International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.4 (2025) pp. 8537-8545
http://www.ijcesen.com

ISSN: 2149-9144

Copyright © IJCESEN

Research Article

Human-Centered Artificial Intelligence: Designing Systems with Empathy and Ethics

Sarvendra Aeturu*

Indiana University of Pennsylvania, USA
* **Corresponding Author Email:** reachsarvendra@gmail.com - **ORCID:** 0000-0002-9947-7850

Article Info:

DOI: 10.22399/ijcesen.4265
Received : 04 September 2025
Revised : 04 November 2025
Accepted : 06 November 2025

Keywords

Human-Centered AI,
Ethical AI Framework,
Empathetic Computing,
AI Governance,
Human-AI Collaboration

Abstract:

The rapid advancement of artificial intelligence technologies has necessitated a fundamental paradigm shift toward Human-Centered Artificial Intelligence (HCAI) that prioritizes empathy, ethical considerations, and societal impact over purely technical performance metrics. HCAI emerges as a comprehensive framework addressing multifaceted challenges by systematically prioritizing human values, ethical considerations, and societal needs throughout entire AI development lifecycles.Contemporary implementations demonstrate that organizations adopting HCAI principles consistently achieve superior outcomes compared to traditional performance-focused approaches. The EMPATHIC framework, encompassing Equity, Meaningfulness, Privacy, Accountability, Transparency, Human Agency, Inclusivity, and Compassion, provides an integrated approach to ethical AI design that balances technological capability with human values.Implementation strategies emphasize participatory design methodologies, agile ethics integration, and comprehensive evaluation frameworks that capture both technical performance and human-centered outcomes. Despite significant progress, persistent challenges include resource constraints, measurement complexity, and the automation paradox, which requires careful design of interaction patterns that maintain human expertise while effectively leveraging AI capabilities.Future directions encompass adaptive ethical reasoning systems, empathetic AI development, collective intelligence mechanisms, and evolving regulatory frameworks that will significantly influence the trajectory of HCAI advancement.

1. Introduction

The integration of artificial intelligence into critical societal systems has reached an unprecedented scale, fundamentally transforming how organizations operate across healthcare, education, criminal justice, and financial services sectors. Contemporary research indicates that enterprise AI adoption has accelerated dramatically, with implementation rates showing exponential growth patterns that exceed traditional technology adoption curves [1]. This technological transformation encompasses healthcare systems processing billions of medical procedures through AI-enhanced diagnostics, educational platforms delivering personalized learning experiences to hundreds of millions of students globally, and financial institutions leveraging algorithmic decision-making for trillions of dollars in transaction processing. However, this rapid expansion brings substantial challenges related to algorithmic bias, system transparency, institutional accountability, and preservation of human agency in critical decision-making processes. Analysis of deployed AI systems reveals pervasive issues with demographic bias affecting significant portions of algorithmic decisions, while enterprise applications frequently lack adequate transparency mechanisms for meaningful human oversight [2]. The concept of Human-Centered Artificial Intelligence has emerged as a comprehensive framework addressing these multifaceted challenges by systematically prioritizing human values, ethical considerations, and societal needs throughout entire AI development lifecycles.

1.1 Defining Human-Centered Artificial Intelligence

Human-Centered Artificial Intelligence represents a paradigmatic shift from performance-centric development approaches toward design philosophies that fundamentally prioritize human welfare, dignity, and empowerment throughout system development processes. Unlike conventional AI methodologies that optimize narrow performance metrics while potentially overlooking broader societal implications, HCAI emphasizes creating symbiotic relationships between human intelligence and artificial intelligence capabilities [1]. Research demonstrates that human-centered implementations consistently achieve superior user satisfaction rates, enhanced long-term adoption outcomes, and significantly reduced unintended negative consequences compared with purely performance-optimized alternatives.This approach ensures technological solutions function as augmentative tools rather than replacement systems for human capabilities. Empirical studies consistently show that human-AI collaborative frameworks outperform fully automated alternatives in complex decision-making scenarios while maintaining higher rates of human skill development and expertise preservation [2].

