Ethical Considerations in Artificial Intelligence: A Comprehensive Overview of Contemporary Challenges and Solutions

1 author:

Velibor Božić

633 PUBLICATIONS   **210** CITATIONS

SEE PROFILE

Ethical Considerations in Artificial Intelligence: A Comprehensive Overview of Contemporary Challenges and Solutions

Velibor Božić
General hospital Koprivnica, Croatia.
veliborbozic@gmail.com

Abstract - This paper provides an overview of the current landscape of ethical considerations in artificial intelligence (AI). The rapid integration of AI technologies across diverse sectors has prompted intense scrutiny of their societal impacts. This review explores key ethical challenges, focusing on issues such as bias in facial recognition, predictive policing, healthcare diagnostics, hiring tools, deepfake technology, autonomous vehicles, AI-powered surveillance, mental health support chatbots, and automated content moderation on social media. Each example elucidates complex ethical dilemmas, ranging from privacy concerns and discriminatory outcomes to the responsible deployment of AI in critical decision-making scenarios. The paper emphasizes the need for transparent, fair, and accountable AI systems. It also underscores the importance of ongoing interdisciplinary research, collaborative stakeholder engagement, and the development of robust regulatory frameworks to navigate the evolving ethical landscape of AI.

Keywords: *Accountability, Bias Mitigation, Ethical AI, Transparency*

I INTRODUCTION

In recent years, the rapid advancement and integration of artificial intelligence (AI) technologies have ushered in transformative changes across various industries. While AI promises unprecedented capabilities and efficiencies, its pervasive influence has prompted a critical examination of the ethical considerations inherent in its application. This introduction seeks to delineate the overarching goal of this study, the key research propositions guiding our inquiry, and the methodology employed to comprehensively explore the multifaceted landscape of ethical considerations in AI.

GOAL

The overarching goal of the study is to provide a comprehensive analysis of the ethical dimensions surrounding AI deployment. Our aim is to deepen the understanding of the challenges posed by AI technologies, emphasizing the exploration of potential biases, transparency issues, accountability mechanisms, and the overarching goal of fostering responsible and ethical AI practices.

RESEARCH PROPOSAL

1. *Bias Identification and Mitigation:* This study posits that identifying and mitigating biases within AI algorithms is imperative for ensuring fair and equitable outcomes. We aim to investigate how biases emerge, their impact on decision-making processes, and the efficacy of proposed mitigation strategies.

2. *Transparency as a Pillar of Ethical AI:* We hypothesize that transparency is a fundamental prerequisite for ethical AI. Our research will explore the importance of explainability in AI models, investigating methods to enhance transparency and promote a deeper understanding of how AI systems arrive at specific decisions.

3. study asserts that clear lines of accountability are essential to address the ethical implications of AI. We will examine existing frameworks and propose mechanisms to establish accountability in AI development, deployment, and the consequences of AI-driven decisions.

4. *Interdisciplinary Exploration:* Recognizing the interdisciplinary nature of AI ethics, our research proposition involves engaging with perspectives from fields such as philosophy, law, computer science,economy and sociology. We seek mixed-methods approach. Quantitative analysis will be employed to assess the prevalence and impact of biases in AI algorithms, while qualitative methods, including case

studies and expert interviews, will be utilized to gain deeper insights into the ethical considerations. Additionally, a comprehensive literature review will inform our understanding of the current state of ethical AI research.

5. Through this methodological approach, we aim to provide a nuanced and holistic examination of the ethical landscape in AI, offering valuable insights for researchers, practitioners, and policymakers grappling with the responsible integration of AI technologies into our rapidly evolving society.

6. to understand how diverse perspectives can contribute to a holistic framework for ethical AI.

METHODS

The methodology used for this study is a systematic literature review (SLR), as an essential tool for summarizing evidence accurately and reliably.

II RESEARCH

In this part, we will talk about the literature dealing with ethical AI systems, we will explain ethical AI systems [1], [2], [3] how they are created, we will list examples of ethical AI systems, as well as dilemmas related to ethical AI systems for different purposes.

A GENERAL OVERVIEW OF THE KEY THEMES AND TRENDS IN THE LITERATURE

A general overview of the key themes and trends in the literature related to ethics in AI up to that point. Keep in mind that the field is dynamic, and new research may have emerged since then.

- Bias and Fairness:

A significant portion of literature focuses on addressing biases in AI algorithms. Researchers explore methods for detecting and mitigating bias, as well as promoting fairness in AI systems, especially in contexts like hiring, lending, and criminal justice [27].

- Explainability and Transparency:

Researchers discuss techniques for making AI systems more interpretable, understandable, and transparent. This area is crucial for building trust and accountability, particularly in high-stakes applications like healthcare and finance [28].

- Privacy and Security:

With the increasing use of AI in handling sensitive data, literature emphasizes the importance of privacy and security. Topics include privacy-preserving AI techniques, secure data handling, and the ethical implications of AI in surveillance and data-driven decision-making [29].

- Human-AI Collaboration:

There is a growing interest in the collaboration between humans and AI systems. The literature explores how to design AI technologies that enhance human capabilities, incorporate human values, and ensure that humans remain in control of critical decisions [30].

- Ethical AI Governance and Regulation:

Researchers delve into the development of ethical guidelines, standards, and regulatory frameworks for AI. The literature discusses the role of governments, industry bodies, and organizations in establishing policies that ensure responsible AI development and deployment [31].

- AI in Healthcare Ethics:

A specific focus is on the ethical considerations in applying AI to healthcare. Topics include patient privacy, informed consent, the responsible use of medical data, and the integration of AI into clinical decision-making [32].

- Social and Economic Impacts:

Scholars examine the broader societal and economic impacts of AI, including issues related to job displacement, inequality, and the digital divide. Ethical considerations in AI deployment within various social contexts are explored [33].

- Examination of Ethical Dilemmas:

Literature often presents case studies and ethical dilemmas arising from AI applications. This helps to highlight the complexities and challenges associated with ethical decision-making in different domains [34].

- Experiences from Industry:

Some literature provides insights into the practical challenges faced by companies in implementing ethical AI. Case studies and real-world experiences offer valuable lessons for organizations navigating the ethical landscape [35].

- Interdisciplinary Perspectives:

AI ethics is inherently interdisciplinary. Literature often incorporates insights from philosophy, law, sociology, and other fields. This multidisciplinary approach enriches the discussion on ethical considerations in AI [36].

It's important to note that the field of AI ethics is rapidly evolving, and ongoing research continues to shape our understanding of ethical challenges and solutions. For the most up-to-date information, I recommend exploring recent academic journals, conferences, and publications in the field of AI ethics.

WHAT IS ETHICAL AI SYSTEMS

Ethical AI systems [3], [4] are artificial intelligence (AI) systems that are designed, developed, and deployed in a way that adheres to ethical principles and values. These principles and values can include fairness, transparency, accountability, non-discrimination, and respect for human rights.

