**Ethics in the Digital Era: Navigating the AI Landscape**

***Student***: Hincu Alice-Ramona

***Specialization***: Distributed Systems in Internet

**ABSTRACT**

This paper presents a comprehensive analysis of the ethical challenges and considerations that arise in the digital era, particularly as they pertain to the development and application of Artificial Intelligence (AI). Amidst the rapid expansion of AI across various sectors, this paper argues for the essential role of ethics in guiding AI towards beneficial outcomes for society. It explores the ethical dimensions of AI in art, the proliferation of deepfakes, and the biases inherent in AI systems, providing real-life examples of how these issues manifest in areas such as employment, healthcare, and media. The paper emphasizes the dual nature of AI as a tool for innovation and a potential conduit for replicating societal prejudices. It advocates for ethical vigilance, transparency, and accountability in AI development and calls for multi-stakeholder collaboration to ensure that AI advances with a moral compass. The journey through the AI landscape is navigated with an emphasis on upholding human dignity, ensuring equitable access to technology's benefits, and fostering an inclusive society. In concluding, the paper underscores the collective responsibility of technologists, policymakers, and society at large to ensure that the path of AI is aligned with ethical principles, ensuring a future where technology amplifies the best of humanity.

**Contents**

[1. Introduction 3](#_Toc1548260812)

[2. Ethical Foundations in AI 5](#_Toc505115583)

[Proportionality and Do No Harm 5](#_Toc1288388464)

[Safety and Security 6](#_Toc1821278429)

[Right to Privacy and Data Protection 6](#_Toc1696759769)

[Multi-stakeholder and Adaptive Governance & Collaboration 6](#_Toc1884232668)

[Responsibility and Accountability 7](#_Toc623380741)

[Transparency and Explainability 7](#_Toc145879420)

[Human Oversight and Determination 7](#_Toc1820194706)

[Sustainability 8](#_Toc1808435031)

[Awareness & Literacy 8](#_Toc1475751917)

[Fairness and Non-Discrimination 9](#_Toc1571381223)

[3. Cases and Applications 9](#_Toc635164)

[3.1 Artificial Intelligence in Art 9](#_Toc1439918340)

[3.1.1 Case Study: Midjourney vs AIVA 11](#_Toc1487095226)

[3.2 Artificial Intelligence and Deepfakes 11](#_Toc117667411)

[3.2.1 Case Study: Deepfakes for media propaganda in Rusia and Ucraina war 12](#_Toc1296142035)

[3.2.1 Case Study: DeepNude 12](#_Toc1810169953)

[4. Examining Biases in AI 13](#_Toc157507750)

[Conclusion 15](#_Toc1908781795)

[Bibliography 16](#_Toc566209536)

**Chapter 1**

# **1. Introduction**

At the heart of human civilization lies a set of principles that govern our conduct: ethics. These are the moral guidelines that help us discern right from wrong, guiding our decisions and actions.

Artificial Intelligence, commonly known as AI, refers to the simulation of human intelligence in machines that are programmed to think and learn like humans. Artificial Intelligence is based on creating intelligent systems that can understand their environment and make decisions to achieve certain goals. AI combines different areas of study like computer science, psychology, and neuroscience. Its main goal is to create machines that can adapt and learn, not just copy human intelligence.

The exponential growth of AI can be attributed to several key factors. Firstly, the explosion of big data from various sources like the internet, business operations, and scientific research has provided AI systems with the raw material needed for advanced learning and analysis. Secondly, the significant improvements in computational power and algorithms have enabled these systems to process and analyze this data more effectively than ever before. Vast data availability and computational advancement has allowed AI to evolve rapidly.

Advancements in machine learning and deep learning are changing many areas of technology. Machine learning allows algorithms to learn from data and make decisions with little human help. Deep learning, a type of machine learning, uses networks similar to the human brain. These networks help machines understand very complex data.

