

# Collaborative Intelligence: A Survey of Human AI Interaction in Professional, Educational and Decision Making Contexts

Sumudu Rangika Samarasinghe<sup>1</sup>

<sup>1</sup>University of Peradeniya Peradeniya, Sri Lanka

sumudurangikasamarasinghe@tutamail.com<sup>1</sup>

## Abstract

The integration of artificial intelligence (AI) into human environments has transformed the dynamics of collaboration across multiple sectors, giving rise to the concept of Collaborative Intelligence, a model that emphasizes synergy between human cognitive capabilities and AI-powered systems. This survey paper examines the evolution of human-AI interaction in professional, educational, and decision-making contexts. It reviews current applications, identifies thematic trends, highlights research gaps, and discusses challenges that impact the seamless cooperation between humans and intelligent systems. In professional environments, AI systems are increasingly used to augment decision-making, automate repetitive tasks, and provide predictive insights. This augmentation allows professional in fields such as healthcare, law, and business to improve efficiency, accuracy, and strategic thinking. In educational settings, AI is used for personalized learning, intelligent tutoring systems, and academic analytics, aiming to support both instructors and learners. Decision-making scenarios ranging from organizational strategy to public policy reveal growing dependence on AI tools for data-driven reasoning, risk assessment, and simulation-based forecasting. Through a thematic survey of scholarly and industry literature, this paper classifies current practices in human-AI collaboration, analyses interface design, trust development, cognitive load sharing, and ethical considerations. A comparative evaluation reveals the distinct requirements and interaction modes in each domain. Additionally, the paper identifies common challenges such as algorithmic transparency, over-reliance, user resistance, and the need for interdisciplinary training. The findings suggest that Collaborative Intelligence is not a static goal but a dynamic and evolving framework. Its success depends on mutual adaptability where humans learn to leverage AI's computational strengths, and AI systems are designed to complement human judgment and creativity. This paper concludes by outlining key research directions aimed at designing more intuitive, ethical, and inclusive human-AI collaborations for the future.

**Keywords:** Collaborative Intelligence, Human-AI Interaction, Professional Applications, Decision-Making Systems, Human-Centered Design, Cognitive Collaboration, Intelligent Systems.

## 1. Introduction

In recent years, the interaction between humans and AI has shifted from passive tool usage to active collaboration. As AI systems become more sophisticated, they are increasingly designed not just to perform tasks independently, but to work alongside humans, enhancing how we think, learn, and make decisions. This paradigm, often referred to as Collaborative Intelligence, represents a growing field of interest across sectors that rely on both human insight and machine precision. Understanding how this collaboration functions in real-world contexts is crucial to shaping the future of intelligent, ethical, and human-centric technology (Dellermann et al., 2021).

### 1.1. Background and Context

The emergence of AI has significantly reshaped the landscape of human-computer interaction. Rather than replacing humans, modern AI systems are increasingly being designed to collaborate with human users, creating hybrid frameworks that combine machine efficiency with human judgment. This convergence is known as *Collaborative Intelligence*, a field that seeks to enhance performance and decision-making through synergy between human intelligence and artificial systems (Seeber et al., 2020). The term extends beyond automation to capture interdependent roles, where humans and AI agents co-perform tasks, adapt to each other's strengths and limitations, and mutually improve outcomes. In contrast to traditional automation that often sidelines human input, collaborative intelligence values human agency, ethical oversight, and creative reasoning, which are not easily replicable by machines (Dignum, 2019; Davenport and Kirby, 2016).

### 1.2. Importance of Human-AI Interaction

As AI becomes embedded in professional systems, educational environments, and high-stakes decision-making processes, the way humans interact with these systems becomes central. Effective human-AI interaction requires intuitive interfaces, mutual trust, transparent communication, and ethical alignment (Amershi et al., 2019; Holstein et al., 2019). From intelligent decision-support in healthcare and law to adaptive learning systems in classrooms and strategic simulations in public policy, the potential of collaborative intelligence lies in its flexibility and domain-specific adaptability. Table 1 below briefly outlines the distinct yet overlapping roles of AI in various domains discussed in this paper.

