Digital Government Authority
هيئة الحكومة الرقمية

Human-Centered AI

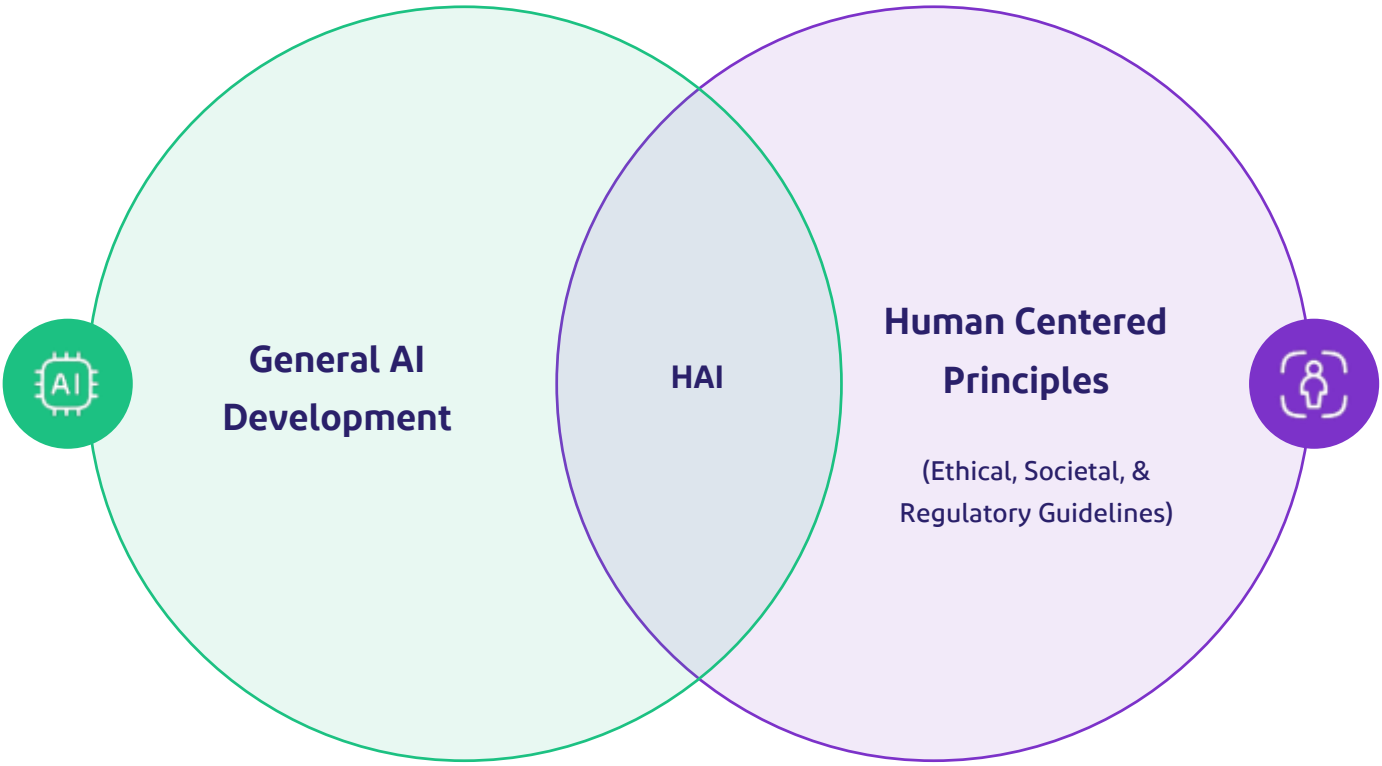
24 August 2025
Document Type: Brief Study
Document Classification: Public
Issue No.: 1.0

Contents

1. Introduction

Human-Centered AI (HAI) represents a transformative approach to artificial intelligence, distinguishing itself from traditional AI by prioritizing human values, ethical principles, and societal well-being. Unlike traditional AI, which focuses on automating tasks and optimizing technical performance, HAI seeks to complement human capabilities, enhance decision-making, and foster collaboration.

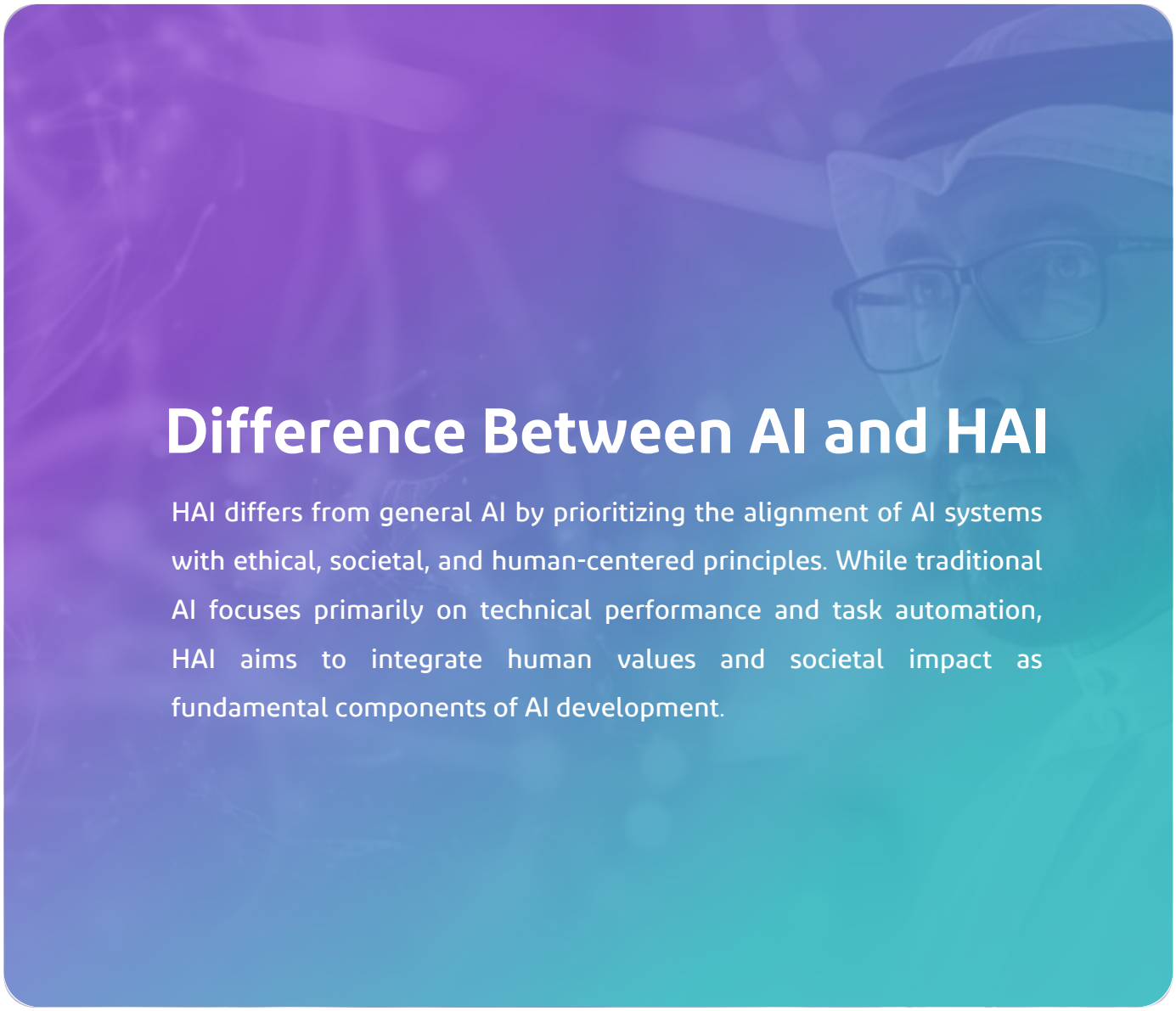
By integrating ethical reasoning, emotional intelligence, and societal alignment, HAI addresses critical challenges such as trust, transparency, and inclusivity. Its applications span sectors like healthcare, education, public services, and business, where ethical and human-centric AI systems can deliver transformative improvements in service delivery, productivity, and citizen engagement.



