International Journal of Science and Research Archive

eISSN: 2582-8185 CODEN (USA): IJSRO2

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Review Article)

A study on privacy of AI chatbots and session hijack

Krithiga B [1, *] and Marikkannan M [2]

[1] Computer Science and Engineering (PG Scholar), Department of Computer Science and Engineering, Government College of Engineering, Erode, India.

[2] Department of CSE, Government College of Engineering, Erode, India.

International Journal of Science and Research Archive, 2025, 15(03), 046–054

Publication history: Received on 21 April 2025; revised on 29 May 2025; accepted on 01 June 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.3.1671

Abstract

The paper revolves around our daily usage of AI(s) and the rise of session hijack attackers on the web. Whatever people surf on the web is potentially vulnerable to attackers because the network is a web of connections, and the attackers are always finding a way to attack users and steal their data. Nowadays, using AI has become a part of daily routine regardless of field, profession, or usage. Most people use AI on the web since it is the most convenient way of using it without installing it on the device. However, most people don't know that using something on the web may lead to session hijacking. So, people want to find a way to do their daily routines with AI without becoming victims of session hijack. As someone who codes from time to time, people may have some knowledge of the Integrated Development Environment (IDE) and environment data. Using something in an IDE has no connection to the internet by any means, even while connected to an internet service, because an IDE is an isolated environment built into the system for coding purposes while also allowing us to connect with the network. This basically works by avoiding things like spiders, cookies from the web, which will be stored if a user uses the browser, but this has no effect on IDEs. Running AI on localhost is also a possibility, but it demands high specifications on the hardware, such as a large amount of RAM, VRAM and external GPU.

Keywords: AI Chatbots; Privacy Concerns; LLMs; Training and Testing; Generative AI; Natural Language Processing

1. Introduction

The rise of large artificial intelligence (AI) models like ChatGPT, AI-generated content (AIGC) is capturing more attention and driving a significant shift in how users create content and represent knowledge. AIGC leverages advanced generative AI algorithms to help or even take over the task of producing vast amounts of high-quality, human-like content quickly and affordably, all based on prompts provided by users. However, despite the impressive strides made in AIGC recently, there are still important issues around security, privacy, ethics, and legality that need to be tackled. This paper offers a comprehensive look at the underlying principles, security and privacy risks, cutting-edge solutions, and future challenges within the AIGC landscape [1]. Viewers review the latest watermarking techniques for AIGC to ensure regulation of both the AIGC models and the content they generate.

The rapid development of AI technologies presents both opportunities and challenges. On the one hand, AI offers significant benefits, such as improved prediction, optimization of operations, and personalized solutions across various industries. On the other hand, the extensive data collection required for AI functionalities introduces substantial risks, including potential misuse of biometric data and other personal information. These concerns are particularly pronounced in applications like emotion recognition, biometric categorization, and remote biometric identification, which necessitate stringent regulatory oversight to prevent abuse and ensure compliance with legal standards [2].

[*] Corresponding author: Krithiga B

The incorporation of Artificial Intelligence (AI) into SE has brought about a significant period of change, transforming the software development cycle with remarkable progress and effectiveness in recent years. This merging of AI with SE signifies a combination of computational ability and inventive issue resolution, altering conventional methods and driving the domain towards advancement and distinction. Across various stages, from initial planning and requirement collection to deployment and maintenance, AI has been instrumental in enhancing efficiency, improving resource management, and tackling intricate issues inherent in SE operations. Initially, merging AI's precise algorithms with the creative problem-solving approach of SE was seen to boost progress and effectiveness. By leveraging AI's capabilities in data analysis, recognizing patterns, and making autonomous decisions, software engineers have gained the ability to address complex design issues, automate routine tasks, and improve overall software quality. This paper's systematic review explores how AI integration transformed SE from 2013 to 2023 [2].

Recent advancements in AI and NLP (natural language processing) have garnered attention, leading to active research in the field of LLM (large language models) related to chatbot technology. As chatbots and AI technologies continue to evolve, their performance has been steadily improving. With the evolution of them, which are trained on massive datasets using LLMs, they have reached a level where they can provide responses at a similar level to that of humans [6].

