

***HUMAN BEHAVIOUR AND ETHICS IN AUTOMATED AI SYSTEMS: A
SURVEY-BASED STUDY***

Vidhi Soni

Psychology Department

C. B. Patel Arts College

Abstract

This study explores human behaviour and ethical concerns related to automated Artificial Intelligence (AI) systems using a survey-based approach. As AI systems are increasingly used in areas such as education, healthcare, employment, and governance, ethical issues like bias, transparency, accountability, privacy, and human responsibility have become more important. A structured questionnaire with 17 questions was given to 106 participants, including students, educators, and working professionals who have basic awareness of AI technologies.

The results show that 72.6% of respondents believe that automated AI systems strongly influence human decision-making. Many participants (67%) feel that relying too much on AI reduces human decision-making skills and critical thinking. In addition, 64.2% of respondents said that people feel less responsible when decisions are made by AI systems. A large majority (78.3%) believe that users have a moral right to understand how AI systems make decisions that affect them. More than half of the participants (61.3%) support the use of Human-in-the-Loop (HITL) systems to ensure ethical outcomes. Concerns about fairness and bias were also common, with 69.8% agreeing that reducing bias in AI requires continuous human involvement. Privacy was another major concern, as 74.5% of respondents believe that large-scale data collection by AI systems threatens individual privacy.

The study concludes that although automated AI systems can improve efficiency and consistency, they also create serious ethical challenges. These challenges cannot be solved by technology alone. Continuous human oversight, clear accountability, and ethical governance are necessary to ensure that AI systems remain human-centred and responsible.

Introduction

The rapid growth of Artificial Intelligence (AI) and automated technologies has dramatically changed how decisions are made and actions are carried out in many areas of life. Today, AI systems are commonly used in education, healthcare, finance, government, hiring, surveillance, and online services, taking over tasks that people used to do themselves. With the help of automation, we can process information faster, analyse huge amounts of data, and make predictions more efficiently and consistently. But while these advances bring many benefits, they also raise important concerns about how people behave and what is considered ethical.

Automation in AI means that smart systems can do complicated tasks on their own, with little or no help from people. These tasks include learning new things, making decisions, and adapting to different situations. While automation makes things easier and more efficient, it also changes how people work with technology and what their roles are. How people trust, accept, or resist AI, and how they think about what is right or wrong, is really important in deciding how these systems are made and used. The way people and organizations interact with AI affects not just how well it works, but also whether it is fair and ethical.

One of the central ethical challenges associated with automated AI systems is algorithmic bias. AI systems are often trained on large datasets that may reflect existing social inequalities, cultural assumptions, or historical discrimination. When these biases are embedded in automated decision-making processes, they can lead to unfair, discriminatory, or exclusionary outcomes. Such risks are particularly concerning in high-stakes contexts, such as educational assessment, hiring, credit allocation, and healthcare diagnostics, where automated decisions can have a significant impact on human lives. Consequently, ethical principles such as fairness, non-discrimination, and justice have become essential considerations in the governance of AI.

Transparency and explainability represent another critical ethical dimension of automated AI systems. Many AI models, particularly those based on complex machine learning and deep learning techniques, operate as “black boxes,” making it difficult for users to understand how specific decisions are reached. This lack of transparency can erode trust, limit accountability, and reduce users’ ability to challenge or contest automated outcomes. From a human behavioural perspective, trust in AI systems is strongly influenced by the

perceived transparency, reliability, and interpretability of automated decisions. Without adequate explainability, users may either over-reliance on AI systems or reject them altogether.

Accountability and responsibility in automated AI systems also pose significant ethical challenges. As AI systems gain autonomy, questions arise regarding who should be held responsible when automated decisions result in harm, error, or injustice. While AI systems themselves cannot bear moral or legal responsibility, accountability must be clearly assigned to human actors, including developers, designers, institutions, and policymakers. The absence of well-defined accountability frameworks can lead to ethical ambiguity and reduced public confidence in automated technologies.

Human behaviour in relation to automated AI systems is further shaped by concerns about privacy and data protection. Automated systems often rely on extensive personal and behavioural data to function effectively. The collection, storage, and processing of such data raise ethical issues related to informed consent, surveillance, and misuse of personal information. Individuals' willingness to engage with AI systems is closely linked to their perception of how safely and ethically their data is handled.

