ChatGPT's opportunities and challenges for privacy protection: A systematic literature review

Qianwen Yu

Dissertation submitted to AUT University in partial fulfilment of the requirements for the degree of Master of Business

2024

AUT Business School

Abstract

With the rapid development of artificial intelligence technology, ChatGPT has become an important platform for information exchange and data acquisition. This is a systematic literature review in compliance with PRISMA guidelines and is also a master dissertation. It organises and evaluates existing research results in the field of "Opportunities and Challenges of Privacy Protection in ChatGPT". As for how to conduct this systematic literature review under PRISMA guidelines is to familiarise myself with the PRISMA 2020 checklist and its essential items. I obtain the checklist and explanations for each item from the official PRISMA website or related publications. I also use a systematic and comprehensive approach to search for relevant studies. I will document my search strategy, databases used, search terms, and any filters applied. This is to ensure my search covers the time frame and study designs specified in my protocol.

The purpose of this dissertation is to provide as comprehensive information as possible on the risks and opportunities faced by ChatGPT in terms of privacy protection. At the same time, this dissertation also puts forward solution suggestions based on existing literature as much as possible. In general, this graduate dissertation is mainly to help ChatGPT provide scientific strategies in the field of privacy protection, help the healthy development of the network, and protect the security of personal information on the network.

As for how to enhance ChatGPT privacy protection, there are many methods. One of the most important thing is how to process data through machine learning algorithms. The second point is whether the network's privacy policy meets the needs. In addition, improve users' privacy protection awareness is also vital. Besides, it is very important to strengthen users' compliance with relevant regulations. At the same time, in response to ChatGPT's privacy protection challenges, the dissertation also discusses how data is collected, how people inadvertently leak their own information, excessive use of artificial intelligence, insufficient transparency of personal information, users' trust in the system, etc.

Based on the above, this dissertation will use prospect theory to analyse the topic for graduate dissertation, which will be introduced in the following chapters. The application of prospect theory can provide valuable reference for user behaviour. At the same time, analysing the topic of the dissertation by combining prospect theory can also help solve research problems and provide directions for future development.

Finally, in response to the research questions of the paper, at the end of the article, various related literature to summarise a series of suggestions for privacy protection of ChatGPT will be combined. Also, The limitations of the current research and proposed

future research directions will also be pointed out. This article contributes to the discussion of privacy issues in the digital generation. The main purpose is to provide users with an easy-to-understand framework and provide some inspiration for future artificial intelligence developers to improve their privacy protection functions.

Table of contents

List of Figures & Tables

Attestation of Authorship

"I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor used artificial intelligence tools or generative artificial intelligence tools (unless it is clearly stated, and referenced, along with the purpose of use), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning. "

Qianwen Yu 08.11.2024

Acknowledgements

Of course, I must express here my utmost favourite words for every one of you who made possible the completion of my dissertation with title “Opportunities and Challenges on Privacy Protection in ChatGPT: A Systematic Literature Review”.

I am primarily thankful to my supervisor, Dr. Hari Srivastava who has been steadfast in his guidance and direction, patience during the course of this dissertation effort. Dr Srivastava's in-depth expertise and pragmatic feedbacks helped me in critically thinking about my dissertation. This has helped me a lot in my research career.

I also thank the entire staff of AUT University for providing such a stimulating and collegial atmosphere within which I could develop myself.

I cannot adequately express my gratitude to my family, whose love and support sustained me through those days when I felt I could not do it anymore. Their comprehension and sacrifices are the reason that I chase with my academic.

Finally, I would like to thank the writers and researchers whose work I had used in this dissertation. To get there, I built on their scholarship.

Looking back at this phase of my academic journey, I only feel like I have achieved something and all thank to sum such role models for guiding me.

1. Introduction

The evolution of artificial intelligence (AI) technology has transformed the way we communicate. AI equipped chatbots like ChatGPT are playing a major role in that change (Al-Amin et al., 2024). Going back to the history of ChatGPT, this advanced chatbot, or what we called "Chat Generative Pre-trained Transformer", a large language model, was first released in November 2022 by OpenAI (Roumeliotis & Tselikas, 2023).

Different from normal chatbots, ChatGPT can not only communicate smoothly with users in a way that is close to natural human dialogue, but also can handle various complex language tasks. According to Roumeliotis and Tselikas (2023), it has advanced features such as automatic text generation, automatic question answering, and automatic content summarisation. These functions allow ChatGPT to show great potential and value in multiple application scenarios such as automated content creation, customer service, and rapid information summary. At the same time, ChatGPT will collect certain data while talking to users, this involves user privacy issues, which means that users have the right to decide whether and to what extent they share their specific personal information with ChatGPT (Naghiyev, 2024).

However, as ChatGPT technology develop, users' privacy protection concerns grow. Since ChatGPT's technology relies on large data sets and user interactions that may contain sensitive information, the integration of ChatGPT into digital interactions in daily life urgently requires us to take a deeper look at their impact on privacy (Roumeliotis & Tselikas, 2023). According to Table 1, since ChatGPT was released in 2022, its users have increased nearly 8 times until 2023. This growth rate was amazing. As the system become more popular, understanding the protection of privacy becomes critical. There is a great need that we evaluate the challenges and opportunities that ChatGPT may face in the context of privacy protection.

Table 1: OpenAI's model history

Model Name	Release Date	Parameters
GPT-1	2018	117 million
GPT-2	2019	1.5 billion
GPT-3	2020	175 billion
InstructGPT	2022	1.3 billion
GPT-3.5	2022	200 billion
ChatGPT	2022	200 billion
GPT-4	2023	8x 220 billion
Code Interpreter	2023	-
GPT-4o	2024	-

Note. Since the advent of ChatGPT, its user parameter data has changed. From *365DataScience*， by Fabio Duarte, 2024, Explodingtopics. Copyright free. Reprinted with permission.

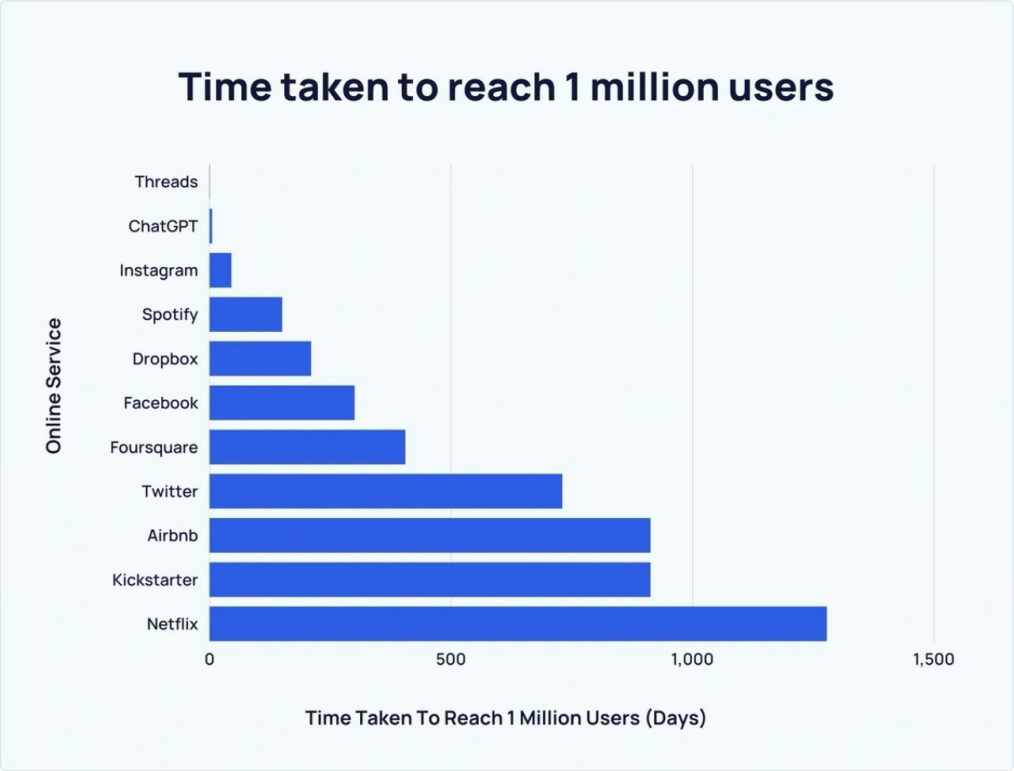
As for why the research is worth studying is mainly because the release of ChatGPT has triggered deep personal concerns about privacy protection. ChatGPT collects and processes large amounts of user data, which raises concerns about data breaches and misuse. For example, the users of ChatGPT, when they communicate and interact with it, they will inadvertently share a lot of personal information, such as personal preferences, personality traits and even sensitive data about personal privacy (Xu et al., 2014). User trust is a crucial factor for the widespread adoption of AI technologies. If users do not trust that their data will be handled securely, they are less likely to use these platforms (Elman, 2019).According to the Edelman Trust Barometer Report, data breaches can significantly reduce consumer trust in brands, and therefore reduce users (Elman, 2019). If a similar incident occurs in ChatGPT, in the process of these users using ChatGPT, if they discover that the content of his or her conversation with ChatGPT has been leaked or abused, other users may also have concerns about the technology, causing ChatGPT to lose users (Charfeddine et al., 2024).

Additionally, ChatGPT's decision-making process may be uncharted territory for users. The complexity of ChatGPT privacy policies often makes it difficult for the normal user to understand, which partly reflects a lack of transparency (Khowaja et al., 2024). Users of ChatGPT may want to know how ChatGPT uses their data to generate responses and whether the data is used for other purposes, and this lack of transparency may lead to

users feeling distrustful of how their data is handled (Choudhury & Shamszare, 2023). With the development of technology, ChatGPT's technology may be used to create wrong information or conduct some serious network attacks, and these attacks will threaten the privacy and security of users. For instance, some people with worse intentions may use ChatGPT to generate realistic fake news or scam emails to induce users to disclose sensitive information (Falade, 2023).

In addition, as ChatGPT technology continues to advance, its functions have become powerful. That is to say, ChatGPT's ability to handle complex tasks is constantly increasing. It enables of processing and generating text up to 200,000 words long for now, making it capable of complex tasks such as long-form writing, translation, and coding (Haleem et al., 2022). Also, according to Figure 1, it is obvious that compared to other platforms, ChatGPT's user growth rate is very fast. This means that ChatGPT not only improves efficiency, but also makes its users increasingly rely on the intelligent systems. However, this dependency also means that ChatGPT users must have greater trust in these systems to properly handle their personal information (Zaman, 2023).

Figure 1: Time taken to reach 1 million users



Note. The time it takes for different mainstream programs to reach 100 million users since the release of ChatGPT. From *Reuters*, *Similarweb*, by Fabio Duarte, 2024, Explodingtopics. Copyright free. Reprinted with permission.

Therefore, there is a great need that we conduct a comprehensive evaluation of ChatGPT to ensure that it can respect and protect user privacy while providing efficient services.