Artificial intelligence-generated content (AIGC) refers to the use of generative AI algorithms to assist or replace humans in creating rich personalized and high-quality content at a faster pace and lower cost, based on user inputs or requirements. By January 2023, nearly 13 million users were interacting with ChatGPT daily **Error! Reference source not found.**. ChatGPT is a variant of the generative pre-training transformer (GPT), a transformer-based large language model (LLM) that can understand human languages and create human-like text [1]. Many reusable solutions have been proposed to tackle various challenges in designing FM-based systems. However, there is a lack of systematic guidance on the architecture design of FM-based systems. The impact of integrating FMs into software architecture is not fully studied yet. Additionally, the FM's growing capabilities can eventually absorb the other components of AI systems, introducing the moving boundary and interface evolution challenges in architecture design [8].

2. Literature Review

The right to privacy has deep legal, moral, and historical roots in American society. It is considered a fundamental right protected by the Constitution of the United States, although it is not explicitly mentioned in the document itself. Privacy serves essential human needs by creating zones for individual liberty, autonomy, seclusion, and self-definition, which includes the exercise of free expression, family life, intimacy, and other personal relationships. Moreover, it protects marginalized or vulnerable individuals and groups, safeguards democratic values, and maintains the integrity of democratic institutions and processes, including elections [2].

Since the 1990s, chatbots have been gaining space in the market and after the 2000s, especially after 2016, there was an even faster growth of interest on the subject [3]. This growth, consequently, brought new challenges such as how to design conversations and manage chatbot content. Concerning chatbot design, scalability and usability can be major issues since they have a direct impact on the user experience of a chatbot [3]. Back in 1950, Alan Turing posed a thought-provoking question that would become later pivotal in the evolution of chatbots. He raised the question of whether a computer program could engage in communication with a group of individuals without them being aware that they were conversing with an artificial entity. This inquiry, known as the Turing test, served as a seminal idea that laid the foundation for the advancement of chatbot technology [4].

The first chatbot named ELIZA appeared in 1966 and it simulated the role of a therapist, although its ability to interact with users was restricted. Irrespective of its limitations, the Turing test served as a great source of inspiration for the subsequent development of other agents. One notable example is ELIZA, which utilizes pattern matching and a pattern-based response selection scheme [4]. The rise of artificial intelligence, particularly the emergence of large language models (LLMs) like ChatGPT, continuously reveals numerous advantages across various domains. However, the area of project management has not yet been sufficiently explored. This study fills the research gap by conducting an empirical evaluation of three well-known LLMs: OpenAI's ChatGPT-3.5 and ChatGPT-4, as well as Google's Gemini [7].

More than 50 years have passed since the development of the first chatbot. Weizenbaum created the first chatbot, ELIZA, to simulate a psychotherapist speaking with a real patient using a pattern-matching process. The mid-1990s witnessed the development of another prominent chatbot called ALICE. It used knowledge records and artificial intelligence markup language (AIML) to determine an appropriate response to user input. The first program to pass the Turing Test was the natural language program Parry, created by Stanford University psychiatrist Kenneth Mark Colby in 1972. It was not until then that chatterbot technology began to take off. Parry, portraying a schizophrenia patient, was seen to be more educated than ELIZA. It had a ‘‘personality’’ and a better controlling structure that determined replies based on an assumption-based framework and ‘‘emotional responses’’ that were triggered by changes in a user’s utterances. However, Parry was still viewed as a chatbot with limited capabilities that could not pick up new information from conversations. However, with the development of Artificial Intelligence (AI) technologies, innovative chatbots such as ChatGPT were developed, which are now being used in various sectors, including business, retail, healthcare, education, etc [17].

3. Technological Overview

The increasing complexity and diverse performance requirements of modern analog systems create a high dimensional design space. In response, full-flow automation has become necessary to handle the intricate trade-offs between numerous performance parameters, as traditional approaches are time-consuming and heavily reliant on expert knowledge. While digital design automation has seen extensive development and adoption in both industry and academia [5].

Unlike traditional programming, ML enables algorithms to learn from data to perform tasks. Data, regardless of complexity, is ultimately represented as tensors: text and audio as vectors, images as matrices, and videos as sequences of images [5]. The CNN model is well-suited for software fault prediction because it can effectively capture local and global patterns within source code structures, aiding in identifying potential defects. The author improved a CNN model for within-project defect prediction and examined it with CNN and empirical results [6].

In [6], they contributed to the field of software defect prediction by utilizing various ML approaches to create multiple categorization or classification models, aiming to improve software quality and reduce testing costs.