Table 1: Roles of AI in various domains

| Domain          | Primary AI Function                       | Human Role                              |
|-----------------|-------------------------------------------|-----------------------------------------|
| Professional    | Data analysis, automation, recommendation | Oversight, decision-making, supervision |
| Educational     | Personalized tutoring, learning analytics | Instruction, mentoring, feedback        |
| Decision-Making | Predictive modelling, scenario simulation | Judgment, ethical reasoning, consensus  |

### 1.3. Objectives of the Survey

This survey paper aims to examine the state-of-the-art in human-AI collaboration across three major domains. Also identify and categorize existing interaction models and their underlying principles. Highlighting practical challenges and unresolved research questions and Propose directions for developing more effective and ethical collaboration frameworks.

### 1.4. Scope and Contribution

This paper does not focus on AI development or architecture in isolation, but rather explores interdisciplinary insights from cognitive science, education, ethics, design, and computer science. The scope includes AI applications that require shared autonomy, interactive feedback, and co-dependence between humans and machines.

The rest of the paper is organized as Section 2 describes the methodology used to conduct the survey. Section 3 provides a thematic analysis of Collaborative Intelligence in professional, educational, and decision-making contexts. Section 4 presents a comparative discussion of trends and differences across domains. Section 5 discusses key challenges and gaps in current implementations. Section 6 proposes future research directions. Finally, Section 7 concludes the paper by summarizing key insights.

## **2. Methodology of Survey**

Understanding how collaborative intelligence operates across diverse contexts requires a systematic approach to gathering, selecting, and analyzing relevant literature. This section outlines the methodological framework used to conduct the survey, including the criteria for source selection, the analytical process, and the limitations of the approach. The goal is to ensure a structured and credible basis for examining how humans and AI interact in professional, educational, and decision-making environments.

### **2.1. Survey Approach and Rationale**

This survey paper employs a qualitative thematic approach to explore the current landscape of human-AI interaction within the framework of collaborative intelligence. Rather than focusing on statistical aggregation or experimental comparisons, the study draws upon a wide range of scholarly and industry-based literature to identify conceptual patterns, practical implementations, and emerging concerns (Amershi et al., 2019; Holstein et al., 2019). The objective is to interpret trends, synthesize insights across sectors, and generate a domain-specific understanding of how collaborative intelligence is functioning in real-world applications. The approach enables a deeper investigation into interaction dynamics, ethical considerations, and human roles in AI-augmented environments (Dignum, 2019).

### **2.2. Source Selection and Inclusion Criteria**

The literature selected for this survey includes peer-reviewed journal articles, institutional reports, white papers, and detailed case studies published between 2015 and 2025. The scope encompasses academic sources, professional insights, and policy recommendations that directly address the integration of AI systems into human-centered workflows. Databases such as Google Scholar, Scopus, IEEE Xplore, and ScienceDirect were consulted using targeted search terms including “Collaborative Intelligence,” “Human-AI Interaction,” “AI in Education,” “AI and Decision Making,” “Trust in AI,” and “Human-Centered AI Design” (Meske et al., 2022; Marcus and Davis, 2019). Each source was selected based on its relevance to one or more of the three identified domains: professional environments, educational contexts, and decision-making frameworks. Special attention was given to literature that discusses the cognitive, functional, and ethical aspects of human-AI collaboration (Rajkomar et al., 2018).

### **2.3. Analytical Framework**

A structured thematic coding method was used to examine the collected literature. The analysis was organized to evaluate the role of AI systems, the nature of human engagement, and the patterns of interaction emerging across different sectors. For each domain, the literature was analyzed with attention to how collaboration is structured, the types of tasks shared between humans and machines, and the interfaces through which these collaborations occur (Shneiderman, 2020). Human involvement was assessed in terms of decision-making authority, oversight, and adaptability, while AI systems were examined for their level of autonomy, responsiveness, and alignment with human goals. The domains were then classified according to recurring themes such as workflow augmentation in professional settings, adaptive learning in educational contexts, and predictive reasoning in decision-making scenarios (Luckin and Holmes, 2016). The table below outlines the thematic focus of each domain used in this classification framework.

