ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202310.00089

Postprint: Development and Application of Artificial Intelligence, Exemplified by AIGC, in the Media Field

Authors: Yang Kongwei

Date: 2023-10-08T00:00:00+00:00

Abstract

Objective: To explore the transformative and innovative forces that artificial intelligence brings to the field of news media. Method: Analysis of the evolution of news content generation models, the reasons behind the AIGC explosion, and the practical applications of AIGC. Results: Technological advancement has facilitated the development of AIGC in the news media sector, with AIGC bringing revolutionary forces to the field. Conclusion: Traditional media must be courageous in innovation and embrace new artificial intelligence technologies represented by AIGC with a more open and positive attitude.

Full Text

Preamble

ChinaXiv Partner Journal: The Development and Application of AIGC-Represented Artificial Intelligence in the Media Field
(Xinhua News Agency Zhejiang Branch, Hangzhou, Zhejiang 310007)

Abstract

[**Objective**] This paper explores the transformative and innovative forces that artificial intelligence brings to the news media industry. [**Methods**] We analyze the evolution of news content generation models, the reasons behind AIGC' s explosive growth, and its practical applications. [**Results**] Technological advancement has propelled AIGC' s development in news media, delivering revolutionary power to the field. [**Conclusion**] Traditional media must embrace innovation and adopt a more open, proactive attitude toward embracing new AI technologies represented by AIGC.

Keywords: artificial intelligence; AIGC; deep learning; content generation; neural network
Classification Code: G213
Document Code: A
Article ID: 1671-0134(2023)05-076-05
DOI: 10.19483/j.cnki.11-4653/n.2023.05.017

Introduction

Artificial intelligence is no stranger to humanity. As early as 1997, the supercomputer Deep Blue defeated chess grandmaster Garry Kasparov, astounding the world with computational power. In 2017, AlphaGo defeated top human Go player Ke Jie 3-0, demonstrating capabilities beyond the strongest human players and causing tremendous shock. Since its inception, artificial intelligence has experienced three major waves of development. In 1956, the scientific community proposed the concept of artificial intelligence, and Alan Turing, the father of AI, introduced the famous "Turing Test." Stimulated by this test, the first wave of AI development emerged globally, though it remained in the scientific exploration phase due to technological limitations. In the 1980s, AI welcomed its second wave as scientists employed new probability-based statistical models to analyze correspondences between phonemes and syllables, achieving breakthroughs in speech recognition and machine translation. However, limited training data and scalability resulted in low accuracy rates, leaving AI with minimal practical value.

It was not until 2006 that Geoffrey Hinton, known as the "father of neural networks," proposed deep learning technology, launching the third wave of AI development and ushering in an era of leapfrog progress. Present-day AI has achieved breakthroughs in computer vision, speech recognition, natural language processing, and data mining, continuously penetrating new innovative domains. It exhibits new characteristics such as deep learning, cross-domain integration, human-machine collaboration, collective intelligence, and autonomous control, significantly impacting news production models and editorial processes [1].

1. Development and Evolution of Generation Modes in News Media

In the traditional media era, production models for newspapers and television represented Professional Generated Content (PGC). In the internet age, individuals and groups with strong knowledge backgrounds and professional qualifications began producing content in ways nearly identical to traditional media, with relatively specialized production workflows yielding refined, quality-controlled content. Platforms like Youku and Tudou were early video sites focusing on PGC, and content platforms such as WeChat Official Accounts and portal websites also adopted this production model.