Recognizing these challenges, global policy frameworks and ethical guidelines, such as those proposed by UNESCO and national education policies, emphasize the importance of human-centric and ethically governed AI.

Despite the growing body of theoretical and policy-oriented literature on AI ethics, there remains a need for empirical research that examines how individuals perceive and respond to ethical issues in automated AI systems. Understanding human behaviour, attitudes, and ethical awareness is essential for designing AI systems that are not only technologically advanced but also socially responsible and ethically sustainable. In this context, the present study adopts a survey-based approach to explore human behaviour and ethical considerations in automated AI systems, aiming to contribute empirical insights to the ongoing discourse on responsible AI adoption.

In the context of education, the integration of automated AI systems has gained significant importance following the introduction of the National Education Policy (NEP)

2020 in India. NEP 2020 emphasizes the strategic use of technology to enhance access, equity, quality, and accountability in education while explicitly advocating for a human-centric and ethical approach to technological adoption. The policy recognizes the potential of AI-driven automation in areas such as personalized learning, assessment, administrative efficiency, and teacher support, while simultaneously cautioning against the risks of excessive automation, bias, and loss of human judgment.

NEP 2020 underscores the necessity of maintaining ethical governance, data privacy, transparency, and human oversight in AI-enabled educational systems. It positions teachers as facilitators and ethical mediators rather than passive users of technology, reinforcing the principle that automation should augment not replace human decision-making. Within this framework, understanding human behaviour, trust, and ethical perceptions toward automated AI systems becomes particularly relevant, as these factors directly influence the successful and responsible implementation of AI in education. Accordingly, the present study aligns with the vision of NEP 2020 by examining human behavioural responses and ethical considerations associated with automated AI systems in educational contexts.

Objectives of the Study

The present study aims to examine human behaviour and ethical concerns associated with automated Artificial Intelligence (AI) systems through a survey-based approach. It seeks to understand how automated AI systems influence human decision-making processes and the extent to which individuals rely on AI-generated recommendations. The study further aims to analyse the ethical challenges that arise from increasing automation, particularly in relation to traditional ethical frameworks, moral responsibility, and accountability when automated decisions lead to harm or negative outcomes. Another key objective is to investigate the presence and impact of bias and discrimination in AI systems, especially with respect to marginalized or vulnerable groups. The study also focuses on understanding user perceptions of transparency, explainability, and trust in AI-driven decision-making. In addition, it examines the role of human intervention, or Human-in-the-Loop mechanisms, in ensuring ethical outcomes in automated systems. The research explores the effects of over-reliance on AI automation on human cognitive abilities, ethical judgment, and sense of responsibility, along with concerns related to data privacy, surveillance, and personal autonomy resulting from large-scale data collection. Finally, the study aims to identify the

most critical ethical challenges posed by automated AI systems to human behaviour across cognitive, social, ethical, economic, and privacy dimensions, as perceived by the respondents.

Research Methodology

This study uses a descriptive, survey-based research approach to explore human behaviour and ethical concerns related to automated Artificial Intelligence (AI) systems. A quantitative method was used to collect data from respondents to understand their opinions, experiences, and perceptions of AI-driven automation and its ethical implications.

Population and Sample

The study targeted students, educators, researchers, and professionals who have some awareness or experience with AI technologies. A convenience sampling method was used to select respondents, as it allowed for easy access and quicker data collection. A total of 106 respondents took part in the survey, providing a variety of perspectives on AI and its ethical challenges.

Research Instrument

Data were collected using a structured questionnaire designed for this study. It included 17 questions, combining close-ended and opinion-based items. Most questions were in a Yes / No / Your Opinion format, allowing respondents to agree, disagree, or share personal views. The questionnaire focused on topics such as decision-making influence, ethical challenges, bias, transparency, accountability, human-in-the-loop, trust, privacy, over-reliance on AI, and wider social impacts. One multiple-response question identified the most important ethical challenges of AI systems.

Data Collection Procedure

The survey was conducted online using Google Forms for easy distribution and response collection. Participation was voluntary, and respondents were informed that the survey was for academic research purposes. Confidentiality and anonymity were maintained, and no personal information was collected.

