# On ChatGPT: Perspectives from Software Engineering Students

Khadija Hanifi[1,2,*], Orcun Cetin[2], and Cemal Yilmaz[2],
[1]Ericsson Research Turkey, Istanbul, Turkey
[2]Sabanci University, Istanbul, Turkey
khadija.hanifi@ericsson.com,
khadija.hanifi,orcun.cetin,cemal.yilmaz@sabanciuniv.edu
*Khadija Hanifi

*Abstract*—ChatGPT, an increasingly popular Large Language Model (LLM), has found widespread acceptance, especially among the younger generation, who rely on it for various tasks, such as comprehending complex course materials and tackling homework assignments. This surge in interest has drawn the attention of researchers, leading to numerous studies that delve into the advantages and disadvantages of the upcoming LLM dominant era. In our research, we explore the influence of ChatGPT and similar models on the field of software engineering, specifically from the perspective of software engineering students. Our main objective is to gain valuable insights into their usage habits and opinions through a comprehensive survey. The survey encompassed diverse questions, addressing the specific areas where ChatGPT was utilized for assistance and gathering students' reflections on each aspect. We found that ChatGPT has garnered widespread acceptance among software engineering students, with 93% of them utilizing it for their projects. These students expressed satisfaction with the level of assistance provided, and most intend to continue using it as a valuable tool in their work. During our investigation, we also assessed the students' awareness of the underlying technologies behind ChatGPT. Approximately half of the students demonstrated awareness of these technologies, while 38.7% had made extra efforts to explore prompt engineering to enhance ChatGPT's productivity. However, an important finding was that 90.6% of the students reported experiencing hallucinations during their interactions with ChatGPT. These hallucinations were shared as examples, raising significant concerns that warrant further exploration and mitigation. Moreover, we delved into potential improvements and gathered valuable recommendations, which could help ChatGPT to become even more effective and dependable in its applications.

*Keywords–ChatGPT; software engineering; academic education; generative AI; Large Language Models*

## 1. Introduction

Over the past few years, many LLMs and Generative AI models have emerged, significantly impacting both industry and society. These models showcase diverse capabilities and have been published to address various tasks. For instance, models like DALL-E2 [1] have the ability to transform texts into images, while Dreamfusion [2] excels in converting text to 3D images. The Flamingo model [3] can adeptly generate text from images, and Phenaki [4] is designed to transform texts into videos. Meanwhile, the AudioLM model [5] can convert texts into audio, and ChatGPT excels in generating texts from other texts. This remarkable progress in the development of Generative AI models has paved the way for a myriad of innovative applications across various domains. However, the ease of use, providing fast, clear, and to-the-point responses, and its versatility in assisting with tasks from simple conversations to complex technical assignments have propelled ChatGPT's popularity, surpassing that of other Generative AI models [6]. As a result, ChatGPT has gained popularity among students who are not yet experts in their fields but require assistance in better understanding their course materials and completing their assignments [7]. This represents a significant development in the education system that could profoundly impact the way schools and research work are conducted [8], potentially necessitating new evaluation and grading systems [9], [10]. Moreover, this technology could even reshape the entire hiring process, potentially replacing certain jobs with AI models [11], [12]. In this study, we focus on the impact of Generative AI models on the education system in the field of software engineering. To assess this impact, we conducted a survey study that reflects students' usage habits and their perspectives and views towards ChatGPT and LLMs. More specifically, we aim to address the following research questions:

- **RQ1:** How do students perceive the assistance provided by ChatGPT across various tasks, and how does this perception align with their usage patterns?
- **RQ2:** To what extent do students place trust in the outputs generated by ChatGPT, and how does this trust impact their subsequent actions and decision-making?
- **RQ3:** Is there a correlation between students' knowledge and experience with ChatGPT and their effective usage of the tool?
- **RQ4:** What are the challenges faced by ChatGPT, and what limitations are linked to its usage?
- **RQ5:** What is the general attitude of students towards ChatGPT and other LLMs? Do they perceive them as valuable tools that facilitate their work, and do they intend to continue relying on them in the future?

## 2. Related Work

Recently, a surge of research studies [13], [14], [15] has focused on exploring the impact of LLMs, with particular attention to ChatGPT, on academia and education systems.

2693-9177/23/$31.00 ©2023 IEEE
DOI 10.1109/QRS60937.2023.00028

These studies have shed light on the opportunities and challenges that arise with the integration of LLMs into educational settings. For instance, in the study by Fraiwan et al. [13], the researchers examined the overall impact of LLMs on educational practices and the evolving role of AI-driven technologies in shaping the learning experience. Sok et al. [16] explored the benefits and risks of employing ChatGPT in educational and research contexts. Lund et al. [15] conducted an interview-based study with ChatGPT in discussions about its potential impact on academia and libraries. These works highlighted many advantages, such as creating learning assessments, improving pedagogical practices, offering virtual personal tutoring, generating essay or research article outlines, and fostering brainstorming activities. However, they also addressed potential risks, including academic integrity issues, unfair learning assessments, inaccurate information, and overreliance on AI for critical tasks. In the realm of academic research, Rahman et al. [17] presented a practical example of ChatGPT's application in initial idea generation for scientific research. They emphasized the potential effectiveness of ChatGPT in this context but cautioned about challenges related to literature synthesis, citations, problem statements, research gaps, and data analysis. Recommendations were made to establish guidelines for appropriate LLM usage in research and publishing. In the pursuit of harnessing the benefits of LLMs in education, authors in [18] proposed mitigation strategies to address the discussed risks. Their work underlined the importance of providing trustworthy and equitable access to LLMs for educational purposes.

Lo et al. [9] conducted a rapid review of 50 articles to understand ChatGPT's capabilities across various domains, including its potential use in education. The findings revealed that ChatGPT's performance varied across subjects, with outstanding results in some areas, such as economics, and more satisfactory outcomes in programming. However, concerns have also been raised regarding the misuse of ChatGPT. Alafnan et al. [19] highlighted that unethical use of ChatGPT by students could potentially lead to a decline in critical thinking and overdependence on automation. The study proposed measures to mitigate these concerns, such as designing assessment tasks that promote personalized, critical, and imaginative thinking.