The proliferation of the internet and maturation of 5G technology have fueled

the vigorous development of mobile internet. The continuous growth of mobile users has driven the rise of social media, where interactive normalization enables users to publish, display, and disseminate original text, images, or audio-video content in various forms. In this era of information explosion, particularly with the rise of the Web 2.0 concept characterized by personalization and the popularity of apps like Douyin and Toutiao, User Generated Content (UGC) production models—where everyone becomes a content producer—began challenging the PGC model universally adopted by traditional media. UGC blurs the boundary between producers and consumers, as users are simultaneously content creators, audiences, and consumers. This approach lowers barriers to content production and product access, satisfies personalized and diverse user needs, enables explosive growth in content products, and greatly enhances media industry prosperity. However, due to uneven professionalism among UGC producers, content quality is inevitably affected. Since the UGC era began, the role of "news disseminator" is no longer exclusive to professional news organizations but can be anyone—or even something non-human—resulting in mixed quality of news information [2].

Even during AI' s second wave, scientists explored using artificial intelligence to generate various types of content such as news, music, and poetry. This early AI relied on rule-based systems with pre-set algorithms to generate simple content. As AI technology iterates, a new content production method—AIGC (AI Generated Content)—has gradually moved from concept to implementation, creating new patterns and ecosystems across industries, particularly in media, at a speed beyond expectations and profoundly changing industry evolution models.

2. Reasons for AIGC' s Explosive Development

2.1 Demand for Content Generation

In communication practice, content production always occupies the upstream position in the media industry chain and value chain. Media that master content advantages can often remain invincible in competition through high-quality content [3]. Under the PGC model, content production relies on specialized teams. To ensure quality, PGC teams must invest substantial human, material, and technical resources, making it difficult to meet user quantity and coverage demands. UGC blurs producer-consumer boundaries, lowering production thresholds and satisfying personalized and diverse needs, but uneven professionalism among producers inevitably affects content quality. Since the UGC era began, "news disseminators" are no longer exclusive to professional news organizations but can be anyone—or even non-human entities—resulting in uneven information quality [4].

From the user perspective, fragmented-time reading habits have also fueled rapid growth in demand for reading resources. In the era of "personal media," users obtain information in real-time through smart terminals while constantly trans-

mitting and sharing it. Information publishing forms are "thousand faces for thousand people," with low thresholds and multiple channels. While providing higher audience participation and interactivity, this also creates an ocean of information that overwhelms users and increases search costs. Compared to traditional print media, smart terminals deliver visual and auditory effects that monotonous print graphics cannot fully express, requiring high-quality audio-video content to provide immersive emotional experiences. Various factors have raised higher requirements for information richness, diversity, and professionalism, yet the strong demand for content consumption and insufficient supply remain prominent contradictions. PGC and UGC gradually struggle to match the rapid expansion of content demand, and AIGC has developed rapidly in response.

2.2 Increasingly Mature Technology

Technological innovation is the primary driver of advanced productivity development, and technology is also the primary productive force driving media transformation. AIGC' s development and maturation cannot be separated from related technological progress. Generally, AI' s three core elements are data, computing power, and algorithms.

2.2.1 Massive Data Foundation Massive data forms the foundation of deep learning algorithms. Deep learning essentially simulates human brain operation mechanisms, and like humans who must continuously practice to acquire skills, AI requires the same. AI algorithm implementation involves training and inference stages. The training stage requires massive data input to train a complex neural network model. Through comprehensive training, this model can quickly and efficiently summarize patterns and "infer"various conclusions from new data. Thanks to multi-layer neural networks, the larger the training data volume and the more scenarios covered, the higher the accuracy. Thus, the richness of model training data largely determines deep learning algorithm quality.

In today' s big data era, networks, cameras, and sensors constantly generate various text, voice, image, and video data. According to IDC reports, as internet technologies continue advancing, global data volume reached 82 ZB in 2021 and is expected to reach 175 ZB by 2025 [5], all of which can serve as abundant training "feed" for AI algorithms. With continuous improvement in big data technology, the cost of obtaining annotatable AI training data has decreased, reducing AI commercialization costs and accelerating AI 普及 applications across industries.

2.2.2 Continuously Improving Computing Power Computing power is the new productive force in the digital economy era, and its magnitude directly relates to data processing capability. Computing power originates from chips, and as the foundation and core hardware of computing power, chip development determines AI progress.

