AI and the Quest of Trustworthy Information

Bokai Lai

Faculty of Science, University of British Columbia, BC, CA

jasolai@student.ubc.ca

Abstract. Artificial Intelligence (AI) is revolutionizing the sectors of medicine and media by improving productivity, precision, and customization. AI is revolutionizing the development, distribution, and consumption of material in the media while also being essential in information verification and moderation. AI in medicine enhances clinical processes, personalizes patient care, and increases diagnostic accuracy. But these developments also bring with them new difficulties, including preserving media originality and guaranteeing the accuracy and openness of AI-generated insights in the medical field. In order to comprehend their impact on content creation and information distribution, this article investigates the approaches and uses of AI, such as Generative Adversarial Networks (GANs) and the Information Adoption Model (IAM). In media, AI automates video editing, content moderation, and tailored recommendations, while in healthcare, it enhances diagnostic accuracy, personalizes care, and simplifies clinical operations. The article also discusses issues like biases, the necessity of human oversight, and data privacy, highlighting how crucial it is to create AI systems that are both morally and practically sound.

Keywords: Artificial Intelligence, media, healthcare, content generation, diagnostic accuracy.

1. Introduction

Artificial intelligence (AI), has become a disruptive force in the ever changing field of digital technology, especially when it comes to the sharing of information. The importance of AI algorithms in producing, publishing, and disseminating news has expanded dramatically with their rising sophistication. But this progress also poses a serious problem: making sure that the data that AI systems offer is reliable. Studying the relationship between AI and reliable information is crucial because it affects both the public's foundational faith in the media and the integrity of media.

AI is being used in the journalism industry in a variety of ways. AI holds great potential to improve journalism's efficiency and accuracy, ranging from the identification of misinformation to automated news authoring and personalized content distribution. Natural language processing and machine learning algorithms, for example, are AI-driven technologies that are used to examine large datasets, find patterns, and even forecast future occurrences. Notwithstanding these advantages, there are worries about prejudice, false information, and the degradation of journalistic standards due to the dependence on AI.

In preserving the veracity of information, research in this area has emphasized both the benefits and drawbacks of artificial intelligence. Research has demonstrated that although artificial intelligence can greatly expedite the news production process and help with fact-checking, it can also reinforce biases

found in training data and be exploited to disseminate misleading narratives. This paradox emphasizes how crucial it is to create reliable AI systems that give transparency and ethical issues first priority.

Despite its impact on media information, Artificial Intelligence is also becoming a pervasive force that is transforming society, industry, and the economy by altering interactions among stakeholders and citizens [1]. Though AI has many advantages, there are worries over the reliability of AI systems due to its broad use. In addition to efforts by businesses and academic organizations to provide standards and toolkits for assessing AI trustworthiness, this has led to a number of legislative policies and principles targeted at supporting trustworthy AI. AI has great potential in areas like medicine, as demonstrated by natural language processing models like ChatGPT. However, a study including 33 doctors from 17 different disciplines found that while ChatGPT typically provided thorough and accurate responses to medical inquiries, there were some noticeable drawbacks, particularly for complicated issues, underscoring the need for more investigation and improvement [2]. AI is being utilized more and more to support human experts in high-stakes situations while they are making decisions. This allows AI and humans to work together to maximize results. Research indicates that elements like confidence scores can help in the calibration of human trust in AI, which is necessary for successful AI-assisted decision-making [3]. However, in order to effectively complement AI's skills with human knowledge, decision-making also depends on people, which highlights the need for novel methods to AI explainability [3].

This essay delves into the quest for trustworthy information in the age of AI, examining the challenges and opportunities that lie ahead. By exploring current research and case studies, the paper aims to shed light on how AI can be both a tool for enhancing information reliability and a potential risk if not properly managed. The discussion will highlight the need for interdisciplinary approaches, combining technological innovation with ethical frameworks, to ensure that AI serves the public good in the quest for truthful and reliable information.

2. Methodology

This section explores two Artificial Intelligence Generated Content (AIGC) methods: enerative Adversarial Networks (GANs) for controlled content generation and Information Adoption Model (IAM) for adopting hidden information in media.

