Mitigating Chatbots AI Data Privacy Violations in the Banking Sector: A Qualitative Grounded Theory Study

John Giordani
Department of Information Technology, University of Fairfax, Salem, VA, USA

ABSTRACT:
This research study examines the impact of Artificial Intelligence (AI) data poisoning on data privacy violations in AI-enabled banking chatbots, employing a qualitative approach grounded in AI, data privacy, and cybersecurity theories. Through qualitative grounded theory research approach, viewpoints were gathered from a group of IT professionals in the banking sector. The research uncovered the impact of AI data poisoning across different professional roles, ranging from direct breaches to indirect exposure. Key findings revealed a spectrum of mitigation strategies, from technical solutions to basic awareness and mixed responses regarding the impact on personally identifiable information (PII), underscoring the complexity of safeguarding customer data [1]. Despite potential limitations stemming from the rapidly evolving AI landscape, this study contributes valuable insights into effective strategies for mitigating AI data poisoning risks and enhancing the security of AI-enabled chatbots in banking. It highlights the critical importance of developing robust security measures to protect sensitive customer data against privacy violations.

Keywords: Chatbot, Artificial Intelligence, customer satisfaction, banking, AI poisoning, risks, bias, data security.

List of abbreviations: AI - Artificial Intelligence; NLP - Natural Language Processing; PII - Personally Identifiable Information; AGI - Artificial General Intelligence; IS - Information Systems; IDS - Intrusion Detection Systems.


INTRODUCTION
This research study is about Artificial Intelligence (AI) and related technologies, like chatbots, focusing on the integration of Chatbots in the banking sector, the risks associated with malicious data or AI data poisoning, and the implications for data privacy. As AI becomes increasingly prevalent in banking operations, it is essential to understand the significance of exploring the intersection of AI applications, such as Chatbots, with customer service in the banking industry and the potential vulnerabilities these technologies face from malicious data inputs [2]. It elaborates on how Chatbots, empowered by advanced natural language processing (NLP) techniques, can enhance customer service by providing automated, efficient responses to inquiries without human intervention. The study also highlights the vulnerability of these systems to data poisoning attacks that can compromise data privacy and security.

The problem statement articulates the core issue under investigation: the impact of AI data poisoning on data privacy violations within Chatbot systems used in the banking sector. The purpose statement further clarifies the research's goal to identify best practices for mitigating AI data poisoning, particularly focusing on protecting personally identifiable information (PII) handled by Chatbots in customer service interactions.
The research questions aim to explore effective strategies to mitigate data privacy violations caused by AI data poisoning in AI-enabled Chatbot systems within the banking sector. The propositions set forth aim to understand the relationship between user trust, data privacy concerns, and AI data poisoning, exploring perceptions and feelings of users and professionals in the field [3].

The research methodology of this study is qualitative, relying on surveys and grounded theory to glean insights from IT professionals and end-users about the experiences and perceptions related to AI enabled Chatbots, data poisoning, and data privacy concerns.

Research on the reasons for and consequences of this impact has focused on objective measures of automation’s impact on employment or job loss [4]. However, little work has explored the potential challenges posed by AI data poisoning and data privacy in the contest of AI-enabled Digital Banking Chatbots [5].

**Topic Overview/Background**

Chatbots in the banking industry refer to computer programs or AI applications designed to interact with customers and provide automated responses and support for various banking-related inquiries and services [6]. These Chatbots are integrated into digital platforms, such as websites, mobile apps, or messaging applications, to offer a convenient and efficient customer service experience.

Malicious Data (also known as Poisoned Data) in the realm of AI applied to Chatbots refers to intentionally manipulated or corrupted input information strategically designed to deceive, disrupt, or compromise the normal functioning of the Chatbot system [7]. This malicious data is meticulously crafted to exploit vulnerabilities within the AI model’s training process, inference mechanisms, or decision-making algorithms, with the intent to trigger unintended and potentially harmful responses or actions from the Chatbot. Malicious data encompasses various forms, such as biased content, misleading queries, or contextually distorted input, and is deliberately injected into the training dataset to deceive the AI model [8].

Examples of malicious or poisoned data in Chatbot systems include the Italian government temporarily blocking ChatGPT in April 2023, citing privacy concerns [9]. AI generally comes with its own set of malicious, unintended disruptions and security privacy concerns. In March 2023, OpenAI’s ChatGPT Chatbot was temporarily disabled due to a bug that exposed the payment information of some users [10]. The bug affected 1.2% of active ChatGPT Plus subscribers during a specific nine-hour window. The exposed payment information included the first and last name, email address, payment address, last four digits of a credit card number, and credit card expiration date [11].

The concept of Chatbots encompasses sophisticated natural language processing (NLP), which is a field of computer science that deals with the interaction between computers and human (natural) language capabilities [12]. NLP allows Chatbots to understand and interpret user queries in a human-like manner. NLP techniques such as keyword extraction, intent recognition, and sentiment analysis are used to help Chatbots understand user queries [13]. Keyword extraction identifies the keywords in a query, which can help the Chatbot determine the query's intent. Intent recognition determines the type of action that the user wants the Chatbot to take, such as providing information, answering a question, or completing a transaction. Sentiment analysis determines the emotional tone of the user’s query, which can help the Chatbot respond more appropriately [14].

Once a Chatbot has understood the user’s query, it can then provide a response, answer the question, complete the transaction, or offer personalized assistance. These tasks can be completed without human intervention, making Chatbots a valuable tool for businesses and organizations [15]. Chatbots can provide responses, answer frequently asked questions, handle basic transactions, and offer personalized assistance, all without human intervention.

Natural language Processing (NLP) enables computers to extract written and spoken language information, facilitating tasks like translation and speech recognition [16]. Examples of NLP applications include translating text between languages or enabling voice-activated commands for devices such as smartphones with NLP capabilities. Other examples are NLP empowering Chatbots to understand and process written text and even powering Global Positioning Systems (GPS) to provide audible directions. The key areas of focus in NLP are text processing, speech recognition, and speech synthesis [13].
In the banking industry, Chatbots are part of the broader trend of using AI and automation to streamline customer interactions and improve service. Chatbots can assist customers with account inquiries, balance checks, transaction history, fund transfers, bill payments, and other routine banking tasks [17]. Advanced Chatbots can handle more complex inquiries, such as loan applications and investment advice, by leveraging machine learning algorithms and accessing vast amounts of customer data [18].

Implementing Chatbots in the banking sector aims to enhance customer engagement, reduce response times, and optimize operational efficiency. By providing round-the-clock support and instant responses, Chatbots contribute to a seamless customer experience, fostering loyalty and satisfaction [19]. Additionally, AI enabled Chatbots efficiently manage a wide array of customer inquiries, lightening the load on human staff and providing valuable insights into customer preferences, further refining banking services.

Regarding the adoption of AI enabled chatbots to be applied in the financial industry, only a few studies examined the acceptance of these applications in the context of the banking sector [20]. Although several studies have examined the factors influencing the acceptance of AI-enabled Chatbots, the findings carried out in different fields may not be transferable to financial services [21]. Specific research is required regarding AI data poisoning of Chatbots in the banking sector.

Data poisoning refers to intentionally injecting incorrect or misleading data into the system’s training data. Malicious data poisoning involves intentionally including fraudulent, biased, or contextually distorted examples, aiming to compromise the Chatbot’s integrity, customer trust, and the overall security of banking operations [22]. Malicious actors may attempt to manipulate the Chatbot’s responses or gain unauthorized access to personally identifiable information (PII) such as credit card numbers [23]; this can have far-reaching consequences, potentially leading to the Chatbot providing inaccurate financial advice, disclosing sensitive customer information, or facilitating fraudulent transactions.

The definition of AI has been a topic of extensive debate, leading to a lack of universal consensus. Over the past few decades, various definitions of AI have emerged. AI is the science and engineering of creating intelligent machines, particularly intelligent computer programs. It is connected to the goal of understanding human intelligence, but it is not limited to biologically observable methods [24]. A more comprehensive definition describes AI as the activity dedicated to making machines intelligent [25].

The challenge with defining AI solely as making machines intelligent is that it does not clearly explain what AI truly entails and what constitutes an intelligent machine [26]. Another way is to describe AI as the ability of a digital computer or computer-controlled robot to perform tasks typically associated with intelligent beings [27]. This includes reasoning, discovering meaning, generalizing, and learning from past experiences. In simple terms, AI refers to the intelligence exhibited by machines, programmed to imitate human actions and later execute activities commonly associated with human minds, such as problem-solving, learning, and performing physical tasks [26].

Furthermore, there are distinctions between Strong AI (also known as general AI or artificial general intelligence - AGI) and Weak AI (also referred to as narrow AI). Strong AI aims to emulate human functions, including reasoning, planning, and problem-solving, to create machines indistinguishable from the human mind [28]. On the other hand, Weak AI focuses on a specific task, such as answering questions or playing chess based on user input. It can handle only one type of activity at a time, while Strong AI is designed to handle a wide range of tasks [29]. Narrow AI relies on human intervention to define learning algorithms’ parameters and provide appropriate training data, eventually enabling it to address new problems. Strong AI, as it progresses, does not require human input and has the capacity to teach itself how to solve new issues.

Today, banks have made the decision to focus on and invest in AI techniques and algorithms, recognizing their importance for future success. According to a survey conducted by Deloitte, which aimed to assess the significance of AI in organizational success over the next five years, a substantial 94% of respondents stated that AI would be either very or critically important [30]. However, it is essential to move beyond the hype surrounding AI and gain a comprehensive understanding of its implications for businesses.
While some banks still view AI as a means for experimenting with certain products and processes, others have taken these experiments to a large scale. They are harnessing intelligent tools to automate back-office tasks, such as Know-Your-Customer (KYC) procedures, credit scoring decisions, and even communication channels through Chatbots. In today’s world, data has become the most valuable resource, surpassing even oil. This represents a significant opportunity for banks due to the need for access to detailed purchasing records from each client [31]. Analyzing this data can provide valuable insights into default and credit risk, cross-selling opportunities (e.g., suggesting insurance products after a car purchase), and numerous other applications.

**Problem Statement**

The problem statement that this research worked on is about AI data poisoning in AI-enabled Chatbots, often leading to data privacy violations [5]. Particularly in the context of Chatbots utilized in customer service interactions requesting and potentially mishandling personally identifiable information (PII) such as credit card numbers [32].

The rapid advancement of AI has led to the widespread adoption of AI-enabled Chatbots systems in various industries, including the financial sector [21]. These AI-enabled Chatbots serve as virtual assistants, providing customer support, processing transactions, and handling sensitive information. However, the increasing use of AI in customer service Chatbots has raised concerns about potential security risks, particularly related to AI data poisoning and its implications for data privacy violations.

AI data poisoning poses a significant threat, often leading to data privacy violations. In particular and within the context of customer service, AI-enabled Chatbots that handle personally identifiable information (PII), such as credit card numbers, pose additional risks to private personal information [33].

The consequences of AI data poisoning are far-reaching. If an AI-enabled Chatbot is deceived into accepting false information, it may lead to data breaches. For instance, when AI-enabled Chatbots are utilized to collect personally identifiable information (PII), such as credit card numbers, Social Security Numbers (SSN), Date of Birth (DOB), or account numbers, false information can be exploited to gain unauthorized access to this sensitive data [34]. Additionally, accepting false information could enable fraudulent activities, such as using AI-enabled Chatbots to collect PII, like Social Security numbers, to open fraudulent accounts or make unauthorized purchases. This can seriously affect data privacy and security, as sensitive information may be mishandled or used maliciously [35]. This study aims to address this problem and find ways to mitigate data privacy violations caused by AI data poisoning in AI-enabled Chatbots systems.

**Purpose Statement**

The purpose of this research is to identify best practices to mitigate AI data poisoning in AI-enabled Chatbots, often leading to data privacy violations [5], particularly in the context of AI-enabled Chatbots utilized in customer service interactions requesting and potentially mishandling personally identifiable information (PII) such as credit card numbers [32]. This study targeted IT professionals securing the banking sector’s AI-enabled Chatbots. This study used qualitative grounded theory methodology.

AI data poisoning is characterized by intentionally injecting false information into an AI-enabled Chatbot, causing it to accept inaccurate data as valid [32]. This poses significant threats to data privacy and security, potentially resulting in the mishandling or malicious use of sensitive information. This can seriously affect data privacy and security, as sensitive information may be mishandled or used maliciously. With the escalating integration of AI technologies in banking operations [36], examining the potential consequences of AI poisoning on customer information is imperative.

By delving into the security aspects of communication with AI-enabled Chatbots and addressing the risks of AI data poisoning, this research endeavors to contribute valuable insights to the field of AI security, data privacy, and customer service in the banking sector [37]. Through rigorous analysis and examination of data privacy violations caused by AI data poisoning, this research seeks to propose mitigation strategies to safeguard sensitive information and bolster customer trust in AI-driven banking interactions. This study aims to identify practical implications for banking institutions.
to enhance their AI-enabled Chatbots deployment and reinforce the protection of customer data, thus contributing to the advancement of AI-enabled Chatbots adoption in the financial domain [36].

The research design for this study is qualitative, focusing on the use of open-ended survey questions for data collection [38]. The population and sample for this research included professionals engaged in developing, managing, or utilizing AI-enabled Chatbots within the banking industry, and end-users who interact with banking platforms. The sampling procedures aimed to ensure that the sample is representative of the wider population of professionals and experts in the field.

**Research Question**

How to mitigate data privacy violations caused by AI data poisoning in AI-enabled Chatbot systems that mishandle personally identifiable information (PII) like credit card numbers, social security numbers (SSN), date of birth (DOB), or account numbers?

Using the qualitative grounded theory research method, the research attempts to address this question and provide recommendations and effective strategies to mitigate data privacy violations caused by AI data poisoning in AI-enabled Chatbots in the banking sector and safeguard sensitive customer data.

**Propositions**

The research aimed to explore the perceptions and feelings of users concerning 'user trust' and 'data privacy concerns' in the context of AI-enabled Chatbots in the banking industry [35]. The intention is to understand if heightened concerns about data privacy are associated with diminished trust in using these AI-enabled Chatbots. This relationship is analyzed qualitatively, critically evaluating responses to determine patterns, themes, and overarching sentiments.

