

Large Language Models (LLM) in Industry: A Survey of Applications, Challenges, and Trends

Zina Chkirbene¹, Ridha Hamila¹, Ala Gouissem², Unal Devrim³

¹ Electrical Engineering, Qatar University, Qatar,

² College of Computing and Information Technology, University of Doha for Science and Technology, Qatar,

³KINDI Center for Computing Research, College of Engineering, Qatar University, Qatar,

Abstract—Large Language Models (LLMs) are transforming industries by automating tasks such as text generation, data analysis, and customer interactions. Their impact spans various sectors, including finance, healthcare, legal services, and education, where they streamline operations and enhance decision-making. Despite these advantages, the adoption of LLMs is hindered by challenges such as high computational costs, data privacy concerns, and the lack of explainability. Existing surveys on LLMs primarily focus on their capabilities and applications, emphasizing their role in generating human-like text, processing unstructured data, and supporting decision-making. However, these studies also highlight the significant limitations of LLMs, particularly around computational expense, privacy, and the "black box" nature of their outputs, which restrict their use in critical, regulated industries. This paper builds on prior work by exploring emerging solutions to address these challenges. It examines innovations such as domain-specific LLMs, LLM-as-a-Service (LLMaaS), and advancements in explainable AI (XAI) to enhance transparency and accessibility. The paper provides practical insights into how businesses can strategically adopt LLMs while mitigating risks, making them more viable for broader industry application.

Index Terms—Large Language Models, automation, bias, data privacy, industry applications.

I. INTRODUCTION

In recent years, Large Language Models (LLMs) have become a cornerstone of advancements in artificial intelligence (AI), revolutionizing how industries engage with language based tasks [1]. These models, built on sophisticated neural networks and trained on vast datasets, have demonstrated remarkable abilities in understanding, generating, and analyzing human language with high levels of coherence and accuracy. LLMs, such as OpenAI's GPT-4, Google's BERT, and PaLM, are no longer limited to conventional tasks like translation and summarization. They now extend their utility across diverse sectors by automating processes, enhancing decision-making, and optimizing operational efficiency [2]. As industries strive to remain competitive in an increasingly digital landscape, many are turning to LLMs to address challenges traditionally managed by humans. This shift is evident in several sectors. For instance, in customer service, LLM-powered chatbots provide real-time support, answering queries and resolving issues 24/7, significantly enhancing customer experience. In content creation, marketing and media companies use LLMs to swiftly generate articles, product descriptions, and social media posts at scale. Similarly, in software development, LLMs

assist developers by generating code and debugging, leading to greater efficiency. In research and development (R&D), LLMs expedite tasks such as data processing, literature reviews, and technical documentation, allowing researchers to derive insights much faster than traditional methods. The scalability and flexibility of LLMs are key advantages, enabling them to be fine-tuned for specific industries and applications, making them adaptable to unique operational requirements.

However, as LLMs are increasingly integrated into business processes, several challenges arise. Key concerns include bias in decision-making, data privacy issues, high computational costs, and difficulties in integrating LLMs into existing workflows. For example, while LLMs excel at processing large volumes of data, they may unintentionally propagate biases inherent in their training data, which raises ethical concerns in sensitive areas like recruitment, law, and healthcare. Furthermore, the substantial computational resources required to train and deploy these models can pose significant barriers, especially for smaller organizations that lack the infrastructure to support large-scale AI systems. The environmental impact of training such large models is also becoming a growing concern, as their energy consumption raises sustainability issues. Despite these limitations, industries are finding innovative ways to overcome barriers and unlock the full potential of LLMs. Solutions include partnerships with cloud-based AI service providers [3], which help organizations access the necessary computational resources, and advancements in energy-efficient AI models, which mitigate some of the environmental and resource-related challenges [4].

This survey aims to provide a comprehensive exploration of how LLMs are being applied across different industries, focusing on their real-world benefits, associated challenges, and emerging trends. By gathering insights from organizations that are actively using LLMs, this study highlights the key applications that have had the greatest impact, analyzes the obstacles companies encounter during implementation, and examines future opportunities for LLM-driven business innovation. Moreover, this survey offers a clear perspective on the evolving role of LLMs in industry, providing actionable insights for businesses seeking to leverage these technologies to remain competitive and future-proof their operations. By understanding both the promises and limitations of LLMs, industry leaders will be better equipped to navigate the complexities of AI adoption and harness its full potential to drive

business success.

