OVERLOOKED BY DESIGN: CHALLENGES AND OPPORTUNITIES OF INCORPORATING HUMAN-CENTRED PERSPECTIVES AND EDI IN IT DEVELOPMENT DURING THE AI BOOM

Åsa Cajander, Beni Suranto, Tony Clear, Ramesh Lal

Introduction

Over the past decades, Information Technology (IT) has transformed nearly every aspect of daily life, changing how people communicate, learn, work, and socialise. IT systems are now deeply embedded in everyday routines, from widespread internet access and social media platforms like Facebook and Instagram to essential tools such as Gmail, WhatsApp, and online shopping services (Campbell-Kelly et al.). However, a problem remains: the frequent oversight of Human-Centered Design (HCD) in IT development. Historically, IT development has focused more on technical performance, efficiency, and rapid deployment— often at the expense of considering user perspectives (Cajander). Integrating Equality, Diversity, and Inclusion (EDI) principles into IT development amplifies these challenges. In this chapter, "integrating EDI into AI" refers to the deliberate embedding of equality, diversity, and inclusion considerations throughout data design, algorithm development, stakeholder engagement, and governance, to ensure that AI systems mitigate bias and serve diverse communities fairly. And yet, integrating EDI is vital to prevent technology from perpetuating or exacerbating structural inequities, particularly for marginalised communities.

The recent rapid advancements in Generative Artificial Intelligence (GenAI) technologies, exemplified by popular platforms like ChatGPT, Google's Gemini, and various AI-driven creative and analytical tools, have intensified ongoing discussions around the ethical, social, and inclusion-related implications of AI in society. These AI systems are celebrated for their unprecedented capabilities, including generating human-like text, automating complex tasks, enhancing productivity, and providing

personalised experiences in fields ranging from education and healthcare to customer service. However, the widespread adoption of AI also brings substantial risks related to EDI. For instance, AI-driven facial recognition systems frequently show biases against people of colour, leading to potential misidentification and unjust treatment (Varsha). Similarly, automated evaluation systems can unfairly penalise non-native English speakers or culturally diverse students due to implicit linguistic biases in the training data (Agarwal et al.). Such scenarios emphasise the necessity of intentionally integrating inclusive and user-centred approaches into AI systems. In the global higher education landscape, the rapid advancement of AI technologies is increasingly transforming academic environments and institutional processes (from assessment systems to student support services), making it imperative to integrate inclusive and human-centered approaches that ensure equitable, transparent, and culturally responsive learning experiences for all students.

Addressing the challenges of HCD and EDI in IT development requires more than just good intentions. It calls for a fundamental shift in how technology is designed, implemented, and governed. This chapter examines how the rise of AI has not only exacerbated existing barriers but also introduced new risks, reinforcing biases, and making inclusivity even more challenging to achieve. While highlighting the problems, the chapter also focuses on practical solutions, advocating for approaches such as Human-Centred AI and policy reforms that can help make technology fairer, more accessible, and genuinely user-focused. Looking ahead, Coimbra Group universities are uniquely positioned to lead collaborative, cross-sector efforts that advance ethical, inclusive, and human-centred digital transformation. By fostering interdisciplinary research, innovating curricula that address the societal and ethical implications of AI, and shaping institutional policies rooted in EDI, these universities can serve as vital agents of change across academia and beyond. Importantly, their contribution extends to engaging with the IT industry (particularly those developing AI systems) by offering critical insights, ethical frameworks, and evidence-based recommendations that support the creation of responsible and inclusive technologies. Through research-industry partnerships, policy advocacy, and public engagement, Coimbra institutions can help ensure that the rapid pace of AI innovation is guided by human values, mitigates systemic bias, and genuinely empowers all users.

AI and Society

AI trends and impacts

Artificial Intelligence (AI) is a branch of computer science focused on creating computers and machines capable of imitating human intelligence. These intelligent machines are designed to perform tasks typically associated with human cognitive functions, such as learning, problem-solving, decision-making, and adapting to new situations (Stryker and Kavlakoglu). Unlike traditional machines, which operate strictly within predefined rules and explicit instructions, AI systems can understand, predict, and respond to complex patterns, even when provided with incomplete or ambiguous information (Ertel). Rather than viewing the world in black-and-white terms, AI operates within the nuanced grey areas, allowing it to interpret and navigate uncertainty.

AI plays a transformative role in contemporary society, profoundly influencing various sectors through capabilities including data mining, image recognition, natural language processing (NLP), and decision-making support (Spangler et al.; Holmström et al.; Cajander et al.). In data mining, AI enables organisations to analyse extensive datasets, revealing insights that inform personalised user experiences, such as Amazon's tailored product recommendations or Netflix's content suggestions (Hallur et al.). In image recognition, AI technology enhances security systems through facial recognition, improves healthcare accuracy by interpreting medical images, and is fundamental to autonomous vehicles, as demonstrated in Tesla's self-driving systems, enabling them to navigate real- world environments (Alnaggar et al.; Bathla et al.). Additionally, AI supports decision- making processes across various sectors. Financial institutions utilise AI to predict market fluctuations, while healthcare providers rely on AI-driven predictions to personalise patient treatments (Cao and Zhai; Cao et al.; Li et al.).

The AI era: accelerating exclusion

While AI offers benefits, it presents ethical, operational, and societal risks. These challenges demand careful attention from researchers, practitioners, and policymakers to prevent unintended harm.

A primary ethical concern is AI bias, which can reinforce systemic discrimination. For example, a widely used healthcare risk prediction algorithm was found to underestimate the medical needs of Black patients

compared to white patients, due to biased cost-based proxies in its training data (Obermeyer et al.). Similarly, AI-driven hiring tools have been found to favour specific demographics over others (Tilmes). These biases stem from the historical data on which AI models are trained, reflecting societal inequalities and perpetuating them in automated decision-making.