Early on, limited chip processing capabilities meant machines could only handle simple programs. AI could only complete basic tasks, and computer performance severely constrained AI development. Over the past 20 years, processor performance has improved at approximately 55% annually, with global computing power scale maintaining rapid growth. Rapid iteration of computing chips has enabled the current AI explosion. Currently, global AI computing power primarily relies on GPU chips. With hardware development and continuous technological updates, dedicated AI ASIC chips and customized FPGA chips are expected to support the next wave of AI computing power development. Diversified technical architectures such as heterogeneous architecture, dedicated computing architecture, and ubiquitous collaborative computing architecture, along with breakthroughs in computing technologies derived from interdisciplinary integration with mathematics and physics—such as quantum computing and in-memory computing—will further drive accelerated computing power enhancement.

2.2.3 Continuous Algorithmic Progress As early as the 1940s, people proposed the concept of neuron mathematical models. In the 1980s, Hopfield neural networks and BT training algorithms further advanced neural network research. Early Hopfield neural networks abstracted human brain structure and external stimulus response mechanisms to explore mathematical models that could simulate learning, association, memory, and pattern recognition functions of the human nervous system for logical operations [6]. However, early neural networks could only solve single problems, and complex data processing was limited. They could not exhaustively enumerate complex scenarios, and when network scale and data volume accumulated to a certain extent, accuracy could not improve even with more data, resulting in limited practical application value.

In 2006, Hinton and others proposed the Deep Learning algorithm for neural networks, liberating humans from complex algorithmic induction and opening new frontiers in machine learning and the third wave of AI development. Compared to traditional machine learning algorithms, Deep Learning features excellent feature learning capabilities, eliminating the need for manual rule-based feature extraction. Machines can automatically extract features, leaving complex inductive algorithms to be completed by machines. Meanwhile, as data scale increases, algorithmic accuracy continuously improves, compensating for the low accuracy defects of traditional BP neural networks. Relevant data shows that before introducing deep learning, speech recognition accuracy remained stable at 76.4% for three consecutive years, while in 2018, Alibaba DAMO Academy' s speech recognition team used deep learning technology to launch a new generation speech recognition model DFSMN, raising the global speech recognition accuracy record to 96.04% [7].

Top AI expert Andrew Ng noted: "Developing AI is like launching a satellite with a rocket—it requires a powerful engine and sufficient fuel. Algorithmic models

are the engine, high-performance computing power is the tool for building the engine, and massive data is the engine' s fuel." Data, algorithms, and computing power are AI' s three most important elements. They mutually promote and support each other, ultimately facilitating AI technology application and value creation [8].

3. AIGC Applications in the News Field

Multimodal pre-training models have endowed AIGC with diverse content production capabilities, enabling not only basic content generation modes such as text, images, audio, and video, but also cross-modal content production among them, bringing a new production revolution to the news media industry.

3.1 AIGC + Text Generation

As the earliest developed AIGC technology, text generation has been widely applied in news reporting. As early as 2014, the *Los Angeles Times* used an earthquake news automatic generation system to broadcast the first news report on a California earthquake, bringing revolutionary force to news media. In recent years, as AI technology has continuously developed and matured, domestic and foreign media have begun applying it throughout news production and dissemination industry chains.

The Associated Press has used the generative AI tool Wordsmith to automatically generate sports news and financial reports since 2018, achieving at least 50,000 automated articles to date. In 2018, Reuters launched an AI news writing tool called Lynx Insight Service to help journalists analyze data, propose story ideas, and automatically generate reports on financial markets and corporate earnings. Xinhua News Agency developed a news robot called "Kuai Bi Xiao Xin" in 2015 that improves news production efficiency through automated data collection, processing, writing, and editing. AI expert Andrew Ng stated: "AIGC can help humans create more high-quality content and better understand complex data and information." Major media organizations domestically and internationally have begun using AIGC to improve news production speed and efficiency, with AIGC-produced content increasingly permeating all aspects of life and providing readers with more personalized news experiences.