The remainder of this paper is organized as follows: Section II provides an overview of LLMs, while Section III presents their applications across various industries. Section IV discusses the benefits of LLMs in industry, and Section V explores the challenges and limitations associated with their deployment. Finally, Section VI concludes the study.

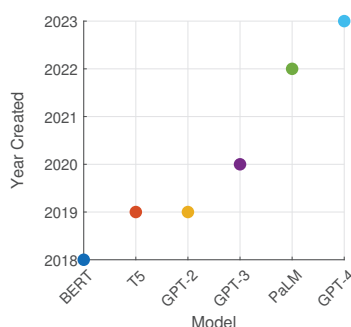
II. OVERVIEW OF LLM TECHNOLOGY

LLMs are a major advancement in artificial intelligence, especially in natural language processing. Built on deep neural networks and transformer architecture, LLMs handle tasks like text generation, translation, and summarization with flexibility and scalability. Their capacity to learn from vast text data enables them to outperform traditional NLP methods across various applications.

A. Popular LLMs

In recent years, numerous LLMs have emerged, each with distinct strengths tailored to diverse applications. OpenAI's GPT, known for text generation, completion, and dialogue, is widely used in chatbots, creative writing, and code generation [5]. Google's BERT specializes in contextual understanding and question answering, making it valuable for text classification, sentiment analysis, and information retrieval. Another model by Google, PaLM (Pathways Language Model), supports scalable performance across tasks and is widely applied in translation, summarization, and general NLP [6]. Google's T5 model transforms all NLP tasks into text generation, suitable for translation, summarization, and text classification [7]. DeepMind's Gopher focuses on reasoning and question answering, making it useful for research and document comprehension [8]. Meta AI's LLaMA, optimized for smaller datasets, finds utility in academic research and smaller-scale NLP applications [9]. Figure 1 presents the release timeline of these influential LLMs, demonstrating their rapid advancement and widespread industry adoption, which has driven innovation and efficiency across sectors.

Fig. 1: LLM models and their creation dates.



B. Recent Advancements in LLMs

The surveys presented in Table I offer extensive discussions on the technical advancements and scaling laws in LLMs,

focusing primarily on the evolution of model architectures, data size, and their impact on performance. These works, while thorough in their analysis of theoretical improvements, largely center around general-purpose applications of LLMs and overlook the specific challenges and use cases encountered in various industries. As such, there remains a need for a survey that delves into the practical, real-world implementations of LLMs across sectors like healthcare, finance, and customer service, where industry-specific requirements and operational complexities demand tailored solutions.

This paper addresses that gap by exploring how LLMs are transforming automation, decision-making, and content generation across different fields, and by highlighting the rise of domain-specific models designed for specialized tasks such as legal document analysis, financial risk assessment, and medical diagnostics. Additionally, it examines emerging trends such as XAI and energy-efficient model designs, which are critical for addressing the ethical, environmental, and regulatory challenges faced by industries adopting LLMs.

III. INDUSTRY APPLICATIONS OF LLMs

LLMs have become transformative tools in a variety of industries, offering new capabilities to automate tasks, enhance productivity, and streamline operations. Their ability to process, understand, and generate human language has made them indispensable in fields ranging from customer service to healthcare. This section explores the key applications of LLMs across different industries, showcasing how they are driving innovation and creating tangible benefits.

A. Healthcare

In the healthcare industry, LLMs have demonstrated great potential in enhancing diagnostic accuracy and automating the management of medical records [15]. IBM Watson Health [16] has developed AI-powered solutions that assist doctors by analyzing patient records and medical literature to suggest potential diagnoses and treatment options. For example, Watson can scan millions of medical papers and clinical trial reports to identify relevant treatments for a patient's specific condition, speeding up diagnosis and improving treatment plans. In telemedicine, platforms like Babylon Health use LLMs to interact with patients, answering basic medical questions, assessing symptoms, and recommending care pathways. This enables healthcare providers to manage patient care more efficiently and remotely, especially in regions where access to healthcare is limited. Additionally, LLMs are used to automate the management of electronic health records (EHRs). By summarizing patient histories, medical professionals can quickly access a comprehensive overview of a patient's condition, allowing for better-informed decisions. Epic Systems, a leader in EHR technology, uses LLMs to streamline medical record-keeping, helping doctors and nurses save time on documentation and focus more on patient care.

