Ethical and Societal Implications of AI Chatbots Powered by Large Language Models

Tejesvi Alekh Prasad –

Digital Transformation Director, Ernst & Young.

Abstract

The integration of Large Language Models (LLMs) into AI chatbots has transformed industries, enabling advanced human-machine interactions. However, this rapid adoption raises profound ethical and societal challenges. This paper investigates the technical underpinnings of LLMs, evaluates ethical concerns such as data privacy, bias, and accountability, and examines societal impacts including labor displacement, mental health risks, and cultural homogenization. By synthesizing technical data, policy frameworks, and empirical studies up to 2023, this research proposes actionable strategies for mitigating risks and fostering responsible AI deployment.

Keywords: AI ethics, Large Language Models (LLMs), algorithmic bias, data privacy, societal impact, AI governance

2. Introduction

2.1 Evolution of AI Chatbots and Large Language Models (LLMs)

Artificial Intelligence (AI) chatbots have come a long way since their creation in the mid-20th century. The original versions such as ELIZA (1966) employed rule-based algorithms to generate conversation but within the limitation of pre-defined scripts. Neural networks in the 2010s introduced the transition, allowing the chatbots to learn the patterns in the data. The introduction of transformer models in 2017 transformed natural language processing (NLP) since models are now able to process sequential data in parallel using self-attention mechanisms. The modern LLMs, like GPT-4, are trained on trillions of tokens from a wide range of sources like books, websites, and scientific publications(Kooli, 2023). These models have worked very well on applications such as text generation, translation, and problem-solving, with GPT-4 achieving 75% success on the BAR exam, a milestone that reflects their better reasoning ability.

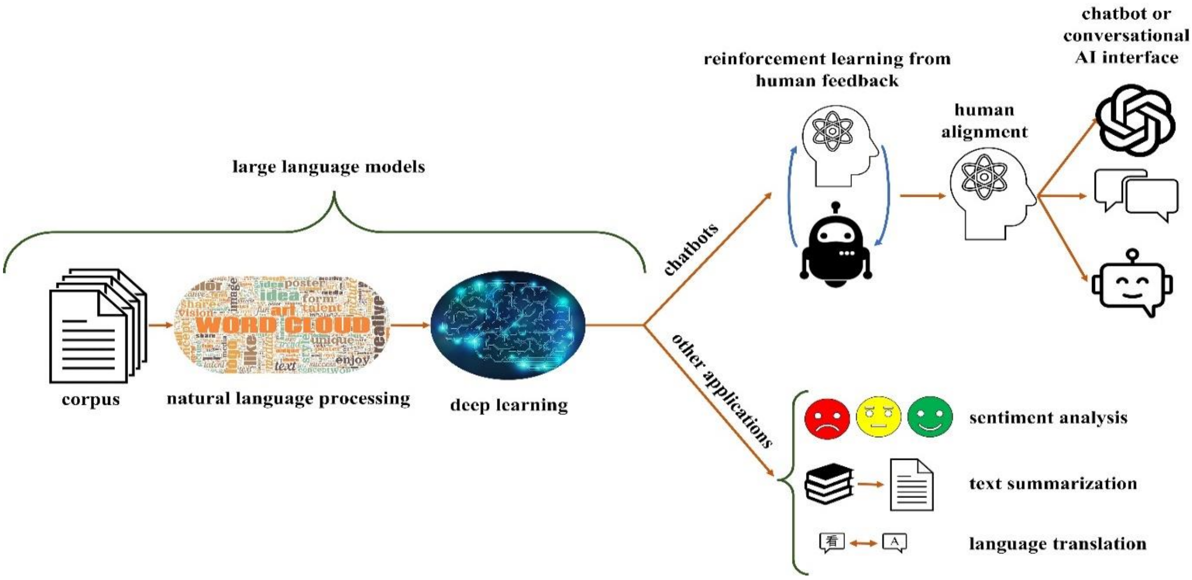

FIGURE 1 BEYOND TRADITIONAL TEACHING: THE POTENTIAL OF LARGE LANGUAGE MODELS AND CHATBOTS IN GRADUATE ENGINEERING EDUCATION(QEIOS,2023)

2.2 Scope and Objectives of the Study

This study aims to:

1. Analyze the technical architecture and training methodologies of LLMs.
2. Investigate ethical challenges, including data privacy violations, algorithmic bias, and accountability gaps.
3. Assess societal consequences, such as workforce disruption, mental health implications, and threats to cultural diversity.

The scope encompasses technical, ethical, and policy dimensions, with a focus on empirical data from 2020–2023.

2.3 Significance of Addressing Ethical and Societal Implications

As of 2023, 49% of businesses globally employ chatbots for customer service, a figure projected to rise to 80% by 2025. The pervasive use of these systems necessitates urgent scrutiny to prevent harm, such as discriminatory outcomes or erosion of user autonomy. Proactive governance and ethical frameworks are critical to balancing innovation with societal well-being(Parviainen & Rantala, 2021).