2.1. AIGC based on GAN

AIGC technology relies heavily on GANs, especially when it comes to producing realistic media content. In this methodology, generative information hiding inside the content generation process is facilitated by using a GAN based approach [4]. The generator and discriminator neural networks, which are the two neural networks that make up the GAN model, compete with one another to improve each other's performance over time.

The generator's job is to turn random noise into believable media outputs in order to create synthetic material that seems like actual data [4]. On the other hand, the discriminator evaluates these outputs in comparison to actual data in an effort to differentiate between produced and authentic material [4]. The generator's outputs can be gradually improved and made more and more similar to genuine material thanks to the iterative feedback loop between these two networks.

This work uses Conditional GANs, which enable the development of content based on particular criteria or limitations, to improve the usefulness of traditional GANs in content creation. Conditional parameters that direct the development process are incorporated into the GAN architecture to enable more targeted and controlled content generation [4]. Furthermore, a pix2pix architecture is employed, in which the generator generates images that conform to particular contours or limitations and the discriminator verifies the authenticity of these generated images [4].

2.2. AIGC based on IAM

IAM, developed by Sussman and Siegel, takes ideas from the dual-path model and the Technology Acceptance Model and combines them to explain how people adopt information and how it influences their decisions. The model views the process of information influencing people as an adoption process

in which people evaluate the information they are given and determine whether or not to act on it. The four main variables in the IAM are argument quality, source credibility, information usefulness, and information adoption. integrates concepts from the dual-path model and the Technology Acceptance Model to explain how individuals adopt information and how it influences their decisions [5]. According to this model, the process of information influencing people is viewed as one of adoption, where individuals assess and decide whether to act on the information they receive. IAM incorporates four key variables: argument quality, source credibility, information usefulness, and information adoption.

Information usefulness acts as a mediating variable, bridging the gap between these independent factors and the final decision-making process. Information that is perceived as useful is more likely to influence the user positively. Finally, information adoption is the dependent variable, representing the decision of whether to accept or act upon the information provided. Argument quality and source credibility serve as the independent variables. These factors reflect the user's assessment of the logical strength of the content and the credibility of the infromation's source [5].

In addition, the IAM integrates the central path, which emphasizes argument quality, and the peripheral path, which focuses on source credibility. This means that it looks at how people assess information from both perspectives to decide whether or not to adopt it [5]. Because it emphasizes the importance of both content quality and source credibility, the IAM provides a thorough framework for comprehending how information influences user behavior and decision making.

3. Application of AI in real life

AI is transforming media and healthcare by revolutionizing content creation and delivery while enhancing the personalization and efficiency of services in both fields.

3.1. Application of AI in Media area

There have been notable developments in the fields of content creation, delivery, and consumption as a result of the incorporation of AI into the media industry. These developments have improved media production quality and efficiency while opening up new avenues for customization, establishing credibility, and creative inquiry.

With the advent of sophisticated generative models, artificial intelligence's contribution to video production and editing has gained significant traction. These artificial intelligence systems are able to create visually appealing and cohesive new video material by analyzing large datasets and identifying patterns [4]. Tasks that formerly required a lot of manual labor, such editing, applying special effects, and creating subtitles, are now automated [4]. AI greatly simplifies the video creation process, opening it up to a wider spectrum of artists by identifying keyframes, evaluating sequences, and making wise editing choices. The end effect is a democratization of video production, enabling a larger audience and more effective creation of high-caliber material [4].

Transparency and reliability in AI-generated content are becoming increasingly important as AI gets more integrated into media creation. AI technologies are being utilized more and more to confirm the veracity of information, guaranteeing that it is accurate and devoid of false information [3]. In journalism, where news content trustworthiness is crucial, this is especially crucial. AI has the ability to cross-reference data from several sources, identify discrepancies, and offer justifications that increase the content's credibility [3]. AI can also support content makers by providing trustworthy sources and real-time fact-checking, which helps to preserve the integrity of media material.

