conducted with 22 respondents to ensure the reliability and validity of the survey instrument. The questionnaire was divided into three parts. Part A is for demographic information of the respondents, followed by part B which has independent variables related to the accuracy level of ChatGPT in information searchers by social science students, and part C, a dependent variable focusing on the perceived accuracy of ChatGPT. The purpose of this data analysis is to provide insights into the true accuracy level of ChatGPT from the perspective of social science students.

4.2 Results of Descriptive Analysis

In this research, the researcher used descriptive analysis to analyze the demographic background of 215 respondents. In this section, the background of respondents is analyzed, including gender, race, course, level of year, device frequently used to access ChatGPT, and frequency of use of ChatGPT.

4.2.1 Gender

Table 4.1: Gender of Respondents

(Sources: SPSS Output)

Gender

		Frequency	Percent (%)
Valid	Male	80	37.2
	Female	135	62.8
	Total	215	100.0

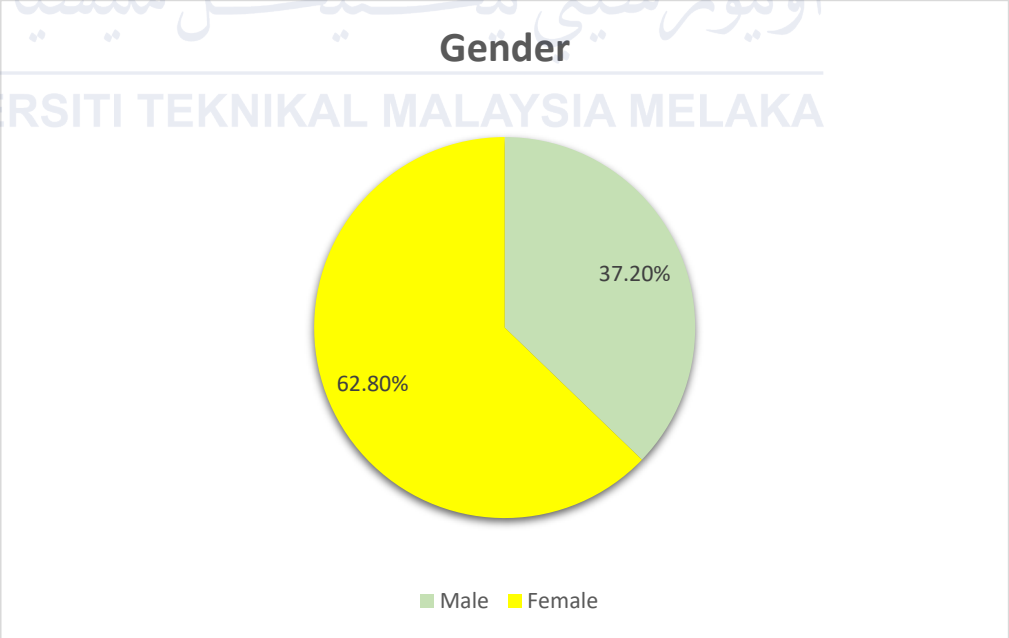

Figure 4.1: Gender of Respondents

Table 4.1 and Figure 4.1 present the frequency and percentage distribution of the respondent's gender. Among the 215 respondents who answered the questionnaire, 62.80%, or 135 female respondents were willing to answer the questionnaires. In contrast, only 80 male students, accounting for 37.20% answered the survey question. These results indicate that female students were more likely to participate in answering the questionnaires compared to male students.

4.2.2 Race

Table 4.2: Race of Respondents
(Source: SPSS Output)

Race

		Frequency	Percent (%)
Valid	Malay	124	57.7
	Chinese	58	27.0
	Indian	28	13.0
	Kayan	1	0.5
	Dusun	1	0.5
	Bugis	1	0.5
	Siamese	1	0.5
	Bumiputera Sabah	1	0.5
	Total	215	100.0

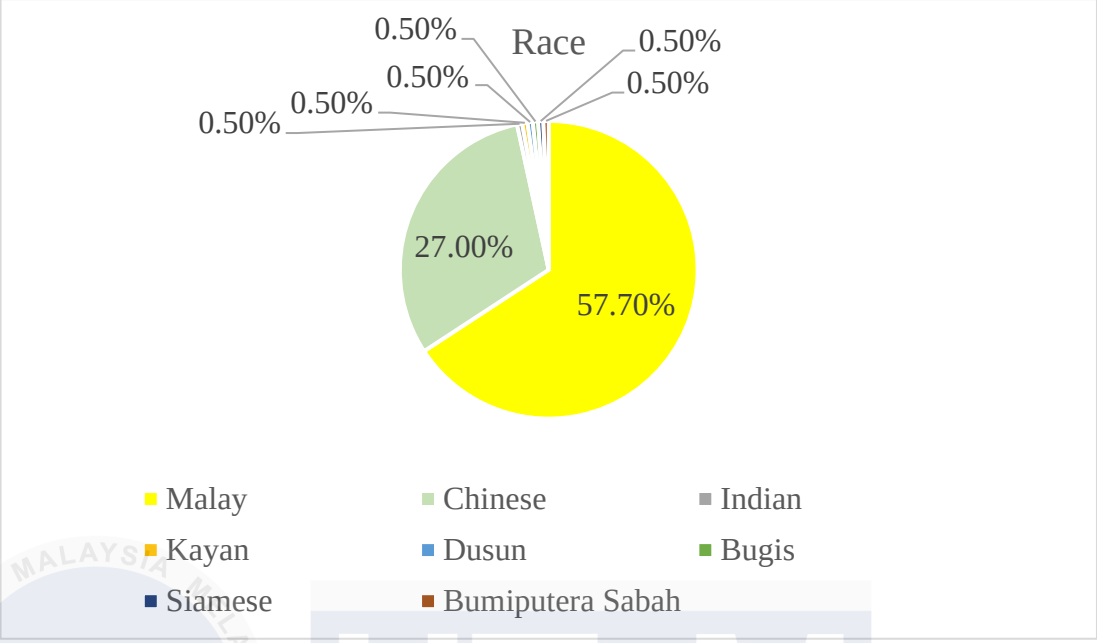

Figure 4.2: Race of Respondent

Table 4.2 and Figure 4.2 illustrate the racial composition of 215 respondents from FPTT, UTeM. It consists of 8 different races. The data reveal that most respondents are predominantly with about 57.70% or 124 respondents. In comparison, Chinese respondents managed to get about 27% (58 respondents) which shows them as the second largest group. Additionally, 28 respondents (13%) are Indian students. The remaining respondents, consisting of Kayan, Dusun, Bugis, Siamese, and Bumiputera Sabah, each 0.5% (1 respondent per group). In summary, most respondents came from the Malay race, reflecting their dominant representation in this survey.

4.2.3 Course

Table 4.3: Course of Respondents

(Source: SPSS Output)

Course

		Frequency	Percentage (%)
Valid	BTMI	68	31.6
	BTMM	69	32.1
	BTMS	44	20.5
	BTEC	34	15.8
	Total	215	100.0

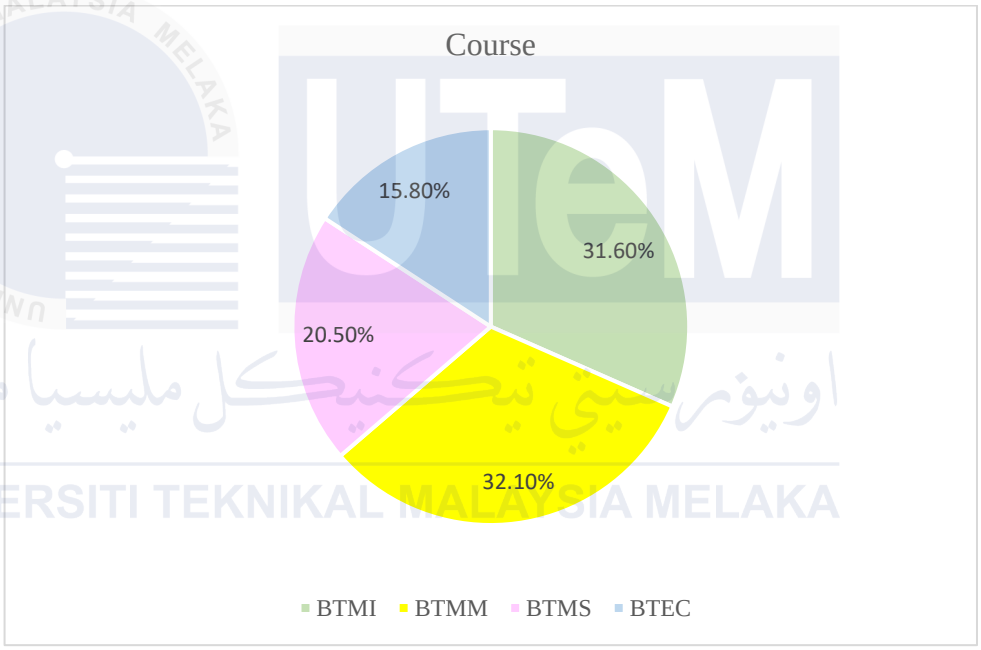

Figure 4.3: Course of Respondents

The course of respondents is shown above in Table 4.3 and Figure 4.3. The majority of respondents are from the Bachelor of Technology Management (High Technology Marketing) (BTMM) program, comprising 69 respondents (32.10%). Following closely, 68 respondents (31.60%) are enrolled in the Bachelor of Technology Management (Innovation Technology) (BTMI) course. Meanwhile, the Bachelor of Technology Management (Supply Chain Management and Logistics) (BTMS) represented accounts for 44 respondents (20.50%). Lastly, the Bachelor of Technopreneurship with Honors (BTEC) program (MQA/FA3623) contributes 34 respondents (15.80%). In summary,

the BTMM course has the highest representation among respondents, followed closely by BTMI, while BTEC has the lowest representation.