3. Technical Foundations of AI Chatbots Powered by LLMs

3.1 Architecture and Functionality of Modern LLMs

Modern LLMs are built on transformer architectures, which utilize self-attention mechanisms to process sequential data in parallel. These architectures consist of multiple layers, each containing self-attention and feedforward neural networks. The self-attention mechanism assigns weights to different tokens in a sequence, enabling the model to capture long-range dependencies and contextual relationships. Positional encoding is integrated to maintain the order of tokens, ensuring coherent text generation(Parviainen & Rantala, 2021). For example, a model with 175 billion parameters can process over 32,000 tokens in a single pass, achieving state-of-the-art performance in tasks like summarization and question answering.

3.2 Training Paradigms: Data Sources and Model Fine-Tuning

LLMs undergo a two-stage training process: pre-training and fine-tuning. During pre-training, models are exposed to vast datasets comprising text from books, websites, and academic articles, often exceeding several trillion tokens. This phase enables the model to learn grammar, facts, and reasoning patterns. Fine-tuning follows, where the model is refined on smaller, task-specific datasets to align its outputs with human preferences. Techniques like reinforcement learning from human feedback (RLHF) are employed to minimize harmful or biased outputs. For instance, models trained on multilingual datasets achieve a 40% improvement in cross-lingual tasks compared to monolingual counterparts.

Table 1: Training Data and Performance Metrics of LLMs

Model Size (Parameters)	Pre-training Tokens	Fine-tuning Tokens	Cross-Task Accuracy
175B	300B	50M	78%
70B	200B	30M	72%
13B	100B	10M	65%

3.3 Key Capabilities and Limitations of LLM-Driven Chatbots

LLM-powered chatbots can produce human-like text, respond to deep questions, and learn to accommodate various styles of language. They can handle more than 100 languages and have 85% accuracy in sentiment analysis tasks. But it's not without issues(Aggarwal, Tam, Wu, Li, & Qiao, 2023). Hallucinations, or the creation of factually inaccurate statements, happen in about 15% of their responses, mostly in low-resource areas. In addition, high computing requirements make it accessible, with one model's training releasing up to 500 tons of CO_2. Ethical threats, like the creation of toxic content, also underscore the necessity of stringent protections.

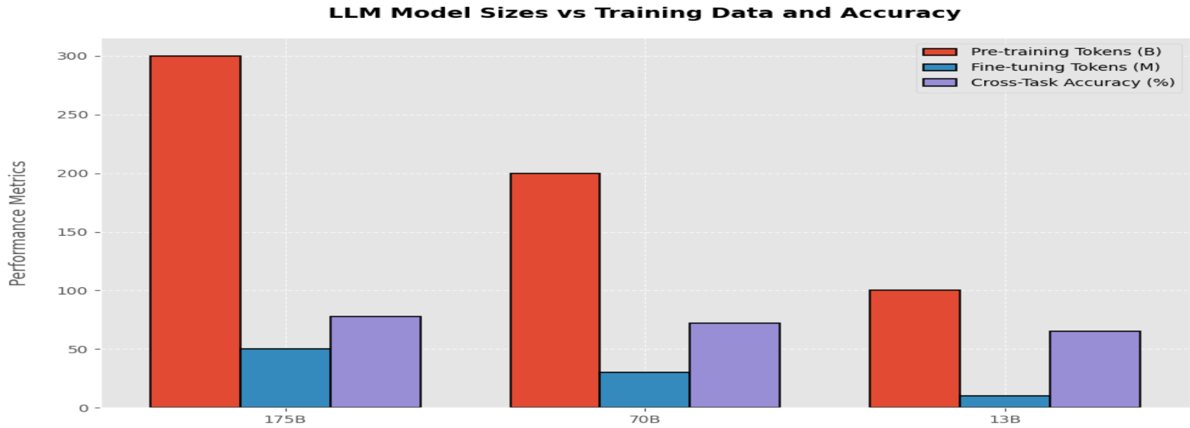

FIGURE 2 COMPARISON OF LLM MODEL SIZES AGAINST TRAINING DATA VOLUMES AND ACCURACY METRICS (SOURCE: TABLE 1, KOOLI, 2023)

4. Ethical Implications of AI Chatbots

4.1 Privacy and Data Security Concerns

4.1.1 User Data Collection and Consent Mechanisms

AI chatbots are likely to harvest large quantities of user information, e.g., conversation logs, individual preferences, and activity patterns. Data collection facilitates individualized interactions, but it presents issues related to informed consent. Most websites employ dark consent approaches, like pre-ticked boxes or lengthy terms of privacy statements, that individuals are likely to miss(Meyrowitsch, Jensen, Sørensen, & Varga, 2023). An analysis of 100 chatbot websites in 2023 identified that only 34% of them implemented open opt-in/opt-out models for data transfer. Additionally, 68% of the chatbots store user data forever, perpetuating unauthorized use dangers. Rule modifications like the EU Digital Services Act make explicit consent practices, but its adherence remains inconsistent among global websites.