Additionally, the opacity of AI-driven decisions raises concerns about security and privacy. Many AI models operate as "black boxes," making it difficult for users and regulators to understand how decisions are made. This lack of transparency increases the risk of malicious exploitation and reduces public trust in AI systems.

Moreover, organisations that adopt AI face several operational challenges, particularly in decision-making processes. The complexity of AI analytics can lead to flawed business strategies when leaders rely on opaque or biased models without a thorough understanding of their limitations (Rana et al.). Over-reliance on AI in business-to-consumer (B2C) settings can create trust issues, as customers may feel uncomfortable with automated decisions. Similarly, in business-to-business (B2B) environments, power imbalances can deepen when access to AI technologies is unevenly distributed (Keegan et al.).

AI's scalability presents challenges for equity, diversity, and inclusion (EDI). Unlike traditional IT systems, AI operates across vast datasets and diverse user bases, making it challenging to implement nuanced, context-specific adjustments. This scalability often prioritises efficiency over inclusion, leading to the marginalisation of underrepresented groups. For example, academic institutions implementing AI chatbots for student support may unintentionally exclude users with non-standard queries or communication styles. Similarly, automated hiring platforms may disadvantage applicants who do not fit predefined patterns, reinforcing existing inequalities.

Level of Social Analysis	"*Equity & Access*" Issues Individual Domain	Professional Domain
Individuals	Personal biases or privileges may remain unnoticed, resulting in limited awareness of how AI systems fail to serve underrepresented users.	AI developers may inadvertently overlook subtle user needs when pressed by tight deadlines, contributing to incomplete accessibility features.
	Lack of accessible feedback channels can prevent individuals from raising concerns about exclusive design choices.	Maintaining fair data collection can be challenging, as it risks underrepresentation of specific demographics.
Communities & Groups	Community members often lack adequate training or resources, which can deepen existing tech gaps when introducing AI solutions.	Project coordinators often struggle with budget and time constraints, which reduces opportunities for in-depth community engagement.
	Early input on local constraints is rarely gathered, limiting the community's influence on design decisions.	Variations in language, literacy, or cultural norms within a group complicate uniform AI design and deployment.
Organisations	Employees may fear reprisals for spotlighting how AI tools disadvantage certain staff or customers.	Leadership often prioritises cost or efficiency over inclusive design, overlooking marginalised user segments.
	Workers with limited influence often cannot effectively advocate for thorough EDI reviews during procurement or implementation processes.	Inconsistent internal standards create confusion about what level of EDI compliance is acceptable.
Cultures	Individuals from minority cultures risk stereotyping or exclusion if AI defaults reflect only a dominant language or worldview. Data sets lacking cultural nuances can misinterpret behaviour or values outside mainstream norms.	Design teams may rely on monolingual or culturally narrow test sets, yielding suboptimal performance in multicultural contexts. Collaborations with cultural experts can be underfunded or disregarded, resulting in culturally tone-deaf AI outputs.

Institutional Sectors	Students, patients, or customers from low- resource settings can find premium AI-driven services inaccessible or unsuitable. When healthcare or educational support is already minimal, digital-only processes can further exclude vulnerable users.	Sector-wide guidelines for EDI often conflict with internal policies, resulting in minimal or inconsistent enforcement of these guidelines. Partnerships to standardise equitable AI may stall due to competition, limited budgets, or reluctance to share resources.
National	Citizens in rural or disadvantaged areas may face slow broadband rollouts, which can hinder the adoption of AI and skill development. Negative perceptions or fears of AI can discourage certain groups from advocating for better digital infrastructure.	Policymakers and organisations often encounter industry lobbying that resists broad accessibility requirements. National AI initiatives may omit minority languages and local contexts if data is not systematically localised.
Global	Individuals in underrepresented regions face minimal influence on global AI products built for wealthier markets. Linguistic barriers hinder collective action across borders in addressing shared equity issues.	Multinational companies may opt for one-size- fits-all solutions, aggravating digital divides between nations. International frameworks for digital equity are experiencing fragmented adoption, with some regions lacking the necessary resources or political will to implement them effectively.

Table 1. Multi-Level Challenges of AI Exclusion and Equity Gaps, adapted from (C. Dianne Martin et al.)

Why Central to Human-Centred Design Matters in IT Development

Central to Human-Centred Design (HCD) in IT development emphasises creating technologies that understand, prioritise, and adapt to users' needs, experiences, and limitations. HCD is ensuring that technology is intuitive, accessible, and beneficial to users. Key principles include clearly understanding user requirements, recognising their context, and enabling systems to evolve dynamically based on continuous user feedback and environmental factors (Giacomin). Additionally, HCD highlights the importance of effective human-machine teaming — balancing human

judgment and machine efficiency to enhance system usability and outcomes. Particularly in interactive intelligent systems, achieving the right balance between human and machine ensures minimal human effort while keeping performance and quality. This user-focused approach is essential for developing technologies that align with real-world expectations and effectively support human activities (Xu).

Adopting HCD in IT development offers substantial benefits by enhancing usability, accessibility, and overall user satisfaction, ultimately leading to broader adoption and greater success of technological products. For instance, Apple's renowned emphasis on intuitive and user-friendly design in products such as the iPhone and macOS has significantly contributed to widespread user adoption and brand loyalty. Similarly, Google's consistent application of HCD principles in platforms like Gmail and Google Maps prioritises simplicity and user engagement, enabling billions of users to interact effortlessly with complex functionalities. Government services have also benefited notably from HCD. The UK government's redesign of its website (https://www.gov.uk/) significantly improved accessibility and usability for millions of citizens by prioritising clarity, ease of use, and user feedback throughout the development process. Another powerful example is Wheelmap (https://wheelmap.org/), a platform developed by the German non-profit Sozialhelden e.V. This platform exemplifies user-centred innovation by enabling wheelchair users to collaboratively map and share accessibility information, illustrating how participatory and inclusive digital solutions can effectively address the needs of marginalised communities.