Key characteristics of ethical AI systems include [5]:

- Fairness: Ethical AI systems should not discriminate against or unfairly disadvantage any particular group of people. This means that they should be trained on data that is representative of the population that they will be used to serve, and they should be designed to avoid bias in their decision-making.
- Transparency: Ethical AI systems should be transparent, meaning that it should be possible to understand how they make decisions. This is important for building trust in AI systems and for ensuring that they are accountable for their decisions.
- Accountability: Ethical AI systems should be accountable, meaning that there should be a clear process for identifying and addressing any problems that arise. This includes having a process for reporting and investigating bias, as well as a process for correcting any mistakes that are made.
- Privacy: Ethical AI systems should respect privacy, meaning that they should not collect or use personal data without the user's consent. This includes taking steps to protect data from unauthorized access and use.
- Safety: Ethical AI systems should be safe, meaning that they should not be used in a way that could cause harm to people or property. This includes taking steps to prevent AI systems from being used for malicious purposes.

BENEFITS OF ETHICAL AI SYSTEMS

Benefits of ethical AI systems are [6],[7]:

- Increased fairness and non-discrimination: Ethical AI systems can help to reduce bias and discrimination in decision-making.
- Improved transparency and accountability: Ethical AI systems can help to build trust in AI systems and ensure that they are accountable for their decisions.
- Reduced risk of harm: Ethical AI systems can help to reduce the risk of harm from AI systems, such as discrimination, privacy violations, and safety hazards.
- Enhanced human well-being: Ethical AI systems can be used to improve human well-being in a variety of ways, such as by providing better healthcare, education, and financial services.

CHALLENGES OF DEVELOPING AND DEPLOYING ETHICAL AI SYSTEMS

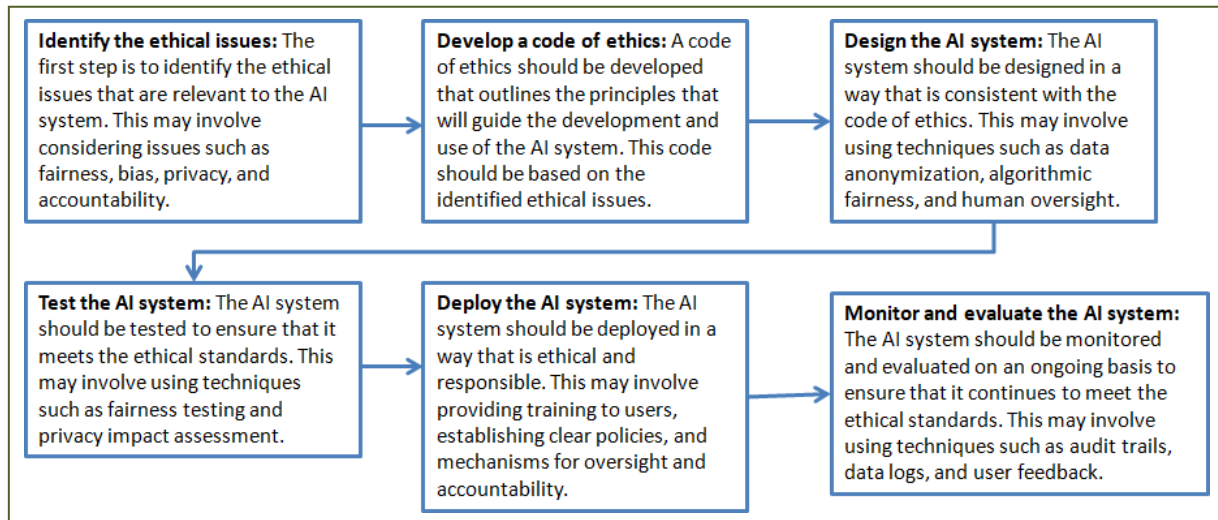Challenges of developing and deploying ethical AI systems are [6], [7]:

- Defining and operationalizing ethical principles: It can be difficult to define and operationalize ethical principles in a way that is both clear and practical.
- Identifying and addressing bias: Bias can be difficult to identify and address, and it can be embedded in data, algorithms, and human decision-making.

- Ensuring transparency and accountability: It can be difficult to ensure that AI systems are transparent and accountable, especially when they are complex and opaque.
- Balancing ethical principles with other goals: Ethical principles may conflict with other goals, such as efficiency, accuracy, and cost.

Despite the challenges, there is a growing consensus that it is important to develop and deploy ethical AI systems. Ethical AI systems can help to ensure that AI is used for good and that it benefits all of society.

PROCESS OF DEVELOPING ETHICAL AI SYSTEMS

Developing an ethical [8] AI system is a complex and challenging task, but it is essential to ensure that AI is used in a responsible and beneficial way. Here are some strategies for coping with the challenges of developing an ethical AI system:

Identify the ethical issues: The first step is to identify the ethical issues that are relevant to the AI system. This may involve considering issues such as fairness, bias, privacy, and accountability.

Develop a code of ethics: A code of ethics should be developed that outlines the principles that will guide the development and use of the AI system. This code should be based on the identified ethical issues.

Design the AI system: The AI system should be designed in a way that is consistent with the code of ethics. This may involve using techniques such as data anonymization, algorithmic fairness, and human oversight.

Test the AI system: The AI system should be tested to ensure that it meets the ethical standards. This may involve using techniques such as fairness testing and privacy impact assessment.

Deploy the AI system: The AI system should be deployed in a way that is ethical and responsible. This may involve providing training to users, establishing clear policies, and mechanisms for oversight and accountability.

Monitor and evaluate the AI system: The AI system should be monitored and evaluated on an ongoing basis to ensure that it continues to meet the ethical standards. This may involve using techniques such as audit trails, data logs, and user feedback.

Picture 1. *Process of developing ethical AI systems*
Source: author's work using [8], [9]

Developing an ethical AI system is a complex task, but it is essential to ensure that AI is used for good and that it benefits all of society. By following these strategies, you can help to develop and deploy AI systems that are fair, transparent, accountable, and safe.

ROLES OF ETHIC IN HOSPITAL AI SYSTEMS

Ethics plays a crucial role in the application of AI-powered business decision-making, especially in sensitive environments such as general hospitals. Here are several key aspects to consider [11], [12], [13]:

1. Patient Privacy and Data Security:
   Ethical AI in hospitals involves safeguarding patient data and ensuring strict adherence to privacy regulations (such as HIPAA in the United States). AI algorithms must be designed to prioritize patient confidentiality and protect sensitive health information.

2. Transparency and Explainability:
   Transparency in AI algorithms is vital. Hospital administrators and healthcare professionals should understand how AI systems make decisions. Transparent AI models help build trust and facilitate better decision-making by providing insights into the rationale behind specific recommendations or actions.

3. Bias and Fairness:
   AI systems can inadvertently inherit biases present in historical data. In a hospital setting, biased AI algorithms could lead to unequal treatment of patients. It's essential to continuously monitor and mitigate biases, ensuring that AI-driven decisions are fair and do not disproportionately impact certain demographic groups.

4. Informed Consent:

Patients should be informed about the use of AI in their healthcare and should have the right to consent or opt-out of AI-powered decision-making processes. Clear communication about the benefits, risks, and limitations of AI technologies is crucial for maintaining ethical standards.

5. Human Oversight and Intervention:

AI systems should complement human decision-making rather than replace it entirely. There should be mechanisms for human oversight and intervention in critical situations. Healthcare professionals must have the ability to challenge or override AI recommendations based on their expertise and ethical considerations.