Table 2: Domain-Based Thematic Classification Used in Survey

| Domain          | Key Themes Identified                                        |
|-----------------|--------------------------------------------------------------|
| Professional    | Workflow augmentation, decision support, AI-human teaming    |
| Educational     | Adaptive learning, personalized tutoring, feedback systems   |
| Decision-Making | Strategic modelling, ethical reasoning, predictive analytics |

## 2.4. Limitations

Despite the wide range of sources consulted, this survey has certain limitations. The fast-paced evolution of AI technologies means that some recent developments may not yet be reflected in published literature. Additionally, this study is limited to English-language sources, which may exclude significant regional innovations or culturally specific approaches to collaborative intelligence (Winfield and Jirotko, 2018). Another limitation is the scope of application-specific data, while this paper provides a broad and thematic overview, it does not undertake in-depth technical assessments or system-level performance evaluations.

## 3. Thematic Classification and Analysis

The analysis of human-AI collaboration reveals distinct thematic patterns within professional, educational, and decision-making domains. Each domain reflects a unique configuration of interaction models, technological functions, and human roles. While the overarching principle of collaborative intelligence shared cognitive labour between humans and machines remains consistent, the practical forms it takes are shaped by contextual demands and institutional goals. This section presents a detailed classification and examination of how collaboration is structured and operationalized across these three major contexts.

### 3.1. Collaborative Intelligence in Professional Environments

In professional settings, AI is widely adopted to enhance human performance rather than replace it. This augmentation model is particularly visible in domains such as healthcare, finance, law, and customer service. For example, in clinical diagnosis, AI systems provide predictive insights based on large-scale datasets, while human practitioners interpret results within the context of patient history and ethical guidelines (Rajkomar et al., 2018). Similar synergies exist in law, where AI-driven tools assist in legal research and contract analysis, while final judgments are rendered by human experts.

The success of collaborative intelligence in professional environments depends largely on the quality of user interfaces, system explainability, and the adaptability of AI models to domain-specific workflows. Research shows that trust-building mechanisms, such as transparent algorithms and feedback loops, are essential for ensuring sustained human engagement (Holstein et al., 2019). Furthermore, AI is most effective when it acts as a decision-support agent handling data-heavy tasks and leaving high-level strategy and judgment to humans. This complementary model reduces cognitive burden while expanding the scope of human decision-making (Davenport and Kirby, 2016).

### 3.2. Collaborative Intelligence in Educational Contexts

In educational settings, collaborative intelligence focuses on personalization, adaptability, and continuous feedback. AI-powered platforms are increasingly being used to tailor learning experiences to individual students. Intelligent tutoring systems, learning analytics dashboards, and virtual learning assistants are designed to interact directly with students, assess their progress, and adjust instructional content accordingly (Luckin and Holmes, 2016). At the same time, these systems offer real-time insights to educators, helping them identify learning gaps and intervene effectively.

Unlike professional contexts, educational applications require AI to be sensitive to emotional, motivational, and developmental factors. Collaborative intelligence in education is therefore not just about delivering content efficiently but about supporting the co-evolution of human learning with

machine feedback. Human teachers play a central role in interpreting AI-generated data, fostering social and emotional development, and ensuring ethical use of technology. Studies also indicate that students respond better when AI tools exhibit conversational transparency and scaffold learning rather than simply automate tasks.

### **3.3. Collaborative Intelligence in Decision-Making Contexts**

In decision-making scenarios, collaborative intelligence plays a strategic role by enhancing the scope, speed, and foresight of human judgment. Whether in government policymaking, emergency response, corporate planning, or environmental forecasting, AI systems are employed to process vast datasets, model scenarios, and predict outcomes (Marcus and Davis, 2019). These systems do not make final decisions but assist human stakeholders in evaluating alternatives, anticipating risks, and simulating potential consequences. Decision-making contexts require a high degree of ethical sensitivity and accountability. Collaborative intelligence here must be explainable, auditable, and aligned with public or organizational values. Human agents are responsible for interpreting AI-generated recommendations and ensuring that decisions reflect social, legal, and environmental considerations. Transparency, traceability, and stakeholder participation are key principles guiding human-AI collaboration in this area (Binns et al., 2018).

