Journal of Computer Science and Frontier Technologies
ISSN Print: 3104-4204
ISSN Online: 3104-4212
CD ISSN: 3104-4220
https://doi.org/10.63313/JCSFT.2003

Universal Village Oriented Solution for Elderly Emotional Support

An-Ting Hsieh
VIS International Experimental School, 3F, No. 175, Sec. 2, Anhe Rd. Taipei, Taiwan 106, China
Email: tinghsieh07@gmail.com

How to cite this paper: Hsieh, A.-T. (2026). Universal Village Oriented Solution for Elderly Emotional Support. Journal of Computer Science and Frontier Technologies, 2(2), 47-58. ISSN Print: 3104-4204; ISSN Online: 3104-4212.
https://doi.org/10.63313/JCSFT.2003
Published: 2026-02-02

Abstract

As the global population continues to age, the number of the elderly significantly grows. The aging of the body, combined with the awareness of physical decline, can negatively affect the elderly's mental health [4]. The negative impacts on physical and mental health caused by loneliness and social isolation could increase the use of healthcare services, leading to burdens on the healthcare system [2]. The gradually decreasing potential support ratio also increases the pressure on society to provide care for the elderly [1]. This means that the problems of elderly care are becoming a growing challenge for the working-age population. Loneliness could be improved by interactions [2], and the emergence of large language models (LLMs) allows smart technologies to manage interactions. In the past, chatbots based on smart technology had limited levels of personalization and emotional interaction. However, achieving greater personalization requires collecting more user information, which can lead to privacy concerns [3]. This study uses a literature review approach to examine existing research in smart technologies from Universal Village's perspectives. We propose an elderly emotional support chatbot with emotional understanding and personalized memory management that can effectively mitigate LLM hallucination, enhance information security, and increase privacy protection. This could enhance elderly's mental health and improve the quality of life, reducing the burden on family caregivers and the healthcare system.

Keywords

Smart healthcare; conversational AI; Emotional Support; Universal Village; elderly care; sustainability; AI hallucination mitigation; closed feedback control loop; information security; multi-model information; data analysis

1. Introduction

The global trend of an aging population has become increasingly evident. Due to social isolation, the elderly are prone to feeling lonely and experiencing negative emotions. Loneliness brings psychological stress to them, influencing mental and physical health [2]. Social robots that provide emotional support and companionship can help reduce loneliness and depression among elderly [7], and the rapid development of AI technology has made it possible to provide personalized and emotionally responsive companionship within smart healthcare systems. However, this poses challenges in information security, privacy, and LLM hallucination. This paper proposes suggestions for applying smart technology to elderly emotional support, along with theoretical and technical frameworks of how to improve personalization, enhance information security, and reduce AI hallucinations.

The study investigates five main dimensions.

1.1. Research Dimensions

Analysis of Elderly's Emotional Support Needs and Technological Gaps: This dimension examines the issue of loneliness among the elderly in the context of global population aging. It also analyzes the limitations and challenges of current smart healthcare technologies in providing effective emotional support.

Exploration of an Integrated Emotional Support Chatbot Framework: Based on existing research, this part explores the design of a chatbot system that combines four key technical modules: emotion recognition, memory management, information security, and hallucination mitigation.

Conceptualization of a Technological Integration Solution: This section investigates how advanced methods—such as recurrent neural network (RNN) and long short-term memory (LSTM)-based emotion recognition, conditional memory management, multi-layer encryption, and Retrieval-Augmented Generation (RAG)—can be combined to build a comprehensive emotional support system for the elderly.

Integration within the Universal Village (UV) Framework: This dimension analyzes how the proposed emotional support chatbot could be incorporated into the UV smart healthcare subsystem, with particular attention to its roles and contributions in information flow and material cycles.

Identification of Research Limitations and Future Directions: This part evaluates the technical, cross-cultural, and ethical limitations of the proposed solution, while also exploring potential pathways for future research and development.

By achieving the above research objectives, this study aims to provide a theoretical foundation and technology guidelines for the applications of smart healthcare technologies in the field of elderly emotion care and lay the foundation for future empirical research and system development.