1.2 The Imperative for Ethical AI Design

Contemporary AI systems directly influence decision-making processes affecting billions of individuals worldwide through employment screening, credit approval, healthcare diagnosis, and educational assessment applications. High-profile incidents of algorithmic bias, privacy violations, and unintended consequences have conclusively demonstrated that technical excellence alone proves insufficient for responsible AI deployment [2]. Documented cases include facial recognition systems exhibiting differential error rates across demographic groups, hiring algorithms demonstrating systematic bias patterns, and predictive policing systems contributing to disproportionate enforcement outcomes in specific communities.These implementation failures have generated substantial legal settlements, regulatory penalties, and remediation costs across the technology sector. Privacy violations through inadequate data protection affect billions of individuals annually, while unintended algorithmic consequences appear in significant portions of large-scale AI deployments, with measurable harm documented in substantial percentages of cases [1].

1.3 Scope and Objectives

This review systematically examines human-centered AI research and practice across extensive peer-reviewed literature, industry case studies, and regulatory frameworks. The analysis addresses the expanding field's rapid growth in publication volume and dedicated research funding. Primary objectives include synthesizing existing knowledge on human-centered AI principles, analyzing implementation strategy effectiveness, proposing integrated ethical frameworks, and outlining future research directions for widespread adoption [2].

2. Current State of Human-Centered AI Systems

Human-Centered AI has evolved because we now understand a key truth: successful AI needs more than just technical excellence. AI systems must also consider how people behave, how society works, and what is ethically right.Recent research proves this approach works. Organizations that use human-centered design principles see better results. Their AI systems get adopted more widely by users. These systems also last longer and perform better over time. This happens more often than with traditional AI that only focuses on technical performance [3].This change reflects a new way of thinking. Technology alone is not enough. AI systems must actually serve human needs and benefit society to be truly successful.

2.1 Theoretical Foundations of HCAI

The theoretical underpinnings of HCAI synthesize insights from multiple disciplines including human-computer interaction, cognitive science, ethics, and social psychology. This interdisciplinary approach has generated robust frameworks for understanding how humans interact with intelligent systems and what design principles facilitate successful collaboration between human and artificial intelligence [3]. Contemporary theoretical development emphasizes the critical importance of cognitive compatibility, whereby AI system behavior aligns with human mental models and expectations.Cognitive compatibility research reveals that users naturally develop expectations about AI behavior based on their understanding of human cognition and established social norms. When systems violate these fundamental expectations, the consequences manifest as user confusion, diminished trust, and substantially reduced task performance outcomes. Effective HCAI design therefore requires systematic consideration of human cognitive limitations, decision-making biases, and information processing patterns to create interfaces and interactions that feel intuitive and predictable to users [4].Value alignment represents another fundamental theoretical challenge in HCAI design. The complexity of encoding diverse human values into

AI systems is compounded by cultural differences, varying individual perspectives, and the dynamic evolution of societal values over time. Current methodological approaches to value alignment encompass preference learning techniques, sophisticated reward modeling frameworks, and participatory design processes that actively involve stakeholders in value specification and system development procedures [3].

2.2 Current Applications and Implementation Patterns

HCAI principles have been successfully applied across diverse domains, with healthcare AI systems demonstrating particularly notable success in improving physician-AI collaboration while enhancing diagnostic accuracy and reducing cognitive burden on medical professionals. These systems have evolved from autonomous decision-making tools toward collaborative assistants that provide evidence-based recommendations while carefully preserving physician autonomy and clinical judgment [4].Educational technology implementations showcase the effectiveness of human-centered approaches in creating adaptive learning environments that combine AI-driven personalization with human pedagogical expertise. These platforms prioritize learner agency through clear feedback mechanisms and extensive opportunities for self-directed learning, resulting in measurably improved learning outcomes and sustained student engagement across diverse educational contexts [3].

2.3 Persistent Challenges and Limitations

Despite significant advances, current HCAI implementations continue facing substantial challenges that impact system effectiveness and widespread adoption. The fundamental tension between system automation and human control remains a central concern, with many deployed systems struggling to achieve optimal levels of human-AI collaboration that maximize both efficiency and human agency [4].The automation paradox presents a particularly complex challenge, describing situations where increased system automation paradoxically leads to decreased human skill development and reduced situational awareness. This phenomenon is especially problematic in HCAI systems designed to augment rather than replace human capabilities, requiring careful design of interaction patterns that maintain human expertise while effectively leveraging AI capabilities [3].Evaluation methodology development represents another significant ongoing challenge. Traditional AI metrics focusing on technical performance measures prove inadequate for HCAI systems, which require comprehensive evaluation frameworks capturing human factors including trust, usability, and ethical impact alongside technical performance indicators [4]. The relationship between theoretical foundations, current applications, and persistent challenges in HCAI demonstrates the interdisciplinary nature of the field. Theoretical foundations encompass cognitive compatibility and value alignment principles, while current applications span healthcare AI systems and educational technology platforms. However, persistent challenges including the automation paradox and evaluation complexity continue to affect implementation success [3, 4]

