

BETWEEN SCYLLA AND CHARYBDIS: A CATEGORIZATION OF ETHICAL CHALLENGES ASSOCIATED WITH THE USE OF ARTIFICIAL INTELLIGENCE IN THE MEDIA

Aleksa Mitić

REVIEW ARTICLE / DOI: 10.20901/ms.16.32.5 / SUBMITTED: 31.08.2025.

ABSTRACT *The increased rate of artificial intelligence use within the media industry has led to the emergence of new ethical problems, as well as the expansion of existing ones. Considering various factors and agents, finding a unique solution to these ethical issues would be a highly complex and arguably impossible undertaking. The general hypothesis of this paper is that current ethical problems related to the use of artificial intelligence in the media can be separated into two main categories: solvable and insolvable. The analysis of the relevant literature will cover the most prominent ethical problems and enable their classification according to the previously mentioned categories. The primary aim of this paper is to present which of the existing ethical issues, categorized as solvable, should be prioritized in future research, in order to eventually find a solution. The group of unsolvable issues should not be completely ignored, but decrease in research frequency is suggested, in line with the advancement of more sophisticated AI models that may be better suited for ethical integration into media practices.*

KEYWORDS

ARTIFICIAL INTELLIGENCE, MEDIA, ETHICS, ETHICAL PROBLEMS, JOURNALISM

Author's note

Aleksa Mitić :: Faculty Philosophy, University of Niš :: miticvaleksa@gmail.com

INTRODUCTION

Artificial intelligence (AI) is probably the most significant technological innovation of the twenty-first century considering its scale of action, which refers to almost all areas of modern social life. The wide range of potential benefits that AI offers in spheres such as economy, ecology, medicine and health, education, etc., (Nasim et al, 2020) already gives it the attribute of a representative technology of this period in the era of modern civilization. The dramatically rapid development and expansion that AI has achieved (Gruetzemacher & Whittlestone, 2022) have resulted in the emergence of a complex situation regarding the supervision of the use of this technology, whose effects and trends cannot be accurately predicted. In recent years, researchers' concerns in this field have been increasingly directed towards negative implications that could arise from the current impacts of this technology on modern society (Zhang & Dafoe, 2019). The prediction of the negative implications arising from application of AI has its roots, among other things, in a diverse range of ethical problems that arise as a result of its use. The irony inherent in the coexistence of the AI's potential benefits and risks is recognizable in areas which it has been implemented and which now face serious language-related challenges and risks brought by AI (Huang et al., 2022). A similar situation exists in the media industry, where the current trends in the use of AI for processing and distributing media content (de-Lima Santos & Ceron, 2021) further complicate the already complex situation of the ethics of media reporting. Ethical issues that the media face on a daily basis gain a new dimension due to the inclusion of technologies such as AI in their work. Researcher Meredith Broussard and colleagues state that "artificial intelligence (AI) represents a new medium through which journalists can express and apply their ethical and normative values through the code they implement" (Broussard et al., 2019, p.7) in order to inform the public correctly and in a timely manner. The metamorphosis and eventual resolution of existing ethical problems in the media will largely depend on the way in which the use of AI is regulated by journalists and other media workers.

The theoretical framework of this paper is based on the assumption that certain types of ethical problems that arise due to the use of AI in the media can be solved through an adequate approach. The aim of the paper is to identify potentially solvable ethical problems, as well as to provide an explanation, i.e. the way in which recognized ethical problems can be solved. The hypothesis that the existing ethical problems associated with AI in the media can be classified into three groups arises from the stated aim: (1) solvable ethical problems; (2) partially solvable ethical problems; (3) unsolvable ethical problems. The stated typology is a modification of the original typology offered by the philosopher Rosalind Hursthouse. In her book "On Virtue Ethics", the author proposes a typology of moral dilemmas from the perspective of virtue ethics: (1) solvable dilemmas; (2) unsolvable dilemmas; (3) tragic dilemmas (Hursthouse, 2000, p. 35-70). The topic of this research necessitates the application of a certain modification to Hursthouse's typology. Accordingly, the first two types of moral dilemmas are retained, while the third, "tragic", dilemmas are replaced with "partially solvable" problems. The term "solvable dilemmas" refers to cases in which certain virtues conflict with each other, but in which a solution can be clearly chosen that will lead to an exclusively positive outcome (Hursthouse,

2000 p. 35-39). On the other hand, the “unsolvable” and “tragic” dilemmas, according to Hursthouse, refer to “situations in which whatever you do, you necessarily come out with dirty hands” (Hursthouse, 2000, p. 57). The primary difference between partially solvable and unsolvable ethical problems is that partially solvable problems can only be solved to a certain extent, while another part of the problem remains unsolvable. By contrast, unsolvable problems cannot be solved in any segment. Thus, unsolvable dilemmas represent cases involving an inevitable choice that will result in a moral violation and damage to at least one of the conflicting parties. The conceptual substitution introduced in this paper, in the form of “partially solvable” problems is explained as a set of situations in which it is possible to solve only part of the problem, that is, to have a completely positive outcome, while the root of the problem remains unsolvable.

The selection the literature to be used will be guided not only by the relevance of the topic of the paper, but also by the research problem and the initial hypothesis. Given the interdisciplinary nature of this paper, which deals with the ethics of AI in the media, the literature also exhibits an interdisciplinary character. A review of the literature combined with analytical approach, in addition to the individual analysis of ethical problems of AI in the media, should also enable the presentation of potential solutions and activities that should be implemented in order to solve the above-mentioned problems. On the other hand, the group that consists of unsolvable ethical problems will serve to highlight the dangers and risks that they entail, which affect researchers in this field, media workers who use AI, and finally the public itself that needs reporting.

