AI in Journalism: The Ethics and Usages of Artificial Intelligence for Facts And Statistics within Journalism

by

Shannon Dooley

"Humanoid Robot in Investigative Journalism." Created March 22, 2024. Digital image. Generated by OpenAI's DALL-E.

University of Pennsylvania

Professor Taransky's Writ 3090

Writing in AI

March 2024

Introduction

The incorporation of Artificial Intelligence into almost all societal functions has also led to its integration in journalism. Journalism currently faces significant challenges posed by the AI industry, particularly affecting its practices and ethics. While AI offers significant benefits for data analysis and content generation, it also introduces complex ethical challenges. AI systems are driven by data and algorithms, which can be inherently biased and, as a result, spread misinformation. A recent study, in 2023, states that AI algorithms use a variety of extensive computational machine learning models for providing data using standard statistical methods[1]. Industry experts suggest that within some journalistic practices such as back end automation, distribution and recommendations, coding and product development, the influence of AI is above 50 percent[2].

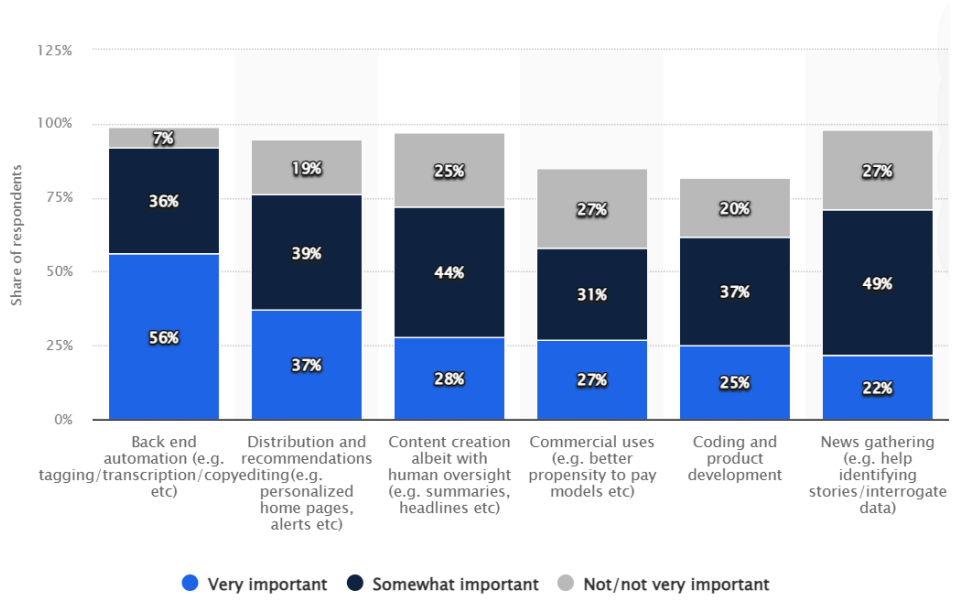

Figure 2. Survey of a group of 314 individuals, which is 56% of the leaders in the journalism industry, responds to how important they think AI is going to be in different tasks of the newsroom in 2024.[3]

Within the newsroom, this presents ethical issues of biases and strategies. The evaluation of AI practices is crucial in order to maintain their trustworthiness to the general public. This white paper delves into these challenges, drawing from a range of scholarly articles and industry reports to propose a multi-faceted approach to journalists in addressing AI bias and misinformation within journalism.

The Problem of AI Bias in Journalism

[1]Nazer, "*Bias in artificial intelligence*"
[2]WAN-IFRA, "*Most important newsroom*"
[3]WAN-IFRA, "*Most important newsroom*"

AI bias in journalism stems from the data used to train AI algorithms. These datasets can reflect already existing societal biases, depending on the collection method. Originally, these models were developed from small datasets due to limited funding and ease of use for scientific study, which underscores the limitations and inherent biases of early AI models[4]. These datasets in general were not representative of a broader population, and instead focused on majority groups. For example, in healthcare, the data collection was disproportionate to all demographic groups, leading to this data to be ingrained into the algorithms themselves and is still commonly used in practice today.[5] [6] This historical reliance on constrained datasets has propagated biases, especially in journalism, where they can significantly influence public perception and exacerbate the pre existing biases already in the fabric of society. In the news room specifically, AI algorithms accomplish a couple things for computational journalists:

