Proceedings of CONF-MLA 2024 Workshop: Securing the Future: Empowering Cyber Defense with Machine Learning and Deep Learning
DOI: 10.54254/2755-2721/94/2024MELB0086

AI-Assisted News Content Creation: Enhancing Journalistic Efficiency and Content Quality Through Automated Summarization and Headline Generation

Fanbu Yuan[1,†], Ruonan Shi[2,†], Ziwei Li[3,4,*]

[1]Monash University, Melbourne, Australia
[2]University of Melbourne, Victoria, Australia
[3]University of Glasgow, Glasgow, Scotland

[4]rara481846778@gmail.com
*corresponding author
[†]Fanbu Yuan and Ruonan Shi have made equally significant contributions to the work and share equal responsibility and accountability for it.

Abstract. Nowadays, Artificial Intelligence (AI) is a buzzword which appears on tech headlines and guides constantly. In a nutshell, AI has been widely implemented in news content generation process, which helped journalists to develop more efficient, diverse and qualitative news production workflows. In this paper, we discuss the applications of AI in automated summarisation, headline generation and content optimisation in newsroom-wide workflow. We analyse this process by looking at how AI tools could potentially reduce human efforts and help humans to produce more concise and accurate summaries and fascinating headlines by also enhancing the engagement of the audience. However, we also argue about the increasing dependency on AI which potentially causes more ethical problems such as algorithmic bias, lack of transparency, and the loss of journalistic integrity. By examining the current AI-driven tools and strategies to avoid the bias produced by machines, we explore both the opportunities and the challenges of AI in journalism. Our findings imply that, although AI substantially improved the production efficiency and content relevance, it requires proper management to prevent the problem such as ignorance of bias and ethical failure. Eventually, humans would still play a part in the story by complementing AI's capabilities.

Keywords: AI-assisted journalism, automated summarization, headline generation, content optimization, algorithmic bias.

1. Introduction

Through the usage of Artificial Intelligence, the media industry is substantially changing the formulation of news content and altering the way news is created and consumed. With the need for improved news accuracy and accessibility in the fiercely competitive digital media realm, numerous news organisations process news gathering and production with the assistance of AI. From automated news summarisation tools that extract and generate complex summaries of long articles to headline generation algorithms that effectively maximise reader engagement, computational journalism has the potential to revolutionise the way newsrooms operate. AI technologies allow journalists to expedite the production

© 2024 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/94/2024MELB0086

of news stories, mitigate manual effort in redundant tasks, and focus more on the editors' discretion and precision. However, the growing incorporation of AI in journalism has come with a cost. Various ethical debates around the issues of algorithmic bias, transparency and loss of human editorial judgement have emerged. For instance, trained on a large dataset, AI systems often inadvertently reproduce or amplify biases of those same datasets, leading to biased or otherwise inaccurate news coverage. Additionally, as sophisticated AI-powered tools for news content creation become mainstream, news organisations are facing questions regarding the transparency of AI's role in news production and the extent to which audiences are informed about these usages [1]. Ultimately, while AI has the potential to maximise efficiency, human involvement is still crucial in maintaining the quality, accuracy and fairness of news content. This paper explores the current landscape of AI-assisted news content creation, highlighting how the technology poses challenges and opportunities for the newsroom, and argues for practical methods towards ameliorating these ethical concerns.

2. Enhancing Journalistic Efficiency

2.1. Automated Summarization Tools

To implement automated summarization tools, Natural Language Processing (NLP) models are typically employed. These models use specific algorithms to extract key sentences or phrases from a text and condense them into a coherent summary. The basic process of automatic summarization can be described using the following simplified formulas:

Text Representation:

The input text T is first represented as a set of word embedding vectors:

$$T = \{w_1, w_2, \dots, w_n\} \tag{1}$$

where w_i represents the vectorized form of the i-th word in the text.

Importance Scoring:

For each sentence S_j, composed of multiple words w_i, the model calculates an overall importance score. The importance score of a sentence can be represented as:

$$\text{Score}(S_j) = \sum_{i \in S_j} f(w_i, T) \tag{2}$$

where $f(w_i, T)$ is a function that measures the importance of the word w_i within the entire text T. Common functions used for this purpose include TF-IDF (Term Frequency-Inverse Document Frequency) or attention-based mechanisms in deep learning models

Sentence Selection:

The top k sentences with the highest importance scores are selected to form the summary:

$$\text{Summary} = \{S_{j_1}, S_{j_2}, \dots, S_{j_k}\} \text{ where } \text{Score}(S_{j_1}) \geq \cdots \geq \text{Score}(S_{j_k}) \tag{3}$$

This approach ensures that the most important sentences are included in the final summary.