MODELS OF AI

The phrase “Artificial Intelligence” refers to a wide spectrum of technological innovations, systems and computer software with an equally wide range of use. An AI system can be explained as “the branch of computer science that deals with designing intelligent computer systems that mimic human intelligence, e.g. visual perception, speech recognition, decision-making, and language translation.” (Sadiku et al., 2021, p. 15). The above-mentioned explanation, which we can also mark as a definition under certain conditions, describes AI in a way that corresponds properly to those models implemented in the media industry. This type of AI, according to the fundamental classification of AI based on capabilities – “narrow” (Artificial Narrow Intelligence), “general” (Artificial General Intelligence) and “super” (Artificial Super Intelligence) (Kumar, 2025) – belongs to the first group, “narrow AI”. Models that belong to this group can be described as “a goal-oriented version of AI designed to perform one specific task at a time, relying on information collected from various databases” (Babu & Banana, 2024, p. 210-211).

The main characteristic of narrow AI is that it performs only one type of tasks, without the capacity to exceed programming limitations, which allows for further classification into several subcategories, among which the most significant for this research is “Natural Language Processing” (NLP) (Saxena, 2024). As the function is explained by the name itself, NLP is the ability of the computer software to recognize and understand human

language in the form of speech and text (Gillis, 2024), thus determining it as the AI models most suitable for the use in the field of media. The most common form of NLP software is the so called “chatbot”, whose fundamental function is to fulfil assigned tasks by interpreting human language.

In most cases, the first association that comes with the mention of the term “chatbot”, or even the term “AI” in general, is the ChatGPT model. Such an association is utterly understandable and, to a considerable extent, justified, both in academic and everyday discourse, with a consideration of the current position and representation that ChatGPT has in modern society (Sidorenko-Bautista et al, 2024). ChatGPT is designed to reflect human conversation by generating responses according to language commands given by users (Atef, 2023). The incorporation of these tools significantly simplifies, speeds up and makes the work of journalists in the newsroom more efficient (Nicoud, 2024). The automation of various activities within media organizations, as well as the generation of a wide range of content in different formats (Bdoor & Habes, 2024, p. 564), gradually confirm the revolutionary potential of models like ChatGPT in the field of media.

THE USE OF AI IN NEWSROOMS

Most of journalists’ activities in newsrooms, where the practice of using AI is available, are related to the shaping of various types of media content. In this case, the term “use”, refers to the methods and purposes of applying AI across various stages of journalistic practice during work within a media organization. The use of AI in media organizations usually encompasses the following processes: (1) data collection; (2) content production; (3) content distribution; and (4) content consumption (Sančanin & Penjišević, 2022, p. 1). The application of AI has enabled numerous benefits such as simplifying and speeding up work, increasing efficiency and quality of content (Ivanovna & Meirzhanovna, 2020), while simultaneously generating various types of ethical problems and dilemmas (Filipović, 2024).

The aggregation of data needed to create certain content is the first step in the process of shaping and generating media content in which journalists can employ AI tools. By searching various sources and extracting large amounts of data from numerous databases, AI allows a journalist to get the necessary information for writing content (Vaishnavi, 2024). Potential ethical problems that may arise at this stage include situations in which databases from which data are retrieved contain unverified, outdated, biased or, in the worst case, completely incorrect information (Guzman & Lewis, 2020). The most significant contribution of the use of AI in the media lies in its ability to create various types of media content. The practice of using AI to convert previously accumulated data and then process them in accordance with the given topic in order to create content (Henestrosa, 2023) has been established in newsrooms for several years. However, modern AI, still not perfected to perform such tasks, can generate biased, discriminatory content (Chancellor et al., 2019; Nasser & Abu-Nasser, 2024; O’Neil, 2017;). In the worst-

case scenario, the created content can be completely imprecise, that is, untrue, which can result in the emergence of ethical problems on a wider scale.

The created content, after verification, needs to be published, which represents another stage where the use of AI can contribute. One of the essential elements of every media organization is finding and reaching the audience. In this context, produced content requires an optimization and personalization in accordance with the preferences of the audience to whom it is delivered (Swankar, 2024). Ethical problems that can occur during content distribution are, to the greatest extent, a reflection of the same problems that arise during the content production stage.

The final stage in this sequence refers to the monitoring of the consumption of media content. Once distributed content elicits or triggers specific audience reactions from the audience, especially when it addresses socially sensitive topics. The use of AI tools at this stage can help analyze the public's reactions to distributed content through the number of views or the general interest in a given topic. The analysis can thus contribute to a better understanding of the existing communication and information trends, as well as to the recognition of future ones (Sančanin & Penjišević, 2022).

FIRST GROUP – SOLVABLE ETHICAL PROBLEMS

Accountability

The essence of the problem of accountability can be explained by asking and answering the question "Who is responsible for the use of AI?" In the media industry, this problem relates to the most appropriate way of determining and dividing responsibilities regarding the use of AI in shaping media content. In academic circles, various approaches proposed by different authors to finding solutions to this problem can be found. According to the theory of deontology, in the relationship between the journalist and AI, the responsibility will always prevail on the side of the journalist, whose obligation will be to ensure the potential for making ethically correct decisions (Barcelo-Ugarte et al., 2021). On the other hand, in addition to journalists and editors, whose responsibility for the use of AI is indisputable, part of the responsibility extends to technical developers and content consumers. Companies that design AI models should take responsibility to a certain degree, taking into account the act that the way an AI model shapes certain content is directly conditioned and determined by its design and training process, which is determined by the developer (Fjeld et al., 2020; Novelli et al., 2024). When it comes to consumers that will review certain content, part of the responsibility takes the form of critical thinking and the tendency to check information, which remains an obligation regardless of whether AI was involved in shaping it (Ahmed, 2024).