At the same time, we also need to consider how to find the right balance between technological innovation and privacy protection to ensure that the development of ChatGPT can benefit users. This is because by identifying and solving privacy protection challenges, we can provide guidance for the innovation of ChatGPT and other similar technologies, promote technological development while protecting user rights (Ali et al., 2023).

Moreover, studying the opportunities and challenges of ChatGPT in privacy protection is not only of great significance to its users, but also has a profound impact on the researchers.

For researchers, ChatGPT and its privacy issues provide them with new research directions, including data security and user behaviour analysis. As ChatGPT model becomes more complex, researchers can study how to ensure the fairness and transparency of data, avoid embedding bias in training data (Ray, 2023). Besides, with the development of ChatGPT, ethical issues have become increasingly prominent. While ChatGPT can continuously enhance its service skills by using a vast quantity of training data, it will never be able to match humans' depth of comprehension and application capabilities (Wu et al., 2023). In the first batch of user tests, ChatGPT even made remarks such as insulting users and inducing users to divorce (McIntosh et al., 2023). Also, although there is discussion about the challenges or opportunities that ChatGPT may bring to users, specific governance strategies and specifications are not yet complete (Gill & Kaur, 2023). Despite privacy issues in AI being a popular research area, a systematic literature review has not been conducted on the privacy issues associated with large language models like ChatGPT(see Table 2) , which indicates that there is a research gap (Zhang & Tur, 2024). By conducting a systematic literature review, this dissertation will also provide a structured overview of the current state of research, identify key trends, and highlight areas that require further investigation.

Table 2: Previous (systematic) literature review related to ChatGPT

Source	**Research type**	**Aims**	**Findings**	**Limitations**
(Alsharida et al., 2023)	Systematic Literature Review (SLR)	Evaluate and synthesise cybersecurity theories and models, independent variables, target variables,	According to research, the two theories that are most frequently applied in the analytical literature are the Theory of	The majority of research is not general and is centred on the individual, particularly students and end users.

		moderator variables, techniques, etc. to provide a multiperspective understanding of human cybersecurity behaviour.	Planned Behaviour (TPB) and Protection Motivation Theory (PMT). Moderator factors were not used in 76% of the studies that were analysed to look at the link between predictor and target variables. The majority of research is done on an individual basis, mostly with end users and students. Mobile devices and social media are the most often used instruments for researching cybersecurity behaviour in people.	
(Ayinde et al., 2023)	Literature Review (LR)	Explore ChatGPT as an important tool in organisational management,	ChatGPT has a significant impact on organisational operations and decision-	While there has been much discussion on ChatGPT's efficacy as an

		and examine its multifaceted impact in organisations through a literature review.	making processes by simplifying data processing and information generation processes. At the same time, it also poses social, economic and legal challenges.	organisational management tool, a thorough literature study is required to deepen the research.
(Kayser & Telukdarie, 2024)	Literature Review (LR)	This article uses the TAM framework to analyse the variables that influence the acceptance and adoption of AI/GPT in the accounting industry, such as external factors, attitudes about AI, and perceived utility and simplicity of use.	The adoption of AI/GPT can bring significant benefits to people, such as improved efficiency, accuracy, and support for decisionmaking, but it also brings challenges to people, such as high investment costs, increased demand for professional talents, and data privacy and security issues.	The scope of the article analysis is in the accounting field and has certain limitations.

(Stahl & Eke, 2024)	Systematic Literature Review (SLR)	Apply the ethical method of analysing emerging technologies to systematically examine the benefits and concerns that ChatGPT may bring.	Although ChatGPT presents significant ethical challenges in the areas of social justice, individual, it can also have positive effects on society and ethics.	As technology use cases develop, a broad, balanced ethical approach must be applied consistently in order to realise benefits and handle ethical dangers.

So studying the privacy issues of ChatGPT will help readers deeply explore and respond to ChatGPT ethical challenges.

Therefore, based on these above, this dissertation raises two research questions are as follows:

1. What are ChatGPT's main challenges and opportunities in the context of privacy protection?
2. What measures can ChatGPT implement to ensure the protection of user information?

Besides, this dissertation adopts the form of a systematic literature review and in accordance with the PRISMA recommendations (Page et al., 2021) to conduct research on "Opportunities and Challenges of Privacy Protection of ChatGPT". The purpose is to comprehensively collect and evaluate the existing literature on the challenges, opportunities and possible solutions faced by ChatGPT in terms of privacy protection. At the same time, providing ChatGPT with scientific, evidence-based suggestions and strategies in the field of privacy protection to promote the healthy development of technology and ensure the security of personal information. This process involves the systematic search, screening, evaluation and synthesis of relevant research evidence to have a thorough awareness of the field's present situation and potential future directions. Therefore, based on the above, the search keywords of this dissertation are set as: ((generative AND "pre-trained" AND transformer* OR chatgpt OR "large language models" OR "LLM*") AND (opportunit* OR barrier* OR challenge*) AND (privacy OR security*)). Since ChatGPT was launched in 2022, the literature evaluation of this

dissertation is set from 2022 to 2024. Also, this dissertation will conduct a systematic literature review based on the following databases: Science Direct, Scopus, Springer Link.

In order to study the topic of the dissertation, the theoretical framework adopted in this paper is based on prospect theory. To clarify, prospect theory is a psychological theory that was proposed by Amos Tversky and Daniel Kahneman in 1979, which aims to describe people's behaviour when they make decisions in the face of risk and uncertainty (Kahneman & Tversky, 2013). People generally have a greater aversion to losses than to an equal gain. In the context of ChatGPT privacy protection, users may be more concerned about the negative impact of data leakage than its potential benefits. It will enable user sharing their data (Wu et al., 2024).

Therefore, under this situation, prospect theory is very suitable for analysing this problem. It can also help us understand users' risk preferences when they make decisions. Additionally, most of the people they make decisions based on their gains or losses. Under ChatGPT's research, users' privacy preferences may be related to their initial expectations for privacy or based on their previous experiences. For example, if users are experienced before, they may have higher expectations for ChatGPT's privacy protection measures (Wang et al., 2023). Also, people are often unclear about the probability of something happening. They may overestimate the possibility of some low-probability things happening. (Goyal & Miyapuram, 2019). When use ChatGPT, users may be more concerned about the minimal risk of privacy leakage, while ignoring more positive results. According to prospect theory, people may be more willing to take some risks when influenced by possible rewards. From this aspect, this can be used to analyse ChatGPT's behaviour towards user privacy protection. For example, they may be more likely to take risks to protect themselves from privacy leakage rather than to obtain additional service features (Zaman, 2023).

Also, prospect theory emphasises the situation of decision-making, which is dependent. That is to say, the same choice by people may lead to different decisions in different situations (Stahl & Eke, 2024). At the same time, prospect theory believes that people's decisions will change over time or their decisions will change when the situation changes. These help me analyse how ChatGPT users' attitudes towards its privacy protection methods together with the development of technology and the occurrence of privacy incidents.

In conclusion, it is very appropriate to use prospect theory to analyse the topic "Opportunities and Challenges of ChatGPT in Privacy Protection". It provides us with a framework for deeply understanding of how users make decisions in terms of privacy protection in ChatGPT.

2.Methodology

In this methodology part, We list the systematic literature review methodology used in this dissertation, to better study the two research questions listed in this dissertation.

2.1.Methodology

This section serves as the basis for a systematic literature review and will provide a detailed description of the research process to better understand the topic. The goal of this part is to fully grasp the topic of ChatGPT in privacy protection area. To meet this goal, this article conducted a systematic collection and structured analysis of existing literature based on the systematic literature review methods and guidelines proposed by Templier and Paré (2015). Besides, this dissertation collected and synthesised relevant academic research and papers that explore ChatGPT in privacy protection. These guidelines help reduce subjective bias in the analysis of the literature on this topic. Moreover, the dissertation adopted the PRISMA framework proposed by Page et al. (2021) as a conceptual structure to guide the dissertation research part. This framework provides a structured perspective for analysing ChatGPT privacy protection and helps identify key topics, trends, and research gaps. Through the combination of the systematic literature review and the framework of the PRISMA guideline, the article will be able to give an objective analysis based on these knowledge.

2.2.Eligibility criteria (Inclusion, exclusion criteria)

This article used several inclusion and exclusion criteria to make sure that the systematic literature evaluation part would only include the most relevant articles. In Table 3 we can see the inclusion and exclusion criteria that must be strictly met by each article during the step of screening.

Table 3: Eligibility criteria (Inclusion and exclusion criteria)

	Inclusion criteria	**Exclusion criteria**
1	The time range of the article should be between 2022-2024.	Articles before 2022 (exclusive 2022).
2	The keywords of the article should include:((generative AND "pre-trained" AND transformer* OR chatgpt OR "large language models" OR "LLM*") AND (opportunit* OR barrier* OR challenge*) AND (privacy OR security*)).	Articles with other irrelevant keywords.
3	Articles should be written in English.	Articles in other languages.
4	Articles which are peer-reviewed journals or conference articles.	Books, magazines
5	The article needs to measure user privacy protection, ChatGPT's privacy protection challenges, ChatGPT's privacy protection opportunities, that is also to say, the abstract of the article must at least mention ChatGPT's privacy protection challenges OR ChatGPT's privacy protection opportunities OR AI Chatbot.	Other measurements.

2.3.Data sources and Search strategy

This stage is focusing on formulating and implementing plans and results for paper retrieval and collection, which is an important part of the research process. The key of this stage is to apply appropriate search terms and select suitable databases to collect the required information and documentation. A well-designed search strategy is essential to collect comprehensive and relevant information and is also important for me to ensure that the research can be conducted efficiently.

First, based on previous research exploration, we carefully selected a series of keywords to search on Google Scholar. The reason why we choose these keywords was that they are more relevant to the research topic. The keywords were specifically designed to capture the main themes and concepts in the research, ensuring that the retrieved information was highly relevant to the research objectives. In addition to building search terms, choosing the right database is also crucial for me. In this dissertation, the selected information platforms include Scopus, Science Direct and Springer Link. All these databases are highly regarded for their extensive collection of academic literature in multiple subject areas.

As a result of the initial search in Google Scholar, following keywords were identified:

- "generative"
- "pre-trained"
- "transformers"
- "ChatGPT"
- "large language models", "LLMs"
- "opportunities"
- "barriers"
- "challenges"
- "privacy"
- "security"

The search results of Google Scholar confirmed that these carefully selected keywords were closely related to the research questions. Based on this finding, these keywords were used for further searching of Scopus, ScienceDirect and Springer Link databases. Also, the resulting search query will be in the following format:

Scopus / Science Direct/ Springer Link	((generative AND "pre-trained" AND transformers OR chatgpt OR "large language models" OR "LLMs") AND (opportunities OR barriers OR challenges) AND ("privacy" OR "security"))

When searching the three databases, we used different search strings. This is because Science Direct does not support the use of the asterisk (*) wildcard character in its search function. In order to solve this problem, we removed these wildcard characters to fit its search requirements.

