Proceedings of the 8th International Conference on Economic Management and Green Development
DOI: 10.54254/2754-1169/122/2024.17803

Research on the Privacy Risks Associated with ChatGPT and Its Business Enablement Strategies in the Chinese Context

Kun Liu[1,a,*]

[1]*Bachelor of Arts in English, Northeastern University at Qinhuangdao, Hebei, 066004, China*
a. liu_kunwy163@163.com
**corresponding author*

Abstract: This study centers on the privacy risks and business enabling strategies of ChatGPT within the commercial landscape. Against the backdrop of the rapid advancements in artificial intelligence technologies, particularly the remarkable performance of large language models such as ChatGPT in natural language processing, this paper delves into the privacy challenges inherent in its widespread adoption. Through a questionnaire survey, the research analyzes the perception of privacy risks among diverse user groups (e.g., gender, educational background) and examines the factors influencing these perceptions, further exploring how these factors shape the formulation of business strategies. The findings reveal that factors such as gender and educational background significantly impact users' concerns regarding privacy breaches. Based on these insights, the study proposes differentiated and customized marketing recommendations for the commercial application of ChatGPT. This paper aims to offer strategic guidance for achieving a balance between technological innovation and privacy protection, while also presenting practical suggestions for the future development of AI technologies.

Keywords: ChatGPT, Privacy Risks, Business Enabling, User Survey.

1. Introduction

In the era of rapid advancements in artificial intelligence (AI) technologies, large language model applications have sprung up like mushrooms, deeply integrating and reshaping our daily lives, professional operations, and even the overall social structure. Among these, outstanding representatives such as ChatGPT, ERNIE Bot, and MOSS are leading the wave of technological transformation with unprecedented momentum. ChatGPT, as the latest achievement of OpenAI in large language modeling, has garnered widespread attention from various fields worldwide due to its exceptional natural language processing capabilities and diverse application prospects. However, amidst the powerful functions and remarkable values demonstrated by ChatGPT, numerous users have also expressed concerns about privacy protection, data security, and ethical issues. Therefore, while enjoying the convenience and efficiency brought by AI technologies, ensuring that user privacy is not violated and striking a balance between technological innovation and ethics have become critical issues that need to be addressed urgently.

Moreover, the actions of competitors in the market cannot be overlooked. Their contributions to technological innovation, the setting of user privacy standards, and their impact on the overall industry environment are all indispensable considerations for this study. This research aims to delve

DOI: 10.54254/2754-1169/122/2024.17803

into the privacy risks faced by ChatGPT and similar AI applications in the commercial environment, while exploring ways to maximize their commercial potential under the premise of ensuring user privacy security. It also aims to provide practical suggestions for future technological development. The study seeks to reveal how factors such as gender, educational background, marketing strategies, and market competition jointly influence the privacy risk management of ChatGPT and its users, and to propose corresponding strategic recommendations. In this process, the study not only focuses on the technology itself but also emphasizes the interaction of technology with the social environment, particularly the differences in attitudes and behaviors towards technology among different groups (e.g., by gender and educational level), and how these differences affect the perception and management of privacy risks.

2. Research Design

2.1. Literature Review

2.1.1. Development of ChatGPT and Other Similar Technologies

In recent years, the rapid development of artificial intelligence (AI) technologies has significantly transformed various aspects of our lives. ChatGPT, a pioneering achievement in the AI landscape, has emerged as a milestone in the evolution of generative AI with the release of its multimodal version, GPT-4 [1]. Concurrently, China has witnessed the emergence of similar large language model (LLM) applications, notably including Wenxin Yiyan, Tongyi Qianwen, Doubao, and iFLYTEK Spark, among others, which have gained high usage frequency. Wenxin Yiyan, for instance, stands out as one of the most popular LLM applications in China due to its unique advantages within the Chinese linguistic and cultural context. It swiftly comprehends and delves into the nuances of the Chinese language and cultural background, achieving superior results compared to foreign LLM applications in the Chinese domain [2].

Regarding users of ChatGPT and other LLMs, they can be categorized into three distinct groups based on their attitudinal orientations: ardent supporters, who recommend and affirm the potential value of LLMs across multiple domains; neutral observers, who acknowledge both the strengths and limitations of LLM applications and adopt a dialectical perspective; and resistant opponents, who express concerns and resistance to the risks and threats posed by these technologies [3]. The core objective of large AIGC (Artificial Intelligence Generated Content) models is to generate content that resembles human language styles. However, during this process, training data may inadvertently contain sensitive and private personal information, such as account passwords, medical records, and so forth, thereby posing a risk of privacy leakage. Furthermore, LLMs possess the capability to infer potentially sensitive information, including users' preferences, interests, and behaviors, which may expose users to the threat of misleading information that could manipulate their opinions and behaviors [4].