B. Customer Service and Support

One of the most prominent uses of LLMs is in customer service and support. Companies increasingly rely on LLM-powered chatbots and virtual assistants to handle customer

TABLE I: Comparison of LLM Research Contributions

Reference	LLM Overview and Scaling Laws	Applications and Domain-Specific Models	Technical Advancements and Deployment	Challenges and Environmental Impact	Explainability, XAI, and Future Trends
[10]	Overview of LLMs from statistical to neural language models, focusing on the impact of scaling model size. Introduced the scaling law for model size and dataset size.	Focus on LLM performance in text generation and other common NLP tasks. No explicit focus on domain-specific models.	Discusses pre-training and fine-tuning approaches. No detailed discussion on specific model deployment strategies.	High computational demands, privacy concerns, and resource limitations. Limited attention to environmental impact.	Limited to discussions of transparency challenges, not focusing on explainability tools. Emergent abilities and alignment issues noted as key future directions.
[11]	Overview of large pre-trained language models, focusing on scaling effects for improving model performance. Introduced the Chinchilla scaling law.	Focused on general NLP applications, without deep industry-specific examples. No explicit focus on domain-specific customization.	Explores optimal compute-resource distribution, pre-training strategies. No deep focus on industry-specific deployment, more on general training insights.	Highlights challenges in balancing computational resources and model performance. Environmental impact mentioned but not in-depth.	Discusses model alignment but no deep dive into explainability solutions. Future focus on model scaling, emergent abilities, and training efficiency.
[12]	Overview of GPT models and the benefits of pre-training and fine-tuning with large datasets. In-context learning and emergent abilities in GPT-3 discussed.	Focused on general NLP tasks like text generation, translation, and summarization. No explicit focus on domain-specific models.	Highlights breakthroughs in model size and in-context learning. Deployment through OpenAI's API, minimal exploration of industry-specific integration challenges.	Mentioned computational challenges, but focuses on GPT-3 as an API-based model. Limited focus on environmental concerns.	Not focused on explainability tools, only discusses GPT-3's conversational abilities. Future focus on model scaling and reducing biases in model outputs.
[13]	LLM overview focusing on instruction-following models like InstructGPT. Discusses emergent abilities in instruction-following models.	Focused on improving NLP task performance through instruction tuning. No specific focus on domain-specific model customization.	Technical advancements in aligning models with human preferences using reinforcement learning. Deployment mainly through API.	Discusses model alignment issues (biases, toxicity). No extensive discussion on environmental impact.	Mentions alignment but not focused on practical explainability tools. Future trends focus on model alignment with human preferences, exploring safer outputs.
[14]	Overview of LLMs with emphasis on societal implications, and focus on technical breakthroughs in pre-trained models. Discusses emergent abilities but less focus on detailed scaling laws.	Focuses on NLP tasks such as language understanding and generation. Domain-specific models are briefly mentioned but not deeply explored.	Covers pre-training and fine-tuning techniques, with emphasis on interdisciplinary applications. Focuses more on interdisciplinary applications, no deep focus on deployment strategies.	Discusses societal implications of LLMs such as fairness, ethics, and transparency challenges. Mentions sustainability concerns but does not deeply explore energy-efficient models.	Discusses transparency challenges but lacks concrete solutions for explainability. Future trends focus on ethical challenges, transparency, and fairness in LLM usage.

inquiries, providing around-the-clock service without human agents. For instance, Sephora utilizes a chatbot powered by an LLM to recommend products to users based on their preferences and previous purchases. This chatbot can answer questions about product availability, ingredients, and more, all in real-time. LLMs are not just limited to handling simple queries; they are also capable of engaging in complex, context-aware conversations. For example, Amazon's Alexa and Google Assistant leverage LLM technology to offer conversational interfaces, enabling users to make purchases, set reminders, or ask about their delivery status. Additionally, LLMs can analyze the sentiment behind customer messages, allowing businesses to tailor responses based on the emotional tone of the conversation and prioritize urgent issues for faster resolution. For instance, H&M uses chatbots that can detect if a customer is frustrated and escalate the case to a human agent if necessary. By automating repetitive tasks, such as answering frequently asked questions or managing returns, LLMs help companies significantly reduce operational costs while improving the overall customer experience. They ensure consistency in responses and provide personalized experiences by remembering customer preferences over time. This is especially impactful in industries like e-commerce, where personalized recommendations and immediate support are key to retaining customers.