An investigation by Forman et al. [20] delved into the usage patterns and future perceptions of ChatGPT among high school students. They highlighted the integration of technology, particularly NLP, in contemporary life and the younger generation's enthusiasm for adopting emerging technologies. High school students reported using ChatGPT for various purposes, including academic support, social communication, and personal management, demonstrating its growing significance in both educational and social contexts.

Regarding awareness and content promotion, Haensch et al. [7] analyzed the content of the 100 most popular videos related to ChatGPT on TikTok in February 2023. The majority of these videos promoted ChatGPT's use in writing essays or code, often focusing on how to transform ChatGPT's output to deceive AI detectors. Wagholikar et al. [21] examined the level of awareness of ChatGPT among students, and their results indicated a consistent level of awareness across genders and academic disciplines.

While existing literature has explored the benefits and limitations of LLMs from the perspectives of academicians and researchers, this study offers a closer look at the usage of ChatGPT by software engineering students. Through a case study conducted via a survey-based approach, we aim to understand how ChatGPT is assisting them in their project assignments and its implications on their education journey.

## 3. CHATGPT

ChatGPT is a state-of-the-art artificial intelligence chatbot developed by OpenAI [1], introduced in November 2022 [22]. It is powered by GPT-3.5 and GPT-4, two models from OpenAI's proprietary series of foundational GPT models. The main goal of ChatGPT is to provide users with cohesive and informative human-like responses to their input.

### 3.1. Capabilities and Training:

ChatGPT's capabilities stem from its extensive training on over 150 billion human-generated items, including books, articles, blog posts, conversations, and reviews [23]. This training enables the model to understand the context of conversations and generate accurate responses. The platform has experienced immense popularity, attracting over one million users within the first week of its launch [24].

### 3.2. Improving Accuracy with CGA:

To enhance accuracy and generative capabilities, OpenAI introduced the ChatGPT Improved Accuracy (CGA) model. CGA utilizes deep learning-based artificial intelligence architecture and learns from its own mistakes, allowing it to adapt to new contexts and produce more accurate results. Recent research has shown that CGA outperforms other popular NLP models in terms of accuracy, coherence, and readability [25].

### 3.3. Addressing Hallucination in AI:

Hallucination in AI refers to the generation of plausible but factually incorrect or unrelated outputs due to inherent biases, lack of real-world understanding, or training data limitations. While ChatGPT has been improved to handle hallucinations better than its predecessors, it is still known to present inaccurate information confidently [26].

### 3.4. Freemium Model and Commercial Offering:

Initially released as a research preview, ChatGPT's popularity led OpenAI to adopt a freemium model. Users on the free tier can access the GPT-3.5 based version, while paid subscribers gain access to the advanced GPT-4 based version and priority access to newer features, under the commercial name "ChatGPT Plus."

[1] https://openai.com/

### 3.5. Rapid Growth and Competing Products:

ChatGPT gained tremendous traction and became one of the fastest-growing consumer software applications in history, reaching over 100 million users by January 2023. Its success spurred competitors like Google, Baidu, and Meta to accelerate the development of their own AI chatbot products, such as Bard [2], Ernie Bot [27], and Large Language Model Meta AI (LLaMA) [28].

### 3.6. Prompt Engineering and In-Context Learning:

Prompt engineering is a crucial approach to designing interactions with LLMs like ChatGPT. It involves developing robust and effective prompting techniques that improve the model's performance on various tasks, such as question-answering and arithmetic reasoning. This method allows the model to adapt and improve its responses based on the user's prompts during a conversation, while ensuring it does not carry unwanted old contexts or biases between different interactions [29].

## 4. Experimental Setup

To conduct this survey, students from Sabancı University who were enrolled in the CS 308 Software Engineering course were invited to participate. Participation in the survey was entirely voluntary, and students had the right to refuse or withdraw from the study at any time without facing any penalties or loss of benefits. However, upon successful completion of the survey, participants were eligible to receive a bonus credit in their Software Engineering course, equivalent to 3% of the overall course grade. It should be noted that an alternative assignment was provided for students who chose not to participate in the survey. This alternative assignment awarded the same credit points as the survey, ensuring equal opportunities for all students. The allocation of bonus points is solely intended to encourage student participation in the study and is unrelated to their academic performance. Similarly, participation in the survey and the responses provided do not have any impact on their scores or course/project performance. Moreover, to ensure anonymity and minimize bias, the responses provided by students were kept confidential, and participants were unable to view each other's replies.

In crafting our survey questions, we employed a structured approach, dividing them into two primary categories. The first set focused on gathering demographic information, encompassing participants' majors, grades, work experience, and software-related background. The second set delved into more specific inquiries concerning ChatGPT usage habits, satisfaction levels, and solicited suggestions. Additionally, we incorporated questions pertaining to students' knowledge of promoting engineering and their experiences with hallucination. To maintain consistency and facilitate meaningful responses, we utilized a fixed-format approach for answer choices. For instance, queries measuring frequency employed a five-point scale: "Never," "Rarely," "Sometimes," "Often," and "Always." Similarly, questions seeking recommendations featured options such as "Strongly Disagree," "Disagree," "Undecided," "Agree," and "Strongly Agree." To ensure that the responses remained contextually relevant, we carefully orchestrated the sequence and interrelation of questions. As an example, in Part 2 focusing on ChatGPT, we initially inquired, "Have you utilized ChatGPT during your software engineering course project?" Only if the response was affirmative did we proceed to seek input on ChatGPT usage experiences; otherwise, these questions were intelligently skipped.

[2]https://bard.google.com/

## 5. Demographic

The survey conducted for this study involved a total of 113 students, comprising mainly of 3rd and 4th grade students from the Computer Science and Engineering department, with a few pursuing a dual major in other engineering fields. Notably, an overwhelming majority of 93.8% of the participants reported utilizing ChatGPT during their course projects.

Regarding employment status, a significant majority of 68.1% of the respondents were not engaged in any work, while 23% were undertaking internships, and 8.8% were working part-time jobs. When asked about their development roles, 37.2% of the respondents identified themselves as full-stack developers, 14.2% as back-end developers, 18.6% as front-end developers, and 22.1% as data scientists. In terms of development types, the participants indicated a significant inclination towards web development (46%), followed by data analysis (27.4%), mobile app development (15.9%), and system-level (low-level) development (8%).