6. Accountability and Responsibility:

Establishing clear lines of accountability is essential. Hospitals and their staff should be accountable for the decisions made by AI systems. This includes taking responsibility for any adverse outcomes and continually improving AI models to enhance performance and ethical standards.

7. Equitable Access to Healthcare:

Ethical considerations extend beyond the hospital walls. AI applications should aim to improve access to healthcare services, particularly in underserved communities. It is important to avoid reinforcing existing disparities and work towards creating a more equitable healthcare system.

8. Continuous Monitoring and Evaluation:

Regular audits and evaluations of AI systems should be conducted to identify and rectify any ethical issues that may arise over time. This ensures that the technology aligns with evolving ethical standards and remains in compliance with regulations.

9. Professional Development and Education:

Hospitals should invest in training healthcare professionals on the ethical implications of AI technologies. This includes educating staff about the capabilities, limitations, and ethical considerations associated with AI-powered decision-making.

10. Stakeholder Involvement:

Involve various stakeholders, including patients, healthcare professionals, administrators, and ethicists, in the development, deployment, and evaluation of AI systems. A collaborative approach ensures that diverse perspectives are considered in the decision-making process.

In summary, embedding ethics in AI-powered business decision-making within a general hospital setting is crucial for ensuring patient trust, fairness, and the responsible deployment of technology to enhance healthcare outcomes. It requires a multidimensional approach that addresses technical, social, and regulatory aspects.

WHY ETHICS PLAYS A CRUCIAL ROLE IN THE APPLICATION OF AI-POWERED BUSINESS DECISION-MAKING, ESPECIALLY IN SENSITIVE ENVIRONMENTS SUCH AS GENERAL HOSPITALS

Ethics plays a crucial role in the application of AI-powered business decision-making in sensitive environments like general hospitals for several reasons:

Reason	Explanation
Patient Trust	Trust is fundamental in healthcare. Patients trust hospitals and healthcare professionals with their sensitive and personal information. Ethical AI practices reassure patients that their data is handled responsibly and that AI-powered decisions prioritize their well-being.
Patient Safety	Decisions made in a hospital setting, whether by humans or AI, directly impact patient safety. Ethical AI ensures that algorithms are designed and trained to prioritize patient safety, minimizing

	the risk of errors, and adhering to the highest standards of care.
Data Privacy and Security	Hospitals deal with highly sensitive patient data, and ethical considerations are crucial to protect this information. AI applications must comply with data privacy regulations to ensure that patient confidentiality is maintained, preventing unauthorized access or misuse of personal health information.
Fair Treatment and Avoidance of Bias	Ethical AI helps prevent bias in decision-making processes. In healthcare, biased algorithms could lead to unequal treatment, misdiagnoses, or other adverse outcomes. Ensuring fairness in AI systems is essential to uphold the principle of equal and just healthcare for all patients.
Legal and Regulatory Compliance	Hospitals are subject to various regulations and legal frameworks governing healthcare practices. Ethical AI practices help ensure compliance with these regulations, reducing the risk of legal consequences and protecting the hospital's reputation.
Public Perception and Reputation	Hospitals are often seen as pillars of trust in the community. Ethical lapses in AI practices could damage the reputation of the hospital and erode public trust. Conversely, ethical AI contributes to a positive perception, showcasing the hospital's commitment to responsible and compassionate care.
Mitigation of Unintended Consequences	AI systems, if not ethically designed and implemented, can have unintended consequences. Ethical considerations help identify and mitigate potential risks, ensuring that AI systems contribute positively to patient outcomes without introducing new challenges or harm.
Human-AI Collaboration	Ethical AI recognizes the importance of maintaining a balance between human expertise and AI capabilities. It ensures that AI systems complement human decision-making rather than replacing it entirely. This collaborative approach is critical in healthcare, where the human touch and judgment are often irreplaceable.
Avoidance of Discrimination	Ethical AI practices seek to eliminate discrimination in healthcare. Discriminatory algorithms could lead to disparities in treatment based on factors like race, gender, or socio-economic status. Ethical considerations help address these issues and promote healthcare equity.
Continuous Improvement and Learning	Ethical frameworks promote continuous improvement and learning. Hospitals committed to ethical AI regularly assess and update their systems, incorporating feedback from healthcare professionals and patients. This iterative process ensures that AI systems evolve responsibly over time.

Table 1. *Reasons why ethics plays a crucial role in the application of AI-powered business decision-making Source: author's work using [14], [15], [24]*

Ethics in AI-powered business decision-making within sensitive environments like general hospitals is essential for maintaining trust, ensuring patient safety, complying with regulations, and upholding the values of

fairness and equity in healthcare delivery. It aligns AI practices with the ethical standards expected in the medical profession and the broader community.

ENSURING ETHICAL AI IN TECHNICAL SENSE

Ensuring ethical AI in a technical sense involves implementing practices and principles during the design, development, deployment, and ongoing management of AI systems. Here are key technical considerations for ensuring ethical AI [16], [17], [18]:

1. Data Quality and Bias Mitigation
   - Data Collection and Preprocessing: Ensure that training data is representative and diverse. Identify and correct biases in training data to prevent the reinforcement of unfair or discriminatory patterns.
   - Bias Detection and Mitigation: Use techniques to detect and mitigate bias in algorithms. This may involve adjusting data, modifying algorithms, or implementing fairness-aware machine learning methods.

2. Transparency and Explainability
   - Interpretable Models: Choose models that are interpretable and can provide explanations for their decisions. This helps stakeholders, including end-users and regulators, understand how the AI system arrives at specific conclusions.
   - Explainability Techniques: Implement techniques such as feature importance analysis, SHAP (SHapley Additive exPlanations), or LIME (Local Interpretable Model-agnostic Explanations) to enhance transparency.

3. User Consent and Control
   - Informed Consent: Ensure users are informed about the use of AI, its implications, and how their data will be utilized. Obtain explicit consent for sensitive applications.
   - User Control: Provide users with the ability to control their data and preferences, including the option to opt-out of certain AI-driven features.

4. Security and Privacy
   - Secure Data Handling: Implement robust security measures to protect AI systems and the data they process. Encrypt sensitive information, use secure communication protocols, and establish access controls.
   - Privacy-Preserving Techniques: Explore privacy-preserving techniques such as federated learning, homomorphic encryption, or differential privacy to protect individual data.

5. Human Oversight and Intervention
   - Human-in-the-Loop Systems: Design AI systems that involve human oversight and intervention. Ensure that human experts can review and override AI decisions when necessary, especially in critical situations.

6. Continuous Monitoring and Auditing
   - Monitoring Performance: Regularly monitor the performance of AI systems in real-world scenarios. Implement mechanisms to detect and address performance degradation or shifts in data distributions.
   - Auditing and Accountability: Conduct regular audits to assess the fairness, transparency, and overall ethical performance of AI systems. Establish accountability mechanisms for addressing issues that may arise.

7. Ethical by Design
   - Ethical Guidelines and Standards: Develop and adhere to ethical guidelines and standards specific to AI development and deployment. These guidelines should reflect the values and principles relevant to the application domain.
   - Ethics Review Boards: Establish ethics review boards or committees to evaluate the ethical implications of AI projects before and during deployment.