HAI also plays a pivotal role in transforming the public sector. By prioritizing ethical and equitable applications, HAI enhances public services and societal welfare. Unlike traditional AI systems that focus on automation, HAI fosters trust and inclusivity by complementing human expertise, supporting equitable access to resources, and driving innovation in governance, healthcare, and education. Its human-centric approach makes HAI indispensable for addressing societal challenges while improving public sector efficiency and citizen satisfaction.

In Saudi Arabia, HAI aligns closely with the Kingdom's ongoing efforts to integrate cutting-edge AI into critical sectors. Whether through personalized healthcare, AI-driven education programs, or advanced governance systems, HAI offers the potential to preserve cultural heritage, support equitable development, and foster innovation in line with national goals.

This study examines the significance and strategic applications of HAI, focusing particularly on its potential to drive societal and economic progress. It begins by defining HAI and presenting its conceptual representation and critical success factors, which outline a framework for responsible and effective deployment. The subsequent sections delve into the history of HAI, its global and local perspectives, and challenges related to its adoption. Finally, the study concludes with actionable recommendations for embedding HAI into Saudi Arabia's digital transformation strategy to ensure ethical, equitable, and sustainable growth.



Difference Between AI and HAI

HAI differs from general AI by prioritizing the alignment of AI systems with ethical, societal, and human-centered principles. While traditional AI focuses primarily on technical performance and task automation, HAI aims to integrate human values and societal impact as fundamental components of AI development.

2. Definition of Human-Centered AI (HAI)

Stanford founded the Institute for Human-Centered AI in 2019 to guide the future of AI development and deployment. Stanford HAI defines HAI as the design and use of AI systems that prioritize human values, enhance societal welfare, and promote ethical, equitable interactions between AI and society. The overarching goal of HAI is to ensure that AI technologies advance in a way that aligns with societal principles of fairness, transparency, and accountability.

In recent years, there has been growing recognition among global AI leaders that HAI is critical for building community trust in AI. Since its founding, Stanford's HAI Institute has distributed over $40 million in grants, supporting research that enhances human capabilities and furthers ethical AI development. This investment emphasizes the critical importance of advancing HAI as a key driver of AI's positive societal impact. By embedding human values in AI, organizations can ensure that these systems remain effective while serving individuals, communities, and institutions. This emphasis on ethical application is essential for addressing the societal impacts of AI and fostering responsible decision-making processes.

The development and implementation of HAI are driven by three main priorities, each playing a crucial role in ensuring AI technologies are designed to benefit society.

Key Priorities of HAI:

Human Impact

AI must positively influence individuals and societal structures, promoting equity and ethical outcomes.

Augmenting Human Capabilities

AI enhances human abilities in areas like healthcare and education, complementing rather than replacing human expertise.

Intelligence

AI that collaborates with humans on complex tasks, ensuring cognitive alignment with human reasoning.

3. Framework for Effective HAI Deployment

HAI's definition and core priorities establish a vision for ethical and human-centered AI. To transition from vision to implementation, HAI requires a structured framework that integrates ethical principles, societal impact, and practical deployment strategies. This framework focuses on aligning technology with human-centered interactions, ensuring responsible adoption across sectors, and maximizing societal benefits. It emphasizes critical success factors such as data accessibility, regulatory alignment, trust-building, and strategic investments in AI infrastructure. By addressing risks like data privacy and bias while fostering augmented human capabilities, this approach provides a clear roadmap for deploying HAI sustainably and ethically across diverse applications and sectors.

Conceptual Representation and Critical Success Factors:

1. Technology Description

AI systems designed with a focus on ethical, human-centered interactions that enhance decision-making and promote societal well-being.

2. Adoption

Widespread and responsible integration of AI across sectors, prioritizing ethical standards and aligning with regulatory frameworks.

3. Benefits

Augmented human capabilities, improved decision-making, and enhanced outcomes in critical areas such as healthcare, education, and public services.

4. Risks

Potential challenges related to data privacy, bias, and regulatory compliance that require proactive management to ensure responsible use.

5. Critical Success Factors

Key enablers such as data accessibility, regulatory alignment, and trust-building to support sustainable HAI deployment.

6. Investment

Strategic investments in AI infrastructure, research, and ethically aligned technologies to drive HAI's growth and societal impact.

This structured roadmap establishes a clear framework of enablers and safeguards designed to maximize societal benefits while ensuring proactive risk management. By embracing these principles, Saudi Arabia can solidify its position as a global leader in advancing ethical AI, one that aligns seamlessly with human values and drives national innovation goals.

4. History of HAI

The history of HAI traces the gradual integration of ethical standards and empathy in AI systems, reflecting a growing emphasis on creating AI that serves societal interests responsibly. This section explores HAI's evolution from early AI research to current frameworks prioritizing ethical AI.