This change is affecting many industries. For example, in business, AI helps make better decisions by analyzing complex data. In healthcare, it can find patterns in medical images as accurately as doctors. AI is also helping to develop self-driving cars by processing lots of information from sensors. AI is becoming a big part of fields like robotics, where it helps robots act more sophisticatedly, and in natural language processing, which makes computers better at understanding human language. The versatility of AI extends into creative domains, challenging our conventional understanding of art and innovation. AI-generated art, music, and literature are pushing the boundaries of creativity, raising questions about the nature of art and the role of human creativity in the age of machines. This intersection of technology and creativity is not just a testament to AI's capabilities but also a ground for ethical debate over the ownership and originality of AI-generated content.

Yet, as AI continues to permeate various sectors, the ethical implications become increasingly complex. The vast data AI systems rely on often include personal information, raising privacy concerns. The decision-making processes of AI, while efficient, can reflect the biases present in their training data, leading to fairness and discrimination issues. Moreover, as AI takes on roles with greater responsibility, from driving cars to making healthcare decisions, the question of accountability in the case of errors or accidents becomes paramount.

AI, with its vast capabilities, brings forth new scenarios where our traditional ethical understanding might be tested. The decisions made by AI systems can have far-reaching consequences, and it is essential that these systems are guided by ethical principles.

It's essential to understand the full capabilities of AI, but equally important is recognizing the moral and ethical boundaries of it. This paper showcases the ethical challenges presented by artificial intelligence. Chapter 2 discussed the need for principles to guide AI development. Chapter 3 presents in depth some applications of the AI and real-life scenarios of its impact. Chapter 4 confronts the biases inherent in AI systems. Chapter 5 concludes the information while calling for a balanced approach to AI that integrates ethical considerations every step of the way. Lastly, the Bibliography provides a comprehensive list of sources and references that underpin the research and discussions presented in this paper

**Chapter 2**

# **2. Ethical Foundations in AI**

As the capabilities of Artificial Intelligence continue to expand, understanding the ethical foundations that guide its development and deployment becomes crucial. Incorporating ethical principles in the development and application of Artificial Intelligence (AI) is crucial for its responsible use. This chapter will explore ten ethical principles to guide AI practices which UNESCO has outlined. [1]

### **Proportionality and Do No Harm**

This principle asserts that AI should be used only as much as needed to achieve its goals and it should check for any possible risks for harms it could cause. This means looking closely at how AI might affect people and society.

For instance, in healthcare, AI is used for diagnosing diseases. While AI can help doctors make faster and more accurate diagnoses, people have to make sure that the AI systems don't misdiagnose, which could lead to incorrect treatments. Before implementing such AI systems, hospitals and developers must rigorously test them to identify any potential risks, like inaccuracies in diagnosis due to biases in the training data. Similarly, in the case of self-driving cars, this principle requires a thorough assessment of safety risks. AI-driven vehicles must be meticulously tested under various road conditions to ensure they can make safe decisions, such as stopping at a red light or avoiding pedestrians. The aim is to minimize the risk of accidents, ensuring the AI in these vehicles does no harm.

### **Safety and Security**

The principle of Safety and Security emphasizes that AI should not only avoid causing unintended harm but also be secure against potential attacks and vulnerabilities. It's essential for those involved in AI development and deployment to consider both safety and security risks.

In the context of AI, safety risks refer to scenarios where AI systems might inadvertently cause harm due to errors or unforeseen situations. Security risks, on the other hand, involve AI systems being vulnerable to malicious attacks, which could lead to harmful outcomes. For example, un personal assistant devices that use AI, safety risks might include accidental misinterpretation of commands that could lead to unwanted actions, while security risks could involve personal data being accessed by hackers. Ensuring robust data encryption and user authentication methods are in place is critical to mitigate these risks.

### **Right to Privacy and Data Protection**

This principle highlights the importance of safeguarding personal information. Privacy should be a priority at every stage of the AI lifecycle, from design to deployment, and that strong data protection measures must be in place.