Next, in developing the search strategy, inclusion and exclusion criteria were established. These criteria are extremely important in filtering search results, ensuring that only the most relevant and highest quality studies are included. Inclusion criteria

detail the specific criteria that need to be met for retrieved documents, such as type of publication, language, and date of publication. The search will focus on studies be published or will be published between 2022 and 2024, establishing a clear time frame. This time frame has happened from the time when ChatGPT was released to now, which allows us for a comprehensive review of research in this field.

After filtering based on keywords and time (the second and third items in the inclusion and exclusion tables), among them, 6696 were initially screened by Science Direct, 1666 were initially screened by Scopus, and 708 were eligible by Springer Link.

According to the inclusion and exclusion criteria 4, we conducted a screening and selected articles in all the field. we found: Science Direct has 6255 articles, Scopus has 1454 articles, and Springer Link has 369 articles which meet the inclusion criteria.

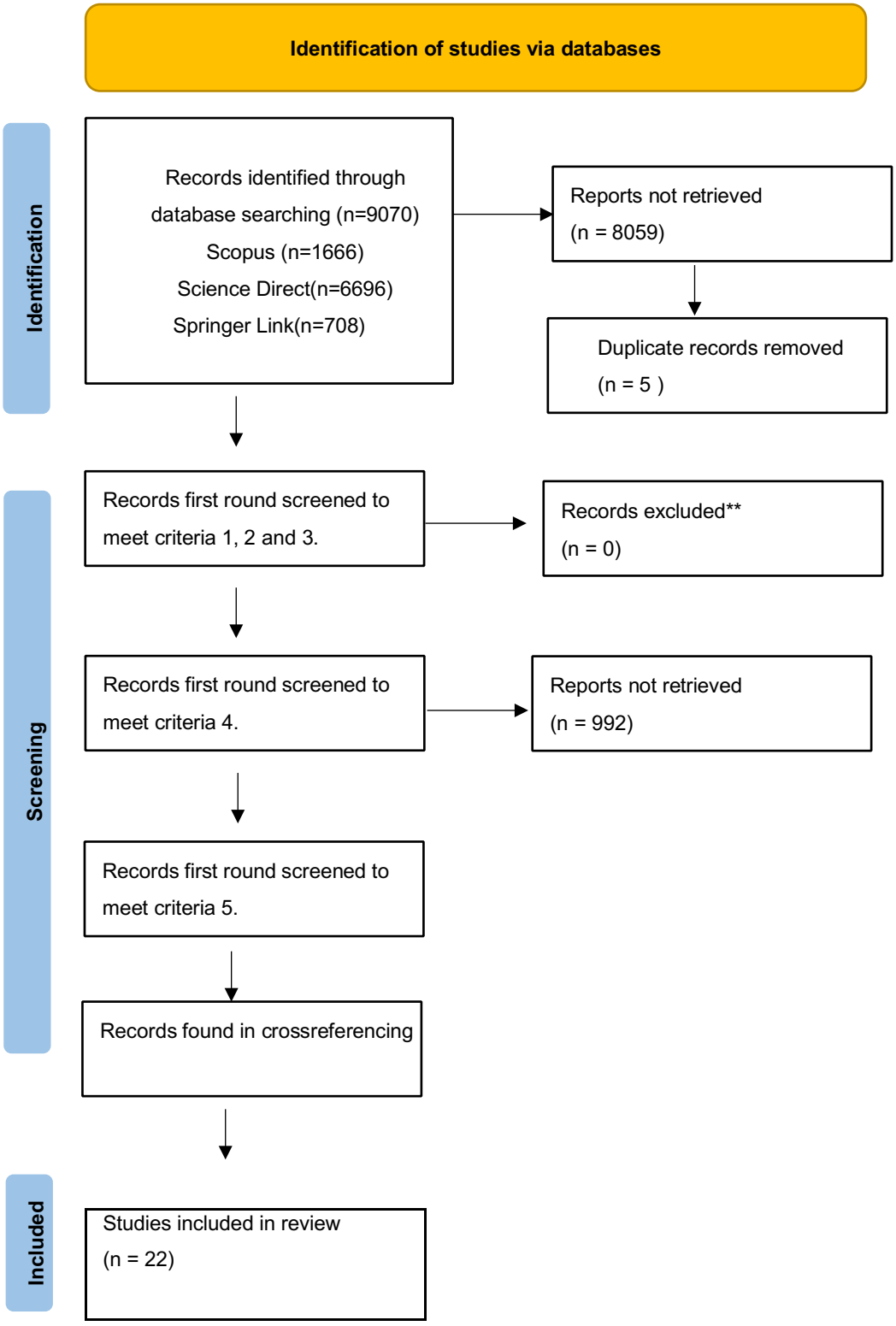
As is showed on the inclusion and exclusion criteria 5, we conducted the final screening and found that: Science Direct has 9 articles, Scopus has 9 articles, and Springer has 1 article.

After screening the reference sections of N = 9 + 9 +1 = 19 articles, we obtained 8 more articles. Then after deleting and merging duplicate articles, there were a total of N = 6 + 7 + 1 +8 = 22 articles. At the same time, since the article studied ChatGPT, when reading the entire article carefully, these 22 articles all meet the requirements.

2.4.Search results

The methods and the related search results of the systematic literature review and the number of eligible articles collected at each stage is shown in Figure 2, and the criteria will meet Table1.

Figure 2: PRISMA flow diagram

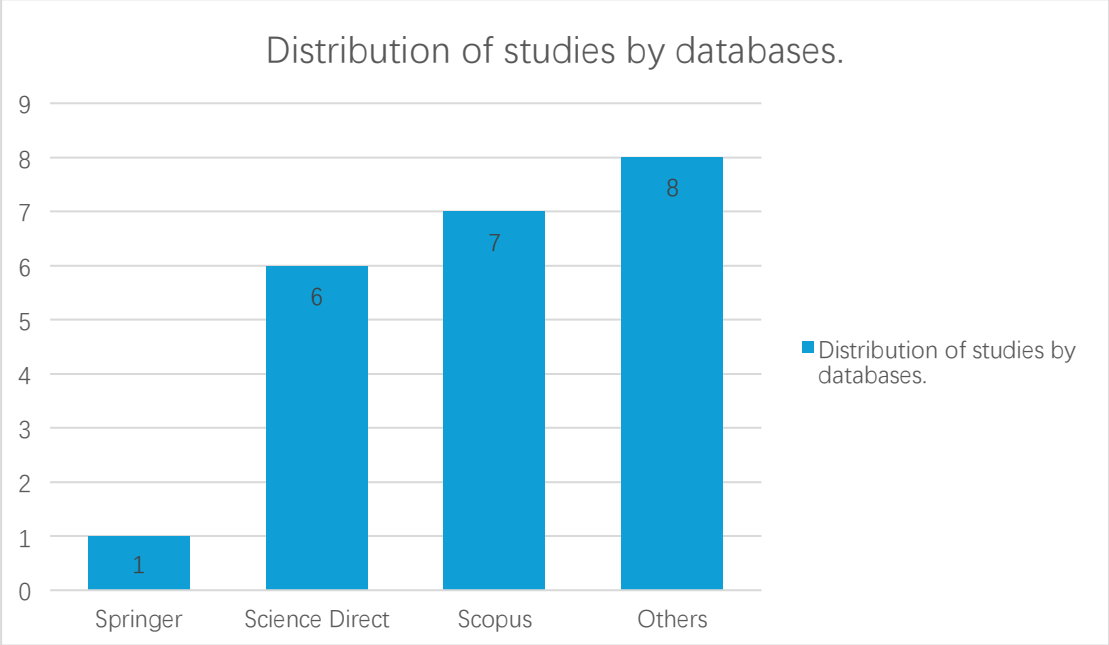


*Since the articles we obtained were not sufficient according to the PRISMA guidelines, in order to make up for this shortcoming, we found some more articles from the references of the finally screened articles to support the topic of the article.

3. Descriptive analysis

During the research journey, three widely recognised databases were used: Scopus, Science Direct and Springer Link to collect academic literature closely related to our research topic. The extensive literature resources held by these databases are crucial for in-depth exploration of the opportunities and challenges in ChatGPT privacy protection field, as shown in the previous table. We identified several documents that were directly related to my research interests, and the frequency of their appearance in Scopus, ScienceDirect and Springer Link highlighted their relevance to my academic topic. The significant relevance of these materials to the aims of my inquiry is evident. According to Figure 3, in particular, other very relevant articles selected from the references following the article contributes the vast majority of literature resources (approximately 36.4%), followed by the documents from the Scopus database, the proportion of articles that meet the requirements is about 31.8%.

Figure 3: Distribution of studies by databases

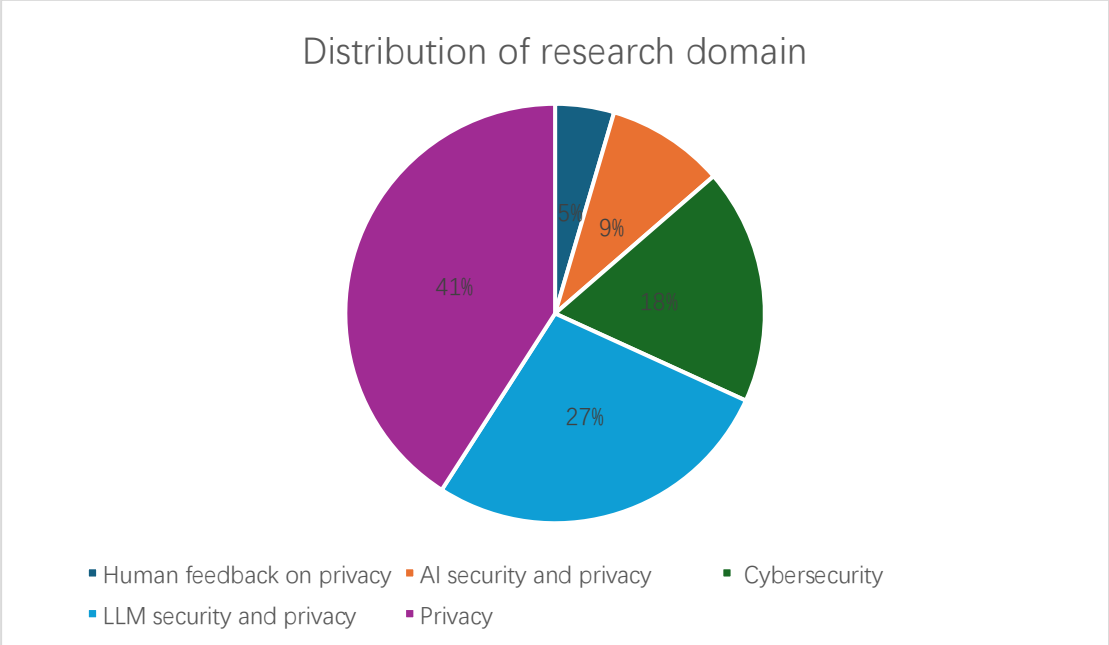


At this stage, this dissertation implemented a two-stage classification process: first by research areas and second by research methods. The first round of classification focuses on domain classification, which allows researchers to observe the distribution of data across different research areas. Such a classification process promotes a deeper understanding of the research topics and assesses their connections to the research questions.

