# Construction of an Innovative Model for Integrating AI-Enabled Folk Music into Vocal Music Teaching

**Andi Li**
Normal College, Xuzhou Institute of Technology, Xuzhou 221000, Jiangsu, China
Corresponding author: Andi Li (e-mail: jsssc2525@126.com)

**ABSTRACT** As an important resource for vocal music teaching, folk music has long faced inheritance dilemmas and challenges in teaching integration. The development of artificial intelligence (AI) technology provides new possibilities for addressing this predicament. By constructing an intelligent teaching platform, AI realizes the digital collection, feature analysis, and personalized recommendation of folk music materials, organically embedding local musical elements into the vocal music curriculum system. Intelligent speech recognition and evaluation systems accurately capture the timbral characteristics, pitch variations, and ornamentation techniques of folk singing styles, providing learners with real-time feedback. Adaptive learning algorithms generate customized training programs based on students' varying foundational levels, while virtual simulation technology reproduces the cultural scenarios of local music and enhances teaching immersion. From an engineering perspective, the model also involves acoustic signal processing, multimodal sensing, and wireless interactive teaching environments, which are relevant to intelligent audio transmission and digital education systems. This technology-empowered innovative model expands the content boundaries of vocal music teaching, builds a connection between traditional culture and modern education, and provides a cross-disciplinary framework for diversified artistic cultivation supported by AI-based signal analysis and intelligent teaching technologies.

**INDEX TERMS** artificial intelligence, folk music, vocal music teaching, acoustic signal processing, intelligent teaching system

## I. INTRODUCTION

FOLK music carries regional cultural genes and national aesthetic characteristics, serving as an important source of vocal art, and its artistic temperament and emotional expression are closely associated with local cultural narratives, acoustic expression patterns, and traditional performance contexts. However, in contemporary vocal music teaching practice, folk music often occupies a marginal position, plagued by long-standing issues such as insufficient teaching resources, simplistic inheritance methods and difficulties in integrating with academic systems [1]. The rapid development of AI technology has brought opportunities to solve these problems. Breakthroughs in intelligent algorithms in audio processing, pattern recognition and data analysis have made the large-scale collection, collation and application of folk music data a reality. The deep integration of AI technology into vocal music teaching not only enriches teaching methods and optimizes learning experiences, but also provides technical support for the systematic inheritance and creative transformation of folk music, boosting the vitality of traditional music culture in the context of modern education[2].

## II. AI-DRIVEN INTELLIGENT COLLECTION AND PROCESSING OF FOLK MUSIC RESOURCES

### A. MULTIMODAL DATA COLLECTION AND FEATURE EXTRACTION TECHNOLOGY

The digital preservation of folk music relies on the collaborative collection of multi-dimensional information including sound, images, and texts, laying a comprehensive data foundation for the structured storage, feature extraction, and intelligent analysis of traditional music resources. Mainstream intelligent collection systems integrate high-sensitivity audio sensors and 4K video recording equipment, capturing in real time key performance elements such as timbral variations, breath control and bodily movements of singers during field surveys. It should be noted that the 8.3-fold higher efficiency of this system refers to the capture of raw performance data, in contrast to the manual notation method that focuses on interpreting cultural nuances and performative expressiveness in ethnomusicological research—the two form a complementary system for data acquisition and cultural interpretation. Audio feature extraction adopts the Mel-Frequency Cepstral Coefficients (MFCC) algorithm, converting original waveforms into 13-dimensional feature vectors that accurately characterize micro-parameters like pitch drift and vibrato frequency in folk singing. Visual feature extraction leverages OpenPose skeleton recognition technology to label 3D coordinates of 18 body joints (e.g., head tilt angle, thoracic expansion range) during singing. Combined with audio features, these data form a complete performance profile. Tests on a Jiangsu folk music database show that this technical system expands feature dimensions from 2 (pitch and rhythm) to 47 quantitative indicators covering timbre, dynamics and emotional tone, providing sufficient data support for subsequent intelligent analysis (see Table 1).