4.2.4 Level of Year

Table 4.4: Level of Year of Respondents

(Source: SPSS Output)

Level of Year

		Frequency	Percentage (%)
Valid	1	24	11.2
	2	35	16.3
	3	27	12.6
	4	129	60.0
	Total	215	100.0

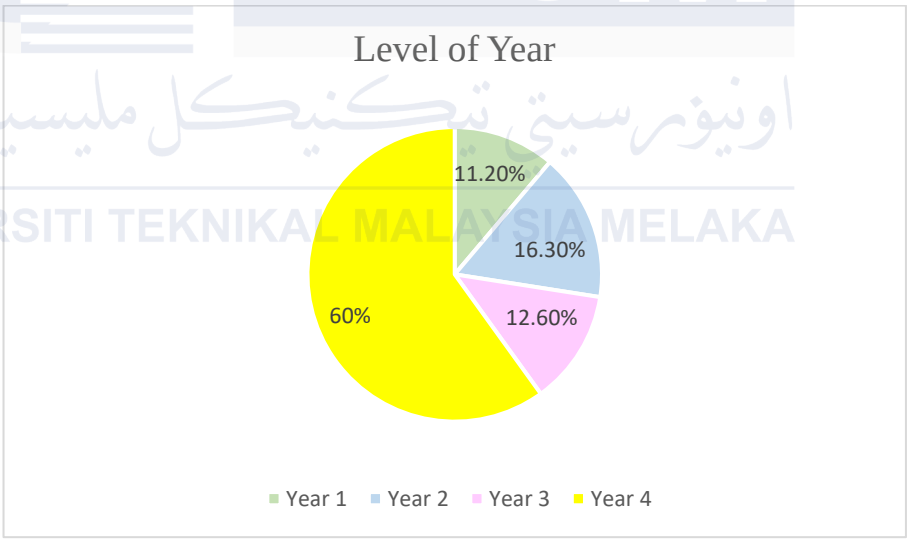

Figure 4.4: Level of Year of Respondent

Table 4.4 and Figure 4.4 present the distribution of respondents based on their academic year. There are 24 respondents (11.20%) from year 1 students. Otherwise, 35 out of 215 respondents (16.3%) from year 2 students, and 27 respondents (12.6%) are year 3 students. Most of the respondents which is 129 (60%) came from 4-year students. The results suggest that final-year students are more familiar with ChatGPT and frequently use AI tools as part of their study method. In contrast, first-year students, representing the lowest

percentage, may be less familiar with AI tools such as ChatGPT due to their limited exposure and experience.

4.2.5 Devices Frequently Use to Access ChatGPT

Table 4.5: Devices Frequently Use to Access ChatGPT of Respondents

(Source: SPSS Output)

Devices Frequently Use to Access ChatGPT

		Frequency	Percentage (%)
Valid	Telephone	30	14.0
	Telephone, Laptop	53	24.7
	Telephone, Laptop, Tablet	50	23.3
	Telephone, Tablet	19	8.8
	Laptop	39	18.1
	Laptop, Tablet	12	5.6
	Tablet	11	5.1
	PC	1	0.5
	TOTAL	215	100.0

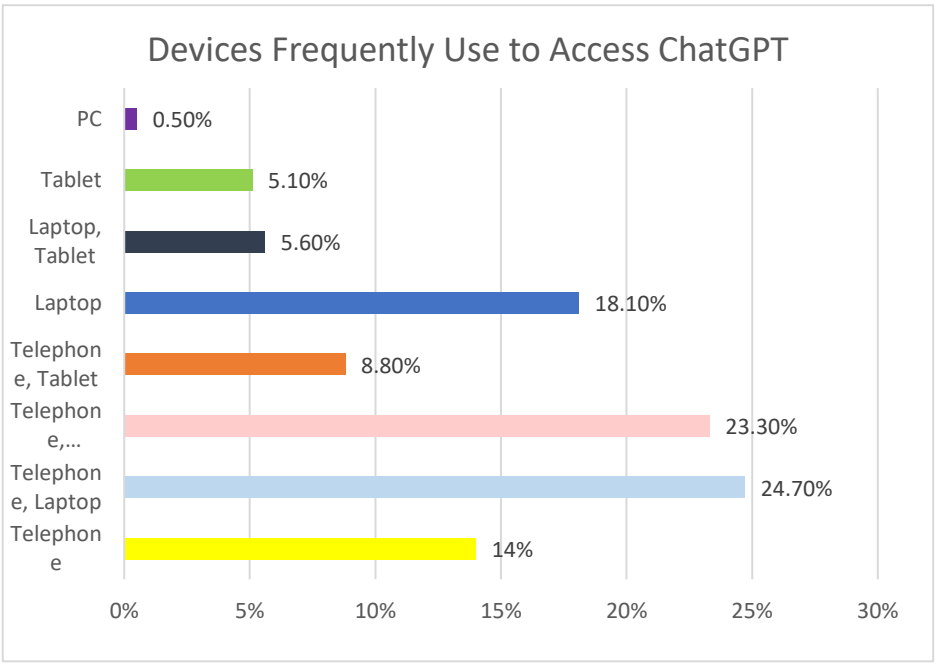

Figure 4.5: Devices Frequently Use to Access ChatGPT

Table 4.5 and Figure 4.5 show the devices frequently used by 215 respondents to access ChatGPT. From the results, the researcher identified that 53 respondents (24.7%) primarily use telephones and laptops to access ChatGPT. Then, it was followed by 50 respondents with a percentage of 23.3% who use their telephone, laptop, and tablet. Moreover, 18.1% of the 39 respondents prefer using their laptop to access ChatGPT. 30 respondents with a percentage of 14.0% reported using a telephone to log in to ChatGPT. Furthermore, respondents preferred to use telephone and tablet which are 8.8% to use ChatGPT while the respondents who used laptop and tablet to access ChatGPT are 5.6%. That is 5.1% who choose to use tablets and only one respondent uses the PC to access ChatGPT. In a nutshell, telephones and laptops are the most preferred devices for accessing ChatGPT, likely due to their convenience and versatility compared to other devices.

4.2.6 Frequently Use of ChatGPT

Table 4.6: Frequently Use of ChatGPT Respondents

(Source: SPSS Output)

Frequently Use of ChatGPT

		Frequency	Percentage (%)
Valid	Daily	41	19.1
	Frequently (4 or more times a week)	73	34.0
	Sometimes (2-3 times a week)	65	30.2
	Occasionally (1-2 times a month)	27	12.6
	Rarely (Less than once a month)	9	4.2
	Total	215	100.0

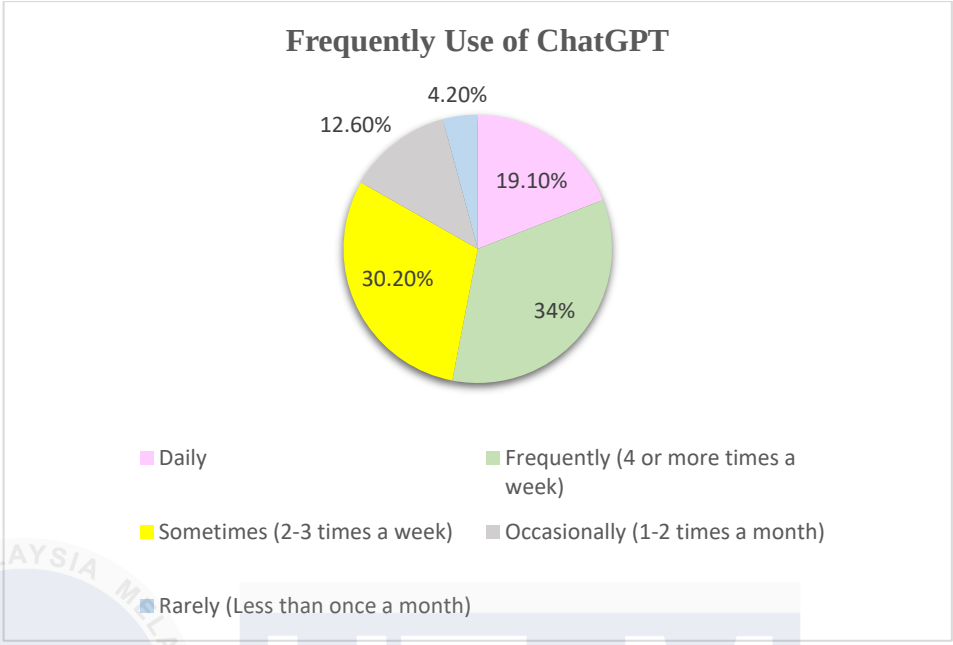

Figure 4.6: Frequently Use of ChatGPT

Table 4.6 and Figure 4.6 illustrate the frequency of ChatGPT usage among the 215 respondents. The data show that 41 respondents (19.1%) use ChatGPT daily. Furthermore, a total of 73 respondents (34%) reported using ChatGPT very often, 4 or more times a week. Next, 65 respondents (30.2%) use ChatGPT sometimes, about 2 to 3 times a week. Among the remaining respondents, 27 (12.6%) use ChatGPT occasionally, about 1-2 times a month, and 9 respondents (4.2%) reported rarely using it, less than once a month. In summary, most respondents use ChatGPT with varying frequency, with daily and frequent use relatively common among students.

4.3 Result of Data Analysis

4.3.1 Normality Test: Skewness and Kurtosis

Skewness measures a distribution's asymmetry. This value may be positive or negative. If the skewness is negative, the tail is on the left side of the distribution and extends to greater negative values. However, positive skew implies that the tail is on the right side of the distribution, leading to higher

positive values. If a value of zero shows that the distribution has no skewness at all, it is symmetrical (Zach, 2022).

According to Zach. (2022, January 6), kurtosis is a measure of the distribution of its strong or weak tails compared to a normal distribution. It states that normal distributions have a kurtosis of 3. If a distribution has a kurtosis less than 3, it is considered playkurtic, which means it produces fewer and less extreme outliers than the normal distribution. However, if a distribution's kurtosis is more than 3, it is considered leptokurtic, which means it produces more outliers than the normal distribution.

In this research, the researcher decides to use -1 and +1 for the skewness value as it shows excellent value (Hair et al., 2022, p. 66). For the kurtosis, the researcher will use -3 and +3 values because they show the normal distribution (Hoskins, Jessica, 2018).

4.3.1.1 Descriptive Statistics

Table 4.7 demonstrates the Skewness and Kurtosis of each item for testing normality distribution.