- Prioritization
- Classification
- Association
- Filtering

Inherently, this will ease tasks and increase productivity, however, with human influence on the criteria, training data, semantics, and interpretation, this could all lead to algorithmic bias and discrimination.[7] This means bias is incorporated at every step of the way; it starts from story selection to the framing of narratives and then how it is spread throughout the media. Consequently, the biased algorithms can lead to a misrepresentation of events, individuals, and communities, thereby influencing the public opinion in a biased manner.[8] [9] The text published by journalists, will therefore, be detrimental to their own industry. Furthermore, examples of AI tools that are harmful to journalism include deep fakes, AI generated versions of people or events that are distorted versions of reality, which are hard to detect by journalists, increasing the spread of misinformation and bias towards certain people or things[10]. This is concerning given journalism's role in providing and reporting factual, unbiased information about the government and events to the public. Lastly, the use of AI in marketing stories towards users is also a cause for concern as it creates personalized bias towards users. In summary, the impact of AI bias can be classified into the following categories:

Writing: AI-driven tools can generate news stories, write headlines, and even adapt articles for different audiences. However, bias is still present, as seen with Twitter's image-cropping algorithm, there is a demonstrated risk that AI technology can perpetuate biases, inadvertently prioritizing certain narratives or demographics over others.[11] Since journalism is founded on truth, accuracy, and objectivity this impacts public perception and trust. Furthermore, the

---

[4]Nazer, "*Bias in artificial intelligence*"
[5]Nazer, "*Bias in artificial intelligence*"
[6]Ivancsics and Hansen, "*Actually, It's about Ethics*"
[7]Diakopoulos, "*Algorithmic accountability*", 398-415.
[8]NAIAC, "*Rationales*"
[9]Zoller, M., and Casteel. "*# March for our lives*:" 813-823
[10]Zoller, M., and Casteel. "*# March for our lives*:" 813-823
[11]Yee, Kyra, Tantipongpipat, and Mishra. "*Image cropping*" 1-24.

inclination of AI algorithms towards sensationalism could further exacerbate the spread of misinformation, as they may favor engaging content over factual accuracy.

Interpretation: The interpretation of AI-generated content, particularly in journalism, can significantly contribute to biases. Recently, in 2024, following the Parkland gun control debate, students advocating for more control used AI to create audio deep fakes of the victims in a campaign for "giving voice to the voiceless".[12] This raises ethical concerns for journalists about the authenticity and truthfulness in reporting. Such applications not only distort public perception and increase societal biases but also challenge journalists' reliance on their traditional verification processes.

Personalization: Misinformation can change public opinion, and fuel national division. Consequently, as political polarization has become increasingly prominent, AI has played a role in this division through its use in spreading advertorials based on the viewer. *Figure 2*, see below, shows the cycle of news personalization from AI and how it can further strengthen sentiment bias. Advertorials on news websites are driven by AI, and promote this effect of sentiment bias based on a brand and their message.[13] Furthermore, studies have shown that based on data Google has collected, called Information Flow Experiments, it will display "personalized" ads geared towards a person's viewership.[14] However, this has led to men receiving significantly more ads about higher paying jobs compared to women.[15] Although this was an unintended consequence, news publications have the capability to adjust their promotional strategies in response to how algorithms influence ad placements and inherently increase sentiment bias.

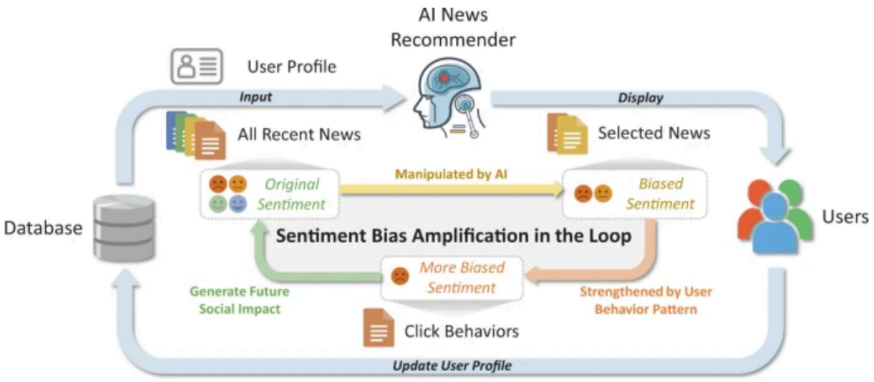

Figure 2; Wu, Chuhan, Wu, Qi, Zhang, Xie, and Huang "Removing AI's sentiment", 1-9

Solutions for News Publications/Journalists to Combat AI Bias

Addressing AI bias in journalism requires transparent guidelines, diversity in AI and human tools, and the implementation of ethical AI practices. Each problem, outlined above, has a

[12]Zoller, M., and Casteel. "*# March for our lives:*" 813-823
[13]Diwanji, Shwetangbhai, Lee, and Cortese. "*Deconstructing the role*" 1-18.
[14]Tschantz, Carl, Datta, and Wing. "*A methodology*", 554-568
[15]Springer-Norris, "*Understand Algorithmic Bias*"

multifaceted approach for mitigation that will overall repair the integrity of the journalism industry. The following solutions are meant to address each problem in multiple ways.