8. Regulatory Compliance
   - Stay Informed About Regulations: Stay informed about and comply with relevant data protection and privacy regulations (e.g., GDPR, HIPAA) and ethical frameworks specific to the industry.

9. Stakeholder Involvement

- Engage Diverse Stakeholders: Involve diverse stakeholders, including end-users, domain experts, ethicists, and community representatives, in the AI development process to ensure a broad perspective on ethical considerations.
10. Education and Training
- Internal Training Programs: Train AI developers, data scientists, and other stakeholders on ethical considerations in AI development. Foster a culture of ethical awareness within the organization.

By incorporating these technical measures into the AI development lifecycle, organizations can build and deploy AI systems that align with ethical principles, address societal concerns, and promote responsible and trustworthy use of AI technologies.

EXAMPLES OF ETHICAL AI SYSTEMS

There are many examples of ethical AI systems in use today. Here are a few [16], [17], [18]:

- AI-powered hiring tools that can help to identify and recruit qualified candidates from a diverse pool of applicants. These tools can help to reduce unconscious bias in the hiring process and ensure that all candidates are given a fair chance to be considered for open positions.
- AI-powered chatbots that can provide customer support 24/7. These chatbots can be trained to answer a wide range of questions and provide assistance in multiple languages. They can also be used to collect feedback from customers and improve the overall customer experience.
- AI-powered fraud detection systems that can help to identify and prevent fraudulent activity. These systems can be used in a variety of industries, including finance, insurance, and retail. They can help to reduce losses and protect customer data.
- AI-powered medical diagnosis tools that can help doctors to diagnose diseases more accurately. These tools can be used to analyze medical images, such as X-rays and MRI scans, and provide doctors with a second opinion. They can also be used to develop personalized treatment plans for patients.
- AI-powered environmental monitoring systems that can help to track changes in the environment and identify potential threats. These systems can be used to monitor air and water quality, deforestation, and climate change. They can help to protect the environment and human health.

These are just a few examples of the many ethical AI systems that are being used today. AI has the potential to make a positive impact on society in a number of ways, and it is important to develop and use AI systems in an ethical and responsible manner.

DILEMMAS IN DEVELOPING DIFFERENT ETHICAL AI SYSTEMS

Dilemmas in developing different ethical AI systems are: [20],[25],[26]

DILEMMAS IN FACIAL RECOGNITION TECHNOLOGIES

The use of facial recognition technology presents significant ethical dilemmas, particularly concerning biases inherent in these systems [39]. Key ethical dilemmas associated with bias in facial recognition are: Accuracy Disparities, Racial and Gender Bias, Impact on Vulnerable Communities, Privacy and Consent, Surveillance State Concerns, Lack of Regulation and Standards, Bias Amplification, Informed Consent and Public Awareness, Legal and Human Rights Implications and Algorithmic Accountability. Addressing these ethical dilemmas requires a multifaceted approach involving technological advancements, regulatory frameworks, public discourse, and a commitment to fairness and equity in the development and deployment of facial recognition technology.

ETHICAL DILEMMAS IN PREDICTIVE POLICING

Predictive policing, which involves using algorithms and data analysis to forecast and prevent potential criminal activity, presents several ethical dilemmas [40]. Here are key ethical challenges associated with predictive policing: Bias in Predictive Models, Risk of Discriminatory Policing, Transparency and Lack of Explainability, Privacy Concerns, Self-Fulfilling Prophecies, Ineffective Crime Prevention, Lack of Community Input, Accountability and Responsibility, Unintended Consequences andAlgorithmic Fairness and Equity. Addressing these ethical dilemmas in predictive policing requires careful consideration of the societal impacts, continuous monitoring and evaluation of algorithmic outcomes, transparency in algorithmic decision-making, and active

engagement with the communities affected by these technologies. Balancing crime prevention with the protection of individual rights remains a complex and evolving ethical challenge in the realm of predictive policing.

ETHICAL DILEMMAS IN HEALTHCARE DIAGNOSTICS

Healthcare diagnostics powered by artificial intelligence (AI) introduces several ethical dilemmas that require careful consideration [41]. Here are key ethical challenges associated with AI in healthcare diagnostics: Accuracy and Reliability, Bias in Training Dana, Interpretability and Explainability, Informed Consent and Patient Autonomy, Impact on Healthcare Professionals, Data Privacy and Security, Equity in Access to AI Diagnostics, Over-Reliance on Technology, Liability and Accountability and Long-Term Impact on Public Health. Addressing these ethical dilemmas in healthcare diagnostics requires a comprehensive approach that involves collaboration among healthcare professionals, technology developers, policymakers, and patients. Developing ethical guidelines, ensuring transparency, and prioritizing patient well-being are essential components of navigating the ethical landscape of AI in healthcare diagnostics.

ETHICAL DILEMMAS IN HIRING TOOLS

The use of artificial intelligence (AI) and automated hiring tools in the recruitment process introduces several ethical dilemmas [42]. Here are key ethical challenges associated with the deployment of hiring tools: Bias in Hiring Algorithms, Lack of Transparency, Fairness and Equity, Impact on Diversity and Inclusion, Data Privacy and Security, Reliability and Validity of Assessments, Exclusion of Human Judgment and Accessibility and Digital Divide. Addressing these ethical dilemmas in hiring tools requires ongoing scrutiny, transparency, and a commitment to fairness. Employers, developers, and regulatory bodies must collaborate to establish ethical guidelines that prioritize diversity, equity, and accountability in the recruitment process.

ETHICAL DILEMMAS IN DEEPFAKE TECHNOLOGY

Deepfake technology, which utilizes artificial intelligence to create highly realistic fake videos or audio recordings, introduces profound ethical dilemmas [43]. Here are key ethical challenges associated with deepfake technology: Misinformation and Disinformation, Manipulation and Deception, Privacy Invasion, Political Implications, Impersonation and Identity Theft, Erosion of Trust, Legal and Regulatory Gaps, Consent and Authenticity, Emotional and Psychological Impact andCombating Deepfake Technology. Addressing these ethical dilemmas requires a multidimensional approach involving technology development, legal frameworks, public awareness, and ethical guidelines. Striking a balance between technological innovation and safeguarding against misuse is essential to ensure responsible development and deployment of deepfake technology.

ETHICAL DILEMMAS IN AUTONOMOUS VEHICLES

The deployment of autonomous vehicles introduces a range of ethical dilemmas that center around the intersection of technology, safety, and decision-making. Here are key ethical challenges associated with autonomous vehicles [44]: Decision-Making in Critical Situations, Trolley Problem and Moral Decision-Making, Responsibility and Liability, Data Privacy and Security, Impact on Employment, Algorithmic Bias and Fairness, Interactions with Human-Driven Vehicles, Accessibility and Equity, Public Trust and Acceptance and Regulatory Frameworks and Standardization. Addressing these ethical dilemmas in autonomous vehicles requires collaboration among technologists, policymakers, ethicists, and the broader public. A holistic approach that considers the societal, economic, and safety implications is essential for navigating the ethical landscape of autonomous transportation.

