

Chapter 16

Beyond the Hype: How Explainable AI Can Ensure Ethical Use of Generative AI in Digital Media

Amaresh Jha

 <https://orcid.org/0000-0002-6083-1610>

University of Petroleum and Energy Studies, India

Sanjeev Ratna Singh

 <https://orcid.org/0000-0002-5899-2749>

Bennett University, India

ABSTRACT

The rise of Generative Adversarial Networks (GANs) in digital media presents both opportunity and challenge. While GANs can create engaging content, their ability to generate realistic yet fabricated media raises ethical concerns about information authenticity and user trust. This paper proposes Explainable AI (XAI) as a solution. By demystifying GAN decision-making processes, XAI can promote transparency in content development, enhance fact-checking, facilitate deepfake detection, and mitigate bias within GAN training data. This multifaceted approach using XAI can foster responsible communication practices and build a foundation of trust between digital media platforms and audiences.

INTRODUCTION

The integration of AI in newsrooms is increasingly shaping modern journalism, with organizations like The New York Times and The Washington Post leading the way. The New York Times has pioneered AI applications such as its 'News

DOI: 10.4018/979-8-3693-7235-7.ch016

Engine,' which aids journalists in identifying trends, sourcing, and drafting stories. The Times also utilizes natural language processing to improve search functionality and personalize content recommendations (Broussard et al., 2019). Similarly, The Washington Post has focused on automating routine tasks, such as data analysis and content distribution, enabling journalists to concentrate on investigative reporting. The Post has also employed AI tools for fact-checking and debunking misinformation (Borges-Rey, 2016).

JournalismAI, a collaborative initiative launched by the LSE's Polis think tank and supported by the Google News Initiative, plays a key role in empowering news organizations to use AI responsibly. Since 2019, the project has facilitated knowledge sharing and capacity building, ensuring that journalists worldwide can navigate AI's ethical, practical, and editorial challenges (Biswal and Gouda, 2020). JournalismAI aims to bridge the technological gap in media, promoting equitable access to AI, particularly for smaller or resource-constrained outlets. It also advocates for ethical AI use, emphasizing transparency, accountability, and human oversight in news production.

While AI is revolutionizing content creation, audience engagement, and business operations in the news industry, its integration presents ethical challenges such as bias, privacy concerns, and misinformation. To address these issues, news organizations must balance AI capabilities with human judgment and invest in training to mitigate risks and harness AI's full potential.

From content creation and curation to audience engagement and business operations, AI is being leveraged across the industry. News organizations are employing AI for tasks such as generating articles, analyzing audience data, and optimizing advertising campaigns (Table1). However, the integration of AI is not without challenges. Ethical considerations, including bias, privacy, and misinformation, must be carefully addressed. Striking a balance between human judgment and AI capabilities is crucial to ensure the responsible and effective use of this technology. As AI continues to evolve, news organizations must invest in training their staff to harness its potential while mitigating its risks. By doing so, the industry can leverage AI to enhance journalism, improve efficiency, and better serve audiences.

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/beyond-the-hype/366361

Related Content

Group Process Losses in Agile Software Development Decision Making

Sharon Coyle, Kieran Conboy and Thomas Acton (2013). *International Journal of Intelligent Information Technologies* (pp. 38-53).

www.irma-international.org/article/group-process-losses-agile-software/77873

The Holistic Framework for the Integration of Learning Analytics to Decision-Making Process in the Context of Stakeholder Theory

Atajan Rovshenov and Erman Cokun (2026). *AI-Augmented Creativity in Learning Analytics* (pp. 123-154).

www.irma-international.org/chapter/the-holistic-framework-for-the-integration-of-learning-analytics-to-decision-making-process-in-the-context-of-stakeholder-theory/393422

Examining the Reliability of ChatGPT: Identifying Retracted Scientific Literature and Ensuring Accurate Citations and References

Rosy Jan (2025). *Impacts of Generative AI on the Future of Research and Education* (pp. 367-392).

www.irma-international.org/chapter/examining-the-reliability-of-chatgpt/358780

Integrating AI and Blockchain into Humanitarian Air Logistics: A Case Study of Turkish Red Crescent's Aviation Response in Disaster Relief

Umit Kanmaz (2026). *Reshaping Humanitarian Logistics With AI, Predictive Analytics, and Autonomous Systems* (pp. 345-372).

www.irma-international.org/chapter/integrating-ai-and-blockchain-into-humanitarian-air-logistics/388225

DeTER Framework: A Novel Paradigm for Addressing Cybersecurity Concerns in Mobile Healthcare

Rangarajan (Ray) Parthasarathy, David K. Wyant, Prasad Bingi, James R. Knight and Anuradha Rangarajan (2021). *International Journal of Intelligent Information Technologies* (pp. 1-24).

www.irma-international.org/article/deter-framework/277070