Large Language Model-Enabled Medical Chatbots

Contributor: James C. L. Chow, Valerie Wong, Kay Li

The trajectory of artificial intelligence (AI) development spans decades, with machine learning (ML) emerging as a pivotal force in propelling AI's evolution. The adoption of AI and ML in the medical field has experienced significant growth, particularly in ML-enabled medical devices. Chatbots, AI-driven conversational agents prevalent in online interactions, have found extensive utility in disseminating healthcare information and enhancing customer services. These features encompass accurate information retrieval, symptom assessment, and diagnosis support to help in understanding and addressing health concerns.

1. Introduction

The trajectory of artificial intelligence (AI) development spans decades, with machine learning (ML) emerging as a pivotal force in propelling AI's evolution [1][2][3][4]. The adoption of AI and ML in the medical field has experienced significant growth, particularly in ML-enabled medical devices. Joshi et al. focused on 691 FDA-approved AI/ML-enabled medical devices, revealing a substantial surge in approvals since 2018, predominantly in radiology. The prevalence of the 510(k) clearance pathway, relying on substantial equivalence, is notable [5]. A specific ML facet is focused: the Large Language Model (LLM) within Natural Language Processing (NLP) [6][7]. Particularly, the researchers delve into the integration of LLMs like Chat Generative Pre-trained Transformer (ChatGPT, version 3–4) into chatbots, augmenting their capacity for seamless user engagement [8][9][10].

Chatbots, AI-driven conversational agents prevalent in online interactions, have found extensive utility in disseminating healthcare information and enhancing customer services [11][12][13][14][15]. **Table 1** summarizes the general features that medical professionals would expect a medial chatbot to have. These features encompass accurate information retrieval, symptom assessment, and diagnosis support to help in understanding and addressing health concerns. Moreover, the chatbot is expected to provide treatment guidance, medication information, and assistance with appointment scheduling, ensuring a comprehensive healthcare experience. Health monitoring features, emergency response capabilities, and patient education contribute to a holistic approach. Privacy and security measures, multilingual support, and integration with electronic health records uphold standards of confidentiality and accessibility. Personalized recommendations, follow-up mechanisms, and a user-friendly interface tailor the chatbot experience to individual needs, while features like adherence support and mental health resources further enhance its utility. Continuous feedback mechanisms ensure ongoing improvement, making the chatbot a valuable tool to promote patient well-being. The advent of ChatGPT has notably elevated the appeal of chatbots, facilitating more human-like interactions through adaptive text learning [16][17]. However, the precision of healthcare information dispensed by ChatGPT still raises some concerns, prompting inquiries into potential user misguidance [18][19][20].

Table 1. Features of AI chatbot expected by the medical professional.

Feature	Description
Accurate Information Retrieval	Provide accurate and up-to-date medical information from reliable sources.
Symptom Assessment	Analyze and assess user-described symptoms to suggest potential health conditions.
Diagnosis Support	Offer preliminary assistance in suggesting potential diagnoses, understanding its limitations.
Treatment Guidance	Provide general information on treatments, medications, and lifestyle recommendations.
Medication Information	Offer details about medications, including dosage, side effects, and potential interactions.

Feature	Description
Appointment Scheduling	Assist users in scheduling appointments with healthcare providers and send reminders.
Health Monitoring	Support users in tracking and monitoring health metrics like blood pressure or blood sugar.
Emergency Response	Recognize urgent situations and provide emergency response information or facilitate contacts.
Patient Education	Offer educational content to enhance users' understanding of medical conditions and prevention.
Privacy and Security	Ensure strict adherence to data privacy regulations and maintain the confidentiality of user health information.
Multilingual Support	Provide communication in multiple languages to cater to diverse patient populations.
Integration with EHR	Facilitate integration with existing healthcare systems to access relevant patient data.
Personalized Recommendations	Offer personalized health advice based on user data, preferences, and lifestyle.
Follow-up and Continuity of Care	Implement features for follow-up interactions, reminders, and maintaining continuity of care.
User-Friendly Interface	Ensure an intuitive and user-friendly interface for easy interaction.
Adherence Support	Assist patients in adhering to prescribed treatment plans and medications.
Mental Health Support	Include features for mental health assessments, stress management, and access to mental health resources.
Feedback and Improvement	Incorporate mechanisms for users to provide feedback on the chatbot's performance.

