4 An Overview and Comparative Analysis of ChatGPT Versions: From GPT-3 to GPT-4o

Dr. Nitin[1*], & Dr. Reena Hooda[2]
[1]Department of CSE, Dr. Akhilesh Das Gupta Institute of Professional Studies, Delhi, India.
[2]DCSE, Indira Gandhi University, Meerpur-Rewari, Haryana, India.
*Corresponding Author: bansal.nitin12@gmail.com

Abstract

This chapter shows a broad general idea of the progression of conversational AI, with a precise emphasis on the evolution of OpenAI's GPT models from GPT-1 to GPT-4o. It sightsees the significant technologies that support human-like, natural dialogue, for example the components of natural language processing—understanding and generation and transformer-based neural networks. The conversation extents practical applications across different sectors like healthcare, education, customer support, and software engineering, emphasizing improvements in language fluency, contextual awareness, and multimodal interaction. A comprehensive comparison of different model versions elucidates progresses in user experience, speed, and overall proficiency. Challenges in deployment, ethical concerns, and the impending trajectory of large language models (LLMs) are also inspected, cross-industry adoption and emphasizing responsible innovation.

Keywords: Conversational AI, Large Language Models(LLMs), GPT, ChatGPT.

Introduction

Conversational artificial intelligence (AI) is a subdivision of artificial intelligence which permits machines to converse with human by means of written or oral natural language in a way that simulates human speech. By uniting numerous technologies of

AI, as well as Natural Language Processing (NLP), Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Machine Learning (ML) and Natural Language Generation (NLG), it consents systems to recognize user intent and respond in a contextual and useful way [4; 5].

Operational Structure Steps of a Conversational AI System

- **Input Generation:** The user provides written or audio input to start the interaction.
- **Input Processing:** NLP and NLU approaches are used to extract intent and semantic meaning from textual input. For voice-based inputs, ASR primarily transforms speech into text, which is consequently managed to conclude the user's intent [6].
- **Answer Evaluation:** The system constructs a pertinent and reasonable answer by using NLG which is based on the assumed intent. Text-to-speech (TTS) technology is employed [7] to deliver an audible response in voice systems.
- **Learning and Adaptation:** Reinforcement learning (by learning from user interactions) supports the system constantly achieve better and increasing the relevancy and accuracy of responses over time [2].

Chatbots and Virtual assistants are eminent instances of applications that are proficient of a extensive kind of functions, including real-time information transmission, transaction processing, customer assistance and appointment scheduling [3]. conversational AI is transfiguring various of industries by refining service accessibility and efficiency, for example banking, e-commerce, education and healthcare [1].

Enlargement of large language models (LLMs)

Large language models (LLMs), a division of generative artificial intelligence (GenAI), are prepared to understand and generate human language with contextual accuracy and coherence. Natural language processing has been significantly enriched by LLMs, such as Claude, ChatGPT, LLaMA, and Gemini which are created on deep learning architectures, which is transformer-based [8]. Their global influence is still varying, however, and their efficiency is dependent on difficulty of task, data quality and model suitability [9].

Evolution of Large Language Models (LLMs)

Large Language Models (LLMs) has a key role to shape the history of Natural Language Processing (NLP), through key turning points:

- **Semantics and Rule-Based Systems (1883–1960s):** Michel Bréal's effort on semantics in 1883 is accredited with positioning the intelligent foundation for language conception [10]. In early language systems, grammatical rules were shaped manually. for example: 1966 ELIZA program, which simulate human-like speech by using pre-set pattern-matching techniques [11].