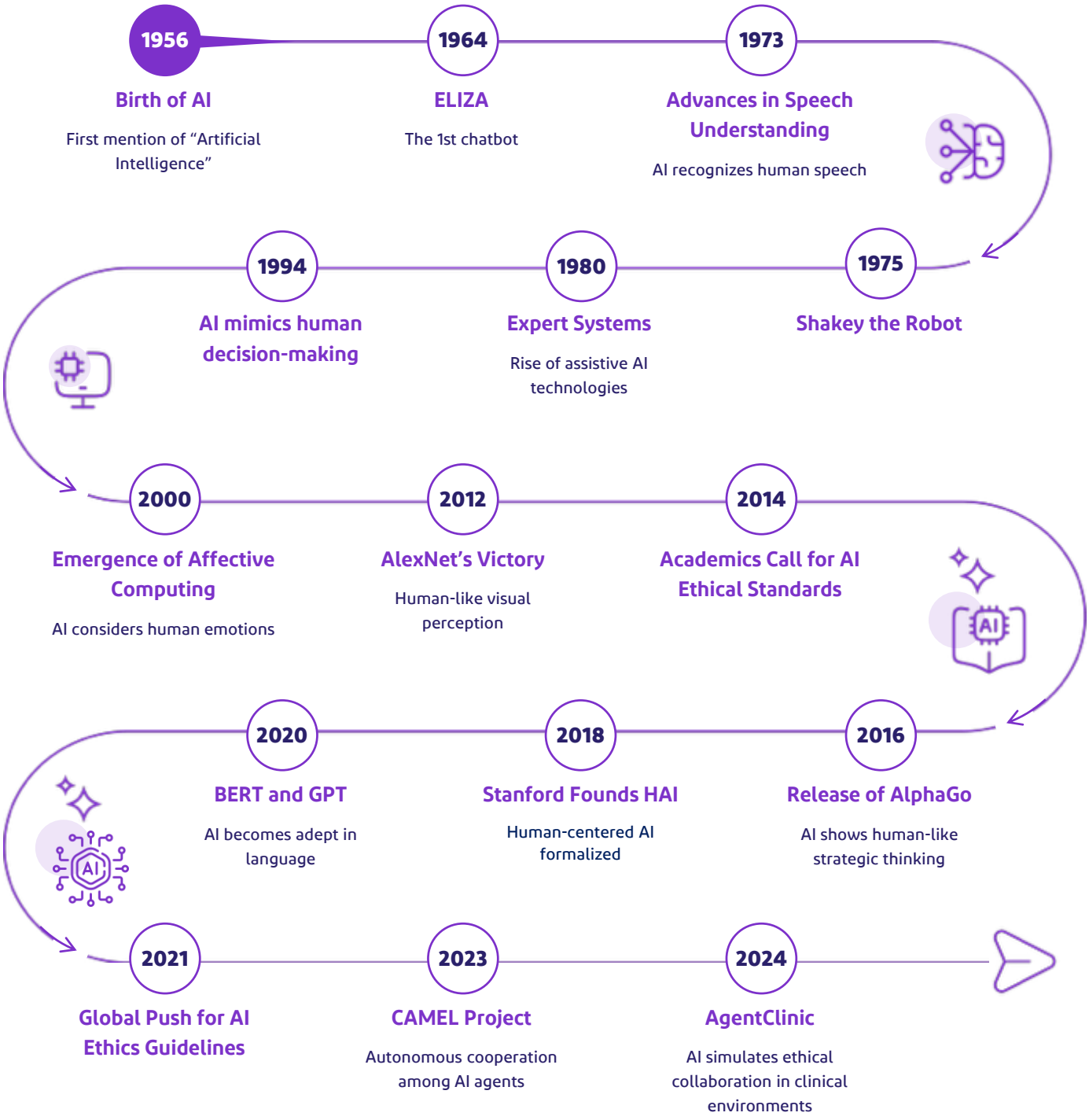
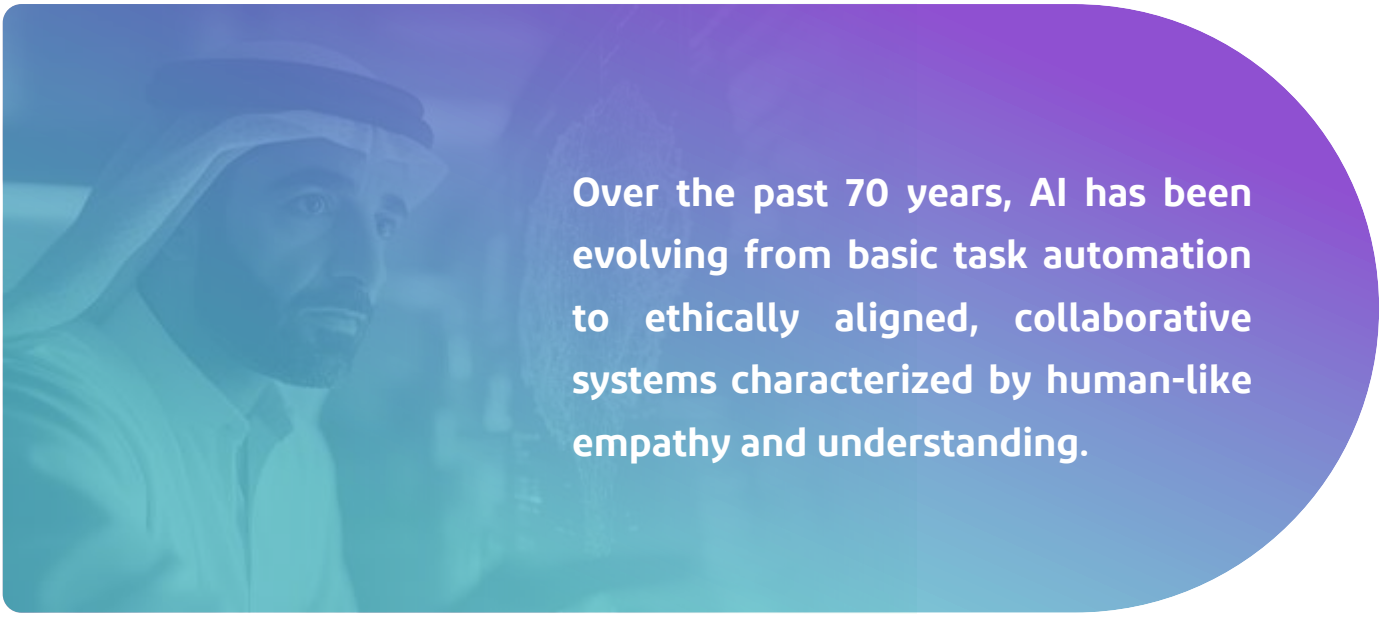


Figure 1: History of HAI (Timeline)

The evolution of HAI reflects a gradual alignment of AI capabilities with human-centric values. Starting in 1956, AI research initially focused on mimicking human intelligence. The first chatbot, ELIZA (1964-1966), demonstrated the potential of conversational AI, a core feature in HAI. By 1973, DARPA's speech recognition advances and Shakey the Robot (1975) underscored AI's ability to interact verbally and physically, laying groundwork for HAI applications. The 1980s saw the rise of Expert Systems that supported complex human decision-making, while the Belief-Desire-Intention model (1994) further deepened AI's alignment with human goals. Affective Computing (2000) brought emotional intelligence to AI, creating a framework for AI systems that can respond to user emotions.