For instance, ChatGPT was found to exhibit skills that allow it to perform at par with humans in algorithmic tasks, attend multiple questions in a query and answer them, create good summaries of text, code very well, is more ethical and truthful than previous models but also is worse than potential single-task fine-tuned models, might provide different results for the same prompt depending on its version, may underperform in underrepresented languages and sometimes consider only utilitarian morality to ethical dilemmas [7].

The release of ChatGPT, Gemini, and other large language model (LLM)-based chatbots has drawn huge attention on foundations models (FMs) worldwide. FMs are massive AI models that are pretrained on vast amounts of broad data and can be adapted to perform a wide variety of tasks.1 With numerous projects already underway to explore their potential, it is widely predicted that FMs will serve as the fundamental building blocks for most future AI and artificial generative intelligence (AGI) systems [8].

4. System Architecture

Let’s understand the basic working of a simple AI chatbot. People, as users, will send their queries and input through our system to the Large Language Model (LLM) (Figure 1). The LLM contains multiple entities, namely the Internal Processing Unit (IPU) and Natural Language Processing (NLP). The role of the IPU is to process the given queries so that the NLP can understand what the user wants or needs. This is done through tokenization and keyword extraction by the IPU, which acts as a connection between input (user's) and output (system's). Tokenization means creating tokens for each keyword generated, and a keyword is a specific word that holds the meaning of the sentence.

For a simple explanation, let us consider an example:

- The user sends “What colour is mango?”
- The AI’s IPU doesn’t take the entire sentence but instead only takes the keywords from it. In this example, “colour” and “mango” are the keywords, and that is enough for AI models to understand and interpret what the user is supposedly saying or asking.
- With those keywords processed by the NLP, it gets the meaning and result, then sends those results back to the user as “yellow.” The output is not sent vaguely as just “yellow,” but in a proper sentence format like “The colour of mango is yellow.”

NLP is the core entity of any LLM because NLP Figure (1) is the unit that understands the extracted keywords and fetches the appropriate results. NLP does this by using a whole database of collected and trained data. Every NLP will be trained with a set of datasets for the AI to understand what the user is asking. So, proper training will be undergone for every AI model with its appropriate dataset in order to provide accurate information. This training process is indeed an important phase while building an AI because only through training can the AI model gain the capabilities such as understanding what the user is saying and what the appropriate thing is to fetch from the dataset.

The dataset typically consists of many variables like words, colors, shapes, objects, etc. Everything inside the dataset is linked as to how a human would perceive the world with their understanding so that the model being trained can reach the potential that the creators wanted it to achieve. Training an AI model also comes with its own complications, such as the dataset not being enough or having errors/inaccurate data. A good dataset in the database means a well-trained AI model as the reward. It is also important that the model knows how to link and think of related things for an input rather than just using narrow, streamlined thinking. This gives the AI the potential to be a compatible assistant for the user.

The understanding of how AI works is very simple yet complex at the same time, as if it is like two sides of a coin because it comes with its own complications and rewards, as everyone already knows. Creating an entire chatbot from scratch takes an enormous number of resources and time for training. But there are several ways that someone can use a pre-trained AI model for their own purposes, such as in this project. The basic representation of LLM is given below in Figure (1).