- **Statistical Language Models (1980s–1990s):** Throughout the 1980s–90s, the research stimulated in the direction of developing methods, probabilistic models like Hidden Markov Models (HMMs) and n-gram models. By breaking from strict rules, these techniques made it potential to symbolize language by means of predictions based on probability and statistical inference [12].
- **Neural Network-Based Language Models (2000–2012):** With the help of developments in backpropagation, neural networks saw a resurgence and were eventually used in natural language processing. Bengio et al.'s 2003 introduction of the feedforward neural network language model was a noteworthy development [13]. Profounder architectures were made imaginable by the practicalities recognized during this time for learning distributed demonstrations of words.
- **Word Embeddings and Semantic Vector Space (2013–2015):** Word embeddings—vector exemplifications in which the words which are semantically comparable are positioned closer in space—were made well-known by Mikolov et al.'s 2013 Word2Vec proposal [14]. Models were capable to efficiently seizure contextual and relational semantics.
- **Attention Mechanisms and Transformers (2016–2017):** in 2016, The model's capability to deliberate on pertinent input segments during processing was upgraded with the accumulation of attention mechanisms [15]. From 2017, Vaswani et al.'s "Transformer" design striking a radical improvement in scalability and performance by leveraging self-attention to internment long-term dependences while easing parallel computing [16].
- **Emergence of GPT and Scaling of LLMs (2018–2023):** Beginning with GPT-1 in 2018, OpenAI's Generative Pre-trained Transformer (GPT) series advanced quickly, reaching GPT-4 in 2023. The scalability and generalization capabilities of LLMs across workloads were demonstrated by GPT-3, which has 175 billion parameters [17]. Google's Gemini and Anthropic's Claude are examples of parallel innovations that are adding to the expanding ecosystem of potent LLMs.
- **Democratization and Ethical Considerations (2023):** Known as the “year of AI” [18], 2023 was a watershed year in which LLMs were much more accessible. More experimentation and customisation were made possible by open-source models like LLaMA, Falcon, Vicuna, Mistral, and Gemma. Calls for regulation and responsible deployment have grown as a result of the potential and concerns brought about by this democratization, including high training costs and ethical risks [19].

- **Integration with Productivity and Development Tools (2023–Present):** Integrating LLMs into production ecosystems is a recent development. Examples of how LLMs are incorporated into practical workflows are Microsoft's Copilot (Office 365), Google's Workspace AI, GitHub Copilot, and StarCoder, which have transformed software development, documentation, and communication [20].

Role of ChatGPT in generative AI landscape

Prominent Example of Generative AI

Based on the transformer architecture presented by Vaswani et al. (2017) [16], ChatGPT, created by OpenAI, is a notable illustration of a generative AI system driven by a Large Language Model (LLM).

Human-like Text Generation: It exhibits remarkable ability to generate text that is grammatically correct, contextually relevant, and cohesive, fluently simulating human speech patterns [17].

Wide Range of Applications: ChatGPT has been effectively incorporated into a number of fields, such as:

- **Customer Service:** Automating help interactions to improve user satisfaction and response times [23].
- **Education:** Helping with academic writing, subject summarization, and one-on-one teaching [24].
- **Content Creation:** Producing emails, reports, and imaginative writing for fields such as journalism and marketing [25].
    - **Enhancing Human–Machine Interaction**: It is a conversational agent that enhances the naturalness and utility of human–AI communication due to its capacity to comprehend complex inquiries and preserve context throughout turns [22].
    - **Driving Productivity and Innovation**: ChatGPT is a tool for increasing productivity and innovation in knowledge-intensive fields by automating repetitive processes, strengthening human decision-making, and improving information accessibility [26].
    - **Catalyst for Future AI Integration**: ChatGPT is anticipated to be a key player in forming next-generation applications, such as multimodal systems and AI-augmented work environments, as generative AI technologies advance [27].

Background: GPT Architecture

Overview of the GPT Model Family

Since the early days of rule-based systems and basic chatbots, artificial intelligence has advanced significantly. At the vanguard of this development are OpenAI's Generative Pre-trained Transformer (GPT) models, which provide strong language production and understanding skills that can be used for a variety of tasks, including coding and writing support. GPT is based on the transformer architecture, which fundamentally relies on self-attention mechanisms to process each word in the context of all others within a sentence—unlike traditional sequential models that process tokens in order [16].

Transformer architecture

This design enables the model to assign varying levels of importance to each word, regardless of its position, resulting in a more nuanced and flexible understanding of language.

- **GPT Architecture**

An **embedding layer** is employed to convert the raw input—after tokenization into words or subwords—into dense vector representations. These vectors are then augmented with **positional encodings** to retain sequence information, which transformers otherwise lack inherently [29]. To reduce the risk of overfitting during training, a **dropout layer** is applied at this stage [30].

Through **multiple stacked transformer blocks**, the processed input is consequently passed. Respectively block usually contains:

- Stabilizing training by **Layer normalization**
- To capture contextual dependencies, A **multi-head self-attention mechanism**
- **Residual connections** to preserve information flow,
- And for further transformation a **feed-forward network** [31]. Within these blocks, Dropout is applied to progress generalization and diminish reliance on precise neurons throughout training [32].