2. Background and Motivation

2.1. Global Population Aging Trend

Population aging, population growth, urbanization, and migration are the four megatrends that significantly affect the change in the global population, significantly impacting the economy, social development, and environmental sustainability [1]. Due to the continuous decrease in the fertility rate and increase in life expectancy, the populations in almost all countries and regions are aging [4]. In 2018, the global population of people aged 65 and over surpassed the number of children under the age of 5 for the first time. In 2050, people aged 65 and over are estimated to be double the number of children under 5, also exceeding the population aged between 15 and 24 [1]. The population aged 80 or over grows faster. It grew from 54 million to 143 million from 1990 to 2019, and it is estimated to be 426 million in 2050, which is triple the population in 2019 [1]. The aging population is expected to have a huge impact on the labor market, economic performance, public healthcare, pensions, and social services [1].

2.2. Elderly Loneliness and Influence on Health

Research shows that a considerable portion of the elderly population faces loneliness and social alienation, with some consistently in a state of loneliness. When they have chronic diseases, the occurrence of loneliness significantly rises. Loneliness leads to multiple negative impacts on the older population. In the mental health aspect, loneliness is closely linked to emotional distress. Moreover, lonely elderly tend to use the medical service more frequently, and their quality of life is also affected. The consequences of long-term loneliness are much more serious, potentially leading to a series of vicious circles of health issues, including abnormal physiological indicators, increasing unhealthy habits, and degeneration in important body functions. These findings emphasize that the loneliness of the elderly is not just a societal issue but a critical health matter that requires attention from both the healthcare and social service systems [2] [4] [5].

2.3. Current Challenge in the Application of Smart Technology in Mental Support

Some smart healthcare technologies have already been applied to wearable devices, sensors, computer vision, AI, VR, chatbots, and robots [3][4]. However, these technologies primarily focus on tele-healthcare, monitoring/ system, environmental control, emergency response, and smart hospital [3][4]. The past research mainly addressed health monitoring, preventive medicine, laboratory applications, clinical decision support, remote monitoring, and various medical applications. Less attention has been given to emotional support for the elderly. Although social robots and information robots have been applied in various areas, there are still several limitations in providing mental support for the elderly.

2.3.1. Personal Information and Emotion Authenticity

There is a lack of personalized information and humanity [3][5]. Even by applying prompt engineering, the caregiver's interventions are still needed [5], as personal information cannot be accumulated. While the current LLM has great performances on fluency and making sense, improvements are needed in engagingness, interestingness, inquisitiveness, and avoiding repetition [5]. These improvements can help elderly receive more authentic emotional feedback. Additionally, the current LLMs are based on English training, lacking support for other languages and cultures [5].

2.3.2. Information Security and Privacy

There is the issue of personal identity information and privacy protection [3]. LLM, such as ChatGPT, collects data [5]. Research has shown that the ChatGPT model has a high tendency for Input Regurgitation [15], displaying personal sensitive information in the generated summary. This raises the need to balance personalization and privacy in the application of chatbots. In addition, due to the limitations in the elderly's digital literacy, threat actors may extract private data or produce harmful outputs through designed prompts [11]. Prompt injection could even work as a computer virus, leading to prompt infection and data leakage within LLMs [14].

2.3.3. AI Hallucination and Bias

Although LLMs have made significant progress, their tendency to generate hallucinated content has raised concerns. These contents may seem plausible but lack factual accuracy. Because LLMs are able to generate highly realistic and human-like responses, hallucination detection becomes difficult, affecting their practical deployments in chatbots, search engines, recommendation systems, and other information retrieval systems [8]. Furthermore, limitations in training data and the echo chamber effect can lead LLMs to produce biased or incorrect data.

3. Proposed Elderly Emotional Support Chatbot

Current smart technologies are limited in how they provide emotional support. LLMs also have limitations when it comes to mental support. To address these issues, this paper proposes an emotional support chatbot that improves on these weaknesses and explores how it can be integrated into the Universal Village framework as part of a smart healthcare system. Unlike past UV studies that focused on dementia care [3], this research looks at general emotional support for elderly. It also expands the discussion to include emotion regulation, privacy protection, and preventing AI's incorrect responses. The chatbot design uses emotion recognition and memory management to improve personalization and emotional authenticity, multi-layer encryption to ensure information security and privacy, and retrieval-augmented generation (RAG) to reduce hallucinations in LLMs.