In the gadget and hardware domain, adopting HCD principles has notably enhanced accessibility and inclusivity, effectively addressing the diverse needs of users, including those of elderly individuals and people with disabilities (Giaccardi and Redström). For instance, Amazon's Echo smart speakers, equipped with Alexa voice technology, provide crucial assistance for elderly or visually impaired individuals, allowing for effortless interaction through intuitive voice commands for tasks such as reminders, home automation, and easy communication (Amazon). Apple's introduction of accessibility-focused features in gadgets such as the Apple Watch, which includes fall detection and emergency alerts, directly supports elderly users in maintaining their independence and safety. Additionally, Microsoft's Xbox Adaptive Controller demonstrates an exceptional commitment to diversity and inclusion by enabling gam-

ers with limited mobility or physical disabilities to participate fully in gaming experiences that were previously inaccessible to them (Warren).

The experiences of Prague, Zurich, Barcelona, and Stockholm demonstrate how Human-Centered Design (HCD) serves as a foundational strategy for successful digital innovation in urban environments. Across these cities, HCD principles (i.e., participatory design, contextual awareness, inclusivity, and long-term usability) are embedded not just in individual projects but in broader smart city frameworks. Citizen engagement platforms, inclusive digital policies, and the integration of social and environmental concerns into design processes demonstrate a shared commitment to aligning technological development with the lived realities, values, and aspirations of diverse communities (Carboni; Dashkevych and Portnov; Calzada et al.).

The relevance of HCD principles in academic environments is also becoming increasingly evident. For example, the University of Leeds in the UK has applied human- centred design approaches in developing its digital education systems to ensure accessibility and inclusion across diverse student populations. By actively involving students and faculty in the iterative design of platforms like its Virtual Learning Environment (VLE), the university improved usability for learners with disabilities and those from varying digital literacy backgrounds (Lonsdale). Similarly, the Open University has adopted inclusive design practices by co-creating course materials with students from underrepresented groups, ensuring that content delivery methods accommodate varied learning preferences, socioeconomic circumstances, and cultural contexts (The Open University). These cases demonstrate how embedding HCD in university digital infrastructures can promote equitable access to learning, foster student engagement, and enhance educational outcomes, contributing directly to institutional efforts to uphold EDI values in both pedagogy and technology.

These examples show that HCD is more than a design method—it's a key strategy to ensure IT systems serve diverse groups equitably and meaningfully. By embracing values such as inclusion, empathy, adaptability, and iteration, HCD steers IT projects away from narrow, top-down solutions that overlook real user needs. Instead, it promotes systems rooted in the social, cultural, and practical realities of their users.

Challenges in Adopting HCD

Structural challenges in human-centred IT design

Adopting HCD in IT projects involves substantial complexities, primarily due to the difficulty of integrating human factors into established technology development practices. One of the most persistent challenges is managing collaboration across large, interdisciplinary, and multi-stakeholder environments. Effective HCD necessitates clear communication, aligned objectives, and seamless collaboration among technologists, researchers, policymakers, and end-users. However, competing priorities and communication barriers often lead to inconsistencies in applying user-centred design principles (Larusdottir et al.). For instance, in academic-industry partnerships, businesses may prioritise profitability and efficiency, whereas researchers focus on user engagement and long-term usability. These misalignments can result in IT solutions that, despite their technical sophistication, fail to meet the actual needs and expectations of users.

Cultural and organisational barriers further complicate the adoption of HCD, particularly in industries where technological adaptation is expected of users rather than the other way around (Van der Bijl-Brouwer and Dorst). In healthcare, for example, electronic health record (EHR) systems are often designed for administrative efficiency rather than clinical usability, leading to workflow disruptions, increased cognitive load for healthcare professionals, and potential risks to patient safety (Hertzum et al.). Similarly, digital learning platforms in education often struggle with low adoption rates because they are developed without sufficient input from teachers and students (de Souza Rodrigues et al.; Greenhow et al.). This persistent disconnect between developers and end-users highlights how IT solutions that do not actively incorporate user feedback risk becoming ineffective and underutilised.

Despite the abundance of methodologies developed to integrate HCD into systems development, a persistent challenge remains in their limited scope and application. Many of these frameworks focus narrowly on individual IT projects, overlooking the broader organisational, institutional, and legal dimensions required to sustain meaningful Human-Centred Design (HCD) practices. As a result, while iterative prototyping, usability testing, and continuous user feedback are widely advocated, their consistent implementation is often hindered by a lack of organisational support, leadership commitment, and policy integration.

Emerging evidence suggests that the inherent complexity of AI-driven systems can obscure their decision-making processes, often leaving even expert users uncertain about the rationale behind autonomous decisions. Designing AI systems that align with human needs requires a deep understanding of human cognition and AI's evolving capabilities. However, the lack of standardised guidelines for AI-human interaction often leads to the development of opaque, inscrutable systems that users struggle to trust. This detachment of HCD from AI development can lead to biased algorithms, ethical concerns, and the exclusion of marginalised user groups that are disproportionately affected by AI-driven decision-making (Riedl). To better visualise and summarise these interconnected structural barriers to adopting HCD in IT projects, see **Figure 1**.