In the final stage, the model output undertakes one last **layer normalization**, is passed through a **linear layer** to map it to the vocabulary space, and then a **softmax function** is functional to produce a probability distribution over conceivable succeeding tokens [33].

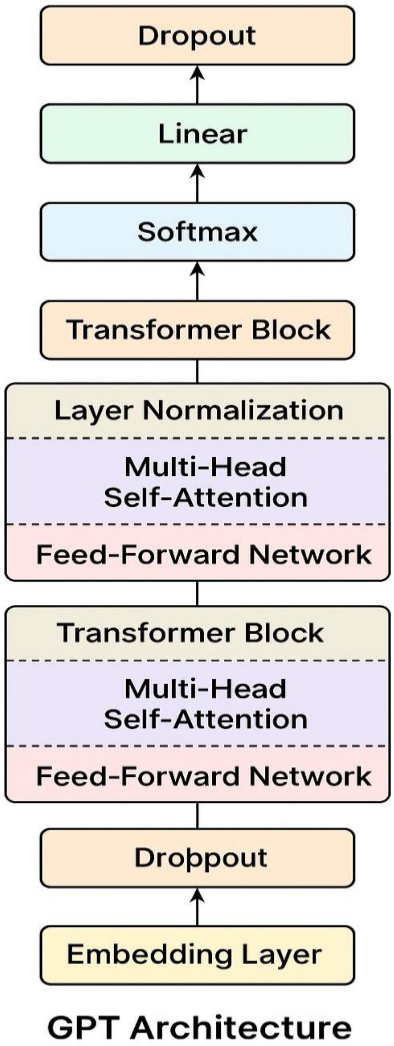

Figure 1: GPT Architecture

Pretraining and Fine-Tuning

In the development of Large Language Models (LLMs), Pre-training and fine-tuning signify two complementary and essential stages. **Pre-training** assists as the introductory learning stage, where the model is visible to a large-scale, different corpus of unlabelled text data—including articles, books, web content etc. The objective throughout this stage is to permit the model to acquire universal language demonstrations, semantic relationships, capturing syntactic structures, and contextual patterns through an extensive range of topics and linguistic styles.

Fine-tuning is an important step in acclimating pre-trained Large Language Models (LLMs) to particular responsibilities. It includes added training a general-purpose model on a slighter, domain-specific dataset, augmenting its relevance and accuracy without demanding large-scale retraining [34]. As a procedure of transfer learning, fine-tuning maintains the model's broad linguistic competence while bring into line its behaviour with task-specific goals [35].

This technique enables:

- Improved performance on niche applications by optimizing outputs for specific objectives.
- Greater alignment of responses with real-world expectations and reduced hallucination tendencies.
- Efficient model customization for domains such as healthcare, law, education, and customer support [36].

The typical fine-tuning process includes selecting a suitable pre-trained base model, preparing a task-relevant dataset, choosing an appropriate tuning method (e.g., supervised fine-tuning, parameter-efficient fine-tuning), and using frameworks like TensorFlow, PyTorch, or Hugging Face Transformers for training. Subsequent evaluation and iterative refinement ensure the model meets application-specific requirements [37].