By including some of these model steps, automated summary generators can save enormous time and energy in journalistic tasks, by allowing the quick condensing of large amounts of information into short summaries. NLP algorithms utilise word embeddings and importance scores for ranking and selecting the most relevant sentences from long texts, such as interviews, research reports or event transcripts. This feature saves lots of time, which would otherwise be spent by a human to review the content, allowing the journalists to write in-depth reports or to follow up with some new investigations. In the quick and frenetic scenario of newsrooms, where the distribution of news is crucial and happens in real-time, these tools can help to produce summaries in a matter of minutes, allowing a journalist to keep up with their competitors and to stay in the game of a highly competitive and fast-paced media landscape. An important limitation of such tools is the inability to replicate a human-sounding text in articles that require subjective interpretation and complex contextual knowledge [2]. As can be seen in the formulas, the importance scoring of words and sentences involves a measurable quantification of summary selection, but this does not suffice to replicate human intuition about tone, context or subtlety.

DOI: 10.54254/2755-2721/94/2024MELB0086

So, although these tools save a great deal of time and enhance efficiency, there's still a long way to go towards obtaining better quality of such machines – and fixing the issue of misrepresentation of the original content.

2.2. Headline Generation Algorithms

Given the importance of headlines to digital media a generator of sensationalism Continuing on from the example above, the AI-based headline generator algorithm analyses the content of the article to offer a number of options for headlines that are deemed attention grabbing as well as informative. Such algorithms typically employ a combination of keyword extraction and sentiment analysis, along with click-through-rate (CTR) optimisation, drawing on a large dataset of previously effective headlines. Optimising headlines for attention, AI tools capitalise on reader engagement to increase traffic to websites, which is essential to both ad revenue and reader retention. Keeping in mind the influence of AI-generated headlines over the audience's attention, a major risk is that they may be incentivised to prioritise clickability over content accuracy. Sensationalism in AI-generated headlines can subvert journalistic integrity, ultimately leading to a lack of trust among the audience in news organisations. For instance, a focus on misleading headlines might increase traffic but hinder readers from staying on a page long understanding of the story [3]. It is thus imperative to design AI algorithms with-grabbing but also true to the content of the article. Moreover, human editors should continue playing a central role in tuning and moderating AI suggestions to adhere to editorial standards.

2.3. Integration with Content Management Systems (CMS)

This type of integration brings AI-assisted tools directly into workflow of a CMS, embedding summarisation, headline generation and other content optimisation tools into one platform. Journalists using an AI-infused CMS can have multiple tools at their fingertips and, without having to switch between applications, they produce more content. For example, built-in AI tools within a CMS can automatically generate summaries of articles, provide suggestions to strengthen headlines for search engine optimisation (SEO) and provide real-time feedback on how to improve article readability and engagement potential. This integration assures that AI-enriched outputs are immediately incorporated into workflow, allowing journalists to focus on storytelling – not technical optimisations. However, if journalists heavily rely on such AI-infused CMS systems, the risk of AI over-optimisation could lead to a lack of human creativity and editorial oversight. In the best of scenarios, efficient content generation and distribution would be enhanced by AI-based tools. However, that very efficiency could also lead to homogenisation of content, with AI-suggested texts being more influenced by predictable linguistic and structural patterns, rather than by innovative or nuanced storytelling. As CMS systems integrate more AI-based tools, the goals of such tools need to be reinforced by humans as well, to assure that AI complements, not replaces, human editorial oversight [4].

3. Improving Content Quality

3.1. Contextual Understanding in AI Summarization

The key challenge for an AI-assisted summarisation tool is to have a deep understanding of context within the content it is processing. AI systems can often be very good at ascertaining surface-level facts, such as the presence of a keyword or common theme, but it can struggle when it comes to identifying more complex layers of interpretation, such as tone, implied meaning or subtle rhetorical devices. When summarising opinion pieces, investigative reports or cultural critiques, the tool needs to understand not just what is being communicated, but the context or worldview through which the story is being told. Traditional AI models can struggle to capture this nuance, often generating summaries where interpretive or emotional elements are given less importance than mere facts. This can lead to summaries that lack depth or misrepresent the original context of an article. Recent advancements in deep learning models based on transformer architectures – such as GPT – have started to address some of these limitations, as they have been trained on large and diverse datasets. These models can now pick up on

DOI: 10.54254/2755-2721/94/2024MELB0086

linguistic cues that signal a context or subtext, which is allowing them to generate summaries that are more reflective of the underlying message of the article [5]. We are moving in the right direction, but news organisations should exercise caution, as chances are that an AI tool will still miss a subtext that might fundamentally alter the way a reader thinks about a topic.

