

Research on the Application of Artificial Intelligence Voice Technology in Journalism and Communication in the Digital Era

Ning Fang

Beijing Institute of Graphic Communication, Beijing, 102600, China

ABSTRACT

The purpose of this text is to explore the application of artificial intelligence voice technology in the context of journalism and communication in the digital era. The potential of artificial intelligence voice technology lies in its capacity to enhance efficiency and facilitate seamless communication. The article puts forward a proposal for the establishment of a solution to the potential crisis of reliability and ethical issues that may arise. Besides, it will examine the hypothesis that the core competitiveness of the industry in question would be transferred to a human-machine co-existence news ecosystem.

KEYWORDS

Artificial intelligence; Voice synthesis; News ethics; Journalism and communication; Human-machine collaboration

1 Introduction

In an era characterized by the digital intelligence, the development of artificial intelligence voice technology is driving significant changes in the landscape of the news media industry. This technology has been shown to enhance the efficiency of news production and to innovate in the format of news reporting. However, it has also been demonstrated that the technology has the potential to challenge the fundamental principles upon which the credibility of news is established. Furthermore, it has given rise to a series of debates concerning issues such as the authenticity of news, its responsibility, and its humanistic value. In light of the prevailing circumstances, it is imperative to establish a framework that ensures the credibility and ethical integrity of technological applications. This necessitates a re-evaluation of the fundamental principles that guide the role of the news anchor, with a view to fostering a collaborative and harmonious environment for the dissemination of news. This approach is of paramount importance in the construction of a sophisticated and cohesive news ecosystem.

The artificial intelligence voice technology necessitates the capacity for such systems to acquire the ability to comprehend and articulate human language, in addition to the capacity to interpret and analyse human language.

The primary function of the system during operation is the implementation of voice synthesis technology(which facilitates the capacity of the artificial intelligence robot to articulate and express itself), the technology of voice recognition (the capacity of artificial intelligence to interpret verbal communication), the technology of comprehending natural language (the capacity of an artificial intelligence robot to comprehend human language expression to the greatest extent possible) and so on, which can engage in further reflection, communication and listening as a significant future development target for human-machine interaction^[1]. As the field of deep learning algorithms continues to evolve and database applications are increasingly developed, the advent of advanced AI voice technology has precipitated a paradigm shift in the realm of voice synthesis, marking a radical transition from the realm of mechanical to that of highly sophisticated and authentic auditory experiences. The degree of naturalness and expressiveness exhibited by these technological advancements has approaches a level comparable to that of a human voice.

This technological innovation has had a profound impact on the news industry. On the one hand, it has significantly enhanced the efficiency of news production, thereby facilitating the automatic and precise dissemination of large quantities of news. On the other hand, it has been instrumental in driving innovation in the domain of dissemination. This has given rise to a plethora of novel forms of dissemination, including virtual news anchors, thereby redefining the manner in which news content is presented and received. However, the credibility of conventional news broadcasting may be undermined by the fact that its voice is no longer solely that of a human broadcaster. Institutional authority weakened by the blurring of sources caused by AI speech and the resulting limitations in the expression of emotion affect the effectiveness of news dissemination. The potential for falsification has given rise to a number of contentious issues, including the question of the attribution of responsibility for ethical considerations in news reporting. This has compelled a re-evaluation of the role of human-machine relationships in the dissemination of news.

2 The present status and application of AI voice technology in the field of journalism and communication

2.1 Enabling efficiency improvement: automated broadcasting of massive volumes of standardized news

The implementation of artificial intelligence voice technology in the context of large-scale standardized news