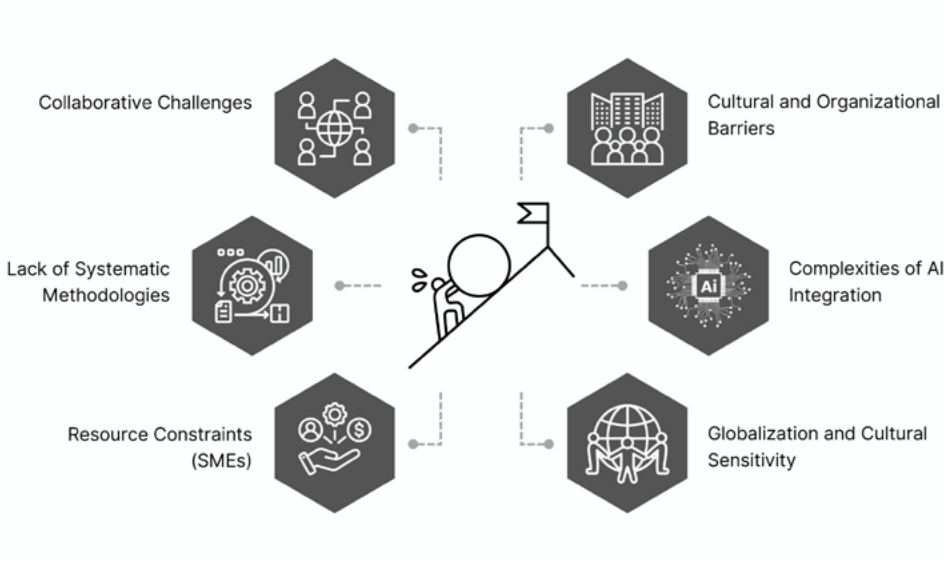

Figure 1. Structural Challenges in HCD (Graphic created by the author using Canva Pro).

Resource constraints further exacerbate the challenges of implementing Human-Centred Design (HCD), particularly for small and medium-sized enterprises (SMEs). Unlike large corporations with dedicated research teams, SMEs often lack the financial and technical resources to invest in extensive user research, usability testing, and iterative design. AI-driven systems, for example, require significant investments in

data quality management, explainability tools, and ethical AI frameworks costs that are often prohibitive for smaller organisations. Moreover, a lack of awareness and understanding of HCD principles among SME leadership frequently results in the undervaluation of user-centred design, leading to systems that fail to accommodate diverse user needs effectively (Gorichanaz).

In an era of globalised systems, cultural sensitivity remains another crucial yet often overlooked aspect of Human-Centred Design (HCD). Culture influences user behaviour, preferences, and interaction styles, but many IT projects fail to incorporate cultural concerns unless explicitly requested by clients. In regions such as India, HCI designers have reported that including cultural considerations in digital product development largely depends on the awareness and commitment of project leadership rather than being an inherent part of the design process. IT solutions risk alienating or being ineffective in diverse markets without cultural adaptability, limiting their global reach and adoption.

Another critical issue is balancing human and technical factors in automation. While automation is often introduced to enhance efficiency and safety, poorly designed automation systems can lead to user frustration, reduced trust, and even catastrophic failures. Studies have shown that systems that fail to account for human behaviour and cognitive limitations can result in over-reliance on automation, making human intervention ineffective when needed (Xu). A collaborative approach to automation is crucial for enhancing, rather than diminishing, human capabilities. This method considers cognitive processes, emotional responses, and the patterns of interaction between humans and systems. By integrating these elements, automation can be designed to work in tandem with human abilities effectively.

User-centred design must be an organisation-wide commitment, not just an initiative within development teams. However, many organisations resist change, particularly those with deeply ingrained technical or engineering-focused cultures, lacking executive buy-in and structured governance models, which can reinforce the notion that HCD, user-centred design, risks becoming an afterthought rather than an integral component of IT development. Overcoming these barriers requires a cultural shift that prioritises human needs alongside technical objectives, ensuring that IT systems are functional, equitable, accessible, and aligned with the diverse needs of their users.

Why EDI presents an even greater challenge

Integrating EDI into IT development is vital not only for ethical and social justice reasons but also as a strategy for inclusive, sustainable innovation. As societies and universities grow more diverse, technologies must reflect and support that diversity. EDI-driven approaches align with global goals, such as the UN Sustainable Development Goals (SDGs), particularly those focused on equity and inclusive growth. Including EDI helps technologies meet demands for fairness, sustainability, and responsibility, boosting long-term acceptance and effectiveness. It also enhances an organisation's reputation, user satisfaction, and societal impact, supporting sustainable development goals.

Despite its importance, integrating EDI into IT projects is challenging due to systemic barriers that exist within the organisation. A key issue is the lack of diversity within tech teams, limiting awareness and representation of varied perspectives. Homogeneous teams often struggle to understand and address the experiences of marginalised groups. Institutional resistance and cultural inertia further impede meaningful change. Developers also face structural hurdles, such as rigid procurement, tight budgets, and short timelines, that deprioritise EDI. As a result, even when its value is recognised, organisations find it difficult to embed EDI into existing tech frameworks, highlighting the depth of these challenges (Aleem and Ahmed).

Integrating EDI principles into IT projects often encounters significant practical barriers. One prominent challenge is the deeply ingrained biases in historical data, which AI systems use for learning and decision-making. For instance, a widely publicised case involving Amazon's recruitment algorithm revealed how the system unintentionally discriminated against women, as it was trained on historical hiring data dominated by male applicants (Iriondo). Despite good intentions, attempts to correct these embedded biases often prove complex and resource-intensive, requiring deliberate adjustments in data selection, system retraining, and continuous oversight (Belenguer; Nazer et al.).