Solving the problem of responsibilities can be achieved by fulfilling conditions for the creation and initiation of mechanisms that clearly define the level of responsibility assigned to each actor, followed by the establishment of a legal regulatory framework for

protection against damage caused in specific situations (Novelli et al., 2024). It is expected that journalists and editors in the media bear full legal responsibility. When it comes to developers and companies, their responsibility can be reflected in the obligation to provide official guarantees that AI models, if used responsibly, will not cause serious problems. The audience, as a constituent with the least influence, will therefore assume the lowest level of responsibility, which remains only on a moral level.

According to Cheong (2024), legal and regulatory frameworks established by governments are of the crucial importance for the permanent resolution the problem of accountability. Certain countries, recognizing the importance of adequate regulation of the responsible use of AI have introduced documents dealing with this issue. For example, at the beginning of 2023, the Government of the Republic of Serbia adopted a document entitled *Ethical guidelines for the development, application and use of reliable and responsible artificial intelligence* based on a similar document adopted in the previous year. As an example of a document that deals with the regulation of responsibility for the use of AI specifically in the media, latest version of the *Code of Journalists of Serbia* – can be cited, as it contains the norms and guidelines related to the responsibility of journalists when using AI. The Code, although imperfect, represents a significant step to solve the problem of AI accountability in the media. Furthermore, it would be very beneficial for all major journalistic organizations to introduce their own internal codes, regulating the responsible use of AI by their employees. If the regulation of the liability of the use of AI were to be viewed at the international level, a prominent example could be the *EU Artificial Intelligence Act* (AI Act) (European Union, 2024). Although the Act does not apply exclusively to the media industry, it clearly contributes to solving the problem of liability in the use of AI (European Union, 2024).

Content Transparency

By the phrase “content transparency”, we mean distinctively marked various media content and formats that are completely or partially generated by using AI. If the entire content is generated by AI then it should be clearly indicated to distinguish it from other content.

Due to some of its characteristics, the problem of content transparency is often associated with the problem of responsibility (Cheong, 2024). The essence of the transparency does not lie in the functioning of AI itself, but in those who use it. Journalists and editors decide on what it will look like and in what way certain content will be distributed. This directly determines whether and how synthetic content or parts of content will be marked so as to distinguish it from other media outputs. Recognizing the impact of AI technology on the production of synthetic content, the editors of the *Merriam-Webster Dictionary* designated the term “authentic” as the Word of the Year 2023, highlighting the importance of being able to recognize AI-generated content (Italie, 2023). The current situation of media respecting the ethics of content transparency can be characterized as ambivalent. On the one hand, there is concern that the audience will stop following the media if they know that content is produced with the assistance of AI, which discourages the media from fully embracing transparency practices. On the other

hand, the implications can be malign if the audience finds out that synthetic content has been published without adequate marking (Piasecky et al., 2024, p. 19).

The process of addressing the problem of content transparency is very similar to the process of resolving the accountability issues. The introduction of adequate regulations in the form of official government documents, as well as codes adopted by journalistic associations that apply across the entire media system, constitutes a foundational step for solving this problem. The aforementioned *EU Artificial Intelligence Act* contains a section that directly relates to ensuring content transparency (European Union, 2024). As a second example, *Code of Journalists of Serbia*, explicitly outlines the norms and guidelines, which mandate the labelling of synthetic content, are clearly stated. On page 8, two guidelines are particularly relevant: a) The use of artificial intelligence that significantly affects the processing of journalistic content must be clearly marked and must not mislead the public", and "c) In reporting on current events, the content must be marked so that it is clearly distinguished from synthetic (generated or significantly processed by artificial intelligence tools) from authentic content that was created by direct journalistic coverage of the event" (Udruženje novinara Srbije, 2024).

SECOND GROUP – PARTIALLY SOLVABLE ETHICAL PROBLEMS

Misuse

The problem of misuse represents the most dangerous and difficult form of violation of the ethical use of AI in media. A deliberate act of using AI in an incorrect way can lead to serious consequences affecting both the public that consumes media content and the media organizations that publish it. The misuse of AI in the media industry poses significant challenges to modern society at both the national and international levels (Pöhler et al., 2024).

The emergence of negative practice of abusing AI in order to produce content (Stilinic et al., 2025) that is inappropriate in a certain way due to possession of information and narratives that will negatively affect the audience, explains risks of using AI in the process of content production. Inaccurate narratives, as well as disinformation content created by generative AI can form a distorted picture of reality about relevant events (Nazar & Bustam, 2020). The misuse of generative AI also enables the creation of content with discriminatory, violent or racist connotations (Bender et al., 2021), the consumption of which can lead to trauma among more vulnerable segments of the audience, but also act as a motivator for the emergence and development of harmful ideologies (Egunjobi, 2025).

The most striking form of misuse of AI in content production is the emergence of a "deepfake", whose fundamental function is the creation of fake content, usually photos and videos, in which real people are represented doing or saying things they never actually did (Yasar et al., 2025). The delusion in which the audience finds itself is initiated by the broadcast of content in which public figures are presented in a completely different

context from the real one. This is a precise example of the greatest risk of deepfakes, which is that the information embedded in deepfake content is completely or mostly incorrect, while coming from reliable and credible sources (Yasar et al., 2025).