Table 4.7: Measures of Skewness and Kurtosis

(Source: SPSS Output)

Items	Skewness		Kurtosis	
	Statistic	Std. Error	Statistic	Std. Error
1.1 Data produced can be inaccurate	-.647	.166	-.270	.330
1.2 Limitations in Understanding Complex Contexts	-1.116	.166	.823	.330
1.3 Responses focused on smoothness can mistakenly cause wrong information.	1.161	.166	.652	.330
1.4 Lack of source attribution (acknowledgment)from ChatGPT can lead to suspicions of plagiarism.	1.046	.166	1.529	.330

2.1ChatGPT's capabilities are based on its "training data."	.981	.166	1.893	.330
2.2 User interactions with ChatGPT influence the reliability of the information.	.937	.166	.403	.330
2.3 Outdated data delays ChatGPT's ability to understand newer information.	.838	.166	.756	.330
3.1 Keywords are important in guiding ChatGPT in generating accurate responses.	.854	.166	.408	.330
3.2 Keywords are essential in helping ChatGPT to understand the context.	.994	.166	3.030	.330
3.3 Content quality impacts the relevance of the information provided by ChatGPT.	1.078	.166	1.908	.330
3.4 Clearer questions help ChatGPT to generate related responses.	.833	.166	.386	.330
3.5 A poorly organized link structure can cause ChatGPT to access information from irrelevant sources.	1.041	.166	1.664	.330
DV1. Trust in ChatGPT influences the frequency of uses.	.975	.166	2.469	.330
DV2. Social media is the main way you learn about ChatGPT.	1.107	.166	.005	.330
DV3. Students mainly use ChatGPT for assignment purposes.	.920	.166	.739	.330
DV4. Information from ChatGPT is more reliable than other AI tools.	1.172	.166	.097	.330
DV5. Self-assessing (re-checking) ChatGPT's answers can help students avoid untruth.	.907	.166	.220	.330

In Table 4.7, independent variables 1.1 (Data produced) and 1.2 (limitations in Understanding Complex Contexts) have a skewness of -.65 and -1.12, respectively, indicating that the distribution is left-skewed. This means that lower responses are more frequent for these variables. However, the skewness values, ranging between -0.5 and -1, indicate that the data are still close to being symmetrical.

In contrast, most items show a positive skewness, also known as the right-skewed, with values generally ranging between 1 and 0.5. This indicates that the data are slightly skewed, with responses concentrated on the lower end of the scale and some higher values stretching the tail to the right.

For the kurtosis, all items have values less than 3, except for item 3.2 (keywords essential for understanding context), which has a kurtosis of 3.030. A kurtosis less than 3 indicates flatter distributions with lighter tails, meaning fewer extreme values. However, item 3.2, with its kurtosis slightly exceeding 3.030, indicates a sharper peak and heavier tails, suggesting a higher concentration of values around the mean with a greater presence of extreme responses.

4.3.2 Reliability Test

4.3.2.1 Pilot Test Result

A pilot study is a small feasibility study designed to test various aspects of the methods planned for a larger, more rigorous, or confirmatory investigation (Arain et al., 2010). The major objective of the pilot study is not to answer specific research questions but to prevent the researcher from undertaking a larger-scale study with a lack of knowledge of the proposed methods. Essentially, the pilot study is done to avoid the fatal flaw occurring in a study that is expensive in time and money (Polit & Beck, 2017). In this study, the researcher chose 22 participants to answer the questionnaires in the pilot test. The internal validity of the questionnaires was evaluated using a reliability test, and the Cronbach Alpha value was computed. According to Ashley (2017), a value above 0.70 is acceptable, and 0.80 or greater is preferred.

Table 4.8: Case Processing Summary of Pilot Test

(Source: SPSS Output)

		N	PERCENT (%)
Cases	Valid	22	100.0
	Excluded*	0	.0
	Total	22	100.0

Table 4.9: Pilot Test Reliability Analysis

(Source: SPSS Output)

Variables	Cronbach Alpha	No of Items
Validity	.824	3
Reliability	.774	3
Relevance	.898	5
Information search by social science students	.754	5

Table 4.9 shows that Cronbach Alpha for pilot test results is reliable. According to Cronbach's Alpha interpretation in Table 3.4, page 38 in Chapter 3, the independent variables, validity, and relevance, were found to be good reliability (3 items; $\alpha = .824$, 5 items; $\alpha = .898$). Moreover, Cronbach's Alpha values of the 3 items of reliability variable and 5 items of information search by social science students were considered acceptable, where both have .774 and .754 respectively. It indicates that the independent and dependent variables can be used in the actual survey questionnaire once the reliability is valid.

4.3.3 Correlation Test

Correlation measures the degree of relationship between two variables under consideration. The correlation coefficient quantifies this relationship, providing insight into the strength and direction of the association between the variables. Two correlation coefficients that are frequently used in practice are Pearson's Product Moment Coefficient and Spearman's Rank Correlation Coefficient (Senthilnathan,2019).

For this study, the researcher employs Pearson's Simple Linear Correlation to examine the relationship between variables. Table 5 below presents the result of the correlation test between the variables.

Table 4.10: Correlation Result

(Source: SPSS Output)

Correlations					
		Mean: Validity	Mean: Reliability	Mean: Relevance	Mean: Information Search by Social Science Students
Mean: Validity	Pearson Correlation	1	.710**	.729**	.679**
	Sig. (2-tailed)		<.001	<.001	<.001
	N	215	215	215	215
Mean: Reliability	Pearson Correlation	.710**	1	.727**	.780**
	Sig. (2-tailed)	<.001		<.001	<.001
	N	215	215	215	215
Mean: Relevance	Pearson Correlation	.729**	.727**	1	.722**
	Sig. (2-tailed)	<.001	<.001		<.001
	N	215	215	215	215
Mean: Information Search by Social Science Students	Pearson Correlation	.679**	.780**	.722**	1
	Sig. (2-tailed)	<.001	<.001	<.001	
	N	215	215	215	215
****. Correlation is significant at the 0.01 level (2-tailed)**					

Table 4.10 above summarizes the correlation analysis results, which evaluate relationships between the independent variables (validity, reliability, and relevance). These relationships were analyzed using Pearson's Product-Moment Correlation.

Based on the table, a Pearson product-moment correlation was run to determine the relationship between validity and reliability. There was a strong, positive correlation between validity and reliability, which was statistically significant (*r = .710, n = 215, p = <.001*).

The relationship between validity and relevance variables also showed a high positive correlation where $r = .729, n = 215, p = < .001$. A strong

positive correlation was observed between reliability and relevance where ($r = .727, n = 215, p = < .001$).

In conclusion, these results show that the independent variables are in strong, positive relationships. These findings show that relationships are significant at a 0.01 level two-tailed, thus showing ChatGPT's strong accuracy in facilitating the information search for such a group.

4.3.4 Regression Test

4.3.4.1 R-squared

In this research, multiple regression was used to test if the accuracy level of ChatGPT significantly predicted information search by social science students in UTeM. Table 4.11 shows the results of the R-squared.

Table 4.11: Multiple Regression Analysis (R-Squared)

(Source: SPSS Output)

Model Summary				
Model	**R**	**R Square**	**Adjusted R Square**	**Std.Error of the Estimate**
1	.817[a]	.667	.662	.47610
a. Predictors: (Constant), MEAN_IV3, MEAN_IV2, MEAN_IV1				

Table 4.11 shows that the correlation coefficient (R) is 0.817, indicating a very strong relationship between the independent variables, that is, validity, reliability, and relevance to the dependent variable of information search by social science students. The R^2 value of 0.667 infers that 66.7% of the variation in information search by social science students is explained by the model. The adjusted R^2 accounted for some predictors involved in the model, hence the value was 0.662, showing that 66.2% of the variance is explained with minimal loss of explanatory power. In addition, the standard error of the estimate was 0.47610, representing the average distance that observed values fall from the predicted regression line, and suggesting that the model is reasonably good at predicting information search behavior. Overall, the findings highlight that

validity, reliability, and relevance are the major factors influencing the accuracy level of ChatGPT in supporting information search by social science students. These all together explain a substantial portion of the variation in the dependent variable.

4.3.4.2 F-Value

In this section, the F-value in the ANOVA table to tests whether the overall regression model is a good fit for the data. Table 4.12 shows the results of ANOVA.

Table 4.12: Multiple Regression Analysis (F-Value)

(Source: SPSS Output)

ANOVA[a]						
Model		Sum of square	df	Mean Square	F	Sig.
1	Regression	95.718	3	31.906	140.762	<.001[b]
	Residual	47.827	211	.227		
	Total	143.545	214			
a. Dependent Variable: MEAN_DV						
b. Predictors: (Constant), MEAN_IV3, MEAN_IV2, MEAN_IV1						

The table presents the result of the ANOVA test, showing that the F-test value for the regression model is 140.762 with a significant level of p = <.001. This high F-value suggests that the overall regression is a good fit for the data and that there is a significant relationship between independent variables like validity, reliability, and relevance and dependent variable, information search by social science students. The significant level is less than 0.05 confirming that the observed relationship is not due to random chance.

The result indicates that independent variables substantially impact the accuracy level of ChatGPT in information search by social science students. The null hypothesis is rejected since it shows an effect, meaning there is no relationship between independent and dependent variables. This supports the conclusion that validity, reliability, and relevance are important predictors of the accuracy level of ChatGPT in information search by social science students.

4.3.4.3 T- Value

Table 4.13: Multiple Regression Analysis (Coefficients)(T-value)

Source: SPSS Output

Coefficients[a]					
Model		Unstandardized Coefficients		**t**	**Sig**
		B	Std.Error		
1	(Constant)	.558	.172	3.254	.001
	Validity	.120	.054	2.214	.028
	Reliability	.493	.063	7.790	<.001
	Relevance	.253	.061	4.123	<.001
a. Dependent Variable: Information Search by Social Science Students					

Table 4.13 above indicates the result of the Coefficient for multiple regression analysis. The unstandardized coefficient (β) contrasts the strength of the effect of each independent variable (IV) on the dependent variable (DV).

The beta value of validity was 0.120 with a significant value of 0.028. Since this is less than 0.05, the relationship between validity and information search by social science students is statistically significant. The corresponding t-value of 2.214, although relatively low but still significant as its p-value is less than 0.05.

Furthermore, for reliability, the beta value was 0.493, with a highly significant p-value of <.001 and a t-value of 7.790. This indicates a very strong and statistically significant relationship between reliability and the dependent variable.

Lastly, the beta value for relevance is 0.253 with a significant value of <.001 and a t-value is 4.123, showing a significant relationship between relevance and the dependent variable.

