iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VIII **Month of publication:** Aug 2023

DOI: https://doi.org/10.22214/ijraset.2023.55573

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com

Healthcare: A Growing Role for Large Language Models and Generative AI

Saurabh Pahune[1], Noopur Rewatkar[2]
Cardinal Health, Dublin OH 43017, USA
Christian Brothers University, Memphis TN 38104, USA

Abstract: ***Large language models and generative artificial intelligence (GAI) have recently demonstrated significant promise for revolutionizing a range of industries, including healthcare. The paper investigates how these cutting-edge AI developments are transforming healthcare applications. We focus on how big language models, like GPT-3 (generative pretrained transformer), Visual ChatGPT and generative AI, such Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), can be applied to solve important problems in the healthcare sector. Medical text analysis is one of the main uses of massive language models in the visual ChatGPT healthcare industry. These models have impressive natural language processing abilities that make it possible to effectively extract important information from electronic health records (EHRs), biomedical text data from large biobanks, scholarly articles and patient notes. The Biomedical Transformer Model represents a ground-breaking development in natural language processing for the biomedical field, exhibiting outstanding performance in comprehending and producing textual data. It opens up new avenues for biomedical research, diagnosis, and personalized therapy when combined with Multimodal Biomedical AI, which makes use of numerous data sources, including pictures, genomes, and clinical records.***

On the other side, generative AI has made great progress in medical picture analysis such as MRI scans and X-rays. The outstanding performance of GANs in medical picture synthesis and augmentation has helped to increase the precision and accuracy of diagnosis. Due to the issues with small and uneven medical datasets, VAEs have proven crucial in producing realistic medical images for training and research reasons. In addition to describing the various generative AI tools used in healthcare, this paper also provides an overview of multimodal medical LLMs and the biomedical transformer LLMs in the healthcare industry. Large language models and generative AI have great potential, but ethical issues and data privacy are still major problems in healthcare applications. Further, we investigate the potential role of multimodal medical LLMs as the foundation for novel assistive technologies in the fields of professional medicine, medical research, and consumer applications in the healthcare industry.

I. INTRODUCTION

A subclass of artificial intelligence (AI) known as ”Generative AI” is concerned with producing new data or content rather than just identifying patterns or making predictions based on previously collected data. Generative AI’s main objective is to create fresh, unique material that matches the data it has been trained on. Generative Adversarial Networks (GANs) are among the most well-known and effective methods used in generative AI. The field of artificial intelligence has seen a big advancement with the development of generative AI, which allows computers to display creativity and produce content that was previously only achievable through human creativity. Generative AI includes a variety of methods and models that can generate fresh information or content. Here are a few of the most common varieties of generative AI:

As per Aronson et al.[1] rapid government action and meticulous use case selection can reduce the threats generative artificial intelligence presents to health care and put it on the right track for success. Duffourc et al.[2] summarize large data sets can be used to train generative artificial intelligence (AI), a rapidly developing branch of AI, to produce lifelike images, videos, texts, sounds, 3-dimensional models, virtual environments, and even medicinal molecules. By creating visit notes, treatment codes, and medical summaries, generative AI has the potential to reduce the long-lamented burden of medical paperwork. By Barclay et al.[3] have mentioned that according to the UK government, 86 AI solutions have received a total investment of £123 million across three rounds of awards, helping to support over 300,000 patients and improve their care and treatment for ailments like cancer, heart disease, diabetes, mental illness, and neurological disorders.

As per Wang et al.[4] the future of healthcare could be altered thanks to developments in parallel computing and graphics processing unit (GPU) programming.

©IJRASET: All Rights are Reserved

However, real-world implementations will be made easier by utilizing already developed large-scale AI's like GPT-4 and Med-PaLM and incorporating them into multiagent models (like Visual-ChatGPT). The goal of their analysis is to increase awareness of the potential uses of these models in healthcare. Peikos et al.[5] choose an automated method that uses the large language model ChatGPT to pull out patient-related data from clinical notes that aren't structured and create search queries to find clinical trials that might be eligible. Mesk et al.[6] has a detailed orientation as LLMs have been used in a variety of ways, such as facilitating clinical documentation, producing discharge summaries, generating clinic, operation, and procedure notes, summing up research papers, and obtaining insurance pre-authorization. On the basis of patient records, imaging studies, and laboratory findings, LLMs can also help doctors make diagnoses and offer potential courses of treatment. At the same time, by receiving a personalized assessment of their data, symptoms, and worries, patients may become more independent than with earlier search techniques. Eysenbach et al.[7] came up with some suggestions for using chatbots in medical education. It also demonstrated its capacity to create quizzes and a virtual patient simulation for medical students, as well as to criticize simulated doctor-patient interactions and attempts to summarize research articles (which turned out to be made up), comment on ways to identify machine-generated text to maintain academic integrity and create a curriculum for health professionals to learn about artificial intelligence (AI). Nichol et al.[8] has discovered a model called GLIDE which consists of photorealistic image generation and editing using text-guided diffusion models.

Generative AI is an intriguing area of study that focuses on developing AI models that can produce original content, such as writing, graphics, and more. Generative models come in a variety of forms, including large models, large language models, multimodal models, foundation models, and others. Kuzlu et al.[9] have mentioned the fields such as finance, education, marketing, and healthcare are just a few of the industries that are being transformed by generative artificial intelligence (GAI). GAI has the potential to change a number of fields, particularly healthcare, including medical imaging, drug discovery, patient care, and treatment planning. Key parties who will gain from these developments consist of medical facilities, hospitals, and clinics. Liu et al.[10] study presents a thorough review of generative models for three-dimensional (3D) volumes, with a special emphasis on the brain and heart. Unconditional synthesis, classification, conditional synthesis, segmentation, denoising, detection, and registration are included in a brand-new, in-depth taxonomy of unconditional and conditional generative models that is proposed to cover a variety of medical tasks for the heart and brain. Pahune et al.[11] mentioned a strong emphasis on current advancements and initiatives made for various LLM types, including task-based financial LLMs, multilingual language LLMs, biomedical and clinical LLMs, visual language LLMs, and code language models.

LLMs have been used in a variety of ways, such as facilitating clinical documentation, producing discharge summaries, generating clinic, operation, and procedure notes, summing up research papers, obtaining insurance pre-authorization, and acting as chatbots to assist patients with their specific questions and concerns. Additionally, LLMs can help doctors make diagnoses based on lab data, pictures, and medical records, and they can also recommend treatments or care strategies. Patients may also be able to become more independent than they could with earlier search techniques by acquiring. Harrer et al.[12] demonstrate that Large Language Models (LLMs) are a crucial part of generative artificial intelligence (AI) applications that produce new content in response to textual instructions, including text, images, audio, code, and videos. Such generative AI applications will remain a party trick with significant potential for spreading false information or damaging and erroneous content at an unprecedented scale without human oversight, guidance, and responsible design and operation. Kuzlu et al.[9] described the rise of generative artificial intelligence is revolutionizing a number of industries, such as banking, education, marketing, and healthcare. GAI has the potential to change a number of areas, particularly in healthcare, including medical imaging, drug discovery, patient care, and treatment planning. Sezgin et al.[13] mentioned the use of artificial intelligence (AI) in clinical practice has grown, and it is clearly improving patient outcomes through better treatment planning, more accurate diagnosis, and improved patient outcomes. The discussions about AI's possible effects on the healthcare business, particularly about the role of healthcare professionals, have been reignited as a result of the rapid progress of AI, particularly generative AI and large language models (LLMs).

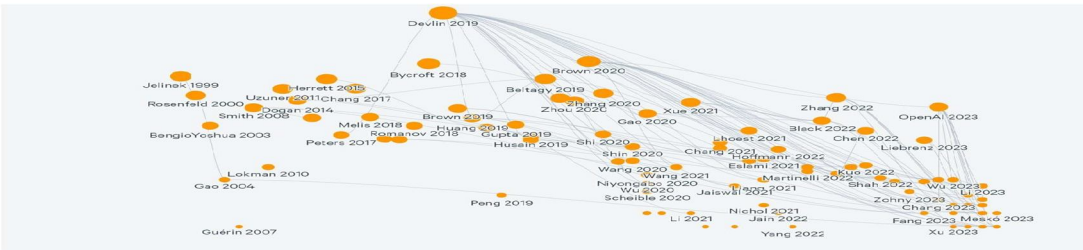


Figure 1: Several related studies on generative AI in healthcare domain

Amazon's new medical transcription service bolsters voice-to-text bid[14], It is intended to translate clinician and patient speech, such as that from physician-dictated notes, drug safety monitoring, telemedicine consultations, or physician-patient discussions, into text. The popular generative AI image models are DALL-E 2[15] (generating high-quality images based on textual prompts), GLIDE[8](focuses on image generation), BiomedGPT[16] (designed for biomedical applications), ChatGPT (primarily focused on generating text-based responses, it can also generate simple images based on textual prompts) and others. These generative AI models have drawn attention for their capacity to produce imaginative and realistic visuals in response to textual cues. They can be used in a variety of fields, including as concept exploration, visualization, and creative endeavors. Beyond the ones stated here, there are additional generative AI paradigms and models, each with unique advantages and uses. In the field of medicine, generative AI enables clinicians to copy patient data and automate form-filling procedures. For documentation tasks, it can also be connected with an EHR. Generative AI is used in healthcare to research concepts. For example, ChatGPT is a helpful resource for coming up with ideas. Users can ask questions or enter a topic of their choice to receive suggestions immediately. Synthetic images, movies, and audio are produced in the healthcare industry using generative AI. The indistinguishability of AI-generated content from actual photos can be deceptive. In recent experiments, the Multimodal Large Language Model (MLLM), which is based on the potent LLM, demonstrated extraordinary emergent abilities like the ability to create poems based on images[17]. Gong et al.[18] offer the MultiModal-GPT vision and language model for conducting multi-round conversations with people. MultiModal-GPT can respond to a variety of human commands, including creating a descriptive caption, counting the number of things that are interesting, and responding to broad questions from the user. Li et al.[19] BLIP-2 (focuses on vision-language pre-training), a general and effective pretraining technique, is suggested. It bootstraps vision-language pretraining from commercially available frozen pre-trained image encoders and frozen big language models.