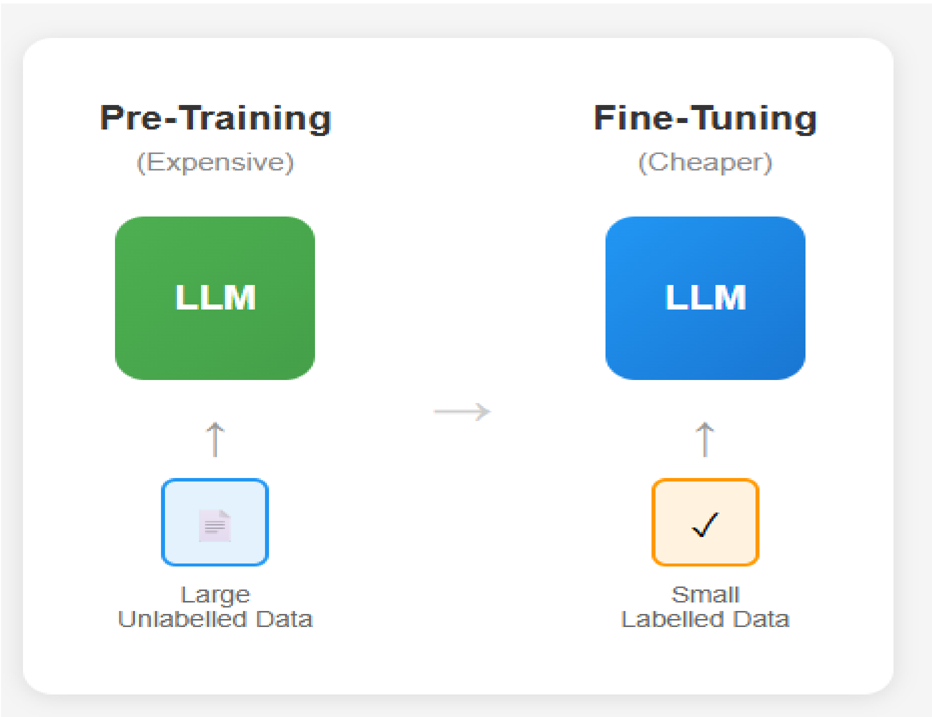

Figure 2: Pre-Training and Fine-Tuning

Pre-training can be likened to imparting broad, foundational knowledge-similar to a child receiving general education-where the model learns language structure, grammar, and semantic relationships from large-scale, diverse datasets. In contrast, fine-tuning represents specialized training, akin to enrolling in a professional program such as medical school, where the model is adapted to perform specific tasks or operate within domain-specific contexts with greater accuracy and relevance.

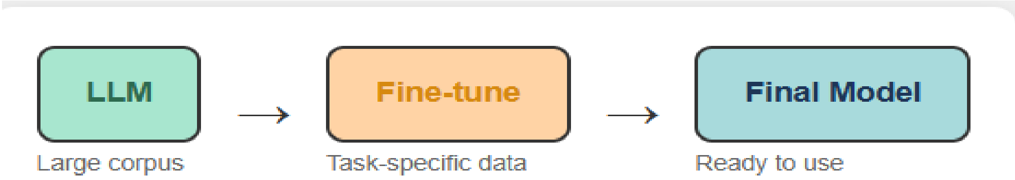

Figure 3: Pre-Training

ChatGPT Evolution: A Version-wise Perspective

This fragment summaries crucial enhancements across versions from GPT-1 to GPT-4o, emphasizing their assistances and restrictions in the larger field of AI.

GPT-1: The Foundational Model

OpenAI announced GPT-1 in 2018, which signalled the commencement of a new era in NLP based on transformer architecture [38]. With almost 117 million factors, GPT-1 used unsupervised pre-training trailed by task-specific fine-tuning, empowering it to creäte contextual input based coherent text.

Key Highlights:

- Acquaint with a two-phase training pattern: pre-training, which is on huge corpora and fine-tuning, it is on precise tasks.
- Demonstrated contextual language generation capabilities.

Limitations

- Limited model size restricted performance in complex scenarios.
- Generated outputs were often shallow and inconsistent.

GPT-2: Scaling Up

GPT-2 is launched in 2019, considerably augmented the parameter count to 1.5 billion, consenting for more contextually rich and fluent text generation [39]. This scale-up made the model adept at tasks like summarization, translation, and creative writing.

Key Highlights

- Improved fluency and diversity in output text.
- Demonstrated zero-shot and few-shot learning capabilities.

Limitations

- Raised ethical concerns due to its potential misuse (e.g., generating fake news).
- Required substantial computational resources for deployment.

GPT-3: A Paradigm Shift

GPT-3, unveiled in 2020, introduced an architectural leap with 175 billion parameters [17]. Its performance in few-shot and zero-shot tasks was widely recognized, making it suitable for a vast array of real-world applications.

Key Highlights:

- Exceptional generalization ability across diverse NLP tasks.
- Delivered state-of-the-art performance in text generation, coding, and dialogue.

Limitations:

- Computationally expensive to train and use.
- Prone to occasional hallucinations or logically incorrect outputs.
- Operated without real-time access to the internet.

GPT-4: Enhanced Reliability

GPT-4, released in 2023, focused on mitigating GPT-3's weaknesses. While specific parameter details were undisclosed, the model emphasized improved factual accuracy, context retention, and task adaptability [41].

Key Highlights

- Greater alignment with user intent.
- Enhanced reliability in factual and technical content.
- Optimized for energy and cost-efficiency in deployment.