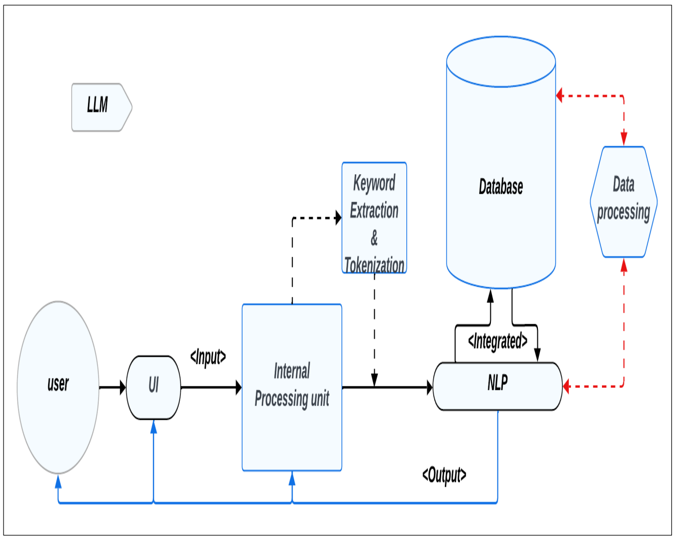

Figure 1 LLM Process

However, not everyone here is building an entire bot from the very beginning. Instead, we are now going to borrow an existing AI model and import it into our Command Line Interface (CLI). The AI model that this paper is going to import is Google's own AI – GEMINI, which was previously known as BARD, and the CLI currently being used is the simplest yet finest CLI – Shell of VS Code. There is a specific reason why people are using VS Code's Shell. It is because of session hijack, and to avoid session hijack while people are conversing with the chatbot. Shell is actually a free isolated space that still has a connection to the internet. It is basically a free safeguard against session hijacks as long as users are inside its isolated space. Although the process is just importing an existing AI model, it still follows the same architectural patterns and functions as a typical AI model on the web and works with the same LLM.

5. Applications

ChatGPT emerged as a significant milestone in chatbot introduction based on the language model GPT. It represents a highly advanced language model trained by OpenAI, designed to assist with various tasks such as generating human-like responses, answering questions, and providing information on a wide range of topics. A few months later, Gemini (previously known as Bard), a chatbot developed by Google based on the Large Language Model (LLM) emerged and is now gaining traction due to its speed and ability to respond to questions in a human-like manner by accessing up-to-date information from the Internet. It employs generative AI for natural conversations across various modalities, including text, voice, and images [4].

Recent advancements in AI and NLP (natural language processing) have garnered attention, leading to active research in the field of LLM (large language models) related to chatbot technology. As chatbots and AI technologies continue to evolve, their performance has been steadily improving. With the evolution of them, which are trained on massive datasets using LLMs, they have reached a level where they can provide responses at a similar level to that of humans [6].

In general, LLM can be associated with the GPT series models. GPT stands for 'generative pre-trained transformer', representing AI models that are pretrained on extensive data through machine learning to generate sentences. In particular, in the recent revolution of Generative AI, ChatGPT has gained prominence for its ability to engage in human-like conversations. ChatGPT can formulate responses to questions in a manner resembling human sentence construction [6].

High-visibility engineering projects often fail due to a lack of proper project management and tools. The Berlin Brandenburg airport started construction in 2006, with a delivery day in 2011, but was only delivered with a delay of 9 years in 2020, an overrun of more than €6.2 bn compared to the original budget, due to problems such as lack of a proper information management system, poor construction planning, execution, management, and corruption. Countless other IT project failures exist due to management reasons, among other factors. While such project failures may result in financial and reputation-loss risks, sometimes they also result in fatal accidents, as recently the two Boeing 737-MAX crashes were linked to the design problems in sensors, technical components, and practice failures [7].

In April 2021, the European Commission introduced the first EU regulatory framework for AI with the aim of overseeing the development and utilization of this groundbreaking technology. The AI Act is part of a broader EU strategy designed to enhance Europe's potential to compete globally in regulating the digital sector. The Commission has been tasked with aiding the co-legislators in concluding the inter-institutional negotiations, commonly known as the "trilogue". Noteworthy negotiation phases include the proposal of a compromise text on the AI Act by the Council of the EU in November 2021. Note that the version approved by the European Parliament is the legal text that was used during the Grand Challenge. Once approved, this regulation could be the world's first legislation governing AI. The AI Act proposal delineates [9].

Through our analysis, it is evident that ChatGPT has shown potential value for various educational applications. According to Dowling and Lucey, ChatGPT can generate impressive, and believable research papers for well-ranked journals by adding data and results analysis. Besides, ChatGPT's language translation capabilities, text summarizing, and QAs make it a powerful tool that can help users with a range of tasks and also provide new opportunities to develop skills in students. Cotton proves that ChatGPT has potential benefits for higher education, such as increasing student engagement and accessibility. For example, students who received personalized and adaptive feedback from AI-based learning platforms showed higher engagement and better performance, particularly those with low prior knowledge and at risk of dropping out ChatGPT also benefits educators. It was stated that machine-generated questions are similar to human-generated questions and can be used in final exams [11].

Natural Language Processing (NLP) for short, is all about how machines and human language interact. It focuses on teaching machines to analyze and interpret vast amounts of raw human language input. Essentially, NLP is a branch of computer science that aims to create systems capable of 'understanding' the content in documents, including the subtle nuances of language used in different contexts. Once a machine can grasp this, it can accurately extract knowledge and insights from those documents. Plus, it can even categorize and organize them on its own [13]. Through NLP, computers can pick up on sentiment and intent, not just the literal meaning of words. This means they can learn and adapt over time, remembering context and information to improve their future responses. For businesses, leveraging NLP can be a game-changer. Without it, they might be missing out on valuable opportunities to automate processes or gain insights about their operations, which could hinder their growth and revenue potential. NLP also plays a crucial role in enabling chatbots to understand and engage in human conversations effectively [13].