Algorithmic Accountability

AI writing and fact checking is used in most content generation methods for news sources, which necessitates algorithmic accountability in order to maintain the integrity of the media coverage. That means a wide range of AI tools that are trained on diverse data sets and with a variety of perspectives. Studies that have audited saliency biased algorithms, suggest that like Twitter's response to their controversy, other companies should follow suit and report all documentation to which how their models were trained and what data sets they used.[16] [17] The integration of objectively sourced AI tools, will mediate the problem of inaccurate/biased AI written work and remove the artificial generation of sensationalism.

AI Implementation Diversity

Assembling a diverse group of people within the news producers will reduce the risk of using inherently biased AI tools. Experts suggest, that by incorporating a specific AI governance council to doctrine and have authority over specific AI policies creates governance over the integration of AI within their industry[18]. In terms of journalism, creating an AI policy and authorization to enforce and verify these tools will remove the ambiguity of its potential uses in writing.

Bias Detection

Along with the integration of AI for ease of production in journalism, AI can also be implemented to flag articles that contain bias. These procedures would include regular audits of AI tools by cross-referencing variations across different demographic groups.[19] Furthermore, this system can also help fact check journalists, themselves, in their own writing. By keeping bias, especially because of the algorithms, out of articles the public will continue to trust the media and cement its role in society.

Collaboration with Tech Companies

The increased collaboration of journalism outlets and tech companies can lead to the development of AI technology that is more catered for the needs of journalists. This construction will help with eliminating bias in writing, interpretation, and personalized content. For example, blockchain technology could be utilized to verify the authenticity of digital content, providing a transparent and immutable record of its origins and modifications.[20] [21] With AI tools geared for

---

[16]Birhane, Abeba, Prabhu, and Whaley, "*Auditing saliency*" 4051-4059
[17]Li et. al, "*Participation and Division*" 1-19
[18]Gellai,. "*Enterprising academics:*", 568-596
[19]Sourlos et al "*Possible bias*"
[20]Ivancsics and Hansen, "*Actually, It's about Ethics*"
[21]Iqbal, "*The relationship*"

the media industry: deep fakes could be recognized quickly, biases will be kept out of algorithms, and advertorials will no longer have underlying sentiment bias.

Verify AI generated Outputs

The battle against misinformation, intensified by AI-generated content like deepfakes, necessitates a thorough, multifaceted verification process of AI content. Journalists must equip themselves with advanced detection technologies and maintain rigorous standards of verification to discern between genuine and AI-manufactured content.[22] It is essential for media organizations to implement and adhere to robust editorial policies that prioritize accuracy and truth.

Transparency for Targeted Advertorials

Targeted Advertorials are implemented through AI systems that track user content and personalize everything the user sees, however, by enacting transparency about personalization algorithms existing user bias can be avoided. Studies suggest that the optimal user experience for media consumption is one in which varying personalized content is made known.[23] Journalism succeeds from having a diverse viewing audience, however, AI complicates this by strengthening personal bias, but with the purposes of AI content generation disclosed this effect can be avoided.

Conclusion

The significant challenges that journalism faces as a result of the integration of AI into most practices require a methodical and thorough approach to remove bias that is inherently built into the algorithms. Following this white paper, journalists and media industry leaders will be more equipped to implement solutions in order to mitigate these effects and maintain the trust of the general public.

[22]Ivancsics and Hansen, "*Actually, It's about Ethics*"
[23]Shin, Donghee, Rasul, and Fotiadis. "*Why am I seeing this?*", 1214-1234

