



Islamic University of Technology (IUT)

**Assignment on Reading papers on
human-centered AI/ML**

Submitted by

Tasfia Tasneem Annesha, 190041220

Submitted to

Dr. Hasan Mahmud

Associate Professor, Department of CSE

Department of Computer Science and Engineering (CSE)

Islamic University of Technology (IUT)

A Subsidiary organ of the Organization of Islamic Cooperation (OIC)

Academic Year: 2022-2023

Jan, 2024

Contents

0.1	Given Task	2
0.2	Answers	2

0.1 Given Task

Read the paper **Human-Centered Artificial Intelligence and Machine Learning** [8] and related research/links and compile a document answering the followings:

1. What is human-centered ML/AI (HCML/HCAI)? Is there any framework or model for HCAI?
2. How does HCAI relate to HCI?
3. What are the roles of humans in the lifecycle of ML/AI?
4. Tech giants like Google, Microsoft, and IBM, what sort of HCAI-based applications they have designed? How they have incorporated humans in the ML/AI applications?
5. Suggest some research area or scope (At least 3) related to the area of HCAI/HCML

0.2 Answers

What is human-centered ML/AI (HCML/HCAI)? Is there any framework or model for HCAI?

Designing artificial intelligence (AI) systems prioritizing human needs, values, and capabilities is known as human-centered artificial intelligence (HCAI) [4]. These systems use data not only to assist people but also acknowledge their own values, biases, constraints, and ethical considerations during data gathering and algorithm application. HCAI emphasizes that for AI to benefit humans, it must understand basic human aspects such as communication and trust. Its research focuses on ethical implications and aims to enhance rather than replace human talents which is inspired by the complexity and diversity of human's socio-cultural situations [3].

The paper **An HCAI Methodological Framework: Putting It Into Action to Enable Human-Centered AI** by *Wei Xu, Zaifeng Gao, and Marvin Dainoff*, published in November 2023, proposes a comprehensive study on a Human-Centered AI (HCAI) Methodological Framework [10].

As for frameworks or models, HCAI emphasizes the integration of AI into human-centric systems, focusing on transparency, equitable outcomes, respecting privacy, and enhancing human capabilities. This framework aims to address the

current lack of methodological guidance in implementing HCAI and presents a practical **three-layer** approach to overcome existing weaknesses and challenges in the field. This framework aims to prioritize humans in the design, development, and deployment of AI systems, ensuring the maximization of AI benefits while avoiding potential adverse effects. It emphasizes the integration of AI into human-centric systems with a focus on transparency, equitable outcomes, privacy, and enhancing human capabilities. It illustrates the integration of various

