

Viewpoint

Developing Effective Frameworks for Large Language Model–Based Medical Chatbots: Insights From Radiotherapy Education With ChatGPT

James C L Chow^{1,2*}, PhD; Kay Li^{3*}, PhD

¹Department of Medical Physics, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

²Department of Radiation Oncology, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

³Department of English, Faculty of Arts and Science, University of Toronto, Toronto, ON, Canada

* all authors contributed equally

Corresponding Author:

James C L Chow, PhD

Department of Medical Physics

Princess Margaret Cancer Centre

University Health Network

Rm 7-606, 7/F, 700 University Ave

Toronto, ON, M5G 1X6

Canada

Phone: 1 416 946 4501

Fax: 1 416 946 6566

Email: james.chow@uhn.ca

Abstract

This Viewpoint proposes a robust framework for developing a medical chatbot dedicated to radiotherapy education, emphasizing accuracy, reliability, privacy, ethics, and future innovations. By analyzing existing research, the framework evaluates chatbot performance and identifies challenges such as content accuracy, bias, and system integration. The findings highlight opportunities for advancements in natural language processing, personalized learning, and immersive technologies. When designed with a focus on ethical standards and reliability, large language model–based chatbots could significantly impact radiotherapy education and health care delivery, positioning them as valuable tools for future developments in medical education globally.

(*JMIR Cancer* 2025;11:e66633) doi: [10.2196/66633](https://doi.org/10.2196/66633)

KEYWORDS

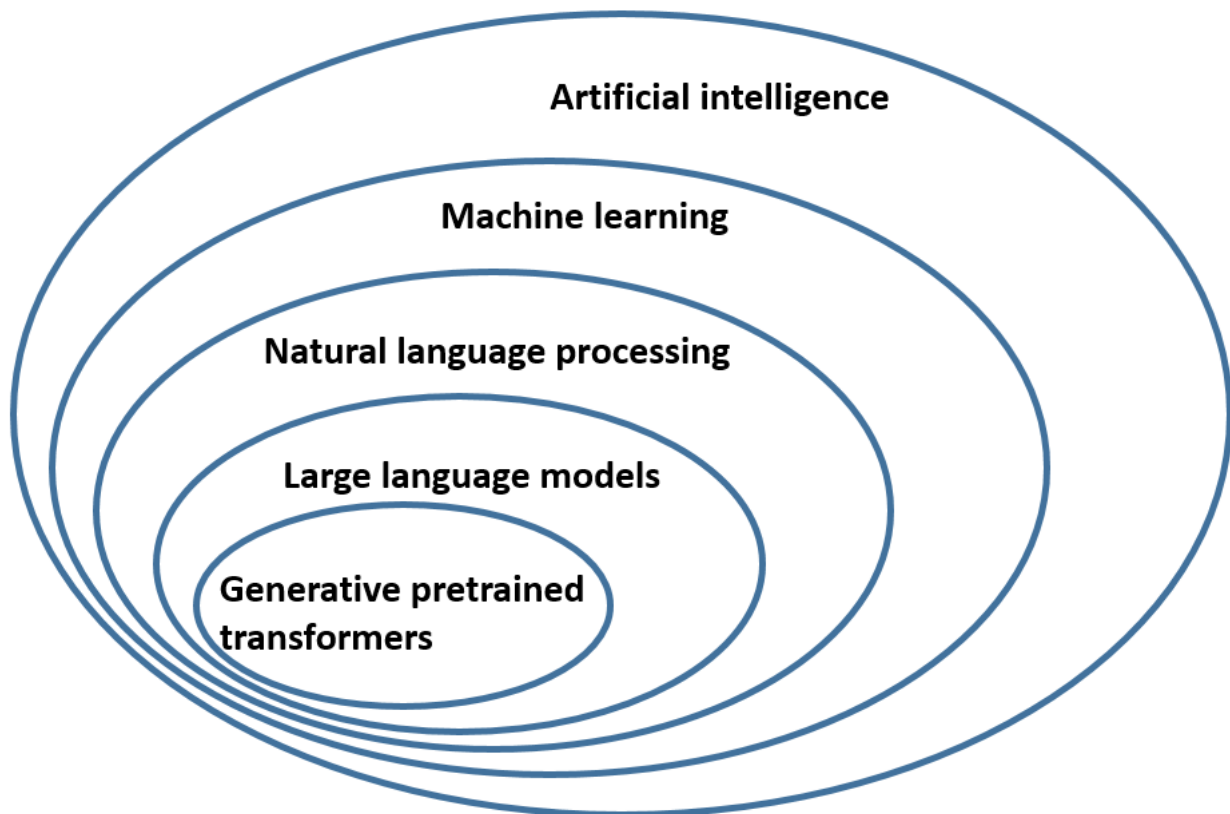
artificial intelligence; AI; AI in medical education; radiotherapy chatbot; large language models; LLMs; medical chatbots; health care AI; ethical AI in health care; personalized learning; natural language processing; NLP; radiotherapy education; AI-driven learning tools

Introduction

The integration of artificial intelligence (AI) in health care has rapidly evolved, with large language models (LLMs) such as ChatGPT at the forefront of this change [1-3]. These models, trained on vast datasets, have shown remarkable potential in various domains, including health care, by understanding and generating human-like text [4]. Unlike traditional rule-based

chatbots, LLM-based systems can engage in complex conversations, answer medical queries, assist in symptom checking, and provide personalized health advice [5]. This capability is largely due to their training on extensive medical literature and patient data [6], which allow them to access and synthesize information in ways that were previously unimaginable [7]. The relationship between AI, machine learning, natural language processing (NLP), LLMs, and generative pretrained transformers is shown in [Figure 1](#).

Figure 1. Relationships between key concepts in artificial intelligence, including machine learning, natural language processing, large language models, and generative pretrained transformers.



The deployment of LLM-based chatbots in health care is driven by the growing demand for accessible information and the need to alleviate the burden on health care systems [8]. These chatbots operate 24-7, providing immediate responses to patient inquiries; enhancing patient education; and supporting telemedicine, mental health, and chronic disease management [9-13]. However, their integration also presents challenges, such as ensuring the accuracy of information, avoiding biased responses, and safeguarding patient privacy [14].

Radiotherapy is a specialized field within oncology that demands precise and accurate information for effective treatment planning and education [15]. Unlike general medical topics, it involves complex concepts such as dose distribution, radiation physics, and the biological effects of ionizing radiation. Inaccurate information can lead to severe consequences, including inadequate treatment or harm to patients, making reliable dissemination critical [16].

Radiotherapy education serves 3 primary groups: patients, radiation staff, and the public. Patients require clear and empathetic guidance to understand their treatment, manage expectations, and make informed decisions, improving adherence and outcomes [17]. For health care professionals, education focuses on advanced training in treatment planning, safety protocols, and technological advancements to ensure high standards of care [18]. Public education, meanwhile, aims to demystify radiotherapy, reduce stigma, and promote awareness of its benefits, fostering trust in cancer treatment systems [19].

Providing accurate and ethically shared information across these groups is vital to ensure safety, informed decision-making, and the effective application of radiotherapy.

This Viewpoint explores a framework for integrating LLM-based chatbots in radiotherapy education, designed to ensure accuracy, reliability, and ethical integrity. The framework incorporates 3 key components: a controlled and curated database, a robust quality control system, and continuous monitoring mechanisms [20]. The curated database includes only verified and relevant medical literature and patient data, ensuring that the chatbot's responses are accurate and up to date. The quality control system involves regular audits, database updates, and algorithms to identify and correct inaccuracies or biases. Continuous monitoring tracks the chatbot's performance, enabling real-time adjustments and improvements [20]. By addressing challenges such as misinformation and patient privacy, this framework establishes a trustworthy and effective tool in radiotherapy education. Furthermore, it sets a new standard for AI-driven tools in specialized health care education by focusing on accuracy, reliability, and ethical integrity [20,21].

Ultimately, the ethical design and use of LLM-based medical chatbots involve collective efforts from policy makers and governments to maximize benefits while mitigating risks such as data privacy breaches and algorithmic bias [22].

Current State of Medical Chatbots

Development and Implementation of LLM-Based Chatbots

The development and implementation of LLM-based chatbots in health care have marked a significant evolution from earlier chatbot technologies. Unlike traditional rule-based systems that operated within a limited scope, LLM-based chatbots such as those powered by OpenAI's ChatGPT have demonstrated the ability to understand and generate natural language with remarkable fluency and contextual awareness [23]. This advancement has led to a broad range of applications in health care, where these chatbots are being used to support patient interaction, provide medical information, assist with diagnosis, and streamline administrative tasks. A review of current LLM-based medical chatbots reveals a diversity of functionalities tailored to various health care needs [24]. One prominent example is ChatGPT itself, which has been adapted for use in several health care settings. ChatGPT can engage in detailed conversations with patients, answering questions about symptoms, explaining medical conditions, and providing guidance on the next steps in care. Its ability to understand nuanced queries and generate contextually appropriate responses has made it a valuable tool for preliminary consultations, particularly in telemedicine where immediate access to health care professionals may be limited [25]. Another notable use case is the deployment of LLM-based chatbots for chronic disease management [13]. These chatbots are designed to assist patients in managing conditions such as diabetes, hypertension, and asthma by providing personalized advice, reminding them to take medication, and offering lifestyle recommendations based on their medical history and real-time data inputs; for example, some LLM-based chatbots integrate with wearable devices to monitor patient vitals and deliver timely interventions, improving adherence to treatment plans and potentially reducing the need for emergency care [26]. In addition to patient-facing applications, LLM-based chatbots have been implemented to support health care professionals. These chatbots can function as web-based assistants, helping clinicians with tasks such as documentation, coding, and accessing up-to-date medical literature; for instance, some systems are designed to summarize patient records, extract relevant clinical information, and even suggest differential diagnoses, thereby enhancing the efficiency of clinical workflows [27]. In research settings, LLM-based chatbots are used to assist in data analysis, generate research hypotheses, and streamline the process of literature review by quickly summarizing large volumes of academic papers.

The flexibility and scalability of LLM-based chatbots have also enabled their adoption in mental health support. Chatbots such as Woebot [28] and Wysa [29] use LLMs to engage in therapeutic conversations, offering cognitive behavioral therapy techniques, mood tracking, and crisis intervention. These tools provide users with immediate access to mental health support, which is particularly valuable in regions with limited access to mental health professionals or during times when in-person therapy is not feasible. While the functionalities and use cases of LLM-based chatbots in health care are vast, the implementation of these tools is not without challenges.

Ensuring the accuracy and reliability of the information provided by these chatbots is crucial, especially in scenarios where they are used for diagnosis or treatment recommendations. Moreover, the ethical implications of using AI in patient care, particularly in terms of data privacy and informed consent, require careful consideration [30]. Despite these challenges, the ongoing development and refinement of LLM-based chatbots continue to push the boundaries of what is possible in health care, offering promising avenues for improving patient care, enhancing health care delivery, and supporting the work of health care professionals.

Strengths and Limitations

The implementation of LLM-based chatbots in health care, particularly in specialized fields such as radiotherapy, offers both significant strengths and notable limitations. As these advanced AI-driven tools become more integrated into medical practice, it is crucial to understand their advantages and challenges to optimize their use while mitigating potential risks.