Human personality can be understood as a unique behavioural characteristic of individuals reflecting the adjustments in attitudes, interests, traits, and emotional patterns. Although there are various taxonomies for categorizing human personalities, the Big Five Factor model is one of the most used taxonomies in psychology studies. These three traits include: (a) Openness: this dimension is used to measure the imaginative, inventiveness, and creative capabilities of the individuals, reflecting curiosity to learn new things and enjoy new experiences; (b) Conscientiousness: this dimension is used to measure the efficiency of individuals and their organization, reflecting a goal-directed and organized behaviour; (c) Extroversion: this dimension is used to measure an individual's sociability, assertiveness, and emotional expressiveness, reflecting an attitude with a concentration of interest on external objects [17].

Systematic mapping studies offer a comprehensive overview of a particular research domain through classification. These studies involve an in-depth exploration of existing literature to assess the coverage of multiple topics, publication frequency, research trends, and relevant publication venues. In the present study, the systematic mapping process adheres primarily to the guidelines proposed by. Following these guidelines specific to SE, the essential steps in the systematic mapping study encompassed defining research questions, searching for pertinent papers, screening the papers, keywording abstracts, extracting data, and mapping. Brereton, Kitchenham, and Bugden are well-known researchers in SLRs on SE introduces a template for creating a case study protocol in SE, aiming to enhance the rigor of case studies [2].

6. Issues in Privacy AI Chatbot

The right to privacy in the age of artificial intelligence (AI) is a pivotal and complex issue that intersects with technology, law, and ethics. As AI continues to evolve and become integral to various sectors—including healthcare, finance, and law enforcement—the privacy of individuals is increasingly at risk. AI systems often require vast amounts of personal data to function effectively, which can reveal intimate details about individuals' lives and potentially lead to privacy invasions if misused or inadequately protected. This has raised significant concerns about mass surveillance, the erosion of privacy rights, and the ethical use of AI [2].

Artificial Intelligence (AI) has become a revolutionary force with deep consequences for various parts of society in the era of rapid technological breakthroughs. Our ability to work, communicate, and engage with the world around us has been completely transformed by the integration of AI technology across a wide range of industries. Unfortunately, worries about data security and privacy have gained more traction as AI develops and permeates more areas of our lives [1].

This lack of data availability poses challenges for replicating and independently validating the findings. Another critical limitation is the absence of a “family of experiments” across the studies, which means that many experiments lack replication and consistent design. This inconsistency ultimately weakens statistical power and reduces the generalizability of the results. Additionally, small sample sizes are common, with many experiments involving fewer than 30 participants, most of whom are students, which restricts the external validity of the findings. Moreover, several studies struggle with unclear experimental design, featuring inconsistencies in task structures, user profiles, and testing environments. Many experiments also overlook external variables, such as user experience or prior knowledge, which could impact the results. Natural language processing (NLP) continues to be a significant hurdle, as users frequently report issues with comprehension, limited personalization, and challenges in handling accents or complex queries. The study also points out that usability evaluations often prioritize user satisfaction over measures of effectiveness or efficiency, which limits a thorough assessment of chatbot performance [14].

The proposal for AI regulation is a key component of the broader EU strategy aimed at bolstering Europe’s ability to compete globally in overseeing the digital sector and other impactful domains associated with AI. In this context, some primary legal challenges arise from potential coordination issues with existing and forthcoming regulations, particularly those related to the protection and governance of both personal and non-personal data, as well as competition law and consumer law. Addressing challenges pertaining to the AI Act’s content, a notable concern arose from the expansive definition of AI systems. The rationale behind adopting such a comprehensive definition lies in prioritizing the “use case” over evaluating the specific technology behind the system. Essentially, the AI Act adopts a broad perspective on what qualifies as an AI system but adopts a more focused stance when considering its application. This approach aligns with the risk-based approach adopted by the AI Act [9].