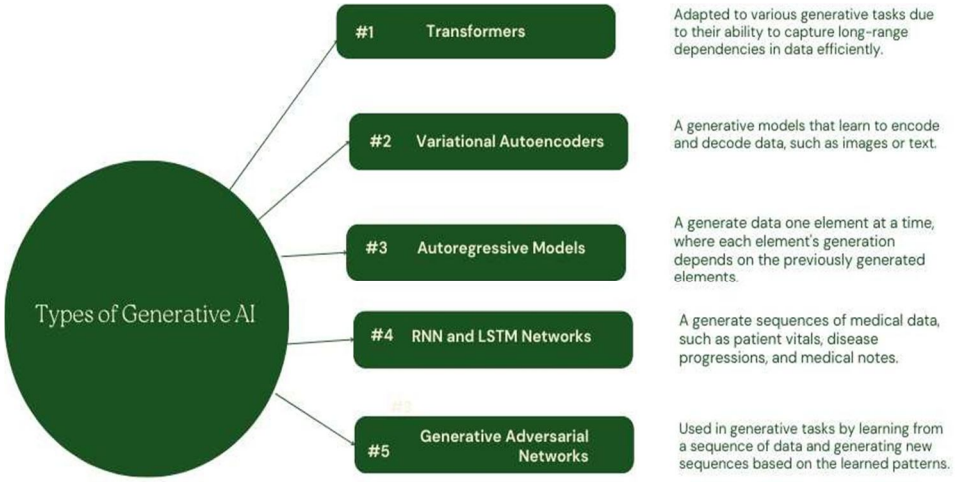


Figure 2: Types of generative AI

II. OVERVIEW OF GENERATIVE ARTIFICIAL INTELLIGENCE IN HEALTHCARE

By giving doctors and other healthcare professionals strong tools for analyzing medical data and developing more precise diagnoses and individualized treatment regimens, generative AI has the potential to revolutionize the healthcare sector. To learn from a vast quantity of data and produce new material that is similar to the input, generative AI algorithms use deep learning techniques and machine learning models. In general, generative AI has the potential to revolutionize the healthcare sector by offering strong tools for assessing medical data, producing more precise diagnoses, and developing individualized treatment regimens. To guarantee this technology's safe and effective usage in healthcare, it is crucial to address the issues and hazards related to it.

A. Clinical and Biomedical Transformer Model in Healthcare

An artificial intelligence system called a clinical and biomedical transformer model was developed specifically for the processing and analysis of clinical and biomedical text data. These transformer models utilize the transformer's architecture, which has proven successful in tasks requiring natural language processing. The large-scale datasets utilized to train the clinical and biomedical transformer models comprise clinical notes, electronic health records, research publications, and other applicable sources of clinical and biomedical data. These models learn the precise terminologies, phrasings, and concepts employed in the medical industry.

Insightful information extraction, text classification, entity recognition, relation extraction, question answering, and other tasks specific to the clinical and biomedical domain are the core objectives of clinical and biomedical transformer models. They may assist healthcare professionals with a variety of tasks, including automated documentation, information retrieval, patient risk assessment, and clinical decision support.

Santosh et al.[42] propose PathologyBERT - a pre-trained masked language model which was trained on 347,173 histopathology specimen reports and publicly released in the Huggingface1repository. Comprehensive experiments demonstrate that pre-training of the transformer model on pathology corpora yields performance improvements in Natural Language Understanding (NLU) and Breast Cancer Diagnose Classification when compared to nonspecific language models.

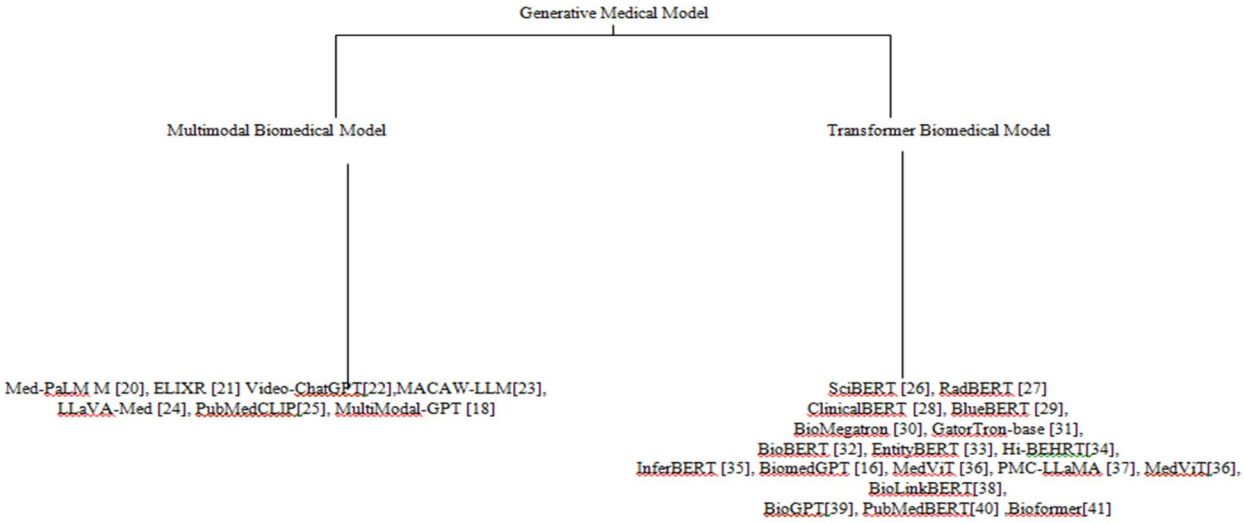


Figure 3: Types of generative models in healthcare

Jaiswal et al.[43] introduce RadBERT-CL which is"Factually-Aware Contrastive Learning For Radiology Report Classification". Also, show that the representations learned by RadBERT-CL can capture critical medical information in the latent space. Gu et al.[40] accelerate research in biomedical and released state-of-the-art pre-trained and task-specific models for the community, and created a leaderboard featuring BLURB benchmark (Biomedical Language Understanding & Reasoning Benchmark). BioGPT is a generative pre-trained transformer for mining and producing biomedical text[39].

The author challenges, the major advantage of domain-specific pretraining from scratch stems from having an in-domain vocabulary. Peng et al.[29] introduce a collection of resources for evaluating and analyzing biomedical natural language representation models. Beltagy et al.[26] SCIBERT leverages unsupervised pretraining on a large multi-domain corpus of scientific publications to improve performance on downstream scientific NLP tasks. Alsentzer et al.[28] released Clinical BERT models for clinical text one for generic clinical text and another for discharge summaries specifically. Also, demonstrate on several clinical NLP tasks the improvements this system offers over traditional BERT and BioBERT. GatorTronGPT improves biomedical natural language processing for medical research [44]. A language model particularly trained on text data from PubMed's database of biomedical and healthcare-related texts referred to as PubMedBERT[40].

Shin et al.[30] come up with BioMegatron consider a large biomedical domain language model. Which show consistent improvements on benchmarks with a larger BioMegatron model trained on a larger domain corpus, contributing to our understanding of domain language model applications. Lee et al.[32] introduce BioBERT (Bidirectional Encoder Representations from Transformers for Biomedical Text Mining), which is a domain-specific language representation model pre-trained on large-scale biomedical corpora. Li et al.[34] Present Hi-BEHRT, a hierarchical Transformer-based model that can significantly expand the receptive field of Transformers and extract associations from much longer sequences. Using multimodal large-scale linked longitudinal electronic health records. Wang et al.[35] propose an innovative causal inference model–InferBERT, by integrating the A Lite Bidirectional Encoder Representations from Transformers (ALBERT). Large language models in health care As per Anmol et al.[45] has already been proposed that LLMs, such as ChatGPT, could have applications in the field of health care due to the large volumes of free-text information available for training models. An LLM trained on more than 90 billion words of text from electronic health records (EHR)[46] Author Yang et al. develop a scratch a large clinical language model GatorTron using more than 90 billion words of text.