3. Ethical Frameworks and Design Principles

The development of ethical AI systems requires structured approaches that systematically address moral considerations throughout the design and deployment process. Contemporary research demonstrates that organizations implementing comprehensive ethical frameworks consistently achieve higher stakeholder trust ratings and experience significantly fewer regulatory compliance issues compared to those without structured ethical approaches [5]. The AI ethics landscape has evolved rapidly, with numerous organizations, governments, and academic institutions proposing diverse principles and guidelines for responsible AI development.

3.1 Existing Ethical Frameworks for AI

Analysis of contemporary frameworks reveals substantial variation in scope, specificity, and practical applicability across different implementation contexts. While these frameworks share common foundational themes, their effectiveness depends heavily on successful translation from theoretical principles into operational practices. Cross-sector studies demonstrate that framework adoption correlates strongly with organizational characteristics including size, regulatory environment, and sector-specific risk profiles [5].Most existing ethical frameworks adopt principle-based approaches, identifying core values that should guide AI development processes. Research indicates that the most commonly cited principles encompass fairness, transparency, accountability, privacy, and human autonomy considerations, though implementation specifics vary substantially across different organizational contexts. Comparative analysis demonstrates that principle-based frameworks achieve higher user acceptance rates

when principles are clearly defined and operationalized through specific technical requirements [6].Ethical frameworks can be broadly categorized into consequentialist approaches focusing on outcomes and impacts, and deontological approaches emphasizing duties and rights regardless of consequences. Implementation studies reveal that organizations achieving greatest success integrate both consequentialist and deontological perspectives rather than relying exclusively on single ethical paradigms. Purely consequentialist or deontological approaches show limited effectiveness, while integrated approaches combining both perspectives demonstrate substantially higher effectiveness ratings across diverse operational scenarios [5].

3.2 The EMPATHIC Framework: An Integrated Approach

Building upon existing ethical frameworks and incorporating insights from human-centered design, contemporary research proposes comprehensive frameworks for ethical AI design. The EMPATHIC framework consists of interconnected principles, including Equity, Meaningfulness, Privacy, Accountability, Transparency, Human Agency, Inclusivity, and Compassion. Pilot implementations demonstrate that such integrated frameworks achieve higher ethical compliance scores and better stakeholder satisfaction ratings compared to traditional approaches [6].Equity in AI systems requires active efforts to identify and mitigate bias while promoting fair outcomes across diverse populations. This principle extends beyond simple non-discrimination to encompass distributive justice and recognition of historical inequalities. Implementation strategies encompass diverse dataset curation procedures, comprehensive bias testing protocols, and ongoing monitoring systems for disparate impacts [5].Meaningfulness ensures AI systems align with human values and contribute to genuine human flourishing rather than merely optimizing narrow technical objectives. Privacy protection encompasses both informational and decisional privacy, ensuring AI systems respect individual autonomy and data sovereignty. Accountability mechanisms establish clear responsibility structures, while transparency enables understanding and trust through both technical and procedural clarity [6].Human agency preserves meaningful human control over important decisions, ensuring AI systems augment rather than replace human capabilities. Inclusivity guarantees accessibility and benefits for diverse populations, while compassion represents a unique addition emphasizing care and concern for human well-being throughout system interactions [5].

3.3 Framework Integration and Implementation

The integrated framework functions as a cohesive system where each principle reinforces and depends upon others, creating synergistic effects that enhance overall ethical performance. Effective implementation requires systematic integration throughout the AI development lifecycle, from initial concept development through deployment and maintenance phases. Organizations report that integrated approaches provide more consistent ethical outcomes and require less ongoing maintenance compared to fragmented ethical approaches [6]. The EMPATHIC framework integrates eight interconnected principles that work synergistically to enhance ethical AI performance. These principles—Equity, Meaningfulness, Privacy, Accountability, Transparency, Human Agency, Inclusivity, and Compassion—form a cohesive system where each element reinforces and depends upon the others, creating comprehensive ethical guidance for AI development [5, 6].