3.1. Emotional Recognition

Understanding emotions depends on context, which means past information needs to be remembered. Deep learning is currently a major trend in research, and transformer models and their variants are seen as more advanced in natural language processing (NLP), including emotion detection. They perform especially well in understanding context and handling complex language tasks [6]. Deep learning techniques allow multimodal data such as video, audio, and different physiological signals [6].

Even so, LSTM models, including Bi-LSTM and hybrid models that combine CNN and LSTM, are still widely used [6] and are helpful in time-series tasks.[6]. Recurrent Neural Networks (RNNs) can process sequential data by remembering inputs of different lengths, which makes them useful for predicting and understanding context [6]. Variants like LSTM and GRU are considered effective for modeling speech and video [6], and LSTM and Bi-LSTM are also common methods in text-based emotion analysis [6]. Overall, the design of RNNs and their variants allows them to handle sequential data, capture context, and automatically extract features related to emotions. This makes them suitable for analyzing text, speech, and video, as well as physiological signals like heartbeat, tremors, and sweating, to better detect emotions [6].

For elderly emotional support, data can be collected from conversation text, facial expressions, and voice. Using RNNs and their variants, such as LSTM, GRU, and Bi-LSTM, a multimodal emotion recognition module can be built. Facial images can be captured with a camera, and OpenCV with the Viola-Jones algorithm can be used for detection and recognition [7]. Models like FER-2013, AffectNet, and CK+ [10] can then be applied to identify emotions. In addition, automatic emotion analysis tools such as Stanford's open-source CoreNLP are already well-developed [7].

Finally, combining physiological data with video and audio can further improve text-based emotion recognition accuracy [6].

3.2. Memory Management

Memories of personalized information are essential in expressing emotions and care. Current AI chatbots lack long-term memory, unable to learn from past interactions to provide customized responses according to users' preferences [12]. Different approaches to memory have their own characteristic:

Historical memory: directly stores entire conversation histories without processing as its memory record; limitations include containing redundant information [12].

Summary-based memory: utilizes LLM to generate summaries for every conversation, often loses important details because of oversimplification [12].

Conditional memory: determines whether a statement includes important information, for example, user preferences and domain knowledge, that is unknown to LLM and stores it, ignoring simple prompts, greetings, and non-essential content [12].

When evaluating the three criteria, continuing past conversations, learning new knowledge, and learning from human feedback, conditional memory has the best performance [12]. However, in continuing past conversations, historical memory has better performance in details [12]. A combined approach using conditional memory plus summary-based memory showed improved results across all three indicators [12].

3.3. Information Security

As information is the core of the smart healthcare system, data security and personal privacy are crucial considerations for the information platform [3][4]. For interoperability with legacy medical systems, APIs compliant with FHIR/HL7 standards can be used [11]. For remote communication with other smart healthcare and smart city, especially when involving sensitive medical information, industry standard protocols such as SSL/TLS are recommended. Information security discussed in this study includes deindividuation and personal information protection. The goal of deindividuation is to protect personal identifiable information (PII) while ensuring the utility of prompts, striking a balance between privacy and utility [14]. Adversarial Generative Desensitization is a method using a local LLM as a protection mechanism, performing PII identification, desensitization, and evaluation [14]. Leveraging the generative ability of LLM to replace semantically related identifiers, and use special characters to disrupt private attributes at the same time, making it difficult to replace by automated tools, but still understandable for LLM. By preserving important semantic meaning while breaking the link between identifier and private attribute, this method more effectively balances the privacy protection and the utility of prompts [14]. Apply techniques, such as deletion, masking, and generative replacement (e.g., replacing "John" with "Alice" and inserting special characters to distort the identifier) [14], to achieve high effectiveness in healthcare, finance, and daily data. Experiment results show that the average PII d accuracy of 95.95% [14]. In addition, Prompt Induced Sanitization, guiding ChatGPT to follow HIPAA or GDPR, can also reduce PII leakage in LLM [15].

All personal information and emotional data are stored on the edge of databases, using virtual environments such as Docker to store personal information separately. The stored data are encrypted using the Advanced Encryption Standard (AES), in compliance with the HIPAA security standard. However, LLMs are vulnerable to prompt injection, causing unauthorized or harmful outputs [11]. To prevent edge LLMs from leaking private information under prompt injection attacks, using prompt injection attack detectors is needed. Building on the foundations of the existing detector, one effective strategy is to apply LLM fine-tuning with open-source models and datasets. This can significantly reduce the false positive rate and improve the true positive rate [13].