Existing biomedical and clinical transformer models for clinical concept extraction and medical relation such as BioBERT[32], ClinicalBERT[28], BioMegatron[30],GatorTron-base[31], GatorTron-medium[31], GatorTron-large [31]. Due to a lack of domain-specific knowledge, they frequently perform poorly in medical applications. Wu et al.[37] introduce PMC-LLaMA, an open-source language model, in this research. It was developed by fine-tuning an open-source language model on a total of 4.8 million biomedical academic publications in order to further incorporate medical information and increase its capabilities in the medical domain. MedViT is a robust vision transformer for generalized medical image classification[36]. Fang et al.[41]has an effective transformer language model for biomedical text mining is called Bioformer. It is intended to support the use of transformer-related language models in practical applications by biological researchers and healthcare providers. In general, biomedical transformer models have the potential to enhance medical picture analysis, electronic health record analysis, and text mining for medical information. They can assist researchers and medical experts in analyzing vast amounts of patient data and enhancing patient outcomes.

B. Multimodal Biomedical AI in Healthcare

LLMs must have the ability to ingest a variety of data modalities that are pertinent to an individual's health status in order to solve personalized health tasks efficiently. Belyaeva et al.[47] developed a framework (HeLM: Health Large Language Model for Multimodal Understanding) that enables LLMs to leverage high-dimensional clinical modalities to assess underlying disease risk, taking a step towards establishing multimodal LLMs for health that are based on individual-specific data. In order to perform illness risk prediction, the author built a framework (HeLM) for translating non-text data modalities into token embedding space and presenting this information as context. Yin et al.[48] has given a contemporary research hotspot is the Multimodal Large Language Model (MLLM), which uses potent Large Language Models (LLMs) as a brain to carry out multimodal tasks. The unexpected emergent skills of MLLM, such as the ability to think mathematically without using OCR and write stories based on images, are unusual in conventional approaches, providing a possible route to artificial general intelligence. Maaz et al.[22] present Video-ChatGPT. It is a multimodal paradigm that combines a visual encoder that can be used with video and an LLM. The model is capable of comprehending and producing dialogues regarding videos that resemble human speech. It has a brand-new dataset of 100,000 video-instruction pairings that was collected using a manual and semi-automated approach and was robust to label noise. A large language model (LLM) and a visual encoder tailored for video are combined in the multimodal model known as "Video-ChatGPT" to enable in-depth video comprehension. Lyu et al.[23] propose MACAW-LLM, a brand-new multi-modal LLM that smoothly combines text, audio, and visual data. The objective of Macaw-LLM is to address the difficulties of multi-modal language modeling by generating texts in natural language and effortlessly integrating several modalities. Acosta et al.[49] developed multimodal artificial intelligence solutions that capture the complexity of human health and disease has been made possible by the growing availability of biomedical data from large biobanks, electronic health records, medical imaging, wearable and ambient biosensors, and the decreasing cost of genome and microbiome sequencing. Med-PaLM Multimodal (Med-PaLM M) [20] is a Multimodal Biomedical AI from Google Research and Google DeepMind. Medicine is a multimodal discipline by nature. Clinicians routinely analyze data from a variety of sources, such as medical pictures, clinical notes, lab tests, electronic health records, genomics, and more, when giving care. Some AI systems interpret CT scans, while others analyze high-magnification pathology slides, and yet others look for uncommon genetic abnormalities. Over the past ten years or so, AI systems have attained expert-level performance on certain tasks within specific modalities. These systems frequently receive complicated data as inputs, such as photographs, and typically produce structured outputs, like dense image segmentation masks or discrete grades[50]. A substantial multimodal generative model called Med-PaLM M can understand and encode biomedical data such as clinical language, imaging, and genetics with the same model weights. ELIXR stands for "Towards a general-purpose X-ray artificial intelligence system through alignment of large language models and radiology vision encoders[21]. This paper describes training a lightweight medical information adapter that re-expresses the top layer output of the foundation model as a series of tokens in the LLM's input embeddings space. The foundation model for understanding chest X-rays has already been shown to be a good basis for building a variety of classifiers in this modality. The final system exhibits abilities for which it was not taught, such as semantic search and visual question responding, despite fine-tuning either the visual encoder nor the language model. LLaVA-Med describes a low-cost method for teaching a conversational assistant for vision-language that can respond to broad research queries about biomedical images. The main concept is to use a large-scale, comprehensive biomedical figure-caption dataset extracted from PubMed Central, use GPT-4 to self-instruct open-ended instruction-following data from the captions, and then refine a sizable general-domain vision-language model using a novel curriculum learning method[24].

Overall, by utilizing a variety of data sources and creating cutting-edge AI models, multimodal biomedical AI has enormous promise for enhancing healthcare outcomes. However, in order for multimodal AI to be fully utilized in healthcare, data issues, privacy concerns, and technical problems must be resolved.

III. GENERATIVE MEDICAL AI BENCHMARKS

A. *Multimodal LLMs data collection and benchmark*

Multimodal Large Language Models (MLLMs) are one kind of language model that combines different modalities, like text, graphics, and audio, to carry out different tasks. There are numerous benchmarks available to assess the effectiveness of MLLM[17, 48, 51]. The MME (Multimodal Model Evaluation) benchmark is one such example.

It is a thorough evaluation standard for MLLMs that incorporates information gathered from actual images[17]. The SEED-Bench is another benchmark used to compare MLLMs with generative comprehension. To evaluate the models, data with images is required[51].

1) Multi-modal pre-training for medical vision-language understanding and generation its a new benchmark, The outcomes demonstrate that the performance of VL models on MedVQA may be greatly enhanced by multi-modal pre-training and that the suggested benchmark dataset can efficiently assess the performance of vision-language (VL) models for report production and medical VQA[52].

2) BenchMD is a standard for integrated learning on sensors and medical pictures. For seven medical modalities, it combines 19 publicly accessible datasets, including 1D sensor data, 2D pictures, and 3D volumetric scans. The benchmark takes into account the difficulties of using medical data in the actual world, such as changes in picture quality, resolution, and modality[53].

3) An publically accessible database called ”A Multimodal Clinical Dataset” comprises de-identified clinical data from intensive care units. It incorporates information from various sources, such as electronic health records, medical imaging, and other modalities[54].

4) MedPerf provides a secure and privacy-preserving platform for benchmarking medical AI models, allowing researchers to evaluate the performance of their models without compromising patient privacy [55].

5) A Multimodal Generative Model for Biomedical Data is Med-PaLM M. This substantial multimodal generative model encodes and interprets biomedical data using a variety of modalities, such as clinical language, imaging, and others[56].

6) DR.BENCH is a benchmark for clinical natural language processing (cNLP) that measures the performance of cNLP models capacity for clinical diagnostic reasoning. It offers a set of six tasks[57].

7) GuacaMol: GuacaMol is a framework for comparing de novo molecular design models. It seeks to harmonize the evaluation of both conventional and neural models for producing compounds with necessary property profiles through virtual design.-make-test loops [58, 59].GuacaMol offers a thorough framework for benchmarking models for de novo molecular design, which might aid researchers in assessing and contrasting the effectiveness of various models.

8) MultiMedBench is a benchmark that covers 14 different biomedical tasks, including question answering, visual question answering, image classification, radiology report generation and summarization, and genomic variant calling. The researchers created MultiMedBench, a brand-new multimodal medical dataset comprising 14 tasks across modalities including text, medical imaging, and genomics, to facilitate the development and benchmarking of Med-PaLM M. For tasks like question answering, report creation, classification, and other therapeutically pertinent tasks, MultiMedBench has over a million instances. This thorough benchmark was essential for developing and assessing Med-PaLM M’s skills in a range of biological applications[56].

9) A sizable dataset called MIMIC-CXR contains 227,835 imaging examinations for 65,379 patients who visited the emergency room at Beth Israel Deaconess Medical Center between 2011 and 2016. One or more images, often a frontal view and a lateral view, may be included in each imaging study. The dataset has 377,110 photos in total. Studies are provided with a semi-structured free-text radiology report that describes the radiological results of the pictures and was created contemporaneously with standard clinical care by a practicing radiologist [21].

10) The UK Biobank is a sizable scientific database and research tool that houses detailed genetic and medical data from over 500,000 volunteer participants. The dataset is useful for public health research and makes it possible to make discoveries that will advance public health [60].

11) MELINDA: A Multimodal Dataset for Classification of Biomedical Experiment Methods. This dataset associates the labels of experiment procedures with the compound graphics and captions from biomedical research papers. It is intended to assist researchers in creating classification schemes for biomedical experimentation techniques[61].

12) PMC-VQA: Contains 227k VQA pairs of 149k images that cover various modalities or diseases [62].

13) BiomedGPT delivers expansive and inclusive representations of biomedical.data, outperforming the majority of preceding state-of-the-art models across five distinct tasks with 20 public datasets spanning over 15 unique biomedical modalities [16].