In addition, this step provides important statistical basis for in-depth analysis. By categorising data into domains, it will be able for us to organise and interpret

information in an orderly manner, which helps uncover patterns, trends, and relationships within the data set. The preliminary classification work laid a solid foundation for subsequent analysis and ensured the consistency of the research methods used with the research objectives. Overall, the classification by research area not only enriches the data set but also provides researchers with key insights in subsequent analysis and interpretation.

Figure 4: Distribution of research domain

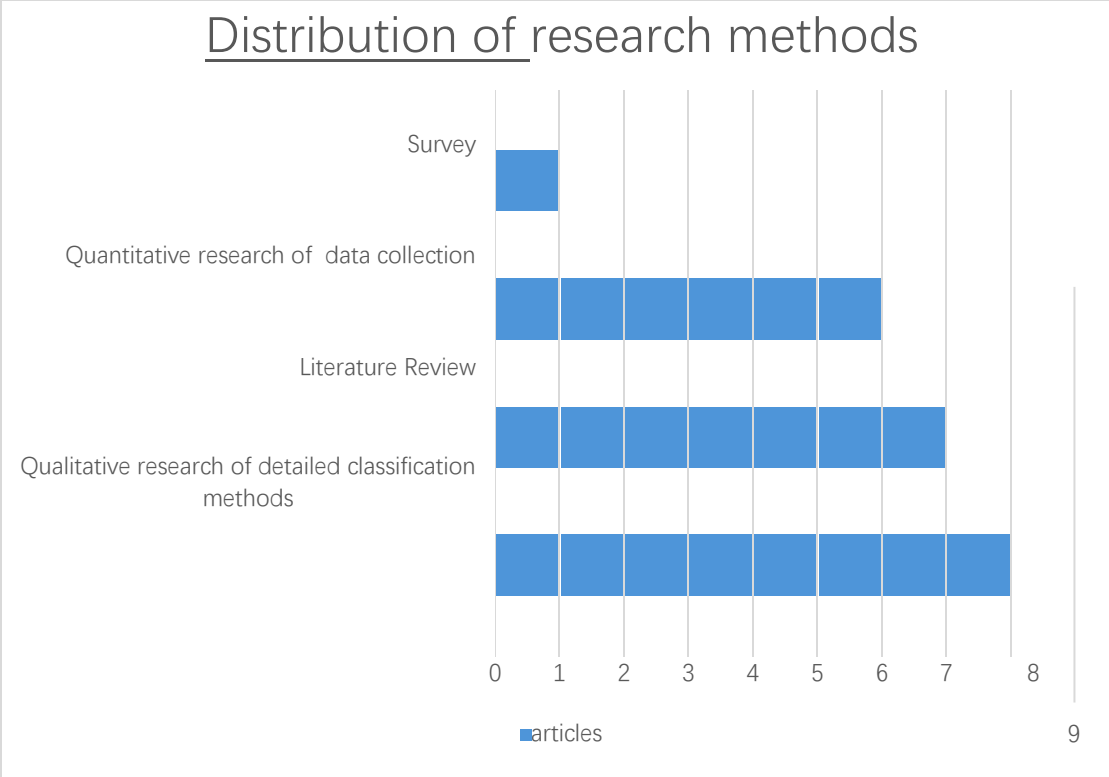


According to Figure 4, among the total 22 articles, the topic "Privacy" was the most popular one, which with 9 articles, accounting for about 41%. Followed by "LLM (Large Language Model) security and privacy", there are 6 articles in total, accounting for 27% of the total selected articles. Next is the topic of "Cybersecurity", with a total of 4 articles that meet the criteria, accounting for 18% of the total. Furthermore, there are two articles under the topic of "AI security and privacy " (accounting for 9%), and one article under the topic of "Human feedback on privacy" (the lowest proportion, only 5%).

In the second round of classification work, we focused on classifying the methods used in the study, which have been adopted by other researchers in research on similar topics. We gained deeper insights into current research trends by identifying each research area and examining the specific research instruments used in their application. This categorisation step allowed us to examine the existing research landscape, identify possible research gaps, and capture the current state of knowledge in the field. By deconstructing the research methods used by other scholars, we gleaned key information about the methods, techniques, and tools used, which was essential to

delving deeper into topics of interest. This methodological categorisation gave us a broader perspective and help us understand the context of the research at a deeper level. It served as a platform for making comparisons, exploring new methods, and possibly improving existing research methods. The insights gained from this classification process provided a solid foundation for the construction of the subsequent framework of this study, ensuring that the research methods used are consistent with current trends and knowledge areas.

Figure 5: Distribution of research methods



According to Figure 5, we can clearly see from the chart that among all the selected articles, the majority of articles use qualitative research of detailed classification methods, with 8 articles. This is followed by 7 articles on literature review as a research method. The least amount of articles are quantitative research of data collection and survey as research methods, with 6 and 1 articles respectively.

4.Synthesis and results

4.1.Synthesis

In the vast field of academic research, the importance of the synthesis stage cannot be underestimated, as it served as the cornerstone for building a comprehensive understanding of the subject matter (Alvesson & Kärreman, 2011). It is vital to the research process itself, bridging the gap between an initial exploration of a topic or claim more comprehensive conclusions that answer my original questions. At this stage of synthesis, the main activity becomes to carefully summarise what are wise lessons learned and relevant for my research area.

But it is not just the process of literally extracting details on dissertation; an intuitive analytical mind must locate those unwritten relationships. This process assists us to organise this huge collection of data and do systematic review on it, one by one information collected can be considered within the study flow. The integration lends a coherence and organisation to the analysis, which was essential for building evidence in making persuasive arguments.

Furthermore, the synthesis stage of the dissertation involved a high level of engagement with the material, where we can critically evaluate the relevance of each source. we can also learn the critical thinking through this stage (Saini & Shlonsky, 2012). Besides, this evaluation was important for us to build the reliability of the research findings.

Additionally, we can identify gaps in the existing literature during synthesis step. It also enabled us to consider how the research might contribute to these gaps. Moreover, this stage provided an opportunity for us to reflect on the research methodology. It allowed us considering whether any adjustments are necessary to ensure the study's objectives were met (Saini & Shlonsky, 2012).

In essence, the synthesis stage was a transformative phase where raw data and literature are transformed into a coherent and meaningful way. It not only answers the research questions, but also contributes to the broader body of knowledge within the academic community (Klag & Langley, 2012). Also, this stage is not just about summarising. It is about synthesising information into a form that is both informative and persuasive, setting the stage for the final presentation of the research findings.

The first task of the synthesis stage is to review and evaluate all the related data. This is a process which need me to be very careful. It also lays the foundation for us to understand the research landscape. This includes not only the original data in the literature, but also the views, theoretical frameworks, and empirical results of other

researchers. This comprehensive review ensures that the view of the topic is relative objective.

Next, we need to categorise this information by several key words : topic, methodology, and findings. This helps us better understand the connections and differences between studies. Categorisation is an important step in the synthesis process, as it organises the various array of data into manageable segments. This organisation can compare research methods more clearly, and can also find something different. Besides, it is helpful for us to find any research gaps during the time when summarising literatures.

In addition to categorisation, researchers need to have critical thinking during the stage of synthesis. As for us, it involves many ways. It includes the comparison of different research methods to evaluate its possible advantages and limitations. Also, it includes analysing the result of the research to find differences. Critical thinking is important, as it challenging assumptions and questioning the status quo. It is also important for us to build a solid foundation on the research.

In conclusion, the synthesis step is a complex but valuable process. It requires us much analytical skills and critical thinking. We can set the stage for a meaningful contribution to the academic research through this step. So, We assessed the information and identified key findings during the synthesis. At the end, we found 22 studies. This may include us comparing different sources, organising data into appropriate categories, and exploring connections between the information been collected.

The results of the analysis showed that a review of these 22 studies are highly related to the research topic. They provided in-depth insights into ChatGPT's opportunities and challenges for privacy protection, and also attract many scholar's attention in this area. These studies, which focus on a wide range of disciplines, are very helpful for us to better understand the research topic. We can also learn the complexities involved in safeguarding user privacy within AI-driven chat platforms like ChatGPT.

These elements are generally considered to be the core of current development of ChatGPT. They are crucial to solving various problems within the privacy protection area. Findings from these studies highlight the need for a multifaceted approach to privacy protection. For example, technological innovation, legal frameworks, ethical considerations, and user education. The integration of these elements is seen as important for the advancement of privacy-preserving technologies. They are also as best practices in the industry for the development.

It is worth us noting that all of these selected articles emphasise the importance of privacy protection, which indicates that privacy protection is an important role. The scholars are agreed on the tide—that privacy is fundamental in an era which they call digital. The studies have highlighted several problems like how data breaches are

becoming increasingly sophisticated, the sheer enormity of personal user-generated data and the necessity to advance encryption & anonymisation methods on this information.

Moreover, the studies have discussed the opportunities that ChatGPT and similar AI technologies present for improving privacy protection. These include the use of machine learning algorithms to detect and prevent unauthorised access to user data. It also includes the development of privacy-improving technologies such as differential privacy, and the process of compliance with privacy regulations.

The systematic analysis showed that, people such as technologists, policymakers, and users should take more attention on the measures of privacy protection. This collaborative effort is important for fostering trust in AI systems and is also essential for ensuring that the benefits of AI are realised by people without compromising the privacy rights of individuals.

In conclusion, the synthesis of these 22 studies has not only showed the current state of privacy protection in the context of ChatGPT but has also pointed the way forward for future research and development. Besides, the findings serve as a call to action for the AI community to put privacy in the first consideration. This can help people innovate in the field of privacy-preserving technologies.

In addition, 11 studies revealed users' continued concern about privacy of ChatGPT and its impact on network security. This finding established an important correlation, indicating that users' trust in their privacy is closely related to information and data sharing practices in the field(Okey et al., 2023; Wu et al., 2023; Pasupuleti et al., 2023; Alawida et al., 2023; Prasad et al., 2023; Khowaja et al., 2024; Goldstein et al., 2023; Voelkel & Willer, 2023; Gupta et al., 2023; Derner & Batistič, 2023; Hu & Chen, 2023). Also, there are 12 studies on the opportunities and challenges of security risks of the LLMs (Large Language Models) model with ChatGPT as a typical example, which shows that ChatGPT under the LLMs model still has many system vulnerabilities and may leak privacy (Yao et al., 2024; Wu et al., 2023; Pasupuleti et al., 2023; Morales et al., 2024; Negri-Ribalta et al., 2024; Khowaja et al., 2024; Goldstein et al., 2023; Voelkel & Willer, 2023; Derner & Batistič, 2023; Li et al., 2023; Zhang et al., 2024; Ferrag et al., 2024). What's more, 8 studies show the interaction between data sharing and trust, law, and personal ethics of GEN AI/AI with ChatGPT as a typical example, which shows that ChatGPT also needs to consider the impact on users at the ethical level (Khalid et al., 2023; M et al., 2024; Sai et al., 2024; Alawida et al., 2023; Prasad et al., 2023; Ouyang et al., 2022; Gupta et al., 2023; Hu & Chen, 2023).