ETHICAL DILEMMAS IN AI-POWERED SURVEILLANCE

The deployment of AI-powered surveillance systems raises numerous ethical dilemmas [45], as it involves the intersection of technology, privacy, and societal values. Here are key ethical challenges associated with AI-powered surveillance: Privacy Invasion, Mass Surveillance and Civil Liberties, Bias and Discrimination, Mission Creep and Function Creep, Lack of Informed Consent, Chilling Effects on Freedom, Security of Surveillance Dana, Automated Decision-Making and Due Process, Community Trust and Policing,Long-term Social Implications. Addressing these ethical dilemmas in AI-powered surveillance requires a careful and thoughtful approach that considers legal frameworks, public opinion, technological limitations, and the potential

for misuse. Ethical guidelines, transparency, and ongoing dialogue between stakeholders are essential to navigate the complex landscape of surveillance technologies responsibly.

## ETHICAL DILEMMAS IN MENTAL HEALTH SUPPORT CHATBOTS

The use of mental health support chatbots introduces unique ethical dilemmas [46] that revolve around the intersection of technology, healthcare, and sensitive personal information. Here are key ethical challenges associated with mental health support chatbots: Accuracy and Reliability of Advice, Informed Consent and User Awareness, Privacy and Data Security, Human-Like Interaction and Emotional Connection, Handling Crisis Situations, Cultural Sensitivity and Diversity, Long-Term Efficacy and User Outcomes, User Autonomy and Empowerment, User Profiling and Targeted Advertising andRegulatory Compliance and Standards. Addressing these ethical dilemmas in mental health support chatbots requires a collaborative effort involving mental health professionals, technologists, regulators, and users. Striking a balance between leveraging technology for mental health support and safeguarding the well-being and rights of users is paramount in the development and deployment of such tools.

ETHICAL DILEMMA IN AUTOMATED CONTENT MODERATION ON SOCIAL MEDIA

Automated content moderation on social media platforms raises several ethical dilemmas [47], as it involves the use of algorithms to monitor, filter, and sometimes remove user-generated content. Here are key ethical challenges associated with automated content moderation: Freedom of Expression vs. Content Regulation, Bias and Discrimination, Transparency and Accountability, Cultural Sensitivity and Context, Overreliance on Automation, Impact on Small Content Creators, False Positives and Content Misclassification,Censorship and Gatekeeping, Erosion of User Trust and Long-Term Impact on Discourse. Addressing these ethical dilemmas in automated content moderation requires ongoing collaboration between technologists, content moderators, platform administrators, and users. Establishing transparent guidelines, incorporating user feedback, and combining automated tools with human oversight are crucial steps in developing responsible content moderation practices.

RESOLVING ETHICAL DILEMMAS ON WRITE WAY

Resolving ethical dilemmas requires a thoughtful and systematic approach. Here are steps you can take to address ethical challenges in a responsible and effective manner [22], [23]:
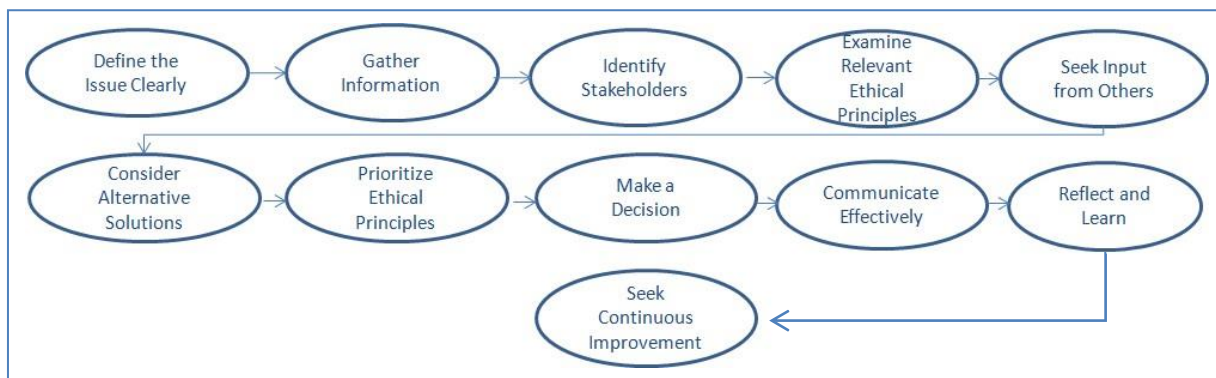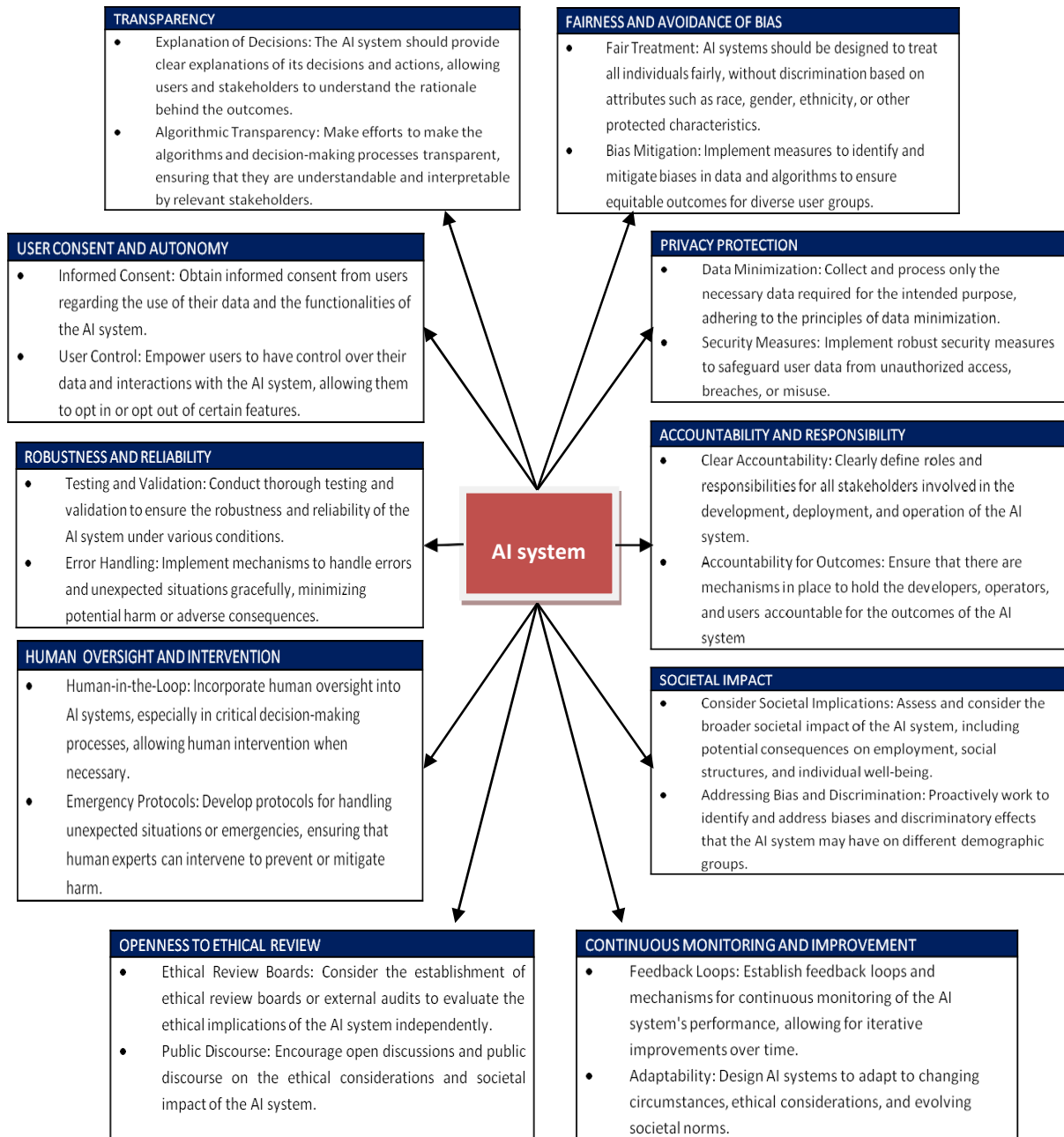

Picture 2. *Process of resolving ethical dilemmas*
Source: author's work using [22], [23]

Remember that ethical dilemmas often involve complex and nuanced considerations, and there may not be a perfect solution. The goal is to make well-informed, ethically sound decisions that align with your values and principles while minimizing harm to stakeholders.