The propositions form the core of the question guiding the study focus on the impact of AI data poisoning in AI-enabled Chatbots systems applied to the digital banking sector. The central research question delves into how to mitigate data privacy violations caused by AI data poisoning in Chatbots handling personally identifiable information (PII), such as credit card numbers, social security numbers (SSN), Date of Birth (DOB) [39].

This study proposed several exploratory questions to delve into the relationship between AI data poisoning and data privacy violations of AI-enabled Chatbots. Questions included: "How does AI data poisoning affect users' perceptions of data privacy in AI-enabled Chatbots systems within the digital banking sector?" or "How do users perceive the mishandling of personally identifiable information (PII), like credit card numbers, social security numbers (SSN), date of birth (DOB), or account numbers, in relation to AI data poisoning?" The extent to which data privacy concerns arise might be tied to the perceptions and reported experiences of AI data poisoning incidents. Furthermore, this research explored perceptions regarding the efficacy of mitigation strategies targeting AI data poisoning in preserving data privacy in AI-enabled Chatbots systems.

Propositions are presented to support these exploratory questions and provide a deeper understanding [40]. These propositions detail the perceived transformational effect of AI-enabled Chatbots in the banking sector, apprehensions about data privacy violations due to AI data poisoning, the nature and understanding of AI data poisoning, and the potential consequences of data mishandling by AI-enabled Chatbots from the users' viewpoint. From a user-centric perspective, ethical considerations are highlighted as a significant factor in AI development and deployment within the digital banking sector [41].

While this research aimed to provide rich and nuanced insights into the perceptions around AI-enabled Chatbots and data poisoning, it acknowledges potential limitations [42]. Constraints might arise from the subjective nature of qualitative data and scope, potentially influencing the depth and breadth of the study. Despite these limitations, the proposed propositions offered a robust framework for analyzing and interpreting the study's findings. The research hopes to enrich the understanding of strategies to mitigate perceived risks associated with AI-enabled Chatbots data poisoning and ensure customer data privacy in digital banking through the users' lens.
Theoretical Perspectives and Conceptual Framework

This research on the impact of AI-enabled Chatbots data poisoning on data privacy violations in AI-enabled Chatbots systems applied to the digital banking sector draws upon theoretical perspectives from the fields of AI, data privacy, and cybersecurity. These perspectives serve as the foundation for understanding the underlying principles and concepts that inform the study's design. Theoretical perspectives include AI and machine learning principles, data privacy ethics and narratives, AI-enabled Chatbots system experiences, and customer service sentiments [20]. These perspectives are essential for grasping the potential risks, perceptions, and vulnerabilities associated with AI data poisoning and its experiential impact on data privacy violations in the digital banking sector. This research looked into the nuances of mitigation strategies and their perceived effectiveness. It captured the essence of AI-enabled Chatbots applications in the digital banking sector, emphasizing customer service interactions and their associated trust dynamics [42].

Overall, the theoretical perspectives and conceptual framework provide a robust foundation for understanding the research design and inform the qualitative exploration into the intricacies of AI data poisoning's experiential impact on data privacy violations in AI-enabled Chatbots systems within the digital banking sector [43]. This research is situated within the broader field of AI, data privacy, and cybersecurity. It zeroes in on the experiential impact of AI data poisoning on data privacy violations in AI-enabled Chatbots systems within the digital banking sector. While building on existing studies that touch on the security of AI-enabled Chatbots in digital banking, it pioneers the qualitative examination of the phenomenon of AI data poisoning, an area not extensively probed from an experiential standpoint in the digital banking context [44].

This research also augments extant studies on customer sentiments in the banking sector by understanding how data privacy narratives related to AI-enabled Chatbots potentially sculpt customer trust and contentment and offers a fresh lens to the ongoing dialogue on data privacy ethics and narratives, especially vis-à-vis AI and machine learning technologies [43]. In terms of methodology, the research adopts a qualitative approach centered around surveys and analyzing survey responses using grounded theory. This approach ensures a deep dive into the lived experiences and perceptions of the specific population - IT professionals safeguarding the banking sector's AI-enabled Chatbots [45].

Overall, this research harmoniously coexists with and robustly enhances the existing research corpus in the field. This study discerns the transformative touch of technology in the banking world, with an accentuated interest in AI's role in shaping customer relationships, privacy perceptions, and challenges [43].

This research respects the monumental shift towards digital platforms in banking, further catalyzed by global phenomena like the COVID-19 pandemic. It leans on the theoretical scaffolding of AI and machine learning principles, data privacy ethics, AI-enabled Chatbots system narratives, and customer interactions, rooted in both time-tested and contemporary literature [46]. This research also underlines the vitality of navigating the legal and ethical terrains in chalking out efficacious data protection tactics and guaranteeing alignment, a central motif in the ongoing scholarly discussions. Besides, the research fathoms the intricacies and confines tethered to cybersecurity, a perpetual topic in classical and emergent literature [47].

Significance of the Study

This research on AI-enabled Chatbots data poisoning, resulting in data privacy violations in AI-enabled Chatbots systems within the digital banking sector, holds significant value for various stakeholders, including financial institutions, customers, policymakers, and the broader research community [5]. This research aimed to provide recommendations and contributions to the body of knowledge related to AI, data privacy, and cybersecurity in the context of digital banking and AI-enabled Chatbots applications.

Financial institutions, especially those employing AI-enabled Chatbots systems for customer service interactions, stand to benefit from the findings of this research. Understanding the potential risks and consequences of AI-enabled Chatbots data poisoning on data privacy violations will help develop robust mitigation strategies and safeguards [1]. Implementing effective measures to protect sensitive
customer information, enhance data security, and reinforce customers’ trust in digital banking services [21].

Ultimately, this knowledge can contribute to a safer and more secure digital banking experience [4]. Policymakers and regulatory bodies responsible for overseeing the digital banking sector and AI technologies will find value in the research findings. The study sheds light on the potential challenges AI data poisoning poses in AI-enabled Chatbots systems, highlighting the importance of developing appropriate mitigation guidelines [5]. Policymakers can use these insights to create comprehensive frameworks that address data privacy concerns and ensure responsible AI use in the financial sector.

For the broader research community, this study contributes to the growing body of knowledge on AI, cybersecurity, and data privacy [32]. The research provides empirical evidence and new perspectives on the impact of AI data poisoning on data privacy violations, particularly in the context of customer service AI-enabled Chatbots in digital banking. The findings may serve as a basis for further exploration and future studies in related fields, enriching the understanding of AI security and privacy issues.

As AI continues to evolve and be integrated into various sectors, including finance, the insights gained from this research can inform the development and implementation of AI systems in other applications [48]. Lessons learned from studying the mitigation of risks and vulnerabilities of AI data poisoning in AI-enabled Chatbots can help shape best practices and guidelines for ensuring data privacy and security in future AI-driven technologies.

Assumptions and Biases

Certain assumptions and potential biases are acknowledged in conducting this research on the impact of AI data poisoning on data privacy violations in AI-enabled Chatbots systems within the digital banking sector. This study assumed the AI-enabled Chatbots systems adhere to standard data privacy and security protocols. The existence of AI data poisoning incidents within the AI-enabled Chatbots systems is assumed, and the selected mitigation recommendations are presumed to be relevant and effective [1].

The researcher’s prior knowledge and experience in AI, data privacy, and cybersecurity could influence the study’s focus and conclusions. The availability of published literature on the topic could create publication bias. Financial institutions providing data might have industry biases, and underreported incidents could create reporting bias. Time bias may limit the study’s ability to capture the evolving impacts of AI-enabled Chatbots data poisoning.

Delimitations

This research is about mitigation recommendations for AI-enabled Chatbots data poisoning. It does not implement or test specific strategies in AI-enabled Chatbots systems [33]. The focus remains on theoretical exploration. This study concentrates on the digital banking sector, and the impact of AI-enabled Chatbots data poisoning on data privacy violations in other sectors has not been extensively explored. The research identified associations between AI-enabled Chatbots data poisoning and data privacy violations but does not establish causality between the variables. Ethical considerations related to AI-enabled Chatbots [49], data poisoning, and data privacy are acknowledged, but the in-depth ethical analysis is beyond the research’s primary scope. Considering these delimitations helps interpret the research findings accurately and understand the context in which the study’s conclusions are applicable.

Limitations

This research study acknowledges certain limitations that may impact the research results. The specific context and sample characteristics may constrain the generalizability of findings. Additionally, the rapidly evolving nature of AI technology and AI-enabled Chatbots may introduce certain uncertainties and potential biases [30]. However, these limitations were addressed through rigorous data collection, analysis, and interpretation to ensure the reliability and validity of the research outcomes.
Definition of Terms

The following terms are widely used in the study and discussion of AI data poisoning in AI-enabled Chatbots, often leading to data privacy violations:

**AI-enabled Chatbots**: A conversational agent utilizes AI to interpret chat text using natural language processing (NLP). Instead of directly communicating with human service personnel, the customer/client can have a conversation via text or voice using the chatbot software [50].

**Digital banking**: All banking services that are carried out with the help of technology, especially the Internet [51].

**Personally Identifiable Information (PII)**: Personally identifiable information is any data that can be used to identify a specific individual. Social Security numbers, mailing or email addresses, and phone numbers are commonly considered PII [52].

**Cyber-attack**: A malicious activity that attempts to collect, disrupt, deny, degrade, or destroy information system resources or the information itself [53].

**Data Privacy**: The protection of personal information from unauthorized access, use, disclosure, disruption, modification, or destruction to ensure confidentiality, integrity, and availability [54].

**AI Data Poisoning**: A form of cyber attack where adversaries intentionally manipulate the training data of an AI system, leading to incorrect or biased outcomes by the model [55].

**Chatbots in Banking**: Automated software tools banks use to simulate conversations with users, typically via text or voice, to handle customer inquiries or perform routine banking tasks without human intervention [20].

**Natural Language Processing (NLP)**: A branch of Artificial Intelligence that helps computers understand, interpret, and respond to human language in a useful way [2].

**Machine Learning (ML)**: A subset of AI that involves using statistical techniques to enable computers to improve at tasks with experience, specifically through the analysis and interpretation of data [15]

**Personal Identifiable Information (PII)**: Any data that could potentially identify a specific individual, such as names, addresses, social security numbers, or any other information that can be used on its own or with other sources to identify, contact, or locate a person [56].

**Cybersecurity**: The practice of protecting systems, networks, and programs from digital attacks aimed at accessing, changing, or destroying sensitive information, extorting money from users, or interrupting normal business processes [53].

**Compliance and Regulation in AI**: The adherence to laws, regulations, guidelines, and specifications relevant to Artificial Intelligence systems, particularly those that protect data privacy and ensure ethical use of AI [53].

**Ethical AI**: The practice of designing, developing, and deploying AI systems in a manner that conforms to widely accepted ethical standards and principles, such as transparency, fairness, and accountability [57].

**AI Transparency**: The quality of an AI system that is clear about its processes, decisions, and actions, making it understandable to humans and capable of providing explanations for its behavior and outputs [58].

**AI Bias**: Systematic and repeatable errors in a computer system that create unfair outcomes, such as privileging one arbitrary group of users over others, often reflecting existing prejudices in society [59].

**General Overview of the Research Design**

This study targeted information technology professionals safeguarding the banking sector’s AI-enabled Chatbots. By opting for a qualitative research method, the study primarily revolved around using surveys designed to capture the intricate experiences, perceptions, and challenges these professionals face [60].
The qualitative approach was specifically chosen to allow participants to delve deep into their experiences, facilitating a richer understanding of the phenomenon. The survey instrument was comprised of open-ended questions to encourage participants to share detailed narratives about their experiences with AI enabled Chatbots, data poisoning, and related data privacy concerns [60]. By analyzing the survey responses, the study aimed to derive themes, patterns, and insights regarding the mitigation strategies and challenges associated with AI data poisoning in AI-enabled Chatbots systems [1]. The qualitative nature of the study allows for a deeper exploration of the nuances, attitudes, beliefs, and potential solutions that might not be captured in a strictly numerical or quantitative analysis [61].

The collected data was then subjected to rigorous thematic analysis, where responses were coded, categorized, and synthesized to identify prevailing trends and sentiments. This approach provided a comprehensive view of the IT professionals’ stance on the issue and offered actionable insights and recommendations for the digital banking sector.

Given the topic's sensitive nature, special emphasis was placed on ensuring the confidentiality and anonymity of the participants. This was vital in obtaining candid and genuine feedback, which is paramount in qualitative research [62].

This study's qualitative approach, focusing on surveys and in-depth analysis of participants’ responses, facilitated a profound understanding of the intricacies of Artificial Intelligence (AI) data poisoning and its repercussions on data privacy in digital banking AI-enabled Chatbots.

**REVIEW AND DISCUSSION OF THE LITERATURE**

Within the context of the banking sector, the literature highlights numerous potential applications of AI, including:

1. **Customer Service and Engagement**: AI-enabled Chatbots have emerged as widely used tools across various industries, offering substantial cost savings. In the banking sector, these AI-enabled Chatbots can provide mini-statements, facilitate transfers, and respond to balance inquiries [5].

2. **Robo-Advisory Services**: AI-driven robo-advisors analyze users’ financial data and history to offer tailored recommendations in specific financial domains [63].

3. **Predictive Analytics**: Leveraging natural language understanding capabilities, AI can identify patterns and correlations within data, uncovering opportunities and insights.

4. **Cybersecurity**: AI can enhance system security by learning from past threats and attacks, monitoring internal breaches, and proposing corrective actions [64].

5. **Credit Scoring and Direct Lending**: AI can assess an individual’s or entity’s creditworthiness by analyzing a wide range of traditional and non-traditional data, thus strengthening credit scoring models, even for those with limited credit histories.

Additionally, the literature highlights how AI is contributing to the competitiveness of banks through various means, including:

- **Enhanced Customer Experience**: AI uses historical data to understand customers’ behavior better [65].
- **Prediction of Future Outcomes and Trends**
- **Cognitive Process Automation**: AI aids in achieving a return on investment and reducing costs.
- **Effective Decision-Making**: AI systems mimic human thinking and responses.
- **Robotic Automation of Processes**: This allows for the automation of up to 80% of repetitive tasks, freeing users to focus on more complex and value-added work [19].
For instance, new literature has emerged discussing the increasing sophistication of AI data poisoning attacks and their potential to bypass existing security measures [55]. This necessitates a more in-depth examination of advanced mitigation strategies and the development of robust AI systems resistant to data poisoning [1].

