Reshaping the Sound Landscape: The Application and Development of AI in the Sound Creation of Chinese Podcasts in the Era of Intelligent Media

Lingyu Kuai[1,a,*]

[1]*School of Journalism and Communication, China West Normal University, No.1 Normal University Road, Shunqing District, Nanchong City, Sichuan Province, China*

a. 2911472972@qq.com

*corresponding author

Abstract: With the rapid development of intelligent media technology, AI has greatly enriched the creation means and expression forms of Chinese podcasts and reshaped the diversity and personalization of speech synthesis, audio editing and content generation, through the diversity and personalization of sound landscape. Through the specific application cases of AI in the creation of Chinese podcast voice, this paper analyzes its positive effects in improving the efficiency of creation, optimizing user experience and promoting the development of the industry, and looks forward to the future development trend and challenges, aiming to provide valuable reference and inspiration for the development of the field of Chinese podcast.

Keywords: artificial intelligence, Chinese podcast, AI technology, sound landscape.

1. Introduction

With the rapid development of intelligent media technology, we are living in an era of information explosion and diversified communication channels. Smart technologies are revolutionizing routine practice, affecting consumers, creators, and industry leaders. Similarly, AI technology is revolutionizing media products, including Chinese-language podcasts. As the number of users and programs grow across the board, podcasts are growing into an important force in the content track. According to the ListenNotes of the global podcasting industry, the number of Chinese podcasts increased sixfold in the three years from 2020 to 2023.[1] In China, Chinese podcasts, as a new form of audio media, have gradually won the favor of the majority of listeners with their unique content charm and flexible way of communication. However, in the face of the increasing content demand and the audience's desire for personalized experience, the traditional voice creation method has been difficult to meet the market demand. In this context, the rise of artificial intelligence (AI) technology has brought unprecedented opportunities and challenges to the sound creation of Chinese podcasts. With its powerful data processing ability, efficient automated operation and innovative creation mode, AI is gradually penetrating into every link of Chinese podcasts, from speech synthesis to audio editing, from content generation to personalized recommendation, showing great application potential.

2. From digital to intelligence: the encounter and symbiosis of AI technology and podcast voice creation

2.1. Definition of relevant concepts

Podcasting (Podcast) is a form of digital media in which users can download or stream audio files over the Internet. This form first appeared around 2004 and initially became popular as a combination of "iPod" and "Broadcast". Before the concept of podcasting returned to the Chinese Internet in 2020, audio books and pay-for-knowledge audio had already blossomed into people's lives. The difference between Chinese podcast and audio books and paid knowledge audio is that it is not reading to users, nor a paid voice course, but mainly about the conversation and sharing of one or more anchors, and the cultural topic is the mainstream content in the current podcast. With the change of the scene and the blessing of the technology, the current Chinese podcast is more like a kind of network audio based on individual expression and the output of personal views. With the introduction of artificial intelligence technology, audio and podcast programs, as the accompanying consumption experience, enter the market with intelligent appearance.[2] In China, 2016 is known as the first year of the outbreak of podcast and other audio streaming media, formed the "Himalaya FM", "Litchi FM" and "Dragonfly FM" three applications the situation, they strengthen the deep connection with intelligent speakers, make full use of artificial intelligence technology from digital podcast to intelligent podcast.

Recently, as a hot application in the field of artificial intelligence, ChatGPT has attracted wide attention from all walks of life. Today, AI is deeply applied in the audio industry, significantly driving audio content retrieval, content recognition, and content generation. Such as AIGC audio creation, AI anchor + AI continuation, ASR automatic voice recognition, SP audio processing, content recommendation, content authentication, etc. The core strengths of the AI-enabled audio industry include accuracy, high efficiency, customization, and accessibility. The new scene of audio communication under 5G and AI technology has formed an effective connection with the industrial chain. With the superposition of AI technology and the 5G era, the production and accurate distribution of content through user data mining has become the norm in the development of the current cultural industry.[3] Based on the rise of automatic language recognition, accessibility, cloud computing technology and natural language processing technology of intelligent speakers and voice assistant, the use of algorithm model and machine technology iteration to improve the accuracy of semantic recognition, but also can draw portraits according to user background data and accurately distribute audio information.[4] AI in the integration of text recognition, voice broadcast application, so that the audio book product content scene design has an advanced breakthrough and upgrade. At present, the development and wide application of artificial intelligence (AI) have brought new opportunities for the development of sound media.