C. Content Creation and Marketing

In the realm of content creation and marketing, LLMs have revolutionized the way businesses produce and distribute content. HubSpot, for instance, uses AI-driven tools for content generation, creating blog posts, product descriptions, and marketing emails tailored to specific customer segments. LLMs help automate this process by quickly generating high-quality drafts, which marketing teams can refine and publish in a fraction of the time traditionally required. Moreover, writing-assistance tools like Grammarly leverage LLMs to help marketers and content creators craft clearer, more effective messages. These tools offer real-time feedback on grammar, tone, and clarity, ensuring that the content is both professional and engaging. LLMs also enable highly personalized marketing campaigns by generating targeted advertisements and email messages based on customer behavior and preferences. For instance, Netflix uses LLMs to create personalized viewing recommendations for its users by analyzing their watching habits and preferences. Similarly, Spotify leverages LLMs to curate personalized playlists and generate tailored marketing messages, boosting user engagement and retention. In addition, LLMs help optimize content for search engines. Tools like Yoast SEO use LLMs to suggest changes that improve search rankings, increasing reach and visibility.

D. Research and Development (R&D)

In research and development (R&D), LLMs play a crucial role by accelerating the process of literature review, data analysis, and technical documentation. In pharmaceuticals, for instance, Pfizer uses LLMs to sift through vast amounts of scientific literature and clinical trial data to identify promising drug candidates. This process, which would take human researchers weeks or even months, can now be completed in a matter of hours, enabling faster innovation and product development. In engineering, companies like Siemens use LLMs to assist in technical documentation, summarizing complex reports and making it easier for engineers to stay updated on the latest technological advancements. LLMs can also identify trends in scientific papers and patent filings, helping companies spot emerging opportunities or threats in their industry. LLMs have also shown great promise in drug discovery. For example, Insilico Medicine uses AI-driven tools to analyze molecular structures and identify potential drug candidates. These LLM-powered systems help accelerate the discovery of new treatments by predicting which molecular compounds are likely to be effective in combating diseases, speeding up the drug discovery process and reducing research costs.

E. Software Development and Code Generation

The impact of LLMs on software development has been transformative, particularly with tools including GitHub Copilot, powered by OpenAI's Codex model. This tool assists developers by generating code snippets based on natural language prompts, enabling them to write code faster and more efficiently. For example, a developer might type a prompt like "create a function that sorts a list in Python," and Copilot will generate the corresponding code. This saves time and reduces the likelihood of errors. LLMs are also used for real-time debugging. Tools like DeepCode scan codebases for potential bugs or security vulnerabilities and provide suggestions for fixes. This AI-driven debugging process significantly reduces the time developers spend on code reviews and enhances the overall quality of the software. In more complex scenarios, LLMs can assist in automating large-scale software refactoring. For instance, IBM has developed LLM-powered tools to assist in migrating legacy software systems to modern architectures, reducing manual effort and minimizing errors during the migration process.

F. Finance, Legal, and Consulting

In industries such as finance, legal, and consulting, where data-driven decision-making is critical, LLMs provide substantial support. In finance, JPMorgan uses LLMs to process vast amounts of financial reports, earnings calls, and market data to generate actionable insights. These LLMs can summarize complex reports in minutes, helping analysts make faster and more informed investment decisions. Moreover, LLMs assist in detecting fraudulent transactions by analyzing patterns in transactional data and flagging suspicious activity, improving risk management. In the legal sector, LLMs are particularly

useful for contract review and legal research. Companies like Luminance leverage LLMs to scan large volumes of legal documents and identify key clauses, reducing the time lawyers spend on tedious document review. Similarly, ROSS Intelligence uses AI to help lawyers find relevant case law and legal precedents more efficiently, transforming the way legal research is conducted. In consulting, LLMs assist firms like McKinsey by generating reports, analyzing market trends, and offering recommendations based on extensive data analysis. This reduces the time spent on manual research and allows consultants to focus on strategic decision-making for clients.