Furthermore, recent studies have highlighted the growing concern among customers about data privacy in digital banking, emphasizing the need for financial institutions to prioritize data protection [66]. This underscores the importance of research in identifying effective strategies to prevent data privacy violations.

Lastly, the literature review also includes recent regulatory developments related to AI and data privacy [67]. Understanding the legal and regulatory landscape is crucial in formulating effective data protection strategies and ensuring compliance.

**Conceptual Framework**

This research on the impact of AI enabled Chatbots data poisoning on data privacy violations in AI-enabled Chatbots systems applied to the digital banking sector draws upon theoretical perspectives from AI, data privacy, and cybersecurity. These perspectives are the foundation for understanding the underlying principles and concepts that inform the study’s design.

Theoretical perspectives include AI and machine learning principles, data privacy laws and regulations, AI-enabled Chatbots system design, and customer service interactions [20]. These perspectives are essential for comprehending the potential risks and vulnerabilities associated with AI data poisoning and its impact on data privacy violations in the digital banking sector.

Overall, the theoretical perspectives and conceptual framework provide a comprehensive foundation for understanding the research design and inform the investigation into AI data poisoning’s impact on data privacy violations in AI-enabled Chatbots systems within the digital banking sector [32]. This study builds on previous research by detailing the conceptual framework, as shown in Figure 1.

**Historical Perspective of Artificial Intelligence and AI-Enabled Chatbots**

The concept of AI has a long history dating back to its initial conceptualization in 1955 within the realm of computer science [68]. Its primary goal was to create intelligent machines capable of emulating human cognitive abilities such as problem-solving and continuous learning [69]. Significant investments from enterprises and organizations have been channeled into AI research and prototype development to address real-world problems that current technology struggles to solve.

![Figure 1. Conceptual Framework](image-url)
AI-enabled Chatbots have found applications in the financial sector, assisting customers with various financial transactions such as account reviews, reporting lost cards, making payments, policy renewals, and refunds [70]. In recent literature, numerous studies have explored the utilization of chatbot technology in the financial industry [71, 20, 22, 23, 37, 72] (see Table 1).

[71] conducted a study on the adoption and diffusion of AI-enabled Chatbots within the insurance context, and the study findings indicated that most participants were familiar with the technology and preferred its use at the beginning of the advisory process. However, one-third of the participants rejected the adoption of AI-enabled Chatbots.

[20] examined customer attitudes toward AI-enabled Chatbots in the banking sector, the study revealed a positive correlation between a favorable attitude toward Chatbots and their perceived utility, accessibility, and associated risks.

[23] studied AI-enabled Chatbots efficiency in Singapore's banking sector, focusing on alignment with customer expectations, the study found comprehensive information to be the top priority for users, followed by swift responses, functionality, interactivity, ease of use, and data protection. Some users were unhappy with AI-enabled Chatbots' response delays.

[22] delved into the acceptance factors of chatbot technology in the Indonesian banking sector, specifically among Millennials, employing the Technology Acceptance Model (TAM). The study identified innovativeness, perceived usefulness, perceived ease of use, and attitudes toward using AI-enabled Chatbots as significant factors influencing behavioral intention.

[37] examined the customer experience of banking AI-enabled Chatbots and its impact on brand loyalty. Findings showed system quality as the top predictor of customer experience, followed by information and service quality. Perceived risks lessened these effects, but a positive AI-enabled Chatbots experience boosted brand loyalty.

[72] found that AI-enabled Chatbots enhance customer service in banking. Key factors include advisory services, ease of use, convenience, cost-effectiveness, customer-friendliness, customization, responsiveness, trustworthiness, and customer security.

### Table 1. Summary of Studies Regarding Banking AI-Enabled Chatbots

<table>
<thead>
<tr>
<th>Authors</th>
<th>Aim of the study</th>
<th>Theories/Studied variables</th>
<th>Research method/sample</th>
<th>Data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>[71]</td>
<td>Adoption and diffusion of chatbots in the German insurance sector</td>
<td>DOI: advantages, compatibility, complexity, trialability, observability. TOE: management support, IS setup, costs, environmental threats, competition, networks. TAM: usefulness, ease of use, behavioral control</td>
<td>Qualitative and quantitative methods, semi-structured expert interviews, and web-based surveys</td>
<td>Descriptive</td>
</tr>
<tr>
<td>[20]</td>
<td>Analysis of customers’ attitude towards the chatbots in banking industry</td>
<td>Observed utility includes ease of use, process simplicity, and customer engagement. Observed accessibility emphasizes basic transaction ease, speed, and user-friendliness. Observed threat and awareness cover data security, social awareness, and usage by peers</td>
<td>Quantitative Data collected via Facebook and WhatsApp</td>
<td>Bivariate analysis</td>
</tr>
<tr>
<td>[23]</td>
<td>Analysis of the effectiveness of the current use of chatbots in Singapore’s banking industry</td>
<td>user experience: response rate, functionality and usability satisfaction: interactivity, informative, data privacy, and protection</td>
<td>Quantitative, qualitative Interviews Qualitative user tests</td>
<td>Descriptive</td>
</tr>
<tr>
<td>[22]</td>
<td>Analysis of the factors that influence millennial’s technology acceptance of chatbots in the</td>
<td>TAM attitude towards usage, behavioral intention innovativeness, perceived usefulness, perceived ease of use</td>
<td>Quantitative Simple random sampling technique</td>
<td>SEM</td>
</tr>
<tr>
<td>Reference</td>
<td>Title</td>
<td>Methodology</td>
<td>Approach</td>
<td>Findings</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>[37]</td>
<td>Examination of customer experience of using banking chatbots and its impact on brand love in Indonesia</td>
<td>Information Systems (IS) success model customer experience, brand love system quality, information quality, service quality, perceived risk</td>
<td>Quantitative Online questionnaire sent to Gen Z individuals</td>
<td>SEM</td>
</tr>
<tr>
<td>[72]</td>
<td>Examination of the role of chatbots in customer service in the banking industry of Bangladesh</td>
<td>Customer service, advisory, ease of use, convenience, cost-efficiency, customer-friendliness, customization, relationship banking, responsiveness, trustworthiness, value, and security</td>
<td>Quantitative Judgment sampling method</td>
<td>Regression</td>
</tr>
</tbody>
</table>

**Note:** Table 1 shows the need for ongoing research into effective mitigation strategies and regulatory compliance [1].

### Key Concepts in Artificial Intelligence

AI is a multifaceted field encompassing fundamental concepts, particularly when applied to AI-enabled Chatbots within the banking sector. At its core, AI refers to the simulation of human intelligence processes by machines, enabling them to perform tasks that typically require human cognitive functions [68]. Several key concepts merit profound exploration in the context of AI-enabled Chatbots in banking.

First and foremost, risk management is a paramount principle. AI-enabled Chatbots, while offering efficiency and convenience, introduce potential risks. These encompass both operational and ethical dimensions [37]. Risk management in AI-enabled Chatbots involves identifying, assessing, and mitigating these risks. Operational risks pertain to system malfunctions, data breaches, or inadequate responses, necessitating robust fault tolerance and contingency planning strategies. Ethical risks encompass bias, discrimination, and privacy infringement. Thus, a meticulous understanding of these risks is imperative for effective AI governance [73].

Threat modeling is another pivotal concept. This process entails understanding potential adversarial actions and developing countermeasures to thwart them [74]. Threat modeling helps to fortify AI-enabled Chatbots against malicious intent, ensuring they operate securely in a dynamic environment.

Data privacy exploitation is a pertinent issue in the context of AI-enabled Chatbots as these systems gather and process vast user data to deliver personalized experiences. However, this practice necessitates stringent data privacy safeguards [54]. Concepts such as informed consent, data anonymization, and user-centric control over data access are vital to balance personalization and privacy.

Data poisoning is a sophisticated threat vector in AI-enabled Chatbots. It involves injecting malicious data into the training datasets to manipulate the AI-enabled Chatbots' behavior [32]. Understanding this concept is crucial to fortify AI-enabled Chatbots against adversarial manipulation. Robust data validation, anomaly detection, and continuous monitoring are essential to counter data poisoning.

### Artificial Intelligence Threat Landscape

The contemporary landscape of AI and Chatbots presents a dynamic and multifaceted arena characterized by the perpetual evolution of technology and the concurrent emergence of novel threats [36]. AI-enabled Chatbots, harnessed for their efficiency and automation in various domains, including customer service, financial institutions, and healthcare, have become a focal point of cyber adversaries. Understanding this evolving threat landscape is critical to mitigating the vulnerabilities inherent in these AI systems [5].

One notable facet of the threat landscape is the continuous development of adversarial AI techniques. It encompasses a spectrum of attacks, including adversarial examples that manipulate input data to mislead AI systems and model inversion attacks, which exploit system outputs to infer sensitive information about users [75]. As AI-enabled Chatbots rely on machine learning models, safeguarding against adversarial attacks has become an imperative research area.
AI-enabled Chatbots inherently necessitate access to vast datasets for training and operation, which, if not handled with the utmost care, can result in data breaches and privacy infringements [54]. The risks associated with data exposure, unauthorized access, and data poisoning attacks are becoming increasingly prevalent and complex.

The proliferation of AI enabled Chatbots also raises concerns about the ethical implications of their interactions. Bias and discrimination in AI responses stemming from biases in training data can lead to unjust or offensive outcomes [57]. Consequently, mitigating bias, ensuring fairness, and fostering ethical behavior within AI-enabled Chatbots systems are pivotal concerns in the contemporary landscape. Adopting AI-enabled Chatbots within critical sectors, such as healthcare [76] and finance, makes it an attractive target for cybercriminals. Threat actors are increasingly focusing on these sectors, seeking to exploit vulnerabilities in AI-enabled Chatbots for financial fraud, identity theft, or even healthcare fraud [77]. As these AI-enabled Chatbots handle sensitive data and financial transactions, robust security measures are imperative to safeguard users and institutions.

**Artificial Intelligence Standards and Frameworks**

The landscape of AI standards and frameworks has emerged as a critical facet of the AI ecosystem, offering a structured approach to address the multifaceted challenges inherent in AI development, deployment, and governance [78]. While AI is rapidly transforming various industries and applications, the absence of universally accepted standards has necessitated the development of a diverse array of frameworks to promote responsible and secure AI innovation.

One of the foremost AI frameworks is the European Union (EU) AI Act [67]. The European Commission (EC) pioneered a first-of-its-kind regulation – the EU AI Act – designed to ensure AI's safe, trustworthy, and lawful use within the European Union (EU). The primary goals of the EU AI Act are to promote consumer trust in AI systems by ensuring they adhere to fundamental rights and ethical standards, ensure safety, mitigate the risks associated with AI by establishing safeguards against potential harm to individuals and society, and foster innovation by creating a regulatory environment within the EU that allows for responsible AI innovation, setting a global standard [79]. The core principle of the AI Act is a risk-based approach. The EC categorizes AI systems based on their potential impact on users, ranging from unacceptable risk (prohibited) to high, limited, and minimal risk. The EU AI Act will significantly impact the heavily regulated finance and banking sectors. Transparency will be paramount, requiring institutions to be clear about leveraging data and utilizing AI [80].

The National Institute of Standards and Technology (NIST) Cybersecurity Framework, which, while originally designed for general cybersecurity, has been adapted to AI [53]. This framework is a robust blueprint for organizations to manage and mitigate cybersecurity risks associated with AI systems. It encompasses core functions of Identify, Protect, Detect, Respond, and Recover, offering a structured approach to addressing AI vulnerabilities and ensuring resilience against cyber threats.

The International Organization for Standardization (ISO) 27001 standard, renowned for its significance in information security management, has been adapted to AI applications [81].

The Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) offer specific standards and regulations for AI applications involving sensitive medical data in the healthcare sector. These frameworks ensure that AI-driven healthcare solutions maintain patient privacy, data security, and ethical data usage [82].

The relevance and applicability of these AI standards and frameworks extend across diverse contexts. The standards provide a structured foundation for organizations to develop secure, ethical, and regulatory-compliant AI systems. As AI continues to permeate critical sectors, such as healthcare, finance, and autonomous systems, these standards and frameworks play a pivotal role in fostering trust among users, stakeholders, and the broader society [78]. They are essential tools in shaping AI’s responsible and sustainable future, wherein innovation is harmonized with ethical considerations and robust security practices.

**Artificial Intelligence-Enabled Chatbots Technologies and Tools**

In the rapidly evolving landscape of AI enabled Chatbots, integrating robust cybersecurity technologies and tools has become imperative to safeguard against an expanding array of cyber
threats [83]. Key technologies and tools include firewalls, intrusion detection systems (IDS), and encryption methods, each playing a critical role in fortifying the security posture of AI-enabled Chatbots systems.

Firewalls are foundational components in the defense against cyberattacks. They are the initial barrier regulating incoming and outgoing network traffic based on an organization’s predetermined security policies [84]. In AI Chatbots, firewalls act as gatekeepers, monitoring and controlling data flows between users and AI-enabled Chatbots servers. By enforcing strict access controls and traffic filtering, firewalls help thwart unauthorized access attempts and potential threats. However, the effectiveness of firewalls in mitigating cyber threats is contingent upon their configuration, regular updates, and alignment with the AI-enabled Chatbots’ specific security requirements [85].

Intrusion Detection Systems (IDS) complement firewalls by actively monitoring network traffic for suspicious patterns or behaviors indicative of cyber threats. AI-enhanced IDS systems can discern anomalous activities by analyzing vast datasets and identifying deviations from normal network behavior [86]. IDS can serve as an intelligent sentry within AI enabled Chatbots, swiftly detecting unauthorized access attempts, malware infiltration, or data exfiltration. Nonetheless, to maximize their effectiveness, IDS systems should be fine-tuned to the AI-enabled Chatbots’ unique traffic patterns, and they necessitate continuous monitoring and alert response to ensure timely threat mitigation [87].