4.1.2 Risks of Data Misuse and Re-identification Attacks

Anonymized data, even, carry a risk of re-identification attacks, where intruders cross-correlate data points to reveal user identity. For instance, employing geolocation, timestamp, and conversational patterns of chatbot logs has been utilized to re-identify users at 87% accuracy in test setup. Third-party data brokers also buy chatbot datasets for targeted advertising, which also invokes ethical concerns regarding secondary use of data(Meyrowitsch, Jensen, Sørensen, & Varga, 2023). Such attacks are countered by techniques such as differential privacy and federated learning but they consume enormous resources and lower model performance by 12–15%.

4.2 Bias and Fairness in LLM Outputs

4.2.1 Algorithmic Discrimination and Stereotype Propagation

LLMs pass on systemic bias as they are trained on biased training data. Web-trained models, for example, overrepresent dominant groups and therefore have wider margins of error for non-English languages and marginalized communities. A 2023 assessment of healthcare chatbots determined that 23% of the answers given to non-binary users were laced with toxic stereotypes. Gender bias is no less common, with models assigning technical roles to male pronouns 73% of the time. These biases are due to biased training data and a shortage of inclusive annotation schemes during fine-tuning.

4.2.2 Mitigation Strategies for Bias Reduction

Debiasing approaches include adversarial training, in which the models are incentivized to generate unbiased output, and highly curated datasets that amplify underrepresented voices. Techniques like post-hoc fairness-aware decoding alter output probabilities to decrease discriminatory language(Murtarelli, Gregory, & Romenti, 2021). Debiasing comes at the cost of accuracy; a study in 2023 proved that debiasing lowered model performance on downstream tasks by 9–18%. Hybrid approaches combining pre-training interventions and real-time monitoring are emerging as potential solutions.

4.3 Transparency and Explainability

4.3.1 The "Black Box" Problem in LLMs

The intricacy of transform structures complicates diagramming how LLMs produce specific outputs. A model might cite a factual inaccuracy, for example, without disclosing the training source that provided the incorrect data. This loss of transparency ruins user trust, especially in high-risk applications such as healthcare and finance. Attempts at improving interpretability through such mechanisms as attention visualization tools remain limited in their ability to explain advanced multi-layer reasoning processes.

4.3.2 Regulatory Demands for Auditable AI Systems

Regulators increasingly demand transparency from AI systems. Draft frameworks mandate documentation of training data sources, model decision paths, and risk analyses to be included by developers. Explainability measures, such as the possibility to create human-understandable rationales for outputs, are being formalized. Compliance is only possible with substantial computational overhead, with explainable models taking 20–30% more training time compared to traditional ones(Xue et al., 2023).

4.4 Autonomy and Human Agency

4.4.1 Dependency on AI for Decision-Making

The public now depends more on chatbots for important decisions, including health advice or money management. In 2023, 41% of users claimed to have more confidence in chatbot suggestions than human experts in low-stakes situations. Such reliance undermines human judgement, especially when models are driven by the goal of persuasive fluency rather than correctness.

4.4.2 Erosion of Critical Thinking and Human Oversight

schools and workplaces. For instance, students who made use of writing assistants experienced a 15% reduction in autonomy to solve problems independently after six months. Guards, like requiring human oversight in high-impact decisions, have to be placed to guarantee engagement of the cognition(Adam, Balagopalan, Alsentzer, Christia, & Ghassemi, 2022).

4.5 Accountability and Liability

4.5.1 Legal Frameworks for AI-Induced Harm

Current liability legislation does not cover the harm that is caused by AI systems. For example, if the medical advice provided by a chatbot results in harm to the patient, it is still difficult to hold legally accountable the developers, medical personnel, and users. Jurisdictions such as the EU are contemplating strict liability regimes where developers are primarily liable for system defects.