Limitations

- Continued absence of live web access.
- Occasional bias or overfitting in task-specific applications.

GPT-4o: A Multimodal Milestone

In May 2024, OpenAI introduced GPT-4o (the "o" standing for "omni"), a multimodal model capable of processing text, images, and audio in a unified architecture [42]. It offered increased responsiveness, creativity, and contextual depth.

Key Enhancements

- Streamlined output with more concise and structured explanations.
- Advanced support for programming, academic writing, and literary analysis.
- Enabled more natural and interactive conversations.
- Significantly faster and more cost-effective than GPT-4 API variants.

Limitations

- Still operates in a static, pre-trained mode without real-time updates.

Significance of the GPT Evolution

The journey from GPT-1 to GPT-4o illustrates the exponential growth in AI model sophistication. Each generation has pushed the boundaries of what is achievable in language understanding and generation, thereby opening new possibilities in education, healthcare, business, and scientific research [43].

GPT-4o, in particular, demonstrates a well-rounded blend of performance, creativity, and practical utility. Understanding this evolution helps researchers and practitioners grasp how AI continues to evolve toward more human-like intelligence. Table 1 shows the comparative analysis of ChatGPT versions.

Table 1: Comparative Analysis of ChatGPT Versions

Aspect	GPT-1	GPT-2	GPT-3	GPT-4	GPT-4o
Year Released	2018 [44]	2019 [45]	2020 [46]	2023 [47]	2024 [42]
Model Size	117M parameters [44]	1.5B parameters [45]	175B parameters [46]	Undisclosed [47]	Similar to GPT-4, optimized [42]
Context Length	~512 tokens [44]	~1,024 tokens [45]	2,048 tokens [46]	8,192–32,768 tokens [41]	128,000 tokens [42]
Training Data	BooksCorpus, Wikipedia [44]	WebText corpus [45]	Web + code + books (up to 2020) [46]	Expanded & cleaned data (pre-2023) [41]	Most diverse & latest (pre-2024) [42]
Capabilities	Basic NLP tasks [44]	Coherent generation [45]	Few/zero-shot learning [46]	Better reasoning, alignment [41]	Multimodal (text, vision, audio) [42]
Benchmarks (MMLU / GSM8K)	Not reported	Low scores [49]	~43% / ~16% [49]	~86.4% / ~92% [50]	~87.2% / ~94% [51]
Speed	Fast, basic model [44]	Moderate latency [45]	Slower than GPT-2 [46]	Slower due to size [41]	Much faster than GPT-4 [42]
Accuracy	Basic text generation [44]	Improved fluency [45]	High but inconsistent [46]	High factual accuracy [41]	Best so far [51]
Cost	Low [44]	Moderate [45]	High [46]	Very high [41]	Lower than GPT-4 [42]
Usability	Internal research use [44]	Limited release [45]	Public API [46]	ChatGPT Plus users [41]	Default model for free & Plus [42]
Limitations	Small capacity [44]	Repetition, bias [45]	Hallucination, cost [46]	No web access, expensive [41]	Still lacks real-time updates [42]
Ethical Issues	Minimal [44]	Misinformation potential [45]	Misuse risks, bias [46]	AI hallucinations, bias [41]	Improved safety, still under scrutiny [51]

Applications of ChatGPT Across Versions

Chatbots and Virtual Assistants

Across its versions, ChatGPT has played a transformative role in powering intelligent conversational agents. These virtual assistants are integrated into customer service systems, help desks, and other interactive platforms, offering context-sensitive, human-like responses. As the models have evolved from GPT-3 to GPT-4 and GPT-4o, there have been marked enhancements in conversation continuity, personalization, and comprehension of user intent, leading to more engaging and effective dialogue systems [52][42].

Education and Tutoring

In the educational sector, ChatGPT is being utilized as a virtual tutor and knowledge assistant. It supports learners by simplifying complex topics, answering academic queries, and assisting with exam preparation and concept clarification. Advanced versions like GPT-4 and GPT-4o deliver better contextual understanding, reasoning, and personalized feedback, making them suitable for adaptive learning environments [54].