The method that content is distributed to viewers has also changed as a result of AI's capacity to customize media consumption experiences. AI is able to customize content recommendations to individual preferences by evaluating user data, including viewing history, preferences, and interaction patterns [1]. AI is used in streaming services to curate playlists, recommend videos, and modify content offers based on the individual interests of each user. This strategy helps media companies maintain their viewers by improving the entire viewing experience, in addition to increasing user engagement through the provision of content that speaks to specific interests [1].

Content moderation is a critical use of AI in the media industry. The growth of user-generated content on social media platforms has made manual regulation and monitoring more difficult [1]. AI-powered content moderation technologies can automatically identify and eliminate offensive, dangerous, or unlawful content, preserving platform security and regulatory compliance. These solutions take proactive measures to stop the spread of harmful content, such as hate speech, misinformation, or graphic violence, by using advanced machine learning algorithms to discover patterns linked to such content [1]. Even if bias and accuracy remain issues, continuing advances in AI are making content moderation systems more equitable and efficient.

AI is also having a big impact on how creative the media industry is. Applications powered by AI, such generative models, can now compose music, write scripts, design images, and even make full films. AI frees up content creators to concentrate more on the creative components of their work by automating the tedious and time-consuming parts of the process. This leads to the creation of distinctive and inventive media that pushes the limits of conventional creative processes, as well as democratizing content creation and making it more accessible to a wider audience.

In summary, artificial intelligence is being applied in the media industry in a variety of ways, impacting everything from content generation and personalization to content moderation and trust-building. AI technologies will probably have a greater influence on the media environment as they develop, spurring more creativity and changing the ways in which media material is created, shared, and enjoyed. The continued development of AI portends a time when media will be more individualized, trustworthy, creatively rich, and efficient in addition to being more diverse and diverse.

3.2. Application of AI in Medical area

By improving diagnostic precision, tailoring patient care, and streamlining clinical procedures, artificial intelligence is revolutionizing the medical industry. The potential of artificial intelligence, fueled by sophisticated algorithms and machine learning models, is being used in a variety of medical fields to enhance patient outcomes and lessen the workload of medical practitioners.

AI systems' very efficient analysis of medical data and visuals has significantly enhanced diagnostic capabilities. Subtle patterns that a human interpreter would miss, including early warning indications of cancers or fractures, can be found by machine learning algorithms [6]. AI is also very good at digesting complicated data, such as genetic sequences, to determine illness risk. Furthermore, clinical decision support systems (CDS) rely heavily on AI. AI is able to give physicians evidence-based recommendations that are customized to each patient's unique needs by real-time analysis of patient data [6]. By lowering treatment approach variability and guaranteeing that clinical decisions are in conformity with the most recent medical guidelines, these technologies aid in standardizing care. AI-driven CDS technologies, for instance, can quickly evaluate patient data in emergency situations to identify patients who may experience difficulties [6]. This allows for the timely and appropriate implementation of therapies. AI is being used in this field to handle treatment result discrepancies and close quality gaps in healthcare. Notwithstanding these advantages, artificial intelligence occasionally produces outcomes that want human validation to guarantee their pertinence and precision. This emphasizes how crucial it is to strike a balance between using AI's skills and consulting with physicians to validate the results because AI's judgments aren't always conclusive.

Personalizing patient care is one of AI's most revolutionary contributions to medicine. To create individualized treatment regimens, AI algorithms can examine a wide range of data, including as a patient's genetic profile, lifestyle, medical history, and even socioeconomic determinants of health [2]. This degree of customization is especially helpful in the management of chronic illnesses, where successful therapy must be customized to each patient's specific demands. For instance, by continuously monitoring blood glucose levels and forecasting the body's reaction to various activities and food consumption, artificial intelligence can assist in the creation of individualized insulin dosage programs for diabetes control [2].

Moreover, AI is a vital tool for individualized patient care due to its ability to learn and adapt over time. AI systems can improve their suggestions as more data from specific patients is gathered,

guaranteeing that the best course of care is continued even as situations change [2]. These systems, however, can struggle to properly comprehend the complexities of clinical circumstances, which might result in either an overabundance of warnings or the overlooking of crucial messages. It is still important to continue improving AI in this field to make sure that healthcare workflows are efficient and free from needless delays or mistakes.