In 2012, AlexNet's success in visual recognition demonstrated the power of deep learning, making AI capable of perceiving the world in ways similar to human vision. By 2016, AlphaGo's victory over Go champion Lee Sedol highlighted AI's strategic potential and human-like intuition, strengthening AI's role as a decision-making partner. The FAT/ML Conference, established in 2014, formalized academic standards around Fairness, Accountability, and Transparency in AI, marking a significant ethical pivot that has since guided AI development. Stanford's founding of its HAI Institute in 2018 further institutionalized HAI, emphasizing a research agenda focused on advancing AI technologies that prioritize human well-being.

Most recently, HAI has shifted towards facilitating human-centric workflows. The release of ChatGPT in 2020 marked a breakthrough in conversational AI, enabling more intuitive and context-aware human interactions across various domains. In 2021, global organizations like the EU and UNESCO established comprehensive AI Ethics Guidelines to formalize principles around transparency, fairness, and accountability in AI. By 2023, the CAMEL Project at King Abdullah University of Science and Technology (KAUST) advanced agent-based AI, enabling autonomous collaboration in complex environments with minimal human oversight. Building on these advancements, in 2024, AgentClinic marked a significant step forward in human-centered AI. AgentClinic evaluates multimodal AI systems in simulated clinical environments, emphasizing patient-centered outcomes and ethical decision-making. These advancements highlight HAI's potential to tackle real-world challenges while fostering trust and collaboration.



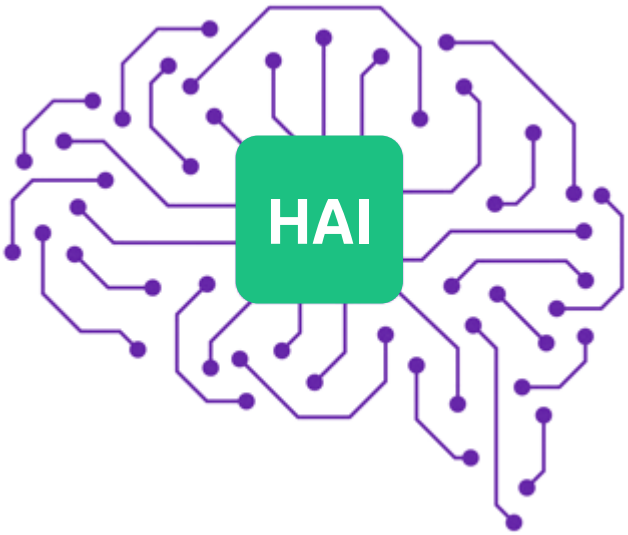
Over the past 70 years, AI has been evolving from basic task automation to ethically aligned, collaborative systems characterized by human-like empathy and understanding.

5. Significance of HAI

HAI is a transformative approach reshaping how artificial intelligence interacts with society, government, and industry. This section explores the significance of HAI from three key perspectives: global, local, and public sector. Each perspective provides insight into how HAI is evolving across the world, how it aligns with Saudi Arabia's strategic goals, and how it is revolutionizing public sector services.

5.1 Global Perspective

HAI is gaining momentum worldwide as nations recognize the need to ensure that AI not only delivers technological advancement but also aligns with ethical, societal, and human-centered values. As of 2023, 127 countries have implemented AI legislation, a significant increase from just 25 countries in 2022. This rapid expansion reflects a global commitment to ethical AI development and the need to mitigate the risks associated with AI misuse, which saw a 32.3% increase in incidents of unethical AI use in the same year. These incidents highlight the importance of creating systems that prioritize fairness, accountability, and transparency.



As explained earlier, HAI differs from traditional AI by focusing on augmenting human decision-making, emotional intelligence, and societal well-being. A striking example is the healthcare sector, where AI-enabled technologies have seen significant regulatory attention. In the United States alone, the FDA approved 139 AI-enabled medical devices in 2022, a 12.1% increase from the previous year and a 45-fold rise since 2012. These advancements are paving the way for more personalized and effective healthcare systems, reinforcing the value of AI that centers on human benefit rather than task automation.

Education is another critical area for HAI development. In North America, the number of computer science graduates has more than tripled since 2010, with a 7.9% growth in 2022 alone. This expanding workforce is essential for developing AI systems that collaborate with humans to improve productivity and decision-making across sectors like healthcare, education, and public services.

One of the key frameworks driving global AI ethics is the UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted by 193 member states in 2021. This landmark recommendation lays out several core principles for responsible AI governance:

Human Rights Protection:

AI must respect and uphold fundamental human rights, ensuring that it enhances societal welfare.

Transparency and Accountability:

AI systems must be designed to be transparent, explainable, and accountable, particularly in decision-making processes.

Fairness and Inclusiveness:

AI must be developed to avoid biases and ensure equitable access to its benefits.

Environmental Sustainability:

AI technologies should be designed with sustainability in mind, reducing environmental impact where possible.

Privacy and Data Security:

AI systems must integrate robust mechanisms to safeguard privacy and ensure data security.

The European Union's AI Act complements these principles by introducing a risk assessment-based regulatory framework, emphasizing stricter requirements for high-risk AI systems in areas such as healthcare, transportation, and public administration. These global ethical standards provide the foundation for responsible AI development across industries.

5.2 Local Perspective

Aligned with Saudi Arabia's National Digital Government Strategy (NDGS), HAI plays a critical role in modernizing public administration and delivering human-centered digital services. By supporting NDGS pillars such as Satisfied Citizen, Enabled Business, Effective Government, and Expedited Transformation, HAI enables tangible advancements in service personalization, efficiency, and transparency. These capabilities position HAI as a cornerstone of Saudi Arabia's efforts to transform its public sector and enhance citizen engagement.

For Citizen Satisfaction

HAI transforms citizen services by delivering personalized, inclusive, and accessible solutions. Practical examples include 24/7 support for healthcare inquiries, license renewals, and educational resources. By reducing wait times, simplifying processes, and addressing diverse needs, HAI fosters trust, equity, and satisfaction in public service delivery.

For Empowering Businesses

HAI empowers businesses through human-centered automation, streamlining regulatory compliance, market analysis, and customer service. By enabling tools that support small businesses and startups, HAI ensures equitable access to opportunities, promotes ethical decision-making, and drives growth that benefits employees, consumers, and society.

For Efficient Governance

HAI advances governance by integrating ethical and explainable AI tools for resource allocation, predictive analysis, and citizen-centric decision-making. Applications such as AI-driven feedback systems ensure policies align with societal needs, enhancing transparency, trust, and equitable service delivery.