2. Fundaments and Evolution of Language Models

2.1. Fundamentals of Natural Language Processing

NLP stands as a cornerstone in the realm of ML, a subset of AI that learns from data to approximate human expectations [21][22]. Particularly, NLP plays a pivotal role in facilitating AI's comprehension of the diverse languages used by individuals. Chatbots integrated with NLP capabilities excel in learning and understanding the natural language patterns employed by users in textual communication, enabling them to respond intelligibly [23][24]. **Figure 1** shows a typical chatbot architecture, including the user interface, user message analysis component, dialog management component, responses generation component, and the database [24]. The NLP is mainly linked to the message analysis component to analyze the context information.
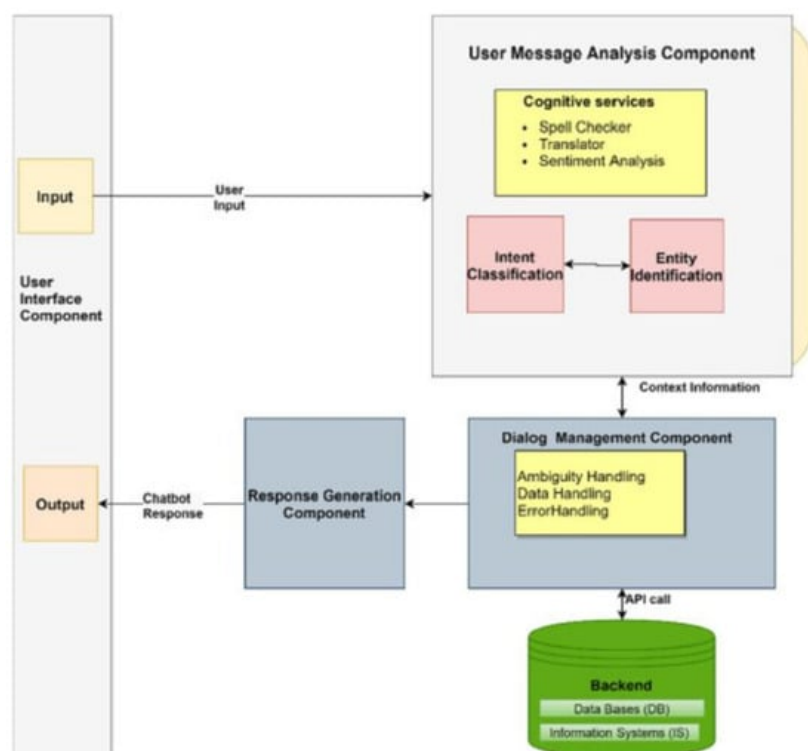
Figure 1. General chatbot architecture. Source: Adapted from [24].

Given the inherent variability in how individuals communicate, lacking a standardized template or exact pattern, ML, especially NLP, strives to analyze free-text and speech through linguistic and statistical algorithms. This analysis aims to extract discernible patterns from the rich tapestry of human expression [25][26]. While pattern analysis forms the foundation, the evolution of AI necessitates its ability to engage in meaningful conversations with users, primarily exemplified in question-answering (QA) scenarios [27][28].

The acquired text patterns are cataloged in a database, empowering the AI to match these learned patterns during user interactions—a process akin to pattern matching and text searching techniques [29]. Crucially, NLP goes beyond mere pattern recognition; it grapples with the nuances of how individuals articulate ideas. This involves understanding that distinct expressions can convey the same meaning, enabling AI to emulate human-like responses, thus enhancing the conversational experience [30][31].

It can be seen that the training of the model involves exposing the algorithm to vast amounts of text data, allowing it to learn the patterns, semantics, and structures inherent in human language. This process typically utilizes large datasets to train the model on tasks such as language understanding, sentiment analysis, or question answering. The NLP model undergoes iterative adjustments during training, refining its ability to recognize and generate meaningful language output. The ultimate goal is to enhance the model's proficiency in understanding and generating human-like text, enabling it to perform diverse linguistic tasks with accuracy and relevance.

In the healthcare domain, NLP demonstrates its prowess by extracting pertinent information from free-text documents such as electronic health records. Beyond symptom examination, NLP's ability to compare, classify, and recommend actions based on vast sets of textual data contributes significantly to disease symptom classification and patient guidance [32][33]. NLP emerges as a linchpin in the intersection of AI and healthcare, fostering a nuanced understanding of language patterns, enhancing conversational dynamics, and contributing invaluable insights in the medical field [34]. The following exploration further delves into the applications and implications of NLP in healthcare conversations.