Our typology of ethical problems classifies the problems of AI misuse in the group of partially solvable problems for several reasons. If the problem of AI misuse is analyzed from a general perspective of all fields, it can be considered to be unsolvable. As is the case with other technologies, AI will always be misused by individuals, regardless of the level of regulation. When the focus is narrowed to the media sphere, the number of actors capable of misusing AI, namely turns out to be limited. By introducing adequate regulatory frameworks, such as internal codes or rulebooks within media organizations, as well as by introducing laws at the national level, it is possible to significantly reduce the rate of misuse. In addition, the secondary solution could be the launch of educational courses and seminars for media workers on the topic of adequate use of AI in work. The problem of misuse is thus partly solved by increasing awareness of the risks of improper use of AI. The insolvable part of this problem comes down to the constant presence of individuals who will knowingly misuse this tool for other goals and interests.

Job Loss

As was the case with the previous examples of ethical problems, the problem of job loss does not occur exclusively in the media industry; it is also present in other areas where AI is included in work processes.

In Till Leopold's research from the first months of this year, statistical data reveal that the rate of emergence of new jobs is higher than the rate of job loss (Leopold, 2025). According to a report of the by World Economic Forum (World Economic Forum, 2025), also from the beginning of this year, 170 million new jobs were created, while the number of job losses stood at 92 million. The interpretation of statistical data drawn from various reports and researches should be conducted with caution, otherwise misinterpretation can produce a distorted view of reality. For instance, the characteristics of new jobs require a much smaller number of people to be employed, compared to traditional jobs that require a much larger number of workers, and are rated for shutdown in the near future. The same research notes that the increased use of AI has resulted in the fact that 41% of global companies have decided to reduce their workforce (Leopold, 2025), heightened concerns, given that AI technology is still treated as a tool for improving productivity and work efficiency (Robinson, 2025).

If we focus primarily on the media sector, similar concerns are present, albeit with different directions. In media organizations, reduction or loss of journalists' jobs becoming increasingly common, due to the automation of the process of shaping and creating media content (McGowan, 2025). The problem of job losses in the media industry due to the advent of AI does not stop at reducing the number of employees or reducing their work tasks, but also goes a step further. What may represent a serious issue in the near future is the emergence of online platforms and websites, in which the management, as well as all activities related to the production and publication of content, is performed

by AI. For example, the first fully automated information portal, called “*AiBalkan.ai*”, was launched by the company JS Guru (JS Guru, n.d.). Furthermore, in November 2024 “*AiPiše - Serbian News Media*” was established as the first online news portal with an editorial staff composed exclusively of AI (Serbian News Media, 2024).

The problem of job loss, compared to all other solvable or partially solvable issues, possesses the status of most complex. One reason for this complexity is the fact that the decision to implement AI in the work of the media is not made by journalists, nor by the government, but by leaders and owners of corporations, which is the essence of the problem. A solution that would partially solve the problem of job losses can be based on the idea of introducing a regulation in the form of a law that will limit and regulate the use of AI in the media sector, thereby protecting the jobs of employed journalists. For example, the regulation of the use of AI can take place in the direction of determining the percentage, that is, a permitted number of publications in the production in which AI was used, where exceeding the permitted limit will be properly sanctioned. In this way the problem of job loss is reduced, or rather slowed down, while the remaining risk is the increase in the number of media and portals that are completely dependent on AI.

Privacy

Respect for the right to privacy is one of the fundamental principles on which modern journalistic reporting is based. Having the right to privacy entails “control over how much and what information about ourselves and others we will share with the public” (Kievčanin, 2022, p. 238). Violation of the right to privacy due to increasingly frequent tabloidization and sensationalism in the last couple of decades is a daily occurrence (Skoko & Bajs, 2007). In addition to journalistic disregard for ethical principles as the root cause, an increasingly frequent cause of privacy violations is the use of AI technologies in content processing. Publishing sensitive data or inappropriate photos when distributing content is the most common example of AI-related privacy violations in the media. A less noticeable, but equally risky form of privacy violation is also found in the data collection segment. AI systems can be used to collect sensitive data – such as health, financial or biometric information – without previous permission of their owners. Such data can also be used to train and improve AI algorithms (Gomstyn & Jonker, 2024). Direct interaction with ChatBots and AI algorithms on different platforms, in which users voluntarily provide personal data, is also a source of sensitive information that AI can later use. Research conducted by Lim and Shim showed the existence of an increasing rate of concern among respondents regarding the security of data voluntarily provided to AI during direct interactions, most often through correspondence (Lim & Shim, 2022). Collected data will thus be used to train algorithms that will later recommend a certain type of content based on these data, which can also violate user privacy.

As previously mentioned, the issue of privacy had been present in media reporting long before the introduction of AI, thus showing its status as a permanently insolvable. Focusing exclusively on AI, there remains space for reducing problems such as the introduction of prohibitions and restrictions on algorithms for the use of sensitive data, making such information about users inaccessible to AI. Furthermore, sensitive

data should be protected through access control measures. An unsolvable segment of this problem remains the practice of individuals voluntarily leaving their sensitive data behind when interacting with an AI chatbot. AI models thus, despite prohibitions on access to databases containing sensitive data, obtain information whose subsequent use may threaten the right to privacy. The solvable segment mainly relates to the way AI technology is implemented, where certain corrections can be made. The permanently unsolvable aspect remains the way AI works, in spite of the fact it can be limited to using certain databases. Once sensitive data is implemented, it becomes difficult to wean the model from its use.