Content Creation and Summarization

ChatGPT has significantly contributed to automating content workflows such as drafting articles, emails, marketing copy, and social media posts. Additionally, it efficiently summarizes lengthy documents, including academic papers, business reports, and legal contracts. As models progressed, especially in GPT-4 and GPT-4o, the outputs have become more coherent, stylistically appropriate, and tailored to the user’s needs, enhancing their utility in professional and creative industries [55].

Code Generation and Debugging (e.g., GitHub Copilot)

When embedded in tools like GitHub Copilot, ChatGPT assists programmers by generating code snippets, recommending solutions, and identifying bugs across multiple programming languages. Newer models show improved natural language comprehension and code reasoning, enabling developers to move from vague problem descriptions to functional code more efficiently. These features benefit both beginners and professionals in software development [56][57].

Medical and Legal Insights (with Disclaimers)

ChatGPT is increasingly explored for use in sensitive domains such as healthcare and law, where it can offer general insights like explaining symptoms, suggesting questions for doctors, or outlining basic legal procedures. However, these uses come with strong disclaimers, emphasizing the non-expert nature of the tool. GPT-4 and GPT-4o have improved factual grounding and domain-specific comprehension, but human verification and ethical caution remain critical [58][59].

Challenges and Ethical Concerns in LLMs and ChatGPT

Bias and Misinformation

Large Language Models (LLMs), including ChatGPT, are trained on vast and diverse datasets that reflect human-generated content from the internet. As a result, these models may inadvertently learn and reproduce societal biases—related to race, gender, or ideology—and propagate misinformation. This presents a significant ethical concern, particularly when LLMs are used in sensitive applications like education, hiring, or healthcare. Research has shown that such biases can be systemic and hard to eliminate entirely, even with fine-tuning and filtering techniques [60][61].

Privacy and Data Usage

The training data for LLMs often includes publicly available text, which may contain personal or sensitive information. This raises questions about data consent, user anonymity, and compliance with data protection regulations such as GDPR. Additionally, real-time usage of ChatGPT might risk capturing user-entered private data, necessitating strict privacy controls and data minimization strategies [62][63].

Over-Reliance and Hallucination

While ChatGPT is highly capable, it is not infallible. One major limitation is **hallucination**—the generation of false or fabricated information presented as fact. This issue becomes critical in domains like medicine, law, or science, where factual correctness is essential. Over-reliance on LLMs without human oversight can lead to misinformation, poor decision-making, or erosion of trust in AI systems. Users must be aware of these risks and apply LLM outputs critically [48][53].

Transparency in Model Updates

Another challenge lies in the limited transparency regarding updates, training data, and internal architectures of proprietary models like GPT-4 and GPT-4o. Without public disclosure of such details, researchers and policymakers face difficulties in evaluating the model's behavior, risks, and trustworthiness. Transparency is critical not only for reproducibility and academic scrutiny but also for building public trust in AI applications [47][22].

Future Directions

The future of large language models (LLMs), including anticipated iterations such as GPT-5, points toward greater contextual understanding, improved long-term memory, and enhanced personalization capabilities. These advancements aim to enable AI systems to remember user preferences across sessions, adapt dynamically, and continuously learn without retraining from scratch—addressing one of the current limitations of static models [47][22]. Additionally, a strong trend involves integrating LLMs with **real-time systems** such as **augmented and virtual reality (AR/VR)** environments, robotics, and Internet of Things (IoT) platforms. This convergence could power more immersive, intelligent human-computer interactions and autonomous decision-making systems [40]. Future models are also expected to support **multi-modal inputs** (combining text, images, audio, and video) and deliver seamless, real-world utility in healthcare, education, and enterprise settings.

To ensure the responsible development and deployment of such powerful models, evolving **regulatory and ethical frameworks** are becoming increasingly essential. Policymakers and AI developers are working together to design safety protocols, enforce transparency, and promote fairness, accountability, and interpretability in AI systems. The European Union's **AI Act** and the **OECD AI Principles** represent key steps in formalizing governance models that balance innovation with societal protection [28][21].

Conclusion

The evolution of ChatGPT from its early versions to GPT-4o highlights significant advancements in conversational AI. Each iteration has brought improvements in language understanding, reasoning, and usability, making AI more accessible and useful across various domains. Looking ahead, the focus will be on enhancing accuracy, ensuring ethical use, and promoting transparency to build more reliable and responsible AI systems.