In terms of the participants' software background, only one individual had not taken any software-related courses apart from software engineering. More specifically, 53.1% of the students had taken six or more additional courses related to software, while 31.9% had completed between three and five courses, and 14.2% had taken one or two such courses. Regarding programming experience, the majority of participants (68.1%) reported having 1-3 years of programming experience, while 19.5% indicated having 3-5 years of experience, and a smaller subset (10.6%) reported having less than one year of programming experience.

## 6. Survey Responses

This section presents an in-depth analysis of the students' responses to the survey questions, exploring various perspectives, including usage and effectiveness, experience and satisfaction, prior knowledge of the technologies behind ChatGPT, as well as challenges and concerns.

### 6.1. Usage and Effectiveness

During the software engineering course projects, 93% of the students utilized ChatGPT. Among them, 80.2% were categorized as active users. The number of prompts required to obtain useful answers varied widely, ranging from 1 to more than 10, as depicted in Figure 1.

The primary task for which ChatGPT was employed by 77.4% of the students was generating source code. Out of these users, 13.4% reported often using the ChatGPT-generated source

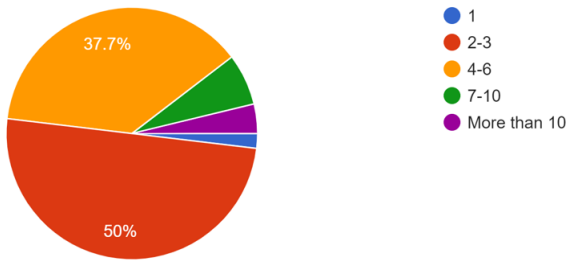


Figure 1. Number of prompts required to get helpful answers from ChatGPT.

code as-is, 32.9% used it sometimes, 34.1% rarely used it, and 19.5% never used it. On the other hand, 11% always made modifications to the generated code, 43.9% often made changes, 30.5% sometimes made changes, 12.2% rarely made changes, and 2.4% never made any modifications.

Figure 2 presents the usage of ChatGPT for different tasks, indicating that it was less frequently used for tasks such as log analysis and resolving security-related issues. Conversely, it was more commonly utilized for tasks involving algorithm suggestions and defect detection. Notably, there was a diverse range of applications for ChatGPT across various tasks, highlighting the wide adoption and multifaceted usage of ChatGPT as an assistant to seek help in different stages of their projects.

Figure 3 illustrates the students' opinions regarding the helpfulness of ChatGPT for each task. Students' reports indicate that ChatGPT demonstrated higher levels of effectiveness in certain tasks. Specifically, it was perceived as more helpful in tasks related to defect detection, generating source code, and debugging. On the other hand, its performance was perceived to be comparatively less effective in log analysis and resolving security-related issues. This assessment sheds light on the varying strengths and limitations of ChatGPT's capabilities across different project tasks, further informing its practical utility and potential areas of improvement.

Note that during the course project students have the flexibility to choose the platforms and programming languages that they feel most comfortable with. This provided an evaluation of a variety of technologies and helped to make a more comprehensive and fair evaluation. From Figure 4, we see that 79.9% of the students used Windows operating system, 26.5% used MacOS, 18.6% used Linux/Unix, 15.9% IOS, and 33.6% Android. The top-5 used front-end technologies are reported as: CSS, HTML, JavaScript, React, and Bootstrap. The top-5 used back-end technologies are reported to be: NodeJS, Python, Java, Spring Boot, and Django. The database technologies used by the students are: MongoDB, SQLite, MySQL, and Firebase.

It is worth noting that throughout the course project, students were granted the flexibility to choose platforms and programming languages that aligned with their comfort and familiarity. This approach facilitated an evaluation encompassing a variety of technologies, leading to a more comprehensive and equitable assessment. This wide array of technology usage further contributes to a comprehensive and well-rounded evaluation of the students' experiences with getting assistance from ChatGPT.

### 6.2. Experience and Satisfaction

Regarding the students' experience and satisfaction, 50% of the students reported satisfaction, with 15.1% indicating they were very satisfied, 34% stating it was okay, and only one student expressing dissatisfaction.

In terms of ChatGPT's impact on productivity, students' feedback was overwhelmingly positive. As depicted in Figure 6-B, 50% of the students agreed that ChatGPT improved their productivity, 38.7% strongly agreed, 9.4% were undecided, while only two students disagreed or strongly disagreed. When it came to trusting in ChatGPT's answers, 49.1% of students trusted it sometimes, 29.2% often, 20.8% rarely, and a mere 0.9% reported always trusting its responses. Interestingly, for 20.8% of the students, ChatGPT was their first choice tool, even surpassing search engines like Google, highlighting its preferred position as an assistance tool during the course project.

Furthermore, the students' inclination to continue using ChatGPT in future software engineering projects was assessed. The results indicate that 37.7% of students expressed a definite intention to use it, 35.8% were likely to use it, and 25.5% considered the possibility of using it.

Moreover, a significant proportion of the students were enthusiastic about recommending ChatGPT to other software engineers. 46.2% strongly recommended its usage, and an additional 43.4% expressed a positive recommendation for ChatGPT to their peers in the software engineering community.

These favorable recommendations underscore the students' trust and satisfaction in using ChatGPT to assist them in their work. The high likelihood of continued usage in future software engineering projects, along with the strong endorsement to other software engineers, reflects the positive impact and effectiveness of ChatGPT as a valuable tool in the students' software development endeavors.

### 6.3. Prior knowledge of the technologies behind ChatGPT

Regarding LLMs, less than half of the students (48.1%) were familiar with them and their usage in ChatGPT. Interestingly, among those who knew about LLMs, half of them had prior experience using other LLMs for software engineering-related tasks. This indicates a level of awareness and prior exposure to similar technologies among a portion of the students surveyed. Among the students surveyed, it was found that 61.3% were aware of prompt engineering, and of those, 58.5% recognized that effective usage of ChatGPT requires prompt engineering. However, only 38.7% of the students had taken the initiative to read about prompt engineering and/or explore some examples of it to enhance their utilization of ChatGPT. Notably, this final group's feedback about the impact of prompt engineering is depicted in Figure 6-C, with 9 students reporting no discernible effect from prompt engineering.