Additionally, cultural and institutional resistance presents substantial hurdles to incorporating EDI effectively. Organisations frequently prioritise rapid technological deployment, cost-efficiency, and immediate results over inclusive and participatory design processes, leaving limited resources for comprehensive diversity assessments or iterative user

testing involving diverse populations (Shams et al.). For example, facial recognition technologies deployed in law enforcement have been criticised for their poor accuracy in identifying people from racial minority groups, resulting in wrongful arrests or misidentifications globally (Bragias et al.).

Several initiatives demonstrate significant efforts to integrate EDI into IT development, particularly in Europe and North America. The Athena SWAN Charter, launched in the UK and internationally recognised, promotes gender equality in STEM by addressing systemic inequities through inclusive policies (Advance HE). Corporate efforts, such as Microsoft's Xbox Adaptive Controller, demonstrate how accessible technology can empower individuals with disabilities to participate in gaming. Google's AI principles emphasise fairness, transparency, and accountability, guiding rigorous evaluation to reduce bias. Its "Fairness Indicators" toolkit offers developers practical tools for embedding fairness in machine learning, proving that inclusive technology is both achievable and essential (Xu and Doshi).

While these efforts highlight notable progress, substantial challenges persist across both academic institutions and the broader IT industry. In academia, despite growing awareness, the integration of EDI principles into curricula, research design, and institutional governance often faces resistance due to deeply embedded traditions, limited resources, or lack of institutional incentives. Similarly, in the IT industry, many organisations remain hesitant to fully integrate EDI principles due to entrenched cultural norms, budgetary constraints, ethical blind spots in the perspectives of IT professionals (Vardi; Pant et al.; Ozkaya), or rigid project management practices. As a result, genuine inclusivity and fairness are inconsistently addressed across the IT sector (Aleem and Ahmed; Kamasak et al.)

Opportunities: Human-Centred AI

The concept of Human-Centred AI (HCAI) presents opportunities to address the intricate challenges associated with integrating Human-Centred Design (HCD) and Equality, Diversity, and Inclusion (EDI) principles into IT projects. At its core, HCAI advocates for developing artificial intelligence systems that prioritise human values, well-being, transparency,

and fairness. For example, IBM has adopted a human-centred approach by developing explainable and trustworthy AI systems, ensuring their products are transparent and understandable to users, thus strengthening trust and usability.

HCAI inherently aligns with HCD principles, particularly through its emphasis on continuous user involvement, adaptability, and iterative design (Shneiderman). AI systems grounded in HCAI practices prioritise user feedback, which helps developers understand how end-users interact with technology in real-world scenarios. Companies like McKinsey emphasise the importance of prioritising people, ensuring that AI systems respond to user needs and adapt dynamically to changing contexts. This user-centred focus improves usability and accessibility, enhancing technology acceptance among diverse populations.

Additionally, HCAI directly addresses critical EDI issues by actively embedding these values into the AI lifecycle. HCAI minimises biases and promotes equitable outcomes through practices involving diverse representation and inclusive stakeholder participation (Régis et al.). For instance, IBM's AI Fairness 360 toolkit provides practical tools for detecting and mitigating bias in AI applications, directly addressing racial and gender biases embedded within training datasets. Similarly, initiatives from Microsoft and Amazon emphasise inclusive design practices, actively involving diverse user groups during the development of AI-powered products, thereby ensuring technology serves marginalised and traditionally underrepresented communities more effectively.

Adopting HCAI principles requires substantial organisational and cultural shifts, particularly within industry and academia. Best practices involve interdisciplinary collaboration, fostering diverse teams capable of addressing broader ethical and social considerations (Auernhammer) Stanford University's Human-Centred Artificial Intelligence (HAI) initiative exemplifies this interdisciplinary approach, bringing together experts from diverse academic fields, including ethics, technology, social sciences, and humanities, to develop AI solutions aligned with societal values. These collaborative environments help ensure that AI technologies are sensitive to social contexts, transparent in their operations, and accountable for their impacts, making them more trustworthy and acceptable across various user demographics.

In Europe, efforts have been undertaken to address HCAI through comprehensive ethical guidelines and policy frameworks. The European

Union's publication of the Ethical Guidelines for Trustworthy AI underscores Europe's commitment to embedding human dignity, fairness, and inclusivity within technological innovation. These guidelines emphasise that AI systems should respect human rights and democratic values, ensure transparency, provide accountability, and prioritise user-centric approaches, making explicit the critical connection between ethical AI and societal well-being (European Commission). Moreover, the EU AI Act, proposed as the world's first legal framework dedicated specifically to AI governance, further solidifies Europe's stance by categorising AI applications according to their potential risks and mandating rigorous assessments and oversight, particularly for systems posing high ethical or societal risks. By enforcing transparency, accountability, and fairness through stringent regulatory measures, the AI Act aims to mitigate bias, protect vulnerable groups, and ensure that AI benefits all segments of society (European Parliament). However, its real-world impact will depend on consistent implementation, effective oversight, and adaptability to rapidly evolving technologies. Nonetheless, the Act marks a significant legislative step toward aligning AI development with ethical, inclusive, and human-centred principles. Collectively, these European initiatives represent steps toward integrating EDI and HCD principles into AI, establishing a clear global benchmark for the development of ethically responsible and inclusive technology.