From the result, reliability has the highest beta value (0.493), showing that reliability is the most influential factor affecting the accuracy level of ChatGPT in information searches by social science students.

Based on Table 4.13, the linear equation was developed as below:

$$\mathbf{Y = .558 + .120X_1 + .493X_2 + .253X_3}$$

Where:

Y = Information search by social science students

X_1 = Validity

X_2 = Reliability

X_3 = Relevance

Based on the linear equation above, there was a strong relation between the validity, reliability, and relevance of the accuracy level of ChatGPT in information search by social science students.

4.4 Results Discussion

4.4.1 Descriptive Analysis

In the descriptive analysis of the demographic respondents, the researcher found that females (62.8%) are more interested in answering survey questions related to this topic. According to findings by Kimbrough et al. (2013), females are more inclined towards mediated types of communication, such as text messaging and social networking, and they also use these means to interact with ChatGPT. Next, in the Malay race, as many as 124 people participated to become respondents than other races. Furthermore, students from the Bachelor of Technology Management (High Technology Marketing) (BTMM) program, achieved the highest participation comprising 69 respondents (32.10%) as they more frequently use ChatGPT as their resources. These capabilities help marketers use various new capabilities and innovative approaches for ideating and executing marketing tasks, including personalization, content creation, and market segmentation (Venture et al., 2024). Respondents from year 4 have 129 participants (60%). This is because they have more knowledge of using ChatGPT as their reference. Moreover, the researcher finds that all the respondents often use their telephones and laptops to interact with ChatGPT because it is more convenient and accessible. Based on Hornby (2024) perspective, while phones are portable and always at our

fingertips, laptops provide the user with a wider screen and keyboard. Laptops are way easier to work with when extended for a longer period or when performing complex tasks. Other than that, it is also because phones are a more suitable and quick way to use ChatGPT anywhere while laptops are preferred for more detailed work, like writing and research. Lastly, the researcher identified that respondents vote for frequent use of ChatGPT as they always use it to find resources for their assignments.

4.4.2 Reliability Analysis

Table 4.14 below shows the case processing summary. In this study, the researcher tested reliability using 215 samples to determine if the survey questions scale was reliable (Laerd Statistics, 2018). Table 4.15 shows the results of reliability statistics.

Table 4.14: Case Processing Summary

(Source: SPSS Output)

Case Processing Summary			
		N	PERCENT (%)
Cases	Valid	22	100.0
	Excluded*	0	.0
	Total	22	100.0
a. Listwise deletion based on all variables in the procedure			

Table 4.15: Reliability Analysis

(Source: SPSS Output)

Variables	Cronbach Alpha	No of Items
Validity	.902	3
Reliability	.854	3
Relevance	.935	5
Information search by social science students	.861	5

According to Table 4.15 above, Cronbach's Alpha for the 3 items of validity and 3 reliability items were .902 and .854 respectively. Both independent variables are acceptable for reliability. The relevance variable consisted of 5 items (α = .935), suggesting that the items are highly reliable and consistently measure the accuracy level of ChatGPT. The information search by social science students consisted of 5 items (α = .861) showing that is a very good scale for reliability.

4.4.3 Correlation Analysis

The correlation analysis revealed strong, positive, and statistically significant relationships between the independent variables which are validity, reliability, and relevance. According to the correlation tests above in Table 4.10, validity strongly correlated positively with reliability, ($r = .710,\ p = <$ $.001$) and relevance, ($r = .729,\ p = < .001$). In turn, reliability was significantly related to relevance, $r = .727,\ p = < .001$. Overall, the findings verify the ideas of validity and reliability along with relevance in supporting information search by social science students, since all relations are statistically significant at the level of ($p < .001$), which may signal that the analysis is robust.

4.4.4 Regression Analysis

The multiple regression analysis assessed the influence of the independent variables: validity, reliability, and relevance on the dependent variable, information search by social science students.

4.4.4.1 R-Squared

The regression model showed a very strong relationship (R=0.817) between the independent and dependent variables. According to the model, the

R^2 of 0.667 explained 66.7% variation in information search by social science students' behavior. The adjusted R^2 was 0.662 when considering the predictors included a minimal loss to the explanation of the phenomena, while the standard error of the estimate (SE = 0.47610) is low enough considering that the model is rather good in predictability about the information search behavior. To summarize, these findings confirm the hypothesis that validity and reliability are influential indicators of the accuracy level of ChatGPT in information searches by social science students.

4.4.4.2 F-Value

The results from the ANOVA test ($F = 140.762$, $p < 0.001$) confirmed that the regression model was a good fit for the data. The high value and significance of the F-value proved the relationship between the independent variables, validity, reliability, and relevance, and the dependent variable, information search by social science students to be statistically significant. This shows that the null hypothesis will be rejected and highlights the importance of the independent variables for assessing the dependent variable.

4.4.4.3 T-Value

From Table 4.13 above, the coefficients revealed the individual contributions of the independent variables where reliability had the highest beta value (0.493), indicating it is the most influential factor in analyzing the accuracy level of ChatGPT in information search by social science students.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter summarizes the study's results. The study concludes with recommendations for further research on the accuracy level of ChatGPT in information searches by social science students. It is also to explore more about the most influencing factors that might be changed due to demographics, society, and environment. Additionally, a review of the literature shows the evidence that the research conducted to identify the validity, reliability, and relevance of ChatGPT in information search by social science students are proved into this research. This study was administered to 215 respondents. The main targets of respondents are UTeM's social science students. In this chapter, this research will try to address the research question and discuss the achievement of the objectives. The researcher also will conclude from the results of Chapter 4. Lastly, the limitations of the study and recommendations for future research will be presented.

5.2 Research Objectives Achievement

5.2.1 Research Objective 1

RO1: To determine the accuracy level of ChatGPT in information search by social science students.

Based on the result in Chapter 4, the first research objective has been achieved. This study identified three accuracy levels of ChatGPT in information search by social science students which are validity, reliability, and relevance. These three accuracy levels are independent variables in this research. These accuracy levels were tested for validity through pilot tests and reliability analysis. The researcher selected 22 respondents from UTeM's social science students and asked them to participate in a pilot test to try the question. Afterward, the questionnaires were tested for their reliability in determining the internal validity, and Cronbach's Alpha value was determined.

According to reliability statistics in Table 4.15. Cronbach's Alpha shows excellent internal consistency, with a value of 0.902 for validity, 0.854 for reliability, and 0.935 for relevance. Based on Howard (2021), the value of Cronbach's Alpha is acceptable when it falls in the region of 0.7 and above. In a nutshell, the survey questions regarding the three independent variables are proven reliable and can be included in the survey.

Moreover, to assess the explanatory power of the independent variables, the researcher chose to use a multiple regression analysis model of R^2 to determine the proportion of variance in the dependent variable that the independent variables can explain. In other words, it shows how well the data fits the regression model (Taylor, 2024). The analysis revealed an R^2 value of 0.667, indicating that 66.7% of the variance in ChatGPT's accuracy levels can be explained by the three predictors. The model fit was strong, supported F (140.762), and a significant level of $P < .001$. These findings suggest a robust relationship between the predictors and the accuracy level of ChatGPT in information search by social science students.

5.2.2 Research Objectives 2

RO 2: To evaluate the impact of the accuracy level of ChatGPT in information search by social science students.

The second objective focuses on the correlation between the accuracy levels of ChatGPT in information search by social science students. This study has three independent variables: validity, reliability, and relevance. In addition to a strong correlation, the pattern of strong relationships was positive. Based on Table 4.10, the validity and reliability were significantly correlated, ($r = .710,\ p = < .001$) while validity and relevance ($r = .729,\ p = < .001$), also were significantly correlated. Furthermore, the result shows reliability and relevance have a correlated relationship ($r = .727,\ p = < .001$). These significant correlations demonstrate that the accuracy levels are not isolated factors but are closely interrelated and influence each other in information quality.

Logically, the nature of information processing calls for such a relation between these independent variables, validity, reliability, and relevance. Valid information will tend also to be reliable since it withstands truth and accurate criteria for some time. Equally, relevance tends to go along with validity and reliability as such information falls within the closer scope of user's needs and provides insights that apply to their work. As Mun (2024) has stated, these three aspects form the pillars of information quality. Working together, they guarantee information precision, consistency, and applicability a must in raising the level of the user's trust and confidence in the tool.

The impact of the accuracy level is important especially for social science students since they tend to use a lot of relevant and credible information in completing their research. At this level of academic training, accessing valid, reliable, and relevant data means their findings are more likely to be accurate, defensible, and aligned with academic standards (Strzelecki, 2024). The ability of ChatGPT to meet these criteria not only helps effectively search for information but also enhances critical thinking, as students can evaluate the quality of their findings with greater confidence. The findings show this objective can be achieved.

5.2.3 Research Objective 3

RO 3: To analyze which measure of accuracy level of ChatGPT has the most significant influence on information search by social science students.

Table 5.1: Coefficients

(Source: SPSS Output)

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	**Sig**
		B	Std.Error	Beta		
1	(Constant)	.558	.172		3.254	.001
	Validity	.120	.054	.139	2.214	.028
	Reliability	.493	.063	.488	7.790	<.001
	Relevance	.253	.061	.266	4.123	<.001
a. Dependent Variable: Information Search by Social Science Students						

The result shows the coefficient between the independent variables and the dependent variable. As can be seen, all three variables have a positive significant relationship as the p-value is less than 0.05, which are validity, reliability, and relevance.

Among these, reliability emerged to be the most significant predictor of effective information search, as evidenced by its highest unstandardized beta value ($\beta = 0.493$) and a highly significant p-value ($p < 0.001$). The result suggests that students searching for information through ChatGPT consider consistency and dependability. It is the nature of the task that reliability determination increases the likelihood that the output generated by ChatGPT will be valid, consistent, and repeatable results, reducing uncertainties in the research process.

Moreover, relevance with an unstandardized beta value of $\beta = 0.253$ and a significant p-value ($p < 0.001$), also has a strong impact on information search that supports their academic goals. This alignment with the user's needs and context not only ensures accuracy but also applicability, thereby reducing the cognitive load associated with filtering irrelevant content (Jela Steinerova, 2024).