3.2 AIGC + Video Editing

With 5G technology development and smart terminal popularization, the formation of audience "fragmented" content consumption habits has enabled rapid short-video development, making it mainstream across content consumption fields. However, standardized content formats and product homogenization also pose fierce competition and challenges for producers. How to improve content production efficiency and quickly launch high-quality, audience-welcome content has become a key focus for media content creators. The convenience and advantages brought by combining AIGC technology with short-video content creation

offer the optimal solution. AIGC technology applications can better improve video quality and creation efficiency, effectively differentiate target audiences, and more efficiently promote relevant works.

Current AIGC applications in video primarily focus on video content attribute editing and automatic clipping functions. For video content editing, AIGC can achieve automatic quality restoration, sensitive figure recognition, theme-based automatic tracking and clipping, visual effects, and automatic beautification. For automatic video clipping, AIGC can parse features based on multimodal information such as visuals and sound in videos, detect according to corresponding semantic constraints, and clip and synthesize segments meeting conditions, thereby achieving intelligent extraction, automatic production, and panoramic live-stream clipping functions.

As early as 2017, Xinhua News Agency and Xinhua Zhiyun, focusing on news content production automation scenarios, jointly launched an AI platform called "Media Brain" that uses AIGC technology to help editors quickly lock shots, rapidly clip highlight reels, and achieve one-click rapid publishing to major platforms, simplifying workflows and saving editors time to achieve "rapid dissemination." China Media Group used an AI intelligent automated production and clipping system during the Beijing 2022 Winter Olympics, leveraging massive event resources to achieve rapid automatic clipping of key moments and large-scale automatic short-video generation and publishing, effectively saving labor costs and liberating editors from tedious manual clipping work. AIGC's clipping capabilities enabled China Media Group to achieve competitive advantages in speed and quality in Winter Olympics video reporting.

AIGC' s widespread application in sports media video content production is the general trend. While greatly improving content production efficiency, it will further extend toward content diversification, creating systematic, structured premium content that satisfies audience dual rigid demands for content quality and quantity.

3.3 AIGC + Digital Humans

AIGC digital humans essentially establish links and interactions between humans and the virtual world while liberating real human labor. AI digital humans can simulate real people' s appearance, movements, expressions, and voice characteristics to achieve vivid realism and can simulate human thinking and behavioral characteristics through natural language models. Since AIGC digital humans are computer-generated, they do not experience birth, aging, illness, or death, nor are they affected by time or environment, making them tireless "model workers."

Empowered by natural language processing, speech synthesis, and speech recognition technologies, AIGC digital humans in broadcasting and hosting can provide excellent hosting work like real anchors while being online 24/7. They can not only create different digital hosts according to program scenarios but also

imitate different avatars of beloved hosts to "play" different roles as hosts or news anchors for various columns, explaining knowledge in different fields such as technology, culture, history, geography, and cuisine. These digital humans are elegant, knowledgeable, and possess comprehensive, extensive knowledge. If connected to dialogue models (ChatGPT), digital humans can also achieve face-to-face communication, answering various questions from audiences and making human-computer interaction more authentic and vivid. Human-machine collaboration manifests not only in information processing workflows but also in deep emotional communication at the physical and mental integration level [9].

As technology matures, AI-driven digital humans will become the mainstream of future digital human markets. Digital humans increasingly approaching real human appearance will provide more friendly, natural, and efficient service experiences for audiences across industries, particularly news media. Under the AI wave, digital humans will also become the link connecting humans to the digital world, accelerating the continuous evolution and diversified development of metaverse industries, enriching digital application scenarios, and promoting digital economy industry development. IDC's "China AI Digital Human Market Status and Opportunity Analysis, 2022" report predicts that by 2026, China' s AI digital human market size will reach 10.24 billion yuan [11].