G. Human Resources and Recruitment

LLMs are making significant strides in human resources (HR) and recruitment. HireVue, for example, uses AI-driven tools powered by LLMs to screen resumes and rank candidates based on job requirements. This automates a traditionally labor-intensive process, allowing HR departments to quickly identify the best candidates for a role. By analyzing thousands of resumes in seconds, these systems can sort candidates by qualifications, experience, and skills, significantly speeding up the hiring process. Additionally, LLMs like Pymetrics assess candidates' soft skills and cognitive abilities through gamified assessments, providing deeper insights into their suitability for specific roles. LLMs can also handle scheduling interviews, sending personalized emails, and even conducting initial interview rounds through AI-powered chatbots, providing a seamless and efficient hiring experience.

IV. BENEFITS OF USING LLMs IN INDUSTRY

The integration of LLMs into industry operations offers numerous benefits that extend across a wide range of sectors. These models provide businesses with advanced tools to automate processes, improve decision-making, and enhance customer experiences. As LLMs continue to evolve, industries are increasingly recognizing their potential to drive efficiency, scalability, and innovation. This section explores the key benefits that LLMs bring to industry.

A. Automation of Repetitive Tasks

One of the most significant benefits of LLMs is their ability to automate repetitive, labor-intensive tasks that traditionally require human intervention. In industries such as customer service, HR, and finance, LLMs can handle large volumes of routine inquiries, document processing, and data entry. By automating these tasks, businesses can free up their human workforce to focus on more strategic and creative activities. For example, LLM-powered chatbots can automate responses to frequently asked questions in customer service, reducing the need for human agents to manage repetitive inquiries. In HR, LLMs streamline processes like resume screening and interview scheduling, allowing HR professionals to focus on more complex hiring decisions and employee engagement. Similarly, in finance, LLMs can automate the review and analysis of financial statements, allowing analysts to concentrate on high-value tasks such as risk assessment and forecasting.

B. Scalability and Efficiency

LLMs enable companies to scale their operations efficiently without a proportional increase in human labor or infrastructure. Due to the high scalability of LLMs, they can handle vast amounts of data and user interactions simultaneously. This is particularly beneficial in industries such as e-commerce, marketing, and customer support, where businesses must manage large-scale, often global, operations. For instance, in E-commerce, LLMs can generate product descriptions, process customer reviews, and personalize shopping recommendations for millions of users at once. This scalability allows businesses to maintain high-quality interactions with their customers regardless of the number of users, improving both operational efficiency and customer satisfaction. Similarly, in marketing, LLMs enable businesses to run multiple personalized campaigns at scale, tailoring messages to different customer segments with minimal additional effort. This helps marketers deliver targeted content to vast audiences, improving engagement rates and return on investment.

C. Improved Decision-Making and Insights

Another key benefit of LLMs is their ability to assist in data-driven decision-making. LLMs are capable of analyzing large datasets, extracting key insights, and summarizing complex information. In industries like legal, healthcare, and financial services, where decisions are often based on extensive document review and analysis, LLMs can help businesses make faster and more informed decisions. In the legal industry, for example, LLMs can review contracts, legal documents, and case law to identify key clauses or relevant precedents. This enables legal professionals to process and review documents more quickly and accurately, improving their ability to deliver timely and precise legal advice. In healthcare, LLMs can analyze patient medical records, research papers, and clinical trial data to assist doctors in making more accurate diagnoses and treatment plans. By providing healthcare professionals with relevant data insights, LLMs improve both the speed and accuracy of medical decision-making. In finance, LLMs assist in analyzing market trends, financial reports, and regulatory changes, helping businesses navigate complex financial landscapes. These insights enable companies to make better investment decisions, optimize financial planning, and manage risks more effectively.

D. Cost Reduction and Operational Efficiency

LLMs have proven to be powerful tools for driving cost reductions and improving operational efficiency across various industries. By automating repetitive tasks and enhancing decision-making processes, businesses can significantly lower their operational expenses while maintaining or even improving service quality. In sectors like customer service, LLMs can automate a substantial portion of customer interactions, reducing the need for large teams of human agents. LLM-powered chatbots offer 24/7 support at a fraction of the cost of employing round-the-clock customer service representatives, allowing companies to provide consistent service while saving

on labor costs. Similarly, in content creation, LLMs reduce the expenses associated with hiring extensive teams of writers and editors. By automating the generation of articles, social media posts, and marketing materials, businesses can scale their content production while keeping costs low. This efficiency extends to human resources (HR) as well, where LLMs automate processes like resume screening, interview scheduling, and employee onboarding. This not only reduces costs but also increases the speed and scalability of recruitment, which is especially beneficial during periods of rapid organizational growth. By incorporating LLMs across these sectors, companies achieve operational efficiencies that directly contribute to cost savings, allowing them to allocate resources more effectively and focus on strategic priorities.