Ethical Considerations

Ethical standards were followed throughout the study. Informed consent was obtained from all participants, who were assured of confidentiality, voluntary participation, and the right to withdraw at any time. The study respected participants' privacy and autonomy while maintaining ethical research practices.

Literature Review:

Previous studies highlight that automated AI systems can improve efficiency and decision accuracy; however, they may also reinforce existing social biases if trained on biased data *Floridi et al., 2018*. Research emphasizes the importance of transparency and explainability to build trust among users *Doshi-Velez & Kim, 2017*.

Several scholars argue that excessive automation may reduce human agency and moral responsibility *Coeckelbergh, 2020*. Ethical AI frameworks proposed by UNESCO and other global organizations stress human-centric AI, accountability, and respect for human rights. Studies also indicate that user behaviour, trust, and acceptance significantly influence the success of AI adoption.

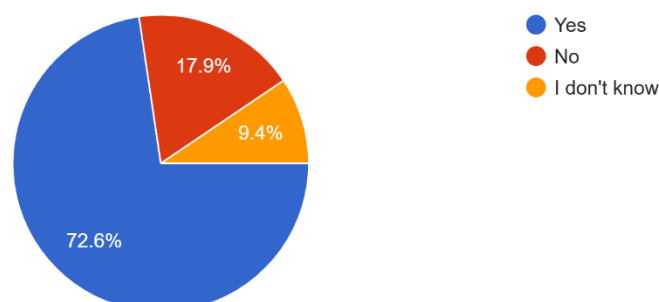
Despite growing research, there remains a gap in empirical studies examining human behaviour and ethical perceptions toward automated AI systems, particularly through survey-based evidence.

Result & Discussion:

Question 1 and 2 were about name and the gender

3. Do automated AI systems significantly influence human decision-making processes?

3. क्या स्वचालित AI प्रणालियाँ मानव निर्णय-निर्माण प्रक्रियाओं को महत्वपूर्ण रूप से प्रभावित करती हैं?

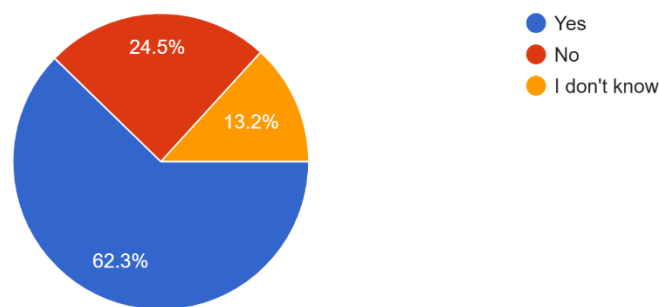


4. Does the increasing automation of AI systems challenge traditional ethical frameworks?

Example - 1 Automated decisions may not fit well with existing human-centred moral or legal rules.

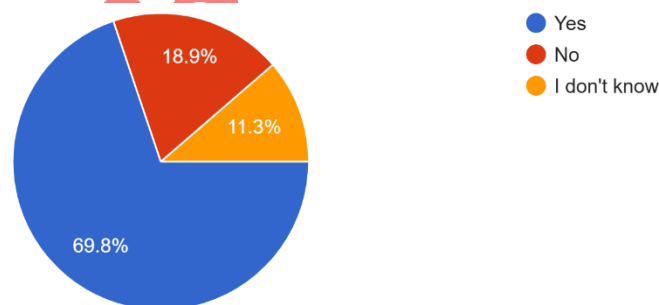
4. क्या AI प्रणालियों में बढ़ता स्वचालन पारंपरिक नैतिक ढाँचों को चुनौती देता है?