3.2. Headline Generation for Audience Engagement

Headline generation using AI technology is one of the ways news publishers now connect with readers. Using machine learning algorithms, which read historical user behavior data, these tools can create headlines that generate the highest CTR and social shares. AI can for example use keywords, emotions and headline design to recommend something that's going to resonate with the right type of people. This allows news outlets to tailor their content to their readers, boosting traffic and branding loyalty. But maximizing for engagement comes at the risk that Ai-generated headlines are tempted towards sensationalism or clickbait and put journalism at the expense of clicks. This is especially true if the headline overstates or misrepresents the content of the article, deceiving the reader and discrediting the outlet. This is the problem developers are tinkering with AI algorithms to get the right combination of stimulation and fact checking, to ensure headlines are still rousing but do not distort the content. Moreover, media outlets need editorial guidelines requiring humans to comment on AI-generated headlines, keeping the possibility of ethical missteps in the automated content generation under scrutiny [6]. Table 1 provides insights on AI-driven headline optimisation for some of the metrics that a research agency in digital media used to evaluate the performance of AI-powered headlines: Keyword extraction, sentiment, CTR prediction and editorial opinion [7].

Table 1. AI-Generated Headline

Headline Option	Keywords Extracted	Sentiment Analysis Score	CTR Prediction (%)	Engagement Level (Clicks /Impressions)	Clickbait Risk (1-5 Scale)	Editorial Review (Pass/Fail)
Economic Growth Slows Amid Global Uncertainty	Economic, Growth, Global, Uncertainty	0.35	12.5	0.125	2	Pass
Tech Giants Face Increasing Scrutiny Over Privacy Concerns	Tech, Giants, Privacy, Concerns	0.42	18.3	0.183673469	3	Pass
New Study Reveals Health Benefits of Plant-Based Diets	Study, Health, Plant-Based, Diets	0.85	25.7	0.254901961	2	Pass
Scientists Discover Breakthrough in Cancer Research	Scientists, Breakthrough, Cancer, Research	0.91	31.4	0.317460317	1	Pass
Government Proposes New Regulations to Tackle Climate Change	Government, Regulations, Climate, Change	0.68	22.1	0.220588235	2	Fail

3.3. AI's Role in Enhancing Multimedia Content

Beyond text, AI has become an invaluable tool for enhancing multimedia elements in news articles, including images, videos, and interactive graphics. For example, AI-powered image recognition systems can automatically suggest or generate visuals that align with the content of an article, improving the visual storytelling experience. Similarly, video summarization tools allow news organizations to condense lengthy video footage into shorter, more digestible clips, highlighting the most important moments while preserving the overall narrative [8]. These AI tools help cater to readers who prefer multimedia content over traditional text-based reporting, thus broadening the reach of news articles. In addition to these enhancements, AI can also be used to create interactive data visualizations that allow readers to explore complex topics more deeply. For instance, AI-driven tools can generate dynamic infographics that update in real-time based on new data, making it easier for readers to understand evolving stories such as election results or financial market trends. However, the growing use of AI in

DOI: 10.54254/2755-2721/94/2024MELB0086

multimedia raises concerns about the loss of human creativity in crafting visually compelling and contextually appropriate content [9]. While AI can automate many aspects of multimedia creation, it is essential that human designers remain involved in the process to ensure that the final product is not only functional but also aesthetically engaging and contextually relevant.

4. Ethical Considerations in AI-Assisted Journalism

4.1. Addressing Bias in AI Algorithms

Perhaps the most urgent ethical issue is the question of bias that could arise within AI algorithms. AI systems are typically 'trained' on large amounts of data. If the data on which they are trained is inherently biased, then an AI system will reproduce, and possibly amplify, that bias in its outputs. This can result in biased headlines, summaries or even full articles being produced, in which particular groups or points of view are negatively misrepresented. For instance, an AI that has been trained on a news dataset in which there is a disproportionate focus on bad news events associated with particular ethnic or demographic groups may produce biased summaries or sensationalist headlines. To minimise this risk, news organisations may want to use more curated training datasets. But there is also a need for greater algorithmic transparency. For example, many AI systems are able to record and log the decisions they make and hence regularly reviewed audits of these logs could help to identify and then correct these biases [10]. Some organisations are already using 'fairness checks' which involve 'giving algorithms a quick but algorithmic assessment before they're passed to humans to take a look.' However, as discussed above, human oversight remains an essential part of detecting and addressing these biases, especially since the most advanced AI systems are still not well-suited to making a distinction between relevant and irrelevant information in social and cultural context. Figure 1 presents an overall view of the contribution of the different strategies adopted in mitigating AI bias.