### **3.4. Cross-Domain Insights**

While the three domains examined differ in their specific applications, they share several unifying themes. Effective human-AI interaction is characterized by shared autonomy, mutual adaptability, transparent interfaces, and ethical alignment. In each context, humans remain central not only as decision-makers but also as co-designers and evaluators of AI systems (Winfield and Jirotko, 2018). The challenges of trust, fairness, interpretability, and user training are recurring across domains and form the foundation for further exploration in the next sections.

## **4. Comparative Discussion**

The comparative analysis of human-AI collaboration across professional, educational, and decision-making contexts reveals both converging trends and distinct variations shaped by domain-specific demands. While all three contexts employ collaborative intelligence to enhance human capacities, the degree of autonomy, interaction design, ethical emphasis, and expected outcomes differ substantially. This section synthesizes the insights drawn from thematic classification to better understand how collaborative intelligence adapts to varying real-world scenarios.

### **4.1. Commonalities in Collaborative Models**

Across all domains, a central theme is the augmentation of human capabilities rather than substitution (Davenport and Kirby, 2016). AI systems are primarily designed to support complex human tasks, offering assistance in data processing, pattern recognition, and adaptive feedback. The role of the human remains crucial, not just in decision-making, but also in interpreting, supervising, and fine-tuning AI outputs. Another common element is the growing emphasis on trust and transparency. Whether it is a healthcare professional interpreting AI diagnostics, a teacher relying on learning analytics, or a policymaker reviewing predictive models, the trustworthiness of AI outputs plays a decisive role in human adoption. Systems that provide clear explanations, visual cues, and opportunities for user input are consistently preferred (Shneiderman, 2020). Similarly, human oversight remains non-negotiable in all contexts, with the most effective models integrating shared control and feedback loops to balance machine autonomy with human judgment. Additionally, each domain places value on mutual adaptability. AI systems must learn from users, just as users adapt to AI suggestions. This reciprocity ensures sustainable and meaningful collaboration, enhancing user satisfaction and system relevance over time.

## 4.2. Differences in Human-AI Interaction Dynamics

Although these shared principles, the nature of interaction and the goals of collaboration vary across domains. In professional settings, human-AI collaboration is often driven by performance efficiency and operational accuracy. Interactions tend to be task-focused, with AI systems acting as assistants or co-pilots that help navigate high-information environments. Time sensitivity and accountability also shape the interaction design, requiring high precision and low error margins (Marcus and Davis, 2019).

In educational contexts, the focus shifts toward personalized learning and cognitive development. The interaction is more sustained and relational, requiring AI to adapt to individual learning styles and emotional states. Here, collaborative intelligence is evaluated not only in terms of knowledge delivery but also in how well it supports motivation, inclusivity, and learner autonomy. The human element typically educators and learners remains emotionally and socially engaged with the AI system, adding layers of complexity to interaction design (Holstein et al., 2019).

Decision-making scenarios present yet another variation, where AI systems are leveraged for strategic insight and scenario analysis. The collaboration is often high-stakes, involving long-term consequences, multiple stakeholders, and high degrees of uncertainty. In such environments, the emphasis is on deliberation, accountability, and ethical reasoning. Human-AI interaction here is characterized by interpretability, multi-perspective presentation, and the need for formal validation procedures. Unlike the more instructional or operational goals of the other domains, decision-making contexts require AI to serve as a cognitive partner in complex, ethically sensitive environments.

## 4.3. Sector-Specific Challenges and Expectations

Each domain also faces unique challenges. Professional sectors often struggle with issues of workflow integration, job displacement concerns, and the explainability of AI recommendations. Educational systems face difficulties related to bias in data, digital divides, and balancing automation with human instruction. In decision-making contexts, the main challenges are related to ethical accountability, data governance, and the potential for algorithmic bias to influence high-level outcomes (Binns et al., 2018; Dellermann et al., 2021).

## 4.4. Implications for Future Collaborative AI Systems

The cross-domain insights suggest that future AI systems must be context-aware, value-sensitive, and designed for co-evolution with human users. Designers must anticipate the social, emotional, and cognitive dimensions of interaction, not just the technical performance. Moreover, human-AI collaboration cannot be one-size-fits-all; systems must reflect the priorities and constraints of the specific domain they serve. A flexible, adaptive approach that combines ethical design, technical robustness, and human-centred values will be essential for building trust, acceptance, and effectiveness across diverse sectors.