3.4. Handling AI Hallucination

Recent studies have proposed a variety of methods to mitigate AI hallucinations. These approaches can be divided into two categories: prompt engineering and model development [8]. Since this paper focuses on using public LLMs to design elderly mental support solutions, the discussion is limited to prompt engineering strategies. There are three main strategies in prompt engineering: Retrieval Augmented Generation (RAG), self-refinement through feedback and reasoning, and prompt tuning. RAG strengthens LLMs' responses by leveraging external authoritative knowledge rather than relying on potentially outdated training data. The self-refinement method improves model outputs iteratively, as feedback guides the LLM to produce more accurate results in subsequent responses [8].

RAG is particularly valuable for enhancing real-time and domain-specific knowledge, making it well-suited for elderly care applications involving both physiological and emotional health information. However, RAG also hallucinates due to retrieval failure and generation deficiency [9]. Current research suggests using reliable sources, optimizing query construction, improving retriever performance, and adopting appropriate retrieval policies [9]. To address generation deficiency, methods include reducing contextual noise, resolving knowledge conflicts, effectively utilizing long-text information, ensuring consistency between generated content and queries, and integrating external tools to support the generation process [9].

4. Elderly Emotion Support Chatbot through UV-Perspective

The UV Healthcare Subsystem forms a closed feedback loop through data acquisition, communication, decision-making, and action.

4.1. UV Closed Feedback Loop

4.1.1. Data Acquisition

The emotional support chatbot collects image, voice, and text information. In addition to gathering information on sleep, medication, and stress from the conversations the elderly have with others, semantic analysis can also be used to obtain subjective data. This is further combined with physiological data such as heart rate and blood pressure from wearable monitoring devices in the UV Smart Healthcare System. The integration of these diverse data sources supports both psychological and physical care for the elderly and can build a comprehensive health data platform for smart cities [3]. It may also serve as part of a patient information sharing system in physician-assisting smart healthcare systems [3][4].

4.1.2. Communication

The emotional support chatbot collects data at the edge, meaning directly in the home, and shares it through a local network with the AI processing layer and knowledge enhancement layer. These layers also connect with cloud-based large language models (LLMs) or knowledge databases to exchange data. The system can also link with instant messaging apps like WhatsApp, Messenger, WeChat, and Line to support social communication. In addition, the emotion recognition module combines its data with information from wearable and IoT sensors in the smart healthcare system and then communicates this with tele-healthcare services.

4.1.3. Decision Making

The emotion recognition module, built on RNNs and their variants like LSTM, GRU, and Bi-LSTM, can understand meaning, facial expressions, and tone of voice. This allows it to judge the emotional state of the elderly and even give clues about their physical condition. When combined with physiological data from smart healthcare systems, it can provide a clearer picture of both their mental and physical health.

4.1.4. Action

The system uses voice, facial expressions, tone, and physiological data to understand the emotional and physical state of the elderly. There are two types of actions. The first is for the elderly without major health or psychological problems. For this group, the chatbot mainly works to reduce loneliness, share accurate information, and support preventive care. The second is for the elderly with conditions like dementia, which has been studied in IEEE UV Smart Healthcare research [3]. In these cases, the chatbot can provide personalized and supportive conversations and connect them with family, friends, or healthcare providers through instant messaging. For medical questions, it uses RAG to pull answers from external databases. It also links with the UV Smart Healthcare Subsystem to provide in-home fall alerts and remote tele-healthcare.

4.2. Information Cycle and Material Cycle

In Universal Village Smart Healthcare, information flow is an integral component of all information systems. It is responsible for the transfer of data, information, understanding, wisdom, and vision, while connecting all important UV elements [4]. The material cycle, as the basis of information flow, connects all components of the Universal Village through the "exchange of physical products, energy, and other medical resources." It includes the movement from raw materials, extracted and processed resources, semi-finished products, end products to waste, and recyclable resources within UV Smart Healthcare systems [4]. The elderly emotional support chatbot proposed in this study is based on these two concepts in the Universal Village framework.

4.2.1. Information Flow

The emotional support chatbot developed in this study aligns with the concept of information flow in the following aspects.
Multi-layer Information Integration: Emotional information collected includes text, video, and voice, as well as ambient and physiological data from existing smart

home and smart healthcare systems. This demonstrates capabilities in memory and understanding, and can be categorized into three types of information. First, information generated through the mental support chatbot functions, including personal states and historical dialogue data obtained via emotion recognition technologies. Second, material-related information within smart healthcare subsystems, such as medication reminders and the use of medical devices. Third, human-related information, particularly the health and lifestyle records of the elderly.