Validity, although having a smaller unstandardized beta value (β = 0.120) and a p-value of 0.028, still demonstrates a significant positive influence. It shows that accuracy and truthfulness are important, they may take secondary importance to reliability and relevance in the context of information search (Hassan, 2023). Therefore, validity ensures that the information retrieved is reliable and correct, serving as a solid foundation for academic and research tasks.

In conclusion, the data above prove that reliability emerges as the most significant influence on information search by social science students. Therefore, this objective has been achieved.

5.3 Research Hypothesis Achievement

Table 5.2: Summary of Hypothesis Testing
(Source: SPSS Output)

Hypothesis	t	Sig	Results
H1: There is a positive significant relationship between validity and information search by social science students.	2.214	0.028 < 0.05	Accepted
H2: There is a positive significant relationship between reliability and information search by social science students.	7.790	<0.001<0.05	Accepted
H3: There is a positive significant relationship between relevance and information search by social science students.	4.123	<0.001<0.05	Accepted

Hypothesis testing is a systematic procedure for deciding whether the results of a research study support a particular theory that applies to a population (Turner, 2020). Table 5.2 above shows the results of hypothesis testing.

5.3.1 Independent variable 1: Validity

Hypothesis 1: There is a positive significant relationship between validity and information search by social science students.

Based on Table 5.2 above, the result showed that the validity affected information search by social science students in hypothesis 1. This is because the t-value was 2.214 and the p-value was 0.028, which is below the threshold of 0.05. This result demonstrates that validity has a statistically significant positive relationship with the information search by social science students.

According to Keshavarz et al., (2016), there is a direct correlation between the validity of information and students' efficient and effective search performance. If the scale of self-efficiency in information searching is high, students will perform a complete and effective search. High self-confidence in the students to get valid information improves the quality and reliability of the search results.

Furthermore, social science students are indeed being trained in informational literacy, which encompasses source evaluation. It not only aids students in finding credible sources but also filters through most irrelevant information on the web and retrieves relevant correct data. Miraj et al. (2021) emphasized that this kind of training in effective search allows individuals to critically evaluate the credibility of the information and avoid untrustworthy or misleading information.

In conclusion, the null hypothesis (H_0) was rejected as the p-value was less than the significance threshold of 0.05 and the alternative hypothesis (H_1) was accepted. There is a significant relationship between validity and information search by social science students. This relationship emphasizes the significance of validity as a critical aspect in the information-seeking process, especially for students who rely on correct and reputable information to achieve successful research outcomes.

5.3.2 Independent variable 2: Reliability

Hypothesis 2: There is a positive significant relationship between reliability and information search by social science students.

In hypothesis 2, the result indicates that reliability significantly influences the information search by social science students. This conclusion is supported by the statistical analysis, where the p-value was found less than 0.001, well below the significance threshold of 0.05. Moreover, the t-value is 7.790. This demonstrates a strong and positive relationship between reliability and information search by social science students.

According to the research (Stevenson University, 2023), choosing reliable sources comes down to good communication. If the user's knowledge is founded on inaccurate information, the user will not be a reliable asset to the company. This concept holds for students engaged in academic work and research, where their ability to source reliable information directly impacts depth of their findings' depth, accuracy, and reliability.

Social science students need to rely on evidence, as they often have to contend with emerging issues that affect society and require sane data for their examination and perhaps interpretation. Furthermore, reliable information helps eliminate the risk of making baseless conclusions due to inconsistencies or ineffective use of sources (Fraser, 2023). Additionally, relying on trustworthy information fosters critical thinking, as learners can confidently support their claims without doubting the reliability of their data.

The focus on reliability is crucial as it helps to improve the effectiveness of information search. Students who prioritize reliability can reduce search time and effort as they can locate information by discarding non-reliable resources. This aligns with good academic practice as information generated is prioritized for its quality, ensuring that research outputs are both meaningful and impactful.

In a nutshell, the findings validate the hypothesis that there is a positive significant relationship between reliability and information search by social science students. By identifying the important reliable information, this study emphasizes the need for students to develop critical evaluation skills and utilize tools like ChatGPT that favor consistent and reliable data. Finally, this proves

that H_2 is accepted and there is a positive significant relationship between reliability and information search by social science students.

5.3.3 Relevance

Hypothesis 3: There is a positive significant relationship between relevance and information search by social science students.

The findings from the analysis demonstrate that relevance plays a significant role in influencing social science students' information search. This conclusion is supported by the statistical result, where the p-value was found to be <0.001 which is below the threshold of 0.05, and the t-value (4.123). This indicates a strong and statistically significant positive relationship between the relevance and information search by social science students, thereby validating hypothesis 3.

Relevance is another important measure of information accuracy since it determines how well the information fits the specific needs, contexts, and goals of the user. Furthermore, relevance ensures that the information retrieved is not only accurate but also applicable to their academic work and research, literally for these social science students (Jordan & Tsai, 2024). According to the demographic section, the data states that most of the respondents are year 4 students. It shows that the relevance variable significantly enhances them to do their quality decision-making since they use these tools regularly to help them address their research questions or academic objectives. Irrelevant data will reduce clarity and distract them from focusing on their research, hence resulting in inefficiencies and potential errors in analysis. Prioritizing relevance helps students refine their search tactics, ensuring that the information they obtain is significant and practical.

To conclude, the statistical analysis supports the acceptance of hypothesis 3, confirming a positive significant relationship between relevance and information search by social science students.

5.4 Significant Contribution (Implication) of the Study

5.4.1 Body of Knowledge (Theory)

This research aims to study the accuracy level of ChatGPT in information search by social science students in the UTeM area. This study focused on social science students at the Faculty of Technology Management and Technopreneurship (FPTT). This research study offers essential insights into integrating AI tools like ChatGPT into educational environments, particularly for social science students. By analyzing the accuracy level of ChatGPT and its value in academic research, the findings indicate practical techniques for improving AI-based learning solutions suited to the specific needs of students in this field.

Therefore, this study investigates the impact of validity, reliability, and relevance in information searching over ChatGPT. This is because ChatGPT has been found to raise concerns about its potential for misinformation and unethical use by early adopters, especially students (CRLT, 2023). The research highlights the importance of addressing these concerns as nowadays ChatGPT is helpful for students in gathering information related to academics, particularly for literature reviews and research (Shubhrajyotsna Aithal & Aithal, 2023).

This emphasizes the necessity of providing students with the ability to identify and utilize the validity, reliability, and relevance of the information. The findings contribute to the theoretical knowledge of ChatGPT's role in academic environments and its implications for promoting ethical and successful use among students.

5.4.2 Industry (Implementation)

This research provides actionable insights for AI development and education industries, emphasizing the implications of ChatGPT's accuracy in academic contexts. By addressing the study's objectives, which are to determine ChatGPT's accuracy level, evaluate its impact, and analyze the most significant accuracy measure, the findings offer clear pathways for enhancing the practical application of ChatGPT in educational settings.

This underlines the need for the developer to enhance the performance of ChatGPT in terms of validity, reliability, and relevance. According to K.sabreena (2023); Article and Biswas (2023); and Christiano (2024), developers can use these insights to enhance this tool in producing valid and reliable responses for academic purposes. Moreover, advanced validation techniques should thus be initiated, reliable source attribution, and refinement of AI models in selecting contextually relevant outputs. These improvements will make ChatGPT a better tool for academic research, especially for social students.

Next, the accuracy level of ChatGPT has indeed brought out the critical impact on students' learning outcomes. These findings are recommended to be used by educational institutions to guide the integration of AI tools into teaching and research frameworks. Research indicates that integrating ChatGPT in educational settings can enhance engagement and learning outcomes for students due to the provision of individualized feedback (Bettayeb et al., 2024; Swargiary, 2024). For instance, developing digital literacy programs will mean an institution is fostering in students the critical thinking needed to evaluate the reliability and relevance of information produced through ChatGPT.

Through Industry-Academic Collaboration, further analysis of which accuracy measures yield the greatest influence gives good insight into potential collaboration among AI applications and educational stakeholders. This would most likely involve developing AI applications that meet social science research needs, filling in the gaps left so far uncovered by current applications, and based on academic standards.

In summary, by focusing on the three areas, this research provides practical recommendations for industries to enhance the usefulness and validity of ChatGPT and other AI tools in educational environments.

5.4.3 Nation

This study also has major consequences for national policies and laws, particularly in regulating AI technologies like ChatGPT and their usage in education. As ChatGPT becomes more widely used in universities and colleagues, the findings of this study provide important insights into how policies can be developed or refined to ensure the ethical and responsible use of AI in higher education.

The present study has established the need for a well-structured guideline and policy in the use of ChatGPT in academic learning. National policies, with concerns for misinformation, bias, and ethical use, should focus on developing frameworks that will guarantee the validity, reliability, and relevance of the content created by ChatGPT. Therefore, policymakers may work out the rules and regulations answering the challenges resulting from ChatGPT and align them with academic standards in advancing equity and integrity in higher education.

The findings of the study also highlight a need for legal frameworks that will protect academic integrity while allowing the beneficial use of AI. These laws do need further development given the challenges brought in by AI tools like ChatGPT, including plagiarism, intellectual property concerns, and the accuracy of ChatGPT-generated content (Mita, 2023). These subjective insights will enable national lawmakers to take protective legal measures to safeguard the students and the educational institution from possible misapplication while encouraging innovation and technological advancement in education.

In summary, by considering these insights, national policies and laws can be better positioned to balance AI's opportunities with the ethical and practical challenges it presents in educational environments.

5.5 Limitations of the study

To accomplish this research, there are some limitations in this research. Firstly, time constraints posed a significant challenge. According to Saunders et al. (2016), time constraint refers to the constraints about the start and end times of every task in a project's critical path, which includes activities that must be completed on schedule to avoid delaying the entire project. The researcher adopted a cross-sectional study design to address these constraints, focusing on data collection within a limited timeframe. The study period spanned only three months, from October 2024 to December 2024.

Due to the time constraints, questionnaires could only be shared via Google Forms. The questionnaires were shared on Telegram and WhatsApp as well, but it is just not convenient to get a quick response. To handle this solution, the researcher had to personally share the link to the survey among potential respondents due to the reason many individuals used to ignore online messages without responding to the survey.

Lastly, the honesty and attention of respondents presented another limitation. In the assessment of ChatGPT's accuracy level in information searches by social science students, some respondents appeared not interested in reading through the questions and responding. As a result, their responses might not be valid and honest, potentially impacting the study's reliability. To address the issue, the researcher designed a straightforward survey and organized instructions to guide respondents through each section.