Encryption methods, encompassing techniques like Transport Layer Security (TLS) and end-to-end encryption, are instrumental in preserving the confidentiality and integrity of data exchanged between AI-enabled Chatbots and users. Encryption ensures that sensitive information remains unreadable to unauthorized entities during transit, rendering intercepted data useless to potential adversaries. In AI-enabled Chatbots, encryption is vital for securing user inputs, responses, and stored data [88]. The strength of encryption methods lies in their adherence to cryptographic standards, appropriate key management, and robust implementation, all of which contribute to their efficacy in mitigating data breaches and unauthorized data access.

The effectiveness of these AI-enabled Chatbots technologies and tools in mitigating cyber threats hinges on a holistic and well-integrated cybersecurity strategy [83]. While firewalls, Intrusion Detection Systems (IDS), and encryption methods form essential components of this strategy, their optimal functioning is contingent upon continual monitoring, regular updates, and alignment with the evolving threat landscape and the ability of AI enabled Chatbots to adapt and learn from emerging threats through machine learning and AI algorithms, further enhancing their resilience against cyberattacks [89].

**Artificial Intelligence Enabled Chatbots Best Practices**

Securing information systems and networks in the AI enabled Chatbot era demands a multifaceted approach encompassing industry best practices and lessons learned from successful cybersecurity strategies [90]. One fundamental best practice is establishing a robust, dynamic cybersecurity framework that aligns with evolving threats and regulatory requirements. This framework should include proactive threat intelligence, risk assessment, and continuous monitoring to identify vulnerabilities and emerging cyber threats [78]. Case studies of organizations like Google, which has successfully leveraged AI-enabled Chatbots for customer support while prioritizing data security, showcase the importance of incorporating AI-specific risk assessments and threat models within this framework.

Encryption and data privacy are paramount in AI-enabled Chatbots implementations. Best practices dictate end-to-end encryption and data anonymization techniques to protect sensitive user information [88]. Apple’s Siri, for instance, employs differential privacy to anonymize data collected from user interactions, setting a benchmark for data privacy preservation in AI Chatbot systems [91]. Additionally, organizations should adhere to stringent access controls, ensuring only authorized personnel can access Chatbot configuration and data repositories. Microsoft’s Azure Bot Service exemplifies a secure implementation by integrating identity and access management protocols and safeguarding access to Chatbot resources [92].

Regular software patching and system updates are other vital practices to mitigate vulnerabilities. Organizations like Facebook, leveraging AI-enabled Chatbots for customer engagement, exemplify
the diligent application of patch management protocols [93]. They swiftly address vulnerabilities to prevent exploitation by cyber adversaries. Furthermore, AI-enabled Chatbots implementations should adopt a ‘security by design’ approach, embedding cybersecurity considerations throughout the development lifecycle. Financial institutions have demonstrated this approach by implementing AI-enabled Chatbots for customer service with a strong focus on stringent security measures [94].

Employee training and awareness are pivotal, as humans often constitute the weakest link in cybersecurity. Prominent organizations like Amazon, which deploys AI-enabled Chatbots for customer interactions, invest substantially in cybersecurity education and training for their workforce, fostering a security-conscious culture [95]. Lastly, organizations should maintain clear incident response plans, exemplified by IBM’s Watson-enabled Chatbots, which integrate incident detection and response capabilities [96].

**Challenges and Limitations**

While this research aims to contribute valuable insights into the impact of AI enabled Chatbots data poisoning impact, it acknowledges potential limitations. Constraints might arise from data availability and scope, potentially limiting the study’s comprehensiveness. Despite these limitations, the formulated propositions provide a solid framework for analyzing and interpreting the study’s findings. The research aims to enhance the understanding of effective strategies for mitigating AI-enabled Chatbots data poisoning risks and preserving customer data privacy in digital banking.

This research study acknowledges certain limitations that may impact the research results. The specific context and sample characteristics may constrain the generalizability of findings. Additionally, the rapidly evolving nature of AI technology and AI-enabled Chatbots may introduce certain uncertainties and potential biases [30]. These limitations will be addressed through rigorous data collection, analysis, and interpretation to ensure the reliability and validity of the research outcomes.

**RESEARCH METHOD**

Understanding the methodologies and data collection processes is pivotal to replicating and comprehending any research [97]. The focal point of this paragraph is to elucidate the data collection methods and the research traditions that underpin this study on AI enabled Chatbots and data privacy violations. This paragraph will clarify the study’s purpose, elaborate on the research traditions relevant to the study, and emphasize how online surveys play an integral role in data collection.

Qualitative methodologies, particularly grounded theory, are employed to explore the intricacies of participants’ experiences, feelings, and perceptions, delving deep into the human experience to extract insights, patterns, and conceptual frameworks [60]. Grounded theory is an inductive methodology that seeks to generate theory from the data rather than testing an existing theory. By continually comparing data, codes, and categories, the research aims to develop a substantive theory grounded in the empirical data collected.

The qualitative approach is rooted in the constructivist paradigm, emphasizing participants’ subjective meanings and experiences, which often emerge from the data [45]. This makes grounded theory particularly suitable for this study, which seeks nuanced insights into the privacy concerns surrounding AI-enabled Chatbots. Through analyzing survey responses, the research aimed to capture Information Technology (IT) professionals’ rich narratives and evolving perspectives safeguarding the digital banking sector’s AI-enabled Chatbots, ultimately contributing to a more comprehensive understanding of data privacy violations. The emphasis here is not on quantifying experiences but on understanding and interpreting the phenomenon’s depth, complexity, and nuances.

This study employed a rigorous triangulation framework to enhance data validity [98]. Key components included pilot testing the survey instrument to refine questions and address biases before wider distribution [98].

Governance documentation and legal frameworks, such as laws and SOPs relevant to AI, chatbots, and data privacy in the banking sector, were integrated to benchmark survey findings [66]. This
alignment ensured a grounded and accurate understanding of AI data privacy issues and mitigation strategies [1]. The carefully designed survey instrument captured comprehensive data on demographics, experiences, and perceptions [99]. These diverse methods and data sources significantly enhanced the study's validity through triangulation.

**Research Tradition(s)**

Given the complex and evolving nature of AI and data privacy concerns, this research primarily adopts a qualitative tradition, with a particular emphasis on grounded theory. Qualitative research emphasizes understanding and interpreting the experiences, feelings, and perceptions of participants, often gathered through interviews, focus groups, and, in this case, surveys [61]. Grounded theory, rooted in the constructivist paradigm, revolves around building theory from data rather than testing pre-existing theory [60]. The essence of this tradition lies in its ability to delve deep into participants' narratives, constantly comparing data, codes, and categories, seeking to derive substantive theories grounded in empirical data [45].

In the context of AI, while vast amounts of data are processed, the human experience and interpretation of these processes, particularly concerning data privacy, cannot be fully understood through numbers alone. Grounded theory, in this context, allowed for a deeper exploration of Information Technology (IT) professionals' lived experiences, challenges, and strategies for safeguarding AI-enabled Chatbots against data privacy violations in the banking sector [60]. This qualitative approach provided a nuanced understanding of the complex interplay between technological safeguards and human factors, highlighting the necessity for ongoing vigilance and adaptability in data protection practices.

Regarding data privacy, qualitative research, especially through surveys, yielded rich insights into IT professionals' perspectives, strategies, challenges, and recommendations. It helped understanding the nuances of data privacy concerns, strategies employed for safeguarding data, and the implications of AI data poisoning. This method uncovered IT professionals' subjective experiences and practical wisdom, providing a more detailed picture of the real-world impact and efficacy of data protection measures. The qualitative data highlighted the critical role of continuous education and training for IT professionals to avoid emerging privacy risks and foster a culture of proactive data security within organizations [34].

The grounded theory tradition is apt for this study on AI-enabled Chatbots and data privacy, emphasizing deep understanding, interpretation, and theory-building. By grounding this research in this tradition, the researcher aims to capture the topic's richness and complexity, providing nuanced insights that are directly grounded in the experiences and perspectives of those at the forefront of this evolving field [100]. This approach facilitates the development of a robust theoretical framework and ensures that the findings are deeply rooted in practical realities. By focusing on the lived experiences of IT professionals, the research offers actionable recommendations and strategies that are both relevant and applicable, bridging the gap between theory and practice in the domain of AI data privacy.

**Research Questions and Propositions**

The propositions form the core of the question guiding the study focus on the impact of AI data poisoning in AI-enabled Chatbots systems applied to the digital banking sector. The central research question inquires about how to mitigate data privacy violations caused by AI data poisoning in AI-enabled Chatbots handling personally identifiable information (PII) like credit card numbers, Social Security Numbers (SSN), Date of Birth (DOB), or account numbers. By addressing this question, the study aimed to provide actionable insights for IT professionals and policymakers to enhance the security and reliability of AI-enabled banking applications, ultimately fostering greater trust and confidence in these technologies [34].

This study proposes to explore the relationship between AI data poisoning and data privacy violations. It is postulated that AI data poisoning significantly affects data privacy violations in AI-enabled Chatbots within the digital banking sector [34]. Additionally, AI data poisoning influences the mishandling of personally identifiable information (PII). The extent of data privacy violations is positively correlated with the prevalence and severity of AI data poisoning incidents, and
Effective mitigation strategies targeting AI data poisoning reduce data privacy violations in AI-enabled Chatbots systems [35].

Several propositions are presented to provide additional context [101]. These propositions outline the transformational effect of AI Chatbots in the banking sector, concerns about data privacy violations due to AI data poisoning, the nature of AI data poisoning, and the potential consequences of data mishandling by Chatbots. Also, ethical considerations are significant in AI development and deployment within the digital banking sector [49]. While this research aimed to contribute valuable insights into the impact of AI enabled Chatbots data poisoning impact, it acknowledges potential limitations. Some data availability and scope constraints might have limited the study’s comprehensiveness. Despite these limitations, the formulated propositions provided a solid framework for analyzing and interpreting the study’s findings. The research aimed to enhance the understanding of effective strategies for mitigating AI-enabled Chatbots data poisoning risks and preserving customer data privacy in digital banking.

Research Design

In AI and data privacy, a qualitative research approach, particularly grounded theory, provides a rich understanding of the intricate nuances of privacy violations and the complexities surrounding AI-enabled Chatbots [100]. Unlike quantitative research, which emphasizes measuring and counting, grounded theory delves deep into the lived experiences, challenges, and strategies associated with data privacy in AI systems [45]. Using grounded theory, this study aimed to comprehend how data privacy violations emerge in AI-enabled Chatbots interactions, the underlying reasons for such breaches, and Information Technology (IT) professionals’ strategies to counteract these challenges [60]. By analyzing survey responses, the research intends to construct a theory grounded in the data, which captures the essence of AI data poisoning, its implications, and the resulting data privacy concerns.

Through in-depth analysis of survey responses and constant comparison, grounded theory seeks to unravel patterns and themes, offering insights that quantitative methods might overlook. The intention was to identify a dynamic perspective on user perceptions, trust levels, and experiences with AI-enabled Chatbots, highlighting the interplay between AI technology and its human users. This approach allows for a rich, contextually grounded understanding of how users interact with AI systems, the trust issues faced, and the overall satisfaction and concerns. The study also contributed to the broader discourse on ethical AI usage, emphasizing the need for transparent and accountable practices in deploying AI technologies within sensitive domains such as banking [102].

Population and Sample

The survey participants for this study were professionals in the banking sector engaged in developing, managing, or utilizing AI-enabled Chatbots. These individuals were selected based on defined criteria and were reached through professional networking platforms and social media channels. The survey also included end-users who interact with banking platforms. These participants were selected to ensure that the sample is representative of the wider population of professionals and experts in the field. The survey was designed to capture their perceptions, attitudes, experiences, and preferences related to AI Chatbots and data privacy in banking. The survey was administered to ensure anonymity, informed consent, and data protection compliance.

Data Protection and Data Access Plan

- All participants were well-informed about the purpose of the research, the use of their data, and any potential risks involved.
- Clear and unequivocal consent was obtained from participants before collecting data.
- Only the researcher had access to the data collected.
- Data was anonymized before publishing.
- Data was securely destroyed after publishing.
Measures were taken to protect the data from unauthorized access, using encryption in a password-protected secure storage to anonymize data. All personally identifiable information (PII), if any, was removed.

**Population Definition and Criteria**

- **Professionals**: Those engaged in developing, managing, or utilizing AI-enabled Chatbots within the banking industry.
- **End-Users**: Individuals who interact with banking platforms, particularly those utilizing AI-enabled Chatbots for their transactions or inquiries.
- **Criteria**: The population must interact directly or indirectly with AI-enabled Chatbots in the banking industry and should hold insights into data privacy considerations.
- **Encompassing**: The population embodies software engineers, data scientists, bankers, customer service representatives, and customers. Expands geographically to include individuals and professionals from various regions to capture a wide spectrum of insights and experiences.

**Sample Criteria:**

- **Professionals**: Members of relevant professional associations holding credentials in fields like data privacy, AI, or banking.

**Criteria Justification**: Focusing on credentialed professionals ensured the collected data was informed and relevant, enhancing the validity of findings.

**Sample Selection:**

- **Professionals**: Professional association members.
- **Experts**: Individuals with experiential knowledge in pertinent areas.

**Sample Justification:**
The sample was purposively structured to encapsulate a depth and breadth of perspectives and insights, enhancing the comprehensiveness and applicability of findings.

**Refinement and Constraints:**

- **Refinement**: The refined population might concentrate on professionals and users in regions with robust AI-enabled Chatbot utilization in banking to ensure the relevance and applicability of findings.
- **Constraints**: Limitations such as accessibility to professionals and experts, geographic variations, and willingness to participate will be considered and transparently communicated in the research documentation.

**Sampling Strategy**

Given the qualitative nature of the study and its reliance on grounded theory methodology, a purposive sampling strategy was employed to deeply understand and explore the experiences and perceptions of individuals within the defined population.