Strengths

One of the primary strengths of LLM-based chatbots is their ability to process and generate natural language with a high degree of fluency and context sensitivity. This capability allows them to engage in detailed and nuanced conversations with both patients and health care professionals, making complex medical information more accessible and understandable. In radiotherapy, where patients often face intricate treatment regimens and technical jargon, an LLM-based chatbot can break down complex concepts into simpler terms, enhancing patient comprehension and engagement. This improved communication can lead to better patient adherence to treatment plans and a more informed patient population, which is essential for the success of radiotherapy [17]. Another significant advantage of LLM-based chatbots is their scalability and availability. These chatbots can provide consistent, round-the-clock support, making them particularly valuable in settings where access to health care professionals may be limited. In radiotherapy, where timely information is critical, a chatbot can offer immediate answers to patient queries, provide pretreatment education, and even guide patients through posttreatment care [31]. This constant availability helps reduce the burden on health care providers, allowing them to focus on more complex cases and personalized care. Furthermore, LLM-based chatbots can be regularly updated with the latest medical research and guidelines, ensuring that the information they provide remains current. In a field such as radiotherapy, where treatment protocols and technologies are continuously evolving, this ability to quickly incorporate new knowledge is a significant strength. It enables the chatbot to serve as a reliable resource for both patients and health care professionals, supporting evidence-based practice and reducing the likelihood of outdated or incorrect information being disseminated [32].

Limitations

Despite their strengths, LLM-based chatbots also present several limitations, particularly in specialized medical fields such as radiotherapy. One of the most pressing challenges is ensuring the accuracy and reliability of the information provided by these

chatbots [20]. While LLMs can generate text that seems plausible and coherent, they may occasionally produce incorrect or misleading information, especially when faced with highly specialized or uncommon queries. In radiotherapy, where precise details about treatment options, dosimetry, and potential side effects are crucial, even small inaccuracies can have serious consequences for patient care. Another limitation is the potential for LLM-based chatbots to oversimplify complex medical information. While simplifying language is essential for patient comprehension, there is a risk that crucial nuances may be lost, leading to misunderstandings or incomplete knowledge. In radiotherapy, where patients need to fully understand their treatment options, the risks, and the benefits, this oversimplification could impact their ability to make informed decisions [3]. Ethical considerations also pose significant challenges for the implementation of LLM-based chatbots in health care. Issues related to patient privacy, data security, and the potential for bias in the AI's responses are critical concerns. In radiotherapy, where patients are dealing with life-altering decisions and sensitive health information, ensuring that the chatbot operates within strict ethical guidelines is essential. There is also the challenge of maintaining transparency in how these chatbots operate and ensuring that patients are fully aware

that they are interacting with an AI, not a human health care provider. Moreover, the reliance on LLM-based chatbots could potentially lead to an overdependence on AI at the expense of human interaction. In specialized fields such as radiotherapy, the human touch is often crucial for providing emotional support and addressing the psychosocial aspects of care. While chatbots can provide information and support, they cannot replace the empathy and personalized care that human health care providers offer [33].

Summary

Table 1 summarizes the key advantages and challenges associated with the deployment of LLM-based chatbots in health care, particularly in the context of radiotherapy, where precision, ethical considerations, and human interaction are critical. It is seen that while LLM-based chatbots offer significant strengths, including enhanced communication, scalability, and up-to-date information, they also come with limitations that must be carefully managed. In specialized fields such as radiotherapy, where there is a need for compassion in addition to accurate information as the users may be in stressful situations, the deployment of these chatbots requires a well-thought-out framework that maximizes their benefits while addressing their inherent challenges.

Table 1. Strengths and limitations of large language model-based chatbots in specialized medical fields such as radiotherapy.

Aspect	Strengths	Limitations
Natural language processing	<ul style="list-style-type: none"> Fluent, context-aware communication Simplifies complex medical information for better patient understanding 	<ul style="list-style-type: none"> Risk of generating incorrect or misleading information Potential oversimplification of complex concepts, leading to misunderstandings
Availability and scalability	<ul style="list-style-type: none"> 24-7 availability for patient and health care provider support Reduces the burden on health care professionals by handling routine queries 	<ul style="list-style-type: none"> Overdependence on AI^a could reduce human interaction, which is vital in emotionally supportive care, particularly in radiotherapy
Up-to-date information	<ul style="list-style-type: none"> Can be regularly updated with the latest research and guidelines, ensuring that the information provided is current 	<ul style="list-style-type: none"> Challenges in maintaining the accuracy of specialized information, especially as new research emerges and medical knowledge evolves
Patient engagement	<ul style="list-style-type: none"> Enhances patient education and engagement through personalized, accessible information Supports informed decision-making in complex treatments 	<ul style="list-style-type: none"> Risk of inadequate personalization, potentially leading to generic advice that does not fully meet individual patient needs
Ethical considerations	<ul style="list-style-type: none"> Can be programmed to follow ethical guidelines, ensuring responsible dissemination of medical information 	<ul style="list-style-type: none"> Issues related to patient privacy, data security, and AI bias Potential lack of transparency in AI operations, affecting trust
Efficiency in health care	<ul style="list-style-type: none"> Improves efficiency by assisting in documentation, coding, and access to medical literature for health care providers 	<ul style="list-style-type: none"> Dependency on chatbot accuracy for clinical tasks might introduce errors if the chatbot provides incorrect or incomplete information

^aAI: artificial intelligence.

Impact on Patient Education and Health Care Delivery

Influence on Patient Education

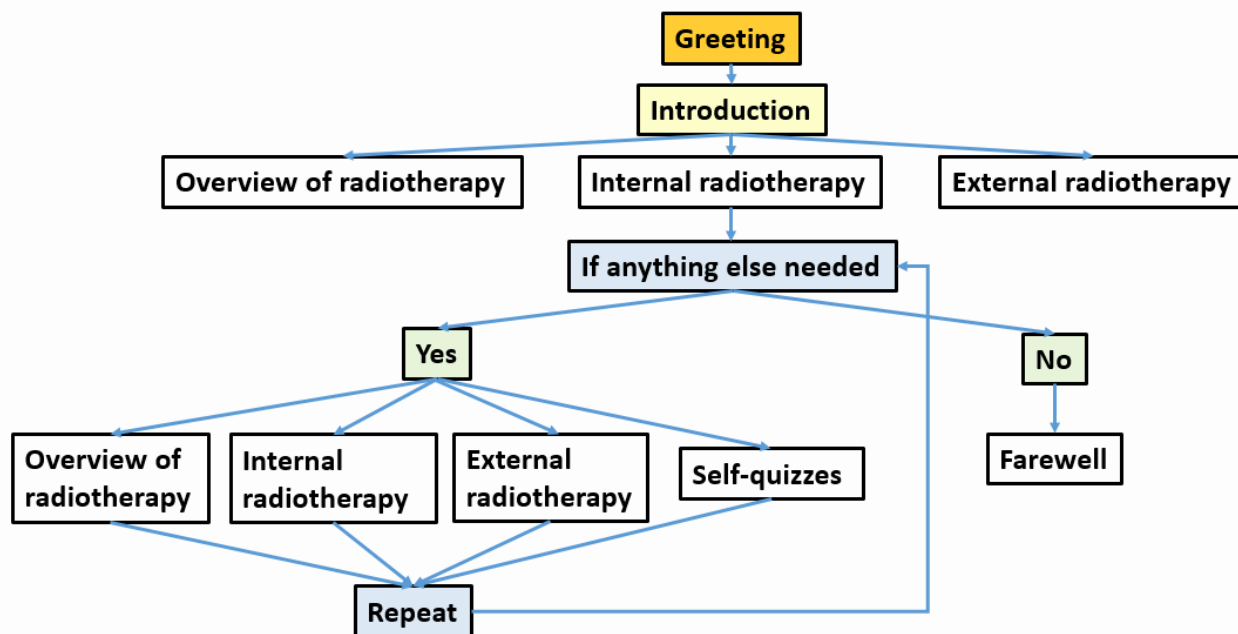
LLM-based chatbots play a significant role in enhancing patient education by making complex medical information more accessible. In specialized fields such as radiotherapy, where

patients must understand intricate treatment protocols and potential side effects, chatbots provide a valuable resource for simplifying and clarifying these concepts. By translating technical jargon into clear language, chatbots help patients grasp the fundamentals of their treatment, which can be crucial for their engagement and adherence. Moreover, chatbots can offer tailored educational content based on individual patient profiles;

for example, a chatbot might provide specific information about the type of radiotherapy a patient is receiving, potential side effects relevant to their case, and strategies for managing these side effects [31,34]. This personalized approach ensures that

patients receive relevant information that directly pertains to their situation, enhancing their understanding and preparedness for their treatment journey. Figure 2 shows a typical dialogue flowchart for a chatbot in radiotherapy education.

Figure 2. Typical dialogue flowchart for a chatbot used in radiotherapy education.



Impact on Decision-Making

In addition to enhancing patient education, LLM-based chatbots significantly influence patient decision-making. By providing timely and accurate information, these tools empower patients to make informed choices about their treatment options; for instance, a chatbot can help patients compare different radiotherapy techniques, discuss the potential benefits and risks of each option, and address any concerns they may have. This support is especially valuable in radiotherapy, where patients often face complex decisions about their care [32]. The chatbot's ability to offer evidence-based information and answer questions in real time helps reduce uncertainty and anxiety, allowing patients to engage more effectively in their treatment planning. Furthermore, by facilitating a deeper understanding of their condition and treatment options, chatbots contribute to a more collaborative decision-making process between patients and their health care providers.

Enhancement of Health Care Delivery

The implementation of LLM-based chatbots also has a notable impact on the overall delivery of health care. These tools streamline administrative tasks, such as appointment scheduling, medication reminders, and follow-up care, thus freeing up health care professionals to focus on more complex patient interactions [35]. In radiotherapy, where treatment regimens are often intensive and require careful management, chatbots can help monitor patient progress, manage treatment schedules, and provide reminders for follow-up appointments or adherence to prescribed protocols. In addition, chatbots can assist in triaging patient queries, directing them to appropriate resources or health care professionals based on the nature of their concerns [36].

This efficient handling of inquiries helps ensure that patients receive timely and relevant support, which can improve their overall experience and satisfaction with their care.

Challenges in Deploying LLM-Based Medical Chatbots

Overview

Despite the promising advancements and benefits of LLM-based medical chatbots, their deployment in health care settings is accompanied by several significant challenges. These challenges span various critical areas, including the accuracy and reliability of the information provided, privacy and data security concerns, ethical considerations, and the potential for misinformation. Addressing these issues is essential to ensure that chatbots contribute positively to health care delivery while mitigating risks.

Accuracy and Reliability

Ensuring the accuracy and reliability of medical information provided by LLM-based chatbots is a critical challenge that affects their effectiveness and safety in health care settings. The complexity of medical knowledge and the rapid evolution of treatment protocols present significant hurdles in maintaining the precision of information delivered by these AI tools. One major challenge is the inherent nature of LLMs, which are trained on vast datasets that include a wide range of information sources. While this breadth of data allows chatbots to generate responses to a variety of queries, it also means that the information they provide can be inconsistent or outdated [37]. In the context of radiotherapy, where treatment guidelines and protocols are continually updated based on the latest research, ensuring that a chatbot reflects the most current and accurate

information is crucial. Failure to do so can result in the dissemination of outdated or incorrect treatment recommendations, potentially compromising patient safety [20].

In addition, LLM-based chatbots rely on pattern recognition and probabilistic inference rather than deep, context-specific understanding [38]. This approach can sometimes lead to the generation of plausible but inaccurate answers, especially when dealing with rare or complex medical scenarios; for example, a chatbot might provide general information about radiotherapy but struggle with specifics related to individual patient conditions or less common treatment modalities [4]. This limitation highlights the need for robust validation processes to ensure that the chatbot's responses are not only accurate but also relevant to the user's specific context. Another aspect of accuracy and reliability involves the management of the chatbot's knowledge base [39]. Regular updates and reviews are necessary to keep the information current and accurate. This process requires a continuous feedback loop from health care professionals who can assess and correct any inaccuracies or gaps in the information provided by the chatbot. Implementing such a system involves collaboration between AI developers and medical experts to ensure that the chatbot's responses align with the latest evidence-based practices. Moreover, there is a challenge in balancing the chatbot's ability to generate detailed responses with the need to provide clear and concise information [40]. Overly complex or technical explanations can be difficult for patients to understand, while overly simplified answers may omit critical details. Striking the right balance is essential to provide information that is both accurate and accessible to users with varying levels of medical knowledge.