THIRD GROUP – UNSOLVABLE ETHICAL PROBLEMS

Bias

Impartiality, as one of the fundamental principles of the journalistic profession, dates back to its beginnings. The conflict between the tendency for a neutral and objective approach to reporting on the one hand, and propaganda-biased fractions on the other hand, is present in all media forms. In the era of digital technologies, the conflict between objective and biased reporting takes on a new form that entails new challenges and perspectives (Gondwe, 2025).

The problem of bias in the media due to the use of AI technology affects to the greatest extent the process of content production in which, due to the application of generative AI, content that carries inaccurate, directed and biased narratives is created (Singh, 2024). The origin of the problem of bias can be found in large databases to which AI have access and collect the necessary information. The above-said databases can carry a lot of unverified, low-quality, outdated, but also completely incorrect data that are later used to produce content (Guzman & Lewis, 2020). Although this may not have been the original intention of the journalist, the result of the use of such data is the creation of content with a biased narrative (Singh, 2024, p. 134). As modern AI models still do not have a sufficiently developed consciousness and moral components, we cannot expect them to independently differentiate between correct, i.e. unbiased, and incorrect, i.e. biased content. That task is still in the domain of the work of journalists.

In addition to the process of content production, the problem of bias affects to a certain extent the process of its distribution. Systems whose function is to recommend certain content to users based on their previous activities and interests may do so with a certain amount of bias. AI algorithms that recommend content to users on various online portals or platforms can motivate the emergence of the so-called “filter bubbles” by deliberately adopting incorrect data, which will later result in proposing content that spreads prejudices, stereotypes and other forms of bias (Singh, 2024, p. 136). The risks of the emergence of social problems due to the presence of bias in the information process can reach proportions that involve law, legal and financial consequences (Jonker & Rogers, 2024).

In the group of solvable ethical problems, the root of the problem lies outside the AI system itself and it is directed towards the users. In the case of the problem of bias, we see a different outcome. The causes of the problem of bias we must search within the AI technology itself, in several specific sequences: (1) biased data used to train AI; (2) the way that model is designed in the beginning; (3) using proxies such as postal codes and addresses in exchange for ethnicity can result in discrimination; (4) failure to include all relevant data can result in the emergence of bias (Jonker & Rogers, 2024). These indicators imply the risk of inherent bias in AI systems that cannot be solved at the moment. Thus, the problem of bias remains intractable even in the AI-driven digital era.

Discrimination

Very similar to the problem of bias is the problem of discrimination. Under certain conditions it can be discussed to classify it as a subtype of bias. In this paper, the problem of discrimination will be presented as a separate problem, arising as a consequence of bias. For example, if a journalist wants to write an article about chauffeurs in city traffic, AI can offer a biased text that will only talk about male chauffeurs, thus raising gender stereotyping. On the other hand, the article also develops the problem of discrimination aimed towards female drivers. Some scholars such as Ferrer and colleagues believe that the problem does not necessarily lead to the emergence of discrimination (Ferrer et al, 2020). The resulting discrimination is a sequence of logical reasoning based on wrong information. Ferrer thus suggests that in some cases being biased is not only justified, but also necessary (Ferrer et al., 2020), which will enable the potential identification of certain hard-to-observe data (Guidotti et al., 2018; Mueller et al., 2019).

Content shaping for which AI can be used in the media is not its only capability. AI can be applied in segments of decision-making related to work organization and the allocation. The problem of discrimination can leave its traces in this domain as well, but not in its traditional forms such as religious, ethnic, racial, etc. In the case of media organizations, there would be the emergence of “industrial discrimination” which emphasizes and encourages disproportionate differences based on seemingly neutral rules (Altman, 2020).

The status of unsolvability, like that of the problem of bias, for the reasons already mentioned, also applies to get the problem of discrimination. The problem remains unsolvable, considering that its root is in the AI technology itself. To solve this problem, we believe that the implementation of newer, more perfect AI models is needed. Also, the logical assumption implies that if AI with more advanced features appear in the future, the problems of bias and discrimination will be solved simultaneously.

Transparency

One of the essential factors that increases the level of trust in the entire AI is the transparency of its work. Transparency can be explained as the possibility of insight into the way the AI system works, how decisions are made, which data are used, showing how certain results are arrived at, etc. (Wren, 2024). For the media industry, it is particularly

important to provide insight into the processes of content generation, as well as to display the databases used to collect the necessary information. Scholars describe AI systems as “black boxes” implying a lack of understanding, explanation, and regulation of how AI works (Jonker et al., 2024). The “black box” problem can have several different consequences which affect not only the media, but all areas where AI is applied. In addition to the discredit, as a basic consequence of the lack of transparency, the precision of the legal regulation of such systems is questionable (Kosinski, 2024). Also, there are problems on a purely ethical basis, such as data privacy or the risks of bias and discrimination towards sensitive structures (Kelly, 2025).

The problem of transparency thus falls into a situation of dead-end in which the only solution is to replace the “black box” model with models that allow insight into their work. Martens and Goethals (2024) recommend the use of simpler AI models, such as linear forecasting models, which allow certain degree of transparency, thereby reducing the problem to a small extent. However processing of large amounts of data, as well as the content generation process, require the use of more complex black box models. The current situation dictates that the problem of transparency is unsolvable, until, according to Kosinski, “white box” models whose work and decision-making systems will be transparent and understandable for users appear (Kosinski, 2024).