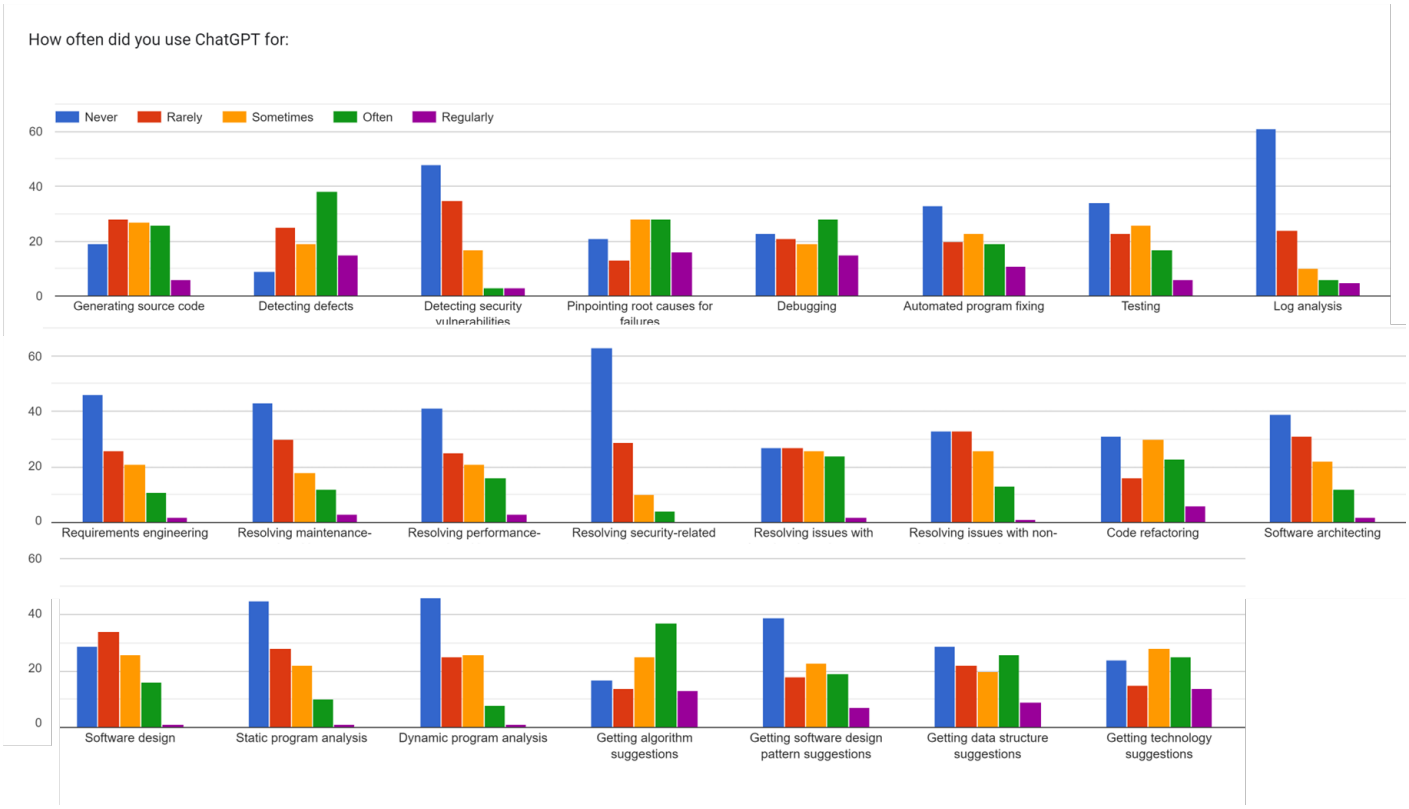


Figure 2. Tasks students used ChatGPT to get help for.

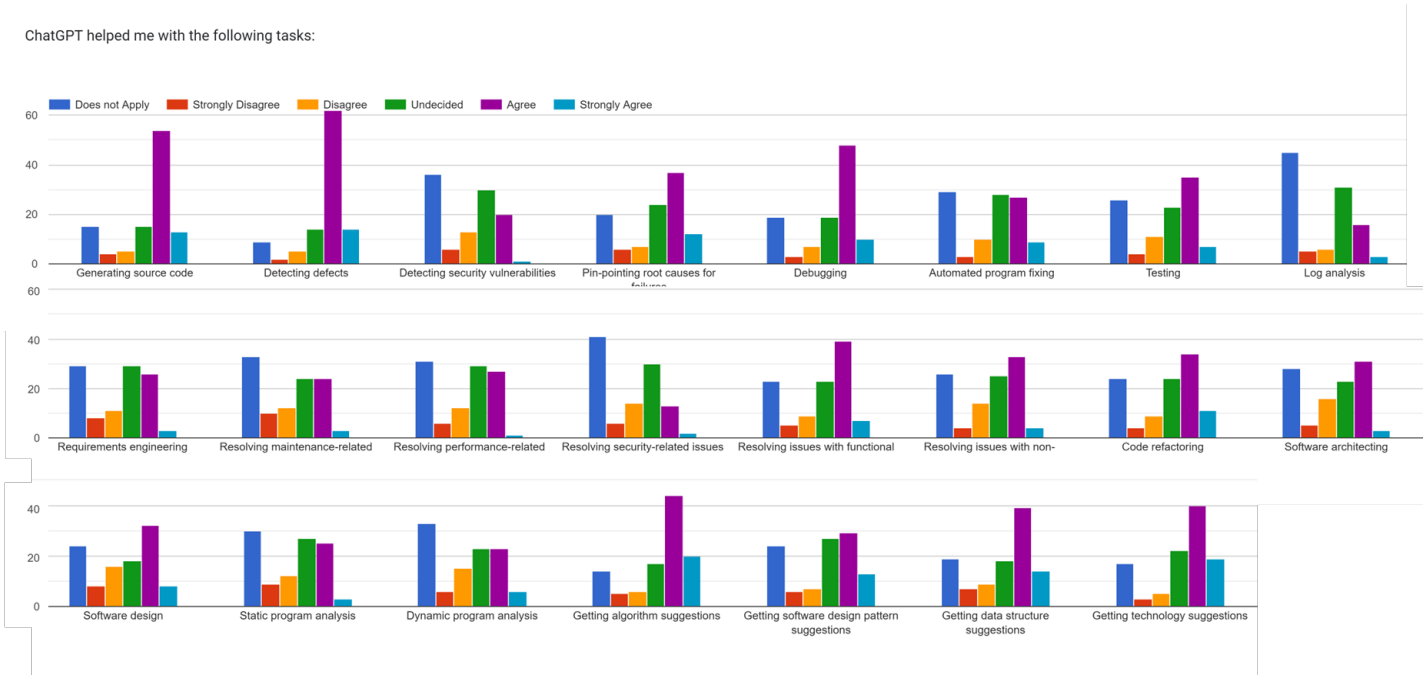


Figure 3. Students' feedback regarding ChatGPT performance in different tasks.