Privacy and Data Security

Privacy and data security are paramount concerns when deploying LLM-based chatbots in health care, given the sensitive nature of patient information handled by these tools. Ensuring the protection of these data against unauthorized access and breaches is essential to maintaining patient trust and complying with regulatory requirements [19]. One of the primary challenges in safeguarding patient data involves ensuring compliance with privacy regulations such as the Health Insurance Portability and Accountability Act in the United States [41] or the General Data Protection Regulation in Europe [42]. These regulations mandate stringent measures for the collection, storage, and handling of personal health information. LLM-based chatbots must be designed to adhere to these regulations, incorporating robust data encryption methods, secure storage solutions, and strict access controls to protect sensitive patient information from breaches and unauthorized access. Data transmission security is another critical aspect. When chatbots interact with patients, the data exchanged—such as medical history, treatment details, and personal identifiers—must be encrypted to prevent interception by malicious actors. Implementing secure communication protocols, such as transport layer security, helps ensure that data transmitted between the chatbot and the user are protected from eavesdropping and tampering [43]. In addition, managing user consent and data use transparency is crucial for maintaining privacy. Patients must be informed about what data are being collected, how the data will be used, and their rights regarding their information. Clear consent

mechanisms should be integrated into the chatbot's design, allowing users to provide informed consent before any data collection occurs. This transparency helps build trust and ensures that patients are aware of how their data are being handled.

Another challenge is ensuring that data anonymization and deidentification practices are effectively applied. Even when data are stored or processed for purposes such as improving chatbot performance or training new models, it is essential to anonymize the data to prevent the identification of individual patients. Proper anonymization techniques reduce the risk of sensitive information being exposed in case of a data breach. Moreover, continuous monitoring and auditing of the chatbot's security systems are necessary to identify and address potential vulnerabilities. Regular security assessments, including penetration testing and vulnerability scanning, help ensure that the chatbot's infrastructure remains resilient against emerging threats. Promptly addressing any identified weaknesses is crucial for maintaining the overall security of patient data.

Ethical Considerations

Overview

The deployment of LLM-based chatbots in health care raises several ethical dilemmas that must be carefully addressed to ensure the responsible and fair use of AI. Key ethical concerns include AI decision-making processes, transparency, and accountability, each of which plays a critical role in maintaining trust and integrity in health care interactions [44]. To address these concerns, transparency can be enhanced by developing clear documentation that explains the chatbot's decision-making processes, including data sources and logic. Bias detection and mitigation tools, such as fairness-aware algorithms, should be used to identify and correct biases in real time. Regular audits by independent third parties can ensure compliance with ethical standards.

User feedback mechanisms are essential for identifying and addressing ethical issues. Users should be able to report concerns or errors, which should be promptly reviewed and acted upon. In addition, clear guidelines for human oversight must be established, defining the roles of health care professionals in monitoring chatbot interactions. By incorporating these detailed suggestions, processes, and tools, we can better address ethical issues and promote the responsible use of AI in health care [18].

Long-Term Societal Impacts

The widespread adoption of ChatGPT and similar technologies could significantly reshape the job market, potentially leading to both job displacement and the creation of new opportunities, necessitating reskilling and upskilling programs to mitigate negative impacts. Moreover, AI technologies might exacerbate economic inequalities if their benefits are not equitably distributed, highlighting the need for policies that promote inclusive growth. The integration of AI into daily life could alter social interactions, reducing human-to-human contact and impacting social skills and relationships, thus requiring a balance between AI use and genuine human connection. These changes underscore the importance of robust ethical governance frameworks involving diverse stakeholders to ensure that AI development aligns with societal values and ethical principles,

addressing emerging challenges and opportunities through continuous dialogue and adaptive regulation.

AI Decision-Making

One of the central ethical issues is the decision-making capability of AI systems. Unlike human health care providers who can apply clinical judgment and empathy, LLM-based chatbots operate based on algorithms and data patterns [45]. This raises questions about the extent to which these systems can make ethical decisions, especially in complex or nuanced medical scenarios; for instance, while chatbots can provide general information and support, they may lack the ability to fully understand the context of a patient's situation or the ethical implications of certain recommendations [46]. Ensuring that AI-driven tools align with ethical guidelines and medical standards is essential to avoid potentially harmful outcomes.

Transparency

Transparency in AI decision-making is another critical ethical consideration. Patients and health care professionals need to understand how chatbots generate their responses and make recommendations. This includes clarity about the data sources and algorithms that drive the chatbot's behavior. Without transparency, there is a risk that users might overestimate the chatbot's capabilities or misinterpret its advice. Providing clear information about the chatbot's operational mechanisms helps build trust and allows users to make informed decisions about how to use the tool effectively [47]. Furthermore, transparency extends to the disclosure of limitations and potential biases inherent in the AI system. Chatbots are trained on datasets that may reflect existing biases or incomplete information, which can affect the fairness and objectivity of their responses [48]. Openly communicating these limitations and actively working to mitigate bias are crucial for maintaining ethical standards and ensuring that the chatbot serves all users equitably.

Accountability

Accountability is a significant ethical issue concerning the use of LLM-based chatbots. Determining who is responsible for the chatbot's recommendations and decisions is essential, particularly in cases where incorrect or harmful information is provided. Clear lines of accountability must be established, involving both the developers who create and maintain the chatbot and the health care providers who deploy it. This includes ensuring that there are mechanisms for addressing errors or issues that arise from the chatbot's interactions with users [49]. In addition, it is important to have protocols in place for managing situations where the chatbot's advice might lead to adverse outcomes [50]. This includes providing a way for users to report issues and seek redress, as well as ensuring that there are processes for continuous improvement based on feedback and incident analysis. Effective oversight and governance structures help ensure that ethical standards are upheld and that the chatbot contributes positively to patient care.

Potential for Misinformation

Overview

The potential for misinformation is a critical concern in the deployment of LLM-based medical chatbots. Given their reliance on vast datasets and sophisticated algorithms, these chatbots can inadvertently spread incorrect or misleading information, which poses significant risks to patient health and safety. As deep learning pioneer Geoffrey Hinton has noted, neural networks can share what they learn instantly [22]; therefore, any erroneous messages can spread instantaneously, making unwanted wide impact.

Sources of Misinformation

One key source of misinformation is the quality and accuracy of the training data used to develop the chatbot. LLMs are trained on diverse datasets that include information from various sources, some of which may be outdated, biased, or inaccurate. If a chatbot's training data contain erroneous or misleading content, there is a risk that the chatbot will replicate and disseminate these inaccuracies in its responses. In fields such as radiotherapy, where precise and current information is crucial, the presence of outdated or incorrect data can lead to harmful consequences for patients relying on the chatbot for guidance [51]. Another potential source of misinformation is the chatbot's ability to generalize information [52]. While LLM-based chatbots can handle a wide range of queries, they may not always be adept at distinguishing between general knowledge and specific, context-sensitive details. This limitation can lead to the generation of responses that are technically correct but fail to address the nuances of individual patient situations [53]; for example, a chatbot might provide general advice on radiation safety but miss specific recommendations tailored to a patient's unique treatment plan.

Risks and Consequences

The spread of incorrect or misleading information by medical chatbots can have serious repercussions for patient care. Patients may make health-related decisions based on inaccurate information, which can result in inappropriate or ineffective treatments. In radiotherapy, where treatment decisions are complex and require careful consideration of numerous factors, relying on incorrect information can lead to suboptimal care or adverse outcomes. Moreover, misinformation can erode trust in health care technology. If patients or health care providers encounter inaccuracies or inconsistencies in the information provided by a chatbot, their confidence in the tool's reliability and usefulness may be undermined. This loss of trust can diminish the chatbot's effectiveness and impact its acceptance and integration into health care practices.

Mitigation Strategies

To mitigate the risks associated with misinformation, several strategies can be used [54]. Regular updates and maintenance of the chatbot's knowledge base are essential to ensure that it reflects the most current and accurate information. Collaboration with medical experts and continuous validation of the chatbot's responses help identify and correct inaccuracies. Implementing mechanisms for user feedback and reporting can also provide valuable insights into potential issues and facilitate prompt

resolution. Furthermore, incorporating a system of checks and balances, such as providing disclaimers that emphasize the chatbot's limitations and the need for professional medical consultation, can help manage user expectations. Ensuring that the chatbot directs users to seek advice from qualified health care professionals when necessary can prevent reliance on potentially flawed or incomplete information [55].

Proposed Framework for a Resilient Medical Chatbot in Radiotherapy Education

Database Management and Control

Effective database management is crucial for developing a resilient medical chatbot in radiotherapy education. The integrity of the chatbot's responses depends heavily on the accuracy and timeliness of the information it accesses. To achieve this, a robust strategy for building and maintaining a controlled database is essential [56]. The first step in this strategy involves curating a comprehensive and authoritative database from verified medical sources, such as peer-reviewed journals, clinical guidelines, and expert consensus statements. This curated database should be dynamic, with automated systems in place for continuous updates. The proposed framework incorporates an AI-driven mechanism for real-time monitoring of new publications and guidelines, ensuring that the database remains current and reducing the risk of disseminating outdated or inaccurate information [57]. To implement this, automated updates could leverage web scraping tools or application programming interface integrations to collect and validate data against trusted sources. The database should use meta-tagging and hierarchical organization for efficient data management and retrieval. Meta-tagging can be achieved using NLP algorithms to assign contextual keywords, improving the chatbot's ability to interpret and respond to user queries accurately. A hierarchical data structure, such as a tree- or graph-based model, can prioritize data by relevance and reliability, ensuring that the chatbot provides the most appropriate responses [58,59]. In addition, periodic audits of the database should be conducted to verify its accuracy and adherence to updated guidelines.

Maintaining control over the database is equally important. This involves setting strict protocols for data entry and modification, with access restricted to authorized personnel who are well versed in radiotherapy and medical education. Regular audits of the database content should be conducted to identify and rectify any inconsistencies or errors [57]. In addition, using version control systems can help track changes, allowing for the restoration of previous database states if needed, which further enhances the reliability of the chatbot. To ensure that the database remains resilient against emerging challenges such as misinformation or biased data, a layered review process should be integrated. This involves cross-referencing new data entries with multiple sources and using machine learning algorithms to detect and flag anomalies or conflicting information [60]. By implementing these strategies, the database management and control framework will serve as the foundation for a resilient medical chatbot capable of providing accurate, reliable, and up-to-date information in radiotherapy education.

Quality Control System

A robust quality control system is vital for ensuring the reliability and effectiveness of a medical chatbot in radiotherapy education. The system must be designed to maintain the highest standards of accuracy, relevance, and trustworthiness in the chatbot's content, which requires regular reviews, timely updates, and stringent validation processes [61]. The cornerstone of this quality control system is the implementation of a multitiered review process. This begins with the periodic assessment of the chatbot's content by a panel of experts in radiotherapy and medical education. These experts should evaluate the chatbot's responses for accuracy, clarity, and consistency with current clinical practices. Regularly scheduled reviews ensure that the content remains aligned with the latest advancements in radiotherapy and adheres to evolving educational standards [4].