WHAT SHOULD AN AI SYSTEM BE IN ORDER TO SATISFY ETHICAL PRINCIPLES

To satisfy ethical principles, an AI system should be designed, implemented, and operated with careful consideration of various ethical dimensions. Here are key aspects that an AI system should embody to align with ethical principles [21], [27]:

TRANSPARENCY
- Explanation of Decisions: The AI system should provide clear explanations of its decisions and actions, allowing users and stakeholders to understand the rationale behind the outcomes.
- Algorithmic Transparency: Make efforts to make the algorithms and decision-making processes transparent, ensuring that they are understandable and interpretable by relevant stakeholders.

FAIRNESS AND AVOIDANCE OF BIAS
- Fair Treatment: AI systems should be designed to treat all individuals fairly, without discrimination based on attributes such as race, gender, ethnicity, or other protected characteristics.
- Bias Mitigation: Implement measures to identify and mitigate biases in data and algorithms to ensure equitable outcomes for diverse user groups.

USER CONSENT AND AUTONOMY
- Informed Consent: Obtain informed consent from users regarding the use of their data and the functionalities of the AI system.
- User Control: Empower users to have control over their data and interactions with the AI system, allowing them to opt in or opt out of certain features.

PRIVACY PROTECTION
- Data Minimization: Collect and process only the necessary data required for the intended purpose, adhering to the principles of data minimization.
- Security Measures: Implement robust security measures to safeguard user data from unauthorized access, breaches, or misuse.

ROBUSTNESS AND RELIABILITY
- Testing and Validation: Conduct thorough testing and validation to ensure the robustness and reliability of the AI system under various conditions.
- Error Handling: Implement mechanisms to handle errors and unexpected situations gracefully, minimizing potential harm or adverse consequences.

ACCOUNTABILITY AND RESPONSIBILITY
- Clear Accountability: Clearly define roles and responsibilities for all stakeholders involved in the development, deployment, and operation of the AI system.
- Accountability for Outcomes: Ensure that there are mechanisms in place to hold the developers, operators, and users accountable for the outcomes of the AI system

HUMAN OVERSIGHT AND INTERVENTION
- Human-in-the-Loop: Incorporate human oversight into AI systems, especially in critical decision-making processes, allowing human intervention when necessary.
- Emergency Protocols: Develop protocols for handling unexpected situations or emergencies, ensuring that human experts can intervene to prevent or mitigate harm.

SOCIETAL IMPACT
- Consider Societal Implications: Assess and consider the broader societal impact of the AI system, including potential consequences on employment, social structures, and individual well-being.
- Addressing Bias and Discrimination: Proactively work to identify and address biases and discriminatory effects that the AI system may have on different demographic groups.

AI system

OPENNESS TO ETHICAL REVIEW
- Ethical Review Boards: Consider the establishment of ethical review boards or external audits to evaluate the ethical implications of the AI system independently.
- Public Discourse: Encourage open discussions and public discourse on the ethical considerations and societal impact of the AI system.

CONTINUOUS MONITORING AND IMPROVEMENT
- Feedback Loops: Establish feedback loops and mechanisms for continuous monitoring of the AI system's performance, allowing for iterative improvements over time.
- Adaptability: Design AI systems to adapt to changing circumstances, ethical considerations, and evolving societal norms.

Picture 2. K*ey aspects that an AI system should embody to align with ethical principles*
Source: author's work using [21], [27]

By embodying these principles, AI systems can contribute to ethical and responsible technology development, deployment, and use. Ethical considerations should be an integral part of the entire AI lifecycle, from design and development to operation and continuous improvement.

USE OF AI IN INCREASE OF BUSINESS ECONOMICS IN THE CONTEXT OF OBSERVING ETHICAL PRINCIPLES

The use of Artificial Intelligence (AI) in business can significantly contribute to the increase of business economics while observing ethical principles. Here are some ways in which AI can be applied to enhance business economics ethically [37],[38]:

1. Data-driven Decision Making
   - Ethical Consideration: Ensure that AI systems use unbiased and diverse datasets to prevent discriminatory outcomes.
   - Impact on Business Economics: Improved decision-making based on accurate and real-time data can lead to cost reductions, increased efficiency, and better resource allocation.

2. Customer Experience Enhancement
   - Ethical Consideration: Respect customer privacy and ensure transparent communication about the use of AI in personalization.
   - Impact on Business Economics: Enhanced customer experiences lead to increased customer loyalty, repeat business, and positive brand perception, ultimately contributing to higher revenue.

3. Supply Chain Optimization
   - Ethical Consideration: Ensure fair labor practices, environmental sustainability, and transparency in the supply chain.
   - Impact on Business Economics: AI can optimize supply chain operations, reduce costs, minimize waste, and improve overall efficiency.

4. Fraud Detection and Prevention
   - Ethical Consideration: Protect user data and privacy while detecting and preventing fraudulent activities.
   - Impact on Business Economics: Reduces financial losses due to fraud, enhances trust in the business, and contributes to a secure economic environment.

5. Workforce Productivity
   - Ethical Consideration: Address concerns related to job displacement and ensure a fair transition for employees affected by automation.
   - Impact on Business Economics: Improved efficiency and productivity through AI-powered automation can lead to cost savings and increased competitiveness.

6. Ethical AI Governance
   - Ethical Consideration: Establish clear guidelines and governance frameworks to ensure responsible and ethical use of AI technologies.
   - Impact on Business Economics: Mitigates risks associated with misuse of AI, protects the company's reputation, and fosters trust among stakeholders.

7. Energy Efficiency
   - Ethical Consideration: Promote the use of AI for sustainability, ensuring energy-efficient algorithms and practices.
   - Impact on Business Economics: Reducing energy consumption through AI applications can result in cost savings and align the business with ethical and environmental considerations.

8. Personalized Marketing
   - Ethical Consideration: Respect user privacy, provide transparency, and avoid intrusive practices in personalized marketing efforts.
   - Impact on Business Economics: Targeted and personalized marketing campaigns based on AI analysis can lead to higher conversion rates and improved return on investment.

By integrating AI into business operations while adhering to ethical principles, companies can unlock new opportunities for growth, innovation, and efficiency while maintaining trust among customers and stakeholders. Ethical considerations should be an integral part of the AI strategy to ensure responsible and sustainable business practices.