उदाहरण - 1 स्वचालित निर्णय मौजूदा मानव-केंद्रित नैतिक या कानूनी नियमों के अनुरूप नहीं हो सकते।



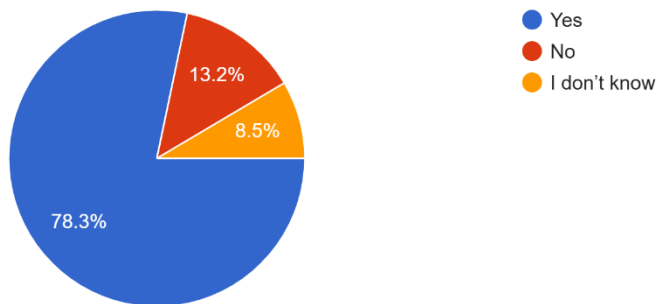
5. Does bias mitigation in AI systems require continuous human intervention rather than purely technical solutions?

5. क्या AI प्रणालियों में पक्षपात को कम करने के लिए केवल तकनीकी समाधान के बजाय निरंतर मानवीय हस्तक्षेप आवश्यक है?



6. Do users have a moral right to understand how automated AI systems make decisions that affect them?

6. क्या उपयोगकर्ताओं को उन स्वचालित AI निर्णयों को समझने का नैतिक अधिकार है जो उन्हें प्रभावित करते हैं?

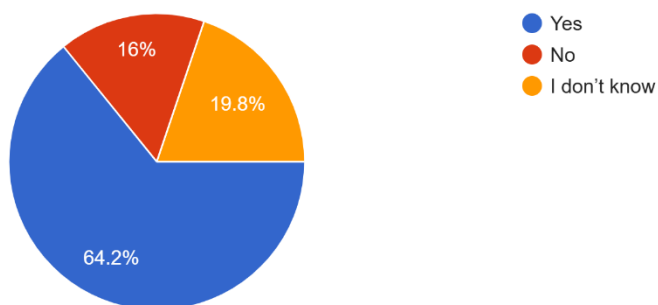


7. Does determining accountability become difficult when harm is caused by an automated AI system?

Example: In the case of a self-driving car accident, it is unclear whether the blame lies with the developer, manufacturer, user, or the AI system.

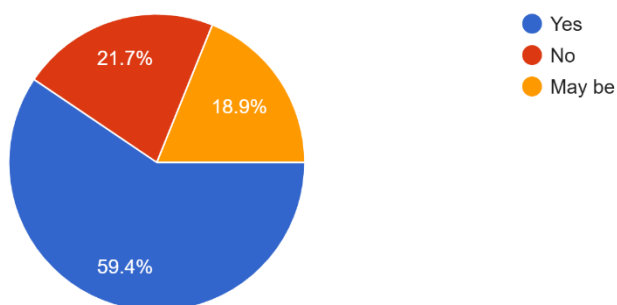
7. क्या स्वचालित AI प्रणाली से नुकसान होने पर उत्तरदायित्व निर्धारित करना कठिन हो जाता है?

उदाहरण: स्वचालित वाहन दुर्घटना की स्थिति में यह स्पष्ट नहीं होता कि दोष डेवलपर, निर्माता, उपयोगकर्ता या AI प्रणाली का है।



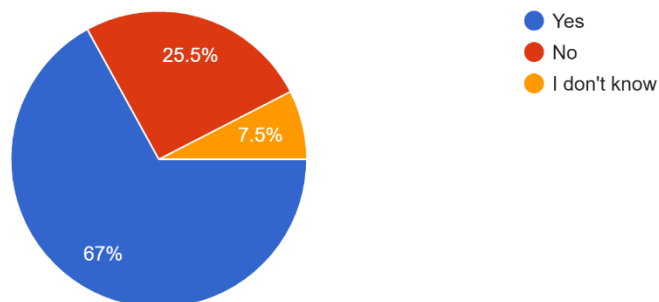
8. Should developers and organizations be held ethically responsible for the outcomes of AI systems?

8. क्या AI प्रणालियों के परिणामों के लिए डेवलपर्स और संगठनों को नैतिक रूप से जिम्मेदार ठहराया जाना चाहिए?



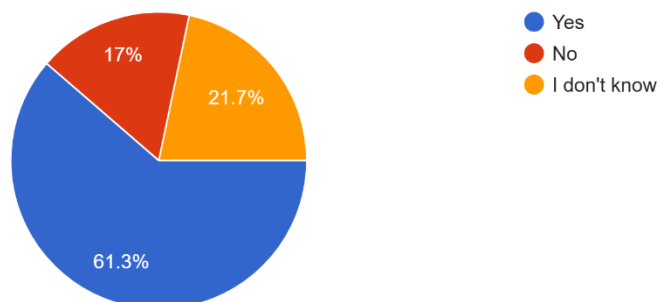
9. Can over-reliance on automated AI systems reduce human decision-making skills?