**Purposive Sampling for Surveys**

- **Target Population**: Individuals and professionals who interact with, develop, or are otherwise engaged with AI-enabled Chatbots in the banking industry.
- **Inclusion Criteria**: Professionals certified in relevant fields, active users of banking AI-enabled Chatbots, and experts in AI, data privacy, and banking.
- **Selection Criteria**: Members of recognized professional associations in banking, AI, and data privacy sectors.
- **Rationale**: By targeting professionals with grounded experiences, the study aims to capture rich insights from those who possess both expertise and firsthand experiences.
• **Sample Size Determination:** The sample size for the survey will be determined by considering the population size of the professional forums and applying a confidence level of 95% and a margin of error of 5% to achieve statistical significance in the findings. The population size was set at a minimum of 20 (twenty) participants.

• **Sampling Procedure:** Invitations to participate in the survey were extended through professional association communication channels, ensuring access to those most versed in the domain.

**Data Collection**

**Open-Ended Questions Surveys**

• **Mode:** Online, facilitated through the platforms of professional associations.

• **Content:** Framed to explore perceptions, attitudes, experiences, and preferences related to AI-enabled Chatbots and data privacy in banking.

• **Administration:** Ensuring anonymity and informed consent and strictly adhering to data protection standards.

**Justification and Rigor**

• **Richness of Data:** The purposive sampling strategy aims to delve deep into the insights of a select group, focusing on the richness of data rather than its breadth.

• **Credibility and Transferability:** The alignment of the sampling method with grounded theory ensures that the findings, while not broadly generalizable, are deeply rooted in participant experiences, making them transferable to similar contexts.

• **Ethical Adherence:** Transparency will be maintained throughout the participant engagement process about research aims, data utilization methods, and participant rights, thereby safeguarding informed consent and maintaining the highest ethical standards.

**Instrumentation**

This study embraces a qualitative research methodology grounded in the principles of the grounded theory [100], employing open-ended surveys to extract deep insights from respondents within the community of Information Technology (IT) professionals in the banking sector regarding the intricacies and narratives of AI enabled Chatbots and data privacy concerns. An open-ended survey questionnaire was disseminated to individuals who met the established criteria, reaching out to them through two professional networking platforms and specialized media channels [61]. Employing surveys in the grounded theory approach presents various merits, including delving deeper into respondents’ experiences and understanding the underlying beliefs, perceptions, and emotions. Open-ended surveys allow for rich data collection, facilitating the extraction of themes and patterns from participants’ narratives. The use of open-ended surveys in research harbors numerous advantages, including a more cost-effective avenue than other data-gathering techniques [61]. Open-ended surveys offer a heightened level of flexibility as they can be conducted online or through social media platforms, and reliability, given the anonymity afforded by surveys, tends to encourage respondents to provide more honest and valid responses [103].

Advantages in the Context of AI AI-enabled Chatbots:

• **Contextual Insights:** Tailored surveys can draw narratives from users of AI-enabled Chatbots, technology experts, and other relevant groups, unearthing rich contextual insights.

• **Global Perspective:** Internet penetration allows for feedback from diverse user demographics, providing a global view of data privacy concerns.

• **Quick Feedback:** Given the swift-paced evolution of AI and data privacy domains, open-ended surveys anchored in grounded theory can swiftly capture evolving narratives, ensuring the research remains current and relevant.
Validity
The validity of a research study signifies the applicability of its outcomes and insights to analogous individuals beyond the scope of the study [104].

Reflexivity Analysis
Reflexivity in qualitative research is a critical examination of the influence of the researcher's role over the research process and outcomes [105]. It involves a conscious reflection on the impact of one's biases, experiences, and perspectives on shaping the research questions, design, data collection, analysis, and interpretations. Essential for ensuring the integrity, authenticity, and ethical conduct of the research, reflexivity was adopted as a core methodology in this study, aimed at deepening the understanding of AI data privacy violations and ensuring findings are grounded in ethical and rigorous inquiry [102].

Researcher's Background and Motivation
The study is driven by a profound interest in the confluence of technology, ethics, and privacy, especially within the dynamic field of AI in banking. The choice of this research topic is influenced by a background in technology ethics, leading to a critical exploration of AI data privacy violations. Recognizing the influence of a predisposition towards emphasizing risks and challenges associated with AI-enabled chatbots, the study strives for a balanced analysis that is open to potential benefits and solutions.

Assumptions and Biases
The research began with assumptions regarding the prevalence and impact of AI data poisoning incidents related to AI-enabled chatbots and the effectiveness of mitigation strategies. A reflexive examination of these assumptions allowed for a critical evaluation of their basis and potential limitations they might impose on the study's scope or skew its findings [102]. This process revealed a bias toward a skeptical view of technological solutions' ability to safeguard data privacy effectively.

Methodological Reflections
Selecting a qualitative grounded theory approach was deliberate, reflecting a preference for depth over breadth in understanding Information Technology professionals' experiences securing banking AI-enabled chatbots.

Impact and Contribution of Findings
The reflexivity analysis encouraged a critical evaluation of the research findings' potential impact on policy, practice, and future studies, contemplating whose voices were amplified and whose might have been marginalized. This reflective process aimed for a balanced discourse acknowledging the complexities of AI in banking, advocating for ethical data practices, and fostering innovation in data privacy safeguards. Ethical considerations were also meticulously followed, and all participants were provided with an informed consent form detailing the research purpose, assurance of confidentiality, and their rights as participants. Rigorous safeguarding of participants’ response privacy and confidentiality was implemented to uphold the integrity of the research process and findings [106]. Through scrupulous attention to the validity and reliability of the survey instrument, this research aimed to yield robust, credible, and generalizable findings within the realm of AI-enabled chatbots and data privacy.

Reliability
Reliability in qualitative research, especially one rooted in grounded theory, encompasses the depth and richness of data and the consistency in interpretation and representation of participants' narratives [107]. Unlike quantitative approaches, where reliability often focuses on instrument stability and repeatability, qualitative reliability centers on the trustworthiness and authenticity of the gathered narratives and the analytical rigor with which they are processed.

In the context of this study, which employed open-ended surveys disseminated through specialized platforms such as professional associations' portals, ensuring reliability revolves around capturing authentic experiences and validating the consistency of emerging themes and patterns. For survey protocols uniquely crafted for this research, strategies were employed to ensure the trustworthiness
of the collected data. One approach involved conducting a pilot study to gauge the clarity and relevance of open-ended questions, thereby ensuring they elicit meaningful narratives. Drawing from a smaller, representative subset of the target audience, the pilot study provided insights into possible areas of improvement for the main survey.

Another significant aspect of grounded theory reliability is the iterative process of data analysis. This means that as data is collected, it's continuously compared and contrasted with new data, ensuring emerging categories and themes are rooted in the data and evolve as more narratives are gathered. Furthermore, the "constant comparison method," a foundational grounded theory technique, will be instrumental. This involves continuously juxtaposing individual participant responses, aiding in identifying and refining emergent categories, and ensuring the thematic analysis is grounded in participants' voices.

Member checking, another strategy, was also pivotal. By revisiting participants with the initial findings and interpretations, researchers can validate that their interpretations align with participants' intended meanings, enhancing the reliability of the research conclusions. Through such layered approaches centered on trustworthiness, authenticity, and analytical rigor, this research sought to ensure reliability in capturing and interpreting participants' experiences, aspirations, and concerns in the domain of AI-enabled Chatbots and data privacy.

**Data Collection**

Qualitative surveys, rooted in the grounded theory approach, served as the cornerstone for data collection in this research [100]. These surveys, designed with open-ended questions, sought to delve deeply into respondents' lived experiences, perceptions, and nuances as they navigate their interactions with AI-enabled Chatbots. Specifically, the online surveys were structured to:

- **Assess User Familiarity:** Through open-ended questions, the surveys endeavor to unravel the depth of users' interactions with AI-enabled Chatbots, probing the platforms they frequently engage with and capturing the richness of their experiences.
- **Gauge Privacy Concerns:** Identify the most common data privacy issues technology experts and users face or are concerned about.
- **Capture Experiences:** Use open-ended questions to document specific incidents of perceived data privacy violations.
- **Solicit Feedback:** Understand from users how AI-enabled Chatbot systems can be improved in terms of data privacy, thus tapping into users' aspirations and solutions.
- **Data Access:** As in all rigorous qualitative research, access to the data was strictly restricted to the primary researcher. Measures were taken to protect the data from unauthorized access, such as encryption, secure storage, and the use of anonymized data.
- **Data Anonymization:** Consistent with grounded theory's emphasis on ethical considerations, the survey was meticulously designed to eschew collecting any personally identifiable information (PII). This circumvents potential ethical quandaries and ensures respondents' narratives remain divorced from their identities, further safeguarding participant privacy. The data collected was anonymized to protect the privacy of the participants.

**Data Analysis**

In the data analysis phase, this research adopted a grounded theory approach characterized by simultaneous data collection and analysis [60]. This method facilitated the emergence of themes and patterns that guided further data collection. Reflecting on the research questions, the researcher critically assessed how the qualitative survey responses could be assimilated and interpreted to comprehensively address and contribute to the research questions. Given the focus on AI-enabled chatbots in the banking industry and the intricate nuances of data privacy, this study employed surveys with a qualitative bent to analyze the rich qualitative data. The analysis was conducted in several structured steps, each designed to progressively refine and deepen our understanding of the survey responses. For detailed explanations of these methods, see Appendix C. To uncover the
nuanced understanding of data privacy in AI-enabled Chatbots, the following structured analytical methods were employed:

- **Open Coding**: Responses from the survey were first subjected to open coding. The data was dissected into discrete parts, closely examined, and compared for similarities and differences. This process helped identify distinct concepts and categories emerging from the participants’ narratives.

- **Axial Coding**: Post open coding, axial coding was employed to identify the relationships between the established categories, focusing on how these categories intersect, overlap, or inform one another.

- **Selective Coding**: In the final coding phase, the research zeroed in on one core category, which was considered central to the research focus. All other categories were related and integrated into this core category, enabling the development of a grounded theory that resonates with the AI-enabled Chatbots data privacy context in the banking sector.

- **Constant Comparative Method**: Throughout the analysis, a constant comparative method was employed, where each piece of data was constantly compared to other data chunks, categories, or properties to identify patterns, relationships, and anomalies.

- **Memo Writing**: To ensure depth in analysis, memo writing was a continuous process during the coding phases. These memos captured the researcher’s thoughts, interpretations, and reflections, forming a crucial part of the grounded theory development.

**General Outline of a Survey Protocol for this Research**

The survey contained a brief introduction of the purpose of the research with an explanation of the survey process, including the estimated duration and the method. Assurance to the participants of the confidentiality of their responses and their right to withdraw at any time. Informed consent from the participant.

**Survey Questions**

**Demographics**

- Can you please share your role and experience in the field of AI, AI-enabled Chatbots, banking, or data privacy?

**Familiarity with AI Chatbots**

- Can you describe your experience with AI-enabled Chatbots in the banking industry?
- How often do you interact with AI-enabled Chatbots?

**Perceptions and Attitudes**

- What are your perceptions of AI-enabled Chatbots in the banking industry?
- How do you feel about the data privacy measures currently in place for AI-enabled Chatbots?

**Experiences and Concerns**

- Can you share any specific incidents where you felt your data privacy was compromised while using AI-enabled Chatbots?
- What are your main concerns regarding data privacy in the context of AI-enabled Chatbots?

**Recommendations Based on Expertise and Experience**

- Based on your expertise and experience, what specific measures would you recommend to enhance data privacy in AI-enabled Chatbots in the banking industry?
- Can you suggest any best practices or standards that should be adopted by banking institutions to ensure data privacy in AI-enabled Chatbots?
- Are there any specific technologies or techniques that you believe could improve data privacy in AI-enabled Chatbots?
What steps should banking institutions take to educate users about data privacy when using AI-enabled Chatbots?

How can banking institutions better communicate their data privacy policies and practices to users of AI-enabled Chatbots?

Assessing the Effectiveness of Various Security Measures

Have you noticed any improvements in data privacy in AI-enabled Chatbots over the past few years? If so, what measures do you believe have contributed to this improvement?

How effective do you believe encryption techniques are in ensuring data privacy in AI-enabled Chatbots?

How effective do you believe user education and awareness campaigns are in enhancing data privacy in AI-enabled Chatbots?

How effective do you believe regulatory oversight and compliance are in ensuring data privacy in AI-enabled Chatbots?

Conclusion

Is there anything else you would like to share about your experience or thoughts on AI-enabled Chatbots and data privacy?

Ethical Considerations

This study employed a survey method to accumulate participants’ insights regarding safeguarding data privacy in the AI-enabled Chatbots contest. Surveys inherently pose potential ethical quandaries encompassing confidentiality, informed consent, anonymity, coercion and undue influence, and non-disclosure of conflicts of interest [106]. As posited by the American Association of Public Opinion Research (AAPOR), respondents must clearly understand the survey’s context, content, objective, and sponsorship before their acceptance to partake. It behooves researchers to divulge to sponsors any limitations and deficiencies inherent in the survey while avoiding any elements that might precipitate bias in the outcomes. The survey report should encapsulate details pertaining to any sponsorship, the entity conducting the questionnaire, the sequence of questions, a delineation of the population under study, and the methodologies deployed for data acquisition [108].

The survey concerning data privacy in the scope of AI-enabled Chatbots encompassed a prefatory segment elucidating the study’s context, content, and objective. This segment also emphasized the voluntary nature of engagement in the survey and the prerogative to withdraw at any juncture, predicated on informed and voluntary consent [108]. Given the sensitivity of data privacy, participants were ensured anonymity, and their responses were encrypted. The study strictly adhered to ethical guidelines to maintain participant trust and ensure data integrity.

RESULTS

Triangulation and Pilot Testing

A methodological framework of triangulation was rigorously applied to bolster the validity of the data collected in this study [98]. This multifaceted approach involved several key components to cross-validate and reinforce the research findings. Initially, pilot testing of the survey instrument was conducted to refine the questions and ensure clarity and relevance. This preliminary phase was critical in identifying potential issues and biases in the survey design, allowing for necessary adjustments prior to the wider distribution of the survey [98].