2.2. The development of AI technology in enabling podcasts

Beginning. AI technology has begun initial attempts in audio processing, but it has not been directly applied to podcast content creation. At this stage, AI technology is mainly used for the research and development of basic technologies such as speech recognition and speech synthesis. In the market, the application of AI technology in the field of podcast is still in the cognitive stage, most people hold a wait-and-see attitude towards its potential value, and no AI technology has been directly applied to the creation of podcast content, but the application of AI in voice assistant, intelligent speaker and other fields has laid the foundation for the subsequent podcast integration.

Fusion exploration stage. AI technology began to deeply integrate with podcast content creation, such as using AI to generate podcast content, auxiliary editing, personalized recommendations, etc. At this stage, the podcasting content generated by AI technology, although limited in number, marks

the substantial application of AI in the field of podcasting technology.podcast.ai Is a podcast generated entirely by AI that digs into a new topic every week. In October 2022, an audio conversation between renowned American podcast host (Joe Rogan) and the late Apple founder Steve Jobs sparked a heated discussion on the Internet.[5] In the 20-minute podcast, the two explored multiple topics, including Jobs's college experience, insights into computers, and personal beliefs. The podcast audio was launched by podcast.ai, and to generate it, podcast.ai used Jobs' biography and all the online recordings about him, heavily trained through the Play.ht AI language model. In addition, the show host Logan's voice is also generated by AI.podcast.ai The AI podcast is a new exploration of generative AI in the field of voice.

Continued the development stage. With the maturity of technology and wide application, users gradually accept and get used to the application of AI in the field of podcasting. The application of AI technology in the field of podcasting is increasingly mature, and significant progress has been made in speech synthesis and natural language processing. AI plays an important role in the creation, distribution and promotion of podcast content and other aspects, such as automatic generation of podcast scripts, intelligent clips, personalized recommendations, etc. During this period, AI synthetic anchors gradually came into people's vision. These anchors can not only imitate the voices, expressions and movements of real people, but also interact and talk in real time. It has been widely used in news broadcast, product introduction, online education and other fields. While these cases are largely focused on news and entertainment, their success provides strong support for the further use of AI in podcasts. In July 2023, the domestic podcast program "Dazhong and Xiaoya" released a podcast completely AI-generated plots and voice, which received more than 5,000 times on the small universe. Some comments in the comment section said that the AI-generated voice was mistaken as the two anchors' "bad emotional state". In the podcast App small universe, the "Hacker News" account produced a program with the voice of "Xiaoxiao", which won the love and reward of the majority of listeners. Xiaoxiao is a woman in Azure's TTS (text to voice) voice library, Microsoft's cloud service platform.

3. Application: the collision of AI technology and Chinese podcast creation

3.1. Audio editing and processing

AI technology can automate the processing of large amounts of audio data, significantly improving editing efficiency. Through preset algorithms and models, AI can quickly identify key parts of audio, such as human voice, background sound, noise, etc., and make accurate editing to reduce the tedious and time-consuming of manual operation. In the process of creation, podcast creators will encounter many problems such as mouth addiction, breath mouth and saliva sound in the program. If you need to listen to all the long audio content every time, it will be time-consuming and laborious. AI "intelligent detection" function can help creators to detect mouth bite, identify air mouth, and delete the unnecessary bite bite, not having to edit word by word, saving a lot of time. At present, there have been a number of intelligent audio editing software based on AI technology on the market, which can automatically identify the key parts of the audio, and provide one-click editing, automatic mixing and other functions, greatly reducing the threshold and difficulty of audio editing. For example, Ximalaya's intelligent creation tool "Cloud clip" created by AI technology has launched many new functions, which greatly reduces the threshold of audio podcast content creation and improves the creative efficiency."Cloud editing" is a powerful online multi-track editing and light application, which integrates intelligent volume, intelligent music, audio to text editing, AI segmentation, intelligent detection, one-click film and other functions.[6] In the past audio clips, one or two hours of audio content needs to be heard from beginning to end, while listening and editing, which greatly consumes the creator's time. Cloud editing is also equipped with the function of "intelligent music

and intelligent volume", which facilitates creators with difficulty in choosing to select music with copyright and matching the content through AI. For audio volume, the volume can also be adjusted with a key, and fade in and out to improve the creative efficiency.