14) SEED-Bench consists of 19K multiple choice questions with accurate human annotations (×6 larger than existing benchmarks), which spans 12 evaluation dimensions including the comprehension of both the image and video modality[51].

15) MME benchmark consist of perception task (OCR task), coarse-grained recognition, fine-grained recognition, cognition task, commonsense reasoning, numerical calculation, text translation[17].

B. Biomedical Transformer LLMs data collection and benchmark

1) CBBQ: A Chinese Bias Benchmark Dataset Curated with Human-AI Collaboration for Large Language Models, [63]Chinese Bias Benchmark is a dataset of approximately 100K questions addressing societal prejudices and stereotypes in 14 social variables connected to Chinese culture and values. The dataset was created by human experts and generative language models.
2) HELM benchmark: The Center for Research on Foundation Models (CRFM) at Stanford University created the HELM (Holistic Evaluation of Language Models) benchmark as a thorough evaluation methodology to evaluate language model performance and capabilities [64].
3) MIMIC-CXR dataset which consists of 377, 110 chest-Xray images of 227, 827 patients along with their corresponding de-identified radiology reports [65].
4) i2b2:"Informatics for integrating biology and the bedside" is what i2b2 stands for, and it was created at Harvard Partners Healthcare. The i2b2 software can be used to query the dataset, which includes selected clinical and billing data from Penn State Health care delivery from 1997 to the present[66].
5) The MIMIC-III dataset is a sizable, publicly accessible database that contains deidentified health-related information about more than 40,000 patients who were admitted to the Beth Israel Deaconess Medical Center's critical care units between 2001 and 2012. Researchers and engineers from all around the world frequently use the dataset to advance work in clinical informatics, epidemiology, and machine learning[67, 68].
6) SciERC: The SciERC dataset is made up of 500 scientific abstracts that have been annotated with coreference clusters, relations, and scientific entities[69].
7) ACL-ARC: The ACL-ARC dataset is a collection of academic papers from the ACL Anthology, a digital repository of journal and conference papers in computer linguistics and natural language processing. 10,920 scholarly articles from the ACL Anthology are included in the ACL-ARC collection[70].
8) The BC5-chem dataset is a collection of annotated biomedical text data used for natural language processing (NLP) activities such as chemical-disease relation extraction[71].
9) NCBI-disease: The NCBI-disease dataset is a collection of annotated biomedical text data used for natural language processing (NLP) activities such as disease name recognition and concept normalization. The dataset, which is a research tool for the biomedical NLP community, consists of 793 PubMed abstracts that have been completely annotated at the mention and concept levels[72].
10) BC2GM: A collection of annotated biomedical text data used for gene mention recognition tasks in natural language processing (NLP), the BC2GM dataset is also known as the BioCreative II Gene Mention Recognition dataset[73].
11) PubMedQA: When it comes to accessing annotated biomedical text data for research and analysis, PubMedQA is a useful tool for researchers in the field of biomedical informatics[74].

IV. GENERATIVE AI TOOLS IN HEALTHCARE

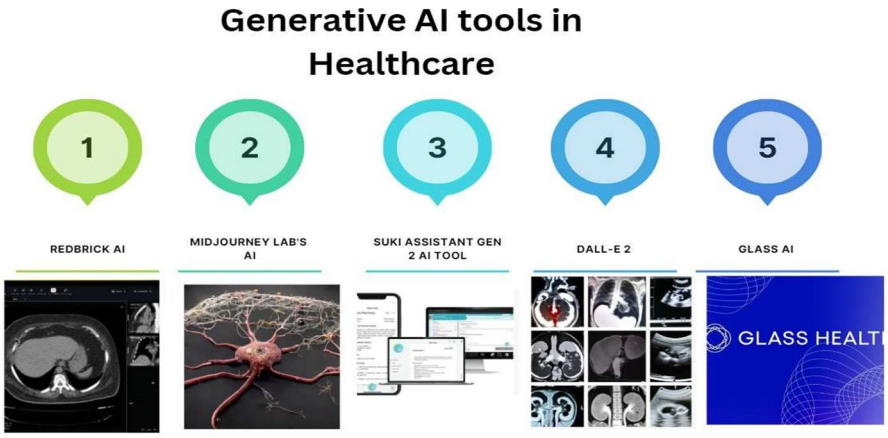


Figure 4: Several generative AI tools in healthcare

1) The Fast Automated Segmentation Tool (F.A.S.T.) from Redbrick AI is intended to help medical practitioners annotate and segment several kinds of medical pictures, including CT scans, MRI images, and ultrasound images. Without the need for extra input, it can segment photos and adapt to new image kinds[75].
2) By listening to exchanges between clinicians and patients, Suki Assistant Gen 2 is an AI tool that automates the writing of clinical notes. Major popular electronic health record (EHR) systems are compatible with it, and no human oversight is necessary[76]. Suki Assistant Gen 2 employs generative AI to receive orders and dictation, write notes ambiently, and respond to queries. It streamlines workflows by making coding easier and integrating with well-known electronic health record (EHR) systems. It is inexpensive, safe, and compliant.
3) With the use of Gridspace, an enterprise solution, parts of patient outreach can be automated by taking calls from patients, providing information, and handling simple administrative duties[77]. A business that specializes in data analytics and speech technology is called Gridspace. They have a diversified group of talented programmers who are driven to create ground-breaking solutions in the area of voice technology. Gridspace focuses on producing valuable data and service metrics from streaming conversational speech audio.
4) A technology called DALL-E 2 uses text instructions to generate artificial medical images in any modality. Since they don't contain any confidential information about actual patients, these visuals are lifelike but not real[78]. DALL-E 2 for AI-driven image generation in radiology. The study investigated DALL-E 2's potential for application in radiography and whether it has learned appropriate representations of medical pictures.
5) ChatGPT is a tool that simulates sophisticated conversation using generative AI. As a chatbot, it can comprehend user commands and respond to them, produce text that sounds like human speech in a variety of languages, and modify its tone according to the information it receives.
6) Doctors may learn, organize, and curate medical knowledge with the aid of Glass AI. It is a test instrument that can assist physicians in developing a list of potential diagnoses and a treatment strategy. It is intended for a clinical audience rather than ordinary internet users[79]. It has capabilities for developing differential diagnoses and clinical plans based on diagnostic problem representations and is specially designed to fit the way doctors learn. To produce precise therapeutic plans and diagnoses, Glass AI blends a large language model (LLM) with a clinical knowledge library.
7) An AI program called Google Bard can produce words and develop new content based on input. It has potential in the healthcare industry because patients might get assistance from it around-the-clock and with ongoing care. Google Bard can help clinicians by offering round-the-clock patient care, responding to inquiries, making diagnoses, and assisting with treatment regimens (with or without the Med-PaLM 2 augmentation). It might also be utilized to offer patients follow-up treatment, assistance with inquiries, or guidance outside regular business hours [80].

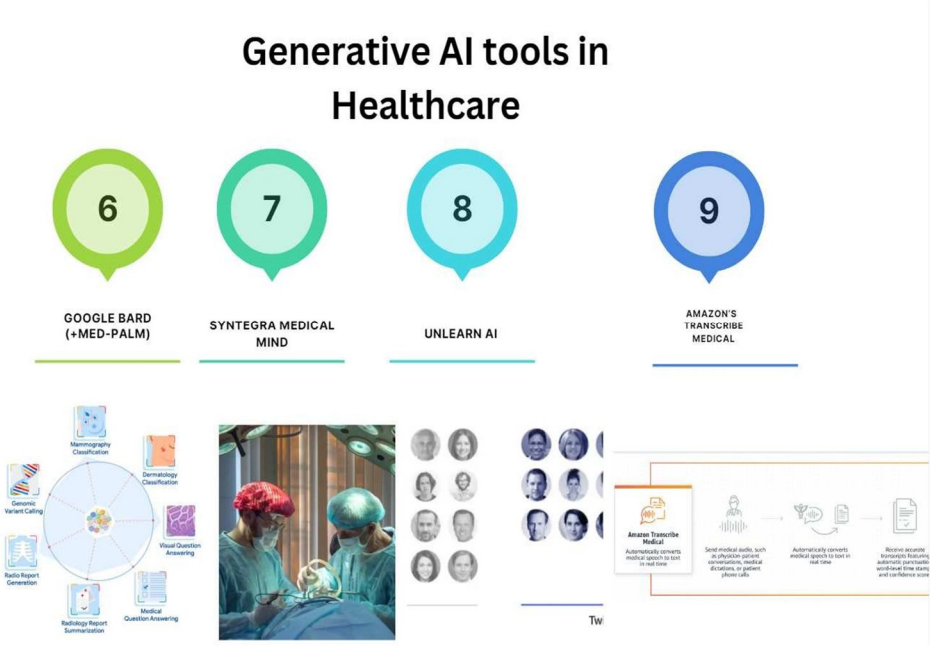


Figure 5: Several generative AI tools in healthcare

8) Regard is a tool that aids in clinical decision-making by assisting physicians in the analysis of patient data. It gathers insights from clinical data that is not structured in order to swiftly deliver pertinent information. Regard serves as an AI co-pilot integrated with the electronic medical record (EMR) system, offering diagnoses and creating clinical notes. It provides excellent patient care, optimizes clinical practice, and streamlines administrative processes[81].