The challenge of Data Privacy and Ethics for ChatGPT is complex and multifaceted. One aspect of this challenge is related to data privacy, which involves protecting personal information collected by ChatGPT. ChatGPT relies on vast amounts of data to train its language model, which often includes sensitive user information, such as chat logs and personal details. Therefore, ensuring that user data is kept private and secure is essential to maintain user trust in the technology. Another aspect of the Data Privacy and Ethics challenge for ChatGPT is related to ethical considerations. ChatGPT has the potential to be used in various applications, including social media, online communication, and customer service. However, the technology’s capabilities also pose ethical concerns, particularly in areas such as spreading false information and manipulating individuals. The potential for ChatGPT to be used maliciously highlights the need for ethical considerations in its development and deployment [11].

The constraints set by course policies and the German, General Data Protection Regulation (GDPR) likely played a role in the lower participation rate. As a result, the sample size was quite small, with only 46 students, and just 30 of them took part in the survey, leading to a limited number of responses. Since participation was voluntary, this may have introduced a selection bias; those who opted in might have been more inclined towards the technology, which could affect how broadly user can apply the results. Voluntary participation is a common hurdle, yet it poses a significant limitation when trying to draw wide-ranging conclusions. Additionally, students who have already completed the course are likely to have a deeper understanding of the material and terminology, allowing them to ask more specific and focused questions. On the other hand, students who haven't finished the course might struggle with the right terminology, which could lead to more general and vague prompts, resulting in less informative responses. These differences have important implications for the design and effectiveness of educational chatbots, particularly in terms of providing adaptive responses that recognize the learner's stage. When it comes to feedback, it's worth noting that a single TA handles the manual evaluation, which could lead to inaccuracies in feedback for more complex responses [12].

Several limitations and future research directions have been identified. First, the users' interactions with the chatbots were bound to a few tasks, potentially limiting the possibility of participants forming a comprehensive perception of the chatbot's personality. Perhaps a longitudinal study supporting a more extensive set of tasks could allow participants to form a more reliable perception of the chatbot's personality. Second, this research has focused on the context of academic advising, where the chatbot is considered a representative of a public institution. Future researchers should consider that an individual's behaviour and interactions may vary depending on the context. Further, research has shown that users' behaviour can be influenced by a public context [12].

Challenges in privacy ai chatbots are listed below**,**

- Vast number ideas but very little technology at the start.
- Session hijack protection can't be given to each and every individual.
- So many ideas but only few ideas are currently become a practical on use case.
- Weaker AI policies and loopholes on policy that is not currently noticed.
- Not enough awareness on AI's privacy and data gathering methods.
- Not enough transparency on data collection from users from the company.
- No clear directions or paths on how they train models using our data.
- LLMs has huge amount of data but how did they gather that much of public data.
- There are several ways to steal user's data online apart from the chatbot and AI.
- Many people are scarifying their own privacy for the so-called 'trends'.

7. Emerging Technologies

Artificial Intelligence (AI) is often seen as a cutting-edge technology, which obviously it is and for a good reason it's evolving quickly, making a significant impact, and presenting a mix of challenges and opportunities. One clear sign of this is the remarkable surge in AI research and publications over the last ten years, with fresh discoveries, techniques, and models popping up at an astonishing pace. New areas like generative AI, explainable AI, and neuromorphic computing are emerging, highlighting that this technology is still venturing into new territories. Moreover, there's a noticeable gap in skilled professionals within the AI field, which suggests that people haven't yet hit a saturation point. Key technical hurdles, such as achieving Artificial General Intelligence (AGI), tackling algorithmic bias, and enhancing energy efficiency, remain unresolved, indicating that AI is still in a state of growth. The full effects of AI on jobs, education, creativity, and society are just starting to emerge, with anticipated disruptions in areas like law, governance, and defence. The thriving AI startup scene and the uptick in venture capital investment further illustrate that innovators and investors view this technology as a frontier. Lastly, the recent surge in public awareness and the integration of AI terms and tools into everyday life show that in a transitional phase AI is expanding rapidly but hasn't yet been fully embraced or understood. All these elements together affirm that AI is more than just a passing trend, it's a transformative force that's still unfolding and shaping our future.