For Efficient Investment

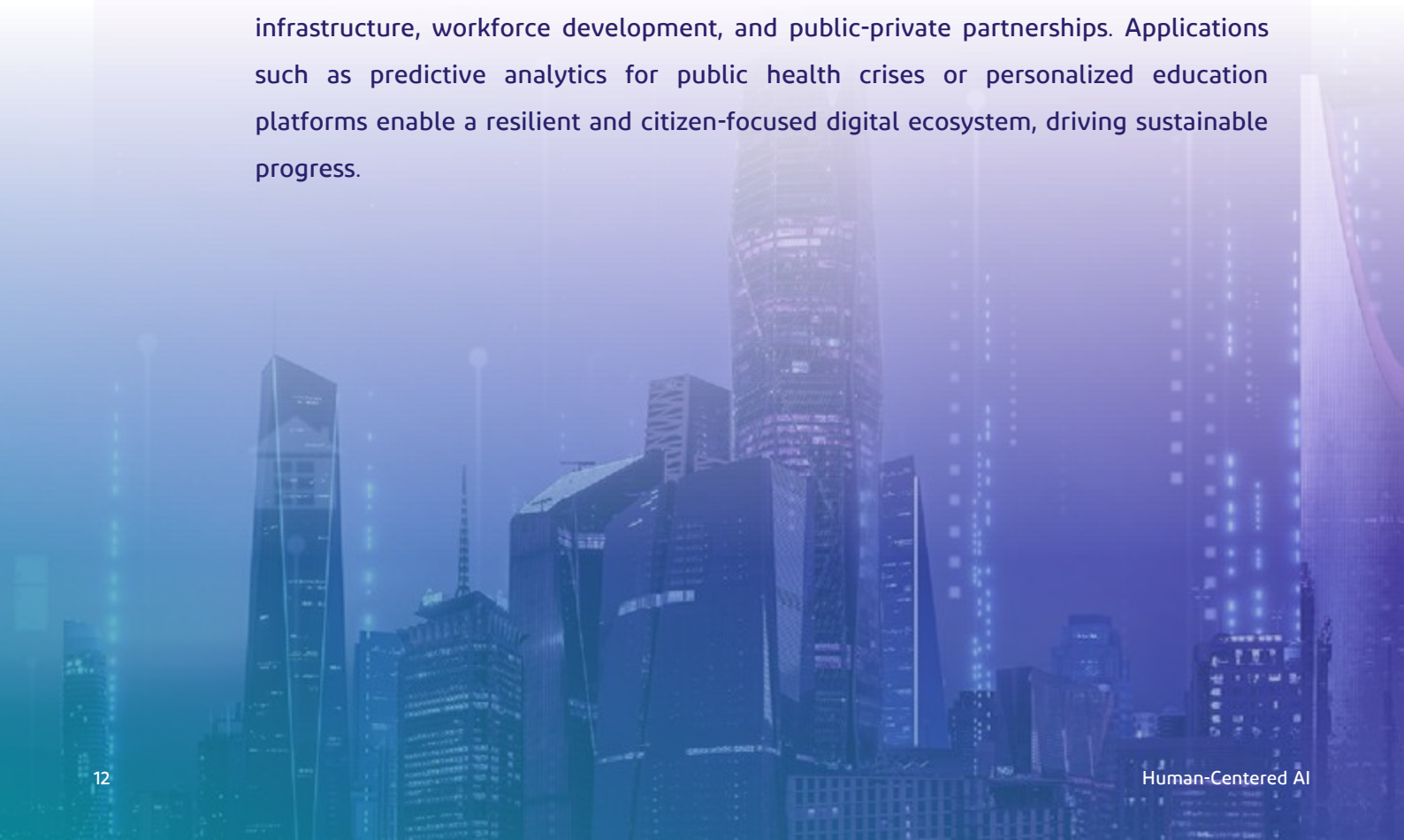
HAI optimizes public investments by leveraging ethical AI tools to prioritize human-centric projects such as smart hospitals, adaptive learning platforms, and sustainable infrastructure. This approach ensures resources are allocated equitably and transparently, fostering long-term societal well-being and sustainable growth.

For Regulation

HAI strengthens regulatory frameworks by embedding fairness, accountability, and inclusivity into compliance systems. For instance, AI tools can monitor healthcare regulations to ensure equitable access and safety standards. By automating oversight and aligning with international benchmarks, HAI promotes adaptive and citizen-focused policy enforcement.

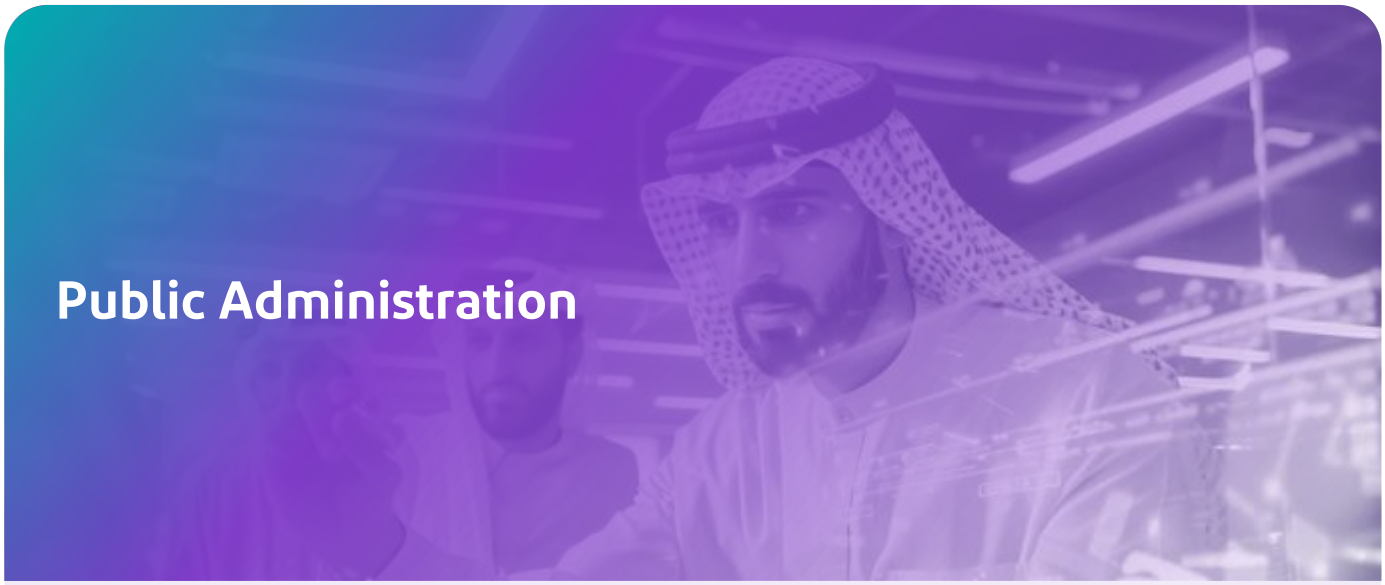
For Expedited Transformation

HAI accelerates digital transformation by embedding ethical principles into AI infrastructure, workforce development, and public-private partnerships. Applications such as predictive analytics for public health crises or personalized education platforms enable a resilient and citizen-focused digital ecosystem, driving sustainable progress.



5.2.1 Local Opportunities for HAI

Building on the local perspective outlined previously, this section provides actionable opportunities for HAI within Saudi Arabia. By focusing on key cross-sectoral strategic opportunities, it identifies how HAI-driven solutions can directly tackle digital government challenges and contribute to the Kingdom's sustainable development and societal well-being.