production serves as an exemplar of the integration of digital technologies in the field of media. The capacity of artificial intelligence voice technology has been demonstrated in a variety of domains, including meteorology, finance and sports. This enables the interpretation of the semantics and emotions in news content, and the generation of expressive voice outputs. This model represents a departure from the conventional linear production process of news reporting. It demonstrates a commitment to the concept of media integration, thereby exhibiting a multifaceted integration of diverse forms of content. The entire process can be completed within seconds using artificial intelligence. The fundamental value of this approach lies in the reduction of redundant manual labour, thereby enabling specialist personnel to focus exclusively on the creation of in-depth content. This, in turn, results in a significant enhancement in the efficiency of news production, thus addressing the challenges posed by the continuous acceleration in news update frequency in the present era. The following example is provided to demonstrate the functionality of the first robot journalist developed by the Xinhua News, who is named 'Xiao Xin'. The aforementioned robot is capable of functioning for a duration of 24 hours per day, 7 days per week. The original requirement for the editing of news articles was between 15 and 30 minutes, whereas the same task performed by the automated system was completed in a mere 3 to 5 seconds. This serves to illustrate the immense value of technology in terms of its ability to enhance efficiency.

2.2 Bridging the information divide: technological practices in accessible communication

The 'use and satisfaction' theory was first proposed by E. Cobb, a sociologist from the United States, in the 1940s. The theory places significant emphasis on the perspective of the audience in evaluating the efficacy of mass communication. It posits that individuals, driven by a diverse array of motivations, actively select and utilize media channels for the purpose of fulfilling their specific needs ^[2]. Throughout the field of journalism and communication, the practice of AI voice technology in accessible communication highlights the humanistic care and social value of technology for good. In the context of news reporting, sign language interpreters facilitate access to crucial information for individuals with hearing impairments. In order to make up for the shortcomings of traditional sign language interpreters in terms of coverage and efficiency, AI sign language anchor technology has emerged ^[3]. The employment of computer vision and motion capture technology enables AI-powered sign language interpreters to translate textual news content into sign language in a manner that adheres to linguistic standards in real time. The efficiency of this process is such that news organisations are able to provide equal access to current news information for both visually and hearing-impaired individuals. This practice is in alignment with the 'use and satisfaction' theory. This is not merely a rudimentary application of technology; rather, it is a significant measure in the construction of an egalitarian society and the fulfillment of media responsibilities, thereby rendering public service information more inclusive and equitable.

2.3 Shaping virtual images: a new strategy for host IP-isation and brand building

In 2018, Xinhua News pioneered the development of an AI-generated news anchor, dubbed 'Xin Xiao Hao', this initiative was followed by the subsequent introduction of a virtual news anchor, named 'Xiao C', by CCTV in 2020, which is utilized for weather reporting and news bulletins. The potential of artificial intelligence to influence the voice technology sector is such that it is now considered a key component in the media's efforts to enhance their brand image and engage younger demographics. In addition to their ability to work continuously for 24 hours, these virtual hosts have been endowed with a distinctive appearance, voice and personality through the integration of audio and visual technologies. They have thus become unique brand symbols, with the capacity to establish a emotional connection with their audience. The aforementioned elements have been elevated to the status of distinctive brand symbols, thereby facilitating the establishment of a emotional connection with the target audience. In the contemporary media landscape, the strategic deployment of numerical hosts with consistent characterization and infallibility represents a pioneering approach to establishing distinct brand identities. This signifies the evolution of the role of the news anchor from mere reporters to public figures, thereby redefining the relationship between the media and its audience.

3 The impact of AI news casting on news authenticity and the path to constructing it

3.1 Technology-induced limitations and the dilemma of journalistic authenticity

The widespread use of AI newscasts will pose a big challenge to news authenticity. The credibility of traditional news is founded on the tangible evidence provided by on-site investigations conducted by journalists and direct communication with sources by broadcasters. The advent of artificial intelligence has obviated the traditional mechanisms, which has had the effect of making it increasingly difficult for recipients to identify the original source of information. Simultaneously, the ambiguity of the source directly undermines the credibility of the news. Secondly, the utilization of artificial intelligence in the domain of voice synthesis remains constrained in complex emotional expression and the analysis of contextual nuances. The potential failure of the media to accurately convey the sentiment of news reports can result in a deficiency of emotional temperature and a misplacement within a specific context. This, in turn, can erode the audience's trust in the objectivity and fairness of the news content, thereby jeopardizing its credibility.