2.2. Evolution of Large Language Models

LLMs have emerged as transformative components within the NLP, significantly influencing the evolution of AI [35][36]. LLMs, belonging to the broader category of machine learning, excel in processing and generating human-like text by leveraging extensive datasets. Their remarkable ability to capture intricate language nuances and generate coherent responses has positioned them as integral players in advancing NLP [37]. In healthcare conversations, LLMs play a crucial role in enhancing the conversational capabilities of AI systems [38]. By understanding and generating contextually relevant responses, LLMs contribute to the humanization of interactions, creating more engaging and effective healthcare dialogues.

Several key LLMs have made a significant impact on the healthcare conversations, revolutionizing the way AI engages with users. One noteworthy exemplar is OpenAI's GPT (Generative Pre-trained Transformer) series, with models like GPT-3 and 4 demonstrating exceptional language understanding and generation capabilities [39][40]. GPT-4, in particular, has garnered attention for its versatility in various applications, including healthcare-related tasks [41]. BERT (Bidirectional Encoder Representations from Transformers) is another influential LLM that has left an indelible mark on NLP [42][43]. Renowned for its bidirectional training approach, BERT excels in grasping contextual nuances, making it particularly adept at understanding the intricacies of medical language and information. Furthermore, models like XLNet [44], T5 (Text-to-Text Transfer Transformer) [45], and BART (Bidirectional and Auto-Regressive Transformers) [46] have played instrumental roles in advancing the sophistication of LLMs in healthcare applications. These models exhibit enhanced capabilities in processing medical literature, extracting relevant information, and generating coherent responses tailored to healthcare-related inquiries [47]. **Table 2** shows some popular LLMs used in healthcare conversations.

Table 2. Some LLMs and their potential applications in healthcare conversations.

Model	Description	Applications in Healthcare
GPT-3 and 4 [39][40][41]	OpenAI's powerful LLM with strong natural language understanding.	Medical documentation, question answering, text-based interactions
BERT, BioBERT and ClinicalBERT [42][43]	Bidirectional processing makes BERT suitable for clinical text analysis.	Clinical text analysis, medical literature understanding, biomedical text mining, information extraction from medical texts, clinical note understanding, medical question answering

Model	Description	Applications in Healthcare
XLNet [44]	OpenAI's model capable of capturing bidirectional context.	Medical literature analysis, clinical documentation
T5 [45]	Text-to-Text Transfer Transformer, designed for various NLP tasks.	Summarization of medical documents, question generation
BART [46]	Bidirectional and Auto-Regressive Transformer, used for text generation.	Text summarization, document generation, paraphrasing

The utilization of LLMs in healthcare conversations signifies a paradigm shift, enabling AI systems to comprehend and respond to user queries with a depth of understanding akin to human-like interactions. As we navigate through the evolutionary trajectory of LLMs, their continued refinement and integration into healthcare dialogue systems hold promise for further augmenting the efficacy and user experience in the realm of medical conversations.

The incorporation of LLMs into medical chatbots introduces significant advantages, revolutionizing the healthcare interactions. LLMs enhance the understanding of contextual nuances in medical queries, enabling chatbots to provide more accurate and relevant responses [48]. This heightened comprehension fosters a humanized interaction, with LLMs proficiently mimicking natural language patterns, creating a more engaging and empathetic user experience. Additionally, LLMs empower medical chatbots to efficiently retrieve and disseminate precise medical information, positioning them as reliable sources of up-to-date healthcare knowledge [49]. **Figure 2** shows the touchpoints of a patient's care journey in which an LLM can be employed to enhance the patient's experience.
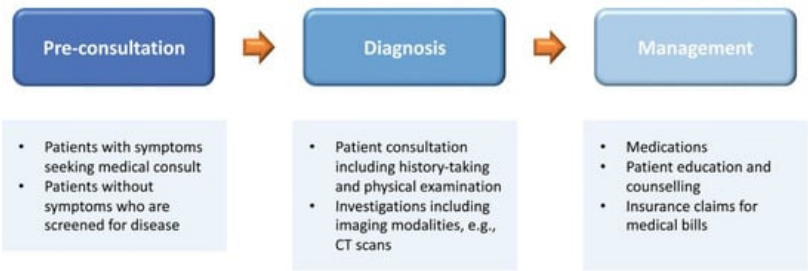

Figure 2. A standard patient journey in healthcare, encompassing three key stages: (1) pre-consultation involves patient registration, medical consultation, or health screening; (2) diagnosis includes patient consultations, examinations, and supplementary investigations; and (3) management comprises medication, patient counseling, education, and reimbursement for medical bills. LLMs exhibit potential to improve the patient experience at each touchpoint in this journey. Source: Adapted from [49].