3.2. Text conception and creation

AI technology can help creators generate content, especially in text-to-voice translation. Creators simply input text, and the AI can automatically generate synchronized voice content, greatly saving time for recording and editing. In addition, AI can also adjust the expression mode of speech according to the emotion and context of the text, so that the dubbing is more in line with the needs of the content. This technology not only improves the efficiency of creation, but also lowers the threshold of content creation, enabling more people to participate in the creation of Chinese podcasts.

AI technology also helps podcast creators generate scripts and conceive program content, such as the "AI segmentation" feature to improve creation efficiency, and is one of the smart technologies that are highly praised by podcast creators. The audio content created by the podcast master contains a large amount of high-density non-visual information. The "AI segmentation" function automatically segments it according to the audio content, and generates a text summary outline for each audio segment. The creator no longer needs to spend time and effort to manually edit and organize, realizing the "shownotes" freedom. More than that, for text writing ability, but because of various reasons is not convenient to use voice expression creators, AI editing software and "text to voice" function, such as TTS (Text-to-speech, speech synthesis) technology, the Himalayan cloud clips can help through AI creators turn text into voice, for text works to audio track clear technical barriers, let more text creators to join audio and podcast creation field.

3.3. Sound editing and optimization

The basic principle of conversion from text to speech (TTS) is that AI technology converts the input text content into natural and smooth speech through deep learning algorithm. This process includes multiple steps such as text preprocessing, speech feature extraction, speech waveform generation and so on. Through deep learning algorithms, AI technology can simulate realistic human voices, providing a variety of voice choices for Chinese podcasts. This technology is not limited to the generation of a single timbre, but also can adjust the intonation, speed and emotion of the sound according to the needs of different scenes and characters, so as to make the content creation more vivid and personalized. For example, the creator function of "Xi Rhyme Workshop" launched by Himalaya platform provides more than 40 tones for creators to choose, so that one person can complete the rich audio programs of multiple character emotions.

As AI-generated voice and cloned voice become more and more real, some content creators are using AI technology to produce highly updated and informative voice broadcast content. For example, the anchor of the podcast program "Crossroads" once revealed in a program that an AI information program called "Fast Dao Radio Station", which was managed by the co-founder, only writes the script, and the voice part is completed by AI. The effect is quite natural. In the production of Chinese podcasts, the creator can directly input the script or copy into the AI voice synthesis system, and the system can quickly generate the corresponding voice content. This has greatly improved production efficiency, especially for podcasts that require a lot of dubbed content. AI technology can also play an important role in the post-production process of Chinese podcasts. Through the analysis and processing of audio data, AI can automatically identify and eliminate noise and noise, and improve the clarity and quality of audio. At the same time, the AI can also mix, balance the audio and other processing operations to make the sound effect more perfect. The application of these technologies

not only reduces the burden of post-production, but also improves the overall quality of Chinese podcasts.

3.4. Personality recommendation and interaction

Podcast program is a choice of content consumption, but due to the linear voice playback characteristics, trying to retrieve and locate their favorite podcast column is a troublesome problem for listeners. AI technology can be used to analyze the content of podcasts and provide insights into the topics discussed, the emotions expressed and the overall tone of the podcast, helping podcasts understand their audience and improve the content they produce. In addition, AI technology can collect and analyze data on user listening behavior, preferences, history, etc., so as to build personalized portraits of users. Based on these portraits, AI can recommend Chinese podcasts that suit their tastes and accurately push personalized content, which can help podcasts reach a wider audience and increase content engagement. For example, in order to help lost in the sound view of the audience accurately search and locate the content they are interested in, technology company Dexa launched a new podcast retrieval AI tool, let the user can ask questions about the robot, and then get the relevant podcast abstract and links, directly into the "most want to hear" progress node. The emergence of Dexa not only makes the search experience easier for podcasting listeners, but also opens up new possibilities for the development of the podcasting industry.