Hierarchical Information Processing: The proposed emotional support chatbot uses a memory management approach that combines conditional memory with summary-based memory, utilizing the concepts of hierarchical, hybrid, and integrated information systems. By obtaining data from different sources, including smart monitoring, intelligent agents, and crowdsourcing, the system can decide what type of data to collect and when, depending on specific needs and context.

4.2.2. Material Cycle

In UV Smart Healthcare, the material cycle is the basis of the information cycle, connecting UV components through the exchange of physical products, energy, and other natural and medical resources [4]. When linked with UV Smart Cities, the material cycle must also consider the reduction of resource demand and waste [4]. The emotional support chatbot proposed in this study aligns with the above concepts in the following aspects.

Physical Device Integration: The emotional support chatbot is deployed on edge devices, including cameras, microphones, and processors, which together form a complete material flow cycle—from raw materials, resource extraction and processing, and semi-finished products to final products, waste, and recyclable resources.

Secured Operation: The use of de-identification techniques, docker virtualization, and AES encryption ensures operational security. Multi-layer encryption further assures the secure operation of the devices.

Emergency Preparedness Mechanism: When the system detects abnormal emotional or physical conditions in elderly, it can automatically connect with relevant personnel to initiate an emergency response, aligning with the material cycle's requirement of ensuring preparedness for critical situations. The application of RAG technology allows timely access to and accuracy of medical information.

Resource Optimization and Environmental Impact: Through the use of edge computing and local AI processing layers, the design reduces the frequency and volume of cloud transmissions. This allows personalization and also reduces the feeling of loneliness and stress, which in turn can lower the burden on healthcare resources and reduce medical demand and waste.

In conclusion, the emotional support chatbot proposed is designed to fit into the UV framework by combining the concepts of information flow and material cycle. Using technologies like multimodal emotion recognition, memory management, information security, and methods to reduce LLM errors, the AI system in chatbot brings together information more effectively and makes resource use safer and more efficient, offering a supportive solution for smart healthcare that follows the idea of the Universal Village.

5. Limitations

5.1. Technical Architecture Limitations

In terms of emotion recognition accuracy, RNNs and their variants, such as LSTM and GRU, are good in sequence data processing, but recent research has shifted toward Transformer architectures like BERT and GPT [6]. These models perform well at handling long-term dependencies, multimodal learning, and end-to-end training, yet studies on complex emotion recognition and cultural adaptability are limited [6].

In terms of facial recognition technology, commonly used training datasets (e.g., FER-2013, AffectNet, CK+) are not composed primarily of elderly but also young to middle-aged adults [10], and the effectiveness of these models in recognizing older adults in real home environments remains to be further validated.

In memory management, conditional memory and summary-based memory architectures can improve system response efficiency, but effective methods for information summarization and reducing memory load still require further investigation.

Although retrieval-Augmented Generation (RAG) systems have been widely used to reduce hallucinations in Large Language Models (LLMs) and to improve response quality and relevance, mechanisms for updating external knowledge bases, verifying the credibility of information sources, and resolving conflicts among external information still require further research.

5.2. Cross-Culture and Language Limitation

Current emotion recognition models may have cultural bias, as most are trained primarily on English-language datasets and therefore lack a deep understanding of emotional expressions, euphemism use, and culture-specific affective patterns in non-English contexts. In many non-Western cultures, for example, older adults tend to express emotions more subtly compared to the more direct style common in Western societies. Such cultural differences may limit the accuracy of existing models in capturing context-specific emotional expressions, resulting in misclassification or inappropriate responses.

5.3. Ethics and Privacy

While privacy concerns are addressed, this paper does not consider how to ensure understanding and consent among the elderly whose cognitive abilities may gradually decline, particularly regarding the collection and use of data. This issue is especially critical when sensitive information, such as emotional and health data, is involved. Furthermore, when the AI system determines that family members or healthcare professionals should be contacted, the basis of such judgments and the triggering mechanisms require further study.

It should also be noted that the proposed design in this paper lacks technical validation and relies entirely on literature-based recommendations.