9. क्या स्वचालित AI प्रणालियों पर अत्यधिक निर्भरता मानव निर्णय क्षमता को कम कर सकती है?



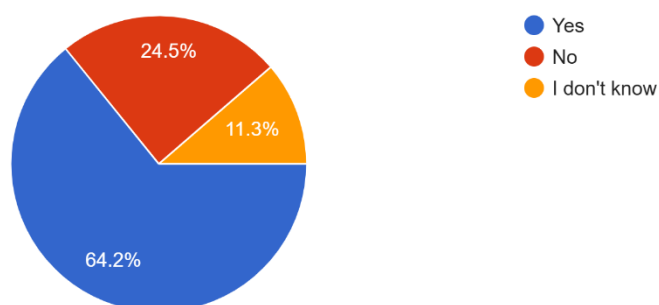
10. Are Human-in-the-Loop (HITL) systems essential to ensure ethical AI outcomes?

10. क्या नैतिक AI परिणाम सुनिश्चित करने के लिए “Human-in-the-Loop” प्रणाली आवश्यक है?



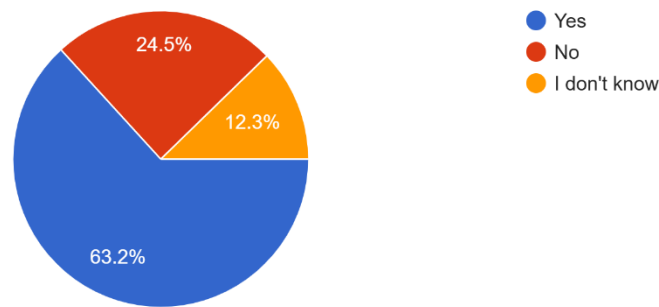
11. Do humans tend to feel less responsible for outcomes when decisions are made by AI systems?

11. क्या जब निर्णय AI प्रणाली द्वारा लिए जाते हैं, तो मनुष्य परिणामों के प्रति कम जिम्मेदारी महसूस करते हैं?



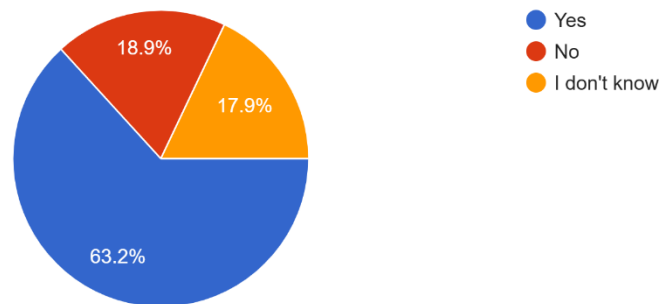
12. Do people often over-trust AI recommendations even when they conflict with human judgment?

12. क्या लोग अक्सर AI की सिफारिशों पर अत्यधिक विश्वास करते हैं, भले ही वे मानवीय निर्णय से टकराती हों?



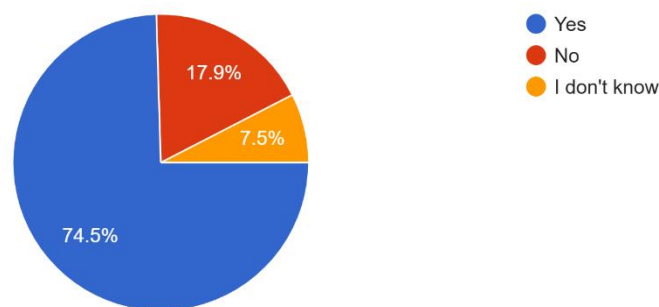
13. Does attributing human-like emotions or intentions to AI systems affect ethical judgment?

13. क्या AI प्रणालियों को मानवीय भावनाएँ या इरादे देना नैतिक निर्णय को प्रभावित करता है?