Delving deeper into the opportunities and challenges of privacy protection in ChatGPT, three prominent features emerged repeatedly. Transparency is a fundamental principle for implementing privacy protection, and privacy issues due to lack of transparency were cited 12 times in the selected studies (Khalid et al., 2023; Okey et al., 2023; Wu et al., 2023; Pasupuleti et al., 2023; Alawida et al., 2023; Prasad et al., 2023; Khowaja

et al., 2024; Ouyang et al., 2022; Voelkel & Willer, 2023; Li et al., 2023; Zhang et al., 2024; Ferrag et al., 2024). Data leakage during model training is another key feature, which was emphasised 22 times, which means all of the articles mentioned that point. Also, users' concerns about the leakage of their personal information led to distrust of AI software such as ChatGPT, which was mentioned 11 times (Okey et al., 2023; Wu et al., 2023; Pasupuleti et al., 2023; Alawida et al., 2023; Prasad et al., 2023; Khowaja et al., 2024; Goldstein et al., 2023; Voelkel & Willer, 2023; Gupta et al., 2023; Derner & Batistič, 2023; Hu & Chen, 2023). These numbers convincingly verify the prevalence of privacy issues in GEN AI/AI products such as ChatGPT. Therefore, it is clear that studying the opportunities and challenges of ChatGPT in privacy protection has made a positive contribution to cultivating users' trust on the Internet.

Through a comprehensive analysis of 22 studies, it can be found that protecting user privacy and limiting real information sharing and data sharing on the Internet are crucial in the field of big data models such as ChatGPT. At the same time, these aspects have attracted widespread attention and are crucial to solving relevant privacy or data protection issues within the industry. In addition, the interaction between information and data sharing of ChatGPT and the concept of user trust has been widely explored in the selected studies, highlighting significant correlations. In the field of ChatGPT technology, transparency, data leakage, and user trust have become prominent features that are always present in ChatGPT-related GEN AI/AI research. These characteristics help to establish a relatively positive impact in LLMs field.

Overall, the comprehensive stage of the research allows me to have a deeper understanding of key findings, patterns, and trends, providing valuable insights for advancing the field of ChatGPT privacy protection.

4.2.Analytic results

In this section, we discuss the research questions listed previously in the paper that are addressed in the final selection of 22 articles. According to the introduction in the methodology section, the selection range of the articles is from 2022 to 2024, and the selected articles should discuss artificial intelligence/ChatGPT security or privacy. This part is the analysis results and is very important for understanding the topic and research questions. Because the analysis results directly demonstrate the value and significance of the research, they are the empirical basis on which research questions or hypotheses are answered, proving the necessity and importance of the research. Additionally, the results of the analysis provide support for the arguments presented in the paper. Without solid data analysis, the dissertation's claims may appear hollow. By analysing the results, we can better understand the data and findings of the study. At the same time, this helps readers evaluate the validity of the study and the practical application of the findings.

Regarding the topic of this dissertation, we raised the following two research questions:

1.What are ChatGPT's main challenges and opportunities in the context of privacy protection?
2.What measures can ChatGPT implement to ensure the protection of user information?

Based on these two research questions, we sorted out the 22 articles that were finally screened, and we came to the following results:

ChatGPT's main opportunities in the context of privacy protection

1.Enhanced privacy measures: Artificial intelligence models like ChatGPT, especially carefully designed machine learning algorithms and deep learning networks, have powerful data processing capabilities and can be trained to identify and delete sensitive information, thereby helping to protect user privacy on multiple levels (Khalid et al., 2023; McIntosh et al., 2023; Wu et al., 2023; Pasupuleti et al., 2023; Negri-Ribalta et al., 2024; Sai et al., 2024; Prasad et al., 2023; Gupta et al., 2023; Hu & Chen, 2023). Artificial intelligence (AI) models have the capacity to learn from vast data sets in order to recognise particular patterns and attributes of sensitive data, including financial, health, and personally identifiable information. (Khalid et al., 2023). At the same time, when processing a huge amount of data, AI tools like ChatGPT can automatically classify the data and distinguish sensitive data from other data so that different processing measures can be taken (Wu et al., 2023). In addition, for data that needs to be shared or made public, using AI models can implement data desensitization techniques. For example, replacing, masking, or generalising personal identifiers to protect personal privacy. Of course, artificial intelligence (AI) can help with the implementation of access control policies while people tracking and recording data. It is to guarantee that only authorised users can access sensitive data. (Ouyang et al., 2022).

2.Customised privacy policy: ChatGPT's AI technology can assist in creating personalised privacy policies and guidelines that match individual user preferences. In other words, in order to understand the user's privacy preferences, an AI model similar to ChatGPT can track and analyse users' adjustments to privacy settings in different situations. In addition, the AI technology used in ChatGPT can dynamically adjust privacy policies based on users' real-time feedback and behavioural changes. It aims to ensure that they always meet users' personal needs (Khalid et al., 2023; Wu et al., 2023; Pasupuleti et al., 2023; Negri-Ribalta et al., 2024; Sai et al., 2024; Alawida et al., 2023; Prasad et al., 2023; Gupta et al., 2023). This AI technology can provide suggestions to help users enjoy a customised service experience. At the same time, it can protect users' personal privacy.

In general, artificial intelligence technologies can provide users with more personalised privacy services through these functions. They can also help users keep their privacy rights while enjoying digital services. In this point of view, it is an opportunity for ChatGPT.

3.Education and awareness raising: ChatGPT can be used to teach its users and inform them about privacy rights. It can also be used to teach users how to protect their data online. That is to say, ChatGPT can help its users understand the basic concepts and importance of privacy by providing tutorials and information (M et al., 2024; NegriRibalta et al., 2024; Sai et al., 2024; Prasad et al., 2023; Goldstein et al., 2023). In addition, as privacy regulations continue to change, AI can help users stay updated with the latest legal requirements. Also, AI can help users make sure that their privacy settings meet the law (Hu & Chen, 2023). This point is another opportunity for ChatGPT.

4.Automated compliance: AI can help organisations make sure privacy regulations by monitoring data flows and recommending necessary actions. Specifically, the AI technology used in ChatGPT can conduct Privacy Impact Assessment (PIA). It can also predict the potential impact of data processing activities on personal privacy (Khowaja et al., 2024; Goldstein et al., 2023). This is very convenient for ChatGPT users or organisations. It can also provide training modules to increase employees' awareness of privacy regulations and reduce the risk of broken rules (Gupta et al., 2023).

ChatGPT's main challenges under privacy protection

1.Data collection：The AI models used by ChatGPT often require large amounts of data to train, which may raise concerns about how user data is collected and used. This model may require collecting data from a variety of sources, including online behavioural data, social media interactions, transaction records, etc. The types of data may include text, images, audio and video (M et al., 2024; Pasupuleti et al., 2023; Sai et al., 2024; Alawida et al., 2023; Gupta et al., 2023; Derner & Batistič, 2023). In addition, to lower the risk of privacy leaks, AI researchers and developers, taking ChatGPT as an example, usually use data anonymisation technologies, such as data desensitisation, differential privacy, etc., to protect user identities from being identified (M et al., 2024; Negri-Ribalta et al., 2024; Alawida et al., 2023; Hu & Chen, 2023).
Since the technology is not particularly popular yet, it is a relatively big challenge for ChatGPT.

2.Inadvertent disclosure: If not managed properly, ChatGPT's language model can inadvertently disclose personal information during interactions. A more detailed explanation is that when processing queries containing sensitive information, if the ChatGPT model is not properly trained and constrained, it may leak the user's personal information, such as name, address, phone number, etc. (Pasupuleti et al., 2023; Voelkel & Willer, 2023). Additionally, users may share private details in conversations with models, expecting to be treated confidentially. If the model fails to properly handle this

information, it may violate the user's privacy (Wu et al., 2023; Negri-Ribalta et al., 2024; Ouyang et al., 2022; Gupta et al., 2023; Derner & Batistič, 2023). These should be developed by ChatGPT.

3.Abuse of AI: ChatGPT and other similar chat systems use a series of advanced artificial intelligence technologies. They include machine learning and deep learning (McIntosh et al., 2023). These technologies can provide useful information and advice for people. However, they can lead a range of privacy and security risks. For example, bad actors may use chat system technology to induce users to provide sensitive information (Pasupuleti et al., 2023; Voelkel & Willer, 2023). In addition, ChatGPT may be used for social attacks to gain users' trust. It may attack their users' system for personal information (Yao et al., 2024; Wu et al., 2023; Pasupuleti et al., 2023; Voelkel & Willer, 2023; Gupta et al., 2023; Derner & Batistič, 2023). What's more, the text generated by the ChatGPT may be used to create fake content. For example, fake news or fake identities, which may have serious results for users and society (Li et al., 2023).

4.Transparency and explainability: The communication and strategic progress between ChatGPT and its users may be unclear. It makes difficult for ChatGPT users to know how their data are used, which is also an important aspect for privacy protection. To explain this, it is often unclear for its users when ChatGPT generating answers through the internal algorithms. This means that ChatGPT users may not fully understand how the model processes their requests (Wu et al., 2023; Pasupuleti et al., 2023; Morales et al., 2024; Gupta et al., 2023; Derner & Batistič, 2023). In addition, ChatGPT may collect users' preferences and behavioural data to provide personalised services and suggestions. This step may cause its users feel conflicted between enjoying a personalised experience and protecting personal privacy. Especially during the training process, ChatGPT may use a large amount of data, including its user-generated content (Li et al., 2023; Zhang et al., 2024; Ferrag et al., 2024). As a result, users of ChatGPT may lack understanding of how this data is used for model training. They may have questions about whether their data will be accessed or shared with third parties. All these points are challenges for ChatGPT privacy protection.

5.Breach of trust: It may lead to the loss of user trust if the ChatGPT system is unstable. This is because a system breach may result in sensitive user data being catched by unauthorised third parties. For example, personal information, communication records and personal preferences (Pasupuleti et al., 2023; Zhang et al., 2024; Ferrag et al., 2024). In addition, ChatGPT users' trust in its system is based on its security and privacy protection capabilities (Wu et al., 2023; Pasupuleti et al., 2023). Once the system is compromised, ChatGPT users may lose confidence in the entire platform or even similar technologies. In order to prevent these risks, relevant personnel of ChatGPT must adopt a security strategy, which is also very big challenge for ChatGPT (Yao et al., 2024; M et al., 2024; Alawida et al., 2023; Prasad et al., 2023; Khowaja et al., 2024; Goldstein et al., 2023; Hu & Chen, 2023; Zhang et al., 2024; Ferrag et al., 2024).