In addition to pilot testing, the study integrated governance documentation and legal frameworks as foundational elements. These documents, including laws and standard operating procedures (SOPs) relevant to the field of AI, chatbots, and data privacy in the banking sector, served as benchmarks against which the survey findings could be compared and contrasted [66]. By aligning the study with established legal and procedural standards, a more grounded and contextually accurate understanding of AI data privacy violations and mitigation strategies was achieved [1].
The survey instrument itself was a critical component of the triangulation process. It was meticulously designed to capture a comprehensive range of data points encompassing respondent demographics, experiences, and perceptions [99]. The thoroughness of the survey ensured that the data collected was robust and multi-dimensional, enabling a nuanced analysis of the research questions. Through the integration of these diverse methods and data sources, the study achieved a level of triangulation that significantly enhanced the validity of its findings.

**Participant Demographics**

The demographics of the participants in the study are represented by a diverse set of participants and a global representation of participants, with a significant number from the United States, followed by South Africa, Poland, and others. The participant pool consisted of 104 individuals. Employed is the most common employment status among participants, with notable counts in countries like the United States (21), South Africa (10), and Mexico (5). Students have a significant presence in Poland (7) and South Africa (3), with the United States having five students. Unemployed participants are found in a few countries, with Portugal having two and the United States having 3. Retired participants are less common, with only 1 in the United States. There were 62 Employed, 27 Students, 1 Retired, and 7 Unknown participants.

The majority of participants are employed, and a significant number of students are also participating in the study. Unemployed and one retired participant comprise a smaller portion of the total. There is also a diverse industry employment representation, with a wide range of industries represented across different nations. Aviation is represented in Bahrain and Germany. Education, Financial, and Healthcare industries have participants in various countries, with a notable presence in the United States. IT (Information Technology) is widely represented across several countries, including the United States, Poland, and South Africa. Unique industries such as Agriculture in Mexico, Transport and logistics in Spain, and science in the United States are also mentioned. This diverse industry representation underscores the wide interest and relevance of the study’s topic across different sectors and geographies.

<table>
<thead>
<tr>
<th>Description</th>
<th>Category</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Common Gender</td>
<td>Male</td>
<td>55</td>
</tr>
<tr>
<td>Most Common Industry</td>
<td>IT (Information Technology)</td>
<td>17</td>
</tr>
<tr>
<td>Most Common Country</td>
<td>USA</td>
<td>32</td>
</tr>
<tr>
<td>Most Common Employment Status</td>
<td>Employed</td>
<td>62</td>
</tr>
<tr>
<td>Most Common Age Group</td>
<td>18-25</td>
<td>36</td>
</tr>
</tbody>
</table>

**Note:** Professionals certified in relevant fields, active users of AI-enabled chatbots in banking, and experts in AI, data privacy, and banking.

As depicted in Table 2, this study on AI data privacy violation in chatbot systems in the banking sector involved diverse participants, as demonstrated by the demographics summary table. The participant pool comprised 104 individuals, offering a comprehensive insight into different perspectives based on gender, age, employment status, industry, and country of residence.

- **Gender Diversity:** The gender distribution among participants included four identified categories, with "Male" being the more prevalent gender, represented by 55 out of 104 participants.

- **Age Groups:** Participants were spread across four age groups, with the majority (36 participants) falling in the "18-25" age category. This suggests a younger demographic predominantly engaged in this study, which might reflect the evolving interest and involvement of younger professionals in the AI and data privacy sectors.

- **Employment Status:** Participants’ employment status varied, with "Employed" being the most common status. This indicates that most participants are actively involved in professional settings, potentially bringing practical insights into the study.
• **Industry Representation:** Participants hailed from five different industries, with a notable concentration on "IT (Information Technology)," represented by 18 participants. This underscores the relevance of IT professionals in discussions about AI and data privacy, especially in the banking sector.

• **Geographical Distribution:** The study included participants from 21 different countries, with the US being the most represented country. Such geographic diversity enhances the study's comprehensiveness, incorporating a wide range of cultural and regulatory perspectives on AI and data privacy.

**Data Analysis Methods**

The data analysis methods adopted the grounded theory approach inherent to qualitative research [60]. The analysis began by reading the survey responses immersively to become familiar with participants’ insights and experiences. Responses from the surveys underwent a qualitative content analysis, extracting dominant themes, narratives, and patterns.

**Presentation of the Data**

This study's qualitative analysis is anchored in a comprehensive survey conducted among professionals in fields intersecting AI, AI-enabled Chatbots, banking, and data privacy. The survey, meticulously designed to align with APA standards, garnered responses from a diverse cohort, ensuring a rich tapestry of insights. Rather than inundating the reader with voluminous data, this summary distills the essence of the findings, emphasizing salient points that speak directly to the research questions.

**Theme Introduction**

The thematic analysis unearthed several pivotal themes, each shedding light on the intricate relationship between AI data poisoning and data privacy in the realm of digital banking chatbots.

**Varied Roles and Exposure**

Respondents’ roles spanned a spectrum from software developers to business owners, each bringing a unique lens to the issue. A representative quote encapsulates this diversity: “I work in an IT company, focusing on integrating AI in our services... my experience in banking is indirect but relevant.”

**Direct and Indirect Experiences with AI Data Poisoning**

Experiences varied significantly, with some respondents having direct encounters with data breaches while others expressed limited exposure. A notable response highlights this: “In my tenure, I've seen AI systems compromised, leading to significant data privacy concerns.”

**Mitigation Strategies - A Range of Perspectives**

The responses indicate a spectrum of strategies, from sophisticated technical solutions to basic awareness-raising measures. One respondent insightfully noted, “Implementing robust encryption and continuous monitoring systems has been key in our approach to safeguarding data privacy.”

**Perceptions on AI Data Poisoning’s Impact on PII**

There was a mixed response regarding AI data poisoning’s influence on handling PII, with some citing it as a growing concern, while others viewed it as under control. A respondent observed, “Currently, the impact seems minimal, but the potential for future risks is undeniable.”

**General Observations on AI Chatbots and Data Privacy**

Many respondents indicated satisfactory contentment with current data privacy measures, though a few anticipated emerging challenges. An interesting perspective was shared: “While we’re not seeing immediate threats, the evolving nature of AI makes constant vigilance essential.”

These themes, enriched by direct quotes from participants, offer a nuanced understanding of the current landscape of AI data poisoning and data privacy in digital banking chatbots. They serve as a foundation for a deeper exploration of the study’s research questions, elucidating the complexities and variances in perceptions and experiences among professionals in the field.
### Table 3. Experiences with AI Data Poisoning

<table>
<thead>
<tr>
<th>Responded ID</th>
<th>Role in Organization</th>
<th>Experience with AI Data Poisoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AI Developer</td>
<td>Encountered minor data breaches</td>
</tr>
<tr>
<td>2</td>
<td>IT Specialist in Banking</td>
<td>No direct experience</td>
</tr>
<tr>
<td>3</td>
<td>Data Privacy Officer</td>
<td>Witnessed AI hacking leading to data leak</td>
</tr>
<tr>
<td>4</td>
<td>Chatbot System Manager</td>
<td>AI system compromised, privacy concerns</td>
</tr>
<tr>
<td>5</td>
<td>Banking Professional with AI Focus</td>
<td>Limited exposure, aware of potential risks</td>
</tr>
<tr>
<td>6</td>
<td>Technology Consultant</td>
<td>Concerns about future risks, no incidents</td>
</tr>
<tr>
<td>7</td>
<td>AI Researcher</td>
<td>Studied cases of AI data poisoning</td>
</tr>
<tr>
<td>8</td>
<td>Digital Security Analyst</td>
<td>Implemented countermeasures for threats</td>
</tr>
<tr>
<td>9</td>
<td>Chatbot Developer</td>
<td>Observed data integrity issues in AI systems</td>
</tr>
<tr>
<td>10</td>
<td>Compliance Officer</td>
<td>Monitored for regulatory breaches</td>
</tr>
<tr>
<td>11</td>
<td>AI Strategy Advisor</td>
<td>Advised on mitigation strategies</td>
</tr>
<tr>
<td>12</td>
<td>Customer Service Manager Using AI Chatbots</td>
<td>Noted customer concerns regarding data security</td>
</tr>
</tbody>
</table>

Note: Professionals certified in relevant fields, active users of AI-enabled chatbots in banking, and experts in AI, data privacy, and banking.

As depicted in Table 3, the array of experiences with AI data poisoning among the survey respondents highlights this issue's diverse and multifaceted nature within the digital banking sector. The table presents a spectrum of encounters, ranging from direct experiences with data breaches and AI system compromises (as reported by Respondents 1, 3, and 4) to roles with no direct experience but a keen awareness of potential risks (such as Respondents 2 and 5). Other respondents, like the Technology Consultant and AI Researcher (Respondents 6 and 7), emphasize concerns about future risks or theoretical aspects of AI data poisoning. This diversity extends to roles like Digital Security Analysts and Chatbot Developers (Respondents 8 and 9), who are actively engaged in implementing countermeasures or dealing with data integrity issues. The table also includes perspectives from a Compliance Officer, an AI Strategy Advisor, and a Customer Service Manager (Respondents 10, 11, and 12), who contribute insights into regulatory compliance, strategic mitigation approaches, and customer-related data security concerns, respectively(Krause, 2023). This range of experiences underscores the complexity and varying degrees of impact that AI data poisoning has on data privacy in the banking sector’s chatbot systems.

### DISCUSSION

This study on AI data privacy violations in chatbot systems within the banking sector involved a detailed analysis of participant responses. This included parsing data from various professionals with diverse roles and experiences in AI, banking, and data privacy. The analysis process focused on identifying patterns and themes in the experiences, perceptions, and strategies related to AI data poisoning. Using qualitative methods, responses were categorized to understand the spectrum of experiences and opinions.

**Key Findings**

- **Varied Roles and Exposure**: Participants from different professional backgrounds, including AI developers and banking professionals, provided a wide lens on AI data poisoning, indicating its broad impact across roles.
- **Direct and Indirect Experiences with AI Data Poisoning**: Experiences ranged from direct encounters with data breaches to limited exposure, highlighting the variability in how professionals confront AI data privacy issues.
- **Mitigation Strategies**: A spectrum of strategies was identified, from advanced technical solutions to basic awareness measures, suggesting varied approaches to combating AI data poisoning.
- **Impact on Personally Identifiable Information (PII):** Mixed responses were observed regarding the influence of AI data poisoning on PII, indicating differing levels of concern and control over this issue.

- **General Observations on AI Chatbots and Data Privacy:** While some respondents were content with current data privacy measures, others anticipated emerging challenges, reflecting a cautious optimism mixed with vigilance.

**Application to Research Question**

These findings address the research question by revealing the prevalence and diverse impacts of AI data poisoning in the banking sector. The study suggests that AI data poisoning is a significant risk to data privacy, particularly in digital banking chatbots. Understanding this complex threat landscape is crucial for developing effective mitigation strategies [1].

The data analysis reveals a nuanced understanding of AI data poisoning, highlighting the need for robust security measures in the banking sector. The diversity in experiences and perspectives emphasizes the complexity of threats and the importance of comprehensive strategies. These findings advocate for an ongoing assessment of risks and the development of adaptive solutions to ensure the protection of data privacy in the evolving field of AI and banking.

**Table 4. Experiences with AI Data Poisoning**

<table>
<thead>
<tr>
<th>Responded ID</th>
<th>Role</th>
<th>Experiences with AI Data Poisoning</th>
<th>Mitigation Strategies Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>creation of a chatbot...</td>
<td>No</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>I work in an IT company,...</td>
<td>No</td>
<td>...I haven't had any incident, ...</td>
</tr>
<tr>
<td>3</td>
<td>AI developer</td>
<td>AI hacking leaked company's data</td>
<td>Data backup and safe storage</td>
</tr>
<tr>
<td>5</td>
<td>I use AI image generation tools, but I don't have...</td>
<td>No, any</td>
<td>I only use generic prompt, without sharing...</td>
</tr>
<tr>
<td>12</td>
<td>We use chatbots and AI in our sector (banking/financial)</td>
<td>Some bots or AI have miscopied some data compilation</td>
<td>...authentication...</td>
</tr>
<tr>
<td>66</td>
<td>...primary role involves researching AI data poisoning and its impacts on chatbots...</td>
<td>...studied cases where chatbot systems received malicious input...</td>
<td>Implementing continuous learning systems with anomaly detection capabilities and frequent data sanitization processes has proven effective.</td>
</tr>
<tr>
<td>67</td>
<td>I am in IT integrating AI chatbot technologies within banking services, ensuring they adhere to strict data privacy standards is one of my role main duties</td>
<td>We had an instance where flawed data input led to a chatbot disclosing account information in an unsecured format, which we immediately corrected.</td>
<td>Implementing rigorous data validation, regular AI model retraining, and deploying anomaly detection systems have been effective against data poisoning.</td>
</tr>
</tbody>
</table>

**Note:** Professionals certified in relevant fields, active users of AI-enabled chatbots in banking, and experts in AI, data privacy, and banking.

The diversity of roles and experiences depicted in Table 4 underscores the multifaceted nature of AI data poisoning in the banking sector. This table reflects the varied impact and awareness levels, from direct encounters with breaches to strategic and regulatory perspectives. These insights form a comprehensive understanding of the challenges in mitigating AI data poisoning risks and emphasize the importance of a multi-pronged approach in addressing data privacy concerns in the context of digital banking chatbots.

The study explored AI data privacy violations within Chatbot systems in the banking sector, employing a robust methodological framework centered on triangulation to enhance the validity of its findings. This approach was meticulously designed to incorporate a variety of perspectives and data sources, including pilot testing of the survey instrument, governance documentation, and legal frameworks pertinent to AI, Chatbots, and data privacy. The initiative aimed to ground the research in the context of existing standards and practices, thus offering a comprehensive view of the state of data privacy and strategies for mitigation in the realm of AI-powered digital banking services [1].
The research drew on the insights of 104 participants from a wide array of countries, with a significant concentration in the United States. The demographics revealed a predominance of employed individuals, particularly in the Information Technology (IT) sector, suggesting a strong representation of professionals directly involved with or interested in AI technologies and data privacy concerns. The analysis, guided by grounded theory and qualitative content analysis, highlighted several critical themes, including the impact of AI data poisoning across different professional roles, underscoring its broad relevance. The analysis unveiled a spectrum of experiences, from firsthand encounters with data breaches to more theoretical concerns about potential risks, illustrating the variability in exposure and awareness among professionals. Responses revealed a range of strategies employed to combat AI data poisoning, from technical solutions like encryption and continuous monitoring to basic awareness and regulatory compliance measures, indicating a multifaceted approach to safeguarding data privacy. Mixed perceptions were observed regarding the effect of AI data poisoning on Personally Identifiable Information (PII), with some participants noting minimal current impact but acknowledging the potential for future risks.