3.4 AIGC + Voice Applications

In news media dissemination, voice has become an indispensable "element" with its unique appeal. Professional voiceovers convey the narrator' s emotions, enhancing audience empathy and experience. However, audio production itself presents challenges: traditional program audio suffers from single creation forms, high voice actor requirements, time consumption, and high costs. With AI technology development, AI speech recognition and synthesis technologies have gradually been applied in news media. Early speech generation systems lacked expressive logical reasoning and causal relationship capabilities, lacked continuous speech prosody, and felt mechanical and monotonous. In recent years, with rapid digital signal processing technology development, speech synthesis technology has made considerable progress. Highly human-like, fluent natural speech synthesis services, speech broadcasting, and simulated human voiceovers have been widely applied in news media, improving audio content user experience.

AIGC has also been applied in voice cloning and generating customized voices for virtual humans, with enhanced interactivity and real-time performance. Generated audio content has emotion and warmth—deep and powerful, playful and cute, strong and forceful, or soft and moving.

Xinhua News Agency jointly launched the world' s first 3D AI synthetic anchor "Xin Xiaowei" with Sogou in 2020. Using ultra-realistic 3D digital human modeling, real-time facial action generation and driving, multimodal recognition and generation, transfer learning, and other AI frontier technologies, the

system can automatically generate highly similar digital human video content based on input text while producing corresponding facial expressions and body language according to semantics during broadcasting [10]. Xinhua Zhiyun began experimenting with digital humans in 2019, pioneering real-time audio and AI real-person image synthesis in the news field. Based on deep learning models, motion simulation, emotion simulation, and other technologies, AI can generate lifelike digital humans with appropriate expressions and lip-syncing after several hours of training from just a few minutes of real-person video. In 2020 local Two Sessions coverage, seven provinces used Xinhua Zhiyun' s virtual anchors. During the 2023 National Two Sessions, Baidu used interactive hyper-realistic digital humans combined with AIGC technology to interpret the Supreme People' s Court work report to the public through a tech-savvy human-computer interactive dialogue format.

On November 1, 2022, Xinhua News Agency released the AI MV "Sheng" featuring digital reporter and world' s first digital astronaut Xiao Zheng, showcasing AI singing and intelligent video creation capabilities. This AI MV was produced by Xinhua News Agency' s Media Convergence Production Technology and Systems National Key Laboratory jointly with Tencent Music Entertainment Group. The singing demonstrated highly human-like synthetic voice technology, generating sweet and deeply emotional speech [12].

The deep integration of AIGC and intelligent voice technology, along with innovative application scenarios in the news media industry, is expected to further promote the development of the intelligent voice industry market.

3.5 AIGC + Video Generation

AI-generated voiceovers are currently the most relied-upon tool across various media industries, particularly short-video content creators. However, with AIGC's leapfrog development, traditional voiceovers can no longer satisfy AIGC' s ambitions. Technological evolution shows that after large language models and large image models, generative diffusion models based on deep learning and multimodal pre-training large models have become the new trend, with video generation being a representative application field.

In 2022, Meta launched a text-to-video model called Make-A-Video, enabling direct video "writing" from text. Its AI model can also convert two static images into video, generate continuous video from two images, or create new videos based on original footage, producing videos with rich aesthetic and artistic styles. Similarly, Google launched an AI model called Phenaki that can generate up to two minutes of story-driven video content from a script prompt alone.

Domestic AI giant Baidu has also kept pace, accumulating AI pre-training model technology since 2019 and applying Wenxin large model capabilities to the intelligent video synthesis platform VidPress. This platform achieves automatic text-to-video conversion through five automated steps: text analysis and summarization, media material collection, intelligent material processing, audio-visual

alignment, and video editing—all based on large-scale Wenxin training.

In 2022, Xinhua News Agency and Baidu Wenxin Yige jointly launched an AIGC video "AI Depicts Heavenly Palace Feast," with all visuals generated entirely by AI. Through magnificent paintings with Eastern imagery, it reviewed the glorious 30-year journey of China' s manned space program.