9) The Ellen AI algorithm improves the conversation functionality of other Generative AI technologies, such as chatGPT. On top of the interactions with generative AI chatbots, it develops the text-to-speech voice interaction layer. It can be applied in healthcare settings to give patients, close friends, and unofficial caregivers auditory explanations to support treatment[82].

10) The AI-powered image generating platform from Midjourney Labs makes it simple to quickly produce high-quality visuals from text descriptions. This is helpful for doctors to create visual aids, teaching materials, and to improve presentations[83].

11) With the help of the automatic speech recognition (ASR) service Amazon Transcribe Medical, can quickly add medical speech-to-text functionality to your voice-activated applications. The basis for a patient's diagnosis, treatment strategy, and clinical recording procedure are conversations between health care professionals and the patient. The accuracy of this information is essential. However, proper medical transcriptions require scribes and dictation recorders, which are pricy, time-consuming, and inconvenient for patients. Although several firms employ the current medical transcribing software, they believe it to be unreliable and inefficient [84].

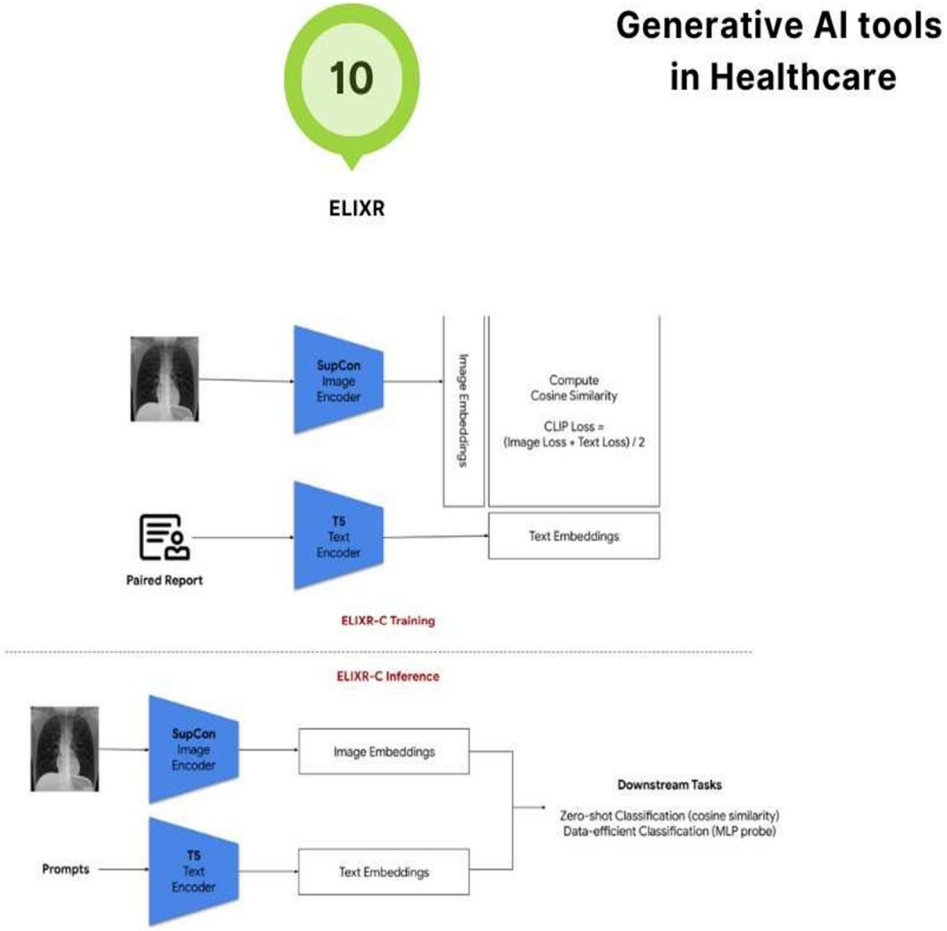


Figure 6: Several generative AI tools in healthcare

12) ELIXR: Toward a broad objective X-ray artificial intelligence system using radiological vision encoders and massive language models[21]. Google Health will create generative AI models for medical imaging that are compact and multimodal. Xu et al.[21] are able to complete formerly challenging tasks by "grafting" language-aligned visual encoders onto a predetermined LLM using a technique known as ELIXR. Along with surprising results in disease classification, ELIXR also exhibits abilities for which it was not specifically trained, such as complex natural language semantic search within Chest X-rays(CXRs), responding to visual questions, and even serving as a checker for the accuracy of radiology reports.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. CHALLENGES AND FUTURE DIRECTION

Although generative AI has great potential for the healthcare industry, there are numerous obstacles and ethical issues that need to be resolved:

1) Data Privacy and Security: Data security and privacy are important considerations when using generative models for healthcare because patient privacy and regulatory compliance are at stake and raising questions about their compliance with existing data privacy legislation, such as the General Data Protection Regulation (GDPR).
2) Interpretability: Many generative models are difficult to read since they are sophisticated and ”black boxes,” which might be crucial in healthcare settings.
3) Data Bias: Generative models are susceptible to picking up biases from the training set, which can result in biased generation. To minimize this problem, careful curation and preprocessing of the data are necessary.
4) Regulatory Permission: To ensure safety and efficacy, the application of generative AI in healthcare may need regulatory permission.
5) Privacy and Security Risks: AI systems have the ability to gather and handle sensitive personal health data, which raises questions about data security and privacy. Healthcare providers must make sure that patient data is stored securely and that only authorized people are permitted access to it[85].
6) Because of their biases, hallucinations, and untruth imitation, AI chatbots are not yet ready for clinical application[86].
7) LLMs should be viewed as imperfect tools that have the potential to significantly increase workflow efficiency but require strict human oversight and intervention at all operational interfaces, input, and output, using the adage ”garbage-in, garbage-out” as a yardstick to measure promptness and response quality, relevance, and appropriateness[12].
8) The privacy and security of ePHI(Protected Health Information)is the main issue facing generative AI. The use of generative AI can analyze data, provide quick responses, and simplify laborious patient documentation tasks. During documentation, it has access to vital patient data and saves all queries made to it. As a result, patient data security and privacy are a serious concern. Additionally, generative AI is subject to bias and prejudice, particularly if it has been trained with care data that is not representative of the population it is intended to serve. As a result, the diagnosis and treatment are incorrect or unfair.
9) The legal issues raised by the usage of LLMs in law require additional study. This entails creating techniques to reduce the likelihood of biases in LLMs and make sure they produce results that are clear and easy to understand. To further increase the precision and efficiency of LLMs in legal duties, specialized data resources and technologies must be created. To guarantee that LLMs are integrated in a responsible and ethical manner, there is also a need to create criteria and norms for their use in the legal field. With these further initiatives, the integration of LLMs into the legal profession holds enormous promise for enhancing legal procedures and access to justice[87].
10) Experiments have shown that BiomedGPT has a number of drawbacks. The model’s sensitivity to instructions is a key issue. There are times when the model misunderstands the instructions and generates dire forecasts, even generating unrelated data types for a VQA work, including picture codes. Increasing the variety of excellent instruction sets during pretraining could be a simple option. Additionally, need to look into ways to balance the diversity of the data[16].
11) Future advancements and content creation applications of generative AI will probably have a big impact on the content creation sector as it develops and gets more advanced. It might result in more automation, more individualized content, and fresh chances for innovation and creativity.
12) The future of generative AI can help with areas like clinical trials, personalized medicine, drug iscovery, natural language processing/understanding, medical imaging, virtual assistants, illness detection and screening, Generative models on medical conversation tasks, Voice generation in healthcare, Video generation in healthcare, Image synthesis and manipulation.

VI. ETHICAL CONCERNS OF USING GENERATIVE AI IN HEALTHCARE

1) Synthetic images, movies, and audio are produced in the healthcare industry using generative AI. Due to their similarity to real photographs, AI-generated content presents ethical challenges. Real healthcare data can be deceived and manipulated by it. People can be harassed or defamed using fake photos or recordings. Additionally, patients ask questions, engage in conversation, and learn more about their medical conditions using generative AI techniques. Because AI may have trouble keeping up with the most recent data, users of generative AI technology must evaluate the correctness and veracity of the information generated. Patients may be misled by inaccurate information, which could be harmful to their health.
2) Who is liable for the effects if Generative AI is used to produce inaccurate or damaging content? It can be difficult to place blame for any undesirable effects because there is no human involvement in the content generation process.

3) The emergence of AI technology that can produce text that ignores plagiarism checks and can appear to have been written by a human author raises significant ethical questions in the field of medicine[88].