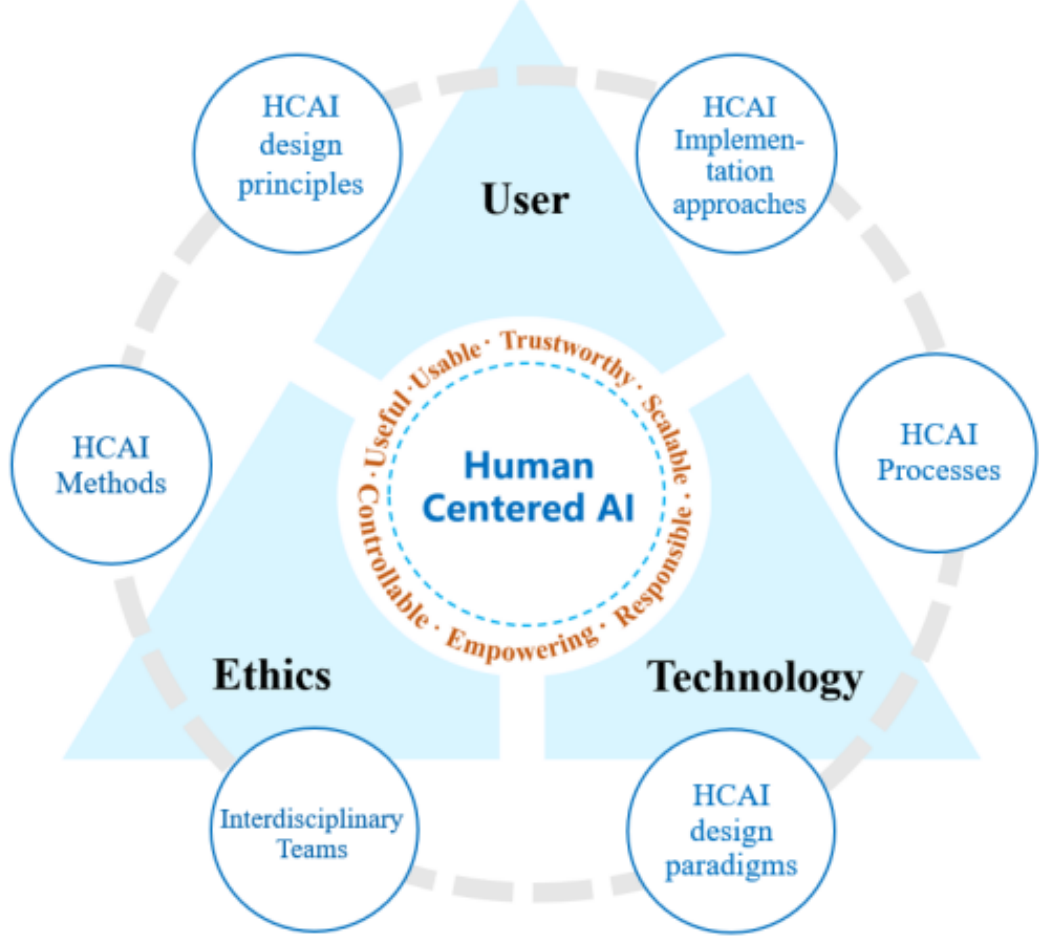


Figure 1: HCAI Framework[10]

components of Human-Centered Artificial Intelligence (HCAI) around three core elements: User, Ethics, and Technology. These elements are interconnected and suggest a holistic approach to AI development. Central to the framework is **Human-Centered AI**, which is surrounded by keywords like *Reliable*, *Trustworthy*, *Inclusive*, *Empowering*, and *Responsible*. This indicates the values that are intended to be at the heart of HCAI. The proposed framework integrates seven key components [10]:

1. **Design Goals:** Clear objectives for intelligent systems, emphasizing the

dimensions of *user-technology-ethics*.

2. **Design Principles:** Strategic principles guiding AI development in a human-centric manner, with a systematic approach across three dimensions: user, technology, and ethics.
3. **Implementation Approaches:** Current approaches are initial and not comprehensive, but the study defines 15 approaches to implement HCAI systematically across the *user-technology-ethics* dimensions.
4. **Design Paradigms:** While current paradigms view AI as a tool or subordinate to human teams, the study proposes alternative design paradigms, such as human-AI joint cognitive systems, human-AI joint cognitive ecosystems, and intelligent sociotechnical systems.
5. **Interdisciplinary Teams:** Emphasis on collaborative efforts from various disciplines to foster ethics in HCAI.
6. **Methods:** This involves specific techniques or sets of procedures for AI design towards HCAI goals. The study recommends consolidating 25 HCAI-based methods to enable HCAI.
7. **Processes:** A comprehensive HCAI process that integrates HCAI and AI processes throughout the AI lifecycle.

This proposed framework enhances the existing HCAI landscape by providing a systematic approach that prioritizes human values and ethical considerations in AI development, offering clear design goals, principles, implementation strategies, methods, and interdisciplinary collaboration. It aims to facilitate the development, transfer, and practical implementation of HCAI while addressing the current limitations in the field.

How does HCAI relate to HCI?

Human-Computer Interaction (HCI) and Human-Centered AI (HCAI) are closely related fields [9]. HCI is about making user-friendly computer systems, while HCAI goes further by combining AI to create systems that are both user-friendly and smart, meeting human needs effectively. The main goal is to develop AI systems that not only understand but also enhance human abilities, promoting better interactions between people and machines.

HCAI emphasizes designing AI systems that match human values and preferences. On the other hand, HCI studies how people interact with computers and creates computing products that suit user needs.

The connection between HCAI and HCI is their blending of AI and HCI principles to build a clear and user-friendly framework. This integration helps create AI intelligence that is more reliable and trustworthy for users. Some worry about AI-powered devices causing potential problems, while others support a balanced integration of HCI and AI for mutual benefits.