- **Opportunities:**

HAI offers ways to enhance public administration across all government sectors through optimized service delivery, real-time insights, and data-driven decision-making. Platforms like "ESTISHRAF" from Saudi Data & AI Authority (SDAIA), provide AI-powered analytics and decision-making support to over 100 government entities, enabling smarter policy implementation and operational efficiency. Similarly, the National Data Bank (NDB) integrates 320+ government systems, hosting over 100 TB of data and offering access to 8,700+ datasets from 249 entities, which can be leveraged for predictive analytics, resource optimization, and transparency initiatives. Adopting these tools, government entities can automate routine processes, improve collaboration, and create actionable insights for public welfare.

- **Impact:**

The integration of HAI into public administration has already shown promising results, with 81% of government entities reporting improved service delivery and 61% noting better decision-making. These advancements not only enhance public trust but also drive efficiency in operations, enabling a more responsive and transparent government aligned with national transformation goals.

Healthcare

- **Opportunities:**

HAI has the potential to address critical challenges in Saudi Arabia's healthcare sector, such as diagnostics, administrative efficiency, and personalized care. With 36% of health consultations already conducted virtually and over 123,000 beneficiaries of the Seha Virtual Hospital, HAI can enhance these services by automating diagnostic tasks like early detection of diabetic retinopathy, potentially increasing the 30,000 annual users of teleophthalmology services. It can also streamline administrative tasks, including generating reports, summarizing consultations, and answering patient follow-up questions in Arabic.

Also, by leveraging the "SEHHATY" application, with over 30 million users, and the "NPHIES" unified health file, which connected 2,000 providers and processed 14 million transactions in 2023, HAI can identify chronic disease patterns nationwide and craft personalized care plans.

- **Impact:**

These advances improve operational efficiency, forecast healthcare demand, and ensure equitable access to healthcare, especially in remote areas.

Education

- **Opportunities:**

HAI can transform Saudi Arabia's education sector by addressing key challenges in personalized learning and teacher development. With over 400,000 children enrolled in kindergarten and a 34% early childhood enrollment rate, HAI-powered adaptive learning systems can tailor content to individual student needs, track developmental milestones, and provide multilingual support for students and parents. Additionally, with over 25,000 schools participating in the "Tamayuz Digital Platform" and 115,000 professional teaching licenses issued, HAI can automate school evaluations, analyze performance data, and provide teachers with personalized professional development plans based on their strengths and students' learning outcomes.

- **Impact:**

These advancements ensure equitable access to high-quality early education, improve teaching effectiveness through targeted development, and foster a data-driven approach to school improvement, aligning with national goals for educational excellence.

HAI's potential further extends to other vital sectors such as transport, telecommunications, energy and utilities, tourism, and entertainment. By prioritizing human values and ethical principles, HAI can optimize transportation systems through predictive traffic management, enhance energy grids with smart, adaptive solutions, and personalize tourism experiences to improve visitor satisfaction. Under the guidance of the Digital Government Authority, HAI-driven tools ensure these sectors deliver citizen-centric, transparent, and efficient services.

5.3 Impact and Added Value of the Public Sector

In the public sector, HAI enables more efficient healthcare, education, and social services by aligning AI systems with societal values and ethical standards. From personalized learning tools in education to AI-driven support in healthcare and disaster response, HAI applications are proving their potential to transform the public sector and enhance community welfare.

Singapore

Singapore's education system is transforming with HAI and EdTech, focusing on personalized learning. AI-driven adaptive learning systems such as Geniebook and Explico tailor educational experiences to individual student needs, enhancing engagement and fostering self-directed learning. These tools align with HAI principles of inclusivity and fairness, promoting equitable educational outcomes.

United States

The Social Security Administration utilizes AI to accelerate disability benefits determinations, identifying medical needs aligned with HAI principles of equity and accessibility. By reducing processing times and ensuring fair access to essential support, these systems prioritize the welfare of individuals and foster trust in public services.

Estonia

Estonia's education system uses HAI tools like the ProgeTiger Program to adapt lessons based on student performance, providing real-time feedback to support individual learning paths. These initiatives exemplify HAI's focus on inclusivity by ensuring that all students, regardless of background, can access tailored educational opportunities that enhance digital literacy and foster equitable learning outcomes.

Australia

Australia's University of Adelaide launched NOBURN, an AI-powered app that predicts bushfire risk by analyzing user-submitted photos of vegetation. NOBURN empowers rural communities to engage in bushfire prevention, aligning with HAI's emphasis on citizen welfare and safety. By expanding data collection in hard-to-reach areas, this initiative highlights how HAI can enhance community resilience and public safety.

Japan

Japan has invested over $300 million in care robots like Robear and Pepper to address eldercare needs, assisting with lifting, mobility, and social engagement. These AI-driven tools aim to ease caregiver workloads and improve eldercare services, aligning with HAI's focus on augmenting human capabilities and supporting aging populations.

Germany

In February 2020, the German Federal Government announced EUR 92 million in funding for 305 AI projects across agriculture, health nutrition, food chain management, and rural areas. This initiative aims to boost efficiency, sustainability, and support rural development, aligning with HAI's goals of enhancing human capabilities & promoting equitable access to advanced technologies.

5.3.1 Advantages of Integrating HAI into the Public Sector

Strengthening Trust and Transparency in Public Services

HAI enhances transparency in government services by making decision-making processes more understandable and accessible to the public. AI systems developed with fairness and transparency reduce biases, improving citizen trust and engagement.

Example: Platforms like Absher could benefit from HAI-driven processes that explain data use and decision-making, reinforcing trust in digital government services and supporting Saudi Vision 2030's goal for a responsive government.

Delivering Personalized and Accessible Services for

HAI customizes public services based on individual needs, ensuring inclusive access. AI-powered personalization adapts interactions for citizens with different backgrounds, capabilities, and locations, promoting equity and inclusivity. By facilitating public access to information, HAI contributes to a fundamental aspect of participatory governance as reflected in the E-Information component of the E-Participation Index.

Example: Through HAI, services on platforms like "TAWAKKALNA" could offer tailored guidance for healthcare, travel, or emergency responses, improving accessibility for citizens in remote areas or those with disabilities.

Ethical AI Practices to Support Workforce Development

HAI fosters a collaborative work environment, where AI tools enhance employees' roles rather than replace them. By prioritizing ethical AI use, HAI ensures that employees are empowered and can upskill to work alongside AI.

Example: The government of Saudi Arabia can integrate HAI principles to train its public sector workforce, enabling employees to work confidently with AI tools and ensuring that government services are delivered ethically and effectively.

Enabling Evidence-Based Decision-Making with Ethical Guardrails

HAI-driven data analytics provide public sector leaders with insights that respect ethical considerations, guiding policy-making that benefits society at large. This allows for more socially responsible decisions.