In addition to expert reviews, the quality control system should include automated checks and balances. Machine learning algorithms can be deployed to continuously monitor the chatbot's interactions, identifying patterns of errors or discrepancies in the responses [62]. These algorithms can flag potential issues for further human review, ensuring that errors are caught and corrected promptly. Automated processes should also include regular updates to the chatbot's knowledge base, triggered by new research findings, clinical guidelines, or changes in medical protocols. Another critical aspect of the quality control system is the validation of the chatbot's content before it goes live. This can be achieved through rigorous testing with simulated user interactions, covering a wide range of scenarios that the chatbot is likely to encounter. Feedback loops from these tests should be analyzed to refine the chatbot's algorithms and content delivery mechanisms, ensuring that it provides accurate and contextually appropriate responses [63]. Finally, the quality control system must be adaptable, with mechanisms in place for continuous improvement. This includes incorporating user feedback to identify areas where the chatbot's performance may be lacking and making necessary adjustments. Moreover, the system should be capable of responding to unforeseen challenges, such as the propagation of misinformation, and adapting to emerging trends in radiotherapy by swiftly updating the chatbot's content and protocols [17].

Monitoring and Feedback Loop

Establishing an effective monitoring and feedback loop is essential for the continuous improvement and reliability of a medical chatbot in radiotherapy education. This system ensures that the chatbot consistently meets user needs, adapts to new information, and addresses any issues that arise during its interactions. The foundation of this system is a comprehensive monitoring mechanism designed to track all chatbot interactions in real time. By recording user queries, responses, and outcomes, this monitoring system can identify patterns in user behavior and detect potential issues, such as inaccurate answers, misinterpretations, or gaps in the chatbot's knowledge base [64]. Advanced analytics tools can be integrated to automatically flag problematic interactions for further review by the development team, enabling swift corrective actions. In addition to tracking interactions, the monitoring system should

incorporate a robust feedback loop that allows users to provide direct input on the chatbot's performance [65]. This feedback can be collected through postinteraction surveys, ratings, or optional comment fields, giving users the opportunity to highlight areas where the chatbot excels or falls short. Aggregating and analyzing this feedback offers valuable insights into the chatbot's strengths and weaknesses, guiding future updates and enhancements.

To maximize the effectiveness of the feedback loop, it is crucial to establish a process for prioritizing and addressing the issues identified. A triage system can be implemented to categorize user feedback based on severity and frequency, ensuring that the most critical issues are addressed promptly [66]; for example, if multiple users report the same error or misunderstanding, this issue would be flagged as a high priority for immediate resolution. Less critical feedback, such as suggestions for improved phrasing or additional features, can be scheduled for consideration in future updates. Moreover, the feedback loop should be designed to foster continuous learning and adaptation. Regularly scheduled reviews of the chatbot's performance, informed by user feedback and monitoring data, should be conducted to identify trends and areas for improvement. This iterative process allows the chatbot to evolve over time, enhancing its ability to provide accurate, reliable, and contextually relevant information in radiotherapy education.

Ethical and Legal Safeguards

Implementing robust ethical guidelines and legal safeguards is paramount in developing a medical chatbot for radiotherapy education. These measures are critical to addressing concerns related to privacy, transparency, and accountability, ensuring that the chatbot operates within the highest ethical and legal standards. To begin with, the chatbot must be designed with stringent privacy protections. Given the sensitive nature of medical information, the chatbot should comply with all relevant data protection laws, such as the General Data Protection Regulation in Europe or the Health Insurance Portability and Accountability Act in the United States [19]. This involves

implementing secure data encryption methods, anonymizing user interactions, and ensuring that any personal or health-related data collected during interactions are stored securely and used only for their intended purpose [67]. Access to these data should be strictly controlled, with clear protocols for who can view, modify, or delete the data, ensuring that user privacy is respected at all times. Transparency is another critical aspect of ethical chatbot design. Users must be fully informed about how the chatbot functions, including the sources of its information, the limitations of its advice, and the nature of its data collection practices. This can be achieved by providing clear and accessible information within the chatbot interface, such as disclaimers before interactions, a detailed frequently asked questions section, or links to privacy policies. In addition, the chatbot should be designed to clearly distinguish between general information and personalized advice, ensuring that users understand the context and limitations of the information they receive [68].

Accountability mechanisms are also essential to uphold ethical standards. The development and deployment of the chatbot should include a clear governance structure, where responsibilities for content accuracy, data management, and user interactions are well defined. Regular audits of the chatbot's performance and adherence to ethical guidelines should be conducted, with the results made available to relevant stakeholders [69]. Furthermore, there should be a clear process for users to report any ethical concerns or grievances, and these issues should be addressed promptly and transparently. To further enhance ethical and legal safeguards, it is important to establish an oversight committee consisting of experts in ethics, law, and medical education [70]. This committee would be responsible for reviewing the chatbot's operations, ensuring that it remains aligned with ethical principles and legal requirements. The committee should also be tasked with evaluating the impact of the chatbot on users, particularly in terms of potential biases, misinformation, or unintended consequences, and recommending corrective actions as needed. The proposed framework is summarized in [Table 2](#).

Table 2. Summary of the proposed framework for a resilient medical chatbot in radiotherapy education. The key components, their essential elements, and their respective purposes in ensuring the chatbot's accuracy, reliability, and ethical operation are outlined.

Component	Key elements	Purpose
Database management and control	<ul style="list-style-type: none"> Curated and dynamic database Automated updates Hierarchical data structure 	Ensures that the chatbot has access to accurate, up-to-date, and relevant information
Quality control system	<ul style="list-style-type: none"> Multitiered expert review Automated checks Content validation 	Maintains high standards of accuracy, relevance, and reliability in chatbot responses
Monitoring and feedback loop	<ul style="list-style-type: none"> Real-time interaction tracking User feedback integration Continuous improvement 	Tracks chatbot performance, identifies issues, and incorporates user feedback for ongoing enhancement
Ethical and legal safeguards	<ul style="list-style-type: none"> Privacy protections Transparency measures Accountability mechanisms 	Addresses concerns related to user privacy, ensures transparency, and upholds accountability

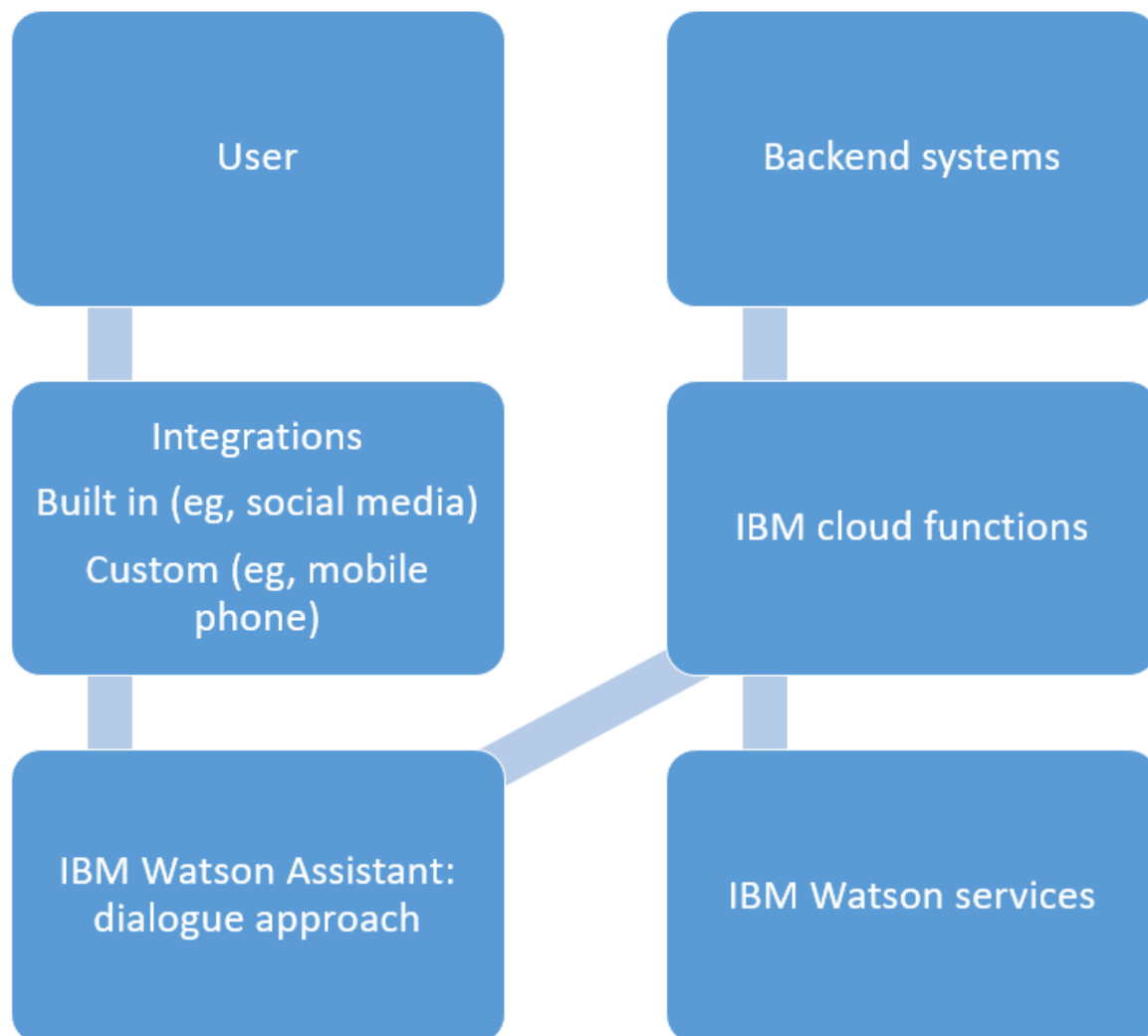
Case Study: Application of the Framework in Radiotherapy Education

Outlining the Process of Developing and Testing the Chatbot Using the Proposed Framework

The development of the radiotherapy education chatbot using the proposed framework commenced with a comprehensive needs assessment. This step involved engaging key stakeholders, including radiotherapy educators, clinical oncologists, medical physicists, and students, to identify the specific educational challenges and gaps that the chatbot aimed to address [32]. The focus was on creating a tool that could effectively supplement traditional education by providing accurate, accessible, and contextually relevant information on complex radiotherapy topics. Once the needs were clearly defined, the next phase involved designing the chatbot's conversational architecture. The framework emphasized the integration of AI-driven NLP

capabilities with a structured knowledge base specific to radiotherapy [71]. This phase required meticulous curation and validation of content, ensuring that the information was not only accurate but also aligned with the latest guidelines and research in the field. A collaborative approach was adopted, involving experts from various disciplines to ensure that the chatbot's responses would be both scientifically robust and pedagogically sound [72]. The chatbot was then built using a hybrid approach that combined rule-based algorithms with machine learning models [73]. The rule-based components ensured that the chatbot could handle critical educational scenarios with precision, while the machine learning aspects allowed for more dynamic and flexible interactions. The integration with the IBM Watson Assistant cloud platform facilitated the deployment of advanced NLP algorithms, enabling the chatbot to understand and respond to user queries effectively [32]. Figure 3 shows the architectural diagram of a chatbot developed using the IBM Watson Assistant platform [74].

Figure 3. Architectural diagram showing how the chatbot is connected to the IBM Watson Assistant cloud platform.



Testing the chatbot was an iterative process. Initial prototypes were subjected to rigorous testing in simulated environments, where various scenarios were presented to evaluate the chatbot's

performance in real time [75]. The testing focused on several key metrics: the accuracy of information, user engagement, response time, and the ability to handle unexpected or complex

queries. Feedback from educators and students was crucial during this phase, providing insights into the chatbot's usability and effectiveness in an educational setting. To ensure the reliability and safety of the chatbot, the testing also included a comprehensive evaluation of ethical considerations. The framework's built-in mechanisms for privacy protection, data security, and bias mitigation were scrutinized to prevent potential harm. The chatbot was tested for its adherence to ethical standards through a Delphi study involving international experts, ensuring that it met global expectations for AI-driven educational tools [76].