2.1.2. Frontier Research on Business Empowerment Strategies

In the realm of business strategies, differentiated and customized marketing have emerged as pivotal factors in enhancing product competitiveness. As their names suggest, differentiated and customized marketing entails the process of designing and individually producing exclusive products tailored to each customer's specifications, while maintaining the efficiency of mass production. By integrating this marketing approach with practical production, enterprises must recognize that implementing customized marketing strategies represents an effective avenue to fundamentally address the issues of high operating costs and low profitability [5].

DOI: 10.54254/2754-1169/122/2024.17803

In the context of today's highly developed market economy, it is essential for corporations to fully understand and prioritize the pivotal role of differentiated marketing in order to enhance customer purchase intention and distinguish themselves in the competitive landscape. Empirical research has demonstrated that brand recognition serves as a mediating factor between market-oriented differentiated marketing and customer purchase intention. Consequently, enterprises need to thoroughly comprehend their competitors, identify breakthrough points for differentiated design, offer diversified and unique products, and tap into broader potential consumer segments. Simultaneously, establishing a distinctive brand image is crucial to fostering customer brand awareness, thereby further augmenting their purchase intention [6].

2.2. Questionnaire

To facilitate the quantitative analysis of users' perceptions and attitudes towards privacy risks associated with ChatGPT, and to encompass a broader range of user groups, this study adopts a questionnaire survey as the primary means of data collection and analysis. The questionnaire comprises two sections: the first section gathers basic user information (including age, gender, region, major, and educational background), while the second section collects information on users' basic usage of ChatGPT (including frequency and duration of use, overall experience, relevance to work/study, continued use of traditional tools, privacy sensitivity, and perceived responsible parties for risks).

In terms of the target population, this study randomly selected users from various digital platforms on the internet, primarily focusing on professionals and student groups. Ultimately, 192 questionnaires from diverse samples were received. Immediately after data collection, the data analysis tool SPSS was utilized for preprocessing, including data cleaning and handling missing values. Subsequently, Excel was employed to conduct regression analysis, examining the influence of factors such as gender, educational background, and age on users' privacy sensitivity.

3. Results

3.1. Overview of Sample Characteristics

This study collected questionnaire feedback from 192 ChatGPT users with diverse backgrounds. Starting with the personal information section, the age distribution was primarily concentrated among individuals aged 19-23, accounting for 47.4% of the sample. The gender breakdown revealed that 45.31% of respondents were male, while 54.69% were female. In terms of geographical origin, 62.5% of users hailed from urban areas. Moreover, 69.27% of respondents were from non-IT majors, and 72.4% did not hold graduate degrees.

Regarding usage frequency, the distribution was relatively wide-ranging, with the highest proportion (18.75% each) belonging to those who used ChatGPT three times or more than five times per week. Conversely, only 14.06% of users reported no usage in a given week. In terms of usage duration, the majority (50%) spent 1-2 hours per week using ChatGPT, with the proportion of users decreasing as usage time increased.

Additionally, 69.27% of users reported using ChatGPT for work/study-related purposes. Notably, a substantial proportion (72.92%) continued to use traditional tools in addition to ChatGPT. When it came to privacy concerns, users displayed a near-neutral attitude towards the potential for privacy breaches during usage, and their opinions on whether they themselves were the primary responsible parties for risks were also relatively balanced.

Finally, regarding overall usage experience and willingness to recommend, the majority of users expressed positive sentiments. Specifically, the average rating for usage experience was 6.4 on a scale

DOI: 10.54254/2754-1169/122/2024.17803

of 1 (very poor) to 10 (excellent). Furthermore, 78.13% of users were willing to recommend this tool to others.

3.2. Sample Results Analysis

After conducting a regression analysis of the survey results using Excel, we derived the following regression equation:

$$\text{Privacy Leaking} = 5.78 + 0.00 \times \text{Age} + 0.65 \times \text{Gender} + 0.40 \times \text{Location} - 0.55 \times \text{Major} - 0.70 \times \text{Education} - 0.14 \times \text{ChatGPT\_Freq} + 0.13 \times \text{ChatGPT\_Len} + 0.09 \times \text{ChatGPT\_Perc} + 0.89 \times \text{ChatGPT\_JobRel} - 0.96 \times \text{Conv\_Tool} + 0.27 \times \text{Liability} - 0.35 \times \text{Recommendation} \quad (1)$$

Where,

Privacy Leaking: The dependent variable, representing the level of concern about privacy leakage.

Age: An independent variable, representing age.