For instance, consider AI systems used in online retail. These systems often analyze customer data to personalize shopping experiences. While this can enhance user experience, it's crucial to ensure that customers' personal data, like browsing history and purchase records, are kept private and secure. Retailers must implement data encryption and strict access controls to protect this information from unauthorized access or breaches, or else retailers are more susceptible to data breaches and exposed data can be used for identity theft.

### **Multi-stakeholder and Adaptive Governance & Collaboration**

This stresses the need to follow international laws and respect each country's rules when using AI and it highlights the importance of involving different groups of people in making decisions about AI.

Consider the use of AI in international research collaborations. In such scenarios, it's crucial that the data used by AI respects the laws of the countries involved. This might include adhering to data protection regulations such as the European Union's General Data Protection Regulation (GDPR). Ensuring compliance with these laws not only respects national sovereignty but also builds trust among international partners

### **Responsibility and Accountability**

Responsibility in AI refers to the obligation of AI developers and users to ensure their systems are used ethically and legally. Accountability means being able to trace decisions made by AI systems back to the entities responsible for them and holding them accountable for the outcomes. Developers are accountable for designing and programming the AI systems. They are responsible for ensuring the system operates as intended, is free from biases, and adheres to ethical and legal standards. If an AI system malfunctions due to a flaw in its design or programming, the developers are held accountable. Users of AI systems, such as businesses or organizations that implement these technologies, are accountable for how they deploy and utilize the AI. They must ensure the AI is used ethically, in compliance with regulations, and for purposes that align with societal values. If an AI system is used in a way that leads to unethical outcomes or harm, the users can be held accountable for these actions.

Take, for example, the use of AI in the criminal justice system, such as in risk assessment tools for sentencing or bail decisions. Here, it’s vital that these AI systems can be audited and their decision-making processes understood. If an AI system suggests a harsher sentence for an individual, there needs to be a clear understanding of how that decision was made. This is to ensure that the system is not unfairly biased against certain groups and adheres to human rights norms.

### **Transparency and Explainability**

This principle means that AI systems should be clear about how they work and make decisions. People using AI should be able to follow what the AI is doing and why. However, how much regular people can see and understand about AI's decision-making process depends on the situation. It's important to find the right balance, making sure that being open about how AI works doesn't create other problems, like risking people's private information, making the AI less safe, or making it easier for someone to break into the system. So, in some cases, AI might need to keep some details hidden to protect privacy and security.

An AI system in healthcare that predicts patient outcomes must not reveal sensitive personal data in its explanations. This requires careful design to ensure that while the AI’s decision-making process is as clear as possible, patient confidentiality is not compromised.

### **Human Oversight and Determination**

This principle emphasizes the importance of keeping humans in the loop, ensuring that AI systems do not completely replace human responsibility and accountability. It's about ensuring that, even as AI systems become more advanced and autonomous, final decisions and responsibilities should still rest with humans.

In practical terms, this principle can be illustrated in various scenarios:

* **In Medical Decision-Making**: While AI systems, like those used for diagnosing diseases or recommending treatments, significantly aid healthcare professionals, the final decisions should always be under human control. For instance, if an AI system suggests a particular treatment plan, a qualified healthcare professional should review this plan, considering all aspects of the patient's condition, before making the final decision. This approach ensures that AI aids, rather than replaces, human expertise and judgment.
* **In Legal Judgments and Law Enforcement**: AI is increasingly used to assist in legal cases, from predicting recidivism risks to aiding in forensic analysis. However, it's crucial that judges and law enforcement officers retain final decision-making authority. An AI might provide insights based on data analysis, but a human should always interpret these insights and make the final judgment, ensuring that decisions are fair, just, and consider all legal and ethical implications.
* **In Workplace Automation**: With AI being used to automate tasks in the workplace, this principle ensures that humans have ultimate control over critical decisions. For example, in HR-related AI systems used for recruiting or performance evaluations, the final decisions should be made by human managers, who can consider factors beyond the AI's analysis.