AI tools like ChatGPT and other similar platforms are quickly becoming a go-to resource for many people, thanks to the immediate and practical benefits they bring to our daily lives. Millions rely on these tools every day to save time, enhance accuracy, and boost productivity whether it's drafting emails, writing code, designing graphics, or getting quick assistance with schoolwork. Major tech companies like OpenAI, Google, Microsoft, and Meta are behind these innovations, and they're being seamlessly integrated into essential applications like Microsoft Office, Google Docs, and various customer service platforms. The popularity of these tools is evident, with ChatGPT hitting 100 million users just months after its launch, making it one of the fastest-growing apps ever. Businesses are incorporating AI into their workflows to automate repetitive tasks, enhance customer experiences, and analyze large datasets with ease. In the education sector, both students and teachers are leveraging AI for personalized learning, homework assistance, and content creation. In software development, AI coding assistants are revolutionizing how developers write, test, and debug their code. The surge in startups focused on AI, the growing job market in AI-related fields, and the billions of dollars being invested globally all highlight that these tools are more than just a passing trend they're becoming a fundamental part of our digital landscape.

Chatbots have become incredibly popular these days, and it’s easy to see why. They provide a quick, efficient, and budget-friendly way to engage with users and automate services across a range of industries. One of the biggest draws is their ability to work around the clock, offering instant support and consistent answers no matter the time of day. This is particularly beneficial for businesses that cater to a global audience or deal with a high volume of inquiries. By taking care of repetitive tasks like answering common questions, scheduling appointments, and handling simple requests, chatbots significantly lessen the need for large customer service teams, which helps cut down on operational costs. Plus, they boost user engagement by facilitating interactive, personalized conversations through natural language processing (NLP), making interactions feel more genuine and human-like. They’re also capable of supporting multiple languages and can easily scale to accommodate thousands of users at once, making them incredibly versatile. On top of that, chatbots are excellent for gathering and analyzing user data, providing organizations with valuable insights into customer preferences and behaviours. Nowadays, you can find chatbots in various sectors, including healthcare, education, banking, e-commerce, and even within businesses, where they assist with everything from checking symptoms to offering product recommendations and HR support. Their ability to integrate seamlessly with other technologies like CRMs, payment systems, and smart devices only adds to their value.

8. Conclusions and Future Direction

The vast majority of people use AI to aid them in work, complete their tasks in less time than usual yet effectively, and use it as a personal assistant. 7 out of 10 people use AI in their daily life tasks regardless of their field of work. But a very small number of people really care about their privacy compared to the vast majority. From the company side, the ones who make these AI models, there isn't enough transparency on how they process user data, and thus there is a huge gap in understanding privacy and user data. The proposal is a small step toward securing our privacy by taking action on our own, because, in all honesty, the companies won't put their user data collection process in public, as it would ruin their business model. So, it's time to spread the word and raise awareness about the privacy concerns while using AI models.

The small step and the idea is having the AI inside a virtual space or isolated space, because having something in this type of space has no effect on the outside environment while still being connected to the internet and able to function as intended. This idea is as simple as using VS Code’s Shell and virtual environment variables during development. Being inside a virtual or isolated space allows users to do whatever they want without restrictions or performance loss, and at the same time, users are securing their privacy within the virtual space itself instead of going out on the web. This also implies that there will be some manual steps to follow, which are obviously absent while using a chatbot on a web platform. In the end, people are using the same AI chatbot, getting the same results, but with some sacrifices such as a basic UI and manual actions. But this is the current cost to pay if someone is truly concerned about their user data and privacy, which should be a basic concern for everyone who uses any new and emerging technologies.

As people now know that there is a potential ‘Session Hijack’ vulnerability in chatbots, users must take steps to safeguard themselves from attackers. By using this method, people can safely stay away from all attacks while also enjoying a safe space with AI. This also opens room for many more future enhancements, such as a scalable isolated environment and diffusion models inside the shell. As already seen, the user needs to import and use an existing AI model from Google, but what the paper uses is Gemini’s LLM only.

Gemini is a powerful AI model that is also trained as a diffusion model as well. So, it is really possible to bring that diffusion model into the shell just like how the LLM was imported into the shell. But diffusion models have a whole other level of complexity issues within an isolated or virtual space, but it can be tackled, and this can be made possible in the future. Users can also try to create their own server and then import Gemini into that server instead of into the shell. In this way, users have absolute control and transparency over all data, and all privacy is kept safe and secure with no one able to access it. This also gives off a clean and appealing interface, which the current version of Shell lacks in terms of visuals and experience.