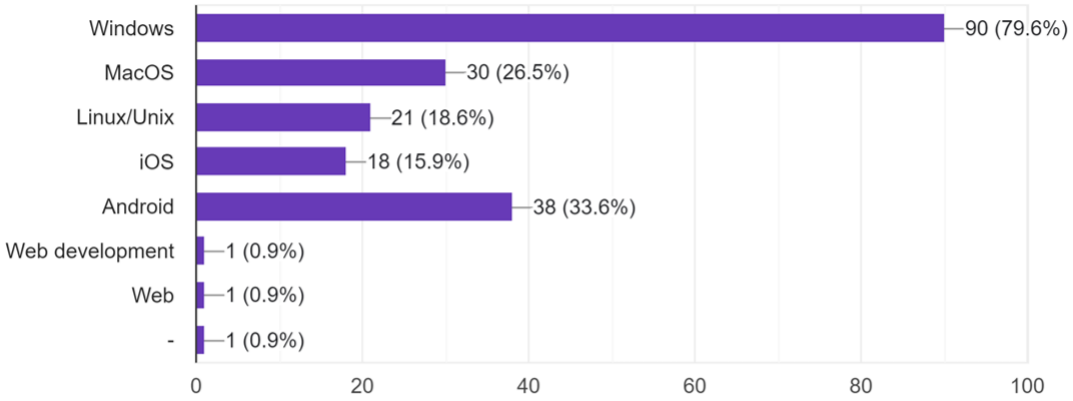


Figure 4. Operating systems students primarily develop software for.

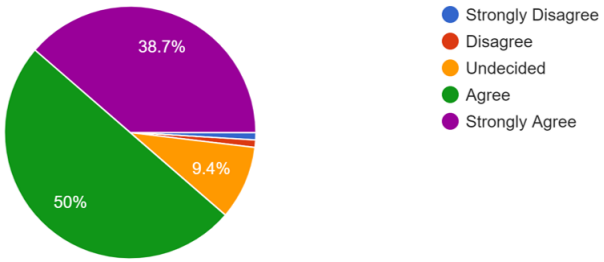


Figure 5. Results regarding if ChatGPT improved the students' productivity in the course project.

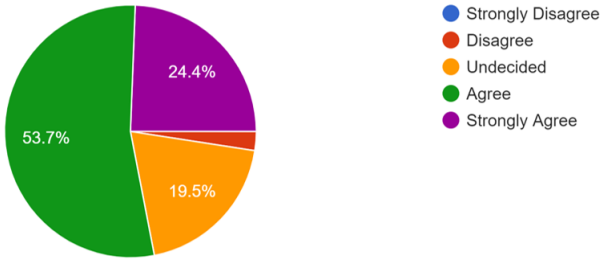


Figure 6. Using ChatGPT in a more effective/productive manner after reading about prompt engineering and/or examining some prompt engineering examples.

### 6.4. Challenges and Concerns

Based on the students' experiences, a majority (61.3%) found ChatGPT's answers to be often helpful, while 33% considered them to be sometimes helpful. However, a notable observation was that 90.6% of the students experienced hallucinations in ChatGPT. These hallucinations varied in frequency, with 27.1% experiencing them often, 38.5% sometimes, and 33.3% rarely. Some of the reported hallucinations include instances where ChatGPT provided incorrect code, repeated responses, gave wrong explanations for true/false questions, and even made up non-existent functions or referred to non-existent

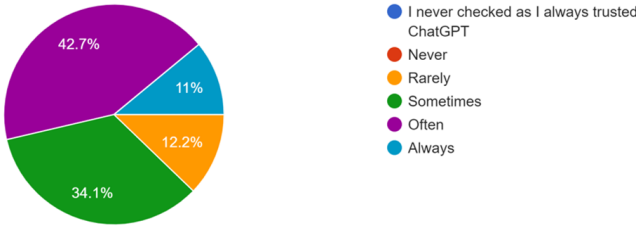


Figure 7. How often did ChatGPT generate code with some bugs (defects) in it.

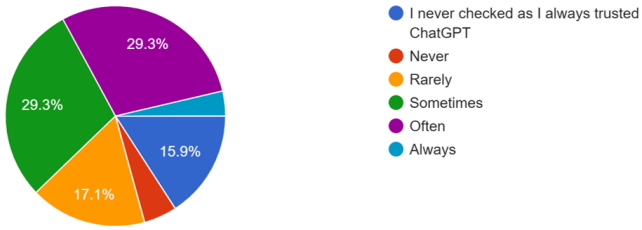


Figure 8. How often did ChatGPT generate code with some security vulnerabilities in it.

packages. The pattern of hallucinations suggests that ChatGPT may face challenges, especially in tasks involving logic and mathematical statements, and when dealing with libraries and functions from implemented libraries.

In terms of generated code quality, 42.7% of the students mentioned that the code produced by ChatGPT often included bugs or defects, while 29.3% highlighted the presence of security vulnerabilities in the generated code. Further details regarding these bugs and vulnerabilities are presented in Figure 7 and Figure 8, shedding light on the specific issues encountered.

These findings emphasize the need for continued improvements in the reliability and accuracy of ChatGPT, especially in handling complex logical and mathematical concepts, as well as ensuring the generation of secure and bug-free code to enhance its overall effectiveness as an assistant for software

engineering tasks.

## 7. Recommendations for Improving ChatGPT

We sought feedback from the students regarding potential suggestions for improving ChatGPT to better support software engineering activities, and 45 (42.5%) of them responded with affirmative suggestions. We manually assessed each recommendation with the help of two coders. Throughout this process, we had a third coder on standby to mediate any disagreements, but no conflicts arose. Ultimately, we pinpointed 53 distinct themes. We found out that suggestion can be mainly summarized into six categories : (i)improving the responses, reasoning, and prompt restrictions; (ii) improving the performance in terms of bug and defect fixing; (iii) does not require any improvement; (iv) prompt sharing and improvement features; (v) access to real-time internet and updated data; and (vi) integration with coding environments. These suggestions cover 88.6% of the total number of suggestions. Figure 9 displays the number of unique students associated with each suggestion.