However, the implementation of LLMs in medical chatbots is not without challenges. Ensuring the accuracy and trustworthiness of information is paramount, as LLMs may inadvertently generate inaccurate responses, posing a risk of misinformation [50][51]. Privacy and security concerns arise, demanding robust measures to safeguard sensitive health-related data [52]. Furthermore, interpreting complex medical terminology and aligning with user expectations present ongoing challenges. Addressing these hurdles is essential to fully harness the potential benefits of LLMs in the dynamic realm of medical chatbots [18].

3. Application of AI Chatbot in Healthcare

3.1. Healthcare Knowledge Transfer with Chatbot

As mentioned, the general public are more likely to use chatbots to look for answers to their questions in daily life or even in medicine. So it is important to note that chatbots might obtain wrong information and mislead users, unless developing a healthcare-based chatbot that is designed to answer people with accurate medical information [53][54]. People would not know if they were misled and believe the provided information blindly, which could lead to accidents. It would be best to develop medical chatbots with professionals like doctors and nurses to ensure the information is accurate and easy to understand for the general public. If LLMs could be trained specifically in healthcare then their trustworthiness would be increased, but this might require professionals to verify that every single piece of information provided by the LLM is correct [55]. LLMs' trustworthiness is controversial; some researchers think they simply gather information from the internet and provide it to the users, while other researchers think they have the potential to be trained specifically for healthcare purposes with related journals [56]. On the other hand, using chatbots for healthcare knowledge transfer can help prevent

clinicians from answering similar questions from various patients repeatedly, and hence could allow clinicians to work on jobs that are highly prioritized [57]. As such, there remains a concern of accuracy and reliability of the information that researchers always keep in mind [18].

3.2. Symptom Diagnosis

According to the study of Kumar et al. [58], training an AI model with a number of papers in the healthcare field means that the model is able to analyze and predict the symptoms of diseases. Aside from simply classifying symptoms, clinicians are responsible for compiling electronic health records (EHRs), which is a patient information managing digital system, for every single patient visiting a hospital [59]. Having an LLM trained for classifying symptoms according to doctor–patient conversations would increase the efficiency of seeing each patient and thus reduce the chance of overcrowding during busy hours [59]. In fact, there is already a ChatGPT-like chatbot created for healthcare, Med-PaLM, which is able to analyze X-ray images according to the examples given in [60]. Moreover, it is possible for the LLM to integrate with telehealth services especially for those who have a disability [49]. In addition, it is believed that LLMs are even able to detect later-life depression, which is a kind of major public health concern that occurs in the older generation, as the name states [61]. NLP analyzes the way people talk and also the speed and pitch they use in order to understand their speech patterns [61]. Having understood their speech patterns, it is possible to discover the speech patterns of people with depression and use these as a kind of template to analyze and compare the template patterns with the target user in order to determine whether they are possibly experiencing later-life depression.

3.3. Mental Healthcare

Having a conversation with someone is the easiest way to help balance mental health, as people are able to express their feelings when they talk to someone. It is not difficult to imagine that people can have conversations with LLM chatbots like ChatGPT especially when there is no one to talk with [62][63]. Stress could simply keep building up when something bad happens in a person's daily life and they cannot talk to someone, and then the stress reaches a limit and their performance at work may worsen and their health might also be affected to some extent. Chatbots can help with daily emotional support; there are studies proving they are capable of helping people get rid of stress and feelings of depression, which also somehow demonstrate better results than traditional mental heath treatments [64]. With NLP analyzing the texts posted on different social media platforms, the AI model is able to perform detection of emotions and monitoring of mental health [65]. It is a sort of text classification scenario in that the NLP technique allows the AI model to analyze texts and compare them with other similar texts and classify them into different cases in order to detect mental illness from those [66]. In social media, AI models most commonly focus on detecting scenarios like suicide; when they detect any wording that might possibly relate to suicide, it recommends users to contact mentors for mental support [67]. Other than that, there are many people living alone and some of them might not have anyone around them with whom they can share their feelings of daily life, for example, they cannot express the negativity they feel at school or in work. People get depressed when negativity and stress keep building up without letting it out by having someone to talk to; so, it would be nice if chatbots can become a sort of a place for people to let out their stresses [68]. It is hard to live alone in society, especially when there are infinite factors that can make people have a rough day and stress builds and bursts out when there is no way to release it; people can get angry at no one without a reason and this might lead to a fight, which could bring down the quality of life around the community.

4. Ethical and Legal Implications

4.1. Patient Privacy and Data Security Concerns

The integration of AI, particularly LLMs, into healthcare conversations brings forth ethical and legal considerations, with the foremost among them being patient privacy and data security. As medical chatbots process sensitive health information, ensuring robust measures for data encryption, storage, and transmission becomes paramount [69]. Ethical considerations demand that patient data are handled with the utmost confidentiality and that stringent protocols are in place to prevent unauthorized access or breaches, safeguarding the trust patients place in AI-assisted healthcare interactions [70][71].