AI is being utilized to enhance healthcare operations in addition to customizing care and increasing diagnostic accuracy. Healthcare workers' administrative workload can be greatly decreased by automating repetitive operations like data input, scheduling, and billing through the integration of AI into electronic health record (EHR) systems [7]. By freeing up doctors to concentrate more on patient care rather than administrative tasks, automation increases the effectiveness of healthcare delivery as a whole.

In therapeutic settings, AI plays a key role in mitigating the "alert fatigue" problem. AI is able to sift through notifications based on relevance and timing, delivering only important alerts to the doctor. By doing this, healthcare workers are better able to respond to critical alerts and experience a reduction in cognitive burden from constantly juggling multiple, frequently irrelevant signals [7]. Moreover, AI-powered solutions can adapt the alerting system to each user's unique preferences and workflow by continuously learning from the interactions between the clinician and the tools [7]. But occasionally, these systems can't completely comprehend the complexities of clinical circumstances, which might result in either too many warnings or the overlooking of important messages [7]. To guarantee that healthcare operations are optimized without needless distractions or oversights, the advancement of AI in this field continues to be a top focus.

For AI-generated medical advice to be successfully integrated into healthcare, trust is essential. Transparency and explainability are becoming more and more important design considerations for AI systems, which benefits patient and doctor trust. AI technologies facilitate the understanding of the reasoning behind certain treatment recommendations by healthcare professionals and make it easier for them to integrate AI-generated insights into their decision-making processes by offering explicit explanations for their recommendations [5]. In high-stakes settings where decisions have serious repercussions, like surgery or intensive care, this transparency is especially crucial.

Lastly, by providing objective and consistent suggestions, AI has the potential to improve the dependability of therapeutic decisions. This holds particular significance in mitigating healthcare inequities, as human decision-making biases may result in uneven treatment outcomes. AI helps guarantee that all patients receive the same quality of care, irrespective of their background or the subjective prejudices of individual practitioners, by offering standardized recommendations based on large datasets [5].

In summary, AI has many applications in the medical profession and is still developing quickly. AI is positioned to take on a more significant role in healthcare, from boosting patient care personalization and diagnostic accuracy to streamlining clinical operations and boosting confidence in medical judgments. However, alongside these developments, there are still issues with the consistency of data produced by AI, and its successful incorporation into clinical practice requires close monitoring and validation.

4. Challenges and Further Expectations

Even though AI is revolutionizing the media and medical industries, a number of obstacles prevent it from reaching its full potential. Keeping automation and creativity in check is the main challenge facing the media industry. Although AI-driven technologies, including those built on diffusion models and GANs, improve content production and personalization, worries about the uniformity and the loss of human originality in media material are intensifying. Furthermore, strong legal frameworks and improved transparency procedures are required to address the serious ethical issues raised by AI-generated material, notably in areas like misinformation and deepfakes [8].

The difficulties are no less complicated in the medical field. Though encouraging, the use of AI in decision-support and diagnostic systems prompts questions regarding the dependability, accuracy, and

bias of insights produced by AI. There is a fundamental need for continuing validation and refining of these systems to ensure that they complement rather than replace human expertise, as illustrated by the variable levels of acceptability and usefulness of proposals provided by artificial intelligence [9]. Furthermore, because AI models require enormous volumes of sensitive patient data, data privacy and security are critical issues in healthcare AI applications that cannot be disregarded. Strict protective measures are therefore necessary.

Going forward, it is probable that additional developments in AI will concentrate on enhancing the interpretability and explainability of AI systems in both domains, guaranteeing that outputs produced by AI are comprehensible and transparent to users. To maintain originality while utilizing AI's skills in the media, it will be crucial to promote a collaborative approach between AI and human artists [10]. Realizing the full potential of AI in medicine to improve patient care and medical decision-making will require ongoing efforts to decrease bias, improve data security, and better AI integration with current clinical workflows.