### **Sustainability**

The principle of Sustainability highlights the importance of assessing AI technologies against their impact on environmental, social, and economic sustainability. This principle aligns with the goals outlined in the United Nations' Sustainable Development Goals (SDGs), ensuring that AI contributes positively to long-term global sustainability.

For instance, in the context of environmental sustainability, AI can be used in climate change research, helping scientists model and predict climate patterns. However, the sustainability principle requires us to also consider the environmental cost of running large AI systems, which often require significant energy consumption. To align with this principle, developers and users of AI in this field must optimize algorithms for energy efficiency and consider using renewable energy sources to power these systems.

### **Awareness & Literacy**

The principle of Awareness & Literacy in the realm of AI focuses on the need for public understanding and education about AI technologies and data, so that individuals can engage with AI responsibly and make informed decisions.

The first aspect of this principle involves promoting open and accessible education about AI. This means offering learning resources and educational programs about AI, making it understandable for people of all ages and backgrounds. For instance, schools and universities could integrate courses on AI and data science into their curriculums, ensuring students have a fundamental understanding of these technologies. Having discussions and workshops in the community can help people learn about AI and share their thoughts. Media literacy is also a big part of this. As AI is used more in things like news and social media, people need to be able to tell what is created by AI. Learning how to spot AI-generated content helps people understand what they're seeing online.

### **Fairness and Non-Discrimination**

This principle is about making sure AI treats everyone fairly and doesn't discriminate. AI should work towards social justice and be accessible to all people, no matter who they are.

For example, in job hiring, AI used to screen applicants should be fair. It shouldn't favor or reject someone based on their gender, race, or age. If an AI system was trained with data from a company that used to hire more men than women, it might keep doing that unless it's carefully checked and corrected.

Also, AI technology should be made so everyone can use it, including people with disabilities. For example, voice recognition should understand different types of speech, even if someone speaks differently because of a disability.

**Chapter 3**

# **3. Cases and Applications**

The rapid advancement and integration of AI into various sectors have also given rise to complex ethical challenges. This chapter analyzes some ethical problems from different fields, followed by examples.

## **3.1 Artificial Intelligence in Art**

In art, AI challenges our perceptions of creativity and originality, questioning the very essence of human artistic expression. Can a machine be creative, or is it merely a tool for human creativity? What happens to the concept of authorship when an AI creates a piece of art?

There are a lot of negative aspects to consider when using Artificial Intelligence in the process of creating art:

* The growing presence of AI in the art world brings up a big issue about human creativity. AI can easily copy many artistic styles, which has resulted in lots of art pieces that look alike. Because there's so much AI-made art online, it's harder for art made by people to get noticed for its unique qualities and originality. The worry is that the special touch and recognition that artists get for their unique skills could be lost among all the art made by algorithms. For artists trying to make a name for themselves with their unique style and creative ideas, this trend of AI art creates a real challenge.
* The easy availability and low cost of AI technology add to the challenges for artists. AI art software, which anyone with a computer can use, makes it much easier for people to create and share art. This means there's a lot more AI-created art out there, and it becomes tougher for artists to make a living just from their artwork. Traditional skills and techniques that artists spend years mastering might be less valued as AI can replicate these skills quickly and easily. This can lead to a decrease in the demand for handcrafted art, impacting the livelihoods of artists who rely on selling their unique creations.
* This situation highlights the question *“what makes art valuable?”*. If art created by humans and AI is hard to tell apart, then what about the importance of the artist's own emotions and the stories they tell through their art? As AI blurs these lines, it forces to reconsider what do the people value in art and how it reflects the culture and society. This isn't just a challenge for individual artists; it affects the whole art community and how society views art in the age of AI.
* AI systems use large amounts of data for training, much of which comes from the internet and might contain copyrighted content used without permission. This brings up concerns about the rights to intellectual property and the ethics of using such content for AI-made art. Also, AI doesn't have the same feelings, emotions, and personal experiences that human artists put into their art. This limits how well art can show and communicate the deep aspects of human life, which lessens its influence on society.
* Moreover, the use of copyrighted materials by AI raises legal and ethical dilemmas. Artists and creators whose work is used without acknowledgment or compensation may feel their rights have been violated, leading to potential legal disputes. This situation complicates the already challenging task of protecting intellectual property in the digital age.