3.2 Strategies for establishing the veracity of news reports

The practice of fact-checking is inextricably linked to the prevailing socio-political context of the time ^[4]. In the context of the current landscape, characterized by the proliferation of data and the prevalence of artificial intelligence, the necessity for factual verification is becoming increasingly evident. In the context of all news articles involving artificial intelligence, it is imperative to ensure transparency in the technology employed. This involves explicitly stating in the articles which images and audio have been generated by AI, thereby fostering trust among the audience in both the veracity of the news content and the credibility of the news institution. On this foundation, it is imperative to strengthen the brand identity of media institutions to a greater extent. In circumstances where the specific source of information becomes ambiguous, the trust placed in the media by the audience will be transferred to the reputation and credibility of the media brand. The formidable influence of media institutions is pivotal in mitigating the uncertainty surrounding technological advancements and laying the foundation for the successful integration of AI into news content. Simultaneously, it is imperative to refine the technical aspects to enhance the user's communication proficiency, thereby facilitating the articulation of more natural and persuasive voice inflections. The establishment of a direct and profound connection between technology and emotion through a design that is imbued with sentimentality serves to better cater to the psychological needs of the audience and their demand for news that is authentic. This process of news reconstruction serves to enhance its veracity and credibility.

4 Risk prevention mechanisms for AI news casting applications

4.1 Attribution of responsibility and accountability dilemmas

The utilization of artificial intelligence in the media industry is an imperative that cannot be overlooked. In the contemporary era, the dissemination of AI-generated news reports has given rise to a series of legal and ethical dilemmas, engendering a state of ambivalence among media institutions. On the one hand, there is a keen desire to capitalize on the technological advantages offered by AI; on the other hand, there is a concomitant need to ensure compliance with existing regulations. In the event of an inaccuracy, bias or ethical violation in the content of the AI announcement, the entity responsible becomes ambiguous and unclear. Technology developers may claim that they are merely providing AI tools and are not responsible for the content generated. Media institutions, on the other hand, may attribute the potential issues to technical malfunctions. Nevertheless, the prevailing legislation continues to exhibit lacunae with regard to the governance of content generated by artificial intelligence. In light of the foregoing, those engaged in the field of news dissemination are required to consider not only the advantages brought about by technological advances, but also the underlying logic of technology and the potential risks that may be associated with its application. They should also be mindful of the potential challenges and take measures to prepare for them ^[5].

4.2 The technological divide and human values

The dissemination of AI-powered news reporting techniques has the potential to exacerbate existing disparities in the field of media. This tendency is indicative of the manifestation of the digital divide within the domain of media. In the context of major metropolitan areas, media entities have the capacity to leverage financial resources to implement advanced AI systems, thereby enlarging disparities in content output efficiency, format, and dissemination impact between major and small media outlets. This dynamic has led to the exacerbation of the 'Matthew effect' within the media industry. Moreover, excessive reliance on intelligent technology has the potential to result in the weakening of the competencies and skills of new industry professionals in the domains of language arts, emotional expression and situational adaptability. Concomitantly, the standardization of AI-generated news reports may result in the loss of the richness and diversity of news content, thereby diminishing the humanistic and ethical dimensions of news communication. This ultimately transforms news reporting into a mere conduit for information, leading to the erosion of its role as a public service.

5 Reshaping the value and future outlook of announcer presenters in the context of human-computer collaboration

5.1 Reconstructing the core competencies of the announcer

In the context of rapid advancements in artificial intelligence voice technology, the professional value of broadcasters is undergoing a significant transformation and elevation. The fundamental and irreplaceable core of the subject will undergo a transition from a basic level of information dissemination to a more advanced level of information transmission.

Firstly, it is imperative that the host of the programme cultivates their capacity to provide in-depth analysis and articulate their ideas. In the event of the successful realization of the task of information transmission by artificial intelligence, the training of professionals in the field of broadcasting and journalism should be redirected towards the

profound interpretation of information, the analysis of background, and the assessment of value. In the context of this discussion, the renowned television presenter, Bai Yansong, serves as a pertinent case study. In his widely regarded programme, entitled 'News 1+1', the primary value of the programme lies not solely in its news reporting, but rather in its incisive analysis and profound commentary on a diverse range of events and issues. This not only ensured the programme's popularity, but also elevated the profile of Bai Yansong, a feat that is currently beyond the reach of artificial intelligence. Whilst it is evident that artificial intelligence has the capacity to enhance the efficiency with which information is obtained, when faced with content of a higher calibre, the general public more frequently opt for the latter.