5. Conclusion

In conclusion, this essay looked at the methods and uses of AI in the media and medical domains, examining how AI tools like GANs and IAM are influencing the production of content and the practice of medicine. It also talked on the difficulties in implementing AI, such as bias, transparency, and the requirement for human supervision. Without a doubt, the application of AI in the media and healthcare industries is a turning point in the development of technology. The potential of AI to automate processes, customize user experiences, and give decision-support has completely changed these sectors and opened them up to new possibilities for creativity and efficiency. AI in media facilitates content control, improves the creative process, and guarantees more individualized customer experiences. AI, on the other hand, improves diagnostic tools, simplifies workflows, and enables more individualized patient care in the medical field.

Though AI still has a lot to offer, it's critical to acknowledge that human oversight is still necessary to preserve accuracy, innovation, and trust. The results of AI need to be regularly assessed, verified, and improved, especially in high-stakes fields like healthcare where patient outcomes and safety are at risk. Similar to this, in the media, careful management of the interaction between human creativity and AI's automated processes is necessary to prevent content uniformity.

As AI technologies develop, encouraging cooperation between AI systems and human specialists will be essential to optimizing AI's potential and guaranteeing that its results are visible, dependable, and morally sound. In the end, this smooth integration will be crucial to the future of AI in various domains, guaranteeing that both sectors continue to gain from the distinct advantages of AI and human intellect.

References

[1]   Mentzas, G., Fikardos, M., Lepenioti, K., and Apostolou, D. (2024). Exploring the landscape of trustworthy artificial intelligence: status and challenges. Intell. Decis. Technol. 18 837–854. https://doi.org/10.3233/IDT-240366.

[2]   Johnson, D., et al (2023). Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the Chat-GPT model. Research Square rs.3.rs-2566942 Preprint. https://doi.org/10.21203/rs.3.rs-2566942/v1.

[3]   Zhang, Y., Liao, Q. V., and Bellamy, R. K. E. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. Proc. 2020 Conf. on Fairness, Accountability, and Transparency (FAT '20)*, Assoc. for Computing Machinery, New York, USA, 295–305. https://doi.org/10.1145/3351095.3372852.

[4]   Di, J. (2024). Principles of AIGC technology and its application in new media micro-video creation. Applied Mathematics and Nonlinear Sciences 9 https://doi.org/10.2478/amns-2024-1393.

[5]   Yu, J. (2024). Research on the influencing factors of user adoption of artificial intelligence generated content (AIGC). Shandong Normal University. https://doi.org/10.27280/d.cnki. gsdsu.2024.001410.

[6]   Vaira, L. A., et al (2023). Accuracy of ChatGPT-generated information on head and neck and oromaxillofacial surgery: a multicenter collaborative analysis. Otolaryngol. https://doi.org/10. 1002/ohn.489.

[7]   Liu, S., Wright, A. P., Patterson, B. L., Wanderer, J. P., Turer, R. W., Nelson, S. D., McCoy, A. B., Sittig, D. F., and Wright, A. (2023). Using AI-generated suggestions from ChatGPT to optimize clinical decision support. J. Am. Med. Inform. Assoc. 30 1237–1245. https://doi.org/ 10.1093/jamia/ocad072.

[8]   Li, T., Shen, K. X., and Fan, S. M. (2024). Brief Analysis of the Impact of a New Generation of Artificial Intelligence on the News Media Industry -- Taking ChatGPT as an Example. China Media Technology (06) 63–67. https://doi.org/10.19483/j.cnki.11-4653/n.2024.06.012.

[9]   Ma, S. J. (2023). Analysis on the Application and Impact of Generative Artificial Intelligence Service in the News Industry. Chinese city newspaper (10) 53–54. https://doi.org/10.16763/j. cnki.1007-4643.2023.10.022.

[10]   Chen, Y. H. and Li, J. (2023). Research on the Identification Risk and Regulation of the Identity of the Content File of Artificial Intelligence: Thinking About the Content Generated by ChatGPT. Research on Archival Science (05) 4–12. https://doi.org/10.16065/j.cnki.issn1002-1620.2023.05.001.