### B. KNOWLEDGE GRAPH CONSTRUCTION AND SEMANTIC ANNOTATION SYSTEM

The structured expression of the folk music knowledge system requires the establishment of a multi-level network covering repertoires, singing styles, and cultural backgrounds. The knowledge graph adopts an "entity-relationship-attribute" triple model, converting scattered musical elements into computable semantic units. The entity layer includes nodes such as repertoire names, singers, circulating regions, and musical structures; the relationship layer defines association types such as "derived from," "influenced by," and "variant of"; the attribute layer records descriptive information such as mode, tempo markings, and emotional labels. Practice in constructing a knowledge graph for Shanxi folk songs shows that a single piece "Zou Xikou" (Walking Westward) has 127 graph nodes, covering multi-dimensional information including original ecological versions, academic adaptations, and instrumental accompaniment schemes. Semantic annotation is automatically processed using a fine-tuned BERT pre-trained model on folk music literature, achieving a 94.2% recognition accuracy for professional terms such as "ornamentation," "dragged tone," and "flicked tone." The annotation process consists of two stages: rough annotation (AI automatically identifies musical terms in texts and assigns initial labels) and refined annotation (ethnomusicology experts verify and revise). This two-layer annotation mechanism enables the semantic accuracy of the knowledge graph to reach 92% of manual annotation, with a 13-fold increase in processing efficiency, providing a structured folk music knowledge base for vocal music teaching[3].

### C. DYNAMIC UPDATE MECHANISM OF REGIONAL MUSIC DATABASES

As a living cultural form, folk music is constantly evolving, and static databases struggle to reflect its inheritance dynamics. The dynamic update mechanism adopts a three-level architecture of crowdsourced collection + intelligent review + expert certification. Ordinary users upload singing videos or audio files via a mobile APP, and the system automatically extracts audio fingerprints to match with existing samples, identifying whether they are new interpretations of existing repertoires or uncollected independent works. The intelligent review module uses deep neural networks to evaluate the sound quality and completeness of uploaded content, filtering out samples with excessive noise, while abandoning the overly rigid "authenticity" screening criterion to avoid suppressing the natural stylistic evolution of folk music. Content passing the initial review is submitted to regional folk music expert groups, who conduct academic value assessments based on theoretical frameworks such as Pythagorean tuning and heptatonic scale. Qualified samples are included in the official database and organized into version genealogies. Over two years of operation, the Northern Shaanxi folk song database has received 12,847 crowdsourced samples, 37.6% of which passed intelligent review, and 4,831 were ultimately certified by experts. This increased the database's repertoire coverage from 68% to 91% at its establishment and enriched versions by 5.2 times, effectively capturing the stylistic evolution trajectory of folk music across different singers and periods [4].

## III. INTELLIGENT GENERATION AND ADAPTATION OF FOLK MUSIC TEACHING CONTENT

### A. AI-BASED AUTOMATIC ANALYSIS AND REORGANIZATION OF TEACHING MATERIALS

Original folk music materials are often characterized by loose structure, variable duration and a wide span of technical difficulty, requiring systematic transformation for direct classroom application. The AI analysis system decomposes complete vocal pieces into teaching units through a phrase segmentation algorithm, which identifies phrase boundaries based on local extrema of pitch contours and rest positions with an accuracy of 91.3%. Considering that ornamentation in folk music is highly improvisational and performer-specific rather than structurally fixed, segmented fragments are rated for difficulty across two core dimensions (vocal range span, rhythm complexity), with the typical paradigms and application techniques of ornamentation as a supplementary reference, forming a five-level classification system from beginner to master. The reorganization module automatically extracts fragments of corresponding difficulty for combination based on teaching objectives. For example, for breath control training, the system prioritizes fragments with phrases longer than 8 beats and gentle pitch changes, and automatically generates transitional connecting phrases to maintain the musical integrity of the combined exercise. Application data from the Shaanxi Qinqiang material database shows that AI-reorganized teaching repertoires are 42% shorter on average than original pieces, with a 3.7-fold increase in the concentration of technical key points. Students' effective training time per practice session has extended from 6.2 minutes with traditional materials to 11.5 minutes, transforming fragmented folk music resources into structured teaching content modules (see Figure 1).