5.6 Recommendation and Future Direction

Based on the study's findings, the following recommendations are given. The researcher should study the development of ChatGPT, focusing on the continuous upgrades that aim to make it more sophisticated and capable of understanding and responding to human language more naturally and nuancedly. Furthermore, the researcher needs to study in-depth how these technological advancements influence its performance in academic information searches. This includes studying natural language processing

(NLP) improvements and their specific impacts on the accuracy level of ChatGPT and the overall user experience for social science students.

It would be worth conducting focused research on how ChatGPT's upgrades affect the accuracy of its information provision. More specifically, future research should consider cases such as fact-based questions, interpretation of qualitative data, reference or citation generation, and proper review of the consistent patterns of inaccuracies or biases that will provide insights for both the developer and users.

Lastly, while the current study limits its participants to social sciences students, further research should also include students and researchers in other educational fields like natural sciences, engineering, humanities, and business studies. In this way, the comparison of the experiences and opinions of users in various academic disciplines will enable the researchers to identify discipline-specific strengths and limitations of ChatGPT and to what extent this AI technology meets the diverging needs of different fields.

5.7 Conclusion

In conclusion, ChatGPT is an innovative technology that has revolutionized how we interact with machines and each other. Its natural language processing capabilities enable it to generate human-like responses to user queries. However, ChatGPT has limitations regarding accuracy level. It is evidence that while ChatGPT is evolving rapidly, it still requires careful analysis and should not be utilized as an independent scholarly resource, as it lacks the critical reliability and depth needed for rigorous academic work. As technology advances, though improving each day, should find a place in education with the basic thought that it needs only to be employed as a supplement tool, and that one must cross-check information and maintain a critical outlook.

REFERENCES

Abdullayeva, M. & Musayeva, Z. M. (2023). The Impact of Chat Gpt on Student's Writing Skills: An Exploration of Ai-Assisted Writing Tools. International Conference of Education, Research and Innovation, 1(4), 61–66.

Abu-Taieh, E., Hadid, I. H. A., & Mouatasim, A. E. (2020). Cyberspace. In *Google Books*. BoD – Books on Demand. https://books.google.com/books?hl=en&lr=&id=eqf8DwAAQBAJ&oi=fnd&pg=PA27&dq=definition++%22research+design&ots=cLN15U9gOb&sig=fl41LsVzXbrX5WWzkky0tdEVe7M

Ahmad, S., Wasim, S., Irfan, S., Gogoi, S., Srivastava, A., & Farheen, Z. (2019). Qualitative v/s. quantitative research-a summarized review. *population*, *1*(2), 2828-2832

ai, D. (2024, March 14). How to Train ChatGPT On Your Own Data (7 Easy Steps). Denser.ai. https://denser.ai/blog/how-to-train-chatgpt-on-your-own-data/

Akhtar, I. (2016, September). *Research Design*. ResearchGate; www.researchgate.net.https://www.researchgate.net/publication/308915548_Research_Design

Alam, M. (2023, September 8). *What is Research Design? Definition, Types, Methods and Examples*. IdeaScale. https://ideascale.com/blog/what-is-research-design/

Alamleh, H., AlQahtani, A., & ElSaid, A. (2023). Distinguishing Human-Written and ChatGPT-Generated Text Using Machine Learning. 2023 Systems and Information Engineering Design Symposium (SIEDS), 154-158. https://doi.org/10.1109/SIEDS58326.2023.10137767.

Alvi, M. (2016). *A Manual for Selecting Sampling Techniques in Research*. Mpra.ub.uni-Muenchen.de. https://mpra.ub.uni-muenchen.de/70218/

Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Medical Research Methodology, 10(1). https://doi.org/10.1186/1471-2288-10-67

Article, R., & Biswas, S. (2023). Evaluating Errors and Improving Performance of ChatGPT. International Journal of Clinical and Medical Education Research, 2(6), 182–188. https://www.opastpublishers.com/open-access-articles/evaluating-errors-and-improving-performance-of-chatgpt.pdf

Ausat, A. M. A., Massang, B., Efendi, M., Nofirman, N., & Riady, Y. (2023). Can chat GPT replace the role of the teacher in the classroom: A fundamental analysis. Journal on Education, 5(4), 16100–16106

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. Available at SSRN 4337484

Bettayeb, A. M., Manar Abu Talib, Zahraa, A., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: conversational AI in education. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1379796

Bevans, R. (2023, June 22). *Multiple Linear Regression | A Quick Guide (Examples)*. Scribbr. Retrieved May 15, 2024, from https://www.scribbr. com/statistics/multiple-linear-regression/

Bhandari, P. (2021, July 15). *Designing a questionnaire*. Scribbr.https://www.scribbr.com/methodology/questionnaire/#:~:text=A%20questionnaire%20is%20a%20list

bomba, dr katarzyna. (2015, February 25). *Journal of Modern Science – Keyword appearance*. Www.jomswsge.com. https://www.jomswsge.com/Keyword-appearance/148626

Bustamante, J. (Jay). (2023, June 9). *Understanding User Interactions with ChatGPT: Lessons for Optimizing the UX of AI*. Medium. https://bootcamp .uxdesign.cc/understanding-user-interactions-with-chatgpt-lessons-for-optimizing-the-ux-of-ai-2542e9fa04cc

BYJU'S. (2023). *Sampling Methods (Techniques) - Types of Sampling Methods and Examples*. BYJUS. https://byjus.com/maths/sampling-methods/

Carlson, M. (2023, January 5). *How does ChatGPT determine whether content on a website is relevant for a specific query?* Www.linkedin.com.https://www. linkedin.com/pulse/how-does-chatgpt-determine-whether-content-website-relevant-carlson

ChatGPT, K. M. and. (2022, December 7). *How ChatGPT Could Transform Higher Education*. Social Science Space. https://www.socialsciencespace.com/ 2022/12/how-chatgpt-could-transform-higher-education/

community, B. (2023, July 19). *How Can I Improve ChatGPT's Accuracy or Relevance In Its Responses? | Botpress Blog*. Botpress.com. https://botpress. com/blog/how-can-i-improve-chatgpts-accuracy-or-relevance-in-its responses

CRLT. (2023). ChatGPT: Implications for Teaching and Student Learning | CRLT. https://crlt.umich.edu/blog/chatgptimplications-teaching-and-student-learning

Cuncic, A. (2022, October 17). *Understanding internal and external validity*. Verywell Mind. https://www.verywellmind.com/internal-and-external-validity-4584479

Darrel, M., Montefalcon, L., Padilla, J., Hernandez, A., Rhald, J., Padilla, C., & Montefalcon, L. (2023). *Information Seeking Behavior in ChatGPT: the Case of Programming Students using Structural Equation Modeling Information Seeking Behavior in ChatGPT: the Case of Programming Students using Structural Equation Modeling*.

Firaina, R., & Sulisworo, D. (2023). Exploring the Usage of ChatGPT in Higher Education: Frequency and Impact on Productivity. *Buletin Edukasi Indonesia*, *2*(01), 39–46. https://doi.org/10.56741/bei.v2i01.310

Formplus Blog. (2021, February 18). *Survey Methods: Definition, Types, and Examples*. Https. https://www.formpl.us/blog/survey-methods

Fraser, C. (2023, September 4). Strategies for Gathering Reliable Information. Pressbooks.pub; Loyalist College. https://ecampusontario.pressbooks.pub/academicwritingsuccess/chapter/strategies-for-gathering-reliable-information/

Frost, J. (2019, March 15). *How t-Tests Work: t-Values, t-Distributions, and Probabilities - Statistics By Jim*. Statistics by Jim. https://statisticsbyjim.com/hypothesis-testing/t-tests-t-values-t-distributions-probabilities/

Fujimoto S and Takemoto K (2023) Revisiting the political biases of ChatGPT. Front. Artif. Intell. 6:1232003. doi: 10.3389/frai.2023.1232003

Fulton, C., & McGuinness, C. (2016). *Information Searching - an overview | ScienceDirect Topics*. Www.sciencedirect.com. https://www.Sciencedirect .com/topics/social sciences/informationsearching#:~:text=Information%20 searching%2C%20also%20referred%20to

Gogtay, N. J. and Thatte, U. M. (2017), Principles of correlation Analysis, Journal of The Association of Physicians of India, 65 (March), pp. 78-81.

Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house." *Administrative Issues Journal Education Practice and Research*, *4*(2), 12–26.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (3 ed.). Thousand Oaks, CA: Sage.

Hammoud, M. (2023, February 16). *ChatGPT, history, practical use, biases, inaccuracies, and outdated information*. www.linkedin.com.https://www. linkedin.com/pulse/chatgpt-history-practical-use-biases-inaccuracies-outdated-hammoud?trk=article-ssr-frontend-pulse_more-articles_related-content-card

Hanum Siregar, F., Hasmayni, B., & Lubis, A. H. (2023). The Analysis of Chat

GPT Usage Impact on Learning Motivation among Scout Students. International Journal of Research and Review, 10(7), 632–638. https://doi.org/10.52403/ijrr.20230774

Harrison, J. (2022, December 9). LibGuides: Fake News: How Fake News Spreads. Libguides.uvic.ca. https://libguides.uvic.ca/fakenews/how-it-spreads

Hassan, M. (2022). *Quantitative Data - Types, Methods, Examples*. Research Method. https://researchmethod.net/quantitative-data/

Haven, J. (2023). ChatGPT and the future of trust. Nieman Lab. https://www.niemanlab.org/2022/12/chatgpt-and-the-future-of-trust/

Heever, M. V. D. (2022). *LibGuides: Research Support: Research Methodology*. Libguides.wits.ac.za. https://libguides.wits.ac.za/c.php?g=693518&p=4914913#:~:text=What%20is%20Research%20Methodology%3F

Hetler, A. (2023, December). *What is ChatGPT? Everything You Need to Know*. TechTarget. https://www.techtarget. com/whatis/definition /ChatGPT

Introducing ChatGPT: Benefits, Future, Importance, Limitations, Uses. (2023, March 30). Careerera.com. https://www.careerera. com/blog / introducing-chatgpt-benefits-future-importance-limitations-uses

JACOBY, D. J. (2023, March 2). *ChatGPT in its own words*. Www. linkedin.com.https://www.linkedin.com/pulse/chatgpt-its-own-words-dr-jack-jacoby

Jones, M. (2024, February 13). Mapping Decisions in ChatGPT from Prompt to Response. Www.linkedin.com.https://www.linkedin.com/ pulse/mapping-decisions-chatgpt-from-prompt-response-mark-jones-mbu8e

Jordan, K., & Tsai, S. P. (2024). Ranking “by Relevance” in Academic Literature Searches: Prevalence, Definitions, and Implications. Postdigital Science and Education. https://doi.org/10.1007/s42438-024-00530-z

Jowarder, M. I. (2023). The Influence of ChatGPT on Social Science Students: Insights Drawn from Undergraduate Students in the United States. *Indonesian Journal of Innovation and Applied Sciences (IJIAS)*, *3*(2), 194-200. https://doi.org/10.47540/ijias.v3i2.878\

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, *7*(4), 396–403. https://doi.org/10.9734/BJAST/2015/14975

Kadir, M. R. A., Johari, N. I. S., & Hussin, N. (2018). Information Needs and Information Seeking Behaviour: A Case Study on Students in Private University Library. International Journal of Academic Research in Progressive Education and Development, 7(3), 226–235.