## 5. Challenges in Implementing Collaborative Intelligence

While the promise of collaborative intelligence is rapidly gaining momentum, its practical implementation across real-world domains remains fraught with complex challenges. These barriers are not solely technical; they span ethical, organizational, and social dimensions that influence the success, scalability, and sustainability of human-AI interaction. Understanding these challenges is essential to designing systems that are not only intelligent but also accountable, inclusive, and contextually appropriate.

### 5.1. Technical Limitations and Design Complexity

One of the foremost challenges lies in the technical limitations of current AI systems. Most collaborative applications rely on machine learning models that are data-dependent and often function as black boxes (Marcus and Davis, 2019). This lack of interpretability hinders human understanding



and can reduce user trust. In high-stakes environments such as medicine or public policy, the inability to explain how an AI system arrived at its recommendation poses a serious risk. Another major issue is interface design. Building intuitive, human-friendly systems that allow seamless cooperation requires extensive user testing, domain-specific adaptation, and continuous calibration. Current systems often fall short in real-time interaction quality, context sensitivity, and adaptability to diverse user behaviours or needs. Additionally, human cognitive models are complex and unpredictable, making it difficult for AI to anticipate user intentions or preferences effectively.

## **5.2. Ethical and Societal Risks**

Ethical concerns play a significant role in shaping public perception and institutional acceptance of collaborative intelligence. The presence of algorithmic bias, often inherited from training data, can result in unfair outcomes, particularly in education and hiring decisions. This is exacerbated when users lack the technical literacy to question or challenge AI outputs. Furthermore, there are growing concerns over the dehumanization of tasks, where over-reliance on AI may lead to deskilling, reduced critical thinking, and erosion of professional agency (Davenport and Kirby, 2016). In educational contexts, for instance, excessive dependence on automated feedback may undermine teacher authority or diminish student engagement. Ethical dilemmas also emerge around consent, data privacy, and surveillance especially when AI systems are embedded in learning platforms or workplace monitoring tools.

## **5.3. Organizational Resistance and Cultural Barriers**

Successful implementation of collaborative intelligence demands significant organizational change, including redesigning workflows, training personnel, and shifting workplace culture (Seeber et al., 2020). Many institutions face internal resistance due to fear of job displacement, lack of technical understanding, or skepticism about AI's reliability. Change management remains a weak point, with many organizations lacking clear strategies to integrate AI while preserving human roles (Dellermann et al., 2021). There is also a cultural dimension to resistance. Different regions and communities perceive AI differently based on cultural values, socio-political history, and levels of technological exposure. As a result, systems designed in one cultural or institutional context may not translate effectively into another, leading to poor user adoption or misuse.

## **5.4. Governance, Regulation, and Accountability**

A consistent challenge across all domains is the lack of robust regulatory frameworks governing collaborative AI systems. Existing legal structures often fail to account for hybrid decision-making scenarios where responsibility is shared between humans and machines. In the absence of clear guidelines, accountability becomes ambiguous, especially in cases of system failure or unethical outcomes.

## **5.5. Sustainability and Inclusivity Concerns**

Finally, the challenge of ensuring that collaborative intelligence is sustainable and inclusive cannot be overstated. AI development and deployment often concentrate in technologically advanced regions, leaving behind under-resourced schools, rural clinics, and low-income communities. This creates a digital divide that undermines the goal of equitable collaboration. Moreover, sustainable AI systems must consider energy consumption, data ethics, and long-term social impact. Designing systems that adapt over time, require fewer resources, and serve diverse populations is a long-term challenge that will define the future viability of collaborative intelligence.

## 6. Future Directions

The evolution of collaborative intelligence will rely on more human-aligned, transparent, and accessible systems. As AI becomes more integrated into professional, educational, and decision-making domains, several focused directions must shape its advancement.

### 6.1. Human-Centred Design and Mutual Learning

Future systems should prioritize co-adaptive design, where humans shape how AI behaves and vice versa (Holstein et al., 2019). Interfaces need to become more intuitive, responsive, and supportive of real-time interaction. Embedding mutual learning, where AI adapts to user behaviour and users develop AI literacy, will strengthen collaboration and trust.