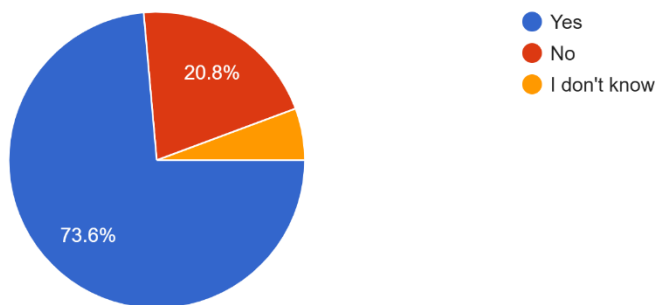
14. Does the large-scale data collection required for AI systems pose serious risks to individual privacy?

14. क्या AI प्रणालियों के लिए आवश्यक बड़े पैमाने पर डेटा संग्रह व्यक्तिगत गोपनीयता के लिए गंभीर जोखिम उत्पन्न करता है?



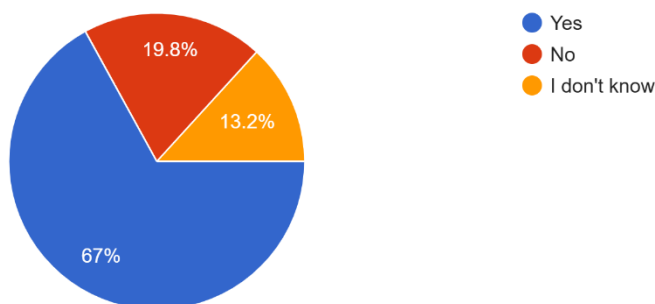
15. Can AI manipulate human behaviour and decision-making?

15. क्या AI मानव व्यवहार और निर्णयों में हेरफेर कर सकता है?



16. Should ethical AI governance prioritise human values over efficiency and economic benefits?

16. क्या नैतिक AI शासन में दक्षता और आर्थिक लाभ की तुलना में मानवीय मूल्यों को प्राथमिकता दी जानी चाहिए?



17. In your opinion, what is the most critical ethical challenge posed by automated AI systems to human behaviour?

(You may select more than one option, if applicable.)

17. आपके अनुसार, स्वचालित AI प्रणालियाँ मानव व्यवहार के लिए सबसे महत्वपूर्ण नैतिक चुनौती क्या प्रस्तुत करती हैं?

(यदि आवश्यक हो तो एक से अधिक विकल्प चुन सकते हैं।)

☐ Cognitive & Mental Impact (Reduced critical thinking, over-reliance, psychological dependency on AI) मानसिक एवं संज्ञानात्मक प्रभाव (आलोचनात्मक सोच में कमी, AI पर अत्यधिक निर्भरता, मानसिक निर्भरता)

☐ Ethical & Moral Judgment (Weakening of moral responsibility, ethical confusion, diffusion of accountability) नैतिक एवं नैतिक निर्णय (नैतिक जिम्मेदारी में कमी, जवाबदेही का भ्रम)

☐ Social & Behavioural Impact (Changes in social interaction, trust, empathy, and human relationships)

सामाजिक एवं व्यवहारिक प्रभाव (सामाजिक संबंधों, विश्वास और सहानुभूति में परिवर्तन)

□ Economic Impact (Job displacement, skill erosion, inequality, economic dependency)

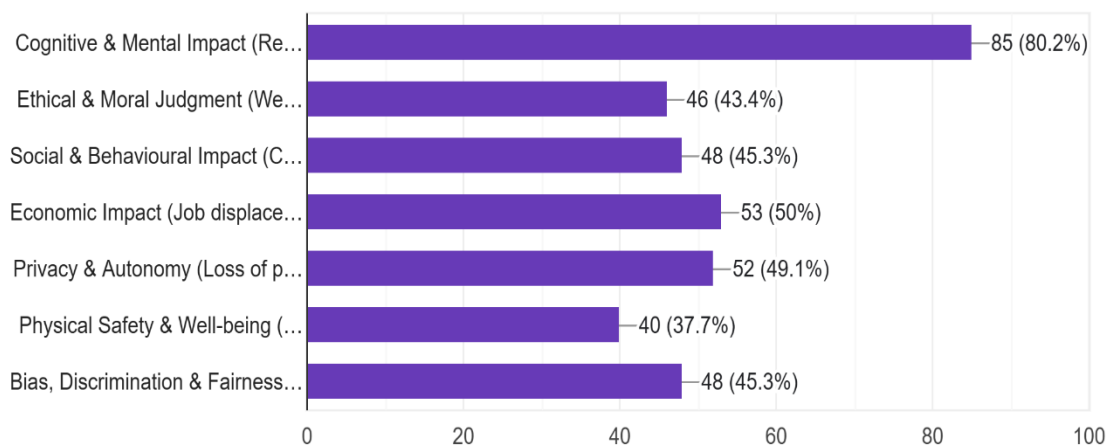
आर्थिक प्रभाव (रोज़गार हानि, कौशल क्षरण, असमानता)