The final phase involved pilot-testing the chatbot in real educational settings, specifically within radiotherapy courses. These studies aimed to evaluate both the chatbot's technical

performance and its educational impact. Observations from the pilot phase revealed notable improvements in student comprehension and engagement; for example, the chatbot, called RT Bot, was deployed on a website accessible by invitation and introduced during workshops and conferences to test its functionality. Participants were invited to complete a survey evaluating the chatbot's functionality and content. This survey, conducted with 60 participants, aimed to gather statistical insights and feedback for improving the chatbot in the future. The results, summarized in Table 3 [32], highlight varying levels of user satisfaction with RT Bot during its testing phase. The assessment used a scale ranging from 1 (lowest satisfaction) to 5 (highest satisfaction). The data indicate that a substantial 70% (42/60) of the respondents rated the content's helpfulness, understandability, and reading duration as above average.

Table 3. Users' satisfaction ratings regarding the contents of RT Bot (N=60) [32].

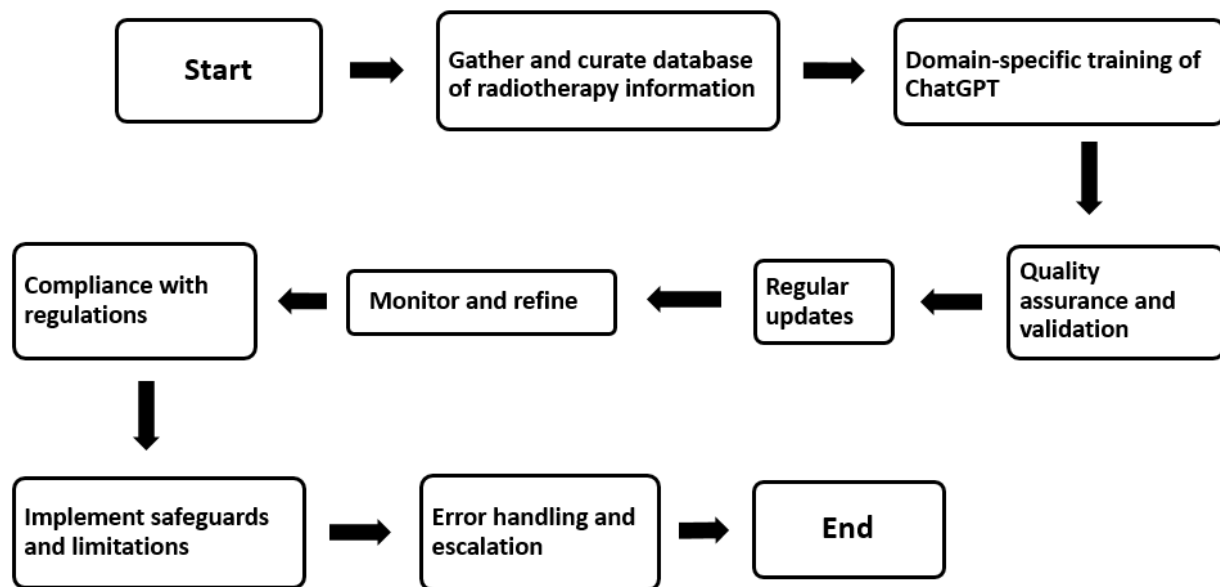
Degree of satisfaction (ranging from 1=lowest to 5=highest)	Helpfulness of content, n (%)	Understandability of content, n (%)	Reading duration, n (%)
1	3 (5)	0 (0)	0 (0)
2	0 (0)	9 (15)	3 (5)
3	15 (25)	9 (15)	15 (25)
4	24 (40)	24 (40)	18 (30)
5	18 (30)	18 (30)	24 (40)

Specific feedback highlighted areas for improvement, such as the need for more intuitive methods to restart conversations and tailored hints for users with diverse backgrounds when answering radiation safety questions. These insights enabled developers to refine the chatbot's adaptability and responsiveness. Data collected from the pilot tests demonstrated a measurable positive effect on learning outcomes, as students achieved higher quiz scores and expressed increased confidence in applying radiation safety principles. The feedback loop established during this phase ensured that the chatbot evolved into a robust, user-centered educational tool, ready to support radiotherapy training effectively [34,77,78].

Figure 4 illustrates the workflow for creating and maintaining an LLM-based medical chatbot powered by ChatGPT, tailored

for radiotherapy. The chatbot relies on a curated database of verified information and undergoes domain-specific training to understand radiotherapy nuances. However, several limitations exist. The quality of training data and the effectiveness of cross-checking mechanisms are crucial; any errors or biases can lead to inaccurate responses. Regular updates are necessary to keep the database current, but this requires significant resources. User feedback helps refine the system, although the feedback can be subjective and challenging to manage. Error-handling mechanisms allow for query escalation, but not all issues may be resolved, potentially affecting user trust. Despite these challenges, the iterative process aims to provide accurate and up-to-date information within the chatbot's expertise.

Figure 4. Flowchart providing a sequential representation of the steps involved in creating, maintaining, and improving a medical chatbot based on the large language model ChatGPT, with a focus on radiotherapy, ensuring accuracy, compliance, and continuous refinement.



This process is crucial because it ensures the reliability, accuracy, and ethical integrity of the medical chatbot in radiotherapy. By curating a precise and verified database, fine-tuning the AI model, and implementing rigorous validation processes, the system can provide trustworthy information. The continuous refinement loop, fueled by user feedback and updates, ensures that the chatbot stays current with evolving practices and advancements in radiotherapy. This reliability and adherence to medical standards instills confidence in users, guiding them with accurate information while emphasizing the chatbot's limitations and the necessity of seeking professional medical advice. Ultimately, this meticulous process not only educates and assists users but also upholds the ethical responsibility of providing reliable health care information through AI-driven technologies.

Performance Evaluation

Evaluating the chatbot's performance was a critical step to ensure that it met the high standards required for educational tools in radiotherapy. The evaluation focused on 3 main areas: accuracy, user satisfaction, and educational impact.

Accuracy

The accuracy of the chatbot was assessed through a multitiered validation process. First, a set of standardized queries covering a wide range of radiotherapy topics was developed. These queries were designed to test the chatbot's ability to provide correct and precise information. Subject matter experts in radiotherapy independently evaluated the chatbot's responses to these queries, comparing them against established medical literature and clinical guidelines [79]. The chatbot's performance was quantified using accuracy metrics such as precision and recall, ensuring that it consistently delivered reliable information [80]. In addition, the chatbot's ability to update and incorporate the latest research findings was tested, emphasizing the need for dynamic content management to maintain long-term accuracy [81].

User Satisfaction

User satisfaction was evaluated through extensive user testing, involving a diverse group of end users, including students, educators, and health care professionals. Surveys and feedback forms were used to gather qualitative and quantitative data on users' experiences with the chatbot. Key aspects such as ease of use, response time, clarity of information, and overall user experience were measured. The feedback loop was integral to refining the chatbot, addressing any issues related to user interaction, and ensuring that the chatbot met the expectations of its target audience. In addition to surveys, usability testing sessions were conducted where users interacted with the chatbot in controlled environments, providing real-time feedback on their experiences. This iterative process helped in identifying and rectifying any usability challenges, enhancing the overall user satisfaction [79].

Educational Impact

The educational impact of the chatbot was evaluated by analyzing its effectiveness as a learning tool in radiotherapy education. This involved conducting controlled studies where the performance of students who used the chatbot as a supplementary learning resource was compared to that of students who did not. Various educational metrics, such as knowledge retention, comprehension, and application of concepts, were measured through pre- and postintervention assessments [82]. The chatbot's impact on these metrics was statistically analyzed to determine its effectiveness in enhancing learning outcomes [77]. Furthermore, longitudinal studies were conducted to assess the sustained educational benefits of the chatbot, ensuring that its use had a lasting positive impact on students' understanding of radiotherapy concepts [83]. In addition to these formal evaluations, the chatbot's performance was assessed in real-world educational settings through pilot studies. These studies provided insights into how the chatbot influenced classroom dynamics, student engagement, and the

overall learning environment [84-86]. Feedback from these settings was crucial in understanding the practical implications of integrating the chatbot into existing curricula and further validated its educational impact.

Lessons Learned and Future Directions

As the radiotherapy education chatbot continues to evolve, there are numerous opportunities for innovation that could significantly enhance its capabilities. Future research and technological advancements hold the potential to expand the chatbot's functionality, improve user experiences, and further integrate it into the educational landscape.

Advanced NLP and Understanding

One of the most promising areas for innovation lies in the continued development of advanced NLP algorithms. Future iterations of the chatbot could benefit from more sophisticated NLP techniques, such as transformer-based models and deep learning architectures, which can better understand and generate human-like responses [87]. These advancements would enable the chatbot to handle more complex queries, recognize nuanced language patterns, and provide more contextually accurate answers. Enhanced NLP capabilities could also improve the chatbot's ability to engage in multiturn conversations, allowing for deeper and more meaningful interactions with users.

Personalized Learning Experiences

Another key area for innovation is the development of personalized learning experiences. Leveraging data-driven insights, the chatbot could be equipped with adaptive learning algorithms that tailor content and responses to individual users' needs, preferences, and learning paces [88]. By analyzing user interactions, the chatbot could identify areas where a student might be struggling and provide targeted educational resources or alternative explanations to enhance understanding. This personalized approach would make the chatbot a more effective tool for diverse learning styles and could significantly improve educational outcomes [89].

Integration With Virtual and Augmented Reality

The integration of the chatbot with virtual and augmented reality (VR/AR) technologies presents exciting possibilities for creating immersive learning environments. By combining the chatbot's informational capabilities with VR/AR platforms, users could engage in interactive simulations of radiotherapy procedures, enhancing their practical understanding of complex concepts [90,91]; for example, students could use the chatbot to guide them through virtual radiotherapy sessions, where they can visualize and manipulate different treatment parameters in a 3D space. This fusion of AI-driven education with immersive technologies could revolutionize how radiotherapy is taught and learned.

Enhanced Multilingual and Cross-Cultural Capabilities

As radiotherapy education becomes increasingly globalized, there is a growing need for the chatbot to support multilingual and cross-cultural communication. Future developments could focus on expanding the chatbot's language capabilities, enabling it to provide accurate and contextually relevant information in

multiple languages. This would not only make the chatbot more accessible to non-English-speaking users but also allow it to adapt its responses to different cultural contexts, ensuring that the information is both accurate and culturally sensitive [92]. Such advancements would make the chatbot a valuable educational tool in diverse global settings.

Ethical AI and Bias Mitigation

As AI continues to advance, addressing ethical concerns and mitigating biases in the chatbot's responses will be paramount. Future research could explore more robust methods for ensuring that the chatbot's algorithms remain unbiased and that its responses adhere to ethical guidelines [93]. This might include developing advanced algorithms for detecting and correcting biases in real time, as well as enhancing the transparency and explainability of the chatbot's decision-making processes. By prioritizing ethical AI, the chatbot could set new standards for responsible AI use in education, fostering trust and reliability among users [20].

Integration With Broader Educational Ecosystems

Finally, future innovations could focus on integrating the chatbot with broader educational ecosystems, including learning management systems and other digital educational tools [94,95]. By doing so, the chatbot could become a seamless part of the educational workflow, providing continuous support to students and educators across various platforms. This integration would enable more comprehensive data collection and analysis, allowing educators to monitor student progress in real time and make data-informed decisions to enhance teaching strategies. The chatbot could also facilitate collaborative learning by connecting users with peers, mentors, or experts, creating a more interactive and supportive educational community.

Adaptability to Other Medical Fields

The framework developed for the radiotherapy education chatbot demonstrates significant potential for adaptation to other medical fields. By leveraging advanced NLP algorithms and personalized learning experiences, the chatbot can be tailored to provide specialized educational content across various medical disciplines, such as cardiology, neurology, and oncology. The integration of VR/AR technologies can further enhance learning by offering immersive simulations of medical procedures relevant to each field. Furthermore, the chatbot's multilingual and cross-cultural capabilities ensure that it can deliver accurate and contextually appropriate information to a diverse global audience. By addressing ethical AI considerations and integrating with broader educational ecosystems [18], the framework can support continuous learning and professional development for health care professionals worldwide, ultimately improving patient care and outcomes.