Kakumbi, G., & Phiri, J. (2022). Adoption of Social Media for SME Growth in the Covid-19 Era: A Case of SMEs in the Clothing industry in Lusaka, Zambia. Open Journal of Business and Management, 10(06), 3202–3229. https://doi.org/10.4236/ojbm.2022.106159

Kalla, D., Smith, N., Kuraku, Dr. S., & Samaah, F. (2023). *Study and Analysis of Chat GPT and Its Impact on Different Fields of Study* , *8*(3). http://www.ijisrt.com/

Kanade, V. (2023, May 16). *ChatGPT Characteristics, Uses, and Alternatives | Spiceworks*. Spiceworks. https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-chatgpt/

Karunaratne, T., & Adesina, A. (2023). Is it the new Google: Impact of ChatGPT on Students' Information Search Habits. *European Conference on E-Learning*, *22*(1), 147–155. https://doi.org/10.34190/ecel.22.1.1831

Keshavarz, H., Givi, M., & Vafaeian, A. (2016). Students' sense of self-efficacy in searching information from the Web: A PLS approach. Webology, 13(2). https://webology.org/data-cms/articles/20200515045508pma149.pdf

Kimbrough, A. M., Guadagno, R. E., Muscanell, N. L., & Dill, J. (2013). Gender differences in mediated communication: Women connect more than do men. Computers in Human Behavior, 29(3), 896–900. https://doi.org/10.1016/j.chb.2012.12.005

Koblyakov, P. (2024, January 26). ChatGPT in Education: An In-depth Guide. Raccoon Gang. https://raccoongang.com/blog/chatgpt-in-the-educational- landscape-a-comprehensive-guide/

Koenders, dr L., & Konning, D. (2023). *ChatGPT in education: What is ChatGPT capable of, and what are its limitations? - Educational Development & Training - Utrecht University*. Www.uu.nl. Retrieved April 29, 2024, from https://www.uu.nl/en/education/educational-development-training/ knowledg e-dossier/what-is-chatgpt-capable-of-and-what-are-its-limitations#:~:text =Outdated% 20data&text= This% 20means%20tha

Kopp, C. M. (2022, October 2). *Why Brand Awareness Matters*. Investopedia. https://www.investopedia.com/terms/b/brandawareness.asp

K.sabreena. (2023, May 8). How to Improve the Reliability of ChatGPT: Techniques and Tips. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2023/05/how-to-improve-the-reliability-of-chatgpt-techniques-and-tips/

Laskowski, N., & Tucci, L. (2022, July 1). *What Is Artificial Intelligence (AI)?* TechTarget.https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence

La Venture, K., An, H. S., & Kim, W. (2024). An exploratory study of marketing students' perception of ChatGPT: Thank you for help! But I am also worried. The 2024 Association of Marketing Theory and Practice Proceedings, 32. https://digitalcommons.georgiasouthern.edu/amtp-proceedings_2024/32

Levitt, A. (2024, January 31). *Can ChatGPT Accelerate Social Science Research?* Yale Insights. https://insights.som.yale.edu/insights/can-chatgpt-accelerate-social-science-research

Loh, J. (2022, May 30). 4 Reasons Why You Should Remove Outdated Content. Internet Removals. https://internetremovals.com/singapore/4-reasons-why-you-should-remove-outdated-content/

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries?

Mariño-Romero, J., Palos-Sanchez, P. R., & Velicia-Martin, F. (2022).

Improving KIBS performance using digital transformation. Study based on the theory of resources and capabilities. Journal of Service Theory & Practice, 33(2), 169–197. https://doi.org/10.1108/ JSTP-04-2022-0095

Mariati, M., Abbas, E. W., & Mutiani, M. (2021). The social science contribution through social studies learning. *The Innovation of Social Studies Journal*, *2*(2), 110-120.

Marr, B. (2023, March 3). The Top 10 Limitations Of ChatGPT. Forbes. https://www.forbes.com/sites/bernardmarr/2023/03/03/the-top-10-limitations-of-chatgpt/?sh=256b0498f355

McCoy, J. (2024, February 6). 10 Most Common ChatGPT Limitations. Content @ Scale. https://contentatscale.ai/blog/chatgpt-limitations/

Meida Rabia Sihite, Meisuri Meisuri, & Berlin Sibarani. (2023). Examining the Validity and Reliability of ChatGPT 3.5-Generated Reading Comprehension Questions for Academic Texts. Randwick International of Education and Linguistics Science Journal, 4(4), 937–944. https://doi.org/10.47175/rielsj. v4i4.835

Metcalf, Ashley. (2017). Re: What is the acceptable range for Cronbach alpha test of reliability?. Retrieved from: https://www.researchgate .net/post/What-is-the-acceptable-range-for-Cronbach-alpha-test-of-reliability/59f09cf3ed99e1b508387762/citation/download.

Mhlanga, D. (2023). Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning. Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (February 11, 2023).

Middleton, F. (2023, June 22). Reliability vs validity in research. Scribbr. https://www.scribbr.com/methodology/reliability-vs-validity/

Miraj, M., Chuntian, L., Mohd Said, R., Osei-Bonsu, R., & Rehman, R. ur. (2021). How Information-Seeking Behavior, Essential Technologies, and Resilience Enhance the Academic Performance of Students. Frontiers in Psychology, 12, 651550. https://doi.org/ 10. 3389 /fpsyg.2021.651550

Mita, S. (2023). AI Proctoring: Academic Integrity vs. Student Rights. https://hastingslawjournal.org/wp-content/uploads/10-Mita_final.pdf

Mun, I. B. (2024). A study of the impact of ChatGPT self-efficacy on the information seeking behaviors in ChatGPT: the mediating roles of ChatGPT characteristics and utility. Online Information Review. https://doi.org/10.1108/oir-01-2024-proceedings_2024/32

Newcastle University. (2024). *Coefficient of Determination, R-squared*. Www.ncl.ac.uk; Newcastle University. https://www.ncl.ac.uk /webtemplate /ask-assets/external/maths-resources/s tatistics/ regression-and-correlation/ coefficient -of-determination-r-squared.html

Ngo, T. T. A. (2023). The Perception by University Students of the Use of ChatGPT in Education. *International Journal of Emerging Technologies in Learning (IJET)*, *18*(17), 4–19. https://doi.org/10.3991/ijet.v18i17.39019

Polit, D. F., & Beck, C. T. (2017). Nursing research: Generating and assessing evidence for nursing practice (10th ed.). Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins.

Potasheva, O. V., Kuzmenko, M. V., & Plutova, M. I. (2019, May). The effective use of digital technologies in education: positive experience of regional innovation platforms. In 1st International Scientific Conference" Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth"(MTDE 2019) (pp. 687-692). Atlantis Press.

Prove that F-statistic follows F-distribution. (2017, January 27). *Prove that F-statistic follows F-distribution*. Cross Validated. https:// stats. stackexchange.com/questions/258461/prove-that-f-statistic-follows-f-distribution

Ray, P. P. (2023). ChatGPT: a Comprehensive Review on background, applications, Key challenges, bias, ethics, Limitations and Future Scope. Internet of Things and Cyber-Physical Systems, 3(1), 121–154. https://doi.org/10.1016 /j.iotcps. 2023.04.003

Rodgers, J. (2023, May 22). *I Asked ChatGPT What It Thinks of "Content Is King" and Its Relevance Today—This Is How It Responded*. Www.linkedin.com. https://www.linkedin.com/pulse/i-asked-chatgpt-what-thinks-content-king-its-how-jeremy-rodgers

Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, *6*(1), Article 1.https://doi.org/10.37074/jalt.2023.6.1.23

Sain, S., & Negi, A. (2024, January). Trends & Challenges in Plagiarism & Text-Matching Software for Academic Integrity. https://www.researchgate.net /publication /377635781_ Trends_ Challenges_in_Plagiarism_Text_Matching_Software_for_Academic_ Integrity

Sambe, M. T. (2017). INFORMATION SEARCH STRATEGY AND RETRIEVAL TOOLS IN LIBRARIES

San, C. (2021, May 20). *Simple Linear Regression: A layman's explanation*. Nerd for Tech. https://medium.com/nerd-for-tech/simple-linear-regression-a-laymans-explanation-d60f1f2fc41b

SAP Help Portal. (2023, December 7). Help.sap.com. https://help.sap.com /docs/SAP_BUSINESSOBJECTS_BUSINESS_INTELLIGENCE_PLATFORM/c95594c101a046159432081ca44d6b18/4723f6e46e041014910aba7db0e91070.html

Saravanan, A. (2022, December 19). The Limitations of Chat-GPT. Medium. https://medium.com/@asarav/the-limitations-of-chat-gpt-8b73f5859bb4

Senthilnathan, S. (2019, July 9). *Usefulness of Correlation Analysis*. Papers.ssrn.com.https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3416918

Shah, J. N., Shah, J., Baral, G., Baral, R., & Shah, J. (2021). Types of plagiarism and how to avoid misconduct: Pros and cons of plagiarism detection tools in research writing and publication. Nepal Journal of Obstetrics & Gynaecology, 16(2).