References

[1] Q. Chen, X. Zhu, Z.-H. Ling, S. Wei, H. Jiang, and D. Inkpen, "A Survey on Applications of Conversational AI in Healthcare," *ACM Computing Surveys*, vol. 54, no. 7, pp. 1–36, 2021, doi: 10.1145/3453155.

[2] J. Gao, M. Galley, and L. Li, "Neural Approaches to Conversational AI," in *The Handbook of Natural Language Processing*, Chapman and Hall/CRC, 2019, pp. 107–143.

[3] M. B. Hoy, "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants," *Medical Reference Services Quarterly*, vol. 37, no. 1, pp. 81–88, 2018, doi: 10.1080/02763869.2018.1404391.

[4] M. Jain, P. Kumar, R. Kota, and S. N. Patel, "Evaluating and Informing the Design of Chatbots," in *Proc. CHI 2018*, doi: 10.1145/3173574.3173996.

[5] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface: Talking to Smart Devices*. Springer, 2016.

[6] R. Sarikaya, "The Technology Behind Personal Digital Assistants," *Microsoft AI and Research Blog*, 2017.

[7] T. Zhao, K. Xie, and M. Eskenazi, "Knowledge-Grounded Dialogue Generation with Pre-trained Language Models," in *Proc. EMNLP 2020*.

[8] Fischer, 2021.

[9] Management Solutions, “The Rise of LLM,” [Online]. Available: https://www.managementsolutions.com/.../rise-of-llm.pdf

[10] M. Bréal, *Essai de sémantique*, Paris, 1883.

[11] J. Weizenbaum, “ELIZA—A Computer Program for the Study of Natural Language Communication between Man and Machine,” *Commun. ACM*, vol. 9, no. 1, pp. 36–45, 1966.

[12] L. R. Rabiner, “A Tutorial on Hidden Markov Models,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[13] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, “A Neural Probabilistic Language Model,” *JMLR*, vol. 3, pp. 1137–1155, 2003.

[14] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient Estimation of Word Representations in Vector Space,” *arXiv preprint*, arXiv:1301.3781, 2013.

[15] D. Bahdanau, K. Cho, and Y. Bengio, “Neural Machine Translation by Jointly Learning to Align and Translate,” in *Proc. ICLR*, 2016.

[16] A. Vaswani *et al.*, “Attention is All You Need,” in *Proc. NeurIPS*, 2017.

[17] T. B. Brown *et al.*, “Language Models are Few-Shot Learners,” in *Proc. NeurIPS*, 2020.

[18] Nature Editorial, “2023: The Year of AI,” *Nature*, 2023.

[19] Open Source Initiative, “Rise of Open Source LLMs,” 2023.

[20] Microsoft, Google, GitHub, “Official Blog Announcements on LLM Integrations,” 2023.

[21] OECD, “OECD Principles on Artificial Intelligence,” 2023. [Online]. Available: https://www.oecd.org/ai

[22] R. Bommasani *et al.*, “Foundation Models in the Era of Customization,” *Stanford CRFM White Paper*, arXiv:2311.08927, 2023.

[23] Y. K. Dwivedi *et al.*, “Artificial Intelligence (AI): Multidisciplinary Perspectives,” *Int. J. Inf. Manage.*, vol. 71, p. 102642, 2023.

[24] O. Zawacki-Richter *et al.*, “Systematic Review of Research on AI Applications in Higher Education,” *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, 2019.

[25] L. Floridi and M. Chiriatti, “GPT-3: Its Nature, Scope, Limits, and Consequences,” *Minds and Machines*, vol. 30, no. 4, pp. 681–694, 2020.

[26] R. Bommasani *et al.*, “On the Opportunities and Risks of Foundation Models,” *arXiv preprint*, arXiv:2108.07258, 2021.

[27] OpenAI, “GPT-4 Technical Report,” 2023. [Online]. Available: https://openai.com/research/gpt-4

[28] European Commission, "AI Act: Europe's New Legal Framework for AI," 2024. [Online]. Available: https://digital-strategy.ec.europa.eu

[29] J. Devlin *et al.*, "BERT: Pre-training of Deep Bidirectional Transformers," in *Proc. NAACL-HLT*, 2019.

[30] N. Srivastava *et al.*, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *JMLR*, vol. 15, pp. 1929–1958, 2014.

[31] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words," in *Proc. ICLR*, 2021.

[32] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[33] A. Radford *et al.*, "Improving Language Understanding by Generative Pre-Training," *OpenAI*, 2018.