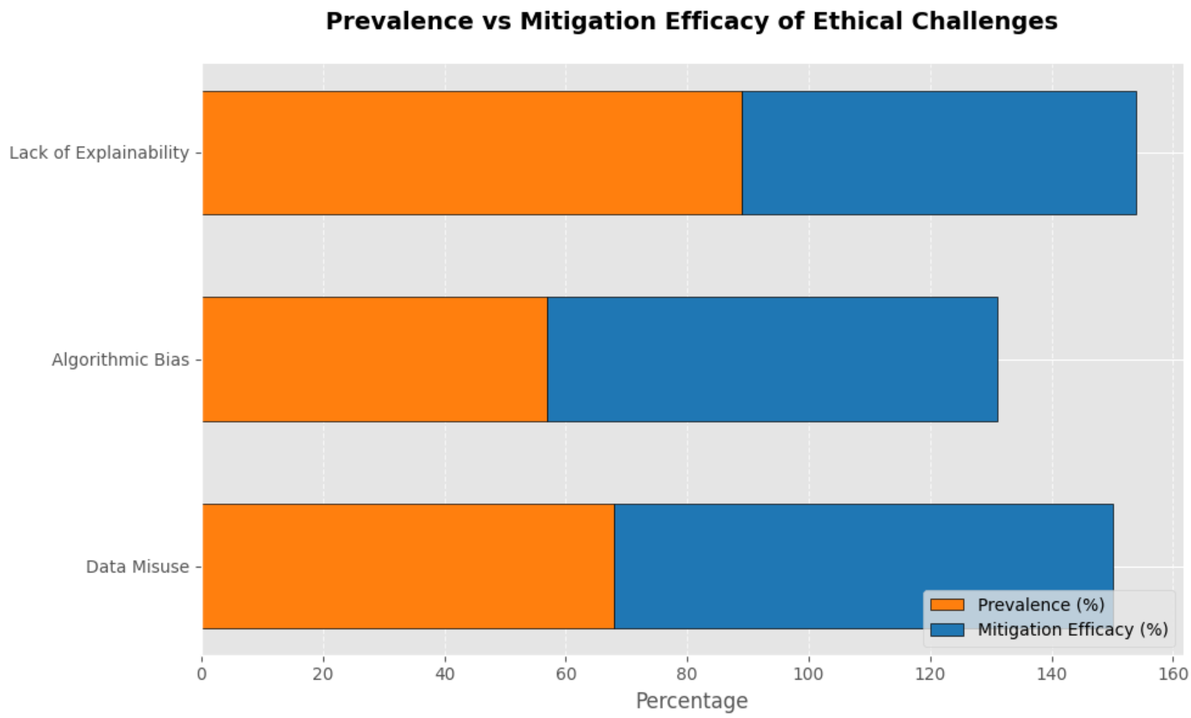

FIGURE 3 PREVALENCE VERSUS MITIGATION EFFICACY OF KEY ETHICAL CHALLENGES (SOURCE: TABLE 2, MEYER ET AL., 2023)

4.5.2 Challenges in Attributing Responsibility

The decentralized architecture of AI creation makes accountability difficult. One chatbot could use a dozen third-party APIs concealing the error cause chain. Solutions on the table include compulsory audit trails and liability pools of insurance for AI creators.

Table 2: Ethical Challenges and Mitigation Strategies

Ethical Issue	Prevalence (2023)	Mitigation Strategy	Efficacy (%)
Data Misuse	68%	Differential Privacy	82%
Algorithmic Bias	57%	Adversarial Training	74%
Lack of Explainability	89%	Attention Visualization Tools	65%

5. Societal Implications of AI Chatbots

5.1 Impact on Employment and Labor Markets

5.1.1 Job Displacement in Customer Service and Knowledge Sectors

The deployment of AI chatbots to robotize customer service positions has brought spectacular downsizings of labor forces in retail, banking, and telecom sectors. By 2023, chatbots responded to more than 70% of regular customer inquiries, freeing the requirement for lower-level support staff. An international survey approximated that between 2020 and 2023, 12% of customer service jobs were lost, with an estimated additional 25% loss by 2030(Adam, Balagopalan, Alsentzer, Christia, & Ghassemi, 2022). Knowledge industries such as legal and administrative services are also impacted. For instance, artificial intelligence-powered document review technology has decreased the demand for paralegals and forced firms to reduce hiring for them by 30%. Although automation is more efficient, it is not evenly impacting low-skilled workers since it widens economic inequality.