Today’s AI chatbots have an appealing visual interface with dynamic or 3D elements throughout the User Interface (UI), which also elevates User Experience (UX). But while inside the shell, the creator can’t provide any sort of UI enhancements for the sake of our own privacy. However, with a personal server, users can use the AI’s original UI or configure and create a new interface that fits their preferences and make it customizable, so that people can change the UI anytime they want based on their mood. Keeping a diffusion model in the field of privacy without sacrificing a whole lot of quality in the output is a bit tricky. For that, there is definitely a need for an external GPU or VRAM to run the diffusion model without any issues.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Y. Wang, Y. Pan, M. Yan, Z. Su and T. H. Luan. A Survey on ChatGPT: AI–Generated Contents, Challenges, and Solutions. IEEE Open Journal of the Computer Society. 2023; 4:280-302.

[2] U. K. Durrani, M. Akpinar, M. Fatih Adak, A. Talha Kabakus, M. Maruf Öztürk and M. Saleh. A Decade of Progress: A Systematic Literature Review on the Integration of AI in Software Engineering Phases and Activities (2013-2023). IEEE Access. 2024; 12:171185-171204.

[3] G. A. Santos, G. G. de Andrade, G. R. S. Silva, F. C. M. Duarte, J. P. J. D. Costa and R. T. de Sousa. A Conversation-Driven Approach for Chatbot Management. IEEE Access.2022; 10:8474-8486.

[4] L. Benaddi, C. Ouaddi, A. Jakimi and B. Ouchao. A Systematic Review of Chatbots: Classification, Development, and Their Impact on Tourism. IEEE Access. 2024; 12:78799-78810.

[5] D. Noori Zadeh and M. B. Elamien. Generative AI for Analog Integrated Circuit Design: Methodologies and Applications. IEEE Access. 2025; 13:58043-58059.

[6] D. Park, G. -t. An, C. Kamyod and C. G. Kim. A Study on Performance Improvement of Prompt Engineering for Generative AI with a Large Language Model. in Journal of Web Engineering. 2023; 22(8):1187-1206.

[7] S. Karnouskos. The Relevance of Large Language Models for Project Management. in IEEE Open Journal of the Industrial Electronics Society. 2024; 5:758-768.

[8] Q. Lu, L. Zhu, X. Xu, Z. Xing and J. Whittle. Toward Responsible AI in the Era of Generative AI: A Reference Architecture for Designing Foundation Model-Based Systems. in IEEE Software. 2024; 41(6): 91-100.

[9] T. Scantamburlo et al. Software Systems Compliance with the AI Act: Lessons Learned from an International Challenge. in IEEE/ACM International Workshop on Responsible AI Engineering (RAIE), Lisbon, Portugal.2024; 44-51.

[10] A. Koubaa, W. Boulila, L. Ghouti, A. Alzahem and S. Latif. Exploring ChatGPT Capabilities and Limitations: A Survey. in IEEE Access, 2023; 11:118698-118721.

[11] O. Mubin, F. Alnajjar, Z. Trabelsi, L. Ali, M. M. A. Parambil and Z. Zou. Tracking ChatGPT Research: Insights from the Literature and the Web. in IEEE Access, 2024; 12:30518-30532.

[12] A. T. Neumann, Y. Yin, S. Sowe, S. Decker and M. Jarke. An LLM-Driven Chatbot in Higher Education for Databases and Information Systems. in IEEE Transactions on Education. 2025; 68(1):103-116.

[13] U. Sehgal and S. Bhardwaj. Building a Chatbot using Natural Language Processing. Second International Conference on Informatics (ICI), Noida, India, in IEEE Access. 2023; 1-5.

[14] R. Ren, M. Zapata, J. W. Castro, O. Dieste and S. T. Acuña. Experimentation for Chatbot Usability Evaluation: A Secondary Study. in IEEE Access. 2022; 10:12430-12464.

[15] P. Haindl and G. Weinberger. Students Experiences of Using ChatGPT in an Undergraduate Programming Course. in IEEE Access. 2024; 12: 43519-43529.

[16] S. Yu, F. Carroll and B. L. Bentley. Insights Into Privacy Protection Research in AI. in IEEE Access. 2024; 12: 41704-41726.

[17] M. Amin Kuhail, M. Bahja, O. Al-Shamaileh, J. Thomas, A. Alkazemi and J. Negreiros. Assessing the Impact of Chatbot-Human Personality Congruence on User Behavior: A Chatbot-Based Advising System Case. in IEEE Access, 2024; 12: 71761-71782.