[34] J. Howard and S. Gugger, "Universal Language Model Fine-tuning for Text Classification," in *Proc. ACL*, 2018.

[35] S. Ruder, "Transfer Learning in Natural Language Processing," in *Proc. NAACL-HLT Tutorials*, 2019.

[36] C. Raffel *et al.*, "Exploring the Limits of Transfer Learning," *JMLR*, vol. 21, no. 140, pp. 1–67, 2020.

[37] P. Liu *et al.*, "Pre-train, Prompt, and Predict: A Survey of Prompt Engineering," *arXiv preprint*, arXiv:2302.11382, 2023.

[38] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," *OpenAI*, 2018.

[39] A. Radford *et al.*, "Language Models are Unsupervised Multitask Learners," *OpenAI*, 2019.

[40] X. Chen *et al.*, "Multimodal Foundation Models for Robotics and XR Applications," *arXiv preprint*, arXiv:2402.06711, 2024.

[41] OpenAI, "GPT-4 Technical Report," 2023. [Online]. Available: https://openai.com/research/gpt-4

[42] OpenAI, "GPT-4o: Introducing Our New Omnimodel," 2024. [Online]. Available: https://openai.com/index/gpt-4o

[43] L. Floridi and M. Chiriatti, "GPT-3: Its Nature, Scope, Limits, and Consequences," *Minds and Machines*, vol. 30, pp. 681–694, 2020.

[44] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," *OpenAI*, 2018.

[45] A. Radford *et al.*, "Language Models are Unsupervised Multitask Learners," *OpenAI*, 2019.

[46] T. Brown *et al.*, "Language Models are Few-Shot Learners," in *Proc. NeurIPS*, 2020.

[47] OpenAI, "ChatGPT Updates: Introducing Memory and Personalization Features," 2024. [Online]. Available: https://openai.com

[48] Z. *et al.*, "Survey of Hallucination in Natural Language Generation," *ACM Comput. Surv.*, vol. 55, no. 12, pp. 1–38, 2023.

[49] D. Hendrycks *et al.*, "Measuring Massive Multitask Language Understanding (MMLU)," *arXiv preprint*, arXiv:2009.03300, 2021.

[50] OpenAI, "GPT-4 Evaluation Benchmarks," 2023. [Online]. Available: https://openai.com/research/gpt-4

[51] OpenAI, "GPT-4o Benchmark Results," 2024. [Online]. Available: https://openai.com/index/gpt-4o

[52] M. B. Hoy, "ChatGPT: OpenAI's Conversational AI System," *Medical Reference Services Quarterly*, vol. 42, no. 1, pp. 84–91, 2023.

[53] H. Nori *et al.*, "Capabilities of GPT-4 on Medical Challenge Problems," *arXiv preprint*, arXiv:2303.13375, 2023.

[54] O. Zawacki-Richter *et al.*, "Systematic Review of Research on AI Applications in Higher Education," *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, 2019.

[55] S. Kreps, R. McCain, and M. Brundage, "All the News That's Fit to Fabricate," *Journal of Media Ethics*, vol. 38, no. 2, pp. 78–89, 2023.

[56] GitHub, "GitHub Copilot Documentation," 2022. [Online]. Available: https://docs.github.com/en/copilot

[57] M. Chen *et al.*, "Evaluating Large Language Models Trained on Code," *arXiv preprint*, arXiv:2107.03374, 2021.

[58] B. N. Patel *et al.*, "AI in Healthcare: Applications and Legal Challenges," *Lancet Digital Health*, vol. 5, no. 4, pp. e203–e212, 2023.

[59] H. Surden, "Artificial Intelligence and Legal Reasoning," *Harvard J. Law & Tech.*, vol. 36, no. 2, pp. 345–378, 2023.

[60] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the Dangers of Stochastic Parrots," in *FAccT '21*, pp. 610–623, 2021.

[61] R. Bommasani *et al.*, "On the Opportunities and Risks of Foundation Models," *Stanford CRFM Report*, arXiv:2108.07258, 2021.

[62] A. S. Luccioni and J. Viviano, "Who Pays for AI?" *Patterns*, vol. 3, no. 7, p. 100570, 2022.

[63] F. Roio *et al.*, "Privacy Risks of Language Models: A Survey," *arXiv preprint*, arXiv:2301.08701, 2023.

☯☯☯