The AI system can monitor user feedback and interaction in real time, such as click-through rate, completion rate, comments, etc., and dynamically adjust the recommendation strategy according to these data. This dynamic adjustment mechanism helps keep the freshness and diversity of recommended content, while improving long-term retention. For example, Google has launched an independent podcast app, based on the complex algorithm and huge data support, which can not only recommend popular podcasts for users, but also provide accurate personalized recommendations according to personal preferences and listening habits. AI technology can also support the podcast platform to build closer community relations through data analysis, and provide suggestions and directions for community construction for the platform by analyzing the interaction mode and relationship network between the users. AI technology can conduct emotional analysis to identify the emotional reactions and attitudes of users when listening to podcasts, which can help podcast creators understand the emotional needs of the audience, so as to adjust the content and style of the creation, and improve the appeal and resonance of the program.

4. Worry: AI technology enables the problem of Chinese podcasts to be created

4.1. Technical bottleneck and breakthrough problem

Despite significant advances in natural language processing, there are limitations in the understanding of complex contexts and deep implications. The podcast content generated by AI technology is often imitated based on the existing data in the large language model, and it is difficult to produce independent, creative and unique content. This imitative content may be novelty to meet the audience's need for novelty and diversity. For example, the creation of literature and talk show podcasts all need the true feelings and unique perspective of the creators, while AI technology is still difficult to completely replace the human ability in emotional expression, creative conception and other aspects. Therefore, the podcast content generated by AI may be insufficient in creativity and inspiration, which can easily lead to the podcast content in semantic expression, and may even appear logical confusion.

In addition, how to ensure the seamless connection of human-machine collaboration between AI technology and human creators is an important issue. At present, AI technology may not fully understand the intentions and needs of human creators, leading to barriers to communication and

collaboration in the creative process. The complexity and high threshold of AI technology may limit its popularization and application in the creation of Chinese podcasts, and many creators may be unable to fully use AI technology to improve their creative efficiency and quality due to the lack of relevant technology and knowledge.

4.2. Content bubble and authenticity test

For many creators, their voice is their only identity, and the podcast voice is one of the most valuable resources for podcast creators. Since AI technology relies on a large amount of data for training, the content it generates may tend to imitate existing successful cases or popular elements, resulting in serious homogenization of podcast content and lack of novelty and diversity. Although AI technology has made significant progress in natural language processing, there are still limitations in understanding complex context and deep meaning, which will lead to AI generated podcast content is not clear, accurate in semantic expression, and even logical confusion.

AI voice generation is a double-edged sword that brings both convenience and risk to creators, because their voices can be abused by unauthorized users, causing the spread of misinformation. A foreign podcast show The Joe Rogan Experience has invited guests from all walks of life to have an in-depth dialogue. But in May, someone used artificial intelligence technology to create a fictional podcast that mimicked the conversation between the podcast's host and some guests who had never appeared on his show. The podcast, called Joe Rogan AI Experience, was produced by an Australian creative director, Hugo, who uses a text-to-voice platform that can clone any voice, and a ChatGPT-based conversation generator. In addition, Hugo produced several other episodes of podcasts mimicking guest conversations between Joe Rogan and Andrew Tate, and Steve Jobs. Hugo said he hopes to demonstrate the development and potential of AI technology, and to be alert to the dangers of fake content. Podcasts as public media, wrong content may bring serious consequences, leading to large-scale dissemination of misinformation. AI technology can generate realistic podcast content, but it also brings the challenge of authenticity discrimination. It may be difficult for listeners to distinguish AI-generated content from reality-created content, which not only affects the authenticity of podcast content, but also impacts the originality of podcast creators.

The strong IP attributes of the podcast sound content, as well as the highly personalized expression, are also issues that creators need to carefully consider when using AI to generate voice. These features require AI not only to convey information, but also to be able to imitate human emotions and intonation and build emotional connections with the audience. For creators who can freely express their views through oral English, a unique accent or intonation can instead form a differentiated advantage that helps to shape the creator's personal style. For example, in the AI podcast experiment of "Big vulgar and Small Elegant", many netizens feedback that the AI generated voice "reading" is obvious. In terms of sound generation, podcast content production pursues not only smooth reading, but also emotional emotion through sound, and enhances the immersion and emotional resonance of the audience. From Jobs' "Resurrection" podcast to the AI podcast experiment, one of the main controversies facing AI-generated voices is the lack of human voice cadence and emotion, such as the monotony and mechanization of sound.