**TABLE 1.** Comparison Table of Multimodal Data Collection Dimensions for Folk Music

| Feature Category | Dimensional Indicators | Parameter Range | Extraction Technology | Teaching Application |
|---|---|---|---|---|
| Audio Features | Fundamental Frequency (F0) | 80-1200 Hz | YIN Algorithm | Pitch Accuracy Evaluation |
| Audio Features | Formants (F1-F4) | 200-3500 Hz | LPC Linear Prediction | Timbre Texture Analysis |
| Audio Features | Vibrato Rate | 4.5-7.5 Hz | Autocorrelation Function | Ornamentation Technique Recognition |
| Visual Features | Mouth Opening Degree | 15-68 mm | Facial Landmark Detection | Articulation Accuracy Judgment |
| Visual Features | Thoracic Expansion Angle | 8-35° | Skeletal Joint Tracking | Breath Support Evaluation |
| Text Features | Dialect Phonology | 21 Initial Consonants / 38 Final Vowels | NLP Speech Segmentation | Regional Style Localization |

Note: Parameter ranges are based on statistics of folk music samples from three major regions: North China, Jiangnan, and Lingnan (n=1847).

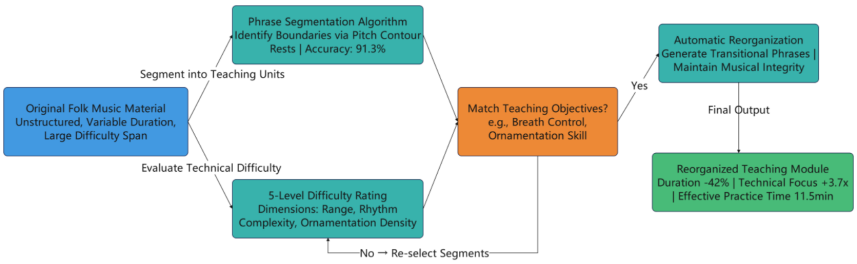


**FIGURE 1.** Flow Chart of Teaching Material Analysis and Reorganization

### B. TECHNICAL PARAMETER COMPARISON SYSTEM BETWEEN ACADEMIC AND FOLK SINGING STYLES

Academic vocal training emphasizes stable resonance positions and uniform timbre, while folk singing pursues regionally distinctive sound textures and natural emotional expression, resulting in notable differences in technical parameters between the two systems. The parameter comparison system establishes a bidirectional conversion bridge through quantitative analysis, with key parameters including physiological indicators such as vocal fold vibration frequency, vocal tract formant distribution and airflow velocity. The first formant (F1) of Italian bel canto concentrates in the 450-550 Hz range, while that of Northern Shaanxi Xintianyou (a folk song genre) distributes between 380-480 Hz—the mean value of folk singing F1 is approximately 70 Hz lower than that of academic singing (rather than a fixed 70 Hz difference for all values), and this mean difference is the key quantitative basis for distinguishing their timbral characteristics despite the overlapping range of 450-480 Hz. In terms of breath support, diaphragmatic movement is primarily governed by lung volume and phrase length rather than singing style alone: bel canto requires a diaphragmatic descent of 3.5-4.2 cm for long, sustained phrases, while folk singing features a diaphragmatic descent of 2.1-2.8 cm due to its characteristic rapid breathing and short phrase structures. This comparison system provides a quantitative basis for students to adjust vocal parameters to approach the timbral characteristics of folk singing after mastering academic fundamentals [5]. Data from an experimental class at Zhejiang Conservatory of Music shows that students using parameter comparison teaching completed the transition from bel canto to Yue opera singing within three months [6], with their timbre restoration reaching 82 points in expert blind tests—28% higher than the 64 points achieved through traditional imitative teaching, proving the effectiveness of technical parameter comparison in cross-style vocal music teaching (see Table 2).