Secondly, the capacity of present-day artificial intelligence technology to facilitate the establishment of emotional connections between hosts and their audience, in addition to its inability to regulate the atmosphere of the event, represents a significant limitation in the field. Despite the capacity of AI voice technology to simulate certain emotional expressions, it remains incapable of achieving authentic emotional resonance and spontaneous interaction. In the context of significant live events, unexpected occurrences, and humanities-focused programmes, the ability of the host to convey authentic emotions and demonstrate adaptability in real-time is of inestimable value. For instance, the renowned presenter, Dong Qing, has demonstrated her ability to engage in profound emotional interactions with guests on her programme, the Readers. This ability to connect with others on a deeply personal level is a testament to her unique emotional intelligence. Different media create different contexts, and human broadcasters create social contexts with real emotional links.

In addition, it is incumbent upon the news anchor to proactively acquire the competencies of both the operator and the content planner of artificial intelligence. In the future, an accomplished news anchor will be expected to possess not only traditional professional competencies, but also a comprehensive understanding of the latest skills and methods in human-machine collaboration. These include the operation and utilization of artificial intelligence content creation systems, as well as the comprehensive planning of human-machine collaborative content, among other competencies. In the present day, the provision of voice samples and artistic guidance for AI presenters is typically undertaken by professional broadcasters. These professionals collaborate with technical teams to optimize the voice characteristics of AI voice assistants. This situation is consistent with the predictions of Paul Levinson in his seminal work *Human Replay: a Theory of the Evolution of Media*, which postulated the evolution of media according to two theoretical frameworks: the 'anthropotropic' hypothesis and the 'remedial media' hypothesis. These two theories are interrelated, and the development of technology is theorized to follow an anthropogenic trajectory^[6].

5.2 A Paradigm exploration of human-computer collaboration

In order to establish an ideal model of human-machine collaboration, it is essential to define the boundaries of human and artificial intelligence capabilities, and to establish a collaborative system in which artificial intelligence is responsible for efficiency and humans for value-added capabilities. This model signifies more than a rudimentary substitute relationship; rather, it is a sophisticated integration of artificial intelligence and human expertise, resulting in a symbiotic relationship that utilizes each party's strengths. From a practical standpoint, a multitude of efficacious collaborative pathways can be forged.

Firstly, the model of 'artificial intelligence production + human correction' has the potential to be applied to the large-scale production of news content. In such a model, the initial transcription of news scripts is the responsibility of an AI system, while the role of the news anchor is to ensure the quality of the content and enhance its linguistic and artistic aspects. The 'Ethereal Sound' platform, operated by the CMG, serves as a case study. The intelligent editing system utilizes artificial intelligence (AI) technology to automatically capture the most prevalent subjects across the internet, and employs natural language processing (NLP) to generate news articles. It is imperative to emphasize that the most crucial news is subjected to rigorous final review and refinement by professional broadcasters. This is to ensure the accuracy and artistic sophistication of the news presentation. Artificial intelligence is good at symbol processing, while humans are good at meaning understanding, and the combination of the two can maximize effectiveness.

Secondly, the model of 'human core creation + AI assistance' is applicable to the production of high-quality and innovative content. This production model establishes the human creator as the primary entity, whilst the implementation of artificial intelligence technology provides a range of auxiliary support. For instance, in the context of certain large-scale special reports, the following sequence of events occurs: firstly, the journalist conducts an on-site investigation; secondly, the news anchor creates the report's core content; and thirdly, the AI system performs tasks such as multilingual translation, content summarisation, and other auxiliary functions. For instance, Scube, the inaugural AIGC application integration tool developed by SMG, was debuted in a series of programmes related to the two sessions. This inaugural debut showcased a variety of novel AI functionalities. The synergy of intelligent engines and manual review mechanisms has been demonstrated to enhance the efficiency of the entire process of news acquisition, editing, broadcasting and dissemination.