On a more positive note, rather than replacing human artists, AI can be seen as a tool to enhance human creativity. It can suggest new ideas, patterns, or compositions that artists can develop further. One use case is creating concept art and brainstorming ideas before creating the final product.

### **3.1.1 Case Study: Midjourney vs AIVA**

This case study examines two AI tools, Midjourney and AIVA, which have been used in the creation of digital art and music, respectively. Midjourney is an AI tool known for generating visual art. It creates images based on textual descriptions provided by users. AIVA (Artificial Intelligence Virtual Artist) is an AI platform specializing in composing music, particularly in the classical genre. Midjourney and AIVA represent two different approaches to AI in the arts. Midjourney's approach raised ethical concerns, while AIVA’s method is more aligned with copyright laws and artist collaboration.

Midjourney faced criticism for using internet-sourced materials, which often included copyrighted content, without explicit permission. This raised questions about the legal and ethical use of existing artworks to train AI systems. The controversy highlighted concerns within the artist community about AI's role in art, especially regarding originality, intellectual property rights, and the potential devaluation of human-created artwork. Popular artists that are known for their unique style and post on platforms like Instagram complained about this tool and even showed how the AI could reproduce their art style by just adding “<username> style” in the prompt. This happened because the tool was fed their artwork, but without their consent, leading to a lot of backlashes from the art community.

Unlike Midjourney, AIVA primarily uses public domain scores for training. This approach respects copyright laws and ethical guidelines, avoiding the use of protected materials without consent. AIVA is known for collaborating with human artists and composers, ensuring a respectful integration of AI in the creative process. This approach not only enhances the quality of the output but also maintains ethical standards in AI-generated art.

The comparison of Midjourney and AIVA offers valuable insights into the ethical considerations of AI in art and music. It underscores the importance of developing AI tools that respect copyright laws and the creative contributions of human artists, ensuring that AI is used as a tool to enhance, not replace, human creativity.

## **3.2 Artificial Intelligence and Deepfakes**

Deepfakes refer to synthetic media where a person's likeness is replaced with someone else's likeness, often using AI algorithms. The quality of deepfakes has seen remarkable advancements due to the rapid development of AI technologies. As deepfakes become more sophisticated, distinguishing them from authentic media becomes more challenging. The AI used to create deepfakes is in a constant arms race with the AI developed to detect them, leading to an ongoing cycle of improvement on both sides.

While this technology has creative and beneficial applications, it also poses serious ethical and societal challenges. One of the most alarming uses of deepfakes is the creation of videos or audio recordings where public figures appear to say or do things that never actually happened. This can be used to spread false information, manipulate public opinion, or discredit individuals. Fabricated videos of politicians making controversial statements or engaging in inappropriate behavior could sway public opinion and influence election outcomes. They can be employed to falsify historical events, create fake news stories, or propagate harmful stereotypes, leading to widespread misinformation and societal discord. The difficulty in verifying the authenticity of videos and audio clips complicates reporting and can undermine public trust in media.

Similarly, in the realm of pornography, AI's capabilities raise critical ethical issues. The technology's ability to create and manipulate images and videos pose serious questions about consent and the representation of individuals. The most significant issue with deepfakes in pornography is the non-consensual use of an individual's likeness. Celebrities, public figures, and private individuals, particularly women, have been targets of deepfake pornography, which is created and distributed without their permission. For victims, the existence of deepfake pornographic content can lead to significant emotional harm, including anxiety, depression, and a sense of violation. It can also have real-world consequences, such as damage to their reputation, career, and personal relationships. It's often difficult to track the creators of such content due to the anonymity afforded by the internet, making legal recourse challenging. Ethically, this use of AI technology crosses boundaries of consent and personal rights. The prevalence of deepfake pornography contributes to a broader culture of exploitation and objectification, especially towards women. It raises concerns about the normalization of such content and its impact on societal attitudes towards consent and sexual abuse.