Conclusions

This Viewpoint paper has explored the development of a resilient framework for a medical chatbot tailored to radiotherapy education, addressing key aspects such as accuracy, reliability, privacy, ethics, and the potential for innovation. The framework emphasizes the importance of maintaining up-to-date and accurate information, ensuring user trust through robust privacy measures, and fostering an ethical approach to AI in

health care. Through performance evaluation, the chatbot demonstrated its capability to enhance learning outcomes and support health care professionals in their continuous education. Challenges such as bias, user engagement, and integration into existing systems were identified, with strategies proposed to overcome these obstacles.

Looking ahead, the future of LLM-based medical chatbots holds significant promise for radiotherapy education and health care as a whole. These technologies have the potential to revolutionize how complex medical knowledge is disseminated, making education more accessible, personalized, and interactive.

By continuing to advance in areas such as NLP, personalized learning, and integration with immersive technologies, LLM-based chatbots can become indispensable tools in both educational and clinical settings. As these chatbots evolve, they will likely play a crucial role in shaping the future of health care, improving patient outcomes, and supporting the ongoing education of health care professionals worldwide. The framework presented in this paper serves as a foundational guide for the responsible and effective implementation of these powerful tools, ensuring that they contribute positively to the field of radiotherapy and beyond.

Acknowledgments

This work was supported by the Planning and Dissemination Grant–Institute Community Support, Canadian Institutes of Health Research, Canada (CHIR PCS–191021). The authors acknowledge the use of ChatGPT, accessed on January 1, 2025, to assist in polishing the English in the paper. After using this tool, the authors reviewed and edited the content as necessary and assume full responsibility for the paper's content.

Authors' Contributions

JCLC and KL were responsible for conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, and reviewing and editing the manuscript. JCLC was responsible for visualization and writing the original draft.

Conflicts of Interest

None declared.

References

1. Siddique S, Chow JC. Artificial intelligence in radiotherapy. *Rep Pract Oncol Radiother*. Jul 2020;25(4):656-666. [FREE Full text] [doi: [10.1016/j.rpor.2020.03.015](https://doi.org/10.1016/j.rpor.2020.03.015)] [Medline: [32617080](https://pubmed.ncbi.nlm.nih.gov/32617080/)]
2. Chow JC. Internet-based computer technology on radiotherapy. *Rep Pract Oncol Radiother*. Nov 2017;22(6):455-462. [FREE Full text] [doi: [10.1016/j.rpor.2017.08.005](https://doi.org/10.1016/j.rpor.2017.08.005)] [Medline: [28932174](https://pubmed.ncbi.nlm.nih.gov/28932174/)]
3. Siddique S, Chow JC. Machine learning in healthcare communication. *Encyclopedia*. Feb 14, 2021;1(1):220-239. [doi: [10.3390/encyclopedia1010021](https://doi.org/10.3390/encyclopedia1010021)]
4. Chow JCL, Wong V, Li K. Generative pre-trained transformer-empowered healthcare conversations: current trends, challenges, and future directions in large language model-enabled medical chatbots. *BioMedInformatics*. Mar 14, 2024;4(1):837-852. [doi: [10.3390/biomedinformatics4010047](https://doi.org/10.3390/biomedinformatics4010047)]
5. Thirunavukarasu AJ, Ting DS, Elangovan K, Gutierrez L, Tan TF, Ting DS. Large language models in medicine. *Nat Med*. Aug 17, 2023;29(8):1930-1940. [doi: [10.1038/s41591-023-02448-8](https://doi.org/10.1038/s41591-023-02448-8)] [Medline: [37460753](https://pubmed.ncbi.nlm.nih.gov/37460753/)]
6. Chang Y, Wang X, Wang J, Wu Y, Yang L, Zhu K, et al. A survey on evaluation of large language models. *ACM Trans Intell Syst Technol*. Mar 29, 2024;15(3):1-45. [doi: [10.1145/3641289](https://doi.org/10.1145/3641289)]
7. Tirumala K, Markosyan A, Zettlemoyer L, Aghajanyan A. Memorization without overfitting: analyzing the training dynamics of large language models. In: *Proceedings of the 2022 Conference on Advances in Neural Information Processing Systems*. 2022. Presented at: NeurIPS '22; November 28-December 9, 2022:35; New Orleans, LA. URL: https://proceedings.neurips.cc/paper_files/paper/2022/hash/fa0509f4dab6807e2cb465715bf2d249-Abstract-Conference.html
8. Dam SK, Hong CS, Qiao Y, Zhang C. A complete survey on LLM-based AI chatbots. *arXiv*. Preprint posted online June 17, 2022. [FREE Full text]
9. Yang R, Tan TF, Lu W, Thirunavukarasu AJ, Ting DS, Liu N. Large language models in health care: development, applications, and challenges. *Health Care Sci*. Aug 24, 2023;2(4):255-263. [FREE Full text] [doi: [10.1002/hcs2.61](https://doi.org/10.1002/hcs2.61)] [Medline: [38939520](https://pubmed.ncbi.nlm.nih.gov/38939520/)]
10. Chakraborty C, Bhattacharya M, Lee SS. Need an AI-enabled, next-generation, advanced ChatGPT or large language models (LLMs) for error-free and accurate medical information. *Ann Biomed Eng*. Feb 27, 2024;52(2):134-135. [doi: [10.1007/s10439-023-03297-9](https://doi.org/10.1007/s10439-023-03297-9)] [Medline: [37368124](https://pubmed.ncbi.nlm.nih.gov/37368124/)]
11. Lai T, Shi Y, Du Z, Wu J, Fu K, Dou Y, et al. Supporting the demand on mental health services with AI-based conversational large language models (LLMs). *BioMedInformatics*. Dec 22, 2023;4(1):8-33. [doi: [10.3390/biomedinformatics4010002](https://doi.org/10.3390/biomedinformatics4010002)]
12. Pool J, Indulska M, Sadiq S. Large language models and generative AI in telehealth: a responsible use lens. *J Am Med Inform Assoc*. Sep 01, 2024;31(9):2125-2136. [doi: [10.1093/jamia/ocae035](https://doi.org/10.1093/jamia/ocae035)] [Medline: [38441296](https://pubmed.ncbi.nlm.nih.gov/38441296/)]

13. Montagna S, Ferretti S, Klopfenstein LC, Florio A, Pengo M. Data decentralisation of LLM-based chatbot systems in chronic disease self-management. In: Proceedings of the 2023 ACM Conference on Information Technology for Social Good. 2023. Presented at: GoodIT '23; September 6-8, 2023:205-212; Lisbon, Portugal. URL: <https://dl.acm.org/doi/10.1145/3582515.3609536> [doi: [10.1145/3582515.3609536](https://doi.org/10.1145/3582515.3609536)]
14. Bengio Y, Hinton G, Yao A, Song D, Abbeel P, Darrell T, et al. Managing extreme AI risks amid rapid progress. *Science*. May 24, 2024;384(6698):842-845. [doi: [10.1126/science.adn0117](https://doi.org/10.1126/science.adn0117)] [Medline: [38768279](https://pubmed.ncbi.nlm.nih.gov/38768279/)]
15. Chandra RA, Keane FK, Voncken FE, Thomas CR. Contemporary radiotherapy: present and future. *Lancet*. Jul 10, 2021;398(10295):171-184. [doi: [10.1016/S0140-6736\(21\)00233-6](https://doi.org/10.1016/S0140-6736(21)00233-6)] [Medline: [34166607](https://pubmed.ncbi.nlm.nih.gov/34166607/)]
16. Shafiq J, Barton M, Noble D, Lemer C, Donaldson LJ. An international review of patient safety measures in radiotherapy practice. *Radiother Oncol*. Jul 2009;92(1):15-21. [doi: [10.1016/j.radonc.2009.03.007](https://doi.org/10.1016/j.radonc.2009.03.007)] [Medline: [19395105](https://pubmed.ncbi.nlm.nih.gov/19395105/)]
17. Xu L, Sanders L, Li K, Chow JC. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. *JMIR Cancer*. Nov 29, 2021;7(4):e27850. [FREE Full text] [doi: [10.2196/27850](https://doi.org/10.2196/27850)] [Medline: [34847056](https://pubmed.ncbi.nlm.nih.gov/34847056/)]
18. Chow JC, Li K. Ethical considerations in human-centered AI: advancing oncology chatbots through large language models. *JMIR Bioinform Biotechnol*. Nov 06, 2024;5:e64406. [FREE Full text] [doi: [10.2196/64406](https://doi.org/10.2196/64406)] [Medline: [39321336](https://pubmed.ncbi.nlm.nih.gov/39321336/)]
19. Li J. Security implications of AI chatbots in health care. *J Med Internet Res*. Nov 28, 2023;25:e47551. [FREE Full text] [doi: [10.2196/47551](https://doi.org/10.2196/47551)] [Medline: [38015597](https://pubmed.ncbi.nlm.nih.gov/38015597/)]
20. Chow JC, Sanders L, Li K. Impact of ChatGPT on medical chatbots as a disruptive technology. *Front Artif Intell*. Apr 5, 2023;6:1166014. [FREE Full text] [doi: [10.3389/frai.2023.1166014](https://doi.org/10.3389/frai.2023.1166014)] [Medline: [37091303](https://pubmed.ncbi.nlm.nih.gov/37091303/)]
21. Abd-Alrazaq A, AlSaad R, Alhuwail D, Ahmed A, Healy PM, Latifi S, et al. Large language models in medical education: opportunities, challenges, and future directions. *JMIR Med Educ*. Jun 01, 2023;9:e48291. [FREE Full text] [doi: [10.2196/48291](https://doi.org/10.2196/48291)] [Medline: [37261894](https://pubmed.ncbi.nlm.nih.gov/37261894/)]
22. Heaven WD. Geoffrey Hinton tells us why he's now scared of the tech he helped build. *MIT Technology Review*. URL: <https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai/> [accessed 2024-04-29]
23. Roumeliotis KI, Tselikas ND. ChatGPT and Open-AI models: a preliminary review. *Future Internet*. May 26, 2023;15(6):192. [doi: [10.3390/fi15060192](https://doi.org/10.3390/fi15060192)]
24. Bedi S, Liu Y, Orr-Ewing L, Dash D, Koyejo S, Callahan A, et al. Testing and evaluation of health care applications of large language models: a systematic review. *JAMA*. Jan 28, 2025;333(4):319-328. [doi: [10.1001/jama.2024.21700](https://doi.org/10.1001/jama.2024.21700)] [Medline: [39405325](https://pubmed.ncbi.nlm.nih.gov/39405325/)]
25. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. Mar 19, 2023;11(6):887. [FREE Full text] [doi: [10.3390/healthcare11060887](https://doi.org/10.3390/healthcare11060887)] [Medline: [36981544](https://pubmed.ncbi.nlm.nih.gov/36981544/)]
26. Giudici M, Padalino L, Paolino G, Paratici I, Pascu AI, Garzotto F. Designing home automation routines using an LLM-based chatbot. *Designs*. May 13, 2024;8(3):43. [doi: [10.3390/designs8030043](https://doi.org/10.3390/designs8030043)]
27. Gupta M, Gupta A, Kaur G. Deep learning for medical healthcare: issues, challenges, and opportunities. In: Gupta M, Gupta A, Kaur G, editors. *Computational Analysis and Deep Learning for Medical Care: Principles, Methods, and Applications*. Hoboken, NJ. Wiley-Blackwell; 2022:28-92.
28. Prochaska JJ, Vogel EA, Chieng A, Kendra M, Baiocchi M, Pajarito S, et al. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. *J Med Internet Res*. Mar 23, 2021;23(3):e24850. [FREE Full text] [doi: [10.2196/24850](https://doi.org/10.2196/24850)] [Medline: [33755028](https://pubmed.ncbi.nlm.nih.gov/33755028/)]
29. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR Mhealth Uhealth*. Nov 23, 2018;6(11):e12106. [FREE Full text] [doi: [10.2196/12106](https://doi.org/10.2196/12106)] [Medline: [30470676](https://pubmed.ncbi.nlm.nih.gov/30470676/)]
30. Chow JC. Artificial intelligence in radiotherapy and patient care. In: Lidströmer N, Ashrafian H, editors. *Artificial Intelligence in Medicine*. Cham, Switzerland. Springer; 2021:1-13.
31. Rebelo N, Sanders L, Li K, Chow JC. Learning the treatment process in radiotherapy using an artificial intelligence-assisted chatbot: development study. *JMIR Form Res*. Dec 02, 2022;6(12):e39443. [FREE Full text] [doi: [10.2196/39443](https://doi.org/10.2196/39443)] [Medline: [36327383](https://pubmed.ncbi.nlm.nih.gov/36327383/)]
32. Chow JC, Wong V, Sanders L, Li K. Developing an AI-assisted educational chatbot for radiotherapy using the IBM Watson assistant platform. *Healthcare (Basel)*. Aug 29, 2023;11(17):2417. [FREE Full text] [doi: [10.3390/healthcare11172417](https://doi.org/10.3390/healthcare11172417)] [Medline: [37685452](https://pubmed.ncbi.nlm.nih.gov/37685452/)]
33. Wang A, Qian Z, Briggs L, Cole AP, Reis LO, Trinh QD. The use of chatbots in oncological care: a narrative review. *Int J Gen Med*. May 2023;Volume 16:1591-1602. [doi: [10.2147/ijgm.s408208](https://doi.org/10.2147/ijgm.s408208)]
34. Kovacek D, Chow JC. An AI-assisted chatbot for radiation safety education in radiotherapy. *IOPSciNotes*. Aug 30, 2021;2(3):034002. [doi: [10.1088/2633-1357/ac1f88](https://doi.org/10.1088/2633-1357/ac1f88)]
35. Clark M, Bailey S. Chatbots in health care: connecting patients to information. *Can J Health Technol*. Jan 22, 2024;4(1). [doi: [10.51731/cjht.2024.818](https://doi.org/10.51731/cjht.2024.818)]
36. Nor A, Nur S, Muhamad K, Rozilawati A, Noorazrul Y, Muhammad F, et al. Radiotherapy communication skills training using chatbot-based prototype (SCIMORT). *J Med Imaging Radiat Sci*. Dec 2022;53(4):S1. [doi: [10.1016/j.jmir.2022.10.005](https://doi.org/10.1016/j.jmir.2022.10.005)]