CONCLUSION

By analysing and classifying ethical problems of AI in the media into solvable and unsolvable ones, we have checked the initial hypothesis of this paper. The analysis of solvable and partially solvable problems points to a significant feature common to all problems of these two categories. The essence of the problem, and therefore its solution, lies outside the AI technology itself. The roots of the problems are in the users, i.e., journalists, rather than in the AI structure itself. Therefore, the ultimate solution to these problems should be sought primarily at the individual level, then in the workplace organization, and finally in the official state authorities, whose adequate response may result in resolving these issues. Efforts to solve these problems are visible on all three levels, primarily through the introduction of AI regulation documents in the media, which extends to the level of the media system of a country.

The unsolvable ethical problems also possess one common characteristics of major importance. The essence of all the unsolvable problems of the use of AI in the media lies within the technology itself. The way algorithms and software is designed and programmed to work determines the cause of content that has discriminatory and biased narratives or disrespects people’s privacy. To solve these problems, we should wait for the appearance of more advanced and better quality models.

The theoretical analysis established that a universal solution encompassing all problems unfortunately does not exist and is unlikely to be discovered in the near future. What can be done and what we strongly recommend is the education of media workers

and the media audience about the effects and risks of using the AI technology to create information with the emphasis on intractable problems. The findings of this paper can contribute to more precise directions for future research addressing each group of these problems individually. Research that would cover solvable problems can test their character of solvability outside the field of media. Research on a group of partially solvable problems can observe each problem individually and more precisely determine to what extent they are solvable. Further research in this area, among other things, can be focused on the analysis of prevention possibilities and ways of reducing negative effects of intractable AI problems, both in the media and in the general public.

References

- >Ahmed, M., (2024, November). Opinion: Navigating Accountability, the Roles and Risks of Using AI in Journalism. *Red Line Project*. <https://redlineproject.news/2024/11/09/opinion-navigating-accountability-the-roles-and-risks-of-using-ai-in-journalism/>
- >Altman, A., (2020, February). Discrimination. *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/discrimination/>
- >Atef, T. (2023, April). How does ChatGPT work?. <https://medium.com/@talal.atef.801/how-does-work-chatgpt-d50e926b26c4>
- >Babu, M. V. S., Banana, K. R. I. S. H. N. A. (2024). A study on narrow artificial intelligence—An overview. *International Journal of Engineering Science and Advanced Technology (IJESAT)*, 24 4, 210–219.
- >Barceló-Ugarte, T., Pérez-Tornero, J. M., & Vila-Fumàs, P. (2021). Ethical Challenges in Incorporating Artificial Intelligence into Newsrooms. In *María Luengo, Susana Herrera-Damas News Media Innovation Reconsidered*. (pp.138–153). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119706519.ch9>
- >Bdoor, S. Y. & Habes, M. (2024). Use chat GPT in media content production digital newsrooms perspective. Artificial intelligence in education: The power and dangers of ChatGPT in the classroom. *Artificial Intelligence in Education: The Power and Dangers of ChatGPT in the Classroom*. Springer Nature Switzerland, 545–561. https://doi.org/10.1007/978-3-031-52280-2_34.
- >Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 610–623. <https://doi.org/10.1145/3442188.3445922>
- >Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C. H. (2019). Artificial intelligence and journalism. *Journalism & mass communication quarterly*, 96(3), 673–695. <https://doi.org/10.1177/1077699019859901>
- >Chancellor, S., Birnbaum, M. L., Caine, E. D., Silenzio, V. M., & De Choudhury, M. (2019). A taxonomy of ethical tensions in inferring mental health states from social media. *Proceedings of the conference on fairness, accountability, and transparency*, 79–88. <https://doi.org/10.1145/3287560.3287587>
- >Cheong, B. C. (2024) Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, 6, 1–11. <https://doi.org/10.3389/fhumd.2024.1421273>
- >de-Lima-Santos, M. F., & Ceron, W. (2021). Artificial intelligence in news media: current perceptions and future outlook. *Journalism and media*, 3(1), 13–26. <https://doi.org/10.3390/journalmedia3010002>
- >Egunjobi, J. P. (2025). The Misuse of AI-Generated Content in Academic and Religious Settings. *International Journal of Research and Scientific Innovation XII*, no. XV, 871–879.
- >European Union (2024) *Artificial Intelligence Act*. <https://artificialintelligenceact.eu/>
- >Ferrer, X., Van Nuenen, T., Such, J. M., Coté, M., & Criado, N. (2020). Bias and discrimination in AI: a cross-disciplinary perspective. *arXiv preprint arXiv:2008.07309*. <https://doi.org/10.48550/arXiv.2008.07309>