4.2. Ethical Considerations in AI-Assisted Healthcare Conversations

Beyond privacy concerns, ethical considerations play a pivotal role in the deployment of AI-assisted healthcare conversations. Ensuring transparency and informed consent becomes crucial when patients engage with medical chatbots. The ethical development and use of AI models, including LLMs, involve addressing biases, avoiding discrimination, and maintaining fairness in the provision of healthcare information [72][73]. Striking the right balance

between technological advancements and ethical principles is essential to build a foundation of trust between patients, healthcare providers, and AI systems [74].

4.3. Regulatory Compliance in AI-Powered Healthcare Applications

Effectively managing regulatory complexities is a multifaceted challenge when incorporating AI-powered healthcare applications. It is crucial to uphold adherence to established healthcare regulations, exemplified by the Health Insurance Portability and Accountability Act (HIPAA) in the United States. Ensuring compliance with these regulatory standards is not only essential for safeguarding patient rights but also forms the bedrock for responsible AI deployment [75]. This adherence mitigates legal risks and fosters a seamless integration of technology into healthcare practices. The evolving dynamics of health-related conversations, driven by AI chatbots, necessitate a thorough understanding of ethical and legal implications. As the regulatory landscape continues to evolve, it is noteworthy to mention the European Union's AI Act [76], which introduces regulations specific to AI systems, emphasizing transparency, accountability, and user safety in the deployment of AI technologies across various sectors [77].

References

1. Confalonieri, R.; Coba, L.; Wagner, B.; Besold, T.R. A historical perspective of explainable Artificial Intelligence. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2021, 11, e1391.

2. Kononenko, I. Machine learning for medical diagnosis: History, state of the art and perspective. Artif. Intell. Med. 2001, 23, 89–109.

3. Siddique, S.; Chow, J.C. Artificial intelligence in radiotherapy. Rep. Pract. Oncol. Radiother. 2020, 25, 656–666.

4. Chow, J.C. Internet-based computer technology on radiotherapy. Rep. Pract. Oncol. Radiother. 2017, 22, 455–462.

5. Joshi, G.; Jain, A.; Araveeti, S.R.; Adhikari, S.; Garg, H.; Bhandari, M. FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape. Electronics 2024, 13, 498.

6. Raiaan, M.A.K.; Mukta, M.S.H.; Fatema, K.; Fahad, N.M.; Sakib, S.; Mim, M.M.J.; Ahmad, J.; Ali, M.E.; Azam, S. A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE Access 2024, 12, 26839–26874.

7. Khan, W.; Daud, A.; Khan, K.; Muhammad, S.; Haq, R. Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends. Nat. Lang. Process. J. 2023, 4, 100026.

8. Thirunavukarasu, A.J.; Ting, D.S.J.; Elangovan, K.; Gutierrez, L.; Tan, T.F.; Ting, D.S.W. Large language models in medicine. Nat. Med. 2023, 29, 1930–1940.

9. Kim, J.K.; Chua, M.; Rickard, M.; Lorenzo, A. ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine. J. Pediatr. Urol. 2023, 19, 598–604.

10. Haupt, C.E.; Marks, M. AI-generated medical advice—GPT and beyond. JAMA 2023, 329, 1349–1350.

11. Siddique, S.; Chow, J.C.L. Machine learning in healthcare communication. Encyclopedia 2021, 1, 220–239.

12. Xu, L.; Sanders, L.; Li, K.; Chow, J.C.L. Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. JMIR Cancer 2021, 7, e27850.

13. Chow, J.C.L.; Wong, V.; Sanders, L.; Li, K. Developing an AI-Assisted Educational Chatbot for Radiotherapy Using the IBM Watson Assistant Platform. Healthcare 2023, 11, 2417.

14. Kovacek, D.; Chow, J.C.L. An AI-assisted chatbot for radiation safety education in radiotherapy. IOP SciNotes 2021, 2, 034002.

15. Lalwani, T.; Bhalotia, S.; Pal, A.; Rathod, V.; Bisen, S. Implementation of a Chatbot System using AI and NLP. Int. J. Innov. Res. Comput. Sci. Technol. IJIRCST 2018, 6, 26–30.

16. Wu, T.; He, S.; Liu, J.; Sun, S.; Liu, K.; Han, Q.-L.; Tang, Y. A brief overview of ChatGPT: The history, status quo and potential future development. IEEE/CAA J. Autom. Sin. 2023, 10, 1122–1136.