The six major challenges of HCAI provide a comprehensive roadmap for developing AI systems that are ethical, inclusive, and aligned with human well-being (Ozmen Garibay et al.). These challenges were identified through a global collaboration of experts from academia, industry, and government, emphasising the need for AI systems that respect human values and prioritise user needs. The first challenge highlights the importance of centring AI development around human well-being, ensuring that technologies enhance, rather than undermine, the quality of life. The second and third challenges focus on responsible AI design and privacy, addressing concerns related to bias, transparency, and data protection. Without these foundational principles, AI systems risk exacerbating existing inequalities and failing to uphold ethical standards. The fourth and fifth challenges highlight the need for robust human-centred design (HCD) frameworks and governance mechanisms. Lastly, the sixth challenge calls for AI-human interaction models that align with human cognitive processes. This acknowledges the complexities of human-AI

collaboration and highlights the necessity for AI systems that support, rather than replace, human decision-making (Ozmen Garibay et al.).

Figure 2 illustrates the critical intersection between the six grand challenges of HCAI and EDI concerns, reinforcing the need for AI systems that uphold fairness, inclusivity, and ethical responsibility. *AI-Human Interaction* is crucial for accessibility and adaptability, ensuring that AI accommodates diverse user needs, including those with different languages, disabilities, and cognitive abilities. *Human Well-Being* must also be a central focus, ensuring that AI benefits all demographics equitably, particularly marginalised and underrepresented groups, rather than exacerbating social disparities. To achieve *Responsible AI Design*, eliminating biases that disproportionately impact specific racial, gender, or socioeconomic groups is essential, ensuring AI-driven decisions remain fair and just.

Additionally, *Privacy* protections are critical for safeguarding vulnerable populations, such as marginalised communities or individuals with disabilities, from risks related to surveillance and data misuse. *Human-Centred Design* emphasises the importance of inclusive participation by involving diverse stakeholders in AI development, ensuring that perspectives from traditionally underrepresented groups inform the usability and functionality of AI. Lastly, strong *Governance* frameworks provide the accountability necessary to align AI with ethical and legal standards, preventing discriminatory practices and promoting transparency. Together, these interconnected elements form the foundation for AI systems that are not only technologically advanced but also socially responsible and genuinely inclusive.

Figure 2. The Intersection of HCAI Grand Challenges and EDI Issues in IT Projects. Concept inspired by (Ozmen Garibay et al.).

Suggested Actions for Advancing EDI through Human-Centred AI Within the Coimbra Group

Based on our analysis and examples from member universities and European policy work, we suggest the following seven actions as ideas for discussion. They are meant to inspire debate, encourage local solutions, and help Coimbra Group universities share good practice in promoting equality, diversity, and inclusion through Human-Centred AI.

1. Design for Inclusion by Default. Human-Centred AI (HCAI) begins with intentional inclusivity, to design systems that serve diverse communities from the outset, particularly those that have been historically excluded. Inclusion should not be a patch; it must be the blueprint.
2. Make Data Justice a Strategic Priority. HCAI demands scrutiny of the data that feeds AI. Coimbra universities must lead in auditing training data for bias, diversifying sources, and building fair data infrastructures. Data justice is foundational to the development of ethical and equitable AI.
3. Institutionalise Interdisciplinary Collaboration. No single discipline can develop truly human-centred AI. EDI-aligned HCAI requires

teams that integrate technical expertise with deep competencies in gender studies, disability studies, race and migration research, and broader social sciences and humanities. Include those with relevant lived experience alongside designers and engineers. Build structures where this collaboration is not only encouraged ad hoc, but institutionally embedded across research, teaching, and governance.

4. Centre the Human in AI Interaction. In HCAI, humans are never passive users. Teach and develop systems that support human decision-making, rather than overriding it. Prioritise explainability, usability, and user control, especially in education, healthcare and governance.
5. Resource HCAI as Core Digital Infrastructure. Ethical and inclusive AI cannot run on goodwill alone. Allocate sustained funding to HCAI research, education, and implementation. Integrate HCAI into existing governance structures and review processes, ensuring it becomes part of the institution's long-term responsibilities.
6. Embed HCAI and EDI Literacy Across the University. AI literacy must be broad, critical, and inclusive. Coimbra Group universities should strive to integrate HCAI and EDI literacy into multiple curricula, spanning from engineering to the arts. This empowers students and staff to become thoughtful users, designers, and critics of the digital systems shaping society.
8. Build Coimbra Group Coalitions for Accountable Innovation. The Group universities can lead Europe in operationalising HCAI. Collaborate across institutions to share best practices, create shared governance models, and hold one another accountable. Collective action is essential for transforming AI into a tool for equity.

Taken together, these ideas show ways that Coimbra Group Universities can share knowledge and take practical steps towards more ethical and inclusive AI development.

Acknowledgments

The authors acknowledge the use of generative artificial intelligence in the development of this chapter. AI tools supported language refinement and conceptual structuring; the authors thoroughly reviewed and approved all content.

Works Cited

Advance HE. *The Equality Charters - Good Practice Initiatives | Advance HE*. https://www.advance-he.ac.uk/equality-charters/qood-practice-initiatives. Accessed 25 Mar. 2025.

Agarwal, Avinash, et al. "Fairness Score and Process Standardization: Framework for Fairness Certification in Artificial Intelligence Systems." *AI and Ethics*, vol. 3, no. 1, 2023, 267–79.

Aleem, Saiqa, and Faheem Ahmed. "Practicing Equity Diversity Inclusion (EDI) in Software Development Teams: A Systematic Literature Survey." *IEEE Access*, vol. 11, 2023, 98977–87.

Alnaggar, Omar Abdullah Murshed Farhan, et al. "Efficient Artificial Intelligence Approaches for Medical Image Processing in Healthcare: Comprehensive Review, Taxonomy, and Analysis." *Artificial Intelligence Review*, vol. 57, no. 8, 2024, 221.

Amazon. "Alexa Is Opening Doors for People with Disabilities. Here's How." *UK About Amazon*, 1 Dec. 2023, https://www.aboutamazon.co.uk/news/devices/alexa-features- people-disabilities.