Example: Saudi Data and Artificial Intelligence (SDAIA) and the Ministry of Health could leverage HAI analytics to improve healthcare resource allocation and address community health disparities, reinforcing data ethics and fairness in decision-making.

Enhancing Proactive Regulatory Oversight

HAI enables continuous, automated monitoring of AI systems to ensure compliance with privacy, fairness, and transparency regulations. This proactive oversight maintains ethical standards and safeguards public interests.

Example: Regulatory bodies can use HAI to monitor and audit AI applications, ensuring they meet Saudi Arabia's legal and ethical standards. For example, HAI could support real-time checks on compliance in financial and educational services.

Promoting Sustainable and Socially Responsible Innovation

HAI encourages sustainable development by aligning AI innovation with societal goals, emphasizing long-term social benefits over short-term gains. This supports Saudi Arabia's vision for a balanced, sustainable society.

Example: Through HAI, initiatives like NEOM and the Saudi Green Initiative can integrate AI that prioritizes environmental and social well-being, promoting responsible innovation that contributes to the Kingdom's sustainable development objectives.

6. Challenges and Considerations for HAI Deployment

While the transformative potential of HAI is clear, its successful integration into society requires addressing several key challenges. Ensuring that HAI aligns with the ethical principles and societal values of Saudi Arabia's Vision 2030 is critical to driving innovation, economic growth, and responsible AI deployment. The global momentum for regulating AI is reflected in the increase in AI legislation, which has grown from 25 countries in 2022 to 127 in 2023 according to the Stanford's HAI Institute. This highlights the importance of aligning AI development with robust legal frameworks to ensure ethical and responsible deployment. This section highlights the main challenges and outlines strategic considerations that can help the Kingdom's leaders build a secure, sustainable, and ethical AI ecosystem.

1. Privacy and Security

Key Concern

HAI often involves handling sensitive data in critical sectors like healthcare and government services. Breaches or misuse of this data can erode trust and harm institutions deploying these technologies. According to the 2024 AI Index Report, incidents of ethical misuse have risen by 32.3%. Moreover, organizations worldwide rank privacy, security, and reliability as top risks, emphasizing the urgency for safeguards.

Mitigation

To address these concerns, adopt privacy-by-design principles, integrating encryption, secure access controls, and differential privacy from the start. Regular audits and transparency reports on data handling can reinforce public trust. Additionally, organizations should align with global standards like GDPR and HIPAA while implementing comprehensive employee training programs to ensure adherence to ethical data practices.

2. Job Displacement

Key Concern

HAI may lead to significant job displacement, particularly in sectors involving repetitive or predictable tasks such as customer service, data entry, and administrative support. As AI automates these functions, lower-skilled workers face heightened risks of unemployment, creating potential economic and social challenges. Such displacement could also contribute to income inequality.

Mitigation

Invest in reskilling and upskilling programs to prepare workers for roles requiring creativity, problem-solving, and interpersonal skills. Promote HAI solutions that augment workers' capabilities rather than replace them, such as AI tools for supporting customer inquiries. Establish public-private partnerships to fund workforce transition initiatives and emphasize lifelong learning to future-proof jobs.

3. Human-Centered Design and Usability

Key Concern

Poorly designed AI systems can frustrate users, reduce adoption rates, and hinder the effectiveness of HAI. If systems are overly complex or fail to integrate into workflows, they may fail to meet their goal of complementing human expertise and improving efficiency. User-friendly designs are crucial to ensure that AI systems support human tasks.

Mitigation

Adopt user-centered design principles by conducting extensive user testing, collecting feedback, and implementing iterative improvements. Ensure interfaces are intuitive, accessible, and aligned with user needs. Engage end-users early in the design process to tailor HAI applications to specific contexts, enhancing usability and maximizing their impact.

7. Future Outlook of HAI

Building on a comprehensive understanding of HAI—its global and local impact, potential challenges, and mitigation strategies, the focus now shifts to its potential to transform public sector services. By integrating advanced capabilities such as predictive analytics, explainable AI, personalized citizen interactions, and human-AI collaboration, HAI is set to enable more efficient decision-making, enhance transparency, and deliver tailored, citizen-centric services.

These advancements are expected to shape the future of digital governance through the following trends:

Human-AI Collaboration:

AI's role in government is not about replacing human workers but augmenting their capabilities. Government employees will be able to leverage AI tools to enhance their decision-making processes, improving both the speed and accuracy of policy implementation. This collaborative approach ensures that while AI handles routine tasks, human oversight remains at the core of strategic government functions.

Explainable AI:

As AI systems take on more complex tasks, there is a growing need for transparency in how decisions are made. Explainable AI will be crucial in building trust between citizens and the government, especially when AI systems are deployed in sensitive areas such as law enforcement or social welfare. HAI ensures that AI's decision-making processes are not only accurate but also easily understood by the public, thus enhancing accountability and transparency.

Predictive Governance:

AI will increasingly be used to forecast societal needs & policy outcomes. For example, predictive analytics could help the government anticipate healthcare demands or optimize infrastructure development by analyzing vast amounts of data. This shift from reactive to proactive governance will allow for more strategic resource allocation and faster responses to emerging challenges.

Enhanced Personalization:

The future of digital government in Saudi Arabia is increasingly citizen-centric, with AI technologies enabling more personalized service delivery. For example, in healthcare, AI could create individualized treatment plans based on a citizen's medical history, while in education, AI-driven systems could tailor learning paths to each student's needs. This level of personalization ensures that citizens receive more relevant and impactful services, improving both satisfaction & engagement.

8. General Recommendations

To fully realize the potential of HAI, strategic actions must be taken to embed HAI into the Kingdom's ongoing initiatives under Saudi Vision 2030. This section presents a set of key recommendations that provide a clear pathway for successfully integrating HAI into Saudi Arabia's digital landscape.

Prioritize Data Privacy and Ethical Standards

Action: Establish specialized AI ethics bodies within key government sectors to oversee the design and deployment of HAI systems, ensuring privacy-preserving techniques like differential privacy are integrated across applications.

Impact: Strengthens ethical oversight, protects sensitive data, and ensures sector-specific alignment of AI systems with societal values, international regulations and user rights.

Invest in AI Workforce Development

Action: Partner with academic and industry leaders to develop HAI-focused training programs that emphasize ethical decision-making, human collaboration, and the augmentation of human capabilities in critical sectors.

Impact: Equips the workforce with skills for ethical and collaborative AI usage, supporting innovation while mitigating job displacement.

Foster AI Innovation Across Critical Sectors

Action: Launch HAI-driven pilot programs in healthcare and education to improve patient outcomes and personalized learning while integrating predictive governance tools for data-driven policy-making.

Impact: Enhances sectoral efficiency and equity, addressing societal challenges with human-centered solutions.

Strengthen AI Infrastructure and Security

Action: Build scalable infrastructure to support HAI applications and deploy security frameworks that protect AI systems against biases, adversarial attacks, and misuse.