17. Li, J.; Dada, A.; Puladi, B.; Kleesiek, J.; Egger, J. ChatGPT in healthcare: A taxonomy and systematic review. Comput. Methods Programs Biomed. 2024, 245, 108013.

18. Chow, J.C.L.; Sanders, L.; Li, K. Impact of ChatGPT on medical chatbots as a disruptive technology. Front. Artif. Intell. 2023, 6, 1166014.

19. Kao, H.-J.; Chien, T.-W.M.; Wang, W.-C.; Chou, W.; Chow, J.C. Assessing ChatGPT's capacity for clinical decision support in pediatrics: A comparative study with pediatricians using KIDMAP of Rasch analysis. Medicine 2023, 102, e34068.

20. Rawashdeh, B.; Kim, J.; AlRyalat, S.A.; Prasad, R.; Cooper, M. ChatGPT and artificial intelligence in transplantation research: Is it always correct? Cureus 2023, 15, e42150.

21. Ayanouz, S.; Abdelhakim, B.A.; Benhmed, M. A smart chatbot architecture based NLP and machine learning for health care assistance. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security, Marrakech, Morocco, 31 March–2 April 2020; pp. 1–6.

22. Olthof, A.W.; Shouche, P.; Fennema, E.M.; IJpma, F.F.; Koolstra, R.C.; Stirler, V.M.; van Ooijen, P.M.; Cornelissen, L.J. Machine learning based natural language processing of radiology reports in orthopaedic trauma. Comput. Methods Programs Biomed. 2021, 208, 106304.

23. Adamopoulou, E.; Moussiades, L. An overview of chatbot technology. In IFIP International Conference on Artificial Intelligence Applications and Innovations; Springer: Cham, Switzerland, 2020; pp. 373–383.

24. Adamopoulou, E.; Moussiades, L. Chatbots: History, technology, and applications. Mach. Learn. Appl. 2020, 2, 100006.

25. Chadha, N.; Gangwar, R.; Bedi, R. Current Challenges and Application of Speech Recognition Process using Natural Language Processing: A Survey. Int. J. Comput. Appl. 2015, 131, 28–31.

26. Malik, M.; Malik, M.K.; Mehmood, K.; Makhdoom, I. Automatic speech recognition: A survey. Multimed. Tools Appl. 2021, 80, 9411–9457.

27. Zaib, M.; Zhang, W.E.; Sheng, Q.Z.; Mahmood, A.; Zhang, Y. Conversational question answering: A survey. Knowl. Inf. Syst. 2022, 64, 3151–3195.

28. Reddy, S.; Chen, D.; Manning, C.D. Coqa: A conversational question answering challenge. Trans. Assoc. Comput. Linguist. 2019, 7, 249–266.

29. Kocaleva, M.; Stojanov, D.; Stojanovik, I.; Zdravev, Z. Pattern recognition and natural language processing: State of the art. TEM J. 2016, 5, 236–240.

30. Fu, T.; Gao, S.; Zhao, X.; Wen, J.-R.; Yan, R. Learning towards conversational AI: A survey. AI Open 2022, 3, 14–28.

31. Sharma, D.; Paliwal, M.; Rai, J. NLP for Intelligent Conversational Assistance. Int. J. Innov. Res. Comput. Sci. Technol. 2021, 9, 179–184.

32. Locke, S.; Bashall, A.; Al-Adely, S.; Moore, J.; Wilson, A.; Kitchen, G.B. Natural language processing in medicine: A review. Trends Anaesth. Crit. Care 2021, 38, 4–9.

33. Lo Barco, T.; Kuchenbuch, M.; Garcelon, N.; Neuraz, A.; Nabbout, R. Improving early diagnosis of rare diseases using Natural Language Processing in unstructured medical records: An illustration from Dravet syndrome. Orphanet J. Rare Dis. 2021, 16, 309.

34. Friedman, C.; Hripcsak, G. Natural language processing and its future in medicine. Acad. Med. 1999, 74, 890–895.

35. Khan, R.; Gupta, N.; Sinhababu, A.; Chakravarty, R. Impact of Conversational and Generative AI Systems on Libraries: A Use Case Large Language Model (LLM). Sci. Technol. Libr. 2023, 42, 1–5.

36. Alberts, I.L.; Mercolli, L.; Pyka, T.; Prenosil, G.; Shi, K.; Rominger, A.; Afshar-Oromieh, A. Large language models (LLM) and ChatGPT: What will the impact on nuclear medicine be? Eur. J. Nucl. Med. 2023, 50, 1549–1552.