□ Privacy & Autonomy (Loss of personal autonomy, surveillance, misuse of personal data)

गोपनीयता एवं स्वायत्तता (व्यक्तिगत स्वतंत्रता में कमी, निगरानी, डेटा दुरुपयोग)

□ Physical Safety & Well-being (Risks due to AI errors in healthcare, transportation, or critical systems) शारीरिक सुरक्षा एवं कल्याण (स्वास्थ्य, परिवहन या अन्य महत्वपूर्ण प्रणालियों में AI त्रुटियों से जोखिम)

□ Bias, Discrimination & Fairness (Unfair treatment of individuals or groups due to biased algorithms) पक्षपात, भेदभाव एवं निष्पक्षता (पक्षपाती एल्गोरिदम के कारण अनुचित व्यवहार)



This research examined human behaviour and ethical concerns in automated AI systems using a sample of 106 participants. The findings show that AI systems have a strong influence on human decision-making, with 72.6% of respondents agreeing that automated AI significantly affects how people make decisions, while 17.9% disagreed. This aligns with the work of Glickman and Sharot (2024), who demonstrated through large-scale experiments that repeated interactions between humans and AI can create a feedback loop where existing biases in AI systems are gradually amplified. Their research shows that humans often internalize these biases over time, leading to increasingly biased judgments. Together, these findings highlight how AI does not merely assist decision-making but actively shapes human thinking and behaviour.

The study also reveals growing ethical concerns surrounding automation. 62.3% of respondents believe that increasing automation challenges traditional ethical frameworks, as automated decisions may not always align with human moral and legal principles. This concern reflects arguments made by Floridi et al. (2018), who emphasize that issues such as accountability, fairness, and transparency are not adequately addressed by existing ethical frameworks designed for human decision-making. As AI systems become more autonomous, there is a clear need for updated, human-centred ethical guidelines. This concern is further supported by 69.8% of respondents who believe that bias mitigation in AI systems requires continuous human intervention rather than purely technical solutions. Research by Barocas, Hardt, and Narayanan (2019) and Selbst et al. (2019) supports this view, arguing that AI systems are sociotechnical in nature and that fairness cannot be achieved without ongoing human judgment, oversight, and contextual understanding.

Transparency and accountability also emerged as key ethical themes. A strong majority (78.3%) believe that users have a moral right to understand how AI systems make decisions that affect them, reinforcing arguments for Explainable AI (XAI). Prior studies by Doshi-Velez and Kim (2017) and Lima et al. (2022) argue that explainability is essential for ethical and legal accountability. However, the results also show that 64.2% of respondents feel accountability becomes unclear when harm is caused by automated AI systems. While explanations may clarify how decisions are made, they do not always clarify who is responsible—whether developers, organizations, or the system itself. This concern is echoed by Lima et al. (2022), who warn that explainability can sometimes shift blame away from human actors and reduce organizational responsibility.

Responsibility and over-reliance on AI were also significant concerns. 59.6% of respondents believe that developers and organizations should be held ethically responsible for AI outcomes, though prior research warns that explainable systems may unintentionally lead people to treat AI as the responsible agent (Lima et al., 2022). Additionally, 67% of respondents believe that over-reliance on AI reduces human decision-making skills, and 64.2% feel that humans tend to feel less responsible when decisions are made by AI. This is supported by findings on automation bias, where users overly trust AI systems even when they are incorrect. In this study, 63.2% agreed that people often over-trust AI recommendations, a pattern consistent with Parasuraman and Riley (1997) and Dietvorst,

Simmons, and Massey (2015), who show that people frequently defer to automated systems due to perceived objectivity and accuracy.