These findings directly respond to the research question by elucidating the prevalence and varied impact of AI data poisoning within the banking sector’s Chatbot systems. The study underscores the significance of this risk to data privacy and highlights the complexity of the threat landscape, advocating for robust and adaptive security measures. The study’s findings point to a nuanced understanding of AI data poisoning challenges in the banking sector. These findings emphasize the need for ongoing vigilance, comprehensive risk assessments, and the development of innovative solutions to protect against data privacy violations. The diversity of experiences and mitigation strategies reported by participants underscores the complexity of the issue and the importance of a collaborative, multi-pronged approach to addressing these challenges [1].

**FINDINGS**

The purpose of the study was to investigate the implications of AI data poisoning on data privacy breaches within AI-enabled chatbot systems in the digital banking sector, with a focus on identifying and suggesting mitigation strategies to address these risks [1]. The research aimed to provide a comprehensive understanding of how AI data poisoning affects data privacy, offering insights that could inform the development of robust data protection measures. To achieve this objective, the study employed a methodological framework centered on triangulation, enhancing the validity of its findings through a multifaceted approach.

The research methodology incorporated several key components. Initially, pilot testing of the survey instrument was conducted to refine the questions, ensuring their clarity and relevance [109]. This was crucial for identifying and addressing potential issues and biases in the survey design before its broader distribution. Additionally, the study integrated governance documentation and legal frameworks relevant to AI, chatbots, and data privacy in the banking sector. These documents served as benchmarks against which the survey findings could be compared and contrasted, thereby grounding the research in the context of established legal and procedural standards.

The survey instrument itself was meticulously designed to capture a comprehensive range of data points, including respondent demographics, experiences, and perceptions [68]. This thorough approach ensured that the data collected was robust and multidimensional, enabling a nuanced analysis of the research questions. Through these diverse methods and data sources, the study sought to achieve a level of triangulation that significantly enhanced the validity of its findings, contributing to a deeper understanding of AI data poisoning and its impact on data privacy in digital banking chatbots [98]. This, in turn, facilitated the identification of potential mitigation strategies to safeguard against such risks.

The most significant discoveries include varied experiences with AI data poisoning among professionals, indicating that the issue is prevalent and has diverse impacts. The study also found a range of perspectives on mitigation strategies and the specific impact of AI data poisoning on personally identifiable information [1]. These findings highlight the need for robust security measures and the importance of understanding the threat of AI data poisoning in the banking sector. The data from the study suggests that AI data poisoning poses significant risks to data privacy in the context...
of digital banking Chatbots. The interpretations of the data revolve around understanding the complexity of the threat landscape and the various factors contributing to data privacy violations. The study interprets Information Technology (IT) professionals and end-users' experiences and perspectives to identify the challenges and best practices in mitigating AI data poisoning risks.

Several mitigation suggestions emerged, which can be integrated into the conclusions to provide actionable insights for mitigating AI data privacy violations in chatbot systems. These suggestions include:

- **Implementing Robust Encryption**: Strengthening the encryption of data stored and processed by chatbots to protect against unauthorized access and data leaks.
- **Continuous Monitoring and Anomaly Detection**: Establishing systems for continuous monitoring of chatbot interactions and data transactions to quickly identify and respond to unusual patterns or potential security threats.
- **Data Backup and Safe Storage**: Regularly backing up important data and ensuring its safe storage as a precautionary measure against data breaches or loss.
- **User Education and Awareness**: Enhancing user awareness about potential data privacy risks and educating them on safe practices when interacting with AI chatbots.
- **Regular Security Audits and Assessments**: Conducting periodic security audits and assessments of chatbot systems to identify vulnerabilities and address them proactively.
- **Adoption of Best Practices and Standards**: Encouraging the adoption of industry best practices and compliance with data protection standards to ensure the chatbot systems are designed and operated securely.
- **Collaboration with Cybersecurity Experts**: Engaging with cybersecurity experts to gain insights into emerging threats and to develop sophisticated defense mechanisms against AI data poisoning.
- **Regulatory Oversight and Compliance**: Ensuring strict adherence to data privacy regulations and guidelines to protect user data and to build trust in AI chatbot systems.

The findings of this study suggest that AI data poisoning represents a significant and multifaceted threat to the security of personal information managed by AI-enabled Chatbot systems in the banking sector. The diverse experiences and perceptions reported by study participants highlight that the challenge of AI data poisoning is not uniform across the banking sector. Different roles within banks—from information technology (IT) specialists and data privacy officers to customer service representatives—have encountered and understood the threat of AI data poisoning in varied ways. This variability affects not only the perceived severity of the threat but also the perceived effectiveness of mitigation strategies implemented to combat it [1]. For instance, technical staff may deeply understand the risks and the technical measures needed to counter them. In contrast, customer-facing staff may prioritize the immediate impacts on customer trust and service quality.

Moreover, the effectiveness of current mitigation strategies against AI data poisoning appears to vary, suggesting that there is no one-size-fits-all solution to this problem. Strategies such as data encryption, robust access controls, and continuous monitoring can provide substantial protection [1]. However, their effectiveness can be contingent upon factors such as the sophistication of the attack, the specific architecture and deployment of the AI system, and the ongoing maintenance and updating of security measures. Additionally, the human element—such as staff training, awareness, and the cultivation of a proactive security culture—plays a critical role in identifying and responding to threats promptly and effectively [110].

This multifaceted challenge underscores the need for a holistic approach to cybersecurity in the banking sector, one that combines advanced technological defenses with strong organizational practices. It also highlights the importance of collaboration across different departments within banks and with external stakeholders, including technology providers, regulatory bodies, and cybersecurity experts. Such collaboration can facilitate the sharing of best practices, the development of industry-wide standards, and the rapid dissemination of information about emerging threats and
vulnerabilities. In light of these considerations, it becomes clear that addressing AI data poisoning in digital banking AI-enabled Chatbots requires ongoing vigilance, investment in cutting-edge security technologies, and a commitment to fostering a culture of security awareness throughout the organization.

LIMITATIONS OF THE STUDY

The potential constraints encountered include limited data availability and the specific context and characteristics of the sample used, which could impact the generalizability of the study's findings. The rapidly changing nature of AI technology and AI-enabled chatbots presents uncertainties and possible biases that could influence the research outcomes. Despite these limitations, the study intended to provide a comprehensive analysis and interpretation of its findings, which were supported by a rigorous approach to data collection, analysis, and interpretation. The ultimate goal is to enhance understanding of how to effectively mitigate the risks associated with AI-enabled chatbots data poisoning, thereby safeguarding customer data privacy in the digital banking sector.

IMPLICATIONS FOR PRACTICE

For practitioners, the study's insights can inform the development of more secure AI enabled Chatbot systems, emphasizing the importance of ethical considerations and legal compliance. Policy-makers can use the findings to craft regulations and standards that protect consumer data privacy and promote transparency in AI-enabled Chatbot deployments. The study's recommendations can also guide the banking sector in implementing best practices for risk management and data protection.

IMPLICATIONS OF STUDY AND RECOMMENDATIONS FOR FUTURE RESEARCH

The knowledge gained from this study can be furthered by exploring additional mitigation strategies, conducting quantitative research to complement the qualitative findings, and examining the effectiveness of different security measures in various banking environments [1]. Future research could also investigate the long-term impacts of AI data poisoning on customer trust and the financial sector's reputation. Future studies could explore the effectiveness of specific mitigation strategies in more detail, examining how different technologies and approaches can be optimized to address AI data poisoning. One promising direction is the in-depth analysis of specific mitigation strategies to combat AI data poisoning. This involves assessing the technological effectiveness of various security measures and understanding how these can be seamlessly integrated into existing banking systems. Such studies could explore advanced encryption techniques, anomaly detection algorithms, and machine learning models that are resilient to poisoning attacks. The goal would be to pinpoint how these technologies can be fine-tuned and applied most effectively to protect sensitive customer data and maintain the integrity of AI-enabled chatbot interactions.

Investigating the psychological and organizational factors that contribute to successful cybersecurity cultures within banking institutions could also yield valuable insights. This area of inquiry could include how awareness, attitudes, and behaviors toward cybersecurity are cultivated and how they influence the overall security posture of an organization. Investigating factors such as leadership commitment to cybersecurity, employee training programs, and the role of incentives or sanctions in promoting security-conscious behavior could provide a comprehensive view of what constitutes a robust cybersecurity culture.

This research contributes to knowledge by providing insights into the challenges posed by AI data poisoning and its implications for data privacy in digital banking AI-enabled Chatbots. It advances the understanding of the threat landscape and offers a conceptual framework for addressing these challenges. This study enriches the body of literature on AI, data privacy, and cybersecurity within the context of the banking industry.
CONCLUSION

The study explored the impact of AI data poisoning on data privacy violations in digital banking chatbots. Employing a qualitative approach grounded in AI, data privacy, and cybersecurity theories, the study targets Information Technology (IT) professionals in the banking sector. It uncovers the varied impacts of AI data poisoning across different professional roles, from direct breaches to indirect exposures. The key findings revealed a spectrum of mitigation strategies, from technical solutions to basic awareness and mixed responses regarding the impact on Personally Identifiable Information (PII), underscoring the complexity of safeguarding customer data [1].

This qualitative grounded theory study uncovered various mitigation strategies to combat AI data poisoning and its impacts on data privacy, particularly regarding personally identifiable information (PII). The conclusions emphasize the critical role of technical solutions in safeguarding against data privacy violations. Advanced techniques such as encryption, anomaly detection, and secure data storage have been pinpointed as essential components of a robust security framework for AI chatbots. These technical measures are foundational in creating a secure environment where customer data is protected from unauthorized access and potential breaches.

Equally important is the focus on awareness and training. The study highlights the necessity for ongoing educational programs targeted at both employees and customers. By fostering an environment of knowledge and vigilance, stakeholders are better equipped to recognize and mitigate the risks associated with AI enabled chatbots, thereby preventing potential data privacy violations.

The significance of regular software updates is also emphasized. Keeping AI systems and their underlying platforms meticulously updated with the latest security patches is critical in addressing vulnerabilities that could be exploited for data poisoning. This practice ensures that AI-enabled chatbots operate within a secure and updated ecosystem, minimizing the risk of security breaches. Strict access controls and authentication procedures form another cornerstone of the recommended mitigation strategies (Krause, 2023). Implementing comprehensive access management systems ensures that only authorized personnel have access to sensitive data and AI systems, significantly reducing the risk of unauthorized data access and potential privacy violations.

Lastly, the study underscores the importance of compliance with regulations, such as the General Data Protection Regulation (GDPR) and the newly released European Union (EU) AI Act. Adherence to data privacy laws and ethical guidelines is essential in guiding the collection, storage, and processing of personally identifiable information (PII). Compliance ensures legal adherence and fosters trust between banking institutions and their customers, ensuring that data privacy remains a top priority in deploying and operating AI chatbots.

Despite the rapidly evolving AI landscape, which poses potential limitations, the study contributes valuable insights into effective strategies for mitigating AI data poisoning risks, enhancing the security of AI chatbots in banking, and highlighting the critical importance of developing robust security measures to protect sensitive customer data against privacy violations.

ACKNOWLEDGMENTS

I would like to express our sincere gratitude to all the participants who participated in this research study and generously shared insights and experiences. The valuable contributions were instrumental in enriching our understanding of the impact of Artificial Intelligence (AI) on customer satisfaction in the banking sector.

I would also like to extend my appreciation to the professional networks that supported this research by providing access to members and facilitating data collection. Collaboration and cooperation were essential in the successful completion of this study.

Furthermore, I would like to acknowledge the guidance and support of my research advisors, Dr. Joseph Issa, Dr. Melinda Lyles, and Dr. Sandra Fonseca, who provided valuable input and feedback throughout the research process. The expertise and encouragement greatly contributed to the quality and rigor of this study.
This research would not have been possible without the collective efforts and contributions of all those mentioned above. Thank you for your valuable support and collaboration.

REFERENCES


2. C. Amama, U. Okengwu, “Smart Chatbot System for Banking using Natural Language Processing Tools,” 2023. DOI: 10.21203/rs.3.rs-3285543/v1


44. B. Monroe, “A Qualitative Examination of the Influence of AI on Social Media Companies,” Journal of Student Research, vol. 12, no. 1, 2023. DOI: 10.47611/jsrshs.v12i1.3890
50. E. Babulak, “Introductory Chapter: Journey to AI Driven Chatbots. Chatbots - The AI-Driven Front-Line Services for Customers”, 2023. DOI: 10.5772/intechopen.112461


91. V. Kepuska, G. Bohouta, “Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home),” 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), 2018. DOI: 10.1109/ccwc.2018.8301638


96. K. Ralston, “Personalizing Chatbot Conversations with IBM Watson,” Inquiry@Queen’s Undergraduate Research Conference Proceedings, 2019. DOI: 10.24908/iqucrp.13261


APPENDIX A – INSTRUMENT UTILIZED

Survey Questions and Informed Consent Form

Al data privacy violation in chatbot systems in the banking sector

Before you decide whether or not to participate, it is important that you understand why the research is being done and what it will involve. Please take the time to read the following information carefully.

The purpose of this study is to identify best practices to mitigate Artificial Intelligence (AI) data poisoning in AI Chatbots, often leading to data privacy violations. Particularly in the context of Chatbots utilized in customer service interactions requesting and potentially mishandling personally identifiable information (PII) such as credit card numbers.

We’ll also collect general demographic information, which is completely confidential and will be used for internal academic research only.

Participation in this study is completely voluntary. You have the right to withdraw at any time without penalty.

Only the researcher has access to the data collected.

Disclaimer:
This research study has been reviewed and approved by the University of Fairfax (UoF) Institutional Review Board (IRB). The IRB has determined that this study meets the ethical obligations required by federal law and institutional policies.