II RESULTS AND DISCUSSION

RESULTS

1. Improved Fairness and Equity:

Ethical AI systems can contribute to fair decision-making by reducing biases and ensuring equitable treatment across diverse demographic groups. This is particularly crucial in contexts such as hiring, lending, and law enforcement.

2. Enhanced Transparency:

Ethical AI systems often prioritize transparency, providing clear explanations for their decisions. This transparency can enhance user understanding, build trust, and facilitate accountability.

3. Minimized Discrimination:

By actively addressing biases in data and algorithms, ethical AI systems aim to minimize discriminatory outcomes, promoting equal opportunities and treatment for all individuals.

4. Human-AI Collaboration:

Ethical AI systems can facilitate effective collaboration between humans and machines, combining the strengths of AI-driven insights with human values, intuition, and ethical judgment.

5. Adherence to Ethical Guidelines:

Organizations using ethical AI systems demonstrate a commitment to adhering to ethical guidelines, industry standards, and legal regulations, promoting responsible and socially acceptable practices.

DISCUSSION

Despite efforts to address bias, challenges remain in achieving algorithmic fairness. Discussions revolve around the limitations of current approaches and the ongoing need for research to increase fairness in artificial intelligence systems.

Ethical AI systems often favor interpretation, but achieving human-understandable decisions, especially in complex models like deep neural networks, remains a challenge. The challenge is to find a compromise between accuracy and interpretability. Building and maintaining user trust is key. It is essential to determine how users perceive and trust ethical AI systems, taking into account factors such as transparency, accountability and the system's track record. The adaptability of AI ethical frameworks to evolving ethical standards is a matter of debate. How well can these frameworks adapt to changing social norms and what mechanisms exist for updating and revisions? Ethical AI systems aim to minimize unintended consequences, but there is the possibility of unforeseen problems occurring during implementation. That is why there should be strategies for proactive identification and resolution. The question is how much autonomy should artificial intelligence systems have and what mechanisms ensure accountability when decisions have significant consequences? The challenges and importance of accommodating different cultural perspectives within ethical AI systems should be considered. How well do these systems respect and accommodate cultural variations in ethical values? Ethical AI decisions often involve trade-offs between competing values. It should focus on the ethical dilemmas inherent in decision-making and how to strike a balance between conflicting principles. The role of education and awareness in ensuring that users understand the ethical considerations in AI decision-making needs to be emphasized. The question to be answered is how can organizations effectively communicate these principles to users? The long-term impact of widespread adoption of ethical AI systems is a key topic. The potential social benefits and challenges associated with the continuous integration of artificial intelligence in different sectors should be explored.

The use of ethical AI systems in decision-making gives positive results in terms of fairness, transparency and minimized discrimination. However, ongoing discussions are necessary to address the challenges, improve ethical frameworks, and ensure that AI systems are aligned with evolving ethical standards and societal expectations. Ethical AI should be viewed as a dynamic and iterative process involving continuous improvement, stakeholder engagement and commitment to ethical values.

III CONCLUSION

Creating AI systems that satisfy ethical principles is a multifaceted endeavor that demands careful consideration of transparency, fairness, privacy protection, accountability, user consent, robustness, and societal

impact. By adhering to these principles, we can foster the development of AI systems that contribute positively to society while minimizing potential risks and ethical concerns.

Transparency ensures that the inner workings of AI algorithms are clear and understandable, promoting trust among users and stakeholders. Fairness and bias mitigation address the critical need for equitable treatment, guarding against discrimination and promoting inclusivity. Privacy protection safeguards individuals' data, emphasizing the principles of data minimization and robust security measures.

Accountability and responsibility entail clearly defined roles and mechanisms for holding stakeholders accountable for the outcomes of AI systems. User consent and autonomy empower individuals to have control over their interactions and data, promoting ethical user engagement. Robustness and reliability, coupled with human oversight and intervention, guarantee that AI systems operate effectively and ethically, with the ability to handle unforeseen circumstances.

Considering the societal impact of AI systems is paramount, with a focus on addressing biases, discrimination, and broader implications on employment and social structures. Continuous monitoring and improvement, facilitated through feedback loops and adaptability, ensure that AI systems evolve responsibly over time.

Openness to ethical review, whether through internal mechanisms or external audits, promotes accountability and allows for independent evaluation of the ethical implications of AI systems. Encouraging public discourse on ethical considerations fosters a collaborative approach to shaping the ethical landscape of AI.

In conclusion, the ethical development and deployment of AI systems require a holistic approach that integrates these principles at every stage of the AI lifecycle. By doing so, we can harness the potential of AI to benefit society, foster innovation, and address complex challenges while upholding ethical standards and values.

Further research in the area of ethical AI systems is crucial to address emerging challenges, refine existing frameworks, and guide the development and deployment of AI technologies responsibly. Here are several potential avenues for future research: Explainability and Interpretability, Fairness and Bias Mitigation, Ethical Decision-Making Models, Human-AI Collaboration, Dynamic Ethical Frameworks, Cross-Cultural Considerations, Robustness and Security, Public Perception and Trust, Legal and Regulatory Frameworks, AI Education and Literacy, Human Rights and AI, Environmental Impact and Ethics in AI Research.

Continued research in these areas will contribute to the development of ethical AI systems that align with societal values, respect human rights, and foster trust among users and stakeholders. It requires interdisciplinary collaboration, involving researchers, ethicists, policymakers, and industry practitioners.

Research on ethical AI systems faces several limitations, and recognizing these constraints is essential for a comprehensive understanding of the field. Here are some common limitations: Rapid Technological, Data Biases and Representativity, Interdisciplinary Nature, Lack of Standardization, Unforeseen Consequences, Human Interpretability of AI Decisions, Resource Constraints, Cross-Cultural Variability, Legal and Regulatory Challenge, Limited User Awareness, Ethical Challenges in Deployment and Long-Term Effects. Recognizing and addressing these limitations is essential for advancing the field of ethical AI. It calls for ongoing interdisciplinary collaboration, robust methodologies, and a commitment to adapt ethical frameworks as technology continues to evolve.

References

[1] KHAN, Arif Ali, et al. Ethics of AI: A systematic literature review of principles and challenges. In: *Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering*. 2022. p. 383-392.

[2] ASHOK, Mona, et al. Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management*, 2022, 62: 102433.

[3] MORLEY, Jessica, et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and engineering ethics*, 2020, 26.4: 2141-2168.

[4] BALASUBRAMANIAM, Nagadivya, et al. Transparency and explainability of AI systems: ethical guidelines in practice. In: *International Working Conference on Requirements Engineering: Foundation for Software Quality*. Cham: Springer International Publishing, 2022. p. 3-18.

[5] STAHL, Bernd Carsten; STAHL, Bernd Carsten. Ethical issues of AI. *Artificial Intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies*, 2021, 35-53.

[6] RODGERS, Waymond; NGUYEN, Tam. Advertising benefits from ethical artificial intelligence algorithmic purchase decision pathways. Journal of business ethics, 2022, 178.4: 1043-1061.