Auernhammer, Jan. *Human-Centered AI: The Role of Human-Centered Design Research in the Development of AI*. 2020.

Bathla, Gourav, et al. "Autonomous Vehicles and Intelligent Automation: Applications, Challenges, and Opportunities." *Mobile Information Systems*, vol. 2022, no. 1, 2022, 7632892.

Belenguer, Lorenzo. "AI Bias: Exploring Discriminatory Algorithmic Decision-Making Models and the Application of Possible Machine-Centric Solutions Adapted from the Pharmaceutical Industry." *AI and Ethics*, vol. 2, no. 4, 2022, 771–87.

Bragias, Adelaide, et al. "'Only in Our Best Interest, Right?'Public Perceptions of Police Use of Facial Recognition Technology." *Police Practice and Research*, vol. 22, no. 6, 2021, 1637–54.

Cajander, Åsa, et al. "The Effects of Automation of a Patient-Centric Service in Primary Care on the Work Engagement and Exhaustion of Nurses." *Quality and User Experience*, vol. 5, no. 1, 2020, 1–13.

---. *Usability–Who Cares?: The Introduction of User-Centred Systems Design in Organisations*. 2010. Acta Universitatis Upsaliensis, PhD Thesis.

Calzada, Igor, et al. "People-Centered Smart Cities: An Exploratory Action Research on the Cities' Coalition for Digital Rights." *Journal of Urban Affairs*, vol. 45, no. 9, 2023, 1537–62.

Campbell-Kelly, Martin, et al. *Computer: A History of the Information Machine*. Routledge, 2023.

Cao, Longbing. "Ai in Finance: Challenges, Techniques, and Opportunities." *ACM Computing Surveys (CSUR)*, vol. 55, no. 3, 2022, 1–38.

Cao, Yi, and Jia Zhai. "A Survey of AI in Finance." *Journal of Chinese Economic and Business Studies*, vol. 20, no. 2, 2022, 125–37.

Carboni, Stefano. "Smart Cities in Comparison: An Analysis of the Best Smart Cities." *ORAȘE INTELIGENTE ȘI DEZVOLTARE REGIONALĂ*, vol. 8, no. 03, 2024, 65–78.

Dashkevych, Oleg, and Boris A. Portnov. "Human-Centric, Sustainability-Driven Approach to Ranking Smart Cities Worldwide." *Technology in Society*, vol. 74, 2023, 102296.

de Souza Rodrigues, Marco Aurélio, et al. "An Exploration of eLearning Adoption in the Educational Ecosystem." *Education and Information Technologies*, vol. 26, no. 1, 2021, 585–615.

Ertel, Wolfgang. *Introduction to Artificial Intelligence*. Springer Nature, 2024.

European Commission. *Ethics Guidelines for Trustworthy AI | Shaping Europe's Digital Future*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy- ai. Accessed 16 Mar. 2025.

European Parliament. "EU AI Act: First Regulation on Artificial Intelligence." *Topics | European Parliament*, 6 Aug. 2023, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first- regulation-on-artificial-intelligence.

Giaccardi, Elisa, and Johan Redström. "Technology and More-than-Human Design." *Design Issues*, vol. 36, no. 4, 2020, 33–44.

Giacomin, Joseph. "What Is Human Centred Design?" *The Design Journal*, vol. 17, no. 4, 2014, 606–23.

Gorichanaz, Tim. "Identifying the Barriers to Human-Centered Design in the Workplace: Perspectives from UX Professionals." *arXiv Preprint arXiv:2412.07045*, 2024.

Greenhow, Christine, et al. "Foundations of Online Learning: Challenges and Opportunities." *Educational Psychologist*, vol. 57, no. 3, 2022, 131–47.

Hallur, Giri Gandu, et al. "Entertainment in Era of AI, Big Data & IoT." *Digital Entertainment: The Next Evolution in Service Sector*, 2021, 87–109.

Hertzum, Morten, et al. "Implementing Large-Scale Electronic Health Records: Experiences from Implementations of Epic in Denmark and Finland." *International Journal of Medical Informatics*, vol. 167, 2022, 104868.

Holmström, Inger K., et al. "Registered Nurses' Experiences of Using a Clinical Decision Support System for Triage of Emergency Calls: A Qualitative Interview Study." *Journal of Advanced Nursing*, vol. 76, no. 11, Nov. 2020, 3104–12. *DOI.org (Crossref)*, https://doi.org/10.1111/jan.14542.

Iriondo, Roberto. "Amazon Scraps Secret AI Recruiting Engine That Showed Biases Against Women - Machine Learning - CMU - Carnegie Mellon Uni-

versity." *Machine Learning | Carnegie Mellon University*, $dateFormat, https://www.ml.cmu.edu//news/news-archive/2016-2020/2018/october/amazon-scraps-secret-artificial-intelligence-recruiting-engine-that-showed-biases-against- women.html.

Kamasak, Rifat, et al. "Emerging Trends of Industry 4.0 in Equality, Diversity, and Inclusion Implementations." *Contemporary Approaches in Equality, Diversity and Inclusion: Strategic and Technological Perspectives*, 2023, 129–48.

Keegan, Brendan James, et al. "Examining the Dark Force Consequences of AI as a New Actor in B2B Relationships." *Industrial Marketing Management*, vol. 115, 2023, 228–39.

Larusdottir, Marta, et al. "A License to Kill–Improving UCSD in Agile Development." *Journal of Systems and Software*, vol. 123, 2017, 214–22.

Li, Yu-Hao, et al. "Innovation and Challenges of Artificial Intelligence Technology in Personalized Healthcare." *Scientific Reports*, vol. 14, no. 1, 2024, 18994.