37. Ethape, P.; Kane, R.; Gadekar, G.; Chimane, S. Smart Automation Using LLM. Int. Res. J. Innov. Eng. Technol. 2023, 7, 603.

38. El Saddik, A.; Ghaboura, S. The Integration of ChatGPT with the Metaverse for Medical Consultations. IEEE Consum. Electron. Mag. 2024, 13, 6–15.

39. Roumeliotis, K.I.; Tselikas, N.D. ChatGPT and Open-AI Models: A Preliminary Review. Future Internet 2023, 15, 192.

40. De Angelis, L.; Baglivo, F.; Arzilli, G.; Privitera, G.P.; Ferragina, P.; Tozzi, A.E.; Rizzo, C. ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. Front. Public Health 2023, 11, 1166120.

41. Waisberg, E.; Ong, J.; Masalkhi, M.; Kamran, S.A.; Zaman, N.; Sarker, P.; Lee, A.G.; Tavakkoli, A. GPT-4: A new era of artificial intelligence in medicine. Ir. J. Med Sci. 2023, 192, 3197–3200.

42. Acheampong, F.A.; Nunoo-Mensah, H.; Chen, W. Transformer models for text-based emotion detection: A review of BERT-based approaches. Artif. Intell. Rev. 2021, 54, 5789–5829.

43. Sayeed, M.S.; Mohan, V.; Muthu, K.S. BERT: A Review of Applications in Sentiment Analysis. HighTech Innov. J. 2023, 4, 453–462.

44. Yang, Z.; Dai, Z.; Yang, Y.; Carbonell, J.; Salakhutdinov, R.R.; Le, Q.V. Xlnet: Generalized autoregressive pretraining for language understanding. Adv. Neural Inf. Process. Syst. 2019, 32. Available online: https://api.semanticscholar.org/CorpusID:195069387 (accessed on 11 March 2024).

45. Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Zhou, Y.; Li, W.; Liu, P.J. Exploring the limits of transfer learning with a unified text-to-text transformer. J. Mach. Learn. Res. 2020, 21, 5485–5551.

46. Hao, Y.; Dong, L.; Wei, F.; Xu, K. Visualizing and understanding the effectiveness of BERT. arXiv 2019, arXiv:1908.05620.

47. Catelli, R.; Pelosi, S.; Esposito, M. Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. Electronics 2022, 11, 374.

48. Chow, J.C. Artificial intelligence in radiotherapy and patient care. In Artificial Intelligence in Medicine; Springer: Cham, Switzerland, 2021; pp. 1–13.

49. Yang, R.; Tan, T.F.; Lu, W.; Thirunavukarasu, A.J.; Ting, D.S.W.; Liu, N. Large language models in health care: Development, applications, and challenges. Health Care Sci. 2023, 2, 255–263.

50. Chakraborty, C.; Bhattacharya, M.; Lee, S.-S. Need an AI-enabled, next-generation, advanced ChatGPT or large language models (LLMs) for error-free and accurate medical information. Ann. Biomed. Eng. 2023, 52, 134–135.

51. Sanaei, M.-J.; Ravari, M.S.; Abolghasemi, H. ChatGPT in medicine: Opportunity and challenges. Iran. J. Blood Cancer 2023, 15, 60–67.

52. Adhikari, K.; Naik, N.; Hameed, B.Z.; Raghunath, S.K.; Somani, B.K. Exploring the ethical, legal, and social implications of ChatGPT in urology. Curr. Urol. Rep. 2023, 25, 1–8.

53. Goodman, R.S.; Patrinely, J.R.; Stone, C.A.; Zimmerman, E.; Donald, R.R.; Chang, S.S.; Berkowitz, S.T.; Finn, A.P.; Jahangir, E.; Scoville, E.A.; et al. Accuracy and reliability of chatbot responses to physician questions. JAMA Netw. Open 2023, 6, e2336483.

54. Walker, H.L.; Ghani, S.; Kuemmerli, C.; Nebiker, C.A.; Müller, B.P.; Raptis, D.A.; Staubli, S.M. Reliability of medical information provided by ChatGPT: Assessment against clinical guidelines and patient information quality instrument. J. Med. Internet Res. 2023, 25, e47479.

55. Fournier-Tombs, E.; McHardy, J. A medical ethics framework for conversational artificial intelligence. J. Med. Internet Res. 2023, 25, e43068.

56. Chang, I.-C.; Shih, Y.-S.; Kuo, K.-M. Why would you use medical chatbots? interview and survey. Int. J. Med. Inform. 2022, 165, 104827.