### 7.1. Improving the responses, reasoning, and prompt restrictions

Many students expressed concerns about ChatGPT's responses, noting they often found them lengthy, confusing, and occasionally inaccurate. They suggested ChatGPT could be enhanced by ensuring answer accuracy and clearly indicating the confidence level of its responses. Furthermore, they felt ChatGPT was more beneficial for back-end tasks than front-end development and emphasized the need for improvements in the latter. Additionally, when working on specific tasks like website development using the Django framework, they noticed that ChatGPT sometimes focused on minor issues rather than providing holistic assistance. They also highlighted the challenge of maintaining codebase context throughout a conversation, as it tends to get lost after several exchanges. While tools like LangChain and PineconeDB allow for the creation of vector databases, they aren't always user-friendly for fresh projects. Many would appreciate a service where the codebase can be connected in real-time, ensuring the language model is updated with the most recent code context.

Reasoning emerged as an another priority. Students mentioned that ChatGPT should offer more justification for its answers and refrain from presenting irrelevant details. Rather than attempting to address every inquiry directly, the tool might benefit from posing questions to obtain a clearer understanding of the user's needs.

They also raised concerns about the length of prompts, suggesting that longer prompts might allow for more comprehensive information exchange. They observed that, at times, the model might overlook details shared earlier in the conversation. Minimizing the number of prompts needed to obtain an accurate answer would enhance efficiency, particularly in intricate scenarios where ChatGPT currently faces challenges. To ensure more pertinent feedback, it would be beneficial to provide ChatGPT with a broader context, insights, and a history of previous interactions. Moreover, feeding the tool with extensive code can prove daunting, reducing its effectiveness for those who might not be well-versed in the subject matter.

### 7.2. Improving the performance in terms of bug and defect fixing

The students' second biggest concern centered around the quality of the generated code, particularly the presence of defects. They noted that ChatGPT's accuracy in detecting and fixing defects and bugs is unreliable, making it unsuitable for fully depending on such tasks. Even when clear defective code is provided and labeled as such, ChatGPT's attempts at fixing it can introduce new problems. Some students reported instances where using ChatGPT required multiple attempts to fix code, and at times, the tool offered identical code without any improvements. To address this, they suggested that ChatGPT should inquire about code connections and propose fixes for related sections. The students emphasized that ChatGPT's generated code often contains numerous bugs and frequently results in a nonfunctional solution in its initial iteration, even when specific requirements are provided. Most notably, the students highlighted a critical caution: users should exercise extreme care when utilizing code generated by ChatGPT. There's a concern that similar code structures could be produced for different users, potentially leading to security risks. In the unfortunate event of a system being compromised, the risk extends to other systems with similar structures that were generated using ChatGPT.

### 7.3. Does not require any improvement

In addition to the 57.5% of surveyed students who expressed that they had no specific suggestions for enhancing ChatGPT's support for software engineering activities, approximately 15.5% of the remaining participants conveyed that ChatGPT had already met their expectations comprehensively. They highlighted its effectiveness for different tasks and saw no immediate need for further enhancements. However, these respondents shared unique and valuable perspectives on ChatGPT's performance and potential refinements. Their insights ranged from recognizing its skill-sustaining impact to the nuances of prompt construction and even extended to advocating for its recognition and integration within educational systems. Some students shared a noteworthy perspective: they shared that they were on the brink of forgetting their Python coding skills, as ChatGPT resolved most of their programming challenges. They even voiced a plea to halt further model improvements, underscoring the considerable impact it had already made in addressing their programming hurdles. Conversely, some respondents held the viewpoint that ChatGPT's performance seemed more robust during its initial release. This highlights the earlier version's effectiveness and encourages consideration of maintaining or restoring certain features. Additionally, a group of students emphasized the importance of crafting effective prompts. They pointed out that while learning to write impactful prompts takes time, it

is ultimately rewarding. Furthermore, a subset of participants stressed the significance of integrating tools like ChatGPT into educational systems. They advocated for the recognition of these technologies by educational institutions and proposed the implementation of necessary updates to facilitate their incorporation into learning environments. This underscores the potential impact of ChatGPT and similar tools in shaping modern educational practices.

### 7.4. Prompt sharing and improvement features

Approximately 15.5% of the suggestions were made to address the lack of practical prompt generation features. These students believed that ChatGPT can give them the results they are looking for, but making it work well takes a lot of time because of how they ask questions. The process of formulating suitable prompts emerged as a time-consuming endeavor for these individuals. To make this process smoother, students suggested that OpenAI could make several improvements. Some of them suggested that OpenAI could provide supplementary resources dedicated to prompt engineering tailored for software developers. This would save them time and help them get better results from ChatGPT. Furthermore, students recommended the establishment of novel communities aimed at sharing impactful prompt formulations. By fostering collaborative environments wherein users can exchange successful prompt strategies, OpenAI could foster a collective learning process. This notion aligns with the concept of creating platforms where individuals can pool their expertise to develop and disseminate effective prompts, enhancing the overall user experience.

### 7.5. Access to real-time internet and updated data

Students have provided various suggestions to enhance ChatGPT by integrating access to real-time internet and updated data. A few have voiced concerns that the current model, trained only up to 2021, can't address issues specific to newer versions of software libraries. They believe real-time updates to the training data would greatly improve its utility. Additionally, the capability for ChatGPT to read and interpret links was a sought-after feature. There's also a notable interest in allowing ChatGPT to access popular coding platforms like StackOverflow and GeeksforGeeks. Drawing from such repositories could generate outputs and even cite its sources, ensuring users receive the most relevant and up-to-date solutions. Lastly, a prevailing sentiment is a desire for more accurate responses, which could be achieved by granting the model internet access, paving the way for an enhanced user experience.