E. Enhanced Customer Experience and Personalization

LLMs play a pivotal role in enhancing the customer experience through highly personalized interactions. In industries such as e-commerce, hospitality, and entertainment, LLMs analyze vast amounts of customer data to deliver tailored recommendations, personalized content, and customized services. This level of personalization not only improves user satisfaction but also drives customer loyalty and engagement.

In e-commerce, for instance, LLMs analyze browsing and purchase history to recommend products that are highly relevant to individual users, leading to increased conversion rates and improved customer retention. Streaming services in the entertainment industry, such as music and video platforms, use LLMs to suggest movies, shows, and songs based on user preferences, making the platform more engaging and personalized. In hospitality, hotels and travel agencies utilize LLMs to offer customized travel recommendations, suggesting destinations, accommodations, and experiences based on past behavior and preferences. By enhancing personalization, LLMs enable companies to create deeper connections with their customers, resulting in higher satisfaction and repeat business. This capacity for personalization at scale is critical in today's competitive market, where customer experience is a key differentiator.

F. Innovation and Competitive Advantage

The rapid adoption of LLMs provides businesses with a clear competitive advantage in industries where automation, personalization, and enhanced decision making are vital for success. Organizations that effectively integrate LLMs into their operations can innovate faster and respond more quickly to market changes, positioning themselves to outpace competitors. In industries like pharmaceuticals, LLMs accelerate the research and development (RD) process by synthesizing vast bodies of scientific literature, identifying potential drug candidates, and optimizing experimental designs. This enables pharmaceutical companies to bring new treatments to market faster and at a lower cost than those using traditional RD methods. In finance and insurance, firms that adopt LLMs early gain a competitive edge by developing more accurate models for risk assessment and fraud detection, enhancing customer service, and offering more personalized financial

products. In summary, the integration of LLMs across industries enables companies to unlock significant benefits such as automation, scalability, enhanced decision-making, cost savings, and improved customer experiences. By leveraging LLM technology, businesses not only optimize their operations but also position themselves to innovate and maintain a competitive edge in an increasingly digital, data-driven economy. The use cases outlined in earlier sections from customer service to (RD) illustrate how LLMs are reshaping industry landscapes and driving the next wave of technological transformation.

V. CHALLENGES AND LIMITATIONS OF LLMs IN INDUSTRY

While LLMs have transformed industries, their deployment brings several challenges. High computational and energy demands raise costs and environmental concerns, which can limit access for smaller organizations. These models also face issues with bias and fairness due to inherent biases in training data, which can lead to unethical outcomes in critical sectors like healthcare and finance. Privacy and security concerns arise as LLMs, trained on sensitive data, risk unintentionally reproducing personal information, creating compliance challenges under regulations such as GDPR and HIPAA. Additionally, LLMs' lack of explainability complicates adoption in regulated fields that require transparency, such as finance, where understanding model decisions is essential.

LLMs also face ethical and societal challenges, including potential job displacement and the risk of misinformation, as they can be used to generate misleading content at scale. Effective governance and standards for responsible AI use are needed to address these concerns. Furthermore, general-purpose LLMs often struggle with domain-specific knowledge, necessitating extensive fine-tuning for industries like medicine and law, which increases the resource demands. Finally, the high costs of implementing and maintaining LLMs, including infrastructure and expertise requirements, can be prohibitive for smaller enterprises, creating a barrier to broader adoption.

