

Human-Computer Collaboration, Free Traversal, and Spiritual Rhyme Return: Ruminating on the Development Path of AI Anchors under the Threshold of Artificial Intelligence

Yiyu Zhu

Guangzhou University, Guangzhou, Guangdong, 510006, China

ABSTRACT

Along with the maturation of digital technology, the commercialization of the landing, from the collective appearance of Xiaosha and Zhu Xiaoxun at the 2019 CCTV Spring Festival Gala to Xiaotian and Xiaodong, who will be online during the National People's Congress in 2024, AI anchors have evolved from the early days of "paper people" to "digital beings" that are almost as real as they are, covering news, variety shows and daily life. However, due to the lack of emotional tension and strain, linear narrative, and mechanical replication, it is still easy to trigger the dilemma of trust, choice, and spirit. In this paper, we will ruminate on the three major themes of human-machine synergy, free traversal, and the return of spiritual rhyme to provide healing countermeasures for the plight of AI anchors under the threshold of artificial intelligence, and to point the way for the development of AI anchors.

KEYWORDS

AI Anchors; Human-machine collaboration; Free traversal; The return of spiritual rhymes

1 Introduction

AI anchor as "image cloning + voice cloning + content generation + real-time drive" four modules composed of multimodal technology, with AIGC as the "underlying engine", with massive data and large models to quickly produce programmes. However, it is also prone to the three major problems of information leakage, small audience choice, and work de-domainisation due to its own lack of human touch, adherence to linear broadcast logic, and mechanical reproduction. In order to solve the above problems, this paper puts forward the following countermeasures: human and AI co-production, from bystanders to side-by-side, to rebuild the subject's trust; AI anchor combined with digital restructuring technology, forming "content nodes arbitrary jump + real-time feedback driven" effect, one-way live broadcasting into "audience-led open Theatre", in saving energy, head between the amplification of the audience's right to choose the content of the programme, enhance the audience's sense of participation; into the broadcasting of the four external skills and internal three major skills, with the "scenario reproduction" to create the scene real, to the sound of the mood changes in the transmission of the national culture and the spirit of China, the audience from the "mechanical reproduction of the era of bystanders" to the "mechanical reproduction of the era of bystanders". The audience will be pulled back from the "spectators in the age of mechanical reproduction" to the "dancers in the age of digital rituals". To sum up, this paper focuses on human-computer collaboration, free traversal, and the return of spiritual rhyme, and explores the trendy technology that can be combined with AI anchors from programme content production to user experience, so as to achieve the unity of "technical instrumentalisation", "audience participation", and "culture-subjectivity". The unity of "technological instrumentalisation", "audience participation" and "cultural subjectivity" will be achieved, and countermeasures and approaches will be provided for the current problems and future development of AI anchors.

2 Human-machine collaboration - from spectator to side-by-side

2.1 Human-computer interaction optimises content production

Traditional AI anchors specialize in a series of repetitive labor such as text, voice, editing, and public opinion, which can achieve cost reduction and efficiency in content production on the production side, but due to their lack of true emotional understanding and empathy, they are unable to fully realize human emotional expression and non-verbal communication. In today's data platform era, human-computer collaboration, as an emerging interaction mode, is undergoing a paradigm shift from a traditional instrumental role to a partner role ^[1]. The addition of people to the AI anchor's work can fill in the emotional gaps in the content of the programme, achieving a balance between efficiency and emotion.

In content creation, humans are responsible for creative planning and value judgement, and AI is responsible for accurate execution and rapid generation, which can ensure efficient content production without losing emotion. For

example, the Bloo project in the Netherlands, through real people to manipulate the virtual image micro-expressions and improvised responses, combined with AI to complete the multi-language dubbing and video editing, through the AI to achieve efficient production while retaining the emotional temperature of the real person interaction, and ultimately to achieve the efficiency of single-video production to the traditional mode of three times, the emotional interaction to increase the click rate of 37 per cent. Another example is Fuzhou Radio and Television Station, which has established a closed-loop mechanism of "Problem-Feedback-Optimisation", using AI intelligent frame filling technology, segment optimisation and intelligent transition strategy, while the editorial team and the AI anchor work in collaboration to ultimately optimise the broadcasting tempo, standardise the language and improve the coherence of the video.