Works Cited

1. Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., Hicklen, R. S., Moukheiber, L., Moukheiber, D., Ma, H., & Mathur, P. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digital Health*, *2*(6), e0000278. https://doi.org/10.1371/journal.pdig.0000278
2. WAN-IFRA. "Most important newsroom uses of AI (artificial intelligence) for publishers in 2024 according to industry leaders worldwide as of December 2023." Chart. January 12, 2024. Statista. https://www.statista.com/statistics/1119232/predictions-ai-initiatives-for-publishers/v
3. WAN-IFRA. "Most important newsroom uses of AI (artificial intelligence) for publishers in 2024 according to industry leaders worldwide as of December 2023." Chart. January 12, 2024. Statista. https://www.statista.com/statistics/1119232/predictions-ai-initiatives-for-publishers/v
4. Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., Hicklen, R. S., Moukheiber, L., Moukheiber, D., Ma, H., & Mathur, P. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digital Health*, *2*(6), e0000278. https://doi.org/10.1371/journal.pdig.0000278
5. Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., Hicklen, R. S., Moukheiber, L., Moukheiber, D., Ma, H., & Mathur, P. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digital Health*, *2*(6), e0000278. https://doi.org/10.1371/journal.pdig.0000278
6. Bernat Ivancsics and Mark Hansen, “Actually, It’s about Ethics, AI, and Journalism: Reporting on and with Computation and Data,” Columbia Journalism Review, November 21, 2019, https://www.cjr.org/tow_center_reports/ai-ethics-journalism-and-computation-ibm-new-york-times.php.
7. Diakopoulos, Nicholas. "Algorithmic accountability: Journalistic investigation of computational power structures." Digital journalism 3, no. 3 (2015): 398-415.
8. Rationales, Mechanisms, and Challenges to Regulating AI: A Concise Guide and Explanation - Rationales-Mechanisms-Challenges-Regulating-AI-NAIAC-Non-Decisional.pdf
9. Zoller, Heather M., and Diana Casteel. "# March for our lives: health activism, diagnostic framing, gun control, and the gun industry." *Health communication* 37, no. 7 (2022): 813-823.
10. Zoller, Heather M., and Diana Casteel. "# March for our lives: health activism, diagnostic framing, gun control, and the gun industry." *Health communication* 37, no. 7 (2022): 813-823.
11. Yee, Kyra, Uthaipon Tantipongpipat, and Shubhanshu Mishra. "Image cropping on twitter: Fairness metrics, their limitations, and the importance of representation, design, and agency." *Proceedings of the ACM on human-computer interaction* 5, no. CSCW2 (2021): 1-24.
12. Zoller, Heather M., and Diana Casteel. "# March for our lives: health activism, diagnostic framing, gun control, and the gun industry." *Health communication* 37, no. 7 (2022): 813-823.

13. Diwanji, Vaibhav Shwetangbhai, Jaejin Lee, and Juliann Cortese. "Deconstructing the role of artificial intelligence in programmatic advertising: at the intersection of automation and transparency." *Journal of Strategic Marketing* (2022): 1-18.
14. Tschantz, Michael Carl, Amit Datta, Anupam Datta, and Jeannette M. Wing. "A methodology for information flow experiments." In *2015 IEEE 28th Computer Security Foundations Symposium*, pp. 554-568. IEEE, 2015.
15. Springer-Norris, Nia. "Understand Algorithmic Bias in Journalism." *Reynolds Journalism Institute*, 21 Nov. 2022. https://rjionline.org/news/understand-algorithmic-bias-in-journalism/
16. Birhane, Abeba, Vinay Uday Prabhu, and John Whaley. "Auditing saliency cropping algorithms." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 4051-4059. 2022.
17. Li, Rena, Sara Kingsley, Chelsea Fan, Proteeti Sinha, Nora Wai, Jaimie Lee, Hong Shen, Motahhare Eslami, and Jason Hong. "Participation and Division of Labor in User-Driven Algorithm Audits: How Do Everyday Users Work together to Surface Algorithmic Harms?." In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1-19. 2023.
18. Gellai, Dániel Béla. "Enterprising academics: Heterarchical policy networks for artificial intelligence in British higher education." *ECNU Review of Education* 6, no. 4 (2023): 568-596.
19. Sourlos, Nikos, Jingxuan Wang, Yeshaswini Nagaraj, Peter van Ooijen, and Rozemarijn Vliegenthart. "Possible bias in supervised deep learning algorithms for CT lung nodule detection and classification." *Cancers* 14, no. 16 (2022): 3867.
20. Ivancsics, Bernat and Mark Hansen, "Actually, It's about Ethics, AI, and Journalism: Reporting on and with Computation and Data," Columbia Journalism Review, November 21, 2019, https://www.cjr.org/tow_center_reports/ai-ethics-journalism-and-computation-ibm-new-york-times.php.
21. Iqbal, Abid, Khurram Shahzad, Shakeel Ahmad Khan, and Muhammad Shahzad Chaudhry. "The relationship of artificial intelligence (AI) with fake news detection (FND): a systematic literature review." *Global Knowledge, Memory and Communication* (2023).
22. Ivancsics, Bernat and Mark Hansen, "Actually, It's about Ethics, AI, and Journalism: Reporting on and with Computation and Data," Columbia Journalism Review, November 21, 2019, https://www.cjr.org/tow_center_reports/ai-ethics-journalism-and-computation-ibm-new-york-times.php.
23. Shin, Donghee, Azmat Rasul, and Anestis Fotiadis. "Why am I seeing this? Deconstructing algorithm literacy through the lens of users." *Internet Research* 32, no. 4 (2022): 1214-1234.