4) The potential for generating false information or fake news is one of the most important ethical implications of utilizing generative AI in content development. Generative AI presents a compelling option for individuals trying to disseminate fraudulent information for their own gain because it may produce realistic-looking content that is challenging to identify as fake.

5) Consequences of utilizing generative AI in content generation in terms of copyright. Who is the rightful owner of AI-produced content? Who trained the algorithm the person or the AI itself? These issues remain unresolved, and there is a chance that using generative artificial intelligence will result in legal challenges involving intellectual property rights.

6) If GAI hurts patients, produces poor material, or is difficult to integrate into a clinic's workflow, clinicians will reject it. As a result, we ought to concentrate initially on straightforward, lower-risk domains where GAI applications are simpler to validate; several use cases that concentrate on reducing clinician burden fall into this category[1].

VII. ACRONYMS USED IN THIS REVIEW AND THEIR FULL FORMS

AI: Artificial Intelligence
GAN: Generative Adversarial Networks GAI: Generative Artificial Intelligence ML: Machine Learning
BLURB: Biomedical Language Understanding Reasoning Benchmark EHR: Electronic Health Records
NLU: Natural Language Understanding VAE: Variational Autoencoders
ASR: Automatic Speech Recognition
ALBERT: A Lite BERT for Self-supervised Learning of Language Representations EMR: Electronic Medical Record
BERT: Bidirectional Encoder Representations from Transformers CXR: Chest X-rays
HeLM: Health Large Language Model for Multimodal Understanding GPT: Generative Pretrained Transformer
GDPR: General Data Protection Regulation PHI: Protected Health Information
MLLM: Multimodal Large Language Model
LLaVA-Med: A Large Language and Vision Assistant for BioMedicine FAST: Fast Automated Segmentation Tool
OCR: Optical Character Recognition MME: Multimodal Model Evaluation VL: Vision Language
GPU: Graphics Processing Unit CT: Computerized Tomography MRI: Magnetic Resonance Imaging EHR: Electronic Health Record
i2b2: Informatics for Integrating Biology and the Bedside BC2GM: BioCreative II Gene Mention

VIII. CONCLUSION

The significance of generative AI and large language models in altering healthcare practices is highlighted by this study. These cutting-edge AI innovations provide never-before-seen opportunities for better healthcare results and improved patient experiences, which have the potential to fundamentally impact medical research, diagnosis, and patient care. However, close collaboration between AI scientists, healthcare professionals, and legislators is necessary to overcome the ethical and legal challenges associated with the integration of generative AI in healthcare.

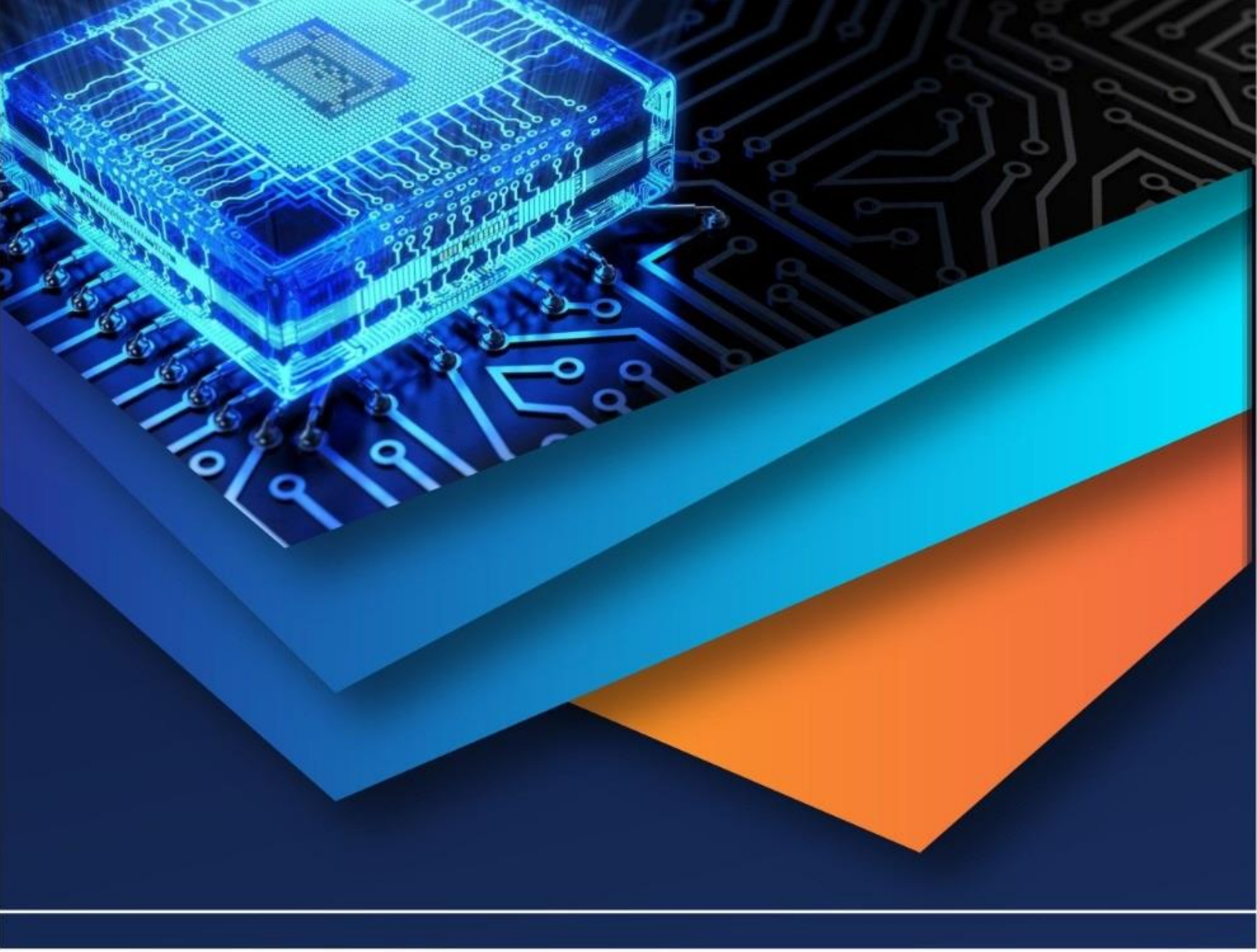
REFERENCES

[1] Samuel Aronson, Ted W Lieu, and Benjamin M Scirica. Getting generative ai right. NEJM Catalyst Innovations in Care Delivery, 4(3), 2023.
[2] Mindy Duffourc and Sara Gerke. Generative ai in health care and liability risks for physicians and safety concerns for patients. JAMA, 2023.
[3] Department of Health and NHS England Social Care. Deepcausalpv-master. https://www.gov.uk/government/news/ thousands-of-patients-to-benefit-from-quicker-diagnosis-more-accurate-tests-from-ground-breaking-ai-research, 2023. [Online; Accessed 06-17-2023].
[4] Ding-Qiao Wang, Long-Yu Feng, Jin-Guo Ye, Jin-Gen Zou, and Ying-Feng Zheng. Accelerating the integration of chatgpt and other large-scale ai models into biomedical research and healthcare. MedComm–Future Medicine, 2(2):e43, 2023.
[5] Georgios Peikos, Symeon Symeonidis, Pranav Kasela, and Gabriella Pasi. Utilizing chatgpt to enhance clinical trial enrollment. arXiv preprint arXiv:2306.02077, 2023.
[6] Bertalan Mesko´ and Eric J Topol. The imperative for regulatory oversight of large language models (or generative ai) in healthcare. NPJ Digital Medicine, 6(1):120, 2023.
[7] Gunther Eysenbach et al. The role of chatgpt, generative language models, and artificial intelligence in medical education: a conversation with chatgpt and a call for papers. JMIR Medical Education, 9(1):e46885, 2023.
[8] Alex Nichol, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew, Ilya Sutskever, and Mark Chen. Glide: Towards photorealistic image generation and editing with text-guided diffusion models. arXiv preprint arXiv:2112.10741, 2021.
[9] Murat Kuzlu, Zhenxin Xiao, Salih Sarp, Ferhat Ozgur Catak, Necip Gurler, and Ozgur Guler. The rise of generative artificial intelligence in healthcare. In 2023 12th Mediterranean Conference on Embedded Computing (MECO), pages 1–4. IEEE, 2023.
[10] Yanbin Liu, Girish Dwivedi, Farid Boussaid, and Mohammed Bennamoun. 3d brain and heart volume generative models: A survey. arXiv preprint arXiv:2210.05952, 2022.