Based on this, human-computer collaboration promotes the transformation of the content production process from "artificial domination" to "human-computer collaboration", allowing the AI anchor to do what it is good at, and letting humans join in to do irreplaceable things, which not only ensures the high efficiency and low cost of content production, but also retains the humanistic and humane care of the programme content in human-computer interaction. In human-computer interaction, it not only ensures efficient and low-cost content production, but also retains the humanistic and humanistic care of the programme content, effectively optimizing the content production.

2.2 Robot anthropomorphism for trust reconstruction

The full text of the article must be typeset in single column. Perceived Value is a key variable in consumers' subjective judgement of "worthiness", and it is also the intermediary bridge between the audience and the AI to establish a trusting relationship. When users perceive more value in their interactions, trust is more likely to develop. This value consists of a combination of practical, emotional and social value. Through the deep integration of anthropomorphism and intelligence, AI anchors create a "human-like" image with personality, which significantly enhances the emotional value; at the same time, their intelligent and timely response strengthens the practical value, and the two synergistically amplify the audience's overall perceived value, which indirectly promotes the generation of trust.

AI anchors create a "quasi-social interaction" experience with the help of appearance, voice and behavioural simulation, so that the audience will see them as warm companions. Intelligent technologies such as natural language understanding, autonomous decision-making and feedback learning guarantee the quality of response and efficiency of interaction, further enhancing the practical value. For example, Hunan TV's digital person Xiaoying opened an account on Weibo, interacting with real people or virtual images with distinctive personas such as "confused ghost", "silly white sweet" and "incomplete voice", realising a leap from one-way communication to anthropomorphic social interaction. The leap from one-way communication to anthropomorphic social interaction. The conferral of identity and personality makes up for the emotional gap of AI, allowing cold machines to become "people", enhancing the double value of emotion and utility in instant interaction, and consolidating the trust of the audience.

All in all, AI anchors "empower the soul" with anthropomorphism to convey emotions, "empower" services with intelligence, and make machines compliant and responsive by responding to audience needs in real time. Driven by both emotion and utility, it enhances the perceived value and strengthens and expands the foundation of trust.

2.3 Human-machine division of labour promotes ethics

AI anchors are driven in parallel by two pipelines: the former digitises real people and the latter allows digital people to speak and interact with each other, relying on massive databases to rapidly produce content. As AI virtual anchors lack the value judgement and social moral consciousness of real anchors, the content they generate and broadcast, if misleading and false, is prone to trigger a crisis of social awareness and trust, seriously undermining the public's trust in the news media^[2] and violating the ethics of information authenticity. When technology is running wild, it is only the inclusion of people - that is, human-machine collaboration - that can keep the "steering wheel" of ethical authenticity in check. AI anchors should take the principle of "human value rationality dominance" as the principle, and establish a closed-loop governance mechanism of "human-machine division of labour + cultural gate-keeping" in the content production process, so as to promote the construction of ethical authenticity and construction.

On the content side, the mechanism of "human-machine division of labour + cultural gatekeeping" is implemented: AI undertakes low-context tasks such as data broadcasting and multi-language translation, while human beings master high-context creative rights such as value interpretation and emotional expression, so as to avoid the erosion of the right to cultural interpretation by technology; a cultural ethics review mechanism is established, and the post of "Human Ethics Establish a cultural ethics review mechanism and set up "human ethics editors" to review the cultural sensitivity of AI-generated content, such as the expression of historical events and the presentation of the image of vulnerable groups, to ensure that it is in line with mainstream values.

On the whole, the AI anchor in the content end and human division of labour, so that humans control the final rights of high-context content production, and then cultural audit to form the last line of defence of information generation, is

conducive to reducing the rate of generation of false content and attenuate the rate of presentation of sensitive information, thus transforming the AI anchor from the efficiency of the generation of tools for trust in the real ethical "dancers". This will transform the AI anchor from an efficiency generation tool to an ethical "co-dancer" who trusts the truth.