Bibliography

Birhane, Abeba, Vinay Uday Prabhu, and John Whaley. "Auditing saliency cropping algorithms." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 4051-4059. 2022.

Diakopoulos, Nicholas. "Algorithmic accountability: Journalistic investigation of computational power structures." Digital journalism 3, no. 3 (2015): 398-415.

Diwanji, Vaibhav Shwetangbhai, Jaejin Lee, and Juliann Cortese. "Deconstructing the role of artificial intelligence in programmatic advertising: at the intersection of automation and transparency." *Journal of Strategic Marketing* (2022): 1-18.

Gellai, Dániel Béla. "Enterprising academics: Heterarchical policy networks for artificial intelligence in British higher education." *ECNU Review of Education* 6, no. 4 (2023): 568-596.

Iqbal, Abid, Khurram Shahzad, Shakeel Ahmad Khan, and Muhammad Shahzad Chaudhry. "The relationship of artificial intelligence (AI) with fake news detection (FND): a systematic literature review." *Global Knowledge, Memory and Communication* (2023).

Ivancsics, Bernat and Mark Hansen, "Actually, It's about Ethics, AI, and Journalism: Reporting on and with Computation and Data," Columbia Journalism Review, November 21, 2019, https://www.cjr.org/tow_center_reports/ai-ethics-journalism-and-computation-ibm-new-york-times.php.

Li, Rena, Sara Kingsley, Chelsea Fan, Proteeti Sinha, Nora Wai, Jaimie Lee, Hong Shen, Motahhare Eslami, and Jason Hong. "Participation and Division of Labor in User-Driven Algorithm Audits: How Do Everyday Users Work together to Surface Algorithmic Harms?." In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1-19. 2023.

Leiser, M. R. "Bias, journalistic endeavours, and the risks of artificial intelligence." In *Artificial Intelligence and the Media*, pp. 8-32. Edward Elgar Publishing, 2022.

Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., Hicklen, R. S., Moukheiber, L., Moukheiber, D., Ma, H., & Mathur, P. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digital Health*, *2*(6), e0000278. https://doi.org/10.1371/journal.pdig.0000278

Rationales, Mechanisms, and Challenges to Regulating AI: A Concise Guide and Explanation - Rationales-Mechanisms-Challenges-Regulating-AI-NAIAC-Non-Decisional.pdf

Shin, Donghee, Azmat Rasul, and Anestis Fotiadis. "Why am I seeing this? Deconstructing algorithm literacy through the lens of users." *Internet Research* 32, no. 4 (2022): 1214-1234.

Sourlos, Nikos, Jingxuan Wang, Yeshaswini Nagaraj, Peter van Ooijen, and Rozemarijn Vliegenthart. "Possible bias in supervised deep learning algorithms for CT lung nodule detection and classification." *Cancers* 14, no. 16 (2022): 3867.

Springer-Norris, Nia. "Understand Algorithmic Bias in Journalism." *Reynolds Journalism Institute*, 21 Nov. 2022. https://rjionline.org/news/understand-algorithmic-bias-in-journalism/

Tschantz, Michael Carl, Amit Datta, Anupam Datta, and Jeannette M. Wing. "A methodology for information flow experiments." In *2015 IEEE 28th Computer Security Foundations Symposium*, pp. 554-568. IEEE, 2015.

WAN-IFRA. "Most important newsroom uses of AI (artificial intelligence) for publishers in 2024 according to industry leaders worldwide as of December 2023." Chart. January 12, 2024. Statista. https://www.statista.com/statistics/1119232/predictions-ai-initiatives-for-publishers/v

Wu, Chuhan, Fangzhao Wu, Tao Qi, Wei-Qiang Zhang, Xing Xie, and Yongfeng Huang. "Removing AI's sentiment manipulation of personalized news delivery." *Humanities and Social Sciences Communications* 9, no. 1 (2022): 1-9.

Yee, Kyra, Uthaipon Tantipongpipat, and Shubhanshu Mishra. "Image cropping on twitter: Fairness metrics, their limitations, and the importance of representation, design, and agency." *Proceedings of the ACM on human-computer interaction* 5, no. CSCW2 (2021): 1-24.

Zoller, Heather M., and Diana Casteel. "# March for our lives: health activism, diagnostic framing, gun control, and the gun industry." *Health communication* 37, no. 7 (2022): 813-823.