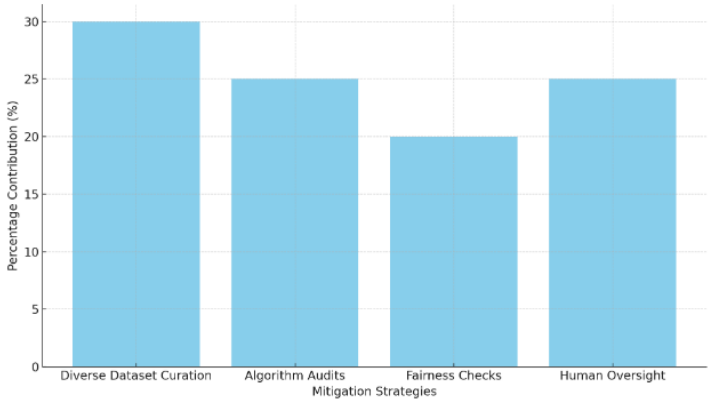

Figure 1. Mitigation Strategies for Addressing Bias in AI Algorithms

4.2. Transparency and Accountability in AI-Generated Content

As AI tools come to play a bigger part in crafting news, questions of transparency are front and centre: audiences have a right to know when and how content we are being served is being created by AI tools. In one sense, it's already happening, as automated systems can alter the framing and interpretation of news. But when news consumers know that a story is the result of machine input, it can boost the trust between them and a news organisation, and ultimately the news itself. It's a matter of trust, but also a question of accountability for the content generated by AI. Some media outlets have already been labelling articles that were generated or assisted by AI tools or algorithms, a practice that could soon become widespread across the industry; right now, though, there's no consensus on best practices for disclosing the presence of AI in content creation.

DOI: 10.54254/2755-2721/94/2024MELB0086

5. Conclusion

News content creation tools using AI have already demonstrated the promise of significantly increasing journalistic productivity by automating parts of the workflow, such as summarisation and headline generation, without compromising on the quality of content, allowing newsrooms to fast-track story delivery in the age of instant information. Besides bringing efficiency, when AI is integrated with working newsrooms, it raises several ethical dimensions, such as avoiding bias, and increasing transparency and user control over editorial decisions. Given the danger of AI-powered news content being considered a form of robot-driven journalism, which lacks the human touch, specific attention needs to be given to how humans can maintain their role in the editorial process to prevent the phenomenon known as 'dehumanisation'. It is clear that while artificial intelligence can improve productivity, it can never replace the nuanced attention that humans with editorial integrity can bring to news production. Cultural and social sensitivities – qualities that cuts across and elevates journalism – cannot be left completely at the mercy of AI. This paper argues that our commitment must be to a hybrid approach that allows AI to boost human productivity without diminishing editorial responsibility.

References

[1] Quinonez, Claudia, and Edgar Meij. "A new era of AI‐assisted journalism at Bloomberg." AI Magazine (2024).

[2] Fridman, M., Roy Krøvel, and Fabrizio Palumbo. "How (not to) Run an AI Project in Investigative Journalism." Journalism Practice (2023): 1-18.

[3] Husnain, Muhammad, Ali Imran, and Hannan Khan Tareen. "Artificial Intelligence in Journalism: Examining Prospectus and Obstacles for Students in the Domain of Media." Journal of Asian Development Studies 13.1 (2024): 614-625.

[4] Lawal, Bashir Abdullahi. "Perception of Registered Journalists on Job Security in the Era of Artificial Intelligence in Katsina State." International Journal of Innovative Social Sciences & Humanities Research 12 (1) (2024): 77-85.

[5] Nishal, Sachita, Charlotte Li, and Nicholas Diakopoulos. "Domain-Specific Evaluation Strategies for AI in Journalism." arXiv preprint arXiv:2403.17911 (2024).

[6] Koniaris, Marios, et al. "Evaluation of automatic legal text summarization techniques for greek case law." Information 14.4 (2023): 250.

[7] Jangra, Anubhav, et al. "A survey on multi-modal summarization." ACM Computing Surveys 55.13s (2023): 1-36.

[8] Zhang, Tianyi, et al. "Benchmarking large language models for news summarization." Transactions of the Association for Computational Linguistics 12 (2024): 39-57.

[9] Jain, Chirag, et al. "A transformer-based approach for abstractive summarization of requirements from obligations in software engineering contracts." 2023 IEEE 31st International Requirements Engineering Conference (RE). IEEE, 2023.

[10] Glazkova, A. V., and D. A. Morozov. "Applying transformer-based text summarization for keyphrase generation." Lobachevskii Journal of Mathematics 44.1 (2023): 123-136.