Sheppard, V. (2020). 1.2 The Process of Undertaking Research. *Ecampusontario.pressbooks.pub*.https://ecampusontario.pressbooks.pub/researchmethodsandethics/chapter/1-2-the-process-of-undertaking-research/

Shubhrajyotsna Aithal, & Aithal, P. S. (2023). Effects of ai-based chatgpt on

higher education libraries. Zenodo (CERN European Organization for Nuclear Research), 8(2). https://doi.org/10.5281/zenodo.7905052

Sidoti, O., & Gottfried, J. (2023, November 16). *About 1 in 5 U.S. teens who've heard of ChatGPT have used it for schoolwork*. Pew Research Center. https://www.pewresearch.org/short-reads/2023/11/16/about-1-in-5-us-teens-whove-heard-of-chatgpt-have-used-it-for-schoolwork/

Singh, A. S., & Masuku, M. B. (2014). Sampling techniques & determination of sample size in applied statistics research: An overview. *International Journal of economics, commerce and management*, *2*(11), 1-22.

somlith, albert. (2023, October 4). *ChatGPT for PPC Keyword Research: Uncovering High-Performing Search Terms*. Ad Mockups for Client Approvals Made Easy | PPC Ad Editor. https://www. Ppcadeditor .com/chatgpt-for-ppc-keyword-research-uncovering-high-performing-search-terms-2/

solmaz. (2023, April 21). *Likert scale: definition and how to use it*. Mentimeter. https://www.mentimeter.com/blog/awesome-presentations/likert-scale-definition-and-how-to-use-it#:~:text=A%20Likert%20scale%20is%20a

Stevenson University. (2023). *How to Identify Reliable Information*. Stevenson University; Stevenson University. https://www. stevenson.edu/online/about-us/news/how-to-identify-reliable-information/

Strzelecki, A. (2024). ChatGPT in higher education: Investigating bachelor and master students' expectations towards AI tool. Education and Information Technologies. https://doi.org/10.1007/s10639-024-13222-9

Surucu, L., & Maslakci, A. (2020). Validity and reliability in quantitative research. *Business & Management Studies: An International Journal*, *8*(3), 694–726. http://dx.doi.org/10.15295/bmij.v8i3.1540

Sutaria, N. (2022, August 29). Bias and Ethical Concerns in Machine Learning.

ISACA. https://www.isaca.org/resources/isaca-journal/issues/ 2022/ volume-4/bias-and-ethical-concerns-in-machine-learning

Swargiary, K. (2024). The Impact of ChatGPT on Student Learning Outcomes: A Comparative Study of Cognitive Engagement, Procrastination, and Academic Performance. https://doi.org/10.2139/ssrn.4914743

team, B. (2023, May 16). *ChatGPT for SEO: Ultimate Guide, Tips & Prompts*. Backlinko. https://backlinko.com/chatgpt-for-seo

Templeton, S. (2024, April 3). ChatGPT Gets Prominent Links: What You Need to Know. Single Grain. https://www. Singlegrain .com /blog /n/chatgpt-prominent-links/

The Problem With Ambiguous Survey Questions – Drag'n Survey. (n.d.). Drag'n Survey. https://www.dragnsurvey.com/blog/en/the-problem-with-ambiguous-survey questions/#:~: text=An%20ambiguous % 20 question%20is%20defined

Tiunova, A. and Muñoz, F., Chatgpt: Using Ai in Social Studies Academic Research. *Available at SSRN 4451612*.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10(1), 15. https://doi.org/10.1186/ s40561-023-00237-x

Tubachi, P. (2018). INFORMATION SEEKING BEHAVIOR: AN OVERVIEW.

University of Newcastle Library. (2023, March 14). *LibGuides: Research Methods: What are research methods?* Newcastle.edu.au. https://libguides.newcastle. edu.au/researchmethods

von Garrel, J., & Mayer, J. (2023). Artificial Intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. *Humanities and Social Sciences Communications*, *10*(1), 1–9. https://doi.org/10.1057/s41599-023-02304-7

Walia, A., & Chetty, P. (2020, February 21). *How to Formulate a Research strategy?* Project Guru. https://www.projectguru.in/how-to-formulate-a-research-strategy/

Wankhede, C. (2024, May 18). How accurate is ChatGPT? Should you trust its

responses? Android Authority. https://www. Androidauthority .com/how-accurate -is-chatgpt-3436192/

Yahaya, Onipe. (2024). Re: What is the minimum number of respondents for a pilot study questionnaire? Retrieved from: https://www.researchgate.net/post/What_is_the_minimum_number_of_respondents_for_a_pilot_study_questionnaire/6642840e4dda7e36a0040e99/citation/download.

Zach. (2022, January 6). The Complete Guide: How to Report Skewness & Kurtosis. Statology. https://www.statology.org/how-to-report-skewness-kurtosis/

Zahreen Mohd Arof, K., Ismail, S., & Latif Saleh, A. (2018). Contractor's Performance Appraisal System in the Malaysian Construction Industry: Current Practice, Perception and Understanding. *International Journal of Engineering & Technology*, *7*(3.9), 46. https://doi.org/ 10. 14 419 /ijet .v7i3. 9.15272

Zhang, S., Liau, G., Tan, M., & Wei Liang Chua. (2024). Evaluating the accuracy and relevance of ChatGPT responses to frequently asked questions regarding total knee replacement. Knee Surgery & Related Research, 36(1). https://doi.org/10.1186/s43019-024-00218-5

APPENDIX A

GANTT CHART OF FYP 1

Year	2024/2025														
Task/Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
FYP 1 Talk															
First Meeting with Supervisor															
Topic Discussion															
Topic Confirmation															
Chapter 1 Lecture and Writing															
Chapter 2 Lecture and Writing															
Chapter 3 Lecture and Writing															
Submission Chapters 1 and 2															
Submission Chapters 1,2,3															
Prepare slide presentation															
Discussion slide presentation															
FYP 1 Presentation															
FYP 1 Correction															

APPENDIX B

GANTT CHART OF FYP 2

Year	2024/2025														
Task/Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Confirm Survey Questionnaires															
Discuss and make corrections															
Collect Data															
Discussion Chapters 4 and 5															
Training SPSS															
Chapter 4 Writing															
Chapter 5 Writing															
Submission Chapters 1-5															
FYP 2 Correction															
Prepare Viva Slides															
Standby for FYP 2 Presentation															
FYP 2 Presentation															
Complete FYP 2 Correction															

APPENDIX C

Assalamualaikum and greetings,

Dear Participant,

I am Nurul Najihah Binti Mazli, a final-year undergraduate student from the Faculty of Technology Management and Technopreneurship (FPTT) at Universiti Teknikal Malaysia Melaka (UTeM), currently pursuing a bachelor's degree in Technology Management (Technology Innovation). As part of my final year project, I am conducting research to evaluate an important research project titled " The Accuracy Level of ChatGPT in Information Search by Social Science Students."

This research has been chosen as nowadays most students, especially university students very efficient in using ChatGPT to help them do their work such as writing assignments, essays, and finding pieces of information. They already made the AI tool one of their reference materials. It is good as they have a lot of sources that students can refer to while doing their work. However, the students should not be too dependent on the information provided by the ChatGPT as it is still a technology, and the accuracy level also is not convincing on several factors faced by each accuracy level.

This survey is designed to gather your insights and experiences on using ChatGPT for academic purposes, focusing on how well the tool provides accurate and reliable information. Your participation in this survey is voluntary and will contribute significantly to my research.

This survey will take approximately 5-10 minutes to complete. All responses are anonymous, and the information gathered will be used solely for academic purposes. Thank you for your time and valuable contribution to this study.

PART A: DEMOGRAPHIC PROFILE

1. **Gender**

   ☐ Male

   ☐ Female

2. **Race**

   ☐ Malay

   ☐ Chinese

   ☐ Indian

   ☐ Other

3. **Course**

   ☐ BTMI

   ☐ BTMM

   ☐ BTMS

   ☐ BTEC

4. Level of Year

- ☐ Year 1
- ☐ Year 2
- ☐ Year 3
- ☐ Year 4

5. Device frequently Use to Access ChatGPT

- ☐ Telephone
- ☐ Laptop
- ☐ Tablet
- ☐ Other

6. Frequency Use of ChatGPT

- ☐ Daily
- ☐ Frequently (4 or more times a week)
- ☐ Sometimes (2-3 times a week)
- ☐ Occasionally (1-2 times a month)
- ☐ Rarely (less than once a month)

PART B: The Accuracy Level of ChatGPT

This section is to seek your opinion about the accuracy level of ChatGPT in information search by social science students. Based on the scale given, please choose which represents for your answer.

Scale:

Strongly Disagree	Disagree	Somewhat Agree	Agree	Strongly Agree
1	2	3	4	5

Independent Variable 1: Validity

No	Statement	1	2	3	4	5
1.	Data produced from ChatGPT can be inaccurate.					
2.	Limitations in understanding complex contexts may result in misleading responses.					
3.	Responses focused on smoothness can mistakenly cause wrong information.					
4.	Lack of source attribution from ChatGPT can lead to suspicions of plagiarism.					

Independent Variable 2: Reliability

No	Statement	1	2	3	4	5
1.	ChatGPT's capabilities are based on its "training data."					
2.	User interactions with ChatGPT influence the reliability of the information.					
3.	Outdated data delays ChatGPT's ability to understand newer information.					

Independent Variable 3: Relevance

No	Statement	1	2	3	4	5
1.	Keywords are important in guiding ChatGPT in generating accurate responses.					
2.	Keywords are essential in helping ChatGPT to understand the context.					
3.	Content quality impacts the relevance of the information provided by ChatGPT.					
4.	Clearer questions help ChatGPT to generate related responses.					
5.	A poorly organized link structure can cause ChatGPT to access information from irrelevant sources.					

PART C: Information Search by Social Science Students

This section is to seek your opinion about the information search by social science students. Based on the scale given, please choose which represents your answer.

Scale:

Strongly Disagree	Disagree	Somewhat Agree	Agree	Strongly Agree
1	2	3	4	5

No	Statement	1	2	3	4	5
1.	Trust in ChatGPT influences the frequency of uses.					
2.	Social media is the main way you learn about ChatGPT.					
3.	Students mainly use ChatGPT for assignment purposes.					
4.	Information from ChatGPT is more reliable than other AI tools.					
5.	Self-assessing (re-checking) ChatGPT's answers can help students avoid untruth.					