4.3. Copyright dispute and attribution issue

As the popularity and application of AI technology expands, the number of AI-generated podcast content will increase significantly. This means that once the infringement occurs, the scope and scale of the infringement will be expanded accordingly, bringing greater challenges to copyright protection. The copyright ownership of the podcast content generated by AI technology is still controversial. On the one hand, some argue that the content generated by AI should be attributed to developers or

technology providers, because they create the algorithms and technologies that generate the content. On the other hand, it is also believed that the copyright of the generated content should belong to the actual users or creators of the content, especially when they have done a lot of editing and personalized creation on the basis of AI generation. At the same time, when generating the podcast content, the AI may unconsciously learn from or imitate the existing original works, leading to the infringement of the copyright of others. Because AI technology works by learning and mimicking large amounts of data, it is difficult to completely avoid similarity to existing works.

The rapid development of AI technology makes the technical means of copyright protection face new challenges, and the traditional copyright protection methods may be difficult to effectively cope with the complexity and diversity of the content generated by AI. The existing copyright legal system may not be fully adapted to the new situation and new problems brought about by AI technology. For example, how to define the originality of AI generated content, how to determine the subject of infringement, etc. In recent years, with the wide application of AI technology, the number of related copyright disputes has gradually increased. For example, the first case in the country was won because its voice was imitated by AI technology and used commercially. The court held that if the voice generated by AI is identifiable, it should be protected by law. Therefore, the court ruled that the defendant company to use its voice without the consent of the plaintiff constitutes infringement, and needed to compensate the infringed for the economic loss of 250,000 yuan and apologize. In the field of AI creation, copyright ownership and infringement are complex and changeable. The copyright ownership of AI-generated content also involves AI platforms, creators, material providers and other aspects, which requires detailed legal analysis and judgment. For podcast creators, understanding and following relevant laws and regulations, respecting originality, and making rational use of AI technology are the key to avoid copyright disputes. At the same time, the platform also needs to establish a sound copyright protection mechanism to actively prevent and timely stop infringement acts.[7]

4.4. Ethical imbalance and blurred subjectivity

The ability of AI technology to mimic and generate content similar to human creation has sparked an ethical controversy over whether AI-generated content constitutes original content. While AI is able to create novel content, it is often based on existing data sets and algorithmic logic, lacking inspiration and unique perspective for human creation, which may lead to a redefinition of the concept of originality and the ambiguity of the identity and rights of creators. AI is still limited by its algorithms and training data in terms of innovation. It may not be able to think creatively across fields, like humans, or to generate completely novel ideas without explicit guidance. How to balance the auxiliary role of AI technology with the subjectivity of human creation, and ensure the diversity and innovation of podcast content, are all issues that should be thought about.

The improved autonomy, versatility and understanding of AI make AI show the ability of "humanoid" and "superman" in language understanding, content generation and social service tasks, but at the same time, it also brings the problems of error or false information, algorithmic discrimination, ability risk and abuse. The decision-making process of AI technology is often complex and difficult to explain, resulting in a lack of transparency in the generation process and results of Chinese podcast content. Listeners may have difficulty understanding the algorithmic logic and data sources behind the content. Intransparency can trigger a crisis of confidence, making listeners question the authenticity and impartiality of the podcast content. When AI-generated podcasts go wrong (such as privacy violations, misleading listeners, etc.), liability is often difficult to define. In addition, the training and application of AI technology requires a large amount of data support, which may contain personal privacy information. If the data protection measures are not in place, it may lead to the risk of privacy leakage. Privacy leakage may infringe on personal rights and

interests, causing a crisis of social trust.[8] The AI system may be the target of hacker attacks, resulting in illegal acquisition or tampering of data, which will pose a threat to the authenticity and security of Chinese podcast content. AI technology may inherit the bias and discriminatory information in the data when learning and generating content, which may be unconsciously reflected in the podcast content, causing unfair treatment to specific groups, such as problems involving social contradictions and gender ambiguity, which will negatively affect the values and perceptions of listeners.