HCI experts are crucial in developing HCAI systems. They deal with unique challenges linked to AI systems and ensure these systems always prioritize human needs and values[7]. So, the focus of HCI work should shift from traditional human-computer interaction to interactions with AI systems. This shift brings new challenges and opportunities for HCI professionals in the field of human-AI interaction (HAII).

What are the roles of humans in the lifecycle of ML/AI?

The roles of humans in the lifecycle of ML/AI are varied and important [2]. Humans are involved in every stage of the process, from data collection and preparation, to model design and development, to evaluation and deployment, to monitoring and maintenance. Humans also provide the ethical and social context for AI systems, ensuring that they are aligned with human values and goals. Here is a brief overview of some of the roles of humans in the lifecycle of ML/AI:

1. Data Collection and Preparation:

- **Data Engineers:** They are responsible for collecting, cleaning, transforming, and storing data for ML/AI applications. They ensure that the data is reliable, consistent, and accessible for analysis and modeling [5].
- **Data Analysts:** They are responsible for exploring, visualizing, and interpreting data for ML/AI applications. They use statistical and analytical tools to identify patterns, trends, and insights from the data.

2. Model Design and Development:

- **Machine Learning Engineers:** They are responsible for designing, developing, testing, and deploying ML/AI models for various applications. They

use programming languages, frameworks, and libraries to implement ML/AI algorithms and techniques [2].

- Machine Learning Researchers: They are responsible for advancing the state-of-the-art in ML/AI by conducting original research and developing new methods and theories [1]. They publish their findings in academic journals and conferences, and collaborate with other researchers and practitioners.

3. Training and Testing:

Humans provide data and feedback for training AI systems, and validate their accuracy and effectiveness.

4. Evaluation and Deployment:

Machine Learning Engineers ensure the successful deployment of models into real-world applications.

5. Usage and Interaction:

End-users interact with AI systems, providing ongoing data and feedback for system refinement.

6. Governance and Ethics:

- AI Ethicists: They are responsible for ensuring that ML/AI systems are ethical, fair, accountable, transparent, and trustworthy [2]. They study the social and moral implications of ML/AI, and provide guidelines and recommendations for ethical design and use of ML/AI.

It's important to highlight that throughout these stages, humans also play a critical role in monitoring and maintaining AI systems to ensure their continued performance, accuracy, and adherence to ethical standards. Additionally, domain experts, project managers, and business stakeholders collaborate across these stages to define objectives, set goals, and guide the development and implementation of AI solutions.

Tech giants like Google, Microsoft, and IBM, what sort of HCAI-based applications they have designed? How they have incorporated humans in the ML/AI applications?

- **IBM:** IBM's *AI for Fashion* is a collaboration between IBM Garage and Bestseller India, a clothing company overseeing 20 fashion brands. Fashion.ai utilizes IBM Watson to predict upcoming trends, relevant styles, colors, and size ratios for diverse markets and segments. By analyzing fashion trends and customer preferences, AI enhances designers' ability to craft personalized collections. IBM also offers various HCAI applications that employ natural language generation, machine learning, and data visualization to assist users across different fields. For instance, *IBM Watson Assistant* serves as a conversational AI platform, aiding users with tasks such as flight booking, food ordering, and customer support. *IBM AutoAI* streamlines machine learning model creation, producing more accurate results, faster. Additionally, *IBM Uncertainty Quantification 360* enhances AI trustworthiness by quantifying and communicating prediction uncertainties.
- **Google:** Google incorporates AI into user experience designs, as seen in *Google Assistant* and *Google Photos*, which personalize and enhance user interactions based on individual preferences. Google has developed numerous HCAI applications employing natural language understanding, computer vision, and speech recognition to assist users across various domains. For example, Google Assistant serves as a virtual assistant for tasks like making calls, scheduling appointments, playing games, and controlling smart devices. *Google Lens* is a visual search tool identifying objects, text, and landmarks through the camera or photos. Google Translate offers language translation for text, speech, images, and websites in over 100 languages.
- **Microsoft:** Microsoft's research encompasses projects like conversational user interfaces, using AI to create more engaging user experiences. Microsoft has developed several HCAI applications employing natural language processing, computer vision, and speech synthesis to support users in various scenarios. *Cortana*, Microsoft's personal assistant, assists users with tasks like setting reminders, sending emails, web searches, and calendar management. *Seeing AI*, a mobile app, narrates the world for individuals who are blind or have low vision by describing scenes, text, faces, and

colors through the camera or photos. *Microsoft Translator* offers language translation for text, speech, images, and conversations in over 70 languages.