3 Free traversal - from unidirectional output to multidirectional interaction

3.1 Expanding audience choice

Traditional AI anchors follow the principle of linear narrative from scripting to content delivery, i.e., the AI writes lines and camera sequences in advance, and cannot dynamically adjust them according to real-time user feedback. The audience can only watch the programme in the order of beginning, development and end, and can only fast-forward/pause, not "point to reach", with very low autonomy and selectivity. Yet, with the rise of digital restructuring and immersive media, AI anchors can take advantage of digital restructuring technologies such as hyperlink mapping and real-time labelling to split live broadcasts into clickable "information rooms", so that audiences can choose their own paths and scenarios in the "node maps", expanding their viewing autonomy. The audience can choose their own paths and scenes in the "node map", expanding their viewing autonomy.

AI anchor fusion "node map" can effectively amplify the audience for the programme content viewing options, for example, Xinhua News Agency "new micro" in the Winter Olympics coverage of each short news, data map, athlete interviews are made into an independent node, the audience clicked on the hands of the anchor. The audience can click on the anchor's hand "ice pier" to instantly move to the athlete interview room, and then click on the medal list and back to the anchor desktop, the whole process of no loading wait; and also such as wisdom to enjoy the AI live broadcasting system allows businesses in the "Kitchen scene - Cyberpunk live room - Outdoor Camping" and other more than 10 3D scenes in a key switch; audience in the mobile phone end of the sliding screen can surround the anchor 360 ° view of the goods, to achieve the "free walking" choice.

In summary, the AI anchor integrates digital restructuring technology to break the linear structure from the content node, so that the audience can arbitrarily turn and jump and freely switch between scenes, and realise the double traversal of perspective and space. Under the guidance of hyperlink mapping and real-time labelling, the audience changes from a one-way information receiver to a chooser of scenes and perspectives, from passive to active, and enhances the audience's right to choose.

3.2 Enhancing audience participation

Real-life anchors, by virtue of their on-the-spot empathy, instantly speed up, pick up stems, and return pop-ups to complete their emotional labour, which enables viewers to consciously integrate themselves into the narrative. The traditional AI anchor is still stuck in the "script-execution" binary framework, following the offline script, lack of real-time semantic perception, resulting in lagging feedback, stereotypical tone, the audience due to the "emotional inaccessibility" and produce a sense of alienation, which in turn This leads to lagging feedback and stereotypical tone, which makes the audience feel alienated due to "emotional inaccessibility" and thus weakens their willingness to participate. In order to bridge the gap between technology and emotional needs, a "pop-up-driven real-time feeder engine" can be embedded in the AI anchor system to algorithmically analyse the emotions of the pop-ups and instantly recalculate the AI anchor's expressions and movements to form timely emotional feedback and enhance the audience's sense of participation.

The AI anchor detects and analyses the emotional state of the pop-up text through the algorithm, and cleverly adjusts the response phrases according to the proportion of emotional words, so as to make a prompt response to ensure a sense of interaction with the audience, and enhance the sense of participation in the programme in the interactive back and forth. Such as B station, AI anchor live, the background real-time traversal of the pop-up screen, when the "emo" type of negative vocabulary accounted for > 30%, the system will automatically trigger the anchor "comfort" expression and words branch. After the real-time feed engine was activated, the average length of stay increased from 7.2 minutes to 13.5 minutes, and the density of pop-ups increased by 42%, with a significant increase in high-frequency sentences such as "the anchor understands me", which systematically enlarged the sense of audience participation.

It shows that when the AI anchor can instantly "emotional echo" through the algorithm, the audience is no longer a passive recipient, but in the open-world live narrative to become a "co-author", free to traverse in the live broadcast. AI anchors are driven by real-time pop-up feedback to meet the audience's emotional needs such as "being seen", "being understood" and "being responded to" in a timely manner, and to enhance the audience's sense of participation.