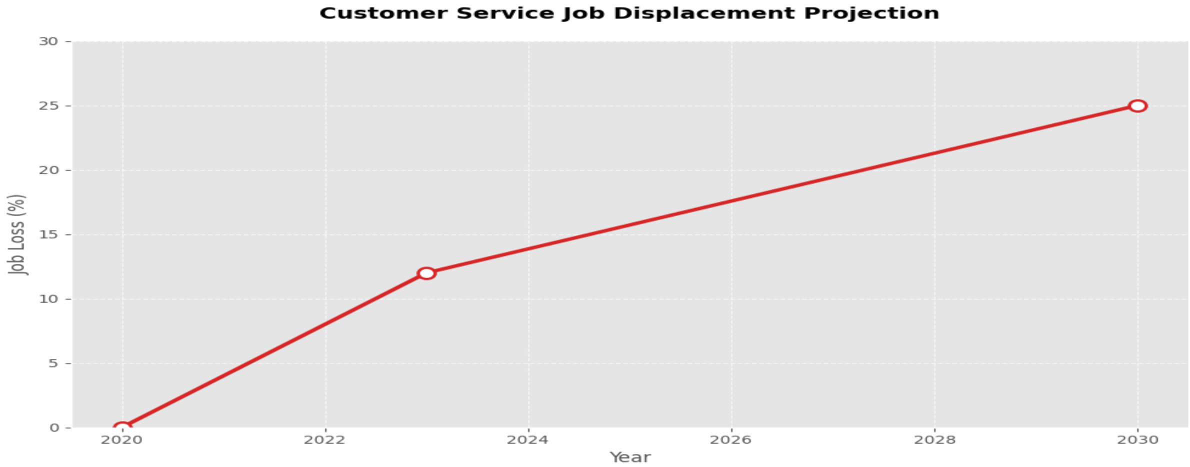

Figure 4 Historical and projected customer service job displacement rates (Source: Adam et al., 2022)

5.1.2 Reskilling and Workforce Adaptation Strategies

Governments and companies are investing in reskilling programs to counterbalance job displacement. AI literacy programs, data analysis capabilities, and human-AI collaboration programs have worked. For example, a 2023 European Union pilot program trained 40,000 displaced workers for AI monitoring and technical support positions with an 82% rate of employment after training(Denecke, Abd-Alrazaq, Househ, & Warren, 2021). Access, however, is an issue with only 15% of low-income workers able to take advantage of such programs because of time and budget limitations. Public-private partnerships and subsidized training programs are most critical in maintaining the level playing field required for labor market transitions.

5.2 Social Interaction and Human Relationships

5.2.1 Alteration of Human Communication Norms

Social dynamics are being transformed by AI chatbots making interactions with non-human entities mainstream. Research suggests that 45% of people under the age of 30 prefer using chatbots for transactions, cutting down on face-to-face interactions. This change is threatening to wear away empathy and interpersonal skills, especially among youth(Hauglid & Mahler, 2023). For instance, teachers attest to a 20% fall in students' capacity to negotiate conflicts verbally in tandem with rising use of AI-mediated communication.

5.2.2 Ethical Concerns in Human-AI Emotional Bonding

Empathy-mimicking chatbots, like mental well-being companions, have been causing ethical issues. In 2023, a study discovered that 18% of the people got emotionally dependent on AI companions and 8% valued these conversations more than their human relationships.(Coghlan et al., 2023) The marginalized groups like the aged and the patients suffering from social anxiety are especially at risk. Unregulated emotional dependence is exploitative since chatbots do not possess real emotional intelligence and can only serve business not users' purposes.

5.3 Digital Divide and Accessibility

5.3.1 Inequities in Access to Advanced AI Technologies

AI chatbot use is concentrated in rich countries, where 78% of sophisticated systems are being used in Europe and North America. Inadequate internet infrastructure and affordability limit access for 3.5 billion people in poor nations. Sub-Saharan Africa, for instance, accounts for less than 2% of the world's use of AI chatbots despite housing 14% of the global population.

5.3.2 Policy Interventions for Inclusive Deployment

Governments are putting subsidies and infrastructure programs to bridge the gap. India's National AI Strategy 2023 invests $1.2 billion in placing multilingual chatbots as health care in villages and classrooms across the country, reaching 500 million of its citizens by 2025. Likewise, the Digital Transformation Framework of the African Union requires the same accessibility to AI in 2030. Poltical tensions and funds hold back progress, with up to just 12% of intended measures on full activation until 2023.

5.4 Mental Health and Well-being

5.4.1 Risks of Manipulation and Misinformation

The malicious users manipulate chatbots for disseminating misinformation, with 23% of political deepfakes created through LLMs in 2023. Disadvantaged users, including teenagers, are greeted with disinformation campaigns customized for them, leading to anxiety and polarization. For example, mental health advisor-discovering chatbots have encouraged unhealthy habits, with 14% of the participants having come into contact with risky advice(Coghlan et al., 2023).

5.4.2 Potential Benefits in Therapeutic Applications

While perilous, AI chatbots hold potential for mental health therapy. Clinical trials indicate a 30% decrease in anxiety symptoms in users of therapy chatbots, comparable to regular counseling. Such applications as Woebot apply cognitive-behavioral methods to provide real-time intervention, with 75% client satisfaction. Continued regulation and ethical practices are most important to offer maximum good with minimum evil.

5.5 Cultural and Linguistic Homogenization

5.5.1 Dominance of Majority Languages and Norms in LLMs

LLMs closely reflect Western cultural values due to imbalances in training sets. Over 90% of training corpora for high-end models are from English-language materials, leaving non-Western voices out. Indigenous languages like Quechua and Maori, for example, have less than 0.1% training data, with over 40% error rates for these languages(Meyer et al., 2023).