VI. CONCLUSION

The adoption of LLMs across industries has transformed business operations by providing automation, scalability, improved decision-making, and cost reduction. Sectors like technology, finance, and customer service have been early adopters, using LLMs to streamline workflows, enhance customer interactions, and process large volumes of data in real time. Healthcare and legal sectors, while slower to integrate LLMs, are beginning to leverage them for tasks such as summarizing medical records and analyzing legal documents. However, challenges like high computational costs, bias, data privacy, and the lack of explainability in LLMs remain significant obstacles, particularly in regulated industries like healthcare and finance, where transparency is crucial. In conclusion, while LLMs offer clear advantages, overcoming their challenges is essential for responsible and effective deployment. Ongoing research, infrastructure investment, and collaboration between businesses and AI experts will be vital in unlocking the full potential of LLMs and driving future innovation.

VII. ACKNOWLEDGMENT

This work was supported by Qatar University Internal Grant IRCC-2023-237. The statements made herein are solely the responsibility of the author[s].

REFERENCES

- [1] Ehsan Ullah, Anil Parwani, Mirza Mansoor Baig, and Rajendra Singh, "Challenges and barriers of using large language models such as chatgpt for diagnostic medicine with a focus on digital pathology a recent scoping review," *Diagnostic pathology*, vol. 19, no. 1, pp. 43, 2024.
- [2] Ala Gouissem, Zina Chkirbene, and Ridha Hamila, "A comprehensive survey on energy efficiency in federated learning: Strategies and challenges," in *2024 IEEE 8th Energy Conference (ENERGYCON)*. IEEE, 2024, pp. 1–6.
- [3] Zina Chkirbene, Aiman Erbad, Ridha Hamila, Ala Gouissem, Amr Mohamed, Mohsen Guizani, and Mounir Hamdi, "Weighted trustworthiness for ml based attacks classification," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2020, pp. 1–7.
- [4] Anabi Hilary Kelechi, Mohammed H Alsharif, Okpe Jonah Bameyi, Paul Joan Ezra, Iorshase Kator Joseph, Aaron-Anthony Atayero, Zong Woo Geem, and Junhee Hong, "Artificial intelligence: An energy efficiency tool for enhanced high performance computing," *Symmetry*, vol. 12, no. 6, pp. 1029, 2020.
- [5] Gokul Yenduri, M Ramalingam, G Chemmalar Selvi, Y Supriya, Gautam Srivastava, Praveen Kumar Reddy Maddikunta, G Deepti Raj, Rutvij H Jhaveri, B Prabadevi, Weizheng Wang, et al., "Gpt (generative pre-trained transformer)—a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions," *IEEE Access*, 2024.
- [6] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al., "Palm: Scaling language modeling with pathways," *Journal of Machine Learning Research*, vol. 24, no. 240, pp. 1–113, 2023.
- [7] Antonio Mastropaolo, Simone Scalabrino, Nathan Cooper, David Nader Palacio, Denys Poshyvanyk, Rocco Oliveto, and Gabriele Bavota, "Studying the usage of text-to-text transfer transformer to support code-related tasks," in *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*. IEEE, 2021, pp. 336–347.
- [8] Felix Brakel, Uraz Odyurt, and Ana-Lucia Varbanescu, "Model parallelism on distributed infrastructure: A literature review from theory to llm case-studies," *arXiv preprint arXiv:2403.03699*, 2024.
- [9] Sophia M Pressman, Sahar Borna, Cesar A Gomez-Cabello, Syed A Haider, Clifton Haider, and Antonio J Forte, "Ai and ethics: a systematic review of the ethical considerations of large language model use in surgery research," in *Healthcare*. MDPI, 2024, vol. 12, p. 825.
- [10] Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei, "Scaling laws for neural language models," *arXiv preprint arXiv:2001.08361*, 2020.
- [11] Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, et al., "Training compute-optimal large language models," *arXiv preprint arXiv:2203.15556*, 2022.
- [12] Tom B Brown, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.
- [13] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al., "Emergent abilities of large language models," *arXiv preprint arXiv:2206.07682*, 2022.
- [14] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al., "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, vol. 35, pp. 24824–24837, 2022.
- [15] Zina Chkirbene, Ridha Hamila, Devrim Unal, Moncef Gabbouj, and Mounir Hamdi, "Enhancing healthcare systems with deep reinforcement learning: Insights into d2d communications and remote monitoring," *IEEE Open Journal of the Communications Society*, 2024.
- [16] Eliza Strickland, "Ibm watson, heal thyself: How ibm overpromised and underdelivered on ai health care," *IEEE Spectrum*, vol. 56, no. 4, pp. 24–31, 2019.