[11] Saurabh Pahune and Manoj Chandrasekharan. Several categories of large language models (llms): A short survey. arXiv preprint arXiv:2307.10188, 2023.
[12] Stefan Harrer. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. EBioMedicine, 90, 2023.
[13] Emre Sezgin, Joseph Sirrianni, and Simon L Linwood. Operationalizing and implementing pretrained, large artificial intelligence linguistic models in the us health care system: outlook of generative pretrained transformer 3 (gpt-3) as a service model. JMIR medical informatics, 10(2):e32875, 2022.
[14] Rebecca Pifer. Med-palm m is a multimodal biomedical ai from google research and google deepmind. https://www.healthcaredive.com/news/ amazons-new-medical-transcription-service-bolsters-voice-to-text-bid/568245/, 2023. [Online; Accessed 07-29-2023].
[15] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. In International Conference on Machine Learning, pages 8821–8831. PMLR, 2021.
[16] Kai Zhang, Jun Yu, Zhiling Yan, Yixin Liu, Eashan Adhikarla, Sunyang Fu, Xun Chen, Chen Chen, Yuyin Zhou, Xiang Li, Lifang He, Brian D. Davison, Quanzheng Li, Yong Chen, Hongfang Liu, and Lichao Sun. Biomedgpt: A unified and generalist biomedical generative pre-trained transformer for vision, language, and multimodal tasks, 2023.
[17] Chaoyou Fu, Peixian Chen, Yunhang Shen, Yulei Qin, Mengdan Zhang, Xu Lin, Zhenyu Qiu, Wei Lin, Jinrui Yang, Xiawu Zheng, et al. Mme: A comprehensive evaluation benchmark for multimodal large language models. arXiv preprint arXiv:2306.13394, 2023.
[18] Tao Gong, Chengqi Lyu, Shilong Zhang, Yudong Wang, Miao Zheng, Qian Zhao, Kuikun Liu, Wenwei Zhang, Ping Luo, and Kai Chen. Multimodal-gpt: A vision and language model for dialogue with humans. arXiv preprint arXiv:2305.04790, 2023.
[19] Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. arXiv preprint arXiv:2301.12597, 2023.
[20] Tao Tu, Shekoofeh Azizi, Danny Driess, Mike Schaekermann, Mohamed Amin, Pi-Chuan Chang, Andrew Carroll, Chuck Lau, Ryutaro Tanno, Ira Ktena, et al. Towards generalist biomedical ai. arXiv preprint arXiv:2307.14334, 2023.
[21] Shawn Xu, Lin Yang, Christopher Kelly, Marcin Sieniek, Timo Kohlberger, Martin Ma, Wei-Hung Weng, Attila Kiraly, Sahar Kazemzadeh, Zakkai Melamed, Jungyeon Park, Patricia Strachan, Yun Liu, Chuck Lau, Preeti Singh, Christina Chen, Mozziyar Etemadi, Sreenivasa Raju Kalidindi, Yossi Matias, Katherine Chou, Greg S. Corrado, Shravya Shetty, Daniel Tse, Shruthi Prabhakara, Daniel Golden, Rory Pilgrim, Krish Eswaran, and Andrew Sellergren. Elixr: Towards a general purpose x-ray artificial intelligence system through alignment of large language models and radiology vision encoders, 2023.
[22] Muhammad Maaz, Hanoona Rasheed, Salman Khan, and Fahad Shahbaz Khan. Video-chatgpt: Towards detailed video understanding via large vision and language models. arXiv preprint arXiv:2306.05424, 2023.
[23] Chenyang Lyu, Minghao Wu, Longyue Wang, Xinting Huang, Bingshuai Liu, Zefeng Du, Shuming Shi, and Zhaopeng Tu. Macaw-llm: Multi-modal language modeling with image, audio, video, and text integration. arXiv preprint arXiv:2306.09093, 2023.
[24] Chunyuan Li, Cliff Wong, Sheng Zhang, Naoto Usuyama, Haotian Liu, Jianwei Yang, Tristan Naumann, Hoifung Poon, and Jianfeng Gao. Llava-med: Training a large language-and-vision assistant for biomedicine in one day, 2023.
[25] Sedigheh Eslami, Gerard de Melo, and Christoph Meinel. Does clip benefit visual question answering in the medical domain as much as it does in the general domain?, 2021.
[26] Iz Beltagy, Kyle Lo, and Arman Cohan. Scibert: A pretrained language model for scientific text. arXiv preprint arXiv:1903.10676, 2019.
[27] An Yan, Julian McAuley, Xing Lu, Jiang Du, Eric Y Chang, Amilcare Gentili, and Chun-Nan Hsu. Radbert: Adapting transformer-based language models to radiology. Radiology: Artificial Intelligence, 4(4):e210258, 2022.
[28] Emily Alsentzer, John R Murphy, Willie Boag, Wei-Hung Weng, Di Jin, Tristan Naumann, and Matthew McDermott. Publicly available clinical bert embeddings. arXiv preprint arXiv:1904.03323, 2019.
[29] Yifan Peng, Shankai Yan, and Zhiyong Lu. Transfer learning in biomedical natural language processing: an evaluation of bert and elmo on ten benchmarking datasets. arXiv preprint arXiv:1906.05474, 2019.
[30] Hoo-Chang Shin, Yang Zhang, Evelina Bakhturina, Raul Puri, Mostofa Patwary, Mohammad Shoeybi, and Raghav Mani. Biomegatron: Larger biomedical domain language model. arXiv preprint arXiv:2010.06060, 2020.
[31] Xi Yang, Aokun Chen, Nima PourNejatian, Hoo Chang Shin, Kaleb E Smith, Christopher Parisien, Colin Compas, Cheryl Martin, Mona G Flores, Ying Zhang, et al. Gatortron: A large clinical language model to unlock patient information from unstructured electronic health records. arXiv preprint arXiv:2203.03540, 2022.
[32] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. Biobert: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 36(4):1234–1240, 2020.
[33] Chen Lin, Timothy Miller, Dmitriy Dligach, Steven Bethard, and Guergana Savova. Entitybert: Entity-centric masking strategy for model pretraining for the clinical domain. Association for Computational Linguistics (ACL), 2021.
[34] Yikuan Li, Mohammad Mamouei, Gholamreza Salimi-Khorshidi, Shishir Rao, Abdelaali Hassaine, Dexter Canoy, Thomas Lukasiewicz, and Kazem Rahimi. Hi-behrt: Hierarchical transformer-based model for accurate prediction of clinical events using multimodal longitudinal electronic health records. IEEE Journal of Biomedical and Health Informatics, 2022.
[35] Xingqiao Wang, Xiaowei Xu, Weida Tong, Ruth Roberts, and Zhichao Liu. Inferbert: a transformer-based causal inference framework for enhancing pharmacovigilance. Frontiers in Artificial Intelligence, 4:659622, 2021.
[36] Omid Nejati Manzari, Hamid Ahmadabadi, Hossein Kashiani, Shahriar B Shokouhi, and Ahmad Ayatollahi. Medvit: a robust vision transformer for generalized medical image classification. Computers in Biology and Medicine, 157:106791, 2023.
[37] Chaoyi Wu, Xiaoman Zhang, Ya Zhang, Yanfeng Wang, and Weidi Xie. Pmc-llama: Further finetuning llama on medical papers. arXiv preprint arXiv:2304.14454, 2023.
[38] Michihiro Yasunaga, Jure Leskovec, and Percy Liang. Linkbert: Pretraining language models with document links, 2022.
[39] Renqian Luo, Liai Sun, Yingce Xia, Tao Qin, Sheng Zhang, Hoifung Poon, and Tie-Yan Liu. Biogpt: generative pre-trained transformer for biomedical text generation and mining. Briefings in Bioinformatics, 23(6):bbac409, 2022.
[40] Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. Domain-specific language model pretraining for biomedical natural language processing. ACM Transactions on Computing for Healthcare (HEALTH), 3(1):1–23, 2021.
[41] Li Fang, Qingyu Chen, Chih-Hsuan Wei, Zhiyong Lu, and Kai Wang. Bioformer: an efficient transformer language model for biomedical text mining. arXiv preprint arXiv:2302.01588, 2023.