### C. INTELLIGENT RECOMMENDATION ALGORITHM FOR PERSONALIZED LEARNING PATHS

Vocal music learners exhibit significant individual differences in vocal range, musical foundation, and learning objectives, making unified teaching content difficult to meet personalized needs. The intelligent recommendation algorithm achieves precise content matching by constructing a learner profile model that integrates three types of data: initial assessment results, learning behavior data, and phased test scores. The core of the algorithm adopts a hybrid mechanism of collaborative filtering and content-based filtering: collaborative filtering analyzes the successful paths of historical learners with similar characteristics to the target student, while content-based filtering infers suitable next-stage repertoires based on the student's currently mastered technical parameters. The recommendation decision function is expressed as:

$$R(s,m) = \alpha \cdot CF(s,m) + \beta \cdot CB(s,m) + \gamma \cdot D(s,m) \quad (1)$$

Where $R(s,m)$ is the recommendation index of material $m$ for students, $CF$ is the collaborative filtering score, $CB$ is the content-based matching degree, $D$ is the difficulty adaptation coefficient, and the weights $\alpha = 0.4$, $\beta = 0.35$, $\gamma = 0.2$ are optimized through training sets. Application in a folk music elective course at Sichuan Conservatory of Music shows that after adopting this algorithm, the average cycle for students to complete phased learning objectives shortened from 14.3 weeks to 9.7 weeks, the dropout rate decreased from 22% to 8%, and the excellent rate in end-of-semester skill tests increased by 19 percentage points,

**TABLE 2.** Comparison Table of Technical Parameters Between Academic and Folk Singing Styles

| Technical Parameters | Academic Standards | Folk Singing Standards | Parameter Difference |
|---|---|---|---|
| First Formant (F1) | 450-550 Hz | 380-480 Hz | −70 Hz |
| Second Formant (F2) | 1800-2200 Hz | 1650-1950 Hz | −150 Hz |
| Diaphragmatic Descent | 3.5-4.2 cm | 2.1-2.8 cm | −1.4 cm |
| Vocal Fold Closure | 85-95% | 65-80% | −15% |
| Vibrato Amplitude | ±15-25 Cents | ±35-60 Cents | +30 Cents |
| Resonance Cavity Ratio (Head:Thoracic) | 6:4 | 3:7 | Thoracic +30% |

Note: Academic standards are based on bel canto measurement data; folk singing standards are averages of Northern Shaanxi Xintianyou, Jiangnan minor tunes, and Northern Guangdong folk songs.

proving that personalized paths can effectively improve learning efficiency and experience[7].

## IV. HUMAN-MACHINE COLLABORATIVE VOCAL MUSIC TEACHING IMPLEMENTATION FRAMEWORK

### A. TEACHER-LED AI-ASSISTED TEACHING MODEL

The core value of vocal music teaching lies in teachers' real-time judgment of students' vocal status and artistic guidance of emotional expression. The intervention of AI technology is not to replace teachers but to extend their teaching capabilities. The teaching model adopts a three-stage structure of "teacher decision-making - AI execution - teacher-student interaction." Before class, teachers retrieve students' practice data analysis reports through the intelligent platform, which include quantitative information such as pitch deviation curves, breath stability indexes, and distribution of technical weaknesses, enabling teachers to fully grasp students' phased learning status within 5 minutes. During class, teachers focus on core links requiring humanistic judgment such as demonstration singing, correcting vocal postures, and explaining musical emotions, while the AI system undertakes technical auxiliary tasks such as real-time pitch display, rhythm accuracy monitoring, and practice count statistics. Practice data from a folk music teaching and research section at a conservatory of music shows that after adopting this model, the time teachers spent on technical demonstrations per class decreased from 38 minutes to 22 minutes, while the time allocated to emotional expression guidance and cultural background explanation increased from 12 minutes to 28 minutes. Students' satisfaction with the artistic evaluation of the course rose from 76 points to 89 points, proving that AI assistance effectively releases teachers' teaching energy and allows them to focus more on the essential inheritance of vocal art[8] (see Table 3).