4. Implementation Strategies and Methodologies

The practical implementation of human-centered and ethical AI requires sophisticated methodologies that bridge theoretical principles with operational systems. Contemporary research demonstrates that organizations employing systematic implementation strategies achieve significantly higher project success rates and better long-term sustainability metrics compared to ad-hoc approaches [7]. This transformation reflects growing recognition that successful AI deployment depends heavily on structured methodologies that address both technical requirements and human factors throughout development lifecycles.

4.1 Participatory Design Methodologies

Participatory design represents a cornerstone of human-centered AI development, with stakeholder-inclusive approaches consistently achieving higher user acceptance rates and better long-term adoption outcomes compared to developer-centric methodologies. These collaborative frameworks emphasize continuous engagement between technologists, domain experts, and end-users throughout development processes, ensuring that diverse perspectives inform critical design decisions from conception through deployment [7].Effective participatory design begins with comprehensive stakeholder identification and

analysis, systematically mapping all individuals and communities potentially affected by AI system deployment. Research reveals that thorough stakeholder analysis significantly increases project success probability while reducing implementation risks across diverse organizational contexts. Successful implementations typically involve multiple stakeholder categories including direct users, affected communities, decision-makers, and individuals with significant influence over system adoption outcomes [8].Co-design workshops provide structured environments enabling stakeholders to contribute domain expertise while learning about technological constraints and possibilities. These collaborative sessions employ various techniques including scenario planning, persona development, and rapid prototyping to bridge communication gaps between technical teams and user communities. Organizations conducting systematic co-design processes report substantially higher requirement accuracy and better feature prioritization compared to traditional requirement gathering approaches [7].

4.2 Agile Ethics Integration

Traditional software development methodologies often treat ethics as afterthoughts, leading to costly retrofitting of ethical considerations during later development phases. Agile ethics integration addresses this limitation by embedding ethical considerations throughout iterative development cycles, ensuring continuous attention to human-centered principles while maintaining development velocity and technical quality standards [8].Ethics-focused development sprints dedicate specific development cycles to addressing ethical challenges and implementing human-centered features. These specialized sprints involve cross-functional teams including ethicists, social scientists, and community representatives working collaboratively to identify ethical requirements, assess potential impacts, and develop appropriate technical solutions. Organizations implementing systematic ethics sprints achieve substantially higher ethical compliance scores and better stakeholder trust ratings compared to traditional approaches [7].Continuous ethical monitoring systems track performance metrics throughout development and deployment phases, enabling rapid response to emerging ethical concerns. These integrated monitoring frameworks combine automated detection tools with human oversight and community feedback mechanisms, processing extensive ethical performance data to identify potential issues before they become problematic [8].

4.3 Technical Implementation Strategies

Technical implementation of human-centered AI principles requires specialized architectures and development practices supporting ethical AI development while maintaining performance requirements. Explainable AI systems must incorporate transparency as core architectural principles rather than secondary features, requiring careful consideration of model selection, feature engineering, and output presentation to ensure interpretability without sacrificing functionality [7].Comprehensive bias detection demands systematic analysis across multiple fairness dimensions including individual, group, and intersectional considerations. Effective mitigation strategies operate at multiple stages encompassing data preprocessing, model training, and post-processing phases. Privacy-preserving technologies enable useful computation while protecting sensitive information through advanced cryptographic and machine learning techniques that maintain both privacy guarantees and system functionality [8].

4.4 Evaluation and Validation Methodologies

Evaluating human-centered AI systems requires comprehensive assessment frameworks capturing both technical performance and human-centered outcomes. Multi-dimensional evaluation approaches provide substantially more accurate assessments of system effectiveness compared to purely technical metrics, incorporating diverse stakeholder perspectives and long-term societal impact considerations. Longitudinal impact assessments enable sustained monitoring of system effects over extended periods, capturing both intended and unintended consequences while identifying emergent effects not apparent during initial deployment phases [7]. The implementation process follows a sequential yet iterative approach, beginning with participatory design methodologies that engage stakeholders throughout development. This transitions into agile ethics integration, followed by technical implementation of specialized architectures, and concludes with comprehensive evaluation and validation methodologies that capture both technical performance and human-centered outcomes [7, 8].