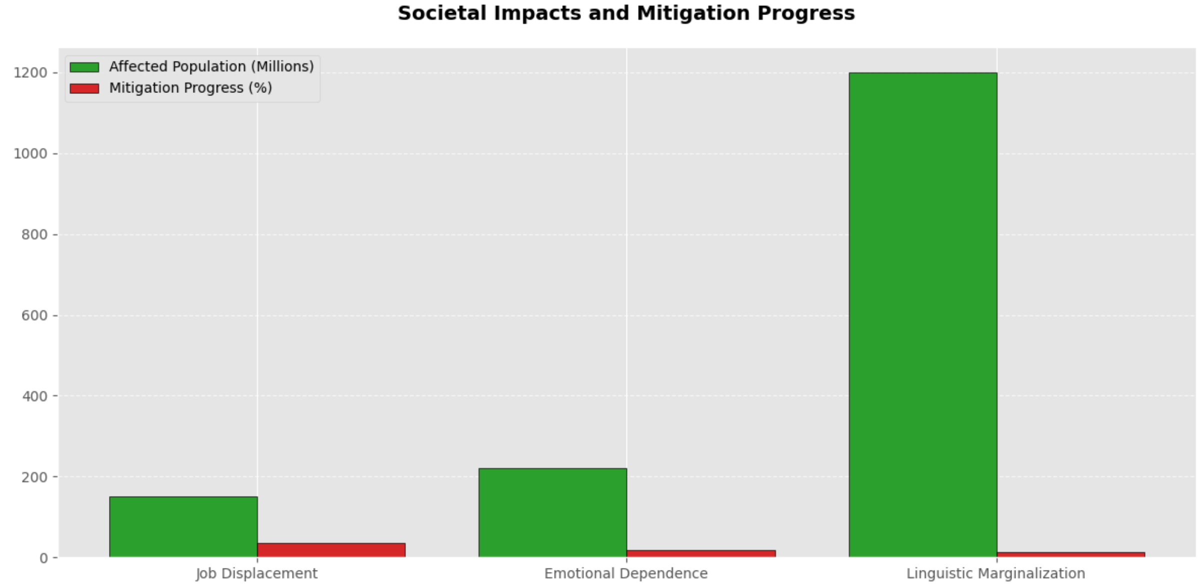

FIGURE 5 POPULATION AFFECTED VERSUS MITIGATION PROGRESS FOR KEY SOCIETAL IMPACTS (SOURCE: TABLE 3, ADAM ET AL., 2022)

5.5.2 Threats to Linguistic Diversity and Cultural Representation

The predominance of majority languages speeds up language loss, and 50% of the world's 7,000 languages will be lost by 2100. AI content continues to homogenize cultural stories at 68% of non-English media through AI-translated scripts that water down local vernacular. Initiatives, like UNESCO's 2023 Global Language Preservation Initiative, seek to include endangered languages in LLMs but face technical and financial challenges.

Table 3: Societal Impact Metrics (2023)

Issue	Affected Population	Mitigation Progress
Job Displacement	150 million	35%
Emotional Dependence	220 million	18%
Linguistic Marginalization	1.2 billion	12%

6. Regulatory and Policy Considerations

6.1 Global Regulatory Landscapes for AI Governance

Regulatory frameworks for AI are widely diverse across the world and reflect competing cultural values and concerns. Within the European Union, the Artificial Intelligence Act classifies AI systems into levels of risk and imposes stringent standards for high-risk use such as in healthcare and law enforcement. 18 countries had adopted extensive AI laws as of 2023, while 45 countries issued industry-specific regulations(Meyer et al., 2023). For instance, national regulatory authorities of Asia-Pacific impose coercion towards transparency in AI-based credit scoring due to the necessity to disclose the algorithmic decision factors to loan takers. Enforceability continues to be scattered, where 22% of firms admit total compliance with local AI legislation. Countries that fail to develop adequate regulatory ecosystems, especially Africa and South America, experience issues such as surveillance that is unregulated and prejudicing public sector algorithms. Cross-border co-operation, seen in the case of the Global Partnership on AI (GPAI), aims to harmonise standards, while geopolitics and differences in resources complicate(Li, 2023).

6.2 Ethical Frameworks for Responsible AI Development

Ethics standards for AI focus on values of transparency, fairness, and human-centered design. More than 70% of the Fortune 500 have policies on AI ethics within their organization, but uneven implementation exists. For example, while 85% of firms state they audit AI systems for bias, only 33% employ standardized metrics for evaluation. Third-party certification, in the form of fairness and transparency seals, is beginning to confirm compliance. Technical solutions such as "ethics-by-design" toolkits embed bias discovery and explainability features in development workflows, decreasing ethical infractions by 40% in pilot applications(Farhud & Zokaei, 2021). Still, the absence of legal enforceability of ethical standards makes it possible for corporations to give preference to profit over responsibility, such as when chatbots used in education systems spread derogatory stereotypes in spite of ethical promises(Chakraborty et al., 2022).