As demonstrated above, the collaborative model between humans and machines has effectively delineated the

contemporary rationale underpinning the division of labour between humans and machines. It provides evidence to the effect that artificial intelligence and human beings are not in a substitute relationship; rather, they complement each other in such a manner as to create a more efficient production system. This collaborative approach, underpinned by a unique blend of competencies and specializations, offers a pragmatic and viable solution to the prevailing challenges in the field of news production. Moreover, it establishes a solid foundation for the future implementation of AI in various domains, paving the way for a more profound and extensive integration of AI technology across diverse sectors.

5.3 Outlook for the future development of the industry

In the future, the construction of a healthy, trustworthy and human-centred news and media ecosystem will necessitate collaborative advancement in technological innovation, talent cultivation and industry. In the dimension of technological development, the future of AI voice technology will develop in a more intelligent and humane direction. AIGC will better understand context and emotion to produce more expressive voice content, enabled by technology that automatically adjusts intonation and cadence based on content. Simultaneously, it is imperative to employ pertinent technical expertise to trace and verify the provenance of artificial intelligence-generated news content, thereby mitigating the potential risks associated with misinformation.

In the context of the cultivation of future broadcast news presenters, it is imperative to overcome the conventional limitations of academic disciplines. This should be achieved by augmenting the existing curriculum with courses that focus on artificial intelligence fundamentals and human-machine interaction design. Concurrently, the moral integrity and critical thinking skills of relevant personnel are to be reinforced. The objective is to cultivate news anchor personnel who are proficient in both the theoretical principles and practical application of the art of broadcasting, and who are also versed in the latest technological developments.

In prevailing industry standards, it is imperative to establish a new set of ethical principles and professional standards that are commensurate with the prevailing technological landscape of the age of artificial intelligence, establishing red lines for the use of AI speech to avoid ethical regulations that may be violated. At the same time, consensus should also be built with the international community when formulating norms, and relevant laws and regulations should be established through full and rational discussions.

6 Conclusions

As demonstrated above, the efficiency of news production has been enhanced by artificial intelligence voice technology due to its unique characteristics. The boundaries of news dissemination have been expanded, and innovative forms of media have been introduced. The potential of this technology to transform the news industry is significant. Despite the present limitations in technology, various solutions are available to address these issues, thereby establishing an advantageous relationship between artificial intelligence and humanity.

It is imperative to acknowledge that both artificial intelligence and other technological tools necessitate a delicate balance between innovation and standardization during their utilization. This equilibrium is paramount to ensure that these technologies serve humanity effectively, thereby facilitating the establishment of a more inclusive and diverse media environment. In order to achieve this, it is not sufficient to merely rely on the continuous development of technology; there is also a need for the relevant legislation to be constantly refined. Furthermore, it is essential for each individual working in the field of communication to embrace change in both theory and practice, to enhance their interdisciplinary skills, and to find their own space for creativity in the new paradigm of human-machine collaboration. In this way, they will be able to fulfil their own potential.

References

- [1] Xie Zhongyang(2019). Application of Artificial Intelligence Speech Technology in Broadcast Media. *Computer Knowledge and Technology* , (05):217-218.
- [2] Zhao Peichuan(2025). An Analysis of the Innovative Strategies of Network Talk Programmes Based on the Theory of Usage and Satisfaction-- Taking the Seventh Season of 'Roundtable' as an Example. *Journalism & Media Studies*,(07):78-80.
- [3] Dong Xin(2025). Exploring the Use of AI Sign Language Anchors in News Reporting. *Journalism & Media Studies*,(08):10-12.
- [4] Yan Wenjie, Liu Yusi, Zhou Ruiming(2022). Fact-checking: What to check, and who is to check? On the Boundary and Normative Ideal of Fact-checking. *Global Journal of Media Studies*,(3):156-174.
- [5] Zhou Junjun(2025). The application of AI technology in the field of news media is analyzed. *Fusionmedia Time*,(04):49-51.
- [6] Paul Levinson(2017). *Human Replay: a Theory of the Evolution of Media*. Chongqing: Southwest China Normal University Press.