### 7.6. Integration with coding environments

We found out that around 13% of the suggestions mentioned the need for ChatGPT and development environment integration. Students mentioned that integrating ChatGPT with IDEs, especially popular ones like VSCode, can significantly enhance the user experience. Such integrations allow the model to better understand the context, project structure, and which code can be repurposed or imported. In addition, one student

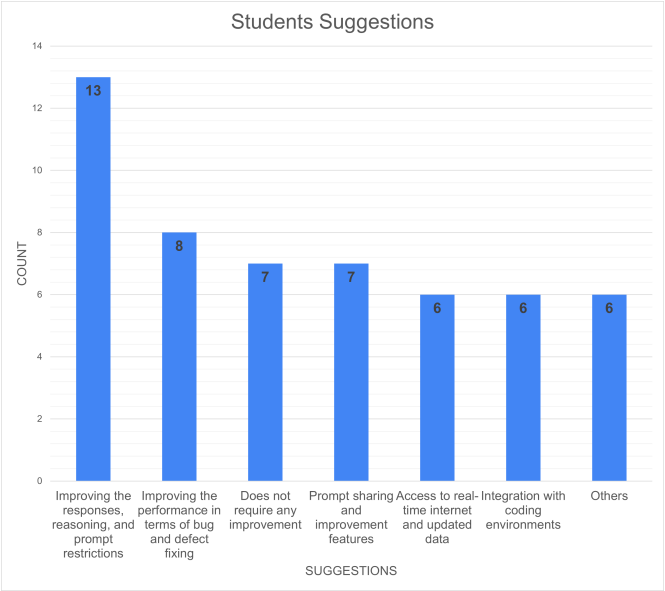


Figure 9. Students' suggestions to improve ChatGPT performance clustered under six main categories.

mentioned that there is an existing GPT extension for VSCode which is currently in beta and there is a waiting list to be a tester. A few students mentioned that a plugin for VSCode that automatically reads and understands all of a user's code files would be very useful for developers. This would lead to more accurate suggestions, considering the limitations when using ChatGPT from a browser, where it's often challenging to send lengthy code and receive accurate advice due to lack of context. Additionally, another student suggest that the idea of having specialized versions of ChatGPT tailored for different subjects can ensure developers receive relevant assistance without distractions from unrelated topics.

### 7.7. Others

Lastly, a handful of additional suggestions were made to refine ChatGPT's code-generation capabilities. One idea emphasizes expanding the interface, allowing users more room to guide the question bot toward a desired answer. Integrating a built-in compiler is another proposal, enabling users to directly run and test the generated code within the platform. To further this interactive experience, there's interest in allowing users to manually tweak the provided code and subsequently run it in the browser, ensuring real-time validation. Additionally, introducing a feature to upload files, such as source code, would streamline the process of code analysis and recommendation, potentially making ChatGPT an even more invaluable tool for developers.

## 8. Threads to Validity

In this section, We discuss two key aspects that require consideration:

**Sample Selection Limitations:** Our sample selection process was mainly focused on undergraduate students majoring in

software engineering. This specific demographic focus may constrain the generalizability of our findings to a more diverse population. Future research could broaden the scope by incorporating a more heterogeneous sample, including graduate students and industry professionals. Furthermore, our sample is drawn from a single academic institution, potentially introducing institutional biases. Different institutions may have varying approaches to software engineering education, which could impact students' familiarity and interaction with AI tools like ChatGPT. Thus, researchers should exercise caution when extrapolating our findings to institutions with significantly different academic contexts.

**Questionnaire Design Limitations:** The design of our questionnaire, while meticulously crafted to gather relevant insights, may still pose limitations. Firstly, the questions were structured in a self-report format, relying on participants' recollections and subjective assessments. This introduces the possibility of recall bias, potentially influencing response accuracy. Additionally, our questionnaire design employed fixed-format response options, such as 5-point Likert scales, for various questions. While these formats provide quantifiable data, they may not capture the full spectrum of participant perspectives or experiences. Future research could consider incorporating open-ended questions or qualitative interviews to delve deeper into participant experiences and viewpoints, thereby enriching the qualitative aspects of our study.

## 9. Conclusion and Future Work

The recent popularity of AI-powered tools like ChatGPT has garnered significant attention within the research community, leading to numerous studies and discussions about the new AI-based future. In this study, our primary objective is to address a fundamental question: How do software engineering students employ ChatGPT? To achieve this, we conducted a survey to analyze the perspectives of software engineering students. Our aim is to understand how they utilize ChatGPT and for what purposes. Our analysis indicates that software engineering students are highly active users of ChatGPT in their projects, expressing overall satisfaction with its performance and intending to continue using it for upcoming endeavors. However, they have also raised concerns about hallucinations in ChatGPT and have offered valuable suggestions for improvements to enhance their experience and obtain more accurate outputs efficiently.

Overall, the insights derived from the responses emphasize several practical implications that could significantly impact the field of software engineering education. 93.8% of respondents reported using ChatGPT for quick problem-solving during software engineering course projects, which underscores the practical role ChatGPT plays in enhancing the learning experience of software engineering students by providing immediate assistance and solutions. Our research suggests that ChatGPT is more than just a supplementary tool; it has the potential to influence curriculum development in software engineering programs. Additionally, ChatGPT can contribute to the development of essential skills among software engineering students. The 77.4% of respondents who reported using ChatGPT for code optimization and debugging indicate that it can serve as a valuable resource for honing critical problem-solving and debugging skills, aligning with industry demands. Educators can take actionable steps based on our findings, including providing guidelines for the effective and ethical use of AI tools like ChatGPT and incorporating AI-assisted problem-solving exercises into their teaching methods to foster adaptability and industry relevance. However, It's important to note that while ChatGPT offers numerous advantages, ethical considerations and potential challenges must also be acknowledged. Educators should be cautious about overreliance on AI tools and ensure that students develop core problem-solving skills independently.

Looking ahead, our upcoming research endeavors aim to elevate this study by investigating the readiness of the education system to adopt this technological revolution. We aim to understand the optimal pace for integrating these changes and identify the necessary steps to facilitate seamless adoption. Additionally, we plan to investigate the impact of ChatGPT on students' work styles, communication, and collaboration with peers. We are curious to ascertain whether the increased reliance on AI models for rapid assistance may limit collaboration or, conversely, provide an opportunity for students to focus and engage in more innovative tasks. By comprehensively examining these aspects, we aspire to contribute valuable insights to the ongoing discourse surrounding AI-based tools like ChatGPT and their implications for the education system and student learning experiences.

## Acknowledgment

This work was funded by The Scientific and Technological Research Council of Turkey, under 1515 Frontier R&D Laboratories Support Program with project no: 5169902.