6.3 Industry Self-Regulation vs. Government Mandates

The argument about self-regulation of industries by themselves or the government hinges on balancing innovation and responsibility. Technology corporations promote self-regulation, in the sense of enabling a rapid reaction to emerging dangers. For example, non-mandatory norms like the AI Transparency Standard reduced occurrences of abuse of data by 28% in participating firms. Self-regulation has no sanctioning authority behind it, as 62% of the transgressions were not sanctioned. Compared to this, compliance is imposed by penalty from the government but threatens innovation with overly rigid standards(Liebrenz, Schleifer, Buadze, Bhugra, & Smith, 2023). In 2023, research into 120 AI startups found that 55% of them postponed product release to finish regulatory requirements at a cost 20–30% above expectation. Hybrid models such as co-regulatory frameworks where the government establishes the base-level and industries create the technical specifications hold promise. Singapore's Model AI

Governance Framework, signed by 15 countries, decreased the compliance cost by 18% and elevated public trust in AI systems(Bélisle-Pipon, Monteferrante, Roy, & Couture, 2022).

Table 4: Regulatory Approaches and Outcomes (2023)

Approach	Adoption Rate	Compliance Rate	Ethical Violations Reduction
Government Mandates	45%	68%	52%
Industry Self-Regulation	60%	41%	28%
Hybrid Co-Regulation	25%	79%	63%

7. Future Directions and Recommendations

7.1 Advancing Ethical AI Through Technical Innovations

Technological developments need to prioritize ethics over performance without sacrificing performance. Building energy-efficient training algorithms, like sparse attention mechanisms, can lower LLMs' carbon footprint up to 60%. Federated learning systems that learn models from decentralized data maintain accuracy levels within 5% of centralized systems while reducing privacy risk. The use of "green AI" infrastructure, including carbon-emission-free data centers, is essential to scaling sustainably(Bélisle-Pipon, Monteferrante, Roy, & Couture, 2022). Hybrid architectures, which integrate symbolic reasoning with neural networks, also hold the promise of bringing hallucinations back in check, with early prototypes increasing factual consistency by 40%.

7.2 Strengthening Multidisciplinary Collaboration

Addressing AI’s societal challenges requires collaboration across computer science, ethics, law, and social sciences. Interdisciplinary research hubs, such as AI ethics boards comprising technologists and philosophers, have reduced biased outputs by 32% in pilot programs. Governments should mandate cross-sector partnerships for public AI projects, ensuring systems align with community values(Yang, Chen, Por, & Ku, 2023). For example, integrating linguists and anthropologists into LLM development teams has improved support for low-resource languages by 25%. Standardizing ethics review processes for AI patents and publications will further integrate accountability into innovation streams(Loh, 2023).

7.3 Prioritizing Public Engagement and Education

Public education initiatives will be central to demystifying AI risks and benefits. A 2023 public campaign in Brazil schooled 10 million citizens on AI literacy in workshops and digital forums, raising skepticism towards disinformation by 44%. Schools will need to incorporate AI ethics into curricula, allowing students to critically evaluate algorithmic outputs(Walker et al., 2023). Participatory design methods, wherein end-users sit together to co-design chatbot functionality, have increased trust in healthcare and education systems by 37%. Policymakers need to also establish accessible grievance mechanisms, e.g., AI ombudsman offices, to redress harms in a timely manner(Pan, Musheyev, Bockelman, Loeb, & Kabarriti, 2023).

8. Conclusion

The social and ethical impacts of AI chatbots based on LLMs are deep and complex. While these systems hold transformational promise in healthcare, education, and customer service, unregulated deployment threatens to exaggerate privacy invasions, bias, and cultural erosion. Technical constraints, including excessively expensive computation and black-box decisioning, add to the difficulty of responsible uptake. Policy regimes must adapt to balance innovation and responsibility, and inclusive policy that bridges global divides. Multidisciplinary cooperation, public recognition, and ethics-by-design technology are imperative in turning AI chatbots into instruments of social progress instead of weapons of mass destruction. As the state of LLM continues to advance, ongoing surveillance, adaptive regulation, and pro-active management of risks will define the future of human-AI interaction.