4 The return of the spiritual rhythm - "inside three, outside four" with "soundscape on the ground"

4.1 Giving the scene real

"Mechanical reproduction can take a copy of the original to places where the original could not go A choral work performed in a concert hall or in the open air can be echoed in the living room.^[3] " With the advent of digital media, information, including light, sound, and text, can be converted into binary code, allowing for the mechanical reproduction of works of art, resulting in works that are decontextualised and lose the spirit of the 'here and now'. Algorithmising the four external skills of broadcasting, namely tone, rhythm, stopping and accenting, and letting them jump in real time with the environmental data, can give a sense of realism to the scene at the sound level, and then dissolve the rigidity of the mechanical reproduction of part of the AI anchor, so as to reproduce the spiritual rhythm.

AI anchor through the use of tone change in the scene, speed of speech changes and other external broadcasting techniques, with a variety of voice changes in the audience auditory level to produce stimulation, the voice to match the scene changes, to achieve the scenario "reproduction", to enhance the sense of realism of the programme scene. Such as New Year's Eve in 2024, "Hangzhou news broadcast" in the AI anchor "Xiaoyu" broadcast cold wave warning, the system synchronised access to the Meteorological Bureau anemometer data, when the gust > 7, AI in the "outdoor body temperature" before inserting 0.2 seconds of gas and tongue trill The audience in the headset "heard" the cold wind into the microblogging real-time pop-up screen brush screen "this sound out I shivered". The technology allows for a live translation of the internal techniques of podcasting, and for the first time an AI programme has a palpable 'here and now'.

In conclusion, integrating the external skills of podcasting into the algorithmic database, and changing the tone and speed of speech to work together and cooperate with the scene factors can, to a certain extent, make the speech flow of the AI anchors more natural and relevant to the context. AI anchor in the voice and the environment in line with the natural flow between, weakened the "man-machine" rigidity, in the sound level to enhance the sense of realism.

4.2 Awakening cultural memory

In the mid-20th century, Frankfurt School scholars Adorno and Horkheimer put forward the concept of "cultural industry" in the Dialectic of Enlightenment, stressing the gradual industrialisation of modern cultural production and consumption, and the formation of a large-scale cultural consumer market through the standardisation of processes and patterned reproduction, resulting in the lack of individuality of cultural products^[4]. In this theoretical perspective, although AI speech synthesis technology can achieve acoustic "perfection" with zero fatigue and zero out-of-tune sound, it is over-standardised and presents the characteristic of "batch replica", which weakens the role of emotional arousal of the differentiated contents of national and regional cultures. The sound is "perfect". If the dialect rhythm, national or regional characteristics of "sound symbols" are embedded in the AI speech model, the local auditory scene can be reconstructed in the acoustic dimension, thus partially alleviating the homogenisation dilemma in the context of the cultural industry, and enhancing the audience's cultural identity and emotional resonance.

AI anchors can integrate two technologies, "dialect voiceprint library" and "environmental sound coupling": the former extracts and reproduces the rhythmic characteristics of the dialect, while the latter embeds region-specific ambient sounds into the narrative as an acoustic background noise. Taking Shanxi Radio and Television's "Jinji Non-legacy Stall" as an example, the AI anchor, when talking about Pingyao Push Light Lacquerware, first calls the Jinzhong accent model to recite the original text of Ming Dynasty's "Painting and Decorating Record", then reduces the speed of speech at the semantic pause and superimposes the sound of Push Light Lacquer Polishing recorded on the spot, forming the multilayered soundscape of the ancient language, the local dialect and the sound of the craftsmen's operation. While personalised dialects and regionalised ambient sounds enrich the sound level, they also strengthen the local sense of national culture, enhance the audience's perception of local culture, and the audience's cultural memory is awakened to a certain extent by the "sound pattern".

It can be seen that the AI anchor through the "dialect voiceprint library" "environmental sound coupling" two major technologies, the national sound symbols cleverly translated into a perceptible cultural experience, so that the user in the rich, personalised hearing "back to the" cultural scene "with technical action to achieve the effect of cultural awakening. It allows users to "return" to the cultural scene in a rich and personalised listening experience, and achieves the effect of cultural awakening with technical actions.