5. Development: the optimization of Chinese podcast voice creation enabled by AI technology

5.1. The road to compliance: the legal framework guarantees the podcast creation

With the continuous development of AI technology, the existing laws and regulations may not fully adapt to the new technology environment. Therefore, it is necessary to timely revise and improve relevant laws and regulations, such as the Copyright Law and the Personal Information Protection Law, to better protect the sound rights and interests and personal privacy. In 2019, the — Development of Responsible AI, proposed in 2019, stressed that respect for privacy, security and control as an important principle of AI. The Civil Code also stipulates the authorization and protection of the voice of natural persons, establishes the legal status of the right of voice, and provides a more solid legal foundation for sound creation.

It is the only way to promote the healthy development of Chinese podcasts to define the copyright ownership of AI-generated content and regulate the copyright authorization and protection of sound products. First of all, the scope of sound rights and interests should be defined, and sound rights and interests should include exclusive rights of sound recording, exclusive rights of sound use, exclusive rights of sound disposal and the protection of sound interests. [9] The clear definition of these rights can help to regulate the use of AI in sound creation. Secondly, the use of AI generated sounds should be regulated, and the recognition standards of AI generated sounds should be formulated. In view of the particularity of AI technology, the recognition standard of AI generated sounds should be formulated, and which AI generated sounds are identifiable, so as to be included in the scope of legal protection. These criteria can take into account the similarity of timbre, intonation, rhythm, and the recognition ability of the general public or related field audience. In addition, the introduction of technical means for infringement monitoring can better make the intelligent creation of Chinese podcasts on the security chain, such as the use of blockchain, digital watermark and other technical means to identify and track the sound generated by AI, so as to quickly locate the source of infringement and collect evidence in case of infringement.

5.2. Autonomy and sharing: the self-discipline convention to maintain the development of the industry

AI technology is not a shield and an umbrella for infringement. Faced with the dilemma of the use of AI technology standards, the industry should jointly formulate and establish industry standards and norms to ensure that the AI generation and auxiliary creation of podcast content reaches a certain professional level and audience acceptance. Through the establishment of industry self-regulatory organizations, the formulation and implementation of self-discipline standards, standardize the scope of AI technology use, authorization methods, copyright ownership and other key issues, and jointly maintain the healthy and orderly development of the Chinese podcast industry. For example, on July 15,2024, SPai, together with many podcast anchors, launched the "Chinese Podcast Voice Data Set" project, aiming to provide high-quality data sets authorized by anchors for the AI industry. This project will not only provide valuable resources for the training of TTS (text-to-speech) and ASR

(speech-to-text) large models, but also play an important role in the emotional dimension training of multimodal large models. Through this project, anchors will be able to gain profits in the licensing process, share the dividends of the rapid development of the AI industry, and curb the infringement of content copyright by some enterprises and developers. Through the joint participation of many podcast anchors, AI is used to evaluate the quality of podcast content, provide creative suggestions, and AI to assist content promotion decisions, help generate the promotion caliber and materials in various platforms, and promote the development of voice generation technology to a more natural and closer to the real human dialogue. At the same time, it can also build a AI technology resource sharing and data protection platform to promote the exchange and sharing of technological achievements, experience and lessons. Through the construction of the platform, the risk cost of the podcasting industry can be reduced, and the efficiency and competitiveness of the whole industry can be improved. The industry can also safeguard the legitimate rights and interests of the originator and the fair playing field of the industry.

5.3. Value return: humanistic rationality guides sound creation

In 2020, at a symposium of scientists, General Secretary Xi Jinping put forward the lofty concept that scientific and technological innovation must be oriented to people's lives and health. The purpose of scientific and technological innovation is to serve the people, which is the fundamental purpose of developing scientific and technological undertakings. People-centered means that technology is not an external blessing, but to help realize the internal connection between content and people, and the cultural behavior of participants. All parties in the field of AI should adhere to value rationality, establish science and technology ethics, and guide science and technology to be good, that is, the development of science and technology should be people-oriented, benefit mankind, and solve the problem of sustainable development with science and technology.[10] Through the regulatory algorithm bias to achieve technology naturalization, adhere to the intelligent communication ethics to reconstruct the man-machine relationship, fundamentally solve the information security problem from the system, and boost and standardize the healthy development of AI technology and audio industry.