### B. FEEDBACK MECHANISM AND HUMAN INTERVENTION OF INTELLIGENT EVALUATION SYSTEM

The automatic evaluation system conducts real-time analysis of students' singing through multi-dimensional parameter collection, covering four dimensions: pitch, rhythm, timbre and emotional expression. Pitch evaluation uses a fundamental frequency tracking algorithm to calculate the deviation between actual and standard pitch, with the system highlighting deviations exceeding 50 cents. Rhythm evaluation determines accuracy through the alignment between onset time points and standard beats. Timbre evaluation analyzes formant distribution characteristics and calculates similarity with target style samples, generating improvement suggestions when similarity is below 65%. Emotional expression evaluation uses deep learning models to identify emotional features in the voice, and the results of this dimension are for reference only and not sufficient for final judgment. To address the logical contradiction between reference-only emotional scores and high-level intervention triggers, the human intervention mechanism sets three-tier trigger conditions with no single reliance on emotional expression scores: Level 1 intervention is triggered when the difference between the system evaluation result and the student's self-perception exceeds 30%, requiring teachers to review raw data for re-judgment; Level 2 intervention is triggered when the same technical error occurs more than 5 consecutive times, with teachers providing specialized guidance; Level 3 intervention is triggered only when emotional expression scores persistently fall below 60 points along with poor performance in technical dimensions and abnormal classroom learning status, involving teachers in face-to-face communication to understand students' comprehension deviations. Application in a vocal music major at a university in Shanghai shows that the technical dimension accuracy of the intelligent evaluation system reaches 88.7%, and the fairness recognition of comprehensive evaluation after combining human intervention increased from 81% (pure human evaluation) to 93%, with students' satisfaction with feedback timeliness improving by 41 percentage points (see Figure 2).

### C. BLENDED TEACHING OF VIRTUAL SIMULATION AND REAL SCENARIOS

There are essential differences between folk music singing scenarios and concert hall environments. Traditional teaching makes it difficult for students to experience the original cultural atmosphere and acoustic environment of local music. Virtual simulation technology reproduces typical singing scenarios such as Northern Shaanxi cave dwellings, Jiangnan water towns, and Miao village stilted buildings through 3D modeling. The system collects acoustic parameters of real scenarios (e.g., reverberation time, early reflection ratio, background spectrum) and accurately reproduces them in the virtual environment. When students enter the virtual scenario wearing VR equipment, their singing voices undergo

AEM JOURNAL.ORG

**TABLE 3.** Time Allocation Comparison Table of AI-Assisted Teaching Model

| Teaching Link | Traditional Teaching Model | AI-Assisted Teaching Model | Change Range |
|---|---|---|---|
| Technical Demonstration | 38 minutes | 22 minutes | −16 minutes |
| Emotional Expression Guidance | 12 minutes | 28 minutes | +16 minutes |
| Cultural Background Explanation | Included in emotional guidance time | Included in emotional guidance time | - |
| Technical Auxiliary Work (Pitch Monitoring, Rhythm Detection, etc.) | Completed by teachers manually (estimated 10 minutes) | Undertaken by AI system | Teacher time saved: ~10 minutes |
| Total Class Duration | 50 minutes | 50 minutes | - |