- > Filipović, I. (2024). The usage of artificial intelligence in media: benefits, disadvantages, and ethical concerns. *Media Studies and Applied Ethics*, 5(1), 121-133. <https://doi.org/10.46630/msae.1.2024.09>
- > Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Research Publication*, (2020-1). <https://dx.doi.org/10.2139/ssrn.3518482>
- > Gillis, A., (2024, August). *What is natural language processing (NLP)?*. <https://www.techtarget.com/searchenterpriseai/definition/natural-language-processing-NLP>
- > Gomstyn, A., & Jonker, A., (2024, September). *Exploring privacy issues in the age of AI*. <https://www.ibm.com/think/insights/ai-privacy>
- > Gondwe, G. (2025). Is AI Bias in Journalism Inherently Bad? Relationship Between Bias, Objectivity, and Meaning in the Age of Artificial Intelligence. *Harvard-Berkman Klein Center*. 1-6.
- > Gorman, B., (2025, January). Different Types of Artificial Intelligence (AI) You Need to Know. <https://www.avast.com/c-types-of-ai-explained>
- > Gruetzemacher, R., & Whittlestone, J. (2022). The transformative potential of artificial intelligence. *Futures*, 135, 1-11. <https://doi.org/10.1016/j.futures.2021.102884>
- > Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5), 1-42. <https://doi.org/10.1145/3236009>
- > Guzman, A., L., & Lewis, S., C., (2020). Artificial Intelligence and Communication: A Human-Machine Communication Research Agenda. *New Media & Society* 22, 70-86. <https://doi.org/10.1177/1461444819858691>
- > Henestrosa, A. L., Greving, H., & Kimmerle, J. (2023). Automated journalism: The effects of AI authorship and evaluative information on the perception of a science journalism article. *Computers in Human Behavior*, Volume 138, 1-56. <https://doi.org/10.1016/j.chb.2022.107445>
- > Huang, C., Zhang, Z., Mao B., & Yao, X (2023). An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence*, 4(4), 799-819. <https://doi.org/10.1109/TAI.2022.3194503>
- > Hursthouse, R. (2000). *On Virtue Ethics*. Oxford University press.
- > Italie, L., (2023, November). What's Merriam-Webster's word of the year for 2023? Hint: Be true to yourself. <https://apnews.com/article/merriam-webster-word-of-year-2023-a9fea610cb32ed913bc15533acab71cc>
- > Ivanovna, S. A., & Meirzhanovna, A. S. (2020). AI in journalism: Benefits of using vs ethical issues. *The 1st International scientific and practical conference "European scientific discussions"* Potere della ragione Editore, 491-498.
- > Jonker, A., Gomstyn, A., & McGrath, A., (2024, September). *What is AI transparency?*. <https://www.ibm.com/think/topics/ai-transparency>
- > Jonker, A., & Rogers, J., (2024, September). *What is algorithmic bias?*. <https://www.ibm.com/think/topics/algorithmic-bias>
- > JS Guru. *AI Balkan*. <https://www.aibalkan.ai/about/>
- > Kelly, J., I., (2025, February). *What Is Black Box AI?*. <https://www.invo.ca/blog/what-is-black-box-ai>
- > Kijevčanin, R. (2022). Pravo na privatnost u medijima. In *Zbornik radova: XXI vek - vek usluga i služnog prava* (pp. 231-244). <http://dx.doi.org/10.46793/XXlv-13.231K>
- > Kosinski, M., (2024, October). *What is black box artificial intelligence (AI)?*. <https://www.ibm.com/think/topics/black-box-ai>
- > Kumar, A., (2025, August). *Types of AI Explained*. <https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/types-of-artificial-intelligence>
- > Leopold, T. (2025, January). *Future of Jobs Report 2025: The jobs of the future – and the skills you need to get them*. <https://www.weforum.org/stories/2025/01/future-of-jobs-report-2025-jobs-of-the-future-and-the-skills-you-need-to-get-them/>
- > Lim, S., & Shim, H. (2022). No secrets between the two of us: Privacy concerns over using AI agents. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 16(4), Article 3, 1-18. <https://doi.org/10.5817/CP2022-4-3>

- >McGowan, C., (2025, May). *One day I overheard my boss saying: just put it in ChatGPT: the workers who lost their jobs to AI*. <https://www.theguardian.com/technology/2025/may/31/the-workers-who-lost-their-jobs-to-ai-chatgpt>
- >Mertens, D., & Goethals, S., (2024, December). *Opening the black box: how 'explainable AI' can help us understand how algorithms work*. <https://theconversation.com/opening-the-black-box-how-explainable-ai-can-help-us-understand-how-algorithms-work-244080>
- >Mueller, S. T., Hoffman, R. R., Clancey, W., Emrey, A., & Klein, G. (2019). Explanation in human-AI systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI, 1-204. *arXiv preprint arXiv:1902.01876*. <https://doi.org/10.48550/arXiv.1902.01876>
- >Nasim, S. F., Ali, M. R., & Kulsoom, U. (2022). Artificial intelligence incidents & ethics a narrative review. *International Journal of Technology Innovation and Management (IJTIM)*, 2(2), 52-64. <https://doi.org/10.54489/ijtim.v2i2.80>
- >Nasser, B. S. A. & Abu-Naser, S. S. (2024). Artificial Intelligence in Digital Media: Opportunities, Challenges, and Future Directions. *International Journal of Academic and Applied Research (IJAR)* ISSN: 2643-9603 Vol. 8 Issue 6, 1-10.
- >Nazar, S., & Bustam, M. R. (2020). Artificial intelligence and new level of fake news. *IOP conference series: Materials science and engineering*, 879(1), 1-7. <https://doi.org/10.1088/1757-899X/879/1/012006>
- >Nicoud, A., (2024, November). *Two years later: Taking stock of ChatGPT's impact on the news media*. <https://wan-ifra.org/2024/11/two-years-later-taking-stock-of-chatgpts-impact-on-the-news-media/>
- >Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: what it is and how it works. *AI & Society*, 39(4), 1871-1882. <https://doi.org/10.1007/s00146-023-01635-y>
- >O'neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- >Piasecki, S., Morosoli, S., Helberger, N., & Naudts, L. (2024). AI-generated journalism: Do the transparency provisions in the AI Act give news readers what they hope for?. *Internet Policy Review*, 13(4), 1-28. <https://doi.org/10.14763/2024.4.1810>
- >Pöhler, L., Schrader, V., Ladwein, A., & von Keller, F. (2024). A technological perspective on misuse of available AI. *arXiv preprint arXiv:2403.15325*. 1-13. <https://doi.org/10.48550/arXiv.2403.15325>
- >Robinson, B. (2025, February). *Fears About AI Job Loss: New Study Answers If They're Justified*. <https://www.forbes.com/sites/bryanrobinson/2025/02/09/fears-about-ai-job-loss-new-study-answers-if-theyre-justified/>
- >Sadiku, M. N., Ashaolu, T. J., Ajayi-Majebi, A., & Musa, S. M. (2021). Artificial intelligence in social media. *International Journal of Scientific Advances*, 2(1), 15-20. <https://doi.org/10.51542/ijscia.v2i1.4>
- >Sančanin, B., & Penjišević, A. (2022). Use of artificial intelligence for the generation of media content. *Social informatics journal*, 1(1), 1-7. <https://doi.org/10.58898/sij.v1i1.01-07>
- >Saxena, A. (2024, November). *Artificial Narrow Intelligence (ANI) – Definition, Challenges*. <https://www.appliedaicourse.com/blog/artificial-narrow-intelligence/>
- >Sidorenko-Bautista, P., Cabezas-Clavijo, Á., Cantón-Correa, J., & Alba-Ruiz, R. (2024). The Impact of ChatGPT on Journalism: Social Listening, Bibliographic Production, and Media Agenda. *BID*, (53) 1-18. <https://doi.org/10.1344/bid2024.53.06>
- >Singh, P. (2024). Media 2.0: A Journey through AI-Enhanced Communication and Content. *Media and AI: Navigating*, 127-149.
- >Skoko, B., & Bajš, D. (2007). Objavlivanje neistina i manipuliranje činjenicama u hrvatskim medijima i mogućnosti zaštite privatnosti, časti i ugleda. *Politička misao: časopis za politologiju*, 44(1), 93-116.
- >SerbianNewsMedia. (2024, November). AI Piše! Prvi srpski informativni portal sa AI redakcijom startuje 11. novembra!. <https://snm.rs/ai-pise-prvi-srpski-informativni-portal-sa-ai-redakcijom-startuje-11-novembra/>
- >Stilinovic, M., Bailo, F., & Hutchinson, J. (2025). Creative Underspheres and democratic challenges: Exploring the implications of generative AI misuse. *New Media & Society*, 1-17. <https://doi.org/10.1177/14614448251338511>