Although AI technology can efficiently generate and process audio content, humanistic care and emotional connection are difficult to completely replace by AI. In the sound creation of Chinese podcasts, the humanized and emotional expression should be emphasized, and the real and warm emotions should be transmitted through the sound, so that the audience can feel the emotional resonance between people. In the process of AI enabling Chinese podcast sound creation, we should focus on improving the quality of content. By optimizing the presentation mode of audio content, improving the fluency of playback, increasing interactive functions and other ways, users' auditory enjoyment and participation are improved. At the same time, pay attention to user feedback and demand changes, timely adjust and optimize the product strategy and service mode. Establishing an effective user feedback mechanism is an important means to ensure the user experience. By setting up a customer service hotline, online message board, social media interaction and other ways, we can collect user opinions and suggestions, and timely respond to and process user feedback. This helps to understand user needs and market trends, and provide strong support for product optimization and upgrading. In addition, strict audit and screening mechanisms can be used to ensure that the audio content generated by AI conforms to the mainstream social values and moral standards, and to avoid the dissemination of vulgar, false and misleading content. As an important carrier of cultural communication, Chinese podcasts should assume the responsibility of guiding the positive value. In the process of content creation, attention should be paid to transmitting positive energy, promoting social integrity, and guiding the audience to form positive values and attitudes towards life. The expected Chinese podcast content in the future needs to match with the development of new

technology and its own cultural mission, make a deep value transformation, and return to the thinking and attention of the deep connection between audio and people.[11]

6. Conclusion

Under the diversified development of technology empowerment and man-machine symbiosis, the discussion of human subjectivity is gradually weakened, and people need to gradually find the equilibrium state of "man-machine coexistence" in the way of technology evolution.[12] By taking positive actions to address the risks associated with the technology, the parties work together to ensure that it is used in a responsible and ethical manner. The introduction of AI technology greatly enriched the sound of Chinese podcast creation means, through deep learning, natural language processing and audio generation and other advanced technology, AI can not only simulate the realistic voice, also can adjust speed, intonation, tone according to demand, make the podcast content more joint style and audience demand. The application of AI in the sound creation of Chinese podcasts has promoted the innovation and development within the industry. However, the application of AI in the sound creation of Chinese podcasts is not plain sailing. The limitations of technology, ethical considerations and audience acceptance all need to be in-depth discussed and solved. To maintain the awe of technology, constantly reflect and examine the role and positioning of AI in sound creation, to ensure the healthy development of technology and the harmonious coexistence of human civilization, which is also the future of artificial intelligence technology to empower the Chinese podcast industry.

References

[1] *Future period | Digital age to social, is it imagination or castles in the air? Sina News. Retrieved from https://cj.sina.com.cn/articles/view/5044281310/12ca99fde02001wlc5*

[2] *Gao, H. (2022). The ternary fusion and harmonious development of "human and machine objects" in intelligent communication. Young Reporter, (16), 28-30.*

[3] *Lan, G., Guo, Q., Wei, J., et al. (2019). 5G + intelligent technology: Building a new intelligent education ecosystem in the era of "intelligent +." Journal of Distance Education, 37(03), 3-16.*

[4] *Liu, Z. (2019). The causes of the rise of mobile audio from the perspective of media. News Front, (24), 36-38.*

[5] *Meng, W. (2021). The outlet of intelligent communication of audio media. China Radio, (12), 25-29.*

[6] *Podcast video trend is rising, sound marketing to expand new boundaries? Interface News. Retrieved from https://www.jiemian.com/article/822489.html*

[7] *Sun, Y. (2020). How to avoid the risk of infringement. Invention and Innovation (Big Science and Technology), (11), 30-31.*

[8] *The 2023 Himalayan Creators Conference was held, and the podcast became the new blue ocean for content creation. Sina Finance. Retrieved from https://finance.sina.com.cn/tech/roll/2023-02-15/doc-imyfukur2310286.shtml*

[9] *Wang, S. (2023). Interpretation and application of voice protection in the Civil Code. Laws, (06), 35-44.*

[10] *Xie, X., & Gao, J. (2021). Evolution mechanism and advanced mode of cultural industry driven by AI technology and system. Social Science Research, (02), 104-114.*

[11] *Yin, L., & Zhu, D. (2019). Intelligent development of sound media — New terminals, new applications, and new relationships. China Radio, (04), 31-35.*

[12] *Yu, J. (2023). The leakage and management of readers' privacy in the meta-cosmic environment — Based on the theoretical perspective of power trust of information flow. Library Research, 53(04), 10-19.*