Gender: An independent variable, representing gender (Male = 0, Female = 1).

Location: An independent variable, indicating hometown location (Urban = 0, Rural = 1).

Major: An independent variable, classifying majors (IT-related = 0, Non-IT-related = 1).

Education: An independent variable, indicating the highest level of education (Graduate = 0, Non-Graduate = 1).

ChatGPT_Freq: An independent variable, representing the frequency of ChatGPT usage per week.

ChatGPT_Len: An independent variable, indicating the duration of each ChatGPT session.

ChatGPT_Perc: An independent variable, representing the perceived experience of using ChatGPT.

ChatGPT_JobRel: An independent variable, measuring the relevance of ChatGPT usage to work or study.

Conv_Tool: An independent variable, indicating whether traditional tools are still being used (No = 0, Yes = 1).

Liability: An independent variable, reflecting the perceived responsible party for risks (Not Me = 0, Me = 1).

Recommendation: An independent variable, representing the willingness to recommend.

This study collected data through a questionnaire survey and employed multiple linear regression analysis to explore the impact of various independent variables on the level of concern about privacy leakage (Privacy Leaking). The regression model (as presented above) reveals the quantitative relationships between each variable and the level of concern about privacy leakage. The model's R Square value of 0.05 indicates that the model explains 5% of the total variation in the dependent variable, suggesting that there are other significant factors not accounted for in the model. Upon further analysis of the P-values, we identified the following four independent variables as having statistically significant effects on the dependent variable (P < significance level, typically 0.05):

Table 1: The results of four independent variables

Variables	P < significance level, typically 0.05
Gender	0.65
Education	-0.70
ChatGPT_JobRel	0.89
Conv_Tool	-0.96

Gender: The coefficient of 0.65 indicates that gender has a significant impact on the perception of privacy leakage. We can infer that female users are more concerned about privacy leakage, which

DOI: 10.54254/2754-1169/122/2024.17803

aligns with our prior theories and may be related to inherent gender-specific personality traits and varying degrees of acceptance of technology.

Education: The coefficient of -0.70 shows that graduate students are indeed more concerned about privacy leakage than non-graduates. This suggests that educational background can influence users' risk-taking and sensitivity levels towards emerging technologies, thereby affecting their worries about privacy security when faced with potential privacy breaches. The higher the educational level, the more concerned users tend to be about their privacy safety.

ChatGPT_JobRel: The coefficient of 0.89 demonstrates that users become more concerned about privacy leakage as their use of ChatGPT becomes more relevant to their work or study. This may be attributed to users being more attuned to content related to work or study, leading them to worry more about the potential repercussions of its exposure.

Conv_Tool: The coefficient of -0.96 indicates that individuals who tend to use both traditional and emerging tools simultaneously are less concerned about privacy leakage. This may stem from the fact that such users do not solely rely on emerging tools but instead utilize a variety of tools to mitigate privacy risks.

While other variables such as age, hometown location, major, ChatGPT usage frequency, and duration were included in the model, they did not exhibit significant statistical correlations in this study. This could be due to the weaker direct connections between these factors and privacy breaches, or the influence of unobserved variables. In summary, this research emphasizes the importance of gender, education, the relevance of ChatGPT usage to work/study, and the use of traditional tools in shaping the level of concern about privacy leakage.

4. Recommendation

Based on the analysis above, it can be concluded that users have a significant concern for privacy protection, particularly among specific demographics such as females and highly educated individuals. Therefore, based on the data analysis results and recommendations for improving privacy security, large model applications like ChatGPT should further refine their market positioning as "Intelligent and Secure AI Large Model Platforms" and intensify their efforts in privacy protection. To address the varying privacy sensitivities of different groups, we can develop differentiated and customized marketing strategies. Below are some suggested improvements:

4.1. Enhancing Technological Transparency

As a cutting-edge, high-tech platform, ChatGPT should cater to a diverse range of users, encompassing people of different ages, genders, and educational backgrounds. This necessitates the platform to openly and comprehensively disclose the entire process of data collection, learning, and storage to users, enabling them to genuinely understand its operational model and alleviate concerns about privacy security. Simultaneously, the platform should introduce more secure authentication and encryption methods to ensure that data remains secure and intact during transmission, thereby enhancing user data security and building trust among a wider user base.

Furthermore, the platform should utilize user-friendly language and visual tools to simplify the user experience, facilitating comprehension and mastery of large language models. By doing so, users' trust in the platform's privacy protection measures will be strengthened.