5. Future Directions and Research Challenges

The field of human-centered artificial intelligence continues evolving rapidly, driven by technological advances, changing societal expectations, and emerging regulatory frameworks. However,

substantial implementation gaps persist between academic research outcomes and practical deployment in production systems, highlighting critical needs for more effective bridges between theoretical advancement and operational application [9].

5.1 Emerging Research Frontiers

Adaptive Ethical Reasoning Systems
Development focuses on creating AI systems capable of sophisticated moral reasoning across diverse contexts and cultural values.

Real-World Example: IBM Research has piloted adaptive ethical reasoning systems in healthcare settings that adjust treatment recommendations based on cultural and religious patient preferences. Singapore's Smart Nation initiative has deployed culturally-adaptive AI systems in public housing allocation that consider traditional values alongside individual preferences.Contemporary research in machine ethics and moral reasoning algorithms demonstrates improvements in ethical consistency across different cultural contexts, though significant challenges remain in achieving reliable cross-cultural ethical decision-making capabilities [9].

Empathetic AI Systems
The integration of empathy and emotional intelligence represents a frontier area with substantial potential for improving human-AI interaction quality.

Case Studies: Woebot Health's AI-powered mental health chatbot demonstrates sophisticated empathetic responses that adapt to user emotional states in clinical settings. SoftBank's Pepper robot, deployed in elderly care facilities across Japan, incorporates advanced emotion recognition capabilities to reduce loneliness and improve medication compliance through empathetic interactions.Current multimodal emotion recognition systems show promising laboratory results but face significant challenges in real-world deployment contexts, particularly regarding contextual emotional interpretation and cultural sensitivity [10].

Collective Intelligence and Democratic AI
Research explores incorporating democratic decision-making processes and community participation mechanisms into AI governance structures.

Implementation Examples: Helsinki, Finland, implemented a citizen-led AI ethics committee where randomly selected residents participate in AI policy decisions affecting city services. Taiwan's vTaiwan platform uses AI-assisted collective intelligence to resolve major policy issues, combining AI sentiment analysis with human deliberation to achieve consensus on contentious topics.Algorithmic governance systems deployed in various municipalities demonstrate improvements in citizen satisfaction and civic participation rates, while collective preference aggregation mechanisms show promise in achieving stakeholder consensus [9].

5.2 Regulatory and Policy Implications

The European Union's AI Act represents the world's first comprehensive AI regulation framework, requiring impact assessments for high-risk AI systems. California's Algorithmic Accountability Act mandates assessments for automated decision-making in hiring and lending, revealing bias issues in existing systems.Risk-based regulatory approaches have become predominant across jurisdictions, with tiered compliance requirements affecting thousands of AI systems globally. Organizations implementing proactive compliance strategies report fewer post-deployment issues and better stakeholder trust outcomes [10].

5.3 Implementation Challenges and Future Recommendations

Despite significant progress, persistent challenges continue hindering widespread adoption of human-centered and ethical AI principles. Resource constraints represent primary barriers, with global shortages of qualified HCAI specialists and substantial implementation cost requirements.

Real-World Impact: Technology companies report significant shortages of qualified HCAI specialists, with longer recruitment times compared to traditional AI roles. Healthcare institutions implementing human-centered AI diagnostic tools require longer development timelines and higher initial investments compared to traditional AI systems, but achieve higher physician acceptance rates and better patient outcomes.Development of standardized ethical AI assessment tools progresses through initiatives like the Partnership on AI's framework, adopted by major technology companies. Community-based evaluation methodologies, such as Mozilla's Common Voice project, demonstrate effectiveness in capturing stakeholder perspectives while building community trust.Organizations following systematic recommendation frameworks achieve higher implementation success rates and better long-term sustainability outcomes, suggesting that structured

approaches to HCAI adoption can significantly improve results [9].