37. Goodman RS, Patrinely JR, Stone CA, Zimmerman E, Donald RR, Chang SS, et al. Accuracy and reliability of chatbot responses to physician questions. *JAMA Netw Open*. Oct 02, 2023;6(10):e2336483. [FREE Full text] [doi: [10.1001/jamanetworkopen.2023.36483](https://doi.org/10.1001/jamanetworkopen.2023.36483)] [Medline: [37782499](https://pubmed.ncbi.nlm.nih.gov/37782499/)]
38. Luo B, Lau RY, Li C, Si YW. A critical review of state - of - the - art chatbot designs and applications. *Wiley Interdiscip Rev Data Min Knowl Discov*. Oct 25, 2021;12(1):e1434. [doi: [10.1002/widm.1434](https://doi.org/10.1002/widm.1434)]
39. Pilato G, Augello A, Gaglio S. A modular system oriented to the design of versatile knowledge bases for chatbots. *ISRN Artif Intell*. Mar 05, 2012;2012:1-10. [doi: [10.5402/2012/363840](https://doi.org/10.5402/2012/363840)]
40. Shalaby W, Arantes A, GonzalezDiaz T, Gupta C. Building chatbots from large scale domain-specific knowledge bases: challenges and opportunities. In: *Proceedings of the 2020 IEEE International Conference on Prognostics and Health Management*. 2020. Presented at: ICPHM '20; June 8-10, 2020:1-8; Detroit, MI. URL: <https://ieeexplore.ieee.org/document/9187036> [doi: [10.1109/icphm49022.2020.9187036](https://doi.org/10.1109/icphm49022.2020.9187036)]
41. English A, Ford CA. The HIPAA privacy rule and adolescents: legal questions and clinical challenges. *Perspect Sex Reprod*. Feb 12, 2007;36(2):80-86. [doi: [10.1363/3608004](https://doi.org/10.1363/3608004)]
42. Hoofnagle CJ, van der Sloot B, Borgesius FZ. The European Union general data protection regulation: what it is and what it means. *Inf Commun Technol Law*. Feb 10, 2019;28(1):65-98. [doi: [10.1080/13600834.2019.1573501](https://doi.org/10.1080/13600834.2019.1573501)]
43. Hasal M, Nowaková J, Ahmed Saghair K, Abdulla H, Snášel V, Ogiela L. Chatbots: security, privacy, data protection, and social aspects. *Concurr Comput Pract Exp*. Jun 03, 2021;33(19):e6426. [doi: [10.1002/cpe.6426](https://doi.org/10.1002/cpe.6426)]
44. Kim B, Park J, Suh J. Transparency and accountability in AI decision support: explaining and visualizing convolutional neural networks for text information. *Decis Support Syst*. Jul 2020;134:113302. [doi: [10.1016/j.dss.2020.113302](https://doi.org/10.1016/j.dss.2020.113302)]
45. Ding R, Palomares I, Wang X, Yang GR, Liu B, Dong Y, et al. Large-Scale decision-making: characterization, taxonomy, challenges and future directions from an Artificial Intelligence and applications perspective. *Inf Fusion*. Jul 2020;59:84-102. [doi: [10.1016/j.inffus.2020.01.006](https://doi.org/10.1016/j.inffus.2020.01.006)]
46. Guan H, Dong L, Zhao A. Ethical risk factors and mechanisms in artificial intelligence decision making. *Behav Sci (Basel)*. Sep 16, 2022;12(9):343. [FREE Full text] [doi: [10.3390/bs12090343](https://doi.org/10.3390/bs12090343)] [Medline: [36135147](https://pubmed.ncbi.nlm.nih.gov/36135147/)]
47. Williams R, Cloete R, Cobbe J, Cottrill C, Edwards P, Markovic M, et al. From transparency to accountability of intelligent systems: moving beyond aspirations. *Data Policy*. Feb 18, 2022;4:e7. [doi: [10.1017/dap.2021.37](https://doi.org/10.1017/dap.2021.37)]
48. Xue J, Wang YC, Wei C, Liu X, Woo J, Kuo CC. Bias and fairness in chatbots: an overview. *arXiv*. Preprint posted online September 16, 2023. [FREE Full text] [doi: [10.1561/116.00000064](https://doi.org/10.1561/116.00000064)]
49. Srinivasan R, González BS. The role of empathy for artificial intelligence accountability. *J Responsib Technol*. Apr 2022;9:100021. [doi: [10.1016/j.jrt.2021.100021](https://doi.org/10.1016/j.jrt.2021.100021)]
50. Tawfeeq TM, Awqati AJ, Jasim YA. The ethical implications of ChatGPT AI chatbot: a review. *J Modern Comput Eng Res*. 2022;2023:49-57. [FREE Full text]
51. Pan A, Musheyev D, Bockelman D, Loeb S, Kabarriti AE. Assessment of artificial intelligence chatbot responses to top searched queries about cancer. *JAMA Oncol*. Oct 01, 2023;9(10):1437-1440. [doi: [10.1001/jamaoncol.2023.2947](https://doi.org/10.1001/jamaoncol.2023.2947)] [Medline: [37615960](https://pubmed.ncbi.nlm.nih.gov/37615960/)]
52. Zhou J, Zhang Y, Luo Q, Parker AG, De Choudhury M. Synthetic lies: understanding ai-generated misinformation and evaluating algorithmic and human solutions. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2022. Presented at: CHI '23; April 29-May 5, 2022:1-20; Virtual Event. URL: <https://dl.acm.org/doi/full/10.1145/3544548.3581318> [doi: [10.1145/3544548.3581318](https://doi.org/10.1145/3544548.3581318)]
53. Monteith S, Glenn T, Geddes JR, Whybrow PC, Achtyes E, Bauer M. Artificial intelligence and increasing misinformation. *Br J Psychiatry*. Feb 26, 2024;224(2):33-35. [doi: [10.1192/bjp.2023.136](https://doi.org/10.1192/bjp.2023.136)] [Medline: [37881016](https://pubmed.ncbi.nlm.nih.gov/37881016/)]
54. Waldrop MM. How to mitigate misinformation. *Proc Natl Acad Sci U S A*. Sep 05, 2023;120(36):e2314143120. [FREE Full text] [doi: [10.1073/pnas.2314143120](https://doi.org/10.1073/pnas.2314143120)] [Medline: [37647368](https://pubmed.ncbi.nlm.nih.gov/37647368/)]
55. Lee JW, Yoo IS, Kim JH, Kim WT, Jeon HJ, Yoo H, et al. Development of AI-generated medical responses using the ChatGPT for cancer patients. *Comput Methods Programs Biomed*. Sep 2024;254:108302. [FREE Full text] [doi: [10.1016/j.cmpb.2024.108302](https://doi.org/10.1016/j.cmpb.2024.108302)] [Medline: [38996805](https://pubmed.ncbi.nlm.nih.gov/38996805/)]
56. Lommatzsch A, Katins J. An information retrieval-based approach for building intuitive chatbots for large knowledge bases. In: *Proceedings of the 2022 Conference on Lernen, Wissen, Daten, Analysen*. 2022. Presented at: LWDA '19; September 30-October 2, 2019:1; Berlin, Germany. URL: https://ceur-ws.org/Vol-2454/paper_60.pdf
57. Minkinen M, Laine J, Mäntymäki M. Continuous auditing of artificial intelligence: a conceptualization and assessment of tools and frameworks. *Digit Soc*. Oct 04, 2022;1(3):1. [doi: [10.1007/S44206-022-00022-2](https://doi.org/10.1007/S44206-022-00022-2)]
58. Davidian K. New space and generative artificial intelligence. *New Space*. Sep 01, 2023;11(3):148-150. [doi: [10.1089/space.2023.29050.editorial](https://doi.org/10.1089/space.2023.29050.editorial)]
59. Bateman S, Brooks C, McCalla G. Collaborative tagging approaches for ontological metadata in adaptive e-learning systems. In: *Proceedings of the 4th international workshop on applications of semantic web technologies for e-learning*. 2006. Presented at: SW-EL '06; June 20, 2006:3-12; Virtual Event. URL: https://www.researchgate.net/profile/Gordon-Mccalla/publication/228464205_Collaborative_tagging_approaches_for_ontological_metadata_in_adaptive_e-learning_systems/links/09e415093093c97b7f000000/Collaborative-tagging-approaches-for-ontological-metadata-in-adaptive-e-learning-systems.