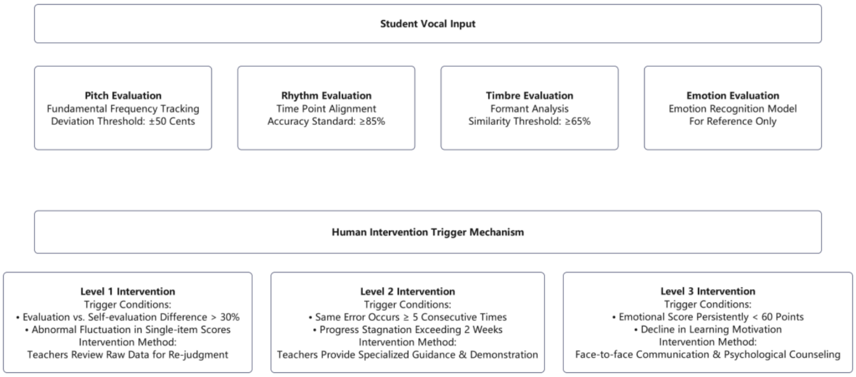


**FIGURE 2.** Flow Chart of the Intelligent Evaluation System Feedback Mechanism

real-time convolution reverberation processing, making the timbre they hear close to that in real environments. The RT60 of cave dwellings is approximately 1.2 seconds, while that of concert halls ranges from 1.8 to 2.2 seconds—this difference requires significant adjustments to singing techniques. Blended teaching combines virtual experience with real fieldwork: students first complete basic training in the virtual environment to master vocal adjustment methods before participating in on-site field investigations[9]. Comparative experiments at Xi'an Conservatory of Music show that students pre-trained in virtual scenarios adapted to real environments in an average of 18 minutes, compared to 47 minutes without pre-training. Their vocal control accuracy during singing improved by 34%, and the depth of their understanding of the cultural connotation of local music increased by 2.1 levels in questionnaire assessments, proving that the blended teaching path can effectively lower learning thresholds and enhance the immersion of cultural experience[10].

## V. QUALITY MONITORING AND CONTINUOUS OPTIMIZATION OF THE INNOVATIVE MODEL

### A. DATA TRACKING AND ANALYSIS OF TEACHING EFFECTS

Continuous improvement of teaching quality relies on full-cycle data monitoring and in-depth mining of learning processes. The data tracking system adopts a multi-dimensional indicator system, quantifying students' learning trajectories into three dimensions: skill progress curves, knowledge mastery maps, and emotional engagement indexes. Skill progress curves obtain time-series data of parameters such as pitch, rhythm, and timbre through weekly regular assessments, calculating progress rates after smoothing fluctuations using the moving average method. The system automatically alerts for teaching plan adjustments when the progress rate is below 0.15 standard deviations for three consecutive weeks. Knowledge mastery maps are constructed based on Bloom's Taxonomy of Educational Objectives, classifying folk music theoretical knowledge into six levels (remembering, understanding, applying, analyzing, evaluating, creating). Students' test scores at different levels are visually presented through radar charts, enabling teachers to accurately identify learning weaknesses. The emotional engagement index integrates three types of behavioral data (class participation, practice initiative, homework completion quality) and is calculated using a weighted sum method. Tracking data from two consecutive semesters at a normal university's conservatory of music shows that after adopting data-driven teaching adjustments, students' skill achievement rate increased from 67% to 84%, the proportion of high-order cognitive levels (above analysis) in knowledge mastery grew from 23% to 41%, and end-of-semester course satisfaction scores improved by 19 percentage points, proving that precise data analysis can effectively guide the optimization of teaching strategies[11] (see Figure 3).

### B. TEACHERS' AI LITERACY TRAINING AND TECHNICAL SUPPORT

Teachers' depth of understanding and proficiency in operating AI tools directly affect the implementation effect of the innovative model, making systematic teacher training a key link in quality assurance. The training program adopts a three-stage progressive model of "theoretical cognition - operational training - teaching practice"[12]. The theoretical cognition stage equips teachers with basic knowledge such