Measures that ChatGPT can implement to ensure the protection of user information

1.Data encryption: It cannot be read because ChatGPT user data is encrypted during transmission and storage to ensure that even if the data is intercepted. Specifically, it is necessary to have regular security audits and have evaluations of encryption measures to ensure that encryption algorithms and key management strategies comply with current security standards (M et al., 2024; Sai et al., 2024; Alawida et al., 2023; Prasad et al., 2023; Khowaja et al., 2024; Li et al., 2023; Hu & Chen, 2023).

2.Access control: Implement strict access control policies to ensure that only authorised person have access to user data. Also, the access is logged and monitored. Specifically, it is to determine that ChatGPT's access control policy complies with relevant laws, regulations and industry standards. For example, the standard GDPR. Or ChatGPT should conduct regular security training for employees who use it to increase their awareness of the importance of data protection. Besides, they can also educate them on how to handle data securely (Pasupuleti et al., 2023; Khowaja et al., 2024; Ferrag et al., 2024).

3.Anonymisation: ChatGPT related personnel can remove personal identifiers from the data used to train and improve ChatGPT. They can also ensure user-specific information not be traced back to an individual. Besides, before the data is used for model training, ChatGPT can also through technical means to reduce privacy risk. For example, the replacement, masking or generalisation, remove or modify personal identification information in the data. During this process, ChatGPT can ensure that the anonymisation step complies with relevant privacy protection laws and regulations, such as the anonymisation requirements in GDPR (Pasupuleti et al., 2023; Alawida et al., 2023; Khowaja et al., 2024; Voelkel & Willer, 2023; Li et al., 2023; Hu & Chen, 2023; Ferrag et al., 2024). By taking some comprehensive methods, the privacy rights of ChatGPT users can be effectively protected. While providing the required data support for the training and improvement of ChatGPT's AI model, ChatGPT can achieve a balance between the reasonable utilisation of its user data and privacy protection.

4.Data minimisation: ChatGPT can only use the minimum amount of the data which is required to complete the task. Also, the time to keep the task should not be too long. ChatGPT can set storage periods for different types of data. Once the data exceeds its validity period or is no longer needed, they will be deleted or anonymised (Wu et al., 2023; Sai et al., 2024). Moreover, ChatGPT takes appropriate security measures during data processing to make sure its users' information been protected. For example, encryption and security protocols. They aim at preventing data from being leaked or misused during processing (Wu et al., 2023; Sai et al., 2024; Khowaja et al., 2024; Li et al., 2023). By implementing these measures, ChatGPT can make sure that data collection meets its business needs. It can also minimise potential risks to the privacy of its users.

5.Privacy by design: It is a fundamental aspect of the technology that ChatGPT's AI system take privacy considerations into the design step. It can make sure the privacy protection. In other words, ChatGPT gradually designs an interface that is easy for its users to understand and operate. Also, it can allow users to easily manage their privacy settings, including data collection sharing. Moreover, ChatGPT provides its users with a clear privacy policy and data processing instructions. It allows users to understand how their data is processed and protected by the AI system (Wu et al., 2023; Khowaja et al., 2024). Through these measures, the ChatGPT system can consider user privacy at every stage of design and implementation. It's also a way of protecting users' privacy.

6.Regular check: ChatGPT's AI system can conduct regular security check to identify and resolve potential security vulnerabilities. It will find and give response to the secure problem according to the complexity of the system (Morales et al., 2024; Alawida et al., 2023).

7.User agreement: ChatGPT's data capture should follow their user agreement, making sure users understand how their data will be used. More important among this is data sharing and transferring. ChatGPT can inform their users and get their agreement before sharing or transferring data to third parties (Alawida et al., 2023; Prasad et al., 2023; Goldstein et al., 2023). At the same time, ChatGPT needs to inform its users when there are changes to the privacy policy (Pasupuleti et al., 2023; Voelkel & Willer, 2023).

8.Transparency: ChatGPT aims to be fully transparent with users on what data practices we follow — both concerning how the data is acquired, used and safeguarded. ChatGPT explains exactly what data is being collected (what kind of data, how often it will be stored and under which conditions). At the same time, informing users in detail about how data is used to provide them with various services; improve user experience, and conduct research and analysis, etc. (Wu et al., 2023; Khowaja et al., 2024; Li et al., 2023; Hu & Chen, 2023; Ferrag et al., 2024).

9.Incident response plan: By adding an advanced monitoring service to capture data for ChatGPT. Therefore, ChatGPT is able to catch it if something were to behave unexpectedly (Alawida et al., 2023; Prasad et al., 2023; Goldstein et al., 2023).

10.Legal Compliance: ChatGPT AI systems are developed to comply will all applicable data protection regulations. Users should trust ChatGPT to ensure the right methods are used when data is transferred between countries (Wu et al., 2023; Pasupuleti et al., 2023; Morales et al., 2024; Alawida et al., 2023; Khowaja et al., 2024; Goldstein et al., 2023; Voelkel & Willer, 2023; Zhang et al., 2024; Ferrag et al., 2024).

11.User control: The ability of users to access their data as well as correct or delete it when using ChatGPT. They can erase the data once they are done with it or permanently when user objects to processing of their personal data (Pasupuleti et al., 2023; Sai et al., 2024; Prasad et al., 2023; Khowaja et al., 2024; Hu & Chen, 2023; Ferrag et al., 2024).

12.Security updates: ChatGPT will keep system updates regularly to make sure user data security and system stability. These updates include, fixing for known security vulnerabilities, improved encryption technology and improved user authentication processes (Morales et al., 2024; Alawida et al., 2023). Through these measures, ChatGPT can respond to new and emerging security threats. In addition, ChatGPT can monitor the latest network security trends and attack methods. It can quickly respond and take corresponding protective measures to make sure that its users' privacy and information security are protected (Morales et al., 2024; Alawida et al., 2023; Khowaja et al., 2024; Goldstein et al., 2023; Hu & Chen, 2023; Ferrag et al., 2024).

5.Discussion

5.1.Theoretical background

We adopted Prospect Theory as an analytical framework to discuss the research topic. The reason why we use this theory is to gain a deeper understanding of users' behavioral decisions when they face privacy risks and benefits. Also, Prospect theory, proposed by psychologists Daniel Kahneman and Amos Tversky, highlights people's nonlinear psychological awareness when evaluating potential losses and gains (Levy, 1992). This theory provides me with a unique perspective to find how users make balances between the risk of privacy leakage. Also, it shows us the convenience brought by using ChatGPT.

The discussion part of this dissertation will be based on the results of 22 systematic literature reviews. We will use prospect theory to explain users' attitudes and behaviors towards privacy protection. At the same time, we will analyse how users perceive the privacy protection measures provided by ChatGPT and explore how these affect their awareness for privacy risks. In addition, we will discuss how ChatGPT uses prospect theory to optimise the user decision-making process.

We aim to find ChatGPT's opportunities in privacy protection through this discussion. We want to explore how to improve user trust by designing privacy protection policies. We will also explore the challenges that may be faced when implementing these strategies. Our goal is to provide a more comprehensive privacy protection framework for ChatGPT.

The result indicates the significant potential shown by ChatGPT for privacy protection, while also raising some key challenges. In discussing these findings, we will focus on how technological advances provide new opportunities for privacy protection. We will explore how existing research can provide guidance for ChatGPT's privacy protection. Also, we will discuss the applicability of these in different cultural and legal contexts.

In conclusion, We hope to provide the academic and practical communities with a comprehensive perspective on the privacy protection issues of ChatGPT. We also hope to provide valuable insights for future research and practice. Our goal is to promote an indepth understanding of the privacy protection of ChatGPT and promote the development of related technologies to ensure that the security and privacy of user data are protected to the greatest extent.

5.2.Discussion of results under prospect theory

ChatGPT's main opportunities in the context of privacy protection

When discussing ChatGPT's opportunities for privacy protection, We draw on prospect theory to analyse how users perceive and respond to privacy protection measures. Prospect theory emphasises people's decision-making behaviour in the face of potential losses (such as privacy leaks) and potential gains (such as personalised services) (Levy, 1992). Similar to the analysis of research results section, We still discuss from four perspectives:

1.Enhanced privacy measures: According to prospect theory, users generally prefer avoiding losses over gaining gains of equal value. Therefore, ChatGPT can significantly reduce users' concerns about privacy leaks by enhancing privacy measures, such as using machine learning algorithms to identify and delete sensitive information. This loss reduction strategy is consistent with users' risk aversion psychology, thereby increasing users' trust and satisfaction with ChatGPT.

Besides, the result resonates with the research by Dwivedi et al. (2019), who emphasised the importance of machine learning algorithms in identifying and handling sensitive information. However, my study goes further to suggest that AI model like ChatGPT's application in data classification and de-identification techniques may be more effective than traditional methods, a point that has not been fully explored in the existing literature.

2.Customised privacy policy: Prospect theory states that individuals' perceptions of options influence their decisions. ChatGPT allows users to adjust privacy settings according to their own preferences. It does this by providing customised privacy policies, which not only meets users' needs for control, but also reflects the emphasis on user choice. This personalised experience can enhance users' sense of gain, thereby increasing their loyalty to the platform.

Similarly, Gerasimou and Limniotis (2024) discussed the necessary of personalised privacy settings. My research extends this viewpoint by demonstrating that users' demands for personalised privacy policies are closely related to their privacy preferences and behavioral patterns.

3.Education and awareness raising: Combined with Prospect theory, users' understanding and awareness of privacy rights often affect their acceptance of privacy protection measures (Acquisti et al., 2020). ChatGPT helps raise users' awareness of privacy by educating users, providing real-time interaction. This kind of education can reduce users' uncertainty about privacy protection, thereby building positive expectations in users' minds. It is consistent with the perspective of reducing decisionmaking uncertainty in prospect theory.

Our findings align with Li et al. (2023), who argued that education is key to raising user privacy awareness. And I further discovered that real-time interaction and feedback provided by AI platforms like ChatGPT can more effectively educate its users. This is also a point that is less frequently mentioned in existing literature.

4.Automated compliance: Prospect theory also highlights the reference point effect in the decision-making process. That is, how users evaluate outcomes based on a certain reference point. Also, ChatGPT helps its users and organisations comply with privacy regulations by automating compliance measures. For example, conducting privacy impact assessments and monitoring data flows in real time. Thereby setting a reference point for compliance. This compliance not only reduces potential legal risks, but also provides its users with security and trust.

Our research meets Smith et al. (2024) in recognising the role of automated tools in making sure privacy regulation compliance. Also, my study further explores the potential of AI technology in conducting privacy impact, which is an important addition to the existing literature.

In summary, the opportunities of ChatGPT in privacy protection can be further understood through prospect theory. By reducing users' concerns about privacy leaks, ChatGPT can not only meet users' needs for privacy protection, but also build a positive image in users' minds. Under these methods, ChatGPT can improve user trust and satisfaction. These measures meet the perspectives of choice perception, education and compliance in prospect theory. They provide strong theoretical support for the future development of ChatGPT. By comparing my research with existing literature, I not only validate current theories but also provide new perspectives and insights into the study of privacy protection. The results highlight the potential of AI technology in privacy protection and point out areas that require further research.