4.3 Transmitting the spirit of China

The Chinese spirit is the core symbol of mainstream ideology, and its dissemination urgently needs to make use of the new media infrastructure to achieve all-weather, all-geographical precision penetration. the AI anchor, with its 24-hour online, zero error, and mass replicable technological features, meets this strategic dissemination demand. Specifically, the AI anchor can convert the "sense of object" in the internal skills of broadcasting into programmable commands, which

can be called up in the cloud in real time, and translate the spirit of China into audible and palpable acoustic events, so as to deliver the voice of China and the spirit of China in a more sincere way.

In practical applications, the AI anchor dynamically adjusts the acoustic perspective by instantly capturing the multidimensional portrait of the audience's geography, age, and interaction history: when facing the youth group, it moderately raises the fundamental frequency and shortens the breath; and when facing the silver-haired group, it lowers the fundamental frequency and lengthens the pause. For instance, Hunan Radio's "Youth Xiang Companion" May 4 special programme is aimed at users aged 18-24, and the AI anchor, when broadcasting "Please rest assured that the party is strong, I have a strong country", raised the base frequency at the end of the sentence by 25 Hz and compressed the breath by 15%, and the pop-up screen then swiped the screen with the phrase "breaking the defence". The AI anchor will raise the base frequency at the end of the sentence by 25 Hz and compress the breath by 15 per cent when broadcasting "Please don't worry about the party, the strong country has me".

In view of this, AI anchors have been upgraded from "information broadcasting tools" to "technical translators of mainstream values". It takes the internal skills of broadcasting as the skeleton of the algorithm, and completes the reassembly of objects in the cloud, not only transmitting the voice of China, but also translating the "Spirit of China" into a common emotional experience that transcends the barriers of age, geography, and media, so as to complete the in-depth embedding of the mainstream values in the digital era and the long-lasting resonance.

5 Conclusions

Currently, the optimisation of AI anchors unfolds along two main lines: one is "human-machine synergy" and the other is "technology-narrative coupling". In the first line, by writing the "human scale" into the algorithm, the AI anchor continuously calibrates the content production, enhances the audience's trust, and completes the self-correction under the framework of ethical supervision in three ways: human-computer interaction, machine anthropomorphism, and human-machine division of labour. The second main line takes "information node + perspective roaming + pop-up drive" as the technology suite, breaking the linear narrative and turning the audience from passive receivers to active co-constructors; the enhanced right to choose and sense of participation in turn generates high-resolution feedback data, further iterating the AI performance. At the same time, we algorithmise internal and external professional paths such as "tone", "rhythm" and "object sense", and make them jump in real time along with the environmental data, so that the programme scene is both realistic and emotionally intense. and emotional concentration. Technology iteration continues to approach the AI "anthropomorphic asymptote", in the material level to the audience to bring more free choice of nodes and more diversified listening experience at the same time, but also in the spirit of the level of the more authentic transmission of regional and national culture, more accurate to the audience to disseminate the spirit of China.

As Heidegger puts it, "Only what is real can bring us into a free relationship, that is, with that which, by its very nature, concerns us." The fascination with the pseudo-true makes people forget the image of the original, lose the ability to recognise the real and the false, and the meaning of human existence is annihilated in the flood of numbers encoded in 0s and 1s. AI anchors and other virtual reality should not just be a plaything to create the illusion of spectacle, in the future how to return more naturally to the poetic quality of life, awakening the numbness and lethargy of the morning bell and evening drum, still need to be explored.

References

- [1] Zhang, H., & Liu, D. (2025). From tool to partner: ethical transformation and power reconstruction of human-machine collaboration. *Young Journalist*, (07), 5-11.
- [2] Li, B., Xu, D., & Hu, L. (2023). Risks and responses to the rapid expansion of AI virtual anchor industry. *Solidarity*, (06), 44-47.
- [3] Benjamin. (2001). *The origin of German tragedy*. Beijing: Culture and Art Press.
- [4] Yang, Y. (2024). Advantages and strategies of short video communication of health science in medical journals. *Journalism Research Guide*, 15 (08), 247-250.