Human oversight remains a critical solution. 61.3% of respondents agree that Human-in-the-Loop (HITL) systems are essential for ethical AI outcomes, while 21.7% were unsure, likely due to limited understanding of HITL concepts. Prior research by Amershi et al. (2019) and Holzinger (2016) emphasizes that human involvement provides moral judgment, contextual awareness, and ethical oversight that automated systems lack. Similarly, 63.2% of respondents believe that attributing human-like emotions or intentions to AI affects ethical judgment. Research by Nass and Moon (2000), Waytz, Gray, and Epley (2010), and Darling (2016) explains that anthropomorphism can distort ethical reasoning by encouraging misplaced trust, empathy, or moral responsibility toward AI systems.

Privacy and behavioural manipulation are also major ethical concerns. A large majority (74.5%) believe that large-scale data collection by AI systems poses serious privacy risks. While Yang et al. (2024) show that techniques such as differential privacy and synthetic data can reduce these risks, this study highlights the continued need for ethical oversight and human responsibility. Furthermore, 73.6% of respondents believe AI can manipulate human behaviour and decision-making, reflecting strong public awareness of AI's influence. This supports calls for stronger ethical governance frameworks that prioritise human values over efficiency and economic benefits. In fact, 67% of respondents agreed that fairness, transparency, accountability, and respect for human dignity should take priority in AI governance.

Overall, the findings show that automated AI systems raise serious ethical concerns across multiple dimensions of human behaviour. Respondents identified cognitive and mental impacts, reduced critical thinking, weakened moral responsibility, privacy risks, economic instability, physical safety, and bias as major challenges. These results strongly support the need for continuous human oversight, clear accountability structures, and human-centred AI governance to ensure that technological advancement does not come at the cost of ethical responsibility and human well-being.

Conclusion & Future work:

This research shows that AI strongly influences human decision-making and behaviour, often shaping judgments, creating over-trust, and reducing a sense of personal responsibility. The findings show that many people are worried about the ethical problems caused by automated systems, especially bias, privacy issues, accountability, and loss of human skills because of too much reliance on AI. The results show that there is an ongoing conflict between explainability and accountability, because responsibility is often shared or unclear between the AI system, its developers, and the organizations. The study highlights the importance of human-in-the-loop approaches to reduce bias, preserve ethical judgment, and maintain human skills.

Future research should focus on advancing human-centred AI governance, ensuring that human dignity, ethical responsibility, and social well-being are prioritized over efficiency and economic gains. Studies could explore ways to strengthen human-in-the-loop approaches, reduce over-trust in AI, and address accountability challenges, providing practical guidance for ethical AI development and deployment.

REFERENCE:

- Amershi, S., Weld, D. S., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3290605.3300233>
- Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities*. <http://fairmlbook.org>
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Darling, K. (2016). Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior toward robotic objects. In *Robot law* (pp. 213–234). Edward Elgar Publishing.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126. <https://doi.org/10.1037/xge0000033>
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint*. <https://arxiv.org/abs/1702.08608>

- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Glickman, M., & Sharot, T. (2024). Human–AI interaction amplifies biases in decision-making. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-024-01878-3>
- Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, 3(2), 119–131. <https://doi.org/10.1007/s40708-016-0042-6>
- Lima, G., Grgić-Hlača, N., Jeong, J. K., & Cha, M. (2022). The conflict between explainable and accountable decision-making algorithms. *arXiv preprint*. <https://arxiv.org/abs/2205.05306>
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT '19)*, 59–68. ACM. <https://doi.org/10.1145/3287560.3287598>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Waytz, A., Gray, K., & Epley, N. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, 14(8), 383–388. <https://doi.org/10.1016/j.tics.2010.05.006>
- Yang, X., Li, Y., Chen, Z., & Wang, H. (2024). Privacy-preserving data anonymization techniques for AI systems: Balancing data utility and individual privacy. *IEEE Access*, 12, 34567–34579.