[7] ALI, Khalda, et al. Artificial Intelligence: Benefits, Application, Ethical Issues, and Organizational Responses. In: Intelligent Sustainable Systems: Selected Papers of WorldS4 2022, Volume 1. Singapore: Springer Nature Singapore, 2023. p. 685-702.

[8] CHAMORRO-PREMUZIC, Tomas; POLLI, Frida; DATTNER, Ben. Building ethical AI for talent management. *Harvard Business Review*, 2019, 21: 1-15.

[9] VAKKURI, Ville, et al. ECCOLA—A method for implementing ethically aligned AI systems. Journal of Systems and Software, 2021, 182: 111067. [10] *Center for Policy on the Future of Computing (CPFC),* Accessed: 12-8-2023, Available: https://cyber.fsi.stanford.edu/

[11] MIRBABAIE, Milad, et al. Artificial intelligence in hospitals: providing a status quo of ethical considerations in academia to guide future research. *AI & society*, 2022, 37.4: 1361-1382.

[12] MORLEY, Jessica, et al. The ethics of AI in health care: a mapping review. Social Science & Medicine, 2020, 260: 113172.

[13] Schönberger, D. (2019). Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. International Journal of Law and Information Technology, 27(2), 171-203.

[14] SAURABH, Kumar, et al. AI led ethical digital transformation: Framework, research and managerial implications. Journal of Information, Communication and Ethics in Society, 2022, 20.2: 229-256.

[15] KIRADOO, Giriraj. Unlocking the Potential of AI in Business: Challenges and Ethical Considerations. *Recent Progress in Science and Technology*, 2023, 6: 205-220.

[16] FLORIDI, Luciano, et al. An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Ethics, governance, and policies in artificial intelligence*, 2021, 19-39.

[17] EITEL-PORTER, Ray. Beyond the promise: implementing ethical AI. *AI and Ethics*, 2021, 1: 73-80.

[18] FLORIDI, Luciano, et al. How to design AI for social good: seven essential factors. *Ethics, Governance, and Policies in Artificial Intelligence*, 2021, 125-151. [19] Tegmark, M. „Life 3.0: Being Human in the Age of Artificial Intelligence", 2017

[20] STRÜMKE, Inga; SLAVKOVIK, Marija; MADAI, Vince Istvan. The social dilemma in artificial intelligence development and why we have to solve it. *AI and Ethics*, 2022, 2.4: 655-665. [21] *VAN DE POEL, Ibo. Embedding values in artificial intelligence (AI) systems. Minds and Machines, 2020, 30.3: 385-409.*

[22] TØNNESSEN, Siri; URSIN, Gøril; BRINCHMANN, Berit Støre. Care-managers' professional choices: ethical dilemmas and conflicting expectations. BMC Health Services Research, 2017, 17.1: 1-10.

[23] ŽYDŽIŪNAITĖ, Vilma, et al. Ethical dilemmas concerning decision-making within health care leadership: a systematic literature review. Medicina, 2010, 46.9: 595.

[24] GUPTA, Kirti, et al. Harnessing AI for Strategic Decision-Making and Business Performance Optimization. *International Journal of Intelligent Systems and Applications in Engineering*, 2023, 11.10s: 893-912.

[25] MA, Lei; ZHANG, Zhongqiu; ZHANG, Nana. Ethical dilemma of artificial intelligence and its research progress. In: *IOP conference series: materials science and engineering*. IOP Publishing, 2018. p. 062188. [26] Johnson, D. R., Prietula, M. J., „A Survey of AI Ethics in Business Literature: Maps and Trends between 2000 and 2021", Frontiers, 2022

[26] DHIRANI, Lubna Luxmi, et al. Ethical dilemmas and privacy issues in emerging technologies: a review. *Sensors*, 2023, 23.3: 1151.

[27] NAZER, Lama H., et al. Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digital Health*, 2023, 2.6: e0000278.

[28] BALASUBRAMANIAM, Nagadivya, et al. Transparency and explainability of AI systems: ethical guidelines in practice. In: *International Working Conference on Requirements Engineering: Foundation for Software Quality*. Cham: Springer International Publishing, 2022. p. 3-18.

[29] LEE, Raymond ST; LEE, Raymond ST. AI Ethics, Security and Privacy. *Artificial Intelligence in Daily Life*, 2020, 369-384.

[30] CAÑAS, José J. AI and Ethics when human beings collaborate with AI Agents. *Frontiers in psychology*, 2022, 13: 836650.

[31] DE ALMEIDA, Patricia Gomes Rêgo; DOS SANTOS, Carlos Denner; FARIAS, Josivania Silva. Artificial intelligence regulation: a framework for governance. *Ethics and Information Technology*, 2021, 23.3: 505-525.

[32] MORLEY, Jessica, et al. The ethics of AI in health care: a mapping review. *Social Science & Medicine*, 2020, 260: 113172.

[33] Vesnic-Alujevic, L., Nascimento, S., & Polvora, A. (2020). Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks. *Telecommunications Policy*, *44*(6), 101961.

[34] NASSAR, Ahmed; KAMAL, Mostafa. Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations. *International Journal of Responsible Artificial Intelligence*, 2021, 11.8: 1-11.

[35] EITEL-PORTER, Ray. Beyond the promise: implementing ethical AI. *AI and Ethics*, 2021, 1: 73-80.

[36] DWIVEDI, Yogesh K., et al. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 2023, 71: 102642.

[37] SCHULTZ, Mario D.; SEELE, Peter. Towards AI ethics' institutionalization: knowledge bridges from business ethics to advance organizational AI ethics. *AI and Ethics*, 2023, 3.1: 99-111.

[38] DU, Shuili; XIE, Chunyan. Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research*, 2021, 129: 961-974.

[39] PAGANONI, Maria Cristina, et al. Ethical concerns over facial recognition technology. *Anglistica AION an Interdisciplinary Journal*, 2019, 1: 83-92.

[40] SHAPIRO, Aaron. Optimizing the immeasurable: On the techno-ethical limits of predictive policing. In: *Artificial Intelligence and the City*. Routledge, 2023. p. 240-257.

[41] GROSEK, Štefan, et al. The first nationwide study on facing and solving ethical dilemmas among healthcare professionals in Slovenia. *Plos one*, 2020, 15.7: e0235509.

[42] CSILLAG, Sára. Ethical dilemmas and moral muteness in the HRM profession. *Society and Economy*, 2019, 41.1: 125-144.

[43] GOSZTONYI, Gergely; LENDVAI, Gergely Ferenc. Deepfake: A Multifaceted Dilemma in Ethics and Law. *Journal of Information Ethics*, 2023, 32.2: 109-121.

[44] ZHU, Xi; GU, Zhiqiang; WANG, Zhen. Ethical challenges and countermeasures of autonomous vehicles. In: *E3S Web of Conferences*. EDP Sciences, 2021. p. 04016.

[45] FONTES, Catarina, et al. AI-powered public surveillance systems: why we (might) need them and how we want them. *Technology in Society*, 2022, 71: 102137.

[46] CABRERA, Johana, et al. Ethical dilemmas, mental health, artificial intelligence, and llm-based chatbots. In: *International Work-Conference on Bioinformatics and Biomedical Engineering*. Cham: Springer Nature Switzerland, 2023. p. 313-326.

[47] GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 2020, 7.1: 2053951719897945.