6. Future Direction

Consider a hybrid use of Transformer variants and LSTM variants to fully utilize their respective strengths in capturing semantic relationships and sequential dependencies.

Develop emotion recognition models for different languages and dialects. Cultural variations in attitudes toward privacy, family dynamics, perceptions of authority, and help-seeking behaviors need to be taken into consideration. In addition, cross-cultural differences in speech prosody, lexical choice, and nonverbal expressions are important factors that should also be addressed.

Incorporate physiological information to enhance multi-model emotion detection capabilities, as physiological signals are directly associated with emotion [6].

Investigate technology acceptance and usage patterns among the elderly at different stages of cognitive ability and develop adaptive interface designs that adjust interaction complexity and feedback mechanisms accordingly.

Establish a long-term tracking mechanism to examine the effects of emotional support chatbots on older adults' mental health, social participation, and overall quality of life, in comparison with traditional care methods.

Develop AI systems that facilitate collaboration among healthcare professionals, social workers, and family members.

References

[1] United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Highlights.

[2] Berg-Weger M, Morley J. Editorial: Loneliness in Old Age: An unaddressed Health Problem. J Nutr Health Aging. 2020;24(3):243-245. doi: 10.1007/s12603-020-1323-6. PMID: 32115602; PMCID: PMC7223173.

[3] Zhang, K., Yuan, H., & Fang, Y. (2022, October). Evaluation of smart agitation prediction and management for dementia care and novel universal village-oriented solution for integration, resilience, inclusiveness and sustainability. In 2022 6th International Conference on Universal Village (UV) (pp. 1-34). IEEE.

[4] Zhang, L., Ren, J., Yuan, H., et al. (2020).Evaluation of Smart Healthcare Systems and Novel UV-Oriented Solution for Integration, Resilience, Inclusiveness and Sustainability.In:Proceedings of the 2020 5th International Conference on Universal Village (UV), 24–27 October 2020, Boston, NY: IEEE, pp. 1–28.

[5] Alessa, A., & Al-Khalifa, H. (2023).Towards Designing a ChatGPT Conversational Companion for Elderly People. arXiv preprint arXiv:2304.09866.

[6] Chutia, T., & Baruah, N. (2024). A review on emotion detection by using deep learning techniques. Artificial Intelligence Review, 57(8), 203.

[7] Abdollahi, H., Mahoor, M. H., … (2022).Artificial Emotional Intelligence in Socially Assistive Robots for Older Adults: A Pilot Study.arXiv preprint arXiv:2201.11167.

[8] Tonmoy, S. M. I., Zaman, S. M. M., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2023).A Comprehensive Survey of Hallucination Mitigation Techniques in Large Language Models.arXiv preprint.

[9] Zhou, Y., Li, X., & Zhang, L. (2025). Hallucination mitigation for retrieval- augmented large language models: A review. Applied Sciences, 15(5), 856.

[10] Khalil, R. A., & Amin, S. (2025). Deep learning-based facial expression recognition for the elderly. arXiv preprint arXiv:2502.02618.

[11] Khalil, R. A., Amin, S., & Smith, J. (2025). Redefining elderly care with agentic AI: Challenges and opportunities. arXiv preprint arXiv:2507.14912.

[12] Yuan, R., Sun, S., Li, Y., Wang, Z., Cao, Z., & Li, W. (2023). Personalized large language model assistant with evolving conditional memory. arXiv preprint arXiv:2312.17257.

[13] Dennis Jacob, Hend Alzahrani, Zhanhao Hu, Basel Alomair, and David Wagner. 2025. PromptShield: Deployable Detection for Prompt Injection Attacks. In Proceedings of the Fifteenth ACM Conference on Data and Application Security and Privacy (CODASPY '25). Association for Computing Machinery, New York, NY, USA, 341–352. https://doi.org/10.1145/3714393.3726501

[14] Sun, X., Liu, G., He, Z., Li, H., & Li, X. (2024). Deprompt: Desensitization and evaluation of personal identifiable information in large language model prompts. arXiv preprint arXiv:2408.08930.

[15] Priyanshu, A., Vijay, S., Kumar, A., Naidu, R., & Mireshghallah, F. (2023). Are chatbots ready for privacy-sensitive applications? an investigation into input regurgitation and prompt-induced sanitization. arXiv preprint arXiv:2305.15008.