### **3.2.1 Case Study: Deepfakes for media propaganda in Rusia and Ucraina war**

A significant real-life instance of deepfakes impacting the media occurred during the Russo-Ukrainian War. Researchers from University College Cork examined tweets and found that deepfake videos were used as a form of wartime misinformation and propaganda. For example, a deepfake video of Russian President Vladimir Putin announcing peace with Ukraine and a deepfaked message of Ukrainian President Volodymyr Zelensky surrendering were circulated. These deepfakes undermined the trust in authentic footage coming from the conflict, to the point where people lost trust in any footage from the conflict. This study showed that the lack of deepfake literacy led to significant misunderstandings of what constitutes a deepfake and the need to encourage literacy in these new forms of media [3]

### **3.2.1 Case Study: DeepNude**

There have been real-life instances where deepfake technology was used to generate non-consensual nude images of individuals, often by sourcing their photos from social media platforms. One notable example is the case of an application known as "DeepNude." This app used AI algorithms to create fake nude images of women by 'undressing' them in photos. The app worked by analyzing photos of clothed women and then replacing their clothing with artificially generated naked bodies.

The emergence of DeepNude raised significant concerns regarding privacy and consent. The app was widely criticized for its potential to enable sexual harassment and for contributing to the problem of non-consensual pornography. It posed a direct threat to the privacy and dignity of any woman whose photo could be found online, as it could generate realistic-looking nude images without their consent.

This instance serves as a stark reminder of the potential for deepfake technology to be used in harmful ways, violating privacy and exacerbating issues of non-consensual imagery and sexual harassment. It stresses the importance of ongoing vigilance and regulatory measures to prevent the misuse of advanced technologies in invading personal privacy and rights.

**Chapter 4**

# **4. Examining Biases in AI**

This chapter addresses the significant and troubling issue of biases in Artificial Intelligence, particularly focusing on racism, misogyny, and other forms of discrimination. AI systems, which are often known and praised for their objectivity and precision, can become conduits for perpetuating societal biases. This happens primarily due to the nature of the data on which they are trained and the way their algorithms are designed. The chapter will explore the origins, manifestations, and impacts of these biases, and discuss potential strategies for mitigation.

Biases in AI are predominantly data-driven. An AI system is only as unbiased as the data it processes. If this data reflects historical prejudices or societal imbalances, the AI will likely replicate these in its decisions and recommendations. For instance, in recruitment and employment, AI systems used for screening job applications have sometimes shown a tendency to favor white male candidates. This could be a reflection of the existing composition of the workforce in the dataset or inherent prejudices in the selection criteria encoded into the algorithm.

An AI recruitment tool used by Amazon, which reportedly showed bias against women. The system was trained on resumes submitted to the company over a 10-year period, most of which came from men, reflecting the male dominance in the tech industry. As a result, the AI learned to prefer male candidates over female candidates, penalizing resumes that included the word "women's," as in "women's chess club captain." [4]

The repercussions of such biases in AI are profound. They perpetuate systemic racism and misogyny, reinforcing and even amplifying existing societal disparities and injustices. The economic and social consequences are far-reaching, ranging from restricting employment prospects for marginalized communities to sustaining income inequality and societal division.

A study published in Science [5] revealed that an AI system used in healthcare in the United States was less likely to refer black people to care programs than white people, despite being equally sick. The algorithm predicted healthcare costs rather than medical needs, and because historically, black patients incurred lower healthcare costs, the AI falsely concluded they required less care.

COMPAS, a software used in the US criminal justice system for assessing the likelihood of a convict reoffending, was found to exhibit racial bias. An investigation by ProPublica found that the algorithm was nearly twice as likely to incorrectly label Black defendants as likely to reoffend compared to White defendants. [6]

The ethical implications of biased AI systems are significant. They raise questions about the responsibilities of AI developers and the companies that deploy these systems. Legally, there's an emerging discourse on how anti-discrimination laws and regulations can be applied and adapted to AI and machine learning systems. These laws are intended to prevent biased outcomes and ensure fairness in automated decision-making processes.