## References

[1] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, "Zero-shot text-to-image generation," in *International Conference on Machine Learning*. PMLR, 2021, pp. 8821–8831.

[2] B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, "Dreamfusion: Text-to-3d using 2d diffusion," *arXiv preprint arXiv:2209.14988*, 2022.

[3] J.-B. Alayrac, J. Donahue, P. Luc, A. Miech, I. Barr, Y. Hasson, K. Lenc, A. Mensch, K. Millican, M. Reynolds *et al.*, "Flamingo: a visual language model for few-shot learning," *Advances in Neural Information Processing Systems*, vol. 35, pp. 23 716–23 736, 2022.

[4] R. Villegas, M. Babaeizadeh, P.-J. Kindermans, H. Moraldo, H. Zhang, M. T. Saffar, S. Castro, J. Kunze, and D. Erhan, "Phenaki: Variable length video generation from open domain textual description," *arXiv preprint arXiv:2210.02399*, 2022.

[5] Z. Borsos, R. Marinier, D. Vincent, E. Kharitonov, O. Pietquin, M. Sharifi, D. Roblek, O. Teboul, D. Grangier, M. Tagliasacchi *et al.*, "Audiolm: a language modeling approach to audio generation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2023.

[6] A. S. George and A. H. George, "A review of chatgpt ai's impact on several business sectors," *Partners Universal International Innovation Journal*, vol. 1, no. 1, pp. 9–23, 2023.

[7] A.-C. Haensch, S. Ball, M. Herklotz, and F. Kreuter, "Seeing chatgpt through students' eyes: An analysis of tiktok data," *arXiv preprint arXiv:2303.05349*, 2023.

[8] D. Baidoo-Anu and L. Owusu Ansah, "Education in the era of generative artificial intelligence (ai): Understanding the potential benefits of chatgpt in promoting teaching and learning," *Available at SSRN 4337484*, 2023.

[9] C. K. Lo, "What is the impact of chatgpt on education? a rapid review of the literature," *Education Sciences*, vol. 13, no. 4, p. 410, 2023.

[10] J. Rudolph, S. Tan, and S. Tan, "Chatgpt: Bullshit spewer or the end of traditional assessments in higher education?" *Journal of Applied Learning and Teaching*, vol. 6, no. 1, 2023.

[11] E. Felten, M. Raj, and R. Seamans, "How will language modelers like chatgpt affect occupations and industries?" *arXiv preprint arXiv:2303.01157*, 2023.

[12] R. Gozalo-Brizuela and E. C. Garrido-Merchan, "Chatgpt is not all you need. a state of the art review of large generative ai models," *arXiv preprint arXiv:2301.04655*, 2023.

[13] M. Fraiwan and N. Khasawneh, "A review of chatgpt applications in education, marketing, software engineering, and healthcare: Benefits, drawbacks, and research directions," *arXiv preprint arXiv:2305.00237*, 2023.

[14] K. Fuchs, "Exploring the opportunities and challenges of nlp models in higher education: is chat gpt a blessing or a curse?" in *Frontiers in Education*, vol. 8. Frontiers, 2023, p. 1166682.

[15] B. D. Lund and T. Wang, "Chatting about chatgpt: how may ai and gpt impact academia and libraries?" *Library Hi Tech News*, vol. 40, no. 3, pp. 26–29, 2023.

[16] S. Sok and K. Heng, "Chatgpt for education and research: A review of benefits and risks," *Available at SSRN 4378735*, 2023.

[17] M. M. Rahman, H. J. Terano, M. N. Rahman, A. Salamzadeh, and M. S. Rahaman, "Chatgpt and academic research: a review and recommendations based on practical examples," *Rahman, M., Terano, HJR, Rahman, N., Salamzadeh, A., Rahaman, S.(2023). ChatGPT and Academic Research: A Review and Recommendations Based on Practical Examples. Journal of Education, Management and Development Studies*, vol. 3, no. 1, pp. 1–12, 2023.

[18] E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günnemann, E. Hüllermeier *et al.*, "Chatgpt for good? on opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, p. 102274, 2023.

[19] M. A. AlAfnan, S. Dishari, M. Jovic, and K. Lomidze, "Chatgpt as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses," *Journal of Artificial Intelligence and Technology*, vol. 3, no. 2, pp. 60–68, 2023.

[20] N. Forman, J. Udvaros, and M. S. Avornicului, "Chatgpt: A new study tool shaping the future for high school students," *International Journal of Advanced Natural Sciences and Engineering Researches*, vol. 7, no. 4, pp. 95–102, 2023.

[21] S. Wagholikar, A. Chandani, R. Atiq, M. Pathak, and O. Wagholikar, "Chatgpt-boon or bane: A study from students perspective," in *2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)*. IEEE, 2023, pp. 207–212.

[22] K. I. Roumeliotis and N. D. Tselikas, "Chatgpt and open-ai models: A preliminary review," *Future Internet*, vol. 15, no. 6, p. 192, 2023.

[23] M. Dowling and B. Lucey, "Chatgpt for (finance) research: The bananarama conjecture," *Finance Research Letters*, vol. 53, p. 103662, 2023.

[24] N. Grant and C. Metz, "A new chat bot is a 'code red'for google's search business," *New York Times*, no. Dec 21, 2022, 2022.

[25] J. Deng and Y. Lin, "The benefits and challenges of chatgpt: An overview," *Frontiers in Computing and Intelligent Systems*, vol. 2, no. 2, pp. 81–83, 2022.

[26] P. P. Ray, "Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope," *Internet of Things and Cyber-Physical Systems*, 2023.

[27] Y. Sun, S. Wang, S. Feng, S. Ding, C. Pang, J. Shang, J. Liu, X. Chen, Y. Zhao, Y. Lu *et al.*, "Ernie 3.0: Large-scale knowledge enhanced pre-training for language understanding and generation," *arXiv preprint arXiv:2107.02137*, 2021.

[28] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar *et al.*, "Llama: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.

[29] C. Zhou, Q. Li, C. Li, J. Yu, Y. Liu, G. Wang, K. Zhang, C. Ji, Q. Yan, L. He *et al.*, "A comprehensive survey on pretrained foundation models: A history from bert to chatgpt," *arXiv preprint arXiv:2302.09419*, 2023.