- [pdf?_cf_chl_tk=abcTxToN94_dEPf9TZkFC1r7ibOKThA6yiPuS5HNGQ-1738762941-1.0.1.1-FB3Ss_5QH_GRML7bRa9ySZbuZRxYEVzqYoOo7vCAP28](#)
60. Woo H, Kim J, Lee W. Analysis of cross-referencing artificial intelligence topics based on sentence modeling. *Appl Sci*. May 26, 2020;10(11):3681. [doi: [10.3390/app10113681](#)]
 61. Chen Q, Gong Y, Lu Y, Tang J. Classifying and measuring the service quality of AI chatbot in frontline service. *J Bus Res*. Jun 2022;145:552-568. [doi: [10.1016/j.jbusres.2022.02.088](#)]
 62. Hsu IC, Chang CC. Integrating machine learning and open data into social chatbot for filtering information rumor. *J Ambient Intell Humaniz Comput*. May 20, 2021;12(1):1023-1037. [FREE Full text] [doi: [10.1007/s12652-020-02119-3](#)] [Medline: [32837593](#)]
 63. Vijayaraghavan V, Cooper JB. Algorithm inspection for chatbot performance evaluation. *Procedia Comput Sci*. 2020;171:2267-2274. [doi: [10.1016/j.procs.2020.04.245](#)]
 64. Shawar BA, Atwell E. Different measurement metrics to evaluate a chatbot system. In: *Proceedings of the 2007 Workshop on Bridging the Gap: Academic and Industrial Research in Dialog Technologies*. 2007. Presented at: NAACL-HLT-Dialog '07; April 26, 2007:89-96; New York, NY. URL: <https://dl.acm.org/doi/10.5555/1556328.1556341> [doi: [10.3115/1556328.1556341](#)]
 65. Janssen A, Cardona DR, Passlick J, Bretnier MH. How to make chatbots productive – a user-oriented implementation framework. *Int J Hum Comput Stud*. Dec 2022;168:102921. [doi: [10.1016/j.ijhcs.2022.102921](#)]
 66. Ko AJ, Lee MJ, Ferrari V, Ip S, Tran C. A case study of post-deployment user feedback triage. In: *Proceedings of the 4th International Workshop on Cooperative and Human Aspects of Software Engineering*. 2011. Presented at: CHASE '11; May 21, 2011:1-8; Honolulu, HI. URL: <https://dl.acm.org/doi/abs/10.1145/1984642.1984644> [doi: [10.1145/1984642.1984644](#)]
 67. May R, Denecke K. Security, privacy, and healthcare-related conversational agents: a scoping review. *Inform Health Soc Care*. Apr 03, 2022;47(2):194-210. [doi: [10.1080/17538157.2021.1983578](#)] [Medline: [34617857](#)]
 68. Shumanov M, Johnson L. Making conversations with chatbots more personalized. *Comput Human Behav*. Apr 2021;117:106627. [doi: [10.1016/j.chb.2020.106627](#)]
 69. Nicolau A. The impact of AI on internal audit and accounting practices. *Int Audit Risk Manag*. 2021:38-56. [FREE Full text]
 70. Følstad A, Araujo T, Law EL, Brandtzaeg PB, Papadopoulos S, Reis L, et al. Future directions for chatbot research: an interdisciplinary research agenda. *Computing*. Oct 19, 2021;103(12):2915-2942. [doi: [10.1007/S00607-021-01016-7](#)]
 71. Xin Q, He Y, Pan Y, Wang Y, Du S. The implementation of an AI-driven advertising push system based on a NLP algorithm. *Int J Cogn Comput Sci Inf Technol*. Dec 30, 2023;1(1):30-37. [doi: [10.62051/ijcsit.v1n1.05](#)]
 72. Durall Gazulla E, Martins L, Fernández-Ferrer M. Designing learning technology collaboratively: analysis of a chatbot co-design. *Educ Inf Technol (Dordr)*. Jun 24, 2023;28(1):109-134. [FREE Full text] [doi: [10.1007/s10639-022-11162-w](#)] [Medline: [35765268](#)]
 73. Kumar R, Ali MM. A review on chatbot design and implementation techniques. *Int J Eng Technol*. 2020;7(11):2791. [FREE Full text]
 74. Chow JC, Sanders L, Li K. Design of an educational chatbot using artificial intelligence in radiotherapy. *AI*. Mar 02, 2023;4(1):319-332. [doi: [10.3390/ai4010015](#)]
 75. Denecke K, Abd-Alrazaq A, Househ M, Warren J. Evaluation metrics for health chatbots: a Delphi study. *Methods Inf Med*. Dec 31, 2021;60(5-06):171-179. [doi: [10.1055/s-0041-1736664](#)] [Medline: [34719011](#)]
 76. Kaur A, Singh S, Chandan JS, Robbins T, Patel V. Qualitative exploration of digital chatbot use in medical education: a pilot study. *Digit Health*. Aug 31, 2021;7:20552076211038151. [FREE Full text] [doi: [10.1177/20552076211038151](#)] [Medline: [34513002](#)]
 77. Gabrielli S, Rizzi S, Carbone S, Donisi V. A chatbot-based coaching intervention for adolescents to promote life skills: pilot study. *JMIR Hum Factors*. Feb 14, 2020;7(1):e16762. [FREE Full text] [doi: [10.2196/16762](#)] [Medline: [32130128](#)]
 78. Stout NL, Silver JK, Raj VS, Rowland J, Gerber L, Cheville A, et al. Toward a national initiative in cancer rehabilitation: recommendations from a subject matter expert group. *Arch Phys Med Rehabil*. Nov 2016;97(11):2006-2015. [doi: [10.1016/j.apmr.2016.05.002](#)] [Medline: [27237580](#)]
 79. Abd-Alrazaq A, Safi Z, Alajlani M, Warren J, Househ M, Denecke K. Technical metrics used to evaluate health care chatbots: scoping review. *J Med Internet Res*. Jun 05, 2020;22(6):e18301. [FREE Full text] [doi: [10.2196/18301](#)] [Medline: [32442157](#)]
 80. Jungmann SM, Klan T, Kuhn S, Jungmann F. Accuracy of a chatbot (Ada) in the diagnosis of mental disorders: comparative case study with lay and expert users. *JMIR Form Res*. Oct 29, 2019;3(4):e13863. [FREE Full text] [doi: [10.2196/13863](#)] [Medline: [31663858](#)]
 81. Shan Y, Ji M, Xie W, Zhang X, Qian X, Li R, et al. Use of health care chatbots among young people in China during the omicron wave of COVID-19: evaluation of the user experience of and satisfaction with the technology. *JMIR Hum Factors*. Jun 09, 2022;9(2):e36831. [FREE Full text] [doi: [10.2196/36831](#)] [Medline: [35576058](#)]
 82. Laymouna M, Ma Y, Lessard D, Schuster T, Engler K, Lebouché B. Roles, users, benefits, and limitations of chatbots in health care: rapid review. *J Med Internet Res*. Jul 23, 2024;26:e56930. [FREE Full text] [doi: [10.2196/56930](#)] [Medline: [39042446](#)]

83. Ren R, Zapata M, Castro JW, Dieste O, Acuna ST. Experimentation for chatbot usability evaluation: a secondary study. *IEEE Access*. 2022;10:12430-12464. [doi: [10.1109/access.2022.3145323](https://doi.org/10.1109/access.2022.3145323)]
84. Montenegro JL, da Costa CA, Janssen LP. Evaluating the use of chatbot during pregnancy: a usability study. *Healthc Anal*. Nov 2022;2:100072. [doi: [10.1016/j.health.2022.100072](https://doi.org/10.1016/j.health.2022.100072)]
85. Ye BJ, Kim JY, Suh C, Choi SP, Choi M, Kim DH, et al. Development of a chatbot program for follow-up management of workers' general health examinations in Korea: a pilot study. *Int J Environ Res Public Health*. Feb 23, 2021;18(4):2170. [FREE Full text] [doi: [10.3390/ijerph18042170](https://doi.org/10.3390/ijerph18042170)] [Medline: [33672158](https://pubmed.ncbi.nlm.nih.gov/33672158/)]
86. Singh S, Mahmood A. The NLP cookbook: modern recipes for transformer based deep learning architectures. *IEEE Access*. 2021;9:68675-68702. [doi: [10.1109/access.2021.3077350](https://doi.org/10.1109/access.2021.3077350)]
87. Jeong H, Yoo JH, Han O. Next-generation chatbots for adaptive learning: a proposed framework. *J Internet Comput Serv*. 2023;24(4):37-45. [FREE Full text]
88. KC GP, Ranjan S, Ankit T, Kumar V. A personalized medical assistant chatbot: Medibot. *Int J Sci Technol Eng*. 2019;5(7):42-46. [FREE Full text]
89. Zhang H, Zheng J. The application analysis of medical chatbots and virtual assistant. *Front Soc Sci Technol*. 2021;3(2):11-16. [FREE Full text] [doi: [10.25236/fsst.2021.030202](https://doi.org/10.25236/fsst.2021.030202)]
90. Lin AP, Trappey CV, Luan CC, Trappey AJ, Tu KL. A test platform for managing school stress using a virtual reality group chatbot counseling system. *Appl Sci*. Sep 29, 2021;11(19):9071. [doi: [10.3390/app11199071](https://doi.org/10.3390/app11199071)]
91. Orosoo M, Goswami I, Alphonse FR, Fatma G, Rengarajan M, Bala BK. Enhancing natural language processing in multilingual chatbots for cross-cultural communication. In: *Proceedings of the 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks*. 2024. Presented at: ICICV '24; March 11-12, 2024:127-133; Virtual Event. URL: <https://www.computer.org/csdl/proceedings-article/icicv/2024/856400a127/1WKFMzaKGcw> [doi: [10.1109/icicv62344.2024.00027](https://doi.org/10.1109/icicv62344.2024.00027)]
92. Fournier-Tombs E, McHardy J. A medical ethics framework for conversational artificial intelligence. *J Med Internet Res*. Jul 26, 2023;25:e43068. [FREE Full text] [doi: [10.2196/43068](https://doi.org/10.2196/43068)] [Medline: [37224277](https://pubmed.ncbi.nlm.nih.gov/37224277/)]
93. Chang CY, Kuo SY, Hwang GH. Chatbot-facilitated nursing education: incorporating a knowledge-based chatbot system into a nursing training program. *Educ Technol Soc*. 2022;25(1):15-27. [FREE Full text]
94. Murad DF, Iskandar AG, Fernando E, Octavia TS, Maured DE. Towards smart LMS to improve learning outcomes students using LenoBot with natural language processing. In: *Proceedings of the 6th International Conference on Information Technology, Computer and Electrical Engineering*. 2019. Presented at: ICITACEE '19; September 26, 2019 26:1-6; Semarang, Indonesia. URL: <https://ieeexplore.ieee.org/document/8904311> [doi: [10.1109/icitacee.2019.8904311](https://doi.org/10.1109/icitacee.2019.8904311)]
95. Parviainen J, Rantala J. Chatbot breakthrough in the 2020s? An ethical reflection on the trend of automated consultations in health care. *Med Health Care Philos*. Mar 04, 2022;25(1):61-71. [FREE Full text] [doi: [10.1007/s11019-021-10049-w](https://doi.org/10.1007/s11019-021-10049-w)] [Medline: [34480711](https://pubmed.ncbi.nlm.nih.gov/34480711/)]

Abbreviations

- AI:** artificial intelligence
- LLM:** large language model
- NLP:** natural language processing
- VR/AR:** virtual and augmented reality

Edited by N Cahill; submitted 18.09.24; peer-reviewed by M Chatzimina, M Dursun, H Polat; comments to author 11.12.24; revised version received 15.12.24; accepted 16.01.25; published 18.02.25

Please cite as:

Chow JCL, Li K

Developing Effective Frameworks for Large Language Model-Based Medical Chatbots: Insights From Radiotherapy Education With ChatGPT

JMIR Cancer 2025;11:e66633

URL: <https://cancer.jmir.org/2025/1/e66633>

doi: [10.2196/66633](https://doi.org/10.2196/66633)

PMID:

©James C L Chow, Kay Li. Originally published in *JMIR Cancer* (<https://cancer.jmir.org>), 18.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR*

Cancer, is properly cited. The complete bibliographic information, a link to the original publication on <https://cancer.jmir.org/>, as well as this copyright and license information must be included.