Mitigating biases in AI is a challenge. It requires the use of diverse and representative datasets in training AI systems to ensure a broad range of perspectives and experiences are accounted for. There's also a pressing need for greater transparency in how AI algorithms make decisions and for mechanisms to hold AI systems accountable for biased outcomes. Furthermore, continuous monitoring and adjustment of AI systems are essential to identify and correct biases that may emerge over time.

**Chapter 5**

# **Conclusion**

In conclusion, this paper analyzed the ethical considerations that are crucial in the age of Artificial Intelligence, establishing not only ethical benchmarks that serve as a guide, but also showcasing real-life examples of ethical problems regarding the use of AI.

The case studies and real-world examples have shown AI's dual capacity to enhance people’s lives and to pose significant ethical dilemmas. From art and media to the creation of disturbingly realistic deepfakes, the technology's potential for both creativity and deception has been starkly highlighted, reinforcing the need for ethical guardrails.

The issue of bias in AI has been a central concern, revealing how AI can inadvertently inherit human prejudices. This reinforces the importance of embedding ethical principles into the very fabric of AI systems to prevent discrimination and ensure fairness.

AI’s integration into society will challenge people to continually reflect on ethical implications. The potential impact on employment, international diplomacy, healthcare, and privacy demands not just technological oversight but ethical foresight.

In the end, this paper stands as a call for constant ethical awareness in the digital age. As society ventures through the evolution of AI, it's crucial to place ethics at the forefront to direct this influential technology towards results that respect human dignity and promote a fair society. The path taken with AI is intertwined with ethical considerations just as much as it is with technological advancements. It is a shared duty to make sure that as AI progresses, it is guided by strong ethical principles, leading people to a future where technology helps them instead of replacing or discriminating them.

# **Bibliography**

**[1]** UNESCO, *Ethics of Artificial Intelligence,*[**https://www.unesco.org/en/artificial-intelligence/recommendation-ethics**](https://www.unesco.org/en/artificial-intelligence/recommendation-ethics)(last opened: 08.01.2024)

**[2]** Girls for Algorithmic Justice (GFAJ), *Negative Impacts of AI Art on Artists in AI Ethics***,** [**https://medium.com/@gfaj/negative-impacts-of-ai-art-on-artists-in-ai-ethics-305df8619b17#:~:text=AI%20algorithms%20are%20trained%20on,to%20create%20AI%2Dgenerated%20art**](https://medium.com/@gfaj/negative-impacts-of-ai-art-on-artists-in-ai-ethics-305df8619b17#:~:text=AI%20algorithms%20are%20trained%20on,to%20create%20AI%2Dgenerated%20art)**.** (last opened: 08.01.2024)

**[3]** University College Cork, *First-ever study of wartime deepfakes reveals their impact on news media*,[**https://techxplore.com/news/2023-10-first-ever-wartime-deepfakes-reveals-impact.html,**](https://techxplore.com/news/2023-10-first-ever-wartime-deepfakes-reveals-impact.html,)27.10.2023(last opened: 08.01.2024)

**[4]** James Vincent, *Amazon reportedly scraps internal AI recruiting tool that was biased against women***,** [**https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report,**](https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report,)10.10.2018(last opened: 08.01.2024)

**[5]** Starre Vartan, *Racial Bias Found in a Major Health Care Risk Algorithm*,[**https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/,**](https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/,)24.10.2019(last opened: 08.01.2024)

**[6]** Matthias Spielkamparchive, Inspecting Algorithms for Bias, [**https://www.technologyreview.com/2017/06/12/105804/inspecting-algorithms-for-bias/**](https://www.technologyreview.com/2017/06/12/105804/inspecting-algorithms-for-bias/) **,** 12.06.2017 (last opened: 08.01.2024)