57. Chung, K.; Park, R.C. Chatbot-based heathcare service with a knowledge base for cloud computing. Clust. Comput. 2019, 22, 1925–1937.

58. Kumar, Y.; Koul, A.; Singla, R.; Ijaz, M.F. Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda. J. Ambient. Intell. Humaniz. Comput. 2023, 14, 8459–8486.

59. Lee, S.; Lee, J.; Park, J.; Park, J.; Kim, D.; Lee, J.; Oh, J. Deep learning-based natural language processing for detecting medical symptoms and histories in emergency patient triage. Am. J. Emerg. Med. 2024, 77, 29–38.

60. Wilkins, A. The robot doctor will see you soon. New Sci. 2023, 257, 28.

61. DeSouza, D.D.; Robin, J.; Gumus, M.; Yeung, A. Natural language processing as an emerging tool to detect late-life depression. Front. Psychiatry 2021, 12, 719125.

62. Farhat, F. ChatGPT as a complementary mental health resource: A boon or a bane. Ann. Biomed. Eng. 2023, 51, 1–4.

63. Cheng, S.W.; Chang, C.W.; Chang, W.J.; Wang, H.W.; Liang, C.S.; Kishimoto, T.; Chang, J.P.; Kuo, J.S.; Su, K.P. The now and future of ChatGPT and GPT in psychiatry. Psychiatry Clin. Neurosci. 2023, 77, 592–596.

64. Zhang, T.; Schoene, A.M.; Ji, S.; Ananiadou, S. Natural language processing applied to mental illness detection: A narrative review. NPJ Digit. Med. 2022, 5, 46.

65. Tanana, M.J.; Soma, C.S.; Kuo, P.B.; Bertagnolli, N.M.; Dembe, A.; Pace, B.T.; Srikumar, V.; Atkins, D.C.; Imel, Z.E. How do you feel? Using natural language processing to automatically rate emotion in psychotherapy. Behav. Res. Methods 2021, 53, 2069–2082.

66. Madhuri, S. Detecting emotion from natural language text using hybrid and NLP pre-trained models. Turk. J. Comput. Math. Educ. (TURCOMAT) 2021, 12, 4095–4103.

67. Pestian, J.; Nasrallah, H.; Matykiewicz, P.; Bennett, A.; Leenaars, A. Suicide note classification using natural language processing: A content analysis. Biomed. Inform. Insights 2010, 3, BII.S4706.

68. Nijhawan, T.; Attigeri, G.; Ananthakrishna, T. Stress detection using natural language processing and machine learning over social interactions. J. Big Data 2022, 9, 33.

69. May, R.; Denecke, K. Security, privacy, and healthcare-related conversational agents: A scoping review. Inform. Health Soc. Care 2022, 47, 194–210.

70. Li, J. Security Implications of AI Chatbots in Health Care. J. Med. Internet Res. 2023, 25, e47551.

71. 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. 2021, 33, e6426.

72. Oca, M.C.; Meller, L.; Wilson, K.; Parikh, A.O.; McCoy, A.; Chang, J.; Sudharshan, R.; Gupta, S.; Zhang-Nunes, S. Bias and inaccuracy in AI chatbot ophthalmologist recommendations. Cureus 2023, 15, e45911.

73. Jin, E.; Eastin, M. Gender Bias in Virtual Doctor Interactions: Gender Matching Effects of Chatbots and Users on Communication Satisfactions and Future Intentions to Use the Chatbot. Int. J. Hum.–Comput. Interact. 2023, 39, 1–13.

74. Kim, J.; Cai, Z.R.; Chen, M.L.; Simard, J.F.; Linos, E. Assessing Biases in Medical Decisions via Clinician and AI Chatbot Responses to Patient Vignettes. JAMA Netw. Open 2023, 6, e2338050.

75. Pearman, S.; Young, E.; Cranor, L.F. User-friendly yet rarely read: A case study on the redesign of an online HIPAA authorization. Proc. Priv. Enhancing Technol. 2022, 2022, 558–581.

76. Ebers, M.; Hoch, V.R.S.; Rosenkranz, F.; Ruschemeier, H.; Steinrötter, B. The European Commission's proposal for an artificial intelligence act—A critical assessment by members of the robotics and AI law society (rails). J 2021, 4, 589–603.

77. Schmidlen, T.; Schwartz, M.; DiLoreto, K.; Kirchner, H.L.; Sturm, A.C. Patient assessment of chatbots for the scalable delivery of genetic counseling. J. Genet. Couns. 2019, 28, 1166–1177.