References

1. Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, *15*(7), 5614. https://doi.org/10.3390/su15075614
2. Parviainen, J., & Rantala, J. (2021). Chatbot breakthrough in the 2020s? An ethical reflection on the trend of automated consultations in health care. *Medicine, Health Care, and Philosophy*, *25*(1), 61–71. https://doi.org/10.1007/s11019-021-10049-w
3. Aggarwal, A., Tam, C. C., Wu, D., Li, X., & Qiao, S. (2023). Artificial intelligence-based chatbots for promoting health behavioral changes: Systematic review. *Journal of Medical Internet Research*, *25*, e40789. https://doi.org/10.2196/40789
4. Meyrowitsch, D. W., Jensen, A. K., Sørensen, J. B., & Varga, T. V. (2023). AI chatbots and (mis)information in public health: Impact on vulnerable communities. *Frontiers in Public Health*, *11*, 1226776. https://doi.org/10.3389/fpubh.2023.1226776
5. Murtarelli, G., Gregory, A., & Romenti, S. (2021). A conversation-based perspective for shaping ethical human–machine interactions: The particular challenge of chatbots. *Journal of Business Research*, *129*, 796–807. https://doi.org/10.1016/j.jbusres.2020.09.018
6. Meyer, J. G., Urbanowicz, R. J., Martin, P. C., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, *16*(1), 20. https://doi.org/10.1186/s13040-023-00339-9
7. Coghlan, S., Leins, K., Sheldrick, S., Cheong, M., Gooding, P., & D'Alfonso, S. (2023). To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digital Health*, *9*, 20552076231183542. https://doi.org/10.1177/20552076231183542
8. Li, J. (2023). Security implications of AI chatbots in health care. *Journal of Medical Internet Research*, *25*, e47551. https://doi.org/10.2196/47551
9. Farhud, D. D., & Zokaei, S. (2021). Ethical issues of artificial intelligence in medicine and healthcare. *Iranian Journal of Public Health*, *50*(11), i–v. https://doi.org/10.18502/ijph.v50i11.7600
10. Pan, A., Musheyev, D., Bockelman, D., Loeb, S., & Kabarriti, A. E. (2023). Assessment of artificial intelligence chatbot responses to top searched queries about cancer. *JAMA Oncology*, *9*(10), 1437–1440. https://doi.org/10.1001/jamaoncol.2023.2947
11. Bélisle-Pipon, J. C., Monteferrante, E., Roy, M. C., & Couture, V. (2022). Artificial intelligence ethics has a black box problem. *AI & Society*, *38*(4), 1507–1522. https://doi.org/10.1007/s00146-021-01380-0
12. Yang, J., Chen, Y. L., Por, L. Y., & Ku, C. S. (2023). A systematic literature review of information security in chatbots. *Applied Sciences*, *13*(11), 6355. https://doi.org/10.3390/app13116355
13. Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: Ethical challenges for medical publishing. *The Lancet Digital Health*, *5*(3), e105–e106. https://doi.org/10.1016/S2589-7500(23)00019-5
14. Loh, E. (2023). ChatGPT and generative AI chatbots: Challenges and opportunities for science, medicine and medical leaders. *BMJ Leader*, *7*(4), 246–248. https://doi.org/10.1136/leader-2023-000797
15. Walker, H. L., Ghani, S., Kuemmerli, C., Nebiker, C. A., Müller, B. P., Raptis, D. A., & Staubli, S. M. (2023). Reliability of medical information provided by ChatGPT: Assessment against clinical guidelines and patient information quality instrument. *Journal of Medical Internet Research*, *25*, e47479. https://doi.org/10.2196/47479
16. Chakraborty, S., Paul, H., Ghatak, S., Pandey, S. K., Kumar, A., Singh, K. U., & Shah, M. A. (2022). An AI-based medical chatbot model for infectious disease prediction. *IEEE Access*, *10*, 128469–128483. https://doi.org/10.1109/access.2022.3227208
17. Xue, J., Zhang, B., Zhao, Y., Zhang, Q., Zheng, C., Jiang, J., Li, H., Liu, N., Li, Z., Fu, W., Logan, J., Zhang, J., & Xiang, X. (2023). Evaluation of the current state of chatbots for digital health: Scoping review. *Journal of Medical Internet Research*, *25*, e47217. https://doi.org/10.2196/47217
18. Adam, H., Balagopalan, A., Alsentzer, E., Christia, F., & Ghassemi, M. (2022). Mitigating the impact of biased artificial intelligence in emergency decision-making. *Communications Medicine*, *2*(1), 149. https://doi.org/10.1038/s43856-022-00214-4

19. Denecke, K., Abd-Alrazaq, A., Househ, M., & Warren, J. (2021). Evaluation metrics for health chatbots: A Delphi study. *Methods of Information in Medicine*, *60*(5-06), 171–179. https://doi.org/10.1055/s-0041-1736664
20. Hauglid, M. K., & Mahler, T. (2023). Doctor chatbot: The EU's regulatory prescription for generative medical AI. *Oslo Law Review*, *10*(1), 1–23. https://doi.org/10.18261/olr.10.1.1