These companies have not only designed user-friendly, transparent, and trustworthy applications but have also integrated human expertise, feedback, and preferences to continuously enhance their AI systems.

Suggest some research area or scope (At least 3) related to the area of HCAI/HCML

Human-Centered Artificial Intelligence (HCAI) and Human-Centered Machine Learning (HCML) are new research topics that aim to create AI and machine learning systems that prioritize human needs, values, and skills. These domains seek to develop AI systems that are not just technically competent, but also accessible, intelligible, and valuable to humans. Here are three promising research areas or scopes within HCAI/HCML:

Explainable AI (XAI)

Overview

XAI aims to make AI decision-making processes transparent and understandable to humans [6]. This is critical in high-stakes environments such as healthcare, banking, and legal systems, where trust and ethical considerations require an understanding of AI reasoning.

Research Scope

Developing algorithms and interfaces that can effectively communicate AI decision-making processes to users, including non-experts. Research also involves investigating how different user groups perceive and understand AI explanations.

AI for Accessibility and Assistive Technologies:

Overview

This area focuses on developing AI systems that enhance the capabilities and experiences of people with disabilities. It includes creating assistive technologies that are adaptive, responsive, and sensitive to diverse user needs.

Research Scope

Designing AI-powered applications to aid in language processing for people with speech impairments, visual recognition systems for the visually impaired, and predictive text and gesture recognition systems for people with mobility or cognitive issues.

Ethical AI and Bias Mitigation:

Overview

Increasing concerns about ethics in AI have garnered significant attention from researchers and practitioners. Their efforts are concentrated on creating algorithms that ensure fairness and eliminate biases. These biases often stem from historical data or skewed training sets, potentially resulting in discriminatory or unjust outcomes.

Research Scope

The research in this domain is dedicated to developing algorithms aimed at mitigating biases within AI systems. This encompasses the creation of techniques for detecting and mitigating biases in various aspects, including data, models, and predictions. Additionally, researchers are actively involved in defining and assessing fairness metrics, with a focus on establishing quantitative measures and developing tools to evaluate bias in machine learning models.

Each of these domains reflects a unique blend of technology innovation and human-centric design, highlighting the growing relevance of integrating human viewpoints and values into the development of advanced AI systems.

Bibliography

- [1] Google Cloud. Making the machine: the machine learning lifecycle. <https://cloud.google.com/blog/products/ai-machine-learning/making-the-machine-the-machine-learning-lifecycle>, 2022. 6
- [2] Tom Coughlin. The role of the human in artificial intelligence. <https://transmitter.ieee.org/role-human-artificial-intelligence/>, Sep 2017. 5, 6
- [3] IBM. What is human-centered ai? 2
- [4] Interaction Design Foundation. Human-centered ai. 2
- [5] Kazuaki Ishizaki. Ai model lifecycle management: Overview, 2020. 5
- [6] Xiangwei Kong, Shujie Liu, and Luhao Zhu. Toward human-centered xai in practice: A survey. *Machine Intelligence Research*, 2024:1–32, 2024. 8
- [7] Sumit Kumar Kushwaha, Anurag Singh, Shubham Singh, Shubham Singh, Shubham Singh, and Shubham Singh. Review on artificial intelligence and human computer interaction. In *2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON)*, pages 1–6. IEEE, 2022. 5
- [8] Mark O Riedl. Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, 1(1):33–41, 2019. 2
- [9] Wei Xu, Marvin J. Dainoff, Liezhong Ge, and Zaifeng Gao. Transitioning to human interaction with ai systems: New challenges and opportunities for hci professionals to enable human-centered ai, 2023. 4
- [10] Wei Xu, Zaifeng Gao, and Marvin Dainoff. An hcai methodological framework: Putting it into action to enable human-centered ai. *arXiv preprint arXiv:2311.16027*, 2023. 2, 3