[42] Thiago Santos, Amara Tariq, Susmita Das, Kavyasree Vayalpati, Geoffrey H Smith, Hari Trivedi, and Imon Banerjee. Pathologybert–pre-trained vs. a new transformer language model for pathology domain. arXiv preprint arXiv:2205.06885, 2022.
[43] Ajay Jaiswal, Liyan Tang, Meheli Ghosh, Justin F Rousseau, Yifan Peng, and Ying Ding. Radbert-cl: factually-aware contrastive learning for radiology report classification. In Machine Learning for Health, pages 196–208. PMLR, 2021.
[44] Cheng Peng, Xi Yang, Aokun Chen, Kaleb E Smith, Nima PourNejatian, Anthony B Costa, Cheryl Martin, Mona G Flores, Ying Zhang, Tanja Magoc, et al. A study of generative large language model for medical research and healthcare. arXiv preprint arXiv:2305.13523, 2023.
[45] Anmol Arora and Ananya Arora. The promise of large language models in health care. The Lancet, 401(10377):641, 2023.
[46] Xi Yang, Aokun Chen, Nima PourNejatian, Hoo Chang Shin, Kaleb E Smith, Christopher Parisien, Colin Compas, Cheryl Martin, Anthony B Costa, Mona G Flores, et al. A large language model for electronic health records. npj Digital Medicine, 5(1):194, 2022.
[47] Anastasiya Belyaeva, Justin Cosentino, Farhad Hormozdiari, Cory Y McLean, and Nicholas A Furlotte. Multimodal llms for health grounded in individual-specific data. arXiv preprint arXiv:2307.09018, 2023.
[48] Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, and Enhong Chen. A survey on multimodal large language models. arXiv preprint arXiv:2306.13549, 2023.
[49] Juli´an N Acosta, Guido J Falcone, Pranav Rajpurkar, and Eric J Topol. Multimodal biomedical ai. Nature Medicine, 28(9):1773–1784, 2022.
[50] Greg Corrado. Multimodal medical ai. https://ai.googleblog.com/2023/08/multimodal-medical-ai.html, 2023. [Online; Accessed 07-29-2023].
[51] Bohao Li, Rui Wang, Guangzhi Wang, Yuying Ge, Yixiao Ge, and Ying Shan. Seed-bench: Benchmarking multimodal llms with generative comprehension. arXiv preprint arXiv:2307.16125, 2023.
[52] Li Xu, Bo Liu, Ameer Hamza Khan, Lu Fan, and Xiao-Ming Wu. Multi-modal pre-training for medical vision-language understanding and generation: An empirical study with a new benchmark, 2023.
[53] Kathryn Wantlin, Chenwei Wu, Shih-Cheng Huang, Oishi Banerjee, Farah Dadabhoy, Veeral Vipin Mehta, Ryan Wonhee Han, Fang Cao, Raja R. Narayan, Errol Colak, Adewole Adamson, Laura Heacock, Geoffrey H. Tison, Alex Tamkin, and Pranav Rajpurkar. Benchmd: A benchmark for unified learning on medical images and sensors, 2023.
[54] Luis R Soenksen, Yu Ma, Cynthia Zeng, Leonard Boussioux, Kimberly Villalobos Carballo, Liangyuan Na, Holly M Wiberg, Michael L Li, Ignacio Fuentes, and Dimitris Bertsimas. Integrated multimodal artificial intelligence framework for healthcare applications. NPJ digital medicine, 5(1):149, 2022.
[55] Alexandros Karargyris, Renato Umeton, Micah J Sheller, Alejandro Aristizabal, Johnu George, Anna Wuest, Sarthak Pati, Hasan Kassem, Maximilian Zenk, Ujjwal Baid, et al. Federated benchmarking of medical artificial intelligence with medperf. Nature Machine Intelligence, pages 1–12, 2023.
[56] CHRIS MCKAY. Med-palm m is a multimodal biomedical ai from google research and google deepmind. https://www.maginative.com/article/ med-palm-m-is-a-multimodal-biomedical-ai-from-google-research-and-google-deepmind/, 2023. [Online; Accessed 07-29-2023].
[57] Yanjun Gao, Dmitriy Dligach, Timothy Miller, John Caskey, Brihat Sharma, Matthew M Churpek, and Majid Afshar. Dr. bench: Diagnostic reasoning benchmark for clinical natural language processing. Journal of Biomedical Informatics, 138:104286, 2023.
[58] Nathan Brown, Marco Fiscato, Marwin HS Segler, and Alain C Vaucher. Guacamol: benchmarking models for de novo molecular design. Journal of chemical information and modeling, 59(3):1096–1108, 2019.
[59] Nathan Brown, Marco Fiscato, Marwin H.S. Segler, and Alain C. Vaucher. GuacaMol: Benchmarking models for de novo molecular design. Journal of Chemical Information and Modeling, 59(3):1096–1108, mar 2019.
[60] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O'Connell, et al. The uk biobank resource with deep phenotyping and genomic data. Nature, 562(7726):203–209, 2018.
[61] Te-Lin Wu, Shikhar Singh, Sayan Paul, Gully Burns, and Nanyun Peng. Melinda: A multimodal dataset for biomedical experiment method classification, 2020.
[62] Xiaoman Zhang, Chaoyi Wu, Ziheng Zhao, Weixiong Lin, Ya Zhang, Yanfeng Wang, and Weidi Xie. Pmc-vqa: Visual instruction tuning for medical visual question answering, 2023.
[63] Yufei Huang and Deyi Xiong. Cbbq: A chinese bias benchmark dataset curated with human-ai collaboration for large language models, 2023.
[64] Percy Liang, Rishi Bommasani, Tony Lee, Dimitris Tsipras, Dilara Soylu, Michihiro Yasunaga, Yian Zhang, Deepak Narayanan, Yuhuai Wu, Ananya Kumar, et al. Holistic evaluation of language models. arXiv preprint arXiv:2211.09110, 2022.
[65] Alistair EW Johnson, Tom J Pollard, Nathaniel R Greenbaum, Matthew P Lungren, Chih-ying Deng, Yifan Peng, Zhiyong Lu, Roger G Mark, Seth J Berkowitz, and Steven Horng. Mimic-cxr-jpg, a large publicly available database of labeled chest radiographs. arXiv preprint arXiv:1901.07042, 2019.
[66] O¨ zlem Uzuner, Yuan Luo, and Peter Szolovits. Evaluating the state-of-the-art in automatic de-identification. Journal of the American Medical Informatics Association, 14(5):550–563, 2007
[67] Alistair EW Johnson, Tom J Pollard, Lu Shen, Li-wei H Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. Mimic-iii, a freely accessible critical care database. Scientific data, 3(1):1–9, 2016.
[68] Laboratory for computational physiology. https://lcp.mit.edu/mimic, 2023. [Online; Accessed 06-17-2023].
[69] Yi Luan, Luheng He, Mari Ostendorf, and Hannaneh Hajishirzi. Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction. arXiv preprint arXiv:1808.09602, 2018.
[70] David Jurgens, Srijan Kumar, Raine Hoover, Dan McFarland, and Dan Jurafsky. Measuring the evolution of a scientific field through citation frames. Transactions of the Association for Computational Linguistics, 6:391–406, 2018.
[71] Jiao Li, Yueping Sun, Robin J Johnson, Daniela Sciaky, Chih-Hsuan Wei, Robert Leaman, Allan Peter Davis, Carolyn J Mattingly, Thomas C Wiegers, and Zhiyong Lu. Biocreative v cdr task corpus: a resource for chemical disease relation extraction. Database, 2016, 2016.
[72] Rezarta Islamaj Do˘gan, Robert Leaman, and Zhiyong Lu. Ncbi disease corpus: a resource for disease name recognition and concept normalization.Journal of biomedical informatics, 47:1–10, 2014.
[73] Larry Smith, Lorraine K Tanabe, Cheng-Ju Kuo, I Chung, Chun-Nan Hsu, Yu-Shi Lin, Roman Klinger, Christoph M Friedrich, Kuzman Ganchev, Manabu Torii, et al. Overview of biocreative ii gene mention recognition. Genome biology, 9(2):1–19, 2008.
[74] Qiao Jin, Bhuwan Dhingra, Zhengping Liu, William W Cohen, and Xinghua Lu. Pubmedqa: A dataset for biomedical research question answering.arXiv preprint arXiv:1909.06146, 2019.
[75] Boardofinnovation. https://healthcare.boardofinnovation.com/redbrick-ai-hero-f-a-s-t/, 2023. [Online; Accessed 07-29-2023].

[76] Boardofinnovation. https://healthcare.boardofinnovation.com/suki-assistant-2/, 2023. [Online; Accessed 07-29-2023].
[77] Boardofinnovation. https://healthcare.boardofinnovation.com/gridspace/, 2023. [Online; Accessed 07-29-2023].
[78] Boardofinnovation. https://healthcare.boardofinnovation.com/dall-e2/, 2023. [Online; Accessed 07-29-2023].
[79] Boardofinnovation. https://healthcare.boardofinnovation.com/glass-ai/, 2023. [Online; Accessed 07-29-2023].
[80] Board of innovation. https://healthcare.boardofinnovation.com/tools/jsf/jet-engine:cases/pagenum/2/, 2023. [Online; Accessed 07-29-2023].
[81] Boardofinnovation. https://healthcare.boardofinnovation.com/regard/, 2023. [Online; Accessed 07-29-2023].
[82] Boardofinnovation. https://healthcare.boardofinnovation.com/ellen-ai/, 2023. [Online; Accessed 07-29-2023].
[83] Boardofinnovation. https://healthcare.boardofinnovation.com/midjourney/, 2023. [Online; Accessed 07-29-2023].
[84] Boardofinnovation. https://aws.amazon.com/transcribe/medical/, 2023. [Online; Accessed 07-29-2023].
[85] Saurabh A Pahune. How does ai help in rural development in healthcare domain: A short survey. IJRASET, 2023.
[86] Joshua Au Yeung, Zeljko Kraljevic, Akish Luintel, Alfred Balston, Esther Idowu, Richard J Dobson, and James T Teo. Ai chatbots not yet ready for clinical use. Frontiers in Digital Health, 5:60, 2023.
[87] Zhongxiang Sun. A short survey of viewing large language models in legal aspect. arXiv preprint arXiv:2303.09136, 2023.
[88] Hazem Zohny, John McMillan, and Mike King. Ethics of generative ai, 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)