Please note that the findings and conclusions of this research are the sole responsibility of the researcher and do not necessarily represent the views of UoF.

⏰ 22 minutes

This is a timed form.

Once you start, you can’t pause the timer. Don’t worry, Forms gives you a final minute reminder before submission. Your answers will be automatically submitted when the time is up. Please prepare before you begin to help manage your submission time.

Click to start

Microsoft 365

This content is created by the owner of the form. The data you submit will be sent to the form owner. Microsoft is not responsible for the privacy or security practices of its customers, including those of this form owner. Never give out your password.

Microsoft Forms | AI-Powered surveys, quizzes and polls. Create my own form
Privacy and cookie policy | Terms of use
Consent to data collection

1. I have read and understood the above information. My questions have been answered to my satisfaction. I agree to participate in this study.

☐ YES

☐ NO

Next
2. Can you please specify your current role and describe your specific responsibilities and experience in the fields of AI, Chatbots, banking, and data privacy? How long have you been involved in these areas?

Enter your answer

3. Can you describe any experiences or incidents where AI data poisoning has impacted data privacy within your organization’s chatbot systems?

Enter your answer

4. What strategies or practices have you implemented or observed that effectively mitigate data privacy violations caused by AI data poisoning in chatbot systems?

Enter your answer

5. In your experience, how does AI data poisoning influence the handling of personally identifiable information such as credit card numbers, social security numbers, dates of birth, or account numbers in chatbot systems?

Enter your answer
The purpose of this study is to identify best practices to mitigate Artificial Intelligence (AI) data poisoning in AI Chatbots, often leading to data privacy violations. Particularly in the context of Chatbots utilized in customer service interactions requesting and potentially mishandling personally identifiable information (PII) such as credit card numbers.

We'll also collect general demographic information, which is completely confidential and will be used for internal academic research only.

Participation in this study is completely voluntary. You have the right to withdraw at any time without penalty.

Only the researcher has access to the data collected.

Disclaimer:
This research study has been reviewed and approved by the University of Fairfax (UoF) Institutional Review Board (IRB). The IRB has determined that this study meets the ethical obligations required by federal law and institutional policies.

Please note that the findings and conclusions of this research are the sole responsibility of the researcher and do not necessarily represent the views of UoF.

* Required

Familiarity with AI Chatbots

6. Can you describe your experience with AI Chatbots in the banking industry? *

Enter your answer

7. How often do you interact with AI Chatbots? *

- Daily
- Weekly
- Monthly
- Yearly

Back  Next
Perceptions and Attitudes

8. How do you perceive the relationship between AI data poisoning and data privacy violations in chatbot systems? Can you share specific examples or scenarios? *

Enter your answer

9. How do you feel about the data privacy measures currently in place for AI Chatbots? *

Enter your answer
Experiences and Concerns

10. Can you share any specific incidents where you felt your data privacy was compromised while using AI Chatbots?

* 
Feel free to share specifics.

Enter your answer

11. What are your main concerns regarding data privacy in the context of AI Chatbots? *

Enter your answer

12. What ethical considerations do you think are important in the development and deployment of AI in the banking sector, especially concerning chatbots? *

Enter your answer

13. How do you think AI chatbots have transformed operations in the banking sector, particularly in relation to data privacy and security? *

Enter your answer

14. How does user trust in AI chatbots get affected by incidents of data privacy violations due to AI data poisoning? *

Enter your answer

15. What are the main challenges you face in ensuring data privacy in AI chatbot systems in the banking sector? *

Enter your answer
Recommendations Based on Expertise and Experience

16. Based on your expertise and experience, what specific measures would you recommend to mitigate data privacy violations caused by Artificial Intelligence (AI) data poisoning in AI-enabled Chatbot systems that mishandle personally identifiable information in the banking industry? *

Enter your answer

17. Can you suggest any best practices or standards that should be adopted by banking institutions to ensure data privacy in AI Chatbots? *

Enter your answer

18. Are there any specific technologies or techniques that you believe could improve data privacy in AI Chatbots? *

Enter your answer

19. What steps should banking institutions take to educate users about data privacy when using AI Chatbots? *

Enter your answer

20. How can banking institutions better communicate their data privacy policies and practices to users of AI Chatbots? *

Enter your answer
Assessing the Effectiveness of Various Security Measures

21. On a scale of 1-5, how effective do you believe the current data privacy measures in AI Chatbots are? *

   1   2   3   4   5

22. Have you noticed any improvements in data privacy in AI Chatbots over the past few years? If so, what measures do you believe have contributed to this improvement? *

   Enter your answer

23. How effective do you believe encryption techniques are in ensuring data privacy in AI Chatbots? *

   Enter your answer

24. How effective do you believe user education and awareness campaigns are in enhancing data privacy in AI Chatbots? *

   Enter your answer

25. How effective do you believe regulatory oversight and compliance are in ensuring data privacy in AI Chatbots? *

   Enter your answer

26. Is there anything else you would like to share about your experience or thoughts on AI Chatbots and data privacy? *

   Enter your answer
All of your information is highly confidential and for internal use only. Please contact us if you have any concerns.

27. What gender do you identify as? *
- Woman
- Man
- Non-binary
- Prefer to not say

28. What is your age? *
- < 18
- 18-25
- 26-35
- 36-45
- 46-55
- > 55
- Prefer to not say

29. Which will best describe your status? *
- Student
- Employed
- Unemployed
- Retired
- Other
APPENDIX B – DATA ANALYSIS

Open Coding of the Survey Responses

Concern for Insufficiency and Uncertainty

A prevailing sentiment among respondents is a concern over the insufficiency of current data privacy measures. Descriptors like "Good but relatively insignificant" and "not safe" articulate a widespread apprehension that existing protocols do not adequately protect against data breaches or misuse. This unease is compounded by a palpable uncertainty, with several participants expressing doubts about the effectiveness and success of the measures in place. Such feelings of insecurity and skepticism underline the need for more robust and transparent data privacy frameworks that can reassure users and stakeholders about the integrity of their data.

Necessity and Positivity Towards Improvement

Despite these concerns, there is also a thread of acknowledgment regarding the necessity of data privacy measures. Some responses hint at a positive outlook towards the possibility of enhancement and refinement of these measures. This recognition of the importance of data privacy, coupled with suggestions for improvement, indicates an openness to evolving and strengthening privacy protocols to address the dynamic challenges posed by AI technologies.

Main Concerns

Respondents articulated specific concerns related to data privacy, notably regarding who has access to sensitive information. Fears that "everyone in the company will have access" spotlight worries about internal controls and the potential for unauthorized data exposure. Additionally, a lack of transparency or understanding of how data is managed emerged as a significant concern, with some participants admitting to not knowing how their data is handled [58]. This highlights a critical gap in communication and education about data privacy practices. Security risks, such as the potential for hacking, were also frequently mentioned, underscoring the ever-present threat of cyber-attacks and the need for vigilant security measures.

Trust Erosion and the Call for Mitigation Measures

Data privacy violations, particularly those due to AI data poisoning, were noted to have a detrimental impact on user trust. The erosion of trust is a critical issue, with respondents stating that such incidents "affect user trust negatively" and "diminish the trust greatly." To counteract this trust deficit, participants offered various suggestions ranging from the implementation of stronger protocols to limiting the data accessible to AI chatbots. These recommendations reflect a desire for a more secure and controlled data environment that can rebuild and maintain user trust.

Technology Gaps and Mixed Perceptions of Effectiveness

When asked about technologies or techniques that could improve data privacy in AI chatbots, many respondents expressed a lack of knowledge or skepticism about the existence of adequate solutions. This response pattern points to a potential gap in awareness or development of technologies specifically geared towards enhancing data privacy in AI systems. Additionally, the mixed perceptions regarding the effectiveness of current data privacy measures, as indicated by the varied ratings on their effectiveness, underscore the subjective nature of satisfaction and confidence in these measures.

Axial Coding of the Survey Responses

Insufficiency & Uncertainty with Current Measures and Concerns Regarding Access Control and Security Risks

There's a clear link between the perceived insufficiency of current data privacy measures and specific concerns like access control and the threat of security breaches. This suggests that enhancing transparency and control mechanisms could directly address the root of these insecurities [58].
Trust Erosion Due to Data Privacy Violations and Suggestions for Mitigation

The negative impact on user trust due to data privacy violations is closely tied to the suggestions for mitigating these issues. Strengthening data protocols and limiting AI access to sensitive data emerge as potential solutions to rebuild trust.

Lack of Specific Technology Suggestions and Mixed Perceptions of Measure Effectiveness

The difficulty in identifying specific technologies for improving data privacy and the mixed perceptions regarding the effectiveness of current measures point to a broader issue of awareness and education. There seems to be a gap in communicating the effectiveness of the existing solutions and potential technological advancements in data privacy.

Core Categories

From these connections, we can identify two core categories that encapsulate the central phenomena represented in the survey data:

1. **Trust and Security in AI Chatbots**: This category encompasses concerns about data privacy measures' insufficiency, the consequent erosion of user trust, and the call for more robust security protocols. It reflects the critical balance between leveraging AI technologies and safeguarding user data privacy, highlighting the need for transparent, secure, and effective privacy measures to maintain and rebuild trust.

2. **Awareness and Knowledge Gap**: This category captures the lack of specific knowledge about technologies to improve data privacy, the mixed perceptions of current measures' effectiveness, and the general uncertainty about data management practices. It underscores the importance of education and communication in bridging the gap between technological capabilities and user understanding, potentially offering a pathway to address concerns and enhance perceptions of data privacy measures.

Axial Coding of the Survey Responses

Central Phenomenon: Balancing Innovation and Privacy

The core narrative that emerges from the data is the ongoing challenge of Balancing Innovation and Privacy in developing and deploying AI chatbots within the banking sector. This central theme encapsulates the tension between leveraging AI for enhanced customer interaction and service delivery and the imperative to protect user data privacy.

Integration of Core Categories

- **Trust and Security in AI Chatbots**: This category directly feeds into the central theme, illustrating the critical importance of trust for adopting and accepting AI technologies. The erosion of trust due to privacy concerns or violations points to a need for robust security measures to reassure users about their data's safety. This includes clear protocols, transparency about data use, and measures to prevent data breaches or misuse [58].

- **Awareness and Knowledge Gap**: The lack of awareness and understanding about data privacy measures and technologies underscores the necessity for education and communication. Bridging this gap is essential for improving the perception of AI chatbots' data privacy measures and fostering an environment where innovation can thrive without compromising privacy. Educating users and developers about privacy-enhancing technologies and practices can empower all stakeholders to contribute to a more secure and privacy-respecting ecosystem.

Constant Comparative Method

This method was applied across the Open, Axial, and Selective Coding stages conducted for the survey on AI chatbots and data privacy in the banking sector.

Open Coding

During Open Coding, initial themes were identified, such as concerns over the insufficiency of current privacy measures, fears regarding access control, and the impact of data privacy violations on user trust. The Constant Comparative Method was applied to continuously compare these initial codes to
identify similarities, differences, and patterns. For example, concerns about the insufficiency of privacy measures and fears about access control both point to a broader theme of security concerns. Similarly, comparing responses across different demographics or roles may reveal varying levels of concern or awareness about data privacy, further refining the understanding of these initial themes.

**Axial Coding**

In Axial Coding, the researcher looked at how these themes connect, focusing on categories like Trust and Security in AI Chatbots and Awareness and Knowledge Gap. By comparing these categories against each other and against the initial codes from Open Coding to refine their definitions and relationships. For instance, the relationship between the Awareness and Knowledge Gap and the erosion of trust becomes clearer when specific examples of how lack of transparency or understanding directly impacts user trust are considered [58]. This comparison helps to refine the categories further, emphasizing the critical role of communication and education in addressing privacy concerns.

**Selective Coding**

Selective Coding identified the central theme of Balancing Innovation and Privacy. Applying the Constant Comparative Method here involves comparing this central theme with all the categories and codes identified earlier to ensure that it fully encompasses and explains the data. This might lead to revisiting some of the connections between trust and the knowledge gap, questioning whether additional categories or themes emerge that could further enrich the central narrative. For instance, comparing the emphasis on security measures with the importance of transparency and education might reveal additional nuances in how these factors interact to influence the balance between innovation and privacy [58].

**Integrating the Constant Comparative Method Across Stages**

Applying the Constant Comparative Method across all stages allowed for a dynamic and evolving analysis. As the researcher refined codes, categories, and central themes, continuously returned to the data to ensure that the analysis remained grounded and reflective of the respondents’ perspectives. This iterative process revealed deeper insights, such as identifying specific types of communication most effective in addressing privacy concerns or highlighting particular technologies that promise to enhance data security without stifling innovation.

**Memos**

**Open Coding Memo**

**Date:** April 7, 2024

**Observations:** Open Coding unearthed a range of concerns and suggestions about AI chatbots’ data privacy in banking. Key themes include the perceived insufficiency of privacy measures, access control fears, and the nuanced impact on user trust. This stage highlighted the complexity of respondents’ views, emphasizing a broad spectrum of experiences and expectations regarding data privacy.

**Insights:** The diversity in responses underscores the challenge of creating universally satisfying privacy measures. Concerns about security and transparency are paramount, indicating areas for immediate attention and improvement [58].

**Axial Coding Memo**

**Date:** April 7, 2024

**Observations:** Axial Coding connected initial themes into broader categories: Trust and Security in AI Chatbots, and Awareness and Knowledge Gap. These categories reflect the intertwined nature of trust, security, and the need for greater transparency and education regarding AI technologies.

**Insights:** The linkage between trust erosion and security concerns points to a critical need for robust privacy protocols and better communication strategies. Addressing the Awareness and Knowledge Gap is essential for fostering a more informed user base and rebuilding trust.
Selective Coding Memo

Date: April 7, 2024

Observations: Selective Coding distilled the analysis into a central theme: Balancing Innovation and Privacy. This theme captures the overarching challenge of integrating AI chatbots in banking without compromising data privacy, emphasizing the necessity of a balanced approach.

Insights: The balance between innovation and privacy is delicate but achievable. Prioritizing both technological advancement and stringent privacy measures is key to sustaining user trust and facilitating the responsible use of AI in banking.