4.2. Gender-Differentiated Marketing Strategy

Gender-differentiated marketing involves tailoring brand messaging and promotions based on the varying levels of privacy sensitivity among different genders. For female users, a "Privacy Safety First" brand image can be cultivated through a series of thoughtful marketing strategies. This includes,

DOI: 10.54254/2754-1169/122/2024.17803

but is not limited to, designing warm and reassuring advertising visuals, utilizing soft and reassuring color schemes, and emphasizing the product's capability to safeguards users' personal information at a technical level, such as through encryption technologies and anonymous data processing.

On the other hand, for male users, the brand can focus on showcasing a "Performance and Convenience First" brand philosophy. The focus here is on how the product, while ensuring privacy, assists male users in accomplishing tasks quickly and efficiently. Highlighting the product's practicality and technological prowess can appeal to male users who seek a balance between efficiency and privacy.

4.3. Education-Tailored Marketing Strategy

Education-tailored marketing involves crafting exclusive brand messaging and service strategies tailored to users with different educational backgrounds. It necessitates a deep understanding of the unique needs, preferences, and differences in information reception and processing among users with different levels of education.

For users with higher educational qualifications, the brand can adopt a "Professional Depth and Quality Emphasis" promotional strategy. These users tend to have a higher level of interest and requirements for a product's professionalism, technical details, and underlying scientific principles. In contrast, for users with relatively lower educational levels, the brand should prioritize a "Simplicity, Clarity, and Practicality" communication approach, which means using more accessible language and engaging examples to explain product functionalities and service benefits, avoiding excessive technical jargon and complex theories.

4.4. Integrated Marketing Strategy Blending Personalization with Scenario-based Needs

Given the significant differences in users' privacy sensitivity and acceptance levels, which are closely tied to personal factors such as gender and education, platforms should offer customizable privacy handling solutions. These include limiting the types of data collected, implementing temporary data collection modes, and developing flexible privacy control panels. These allow users to adjust their data sharing levels based on their individual privacy preferences and usage requirements, fostering a truly personalized privacy protection experience.

Additionally, for user groups with specific scenarios and purposes, such as enterprise or government users, platforms should provide even more tailored privacy solutions. These customized versions enhance specific functionalities, including data encryption, access control, and security audits, ensuring the absolute safety of corporate core information and sensitive data. By adopting these customized versions, enterprises can effectively mitigate the risk of data breaches and significantly reduce the potential for economic losses arising from data security issues. Through this tailored marketing strategy, platforms not only cater to the individualized privacy needs of diverse users but also enhance the broad acceptance and market appeal of large language model applications.

5. Conclusion

In conclusion, the coming era will be one of massive AI involvement and transformation. As humans, our primary task is to research how to best empower AI to enhance our work and increase its engagement in the daily lives of ordinary people. Through an in-depth exploration of ChatGPT's privacy risks and commercial empowerment strategies, this article offers marketing-related suggestions for the healthy development of AI technology.

Research has shown that large language models like ChatGPT face multiple risks in commercial applications, including privacy breaches and data security concerns, and that different user groups perceive these privacy risks differently. Consequently, devising differentiated privacy protection

DOI: 10.54254/2754-1169/122/2024.17803

strategies integrated with effective marketing tactics has become crucial in balancing technological innovation with privacy protection.

The improvements proposed in this article aim to provide valuable insights for ChatGPT and related applications, as well as offer assistance in refining their commercial strategies. Future research could delve deeper into the mechanisms behind these salient factors and consider incorporating more potential variables to enhance the explanatory power of the models.

References

[1] *Chen Gang. Research on communication change and development driven by generative artificial intelligence: an example of ChatGPT[J]. Academic, 2024(06): 62-69.*

[2] *Park Mi-Sun,Huang Bao-Feng,Qing Feng. A study on the user experience of generative AI service Wenshin Yiyin[J]. Technology and Market, 2024, 31(06): 173-176.*

[3] *X. Zhang,Y. Wang. Integration and symbiosis:ChatGPT's risk response and prospect in digital humanities[J/OL]. Library Theory and Practice: 1-14[2024-07-07]. https://doi.org/10.14064/j.cnki.issn1005-8214.20240529.001.*

[4] *Hsu C.W.,Li H.L.,Li B.,et al.An overview of AIGC big model measurement: enabling technologies, security h azards and responses[J/OL]. Computer Science and Exploration: 1-34[2024-07-07]. http://kns.cnki.net/kcms/detail/11.5602.tp.20240523.1947.002.html.*

[5] *Gao ZJ. Precision marketing based on consumer insight in the context of mobile internet[J]. Research on Business Economy, 2020(11): 86-89.*

[6] *Lai Wenwu. Correlation analysis of differentiated marketing and customer purchase intention - based on the perspective of network centrality and network density[J]. Research on Business Economics, 2023(06): 160-164.*