In media, cultural entertainment, education, and many other fields, video content production is currently the most important way to attract audiences. More people are promoting and commercializing through creative short-video content, which will drive strong demand across industries for AIGC-based video generation. As video becomes the primary information carrier, whether in news programs, short videos, live streaming, or film and television, AI video content generation will become AIGC' s main development direction. When video generation quality reaches professional levels, AIGC will open more imaginative space for content creation and serve audiences on a larger scale, meeting modern life needs.

As artificial intelligence and other new technologies continue to improve, media has entered a new era of intelligent media, with media intelligence reshaping the original ecology of the media industry [13]. As Bill Gates stated, this transformative AIGC technology' s impact is no less than the birth of the PC or the internet. The storm AIGC has stirred in the tech world is extending into various fields, bringing enormous changes to people' s work and life patterns. Its emergence presents new challenges and opportunities for all industries, including news media, and compels people to adopt a more open, proactive attitude to adapt to continuous technological innovation and development, constantly enhancing their own innovation capabilities and core competitiveness to maintain advantageous positions in this technological wave.

References

[1] Ren Ruijuan, Wang Baochao, Zhao Yaqian. Evolution and Trends: The Application of Artificial Intelligence in the Media Field[J]. Journalism and Communication Review, 2021(2): 26-35.
[2] Wang Yikun. The Traditional Media-ization of Social Media: Investigation and Analysis from the Perspective of Content Production[EB/OL]. http://media.people.com.cn/n1/2017/0110/c409682-29013077.html. 2017-01-10/2023-03-25.
[3] Zou Haitao. Reflections on "Content is King" in the Era of Converged Media[J]. Today' s Media, 2017(11): 150-151.
[4] Fan Hong, Zhao Xiaobo. "AI + Media" : Reconstructing and Recreating News Production Processes[J]. Military Correspondent, 2020(8): 22.
[5] IDC Innovator: Intelligent Document Processing (IDP), 2022[EB/OL]. https://www.sohu.com/a/630711212_{121124366}. 2023-01-16/2023-03-25.
[6] Introduction to Artificial Neural Networks[EB/OL]. https://blog.csdn.net/fengbingchun/article/details/502
2015-12-12/2023-03-25.

[7] Alibaba Launches New Generation Speech Recognition Model DFSMN, Now Open Source[EB/OL]. https://www.sohu.com/a/235054085_{223764}. 2018-06-11/2023-03-26.
[8] Sun Shi. Analysis of Using Artificial Intelligence Technology to Govern Online Violence[J]. China Media Technology, 2023(1): 64-68.
[9] Zhao Xin. Affective Expression: A Preliminary Discussion on Body Communication in AI-Synthesized News Live Streaming[J]. China Media Technology, 2022(4): 65-67.
[10] Wu Jinfu. "She" is Here! The World' s First 3D AI Synthetic Anchor Makes a Brilliant Debut[EB/OL]. http://www.xinhuanet.com/politics/2020lh/2020-05/20/c_{1126011533}.htm. 2020-05-21/2023-03-23.
[11] IDC Report: China' s AI Digital Human Market Size Will Reach 10.24 Billion Yuan by 2026[EB/OL]. http://isuike.com/archives/15559. 2022-06-28/2023-03-25.
[12] AI Singing, AI Painting: Xinhua News Agency' s Digital Astronaut Releases Single[EB/OL]. https://baijiahao.baidu.com/s?id=1748275441565636192&wfr=spider&for=pc. 2022-11-02/2023-03-26.
[13] Peng Lan. Everything is Media: The Pan-Media Trend Driven by a New Round of Technology[J]. Editors' Friend, 2016(3): 5-10.

Author Information: Yang Kongwei (1977-), male, from Hangzhou, Zhejiang, holds a master' s degree and is an engineer. His research focuses on media technology.
(Responsible Editor: Zhang Xiaojing)

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.