- >Stefanowicz, B., (2025). *What Is NLP Chatbot & How To Build One: Full Guide*. <https://www.tidio.com/blog/nlp-chatbots/>
- >Swarnkar, E. (2024, May). *How AI Helps in Content Distribution: 5 Effective Ways* . <https://www.waveon.io/en/blog/how-ai-helps-in-content-distribution-5-effective-ways>
- >Udruženje novinara Srbije. (2024). *Kodeks novinara Srbije*. Savet za štampu.
- >Vaishnavi, Y., (2024, November). *Data Aggregation in Big Data and AI: What You Need to Know in 2025*. <https://medium.com/@vaishnaviyada/data-aggregation-in-big-data-and-ai-what-you-need-to-know-in-2025-d0229e439fe9>
- >Verma, S. (2019). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. *Vikalpa*, 44(2), 97-98. <https://doi.org/10.1177/0256090919853933>
- >World Economic Forum. (2025). *Future of Jobs Report*. https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf
- >Wren, H., (2024). *What is AI transparency? A comprehensive guide*. <https://www.zendesk.com/blog/ai-transparency/>
- >Yasar, K., Barney, N., & Wigmore, I., (2025, May). *What is deepfake technology?*. <https://www.techtarget.com/whatis/definition/deepfake>
- >Zhang, B., & Dafoe, A., (2019) Artificial Intelligence: American Attitudes and Trends. *Future of Humanity Research*, University of Oxford. <http://dx.doi.org/10.2139/ssrn.3312874>

IZMEĐU SCILE I HARIBDE: KATEGORIZACIJA ETIČKIH IZAZOVA POVEZANIH S UPORABOM UMJETNE INTELIGENCIJE U MEDIJIMA

Aleksa Mitić

SAŽETAK *Povećana stopa uporabe umjetne inteligencije (UI) unutar medijske industrije dovela je do pojave novih etičkih problema, kao i do širenja postojećih. Uzimajući u obzir različite čimbenike i aktere, pronalaženje jedinstvenog rješenja za te etičke probleme bio bi iznimno složen, a vjerojatno i nemoguć pothvat. Opća hipoteza ovog rada jest da se aktualni etički problemi povezani s uporabom umjetne inteligencije u medijima mogu razdvojiti u dvije glavne kategorije: rješivi i nerješivi problemi. Analiza relevantne literature obuhvatit će najistaknutije etičke probleme te omogućiti njihovu klasifikaciju prema tim dvjema kategorijama. Primarni cilj ovog rada jest prikazati koja bi od postojećih etičkih pitanja, koja se smatraju rješivima, trebala imati prioritet u budućim istraživanjima kako bi se naposljetku pronašlo rješenje. Skupinu nerješivih pitanja ne bi trebalo u potpunosti zanemariti, no predlaže se smanjenje učestalosti njihova istraživanja, u skladu s razvojem sofisticiranijih modela UI-ja koji bi mogli biti prikladniji za etičku integraciju u medijske prakse.*

KLJUČNE RIJEČI

UMJETNA INTELIGENCIJA, MEDIJI, ETIKA, ETIČKI PROBLEMI, NOVINARSTVO

Bilješka o autoru _____

Aleksa Mitić :: Filozofski fakultet, Sveučilište u Nišu :: miticvaleksa@gmail.com