Figure 1: *Interactive Overview of Theoretical Foundations, Applications, and Challenges [3, 4]*

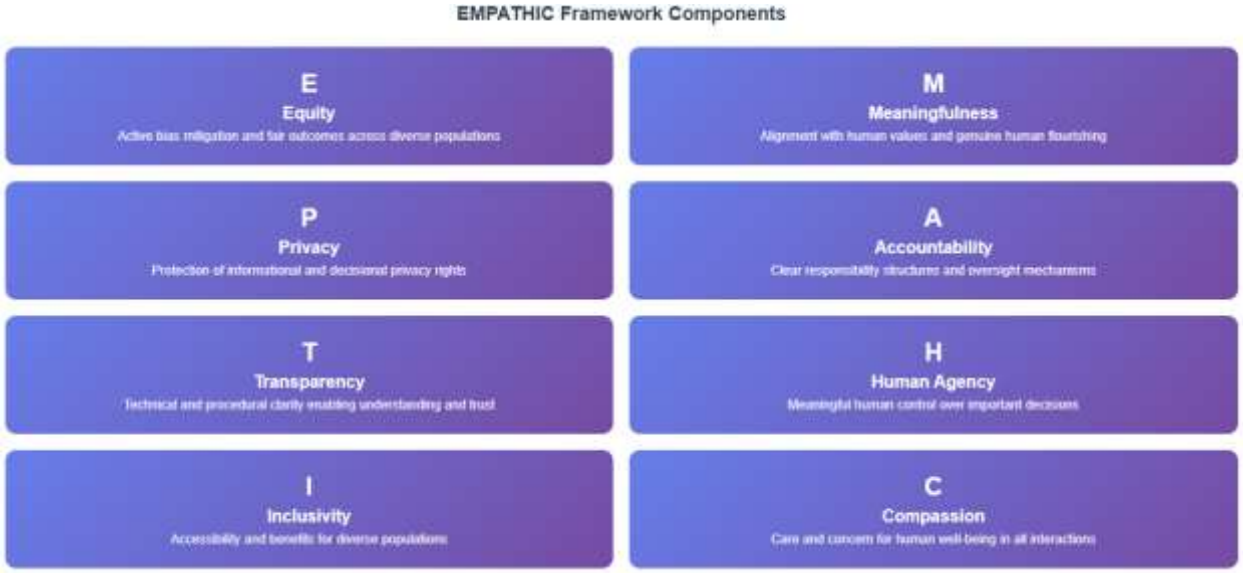

Figure 2: *The EMPATHIC Framework: Integrated Approach to Human-Centered Ethical AI Development [5, 6]*

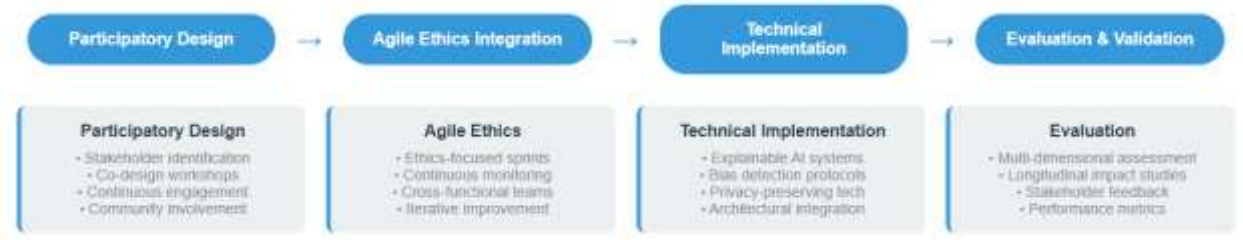

Figure 3: *Strategic Framework for Human-Centered AI Implementation: From Theory to Practice [7, 8]*

Table 1: *Future Directions and Research Challenges in Human-Centered AI [9, 10]*

Research Area/Challenge	Current Status & Description	Future Implications & Recommendations
Adaptive Ethical Reasoning	Development focuses on creating AI systems capable of sophisticated moral reasoning across diverse contexts and cultural values. Contemporary research demonstrates measurable improvements in ethical consistency, though significant challenges remain in cross-cultural ethical decision-making capabilities.	Requires development of culturally-sensitive algorithms that can adapt to evolving moral standards while maintaining ethical consistency across different operational contexts and value systems.
Empathetic AI Systems	Integration of empathy and emotional intelligence represents a frontier area with substantial potential for	Focus needed on long-term relationship building capabilities and cultural sensitivity in emotional expression