Lonsdale, Maria. *Using Information and Instructional Design Principles to Enhance Accessibility and Inclusivity of Course Material on Blackboard VLE | Leeds Institute for Teaching Excellence*. https://teachingexcellence.leeds.ac.uk/current-lite- projects/using-information-and-instructional-design-principles-to-enhance- accessibility-and-inclusivity-of-course-material-on-blackboard- vle/?utm_source=chatgpt.com. Accessed 21 June 2025.

Martin, C. Dianne, et al. "A Framework for Implementing and Teaching the Social and Ethical Impact of Computing." *Education and Information Technologies*, vol. 1, 1996, 101–22.

Martin, C.Dianne, et al. "A Framework for Implementing and Teaching the Social and Ethical Impact of Computing." *Education and Information Technologies*, vol. 1, no. 2, June 1996. *DOI.org (Crossref)*, https://doi.org/10.1007/BF00168276.

Nazer, Lama H., et al. "Bias in Artificial Intelligence Algorithms and Recommendations for Mitigation." *PLOS Digital Health*, vol. 2, no. 6, 2023, e0000278.

Obermeyer, Ziad, et al. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science*, vol. 366, no. 6464, 2019, 447–53.

Ozkaya, Ipek. "Ethics Is a Software Design Concern." *IEEE Software*, vol. 36, no. 3, 2019, 4–8.

Ozmen Garibay, Ozlem, et al. "Six Human-Centered Artificial Intelligence Grand Challenges." *International Journal of Human–Computer Interaction*, vol. 39, no. 3, Feb. 2023, 391–437. *DOI.org (Crossref)*, https://doi.org/10.1080/10447318.2022.2153320.

Pant, Aastha, et al. "Ethics in the Age of AI: An Analysis of AI Practitioners' Awareness and Challenges." *ACM Transactions on Software Engineering and Methodology*, vol. 33, no. 3, Mar. 2024, 1–35. *DOI.org (Crossref)*, https://doi.org/10.1145/3635715.

Rana, Nripendra P., et al. "Understanding Dark Side of Artificial Intelligence (AI) Integrated Business Analytics: Assessing Firm's Operational Inefficiency and Competitiveness." *European Journal of Information Systems*, vol. 31, no. 3, 2022, 364–87.

Régis, Catherine, et al. *Human-Centered AI: A Multidisciplinary Perspective for Policy- Makers, Auditors, and Users*. 1st ed., Chapman and Hall/CRC, 2024. *DOI.org (Crossref)*, https://doi.org/10.1201/9781003320791.

Riedl, Mark O. "Human-centered Artificial Intelligence and Machine Learning." *Human Behavior and Emerging Technologies*, vol. 1, no. 1, Jan. 2019, 33–36. *DOI.org (Crossref)*, https://doi.org/10.1002/hbe2.117.

Shams, Rifat Ara, et al. "AI and the Quest for Diversity and Inclusion: A Systematic Literature Review." *AI and Ethics*, 2023, 1–28.

Shneiderman, Ben. *Human-Centered AI*. 1st ed., Oxford University PressOxford, 2022. *DOI.org (Crossref)*, https://doi.org/10.1093/oso/9780192845290.001.0001.

Spangler, Douglas, et al. "A Validation of Machine Learning-Based Risk Scores in the Prehospital Setting." *PloS One*, vol. 14, no. 12, 2019, e0226518.

Stanford University, United States of America, and Jan Auernhammer. *Human-Centered AI: The Role of Human-Centered Design Research in the Development of AI*. 2020. *DOI.org (Crossref)*, https://doi.org/10.21606/drs.2020.282.

Stryker, Cole, and Eda Kavlakoglu. *What Is Artificial Intelligence (AI)? | IBM*. 9 Aug. 2024, https://www.ibm.com/think/topics/artificial-intelligence.

The Open University. "Evaluating the Use of an Inclusive Curriculum Tool in STEM Modules." *eSTEeM*, 24 June 2022, https://www5.open.ac.uk/scholarship-and-innovation/esteem/projects/themes/access-participation-and-success/evaluating-use- inclusive-curriculum-tool-stem.

Tilmes, Nicholas. "Disability, Fairness, and Algorithmic Bias in AI Recruitment." *Ethics and Information Technology*, vol. 24, no. 2, 2022, 21.

Van der Bijl-Brouwer, Mieke, and Kees Dorst. "Advancing the Strategic Impact of Human- Centred Design." *Design Studies*, vol. 53, 2017, 1–23.

Vardi, Moshe Y. "I Was Wrong about the Ethics Crisis." *Communications of the ACM*, vol. 68, no. 1, Jan. 2025, 5–5. *DOI.org (Crossref)*, https://doi.org/10.1145/3704959.

Varsha, PS. "How Can We Manage Biases in Artificial Intelligence Systems–A Systematic Literature Review." *International Journal of Information Management Data Insights*, vol. 3, no. 1, 2023, 100165.

Warren, Tom. *Microsoft's New Xbox Adaptive Joystick Puts Accessibility at the Center | The Verge*. https://www.theverge.com/2024/8/21/24225264/microsoft-xbox-adaptive-joystick-accessiblity-accessory-release-date-pricing. Accessed 11 Mar. 2025.

Xu, Catherina, and Tulsee Doshi. *Fairness Indicators: Scalable Infrastructure for*

Fair ML Systems. https://research.google/blog/fairness-indicators-scalable-infrastructure-for- fair-ml-systems/. Accessed 22 Mar. 2025.

Xu, Wei. "Toward Human-Centered AI: A Perspective from Human-Computer Interaction." *Interactions*, vol. 26, no. 4, June 2019, 42–46. *DOI.org (Crossref)*, https://doi.org/10.1145/3328485.