ChatGPT's main challenges in the context of privacy protection

We can analyse how users perceive potential privacy risks when discussing the privacy protection challenges of ChatGPT through prospect theory. Prospect theory says that individuals evaluate potential losses differently than gains. It provides us with a framework to understand user behaviour under privacy (Levy, 1992). So, I discuss from the following perspectives based on the analysis results section before:

1.Models like ChatGPT needs good volumes of data for training. It may raise user concerns about data collection and management. Prospect theory posits that people might weigh the costs of privacy invasion due to data collection more heavily than the benefits resulting from enhanced AI performance (Liao et al. 2020). It may not be widely understood because the use of data anonymisation technology is critical. This

technology will also lead to a trust gap. ChatGPT increasing transparency and clearly explaining data practices to users can help solve this perceived loss and build customer trust.

2.The potential disclosure of personal information during interactions with ChatGPT can be seen as a loss in prospect theory (Lily et al., 2023). And if the data is leaked, as users trust them with a high degree of privacy protection. This loss can potentially be avoided by ensuring that ChatGPT is well trained to sequester sensitive information. It is also helpful to keep the user trust.

3.While AI could equally be used for malicious purposes — as a way to mislead users into revealing their private data, deploy social engineering attacks. It is a real loss aversion play for users. It is also possible that the abusiveness is subject to limited scrutiny because prospect theory predicts users are more loss-averse with respect to losses of such abuse as compared to gains in AI (Jhala et al., 2019). Hence, there is need for robust security mechanisms and user education to prevent these losses while retaining that true AI value.

4.Through the view of prospect theory, the decision-making process of ChatGPT can be seen as a privacy risk. Users may feel a loss of control over their data if they cannot understand how their data is being used (Kanbach et al., 2023). It can also help users perceive less loss and feel more in control if ChatGPT give efforts to provide clear information about data usage.

5.It is considered an important negative outcome in prospect theory when the system of ChatGPT brocken. It means that ChatGPT will lose a lot of user. (Remountakis et al., 2023). So, it's important to implement a security strategy to prevent this loss.

The study refers the work of Liu et al. (2021), who listed the importance of system security in maintaining user trust. My research further indicates that security strategies and emergency response plans are important to preventing crises of confidence. It is a point emphasised in existing literature.

The results of the dissertation showed consistency with the existing literature in several aspects. These verified the reliability of the study and the applicability of the existing theory. However, we also found some differences, which may be explained by several factors:

1.The development of technology: New algorithms and applications are constantly emerging with the rapid development of AI technology. This may cause my research results to be different from earlier literature in some aspects.

2.Changes in user behaviour: User attitudes and behaviours toward privacy may change over time and circumstances. This may lead to differences in my findings from past research.

3.Differences in research methods: Different studies may use different methodologies, which may affect the consistency of my research results.

Our research not only verifies existing theories, but also provides new perspectives and insights for ChatGPT research by comparing with existing literature. The results of the discussion also highlight the potential of AI technology in privacy protection. It points out areas that need further exploration in future research.

In summary, by addressing these challenges through enhancing transparency, strengthening security measures, and conducting user education, ChatGPT can strive to minimise losses and establish a stronger foundation of trust with its users.

What measures can ChatGPT take to ensure the protection of user information?

When discussing the measures taken by ChatGPT in terms of privacy protection, we can analyse the impact of these measures on its users. Specifically, these measures manifest in the following aspects:

1.Using data encryption to lower the perceived losses that users of ChatGPT could face (in prospect theory terms). This is because data security with use of encryption by itself reduces the chance of a given hack turning into genuine informational collections (Gundu, 2023). It may give users a perception of security since they are encrypted, which might alleviate privacy breach concerns among ChatGPT user base.

2.ChatGPT users can trust in access control policies of each time. They do this by mitigating unauthorised access risks (Al-Hawawreh et al., 2023).

3.By doing anonymisation, user data of ChatGPT is no longer tied to individual identities. This system does not overlap with any of the problems related to privacy violations which ChatGPT users may have. This is considered as loss aversion in prospect theory (Binhammad et al., 2024).

4.The data minimisation rule reduces users' perception of privacy invasion, avoiding unnecessary amount of collected information. This can be seen as privacy protective by users and relates to the theories of prospect theory (Luo et al., 2021).

5.The improvement of privacy protection should be achieved by the construction and design of ChatGPT itself. Because this shows that users' interests have been considered by developers when designing the product (Baldassarre et al., 2020).

6.Regular security checks can increase user confidence in the ChatGPT system under prospect theory. This is because it shows that ChatGPT is actively looking for and resolving potential security issues (Chowdhury et al., 2023).

7.Respecting ChatGPT user consent according to the prospect theory would be an important aspect involved in reducing how much amount of perceived privacy invasion that users are exposed to. It also makes users feel like they are in control of how their data is being used.

8.Demonstrating transparency in light of the prospect theory can alleviate uncertainty about privacy related risks with ChatGPT users. That is because transparency gives the users to know how their data was used and kept safe.

9.Someone who uses ChatGPT might be less worried about data breaches if they know there is a response plan in place, according to the prospect theory. This is because the plan ensures a way to react instantly and reduce further damages.

10. Adherence to data protection laws should result in greater user confidence that their information is being protected by law. In turn, this means ChatGPT fulfills external norms and requirements (Wang et al. 2023).

11. Users remain in charge of their data and can know what is being done on with it—what they see, correct or delete. With this policy, ChatGPT can increase user satisfaction by protecting their privacy. It also can be a control mechanism. Simultaneously, it is also connected to the prospect theory (Liang & Xue 2009).

12. Regular security updates can reduce user concerns about emerging security threats. This shows that ChatGPT has been putting efforts to keep user data safe and secure.

In general, the measures taken by ChatGPT in terms of privacy protection are consistent with the principles of prospect theory. These measures are designed to reduce users' awareness of privacy risks. Moreover, they also increase users' confidence and satisfaction with privacy protection. Through these ways, ChatGPT can protect user privacy and maintaining user trust and satisfaction. However, we need to note that although these measures significantly reduce the risk of ChatGPT user information, no system can guarantee absolute security.

6.Conclusion, limitations, and future research

6.1.Conclusion

The dissertation has provided an analysis of the opportunities and challenges in ChatGPT privacy protection. It has been built through a systematic literature review way. It shows that ChatGPT offers significant potential for improving privacy measures and it also faces challenges. For example, data collection concerns, abuse of AI and transparency issues.

Additionally, prospect theory enhanced the comprehension of both behaviour and decision-making processes of users. ChatGPT can promote trust and satisfaction by making users less risk-averse and more self-confident. The policies proposed, such as data encryption and privacy by design come naturally from the insights of prospect theory.

To sum up, ChatGPT under privacy protection direction has many opportunities and significant challenges. It should balance the technologies innovations and user privacy rights in its development stage. With a focus on privacy preservation, ChatGPT can still evolve into one of the most reliable platforms in AI.

6.2.Limitations

However, there are some major limitations of this study. First, the research is limited to studies available from 2022–2024. It is not compiling the most recent developments in the field. It also has to be pointed out that using other research to do so makes the dissertation liable as whatever bias and viewpoints of past literature exist. The study was also limited to English-language materials, potentially missing other important input from non-English outlets.

Furthermore, the application of prospect theory as an analytical framework may not fully capture the complexity of all user behaviours and motivations regarding privacy. The study also does not account for cultural differences in privacy perceptions and decision-making, which could influence the result of the findings.

6.3.Future research

Future research can build through this dissertation, which may include:

1. Expanding the scope of literature to include more recent publications and nonEnglish sources. This is to have a more comprehensive and global perspective on ChatGPT's privacy protection.
2. Conducting empirical studies to assess user perceptions and behaviours regarding privacy in ChatGPT directly, which can provide primary data to complement the secondary data analysed in this review.
3. Investigating the impact of cultural differences on privacy attitudes and decision-making in the context of AI chat platforms like ChatGPT.
4. Exploring the long-term effects of privacy protection measures on user trust and the adoption of AI technologies.
5. Assessing the effectiveness of different privacy protection strategies and their implications for the design and implementation of AI systems.
6. Examining the role of policy and regulation in shaping privacy protection in AI, considering the rapid evolution of technology and the legal landscape.

By pursuing these methods, future research can contribute to a deeper understanding of the intricate relationship between ChatGPT, AI, privacy, and user trust, ultimately informing the development of more effective privacy protection strategies in the era of advanced artificial intelligence.

7.References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. Journal of Consumer Psychology, 30(4), 736–758.

Al-Amin, M., Ali, M. S., Salam, A., Khan, A., Ali, A., Ullah, A., ... & Chowdhury, S. K. (2024). History of generative Artificial Intelligence (AI) chatbots: past, present, and future development. *arXiv preprint arXiv:2402.05122.*

Alawida, M., Mejri, S., Mehmood, A., Chikhaoui, B., & Abiodun, O. I. (2023). A Comprehensive study of CHATGPT: Advancements, limitations, and ethical Considerations in natural language processing and Cybersecurity. *Information*, *14*(8), 462. https://doi.org/10.3390/info14080462

Al-Hawawreh, M., Aljuhani, A., & Jararweh, Y. (2023). Chatgpt for cybersecurity: practical applications, challenges, and future directions. *Cluster Computing*, *26*(6), 3421–3436. https://doi.org/10.1007/s10586-023-04124-5

Ali, O., Murray, P., Momin, M., & Al-Anzi, F. S. (2023). The knowledge and innovation challenges of ChatGPT: A scoping review. *Technology in Society, 102402.*

Alsharida, R. A., Al-rimy, B. A. S., Al-Emran, M., & Zainal, A. (2023). A systematic review of multi perspectives on human cybersecurity behavior. *Technology in society, 73, 102258.*

Alvesson, M., & Kärreman, D. (2011). Qualitative Research and Theory Development: *Mystery as method*. https://doi.org/10.4135/9781446287859

Ayinde, L., Wibowo, M. P., Ravuri, B., & Emdad, F. B. (2023). ChatGPT as an important tool in organizational management: A review of the literature. *Business Information Review, 40(3), 137-149.*

Baldassarre, M. T., Barletta, V. S., Caivano, D., & Scalera, M. (2020). Integrating security and privacy in software development. *Software Quality Journal*, *28*(3), 987–1018. https://doi.org/10.1007/s11219020-09501-6

Binhammad, M., Alqaydi, S., Othman, A., & Abuljadayel, L. H. (2024). The role of AI in Cyber Security: Safeguarding Digital identity. *Journal of Information Security*, *15*(02), 245–278. https://doi.org/10.4236/jis.2024.152015

Charfeddine, M., Kammoun, H. M., Hamdaoui, B., & Guizani, M. (2024). ChatGPT's Security Risks and Benefits: Offensive and Defensive Use-Cases, Mitigation Measures, and Future implications. *IEEE Access*, 1. https://doi.org/10.1109/access.2024.3367792

Choudhury, A., & Shamszare, H. (2023). Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *Journal of Medical Internet Research*, *25*, e47184. https://doi.org/10.2196/47184

Chowdhury, M. M., Rifat, N., Ahsan, M., Latif, S., Gomes, R., & Rahman, M. S. (2023). ChatGPT: A Threat Against the CIA Triad of Cyber Security. *IEEE*. https://doi.org/10.1109/eit57321.2023.10187355

Derner, E., & Batistič, K. (2023). Beyond the safeguards: exploring the security risks of ChatGPT. *arXiv preprint arXiv:2305.08005.*

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., . . . Williams, M. D. (2019).

Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Edelman(2019). Edelman Trust Barometer Special Report: *In Brands We Trust.* https://www.edelman.com/research/trust-barometer-special-report-in-brands-we-trust

Fabio Duarte (2024). *OpenAI's model history.*Explodingtopics. https://explodingtopics.com/blog/chatgpt-users

Fabio Duarte (2024). *Time taken to reach 1 million users.*Explodingtopics. https://explodingtopics.com/blog/chatgpt-users

Falade, P. V. (2023). Decoding the threat landscape : ChatGPT, FraudGPT, and WormGPT in social engineering attacks. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 185–198. https://doi.org/10.32628/cseit2390533

Ferrag, M. A., Alwahedi, F., Battah, A., Cherif, B., Mechri, A., & Tihanyi, N. (2024). Generative AI and Large Language Models for Cyber Security: All Insights You Need. *arXiv preprint arXiv:2405.12750.*

Gerasimou, S., & Limniotis, K. (2024). A study on privacy and security aspects of personalised apps. *International Journal of Information Security*. https://doi.org/10.1007/s10207-024-00887-z

Gill, S. S., & Kaur, R. (2023). ChatGPT: Vision and challenges. *Internet of Things and Cyber-Physical Systems*, *3*, 262–271. https://doi.org/10.1016/j.iotcps.2023.05.004

Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., & Sedova, K. (2023). Generative language models and automated influence operations: Emerging threats and potential mitigations. *arXiv preprint arXiv:2301.04246*.

Goyal, S., & Miyapuram, K. P. (2019). Feedback influences discriminability and attractiveness components of probability weighting in descriptive choice under risk. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00962

Gundu, T. (2023). Chatbots: A Framework for Improving Information Security Behaviours using ChatGPT. In IFIP advances in information and communication technology (pp. 418–431).

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. IEEE Access, 11, 80218–80245.

Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks Standards and Evaluations*, *2*(4), 100089. https://doi.org/10.1016/j.tbench.2023.100089

Hu, C., & Chen, J. (2023). A Dimensional Perspective analysis on the cybersecurity Risks and Opportunities of ChatGPT-Like Information Systems. *IEEE*. https://doi.org/10.1109/nana60121.2023.00061

Jhala, K., Natarajan, B., & Pahwa, A. (2019). Prospect Theory-Based active consumer behavior under variable electricity pricing. *IEEE Transactions on Smart Grid*, *10*(3), 2809–2819.

https://doi.org/10.1109/tsg.2018.2810819

Kahneman, D., & Tversky, A. (1988). Prospect theory: An analysis of decision under risk. In *Cambridge University Press eBooks*(pp. 183–214). https://doi.org/10.1017/cbo9780511609220.014

Kanbach, D. K., Heiduk, L., Blueher, G., Schreiter, M., & Lahmann, A. (2023). The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. *Review of Managerial Science*. https://doi.org/10.1007/s11846-023-00696-z

Khalid, N., Qayyum, A., Bilal, M., Al-Fuqaha, A., & Qadir, J. (2023). Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Computers in Biology and Medicine, 158, 106848.*

Khowaja, S. A., Khuwaja, P., Dev, K., Wang, W., & Nkenyereye, L. (2024). Chatgpt needs spade (sustainability, privacy, digital divide, and ethics) evaluation: A review. *Cognitive Computation, 1-23.*

Klag, M., & Langley, A. (2012). Approaching the conceptual leap in qualitative research. *International Journal of Management Reviews*, *15*(2), 149–166. https://doi.org/10.1111/j.1468-2370.2012.00349.x

Levy, J. S. (1992). An introduction to prospect theory. *Political Psychology*, *13*(2), 171–186. https://www.jstor.org/stable/3791677

Li, H., Guo, D., Fan, W., Xu, M., Huang, J., Meng, F., & Song, Y. (2023). Multi-step jailbreaking privacy attacks on ChatGPT. *arXiv*. https://doi.org/10.18653/v1/2023.findings-emnlp.272

Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023). ChatGPT in Education: A Discourse analysis of worries and concerns on social media. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12256-9

Liang, H., & Xue, Y. (2009). Avoidance of information Technology Threats: A Theoretical perspective. MIS Quarterly, 33(1), 71. https://doi.org/10.2307/20650279

Liao, G., Chen, X., & Huang, J. (2019). Prospect Theoretic Analysis of Privacy-Preserving Mechanism. IEEE/ACM Transactions on Networking, 28(1), 71–83.

Lily, A. E. A., Ismail, A. F., Abunaser, F. M., Al-Lami, F., & Abdullatif, A. K. A. (2023). ChatGPT and the rise of semi-humans. *Humanities and Social Sciences Communications*, *10*(1). https://doi.org/10.1057/s41599-023-02154-3

Liu, Y., Gan, Y., Song, Y., & Liu, J. (2021). What influences the perceived trust of a Voice-Enabled smart Home system: an empirical study. *Sensors*, *21*(6), 2037. https://doi.org/10.3390/s21062037

Luo, K., Dang, S., Shihada, B., & Alouini, M. (2021). Prospect Theory for Human-Centric Communications. Frontiers in Communications and Networks, 2.

M, K., Kumar, A., Krishnasamy, L., & Sarveshwaran, V. (2024). Investigation on Preserving Privacy of Electronic Medical Record using Split Learning. *Procedia Computer Science*, *233*, 614–622. https://doi.org/10.1016/j.procs.2024.03.251

McIntosh, T. R., Liu, T., Susnjak, T., Watters, P., Ng, A., & Halgamuge, M. N. (2024). A culturally sensitive test to evaluate nuanced GPT hallucination. *IEEE Transactions on Artificial Intelligence*, 1–13. https://doi.org/10.1109/tai.2023.3332837

McIntosh, T., Liu, T., Susnjak, T., Alavizadeh, H., Ng, A., Nowrozy, R., & Watters, P. (2023). Harnessing GPT-4 for generation of cybersecurity GRC policies: A focus on ransomware attack mitigation. *Computers & Security*, *134*, 103424. https://doi.org/10.1016/j.cose.2023.103424

Morales, G., C, P. K., Jahan, S., Hosseini, M. B., & Slavin, R. (2024). A large language model approach to code and privacy policy alignment. *IEEE*. https://doi.org/10.1109/saner60148.2024.00016

Naghiyev, K. (2024). ChatGPT from a data protection perspective. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4818860

Negri-Ribalta, C., Geraud-Stewart, R., Sergeeva, A., & Lenzini, G. (2024). A systematic literature review on the impact of AI models on the security of code generation. *Frontiers in Big Data*, *7*. https://doi.org/10.3389/fdata.2024.1386720

Okey, O. D., Udo, E. U., Rosa, R. L., Rodríguez, D. Z., & Kleinschmidt, J. H. (2023). Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis. *Computers & Security*, *135*, 103476. https://doi.org/10.1016/j.cose.2023.103476

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in neural information processing systems, 35, 27730-27744.*

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Pasupuleti, R., Vadapalli, R., & Mader, C. (2023). Cyber security issues and challenges related to generative AI and ChatGPT. *IEEE*. https://doi.org/10.1109/snams60348.2023.10375472

Prasad, S. G., Sharmila, V. C., & Badrinarayanan, M. (2023). Role of artificial intelligence based Chat Generative Pre-trained Transformer (CHATGPT) in cyber security. *IEEE*. https://doi.org/10.1109/icaaic56838.2023.10141395

Ray, P. P. (2023). Benchmarking, ethical alignment, and evaluation framework for conversational AI: Advancing responsible development of ChatGPT. *BenchCouncil Transactions on Benchmarks Standards and Evaluations*, *3*(3), 100136. https://doi.org/10.1016/j.tbench.2023.100136

Remountakis, M., Kotis, K., Kourtzis, B., & Tsekouras, G. E. (2023). Using ChatGPT and persuasive technology for personalized recommendation messages in hotel upselling. *Information*, *14*(9), 504. https://doi.org/10.3390/info14090504

Roumeliotis, K. I., & Tselikas, N. D. (2023). Chatgpt and open-ai models: A preliminary review. *Future Internet, 15(6), 192.*

Sai, S., Yashvardhan, U., Chamola, V., & Sikdar, B. (2024). Generative AI for Cyber Security: Analyzing the potential of ChatGPT, DALL-E and other models for enhancing the security space. *IEEE Access*, *12*, 53497–53516. https://doi.org/10.1109/access.2024.3385107

Saini, M., & Shlonsky, A. (2012). *Systematic synthesis of qualitative research*. https://doi.org/10.1093/acprof:oso/9780195387216.001.0001

Smith, M., Torres-Agüero, A., Grossman, R., Sen, P., Chen, Y., & Borcea, C. (2024). A study of GDPR compliance under the Transparency and Consent Framework. *In Proceedings of the ACM on Web Conference 2024 (Pp. 1227-1236)*. https://doi.org/10.1145/3589334.3645618

Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, *74*, 102700. https://doi.org/10.1016/j.ijinfomgt.2023.102700

Templier, M., & Paré, G. (2015). A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems*, *37*. https://doi.org/10.17705/1cais.03706

Voelkel, J. G., & Willer, R. (2023). Artificial intelligence can persuade humans on political issues.

Wang, Y., Pan, Y., Yan, M., Su, Z., & Luan, T. H. (2023). A survey on ChatGPT: AI–Generated Contents, Challenges, and Solutions. *IEEE Open Journal of the Computer Society*, *4*, 280–302. https://doi.org/10.1109/ojcs.2023.3300321

Wu, X., Duan, R., & Ni, J. (2023). Unveiling security, privacy, and ethical concerns of ChatGPT. *Journal of Information and Intelligence*. https://doi.org/10.1016/j.jiixd.2023.10.007

Xu, N. L., Jiang, N. C., Wang, N. J., Yuan, N. J., & Ren, N. Y. (2014). Information security in big data: privacy and data mining. *IEEE Access*, *2*, 1149–1176. https://doi.org/10.1109/access.2014.2362522

Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., & Zhang, Y. (2024). A survey on Large Language Model (LLM) security and privacy: The Good, The Bad, and The Ugly. *High-Confidence Computing*, 100211. https://doi.org/10.1016/j.hcc.2024.100211

Zaman, S. (2023). ChatGPT security risks and solutions. *7th IET Smart Cities Symposium*. https://doi.org/10.1049/icp.2024.0955

Zhang, J., Bu, H., Wen, H., Chen, Y., Li, L., & Zhu, H. (2024). When llms meet cybersecurity: A systematic literature review. *arXiv preprint arXiv:2405.03644.*

Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *European Journal of Education, 59(2), e12599.*