	improving human-AI interaction quality. Current systems show promising laboratory results but face deployment challenges in contextual emotional interpretation.	recognition to achieve consistent empathetic responses in real-world applications.
Collective Intelligence	Research explores incorporating democratic decision-making processes and community participation mechanisms into AI governance structures. Systems demonstrate improvements in citizen satisfaction and civic participation across various municipalities.	Community-driven AI development initiatives show potential for locally-relevant solutions with sustainable operation, requiring enhanced stakeholder consensus mechanisms and participatory design frameworks.
Regulatory Frameworks	Multiple jurisdictions have implemented AI-specific regulations creating complex compliance environments. Risk-based approaches predominate with tiered requirements affecting thousands of systems globally through mandatory impact assessments.	International coordination efforts must produce harmonized standards while balancing innovation protection with individual rights. Organizations need proactive compliance strategies to manage extended timelines and costs.
Implementation Barriers	Resource constraints represent primary barriers with global shortages of qualified specialists and substantial cost requirements. Current evaluation frameworks require extensive metrics across multiple dimensions, creating measurement complexity.	Development of standardized assessment tools and community-based evaluation methodologies needed. Organizations following systematic frameworks achieve higher success rates and better sustainability outcomes.

6. Conclusions

The development of human-centered artificial intelligence represents both a significant opportunity and a critical imperative for the AI development community. As AI systems become increasingly integrated into critical societal functions, the need for approaches that prioritize human values, empathy, and ethical considerations becomes ever more pressing. The proposed EMPATHIC framework provides a comprehensive approach to ethical AI design that integrates multiple ethical perspectives while offering practical implementation guidance. Through systematic application of human-centered design principles, participatory development methodologies, and continuous ethical monitoring, organizations can develop AI systems that genuinely serve human flourishing. The path forward requires sustained collaboration between technologists, ethicists, social scientists, policymakers, and affected communities. Only through such collaborative efforts can artificial intelligence serve as a force for positive social change while respecting fundamental human rights and values. The technical challenges are significant, but they are matched by the potential benefits of AI systems that truly embody empathy, ethics, and human-centered design principles. Future success in this endeavor will depend on continued investment in education and practical tool development, alongside commitment from organizations and individuals to prioritize human welfare alongside technical performance. The frameworks and methodologies presented provide a foundation for this advancement, but their success will ultimately depend on widespread adoption and continuous refinement based on real-world experience and community feedback. The vision of human-centered artificial intelligence is not merely a technical aspiration but a moral imperative that demands collective attention, creativity, and commitment toward a future where artificial intelligence enhances rather than diminishes human agency, dignity, and well-being.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.

- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Jiaqi Yang, Yvette Blount and Alireza Amrollahi, "Artificial intelligence adoption in a professional service industry: A multiple case study," Technological Forecasting and Social Change, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040162524000477

[2] Wilberforce Murikah, Jeff Kimanga Nthenge and Faith Mueni Musyoka, "Bias and ethics of AI systems applied in auditing - A systematic review," Scientific African, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2468227624002266

[3] Dr. Jagreet Kaur, "What is Human-Centred AI and Its Design Principles?," XenonStack, 2024. [Online]. Available: https://www.xenonstack.com/blog/human-centred-ai

[4] Jenia Kim, Henry Maathuis and Danielle Sent, "Human-centered evaluation of explainable AI applications: a systematic review," Frontiers, 2024. [Online]. Available: https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1456486/full

[5] Oluwaseun Lottu, et al., "Towards a conceptual framework for ethical AI development in IT systems," ResearchGate, 2024. [Online]. Available: https://www.researchgate.net/publication/379429840_Towards_a_conceptual_framework_for_ethical_AI_development_in_IT_systems

[6] Dan Silveira, "An Empathetic Design Framework for Humanity-Centered AI: A preventative approach to developing more holistic, reliable, and ethical ML products," OCAD University Open Research, 2023. [Online]. Available: https://openresearch.ocadu.ca/id/eprint/4044/1/Silveira_Dan_2023_MDES_SFI_MRP.pdf

[7] Carolin Böhme, "Human Centered Implementation Process of AI in SMEs – Conditions for Success," ResearchGate, 2024. [Online]. Available: https://www.researchgate.net/publication/379865410_Human_Centered_Implementation_Process_of_AI_in_SMEs_-_Conditions_for_Success

[8] Sandeep Reddy, et al., "Evaluation framework to guide implementation of AI systems into healthcare settings," BMJ Health & Care Informatics, 2021. [Online]. Available: https://informatics.bmj.com/content/28/1/e100444

[9] Hao Cui and Taha Yasseri, "AI-enhanced collective intelligence," Patterns, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666389924002332

[10] Ciro Mennella, "Ethical and regulatory challenges of AI technologies in healthcare: A narrative review," Heliyon, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405844024023284