### 6.2. Enhancing Explainability and Transparency

Explainable AI must become standard, especially in sensitive contexts like healthcare and policymaking. Systems should be able to justify their outputs in clear, non-technical language to support informed human decisions. Transparent systems will foster greater trust and allow for meaningful oversight.

### 6.3. Ethical Regulation and Responsible Deployment

Ethical frameworks should evolve alongside technology. This includes ensuring fairness, privacy protection, accountability, and reducing algorithmic bias. Regulations tailored to specific sectors must be put in place to manage risks and clarify responsibility in shared decision-making processes.

### 6.4. Bridging the Digital Divide

Future efforts must extend beyond advanced institutions to include underserved communities. This means creating affordable, inclusive, and culturally adaptive AI systems, particularly for schools, clinics, and public services in low-resource settings. Ensuring equitable access will prevent further digital inequality.

### 6.5. Interdisciplinary Research and Collaboration

The future of collaborative intelligence requires cross-sector and cross-disciplinary collaboration. Psychologists, technologists, educators, and policymakers must work together to align AI with human values. Shared research agendas should explore not only efficiency, but also social impact and long-term user experience.

## 7. Conclusion

Collaborative intelligence marks a transformative phase in the evolution of human-AI interaction. Rather than replacing human judgment, AI is increasingly being integrated as a cognitive partner that enhances decision-making, personalizes education, and streamlines professional processes. This survey has explored how collaborative intelligence operates in three key domains professional, educational, and decision-making highlighting shared principles such as augmentation, trust, and adaptability, while also recognizing the domain-specific dynamics that shape implementation. Despite significant promise, the practical deployment of collaborative systems still faces challenges. Technical limitations, ethical risks, organizational resistance, and access inequality all affect the effectiveness and fairness of human-AI collaboration. These barriers, if not addressed, may limit the broader societal benefits that collaborative intelligence can offer. Looking forward, the success of collaborative intelligence depends on responsible design, inclusive deployment, and interdisciplinary cooperation. Human-centred approaches, explainable interfaces, and robust governance mechanisms will be essential to creating systems that not only assist but empower users. With thoughtful innovation and ethical foresight, collaborative intelligence can help reshape the future of work, education, and governance anchored in trust, accountability, and shared human values.



## References

- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K. and Teevan, J., 2019, May. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1-13).
- Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J. and Shadbolt, N., 2018, April. 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi conference on human factors in computing systems* (pp. 1-14).
- Davenport, T.H. and Kirby, J., 2016. *Only humans need apply: Winners and losers in the age of smart machines* (Vol. 1). New York: Harper Business.
- Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S. and Ebel, P., 2021. The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems. *arXiv preprint arXiv:2105.03354*.
- Dignum, V., 2019. *Responsible artificial intelligence: how to develop and use AI in a responsible way* (Vol. 2156). Cham: Springer.
- Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudik, M. and Wallach, H., 2019, May. Improving fairness in machine learning systems: What do industry practitioners need?. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-16).
- Luckin, R. and Holmes, W., 2016. Intelligence unleashed: An argument for AI in education.
- Marcus, G. and Davis, E., 2019. *Rebooting AI: Building artificial intelligence we can trust*. Vintage.
- Meske, C., Bunde, E., Schneider, J. and Gersch, M., 2022. Explainable artificial intelligence: objectives, stakeholders, and future research opportunities. *Information systems management*, 39(1), pp.53-63.
- Rajkomar, A., Hardt, M., Howell, M.D., Corrado, G. and Chin, M.H., 2018. Ensuring fairness in machine learning to advance health equity. *Annals of internal medicine*, 169(12), pp.866-872.
- Seeber, I., Bittner, E., Briggs, R.O., De Vreede, T., De Vreede, G.J., Elkins, A., Maier, R., Merz, A.B., Oesterle, S., Randrup, N. and Schwabe, G., 2020. Machines as teammates: A research agenda on AI in team collaboration. *Information & management*, 57(2), p.103174.
- Shneiderman, B., 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6), pp.495-504.
- Winfield, A.F. and Jirotko, M., 2018. Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133), p.20180085.