Impact: Ensures the safe, fair, and transparent operation of HAI systems, fostering sustainable growth and reliability.

Promote Public-Private Partnerships and International Collaboration

Action: Develop joint initiatives with global universities such as MIT and Stanford, along with global industry leaders, to create HAI-specific ethical standards and collaborative frameworks that emphasize societal impact and inclusivity.

Impact: Aligns national HAI practices with international best practices, fostering global leadership and trust in ethical AI development.

9. Conclusion

HAI stands at the forefront of a global shift in technology, reshaping how societies interact with AI systems by embedding ethical principles and prioritizing human values. As outlined in this study, HAI is not merely an alternative approach to AI development; it is an imperative for ensuring that advancements in artificial intelligence serve as a force for societal good.

From the integration of ethical guidelines and regulatory safeguards to its transformative potential in sectors like healthcare, education, and governance, HAI represents a comprehensive strategy for addressing both current and future challenges. The Digital Government Authority (DGA), with its commitment to digital transformation and ethical innovation, is uniquely positioned to leverage HAI to foster innovation, inclusivity, and sustainability across its digital landscape.

The recommendations provided, focusing on data privacy, workforce development, sectoral innovation, and global collaboration, form a clear pathway for embedding HAI into national priorities. For instance, through targeted initiatives like AI-powered personalized healthcare, adaptive learning platforms, and citizen-centric governance tools, DGA can enhance public services while reinforcing trust, transparency, and fairness.

This focus on transparent and citizen-centric governance aligns with Saudi Arabia ranking seventh globally in the UN E-Participation Index, reflecting its leadership in online citizen engagement. By adopting explainable AI and advancing personalized data access through its HAI initiatives, the DGA can build on this success, empowering citizens to better understand government data and further strengthening e-participation.

Looking toward the future, DGA has an unparalleled opportunity to lead by example, showing the world how ethical, human-centric AI can transform societies for the better. By ensuring that AI technologies augment human capabilities rather than replace them, and by fostering equitable access to the benefits of AI, DGA can position itself as a global beacon of responsible innovation.

HAI is more than a framework, it is a vision of how humanity and technology can thrive together. With deliberate and strategic action, DGA can harness this transformative potential, ensuring a legacy of progress that aligns with Saudi Arabia's cultural values, economic aspirations, and societal goals.

10. Definitions

Term	Definition
Human-Centered AI (HAI)	An approach to AI that integrates ethical principles, human values, and societal well-being into technology design and implementation.
Traditional AI	AI systems designed primarily to optimize technical performance and automate processes without necessarily considering ethical or societal impacts.
Augmented Human Capabilities	The enhancement of human skills and decision-making through AI tools that work alongside people rather than replacing them.
Ethical AI	The development and deployment of AI that adheres to fairness, transparency, accountability, and respects human rights.
Explainable AI	AI systems designed to make their decision-making processes transparent and understandable to humans, thereby building trust and ensuring accountability.
Citizen-Centric Services	Public services designed and delivered with a focus on individual needs, ensuring accessibility and fairness.

11. Bibliography

Bibliography	Links
Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2023). Annual report 2023. Stanford University	https://hai.stanford.edu/sites/default/files/2024-04/2023-Annual-Report.pdf
Weizenbaum, J. (1966). ELIZA – A computer program for the study of natural language communication between man and machine. Stanford University.	https://web.stanford.edu/class/cs124/p36-weizenabaum.pdf
Nilsson, N. J. (1984, April 1). Shakey the robot (Technical note no. 323, 135 pp.). SRI International Menlo Park, CA. Approved for public release.	https://apps.dtic.mil/sti/pdfs/ADA458918.pdf
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (NIPS) (pp. 1097-1105).	https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
Google DeepMind. (n.d.). AlphaGo.	https://deepmind.google/research/breakthroughs/alphago/
Fairness, Accountability, and Transparency in Machine Learning (FATML). (n.d.).	https://www.fatml.org
UNESCO. (2021). Recommendation on the ethics of artificial intelligence.	https://unesdoc.unesco.org/ark:/48223/pf0000381137
CAMEL-AI. (n.d.). CAMEL-AI.	https://www.camel-ai.org/
Schmidgall, S., Ziaei, R., Harris, C., Kim, J. W., Reis, E., Jopling, J., & Moor, M. (2025). AgentClinic: A multimodal agent benchmark to evaluate AI in simulated clinical environments. In Proceedings of the International Conference on Learning Representations (ICLR).	https://arxiv.org/pdf/2405.07960
Stanford Institute for Human-Centered Artificial Intelligence (HAI). (2024). Artificial Intelligence Index Report 2024. Stanford University.	https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf
European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Official Journal of the European Union, L 311, 12 July 2024, pp. 1–78.	https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
Health Sector Transformation Program. (2023). Achievements report 2023. Vision 2030, Kingdom of Saudi Arabia.	https://www.vision2030.gov.sa/media/xznnx3vz/hstp-annual-report-2023-en.pdf
Human Capability Development Program (HCDP). (2023). Achievements report for 2023. Vision 2030, Kingdom of Saudi Arabia.	https://www.vision2030.gov.sa/media/dknctbrs/hcdp-achievements-report-2023.pdf
Saudi Data & AI Authority (SDAIA). (2024, September). State of AI in Saudi Arabia. Global AI Summit.	https://globalaisummit.org/Documents/StateofAIinSaudiArabia.pdf
Tech Wire Asia. (2023). How AI is shaping and personalizing education and learning experiences in Singapore.	https://techwireasia.com/2023/08/ai-in-education-geniebooks-personalized-learning-journey/
Geniebook. (n.d.).	https://geniebook.com
Explico. (n.d.).	https://www.explico.sg

Bibliography	Links
U.S. Government. (n.d.). Federal AI use case inventories.	. https://ai.gov/ai-use-cases/
TechNode Global. (2024, May 17). Estonia's Education Minister Kristina Kallas on the challenges and opportunities of AI in learning and empowerment [Q&A].	https://technode.global/2024/05/17/estonias-education-minister-kristina-kallas-on-the-challenges-and-opportunities-of-ai-in-learning-and-empowerment-qa/
University of Adelaide. (2024). New AI-powered bushfire-fighting app NOBURN goes live.	https://www.adelaide.edu.au/research/news/list/2024/01/24/new-ai-powered-bushfire-fighting-app-noburn-goes-live
MIT Technology Review. (2023). Inside Japan's long experiment in automating elder care.	https://www.technologyreview.com/2023/01/09/1065135/japan-automating-eldercare-robots/
European Commission. (n.d.). Germany AI strategy report.	https://ai-watch.ec.europa.eu/countries/germany/germany-ai-strategy-report_en

For continuous development purposes, please fill out the following survey

Survey

SCAN HERE

For more digital studies, please follow the link below

Digital Studies

SCAN HERE

هيئة الحكومة الرقمية
Digital Government Authority

DGA.GOV.SA