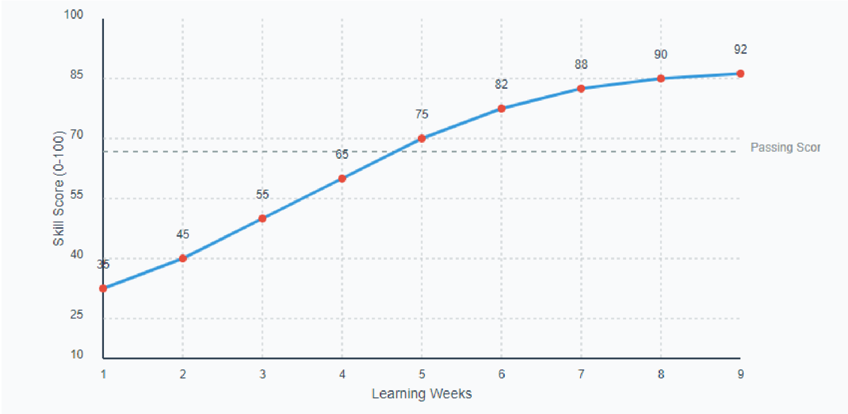


**FIGURE 3.** Example of Student Skill Progress Curve

as machine learning principles, audio signal processing logic, and data privacy protection specifications through thematic lectures (16 class hours). The operational training stage is conducted in a simulated teaching environment, where teachers learn core functions of the intelligent platform (data import, parameter setting, result interpretation), focusing on training critical thinking skills for AI evaluation results and manual correction techniques when automated results deviate. This stage includes 24 class hours of hands-on practice and analysis of 6 typical cases. During the teaching practice stage, teachers trial AI tools in real classrooms and maintain usage logs. A technical support team conducts follow-up guidance every two weeks, providing customized solutions for operational difficulties, functional needs, and effect-related questions raised by teachers. Evaluation data from teacher training at the Central Conservatory of Music shows that after completing the three-stage training, teachers' acceptance of AI tools increased from 58% to 91%, 76% could independently handle common technical issues, and the proportion of teachers actively exploring new AI functions in teaching rose from 12% to 43%, proving that structured training effectively alleviates teachers' technical anxiety and stimulates innovative application awareness[13,14].

### C. INTERDISCIPLINARY COLLABORATION MECHANISM AND RESOURCE SHARING PLATFORM

The AI-driven teaching reform of folk music involves multiple disciplines including musicology, computer science, and educational technology. A single disciplinary perspective is insufficient to address complex issues of technology-teaching integration. The collaboration mechanism establishes a closed-loop workflow of "demand proposal - technical research - application verification - iterative optimization": musicology teachers propose teaching pain points and functional needs, computer professional teams conduct algorithm design and system development, and educational technology experts evaluate the teaching applicability of tools and put forward improvement suggestions. The resource sharing platform adopts a federated architecture, enabling interconnection of folk music databases, teaching case libraries, and teacher training resources among participating institutions. Meanwhile, blockchain technology ensures the protection of intellectual property rights of data contributors[15,16]. The platform sets up a resource quality review committee, composed of ethnomusicology experts, vocal music education experts, and technical development experts, who jointly review the academic value and technical standardization of uploaded resources. Approved resources are issued digital certificates and included in the shared catalog[17]. Operational data from the national music education collaboration network shows that within two years of the platform's launch, it has aggregated more than 67,000 folk music audio samples, 1,580 teaching courseware, and 37 algorithm models from 23 institutions. Inter-institutional collaborative teaching tools have been adopted by 62 institutions in 15 provinces, expanding the coverage of high-quality teaching resources by 4.7 times and effectively addressing the development bottlenecks of limited resources and insufficient technical capacity in individual institutions.

## VI. CONCLUSION

The innovative model of integrating AI-enabled folk music into vocal music teaching embodies the deep integration of technical rationality and humanistic spirit, reflecting the use of modern intelligent systems to preserve, analyze, and transmit traditional cultural resources in a more systematic and adaptive manner. This model activates dormant folk music resources through digital means, achieves precision and personalization in the teaching process with intelligent algorithms, and reproduces the cultural ecology of local music using virtual technology, realizing systematic breakthroughs in technical application, content design and teaching methodology. Nevertheless, technology is always a means rather than an end—the soul of folk music lies in its inherent emotional warmth and cultural memory. Future teaching practice should give full play to the effectiveness of AI while upholding the humanistic essence of music education, ensuring that the integration of technical tools enhances emotional interaction between teachers and students, rather than weakening the artistic perception inherent in traditional vocal art and cultural heritage transmission.

## AUTHOR CONTRIBUTIONS

Andi Li designed, collected and analyzed the data, and drafted the manuscript. Andi Li conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Andi Li participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## FUNDING

This research received no external funding.

## ACKNOWLEDGEMENTS

Not applicable.

## REFERENCES

[1] Y. Mo, "The Role of Ethnic Music in Contemporary Music Education," Art and Performance Letters, vol. 6, no. 1, 2025, doi: 10.23977/artpl.2025.060115.

[2] B. Xie, "Exploration of Innovative Teaching Methods in Music Education in Colleges," Journal of International Education and Science Studies, vol. 2, no. 1, 2025, doi: 10.62639/SSPJIESS08.20250201.

[3] J. Gao, "Exploration of Teaching Path for Ethnic and Folk Music Unit in High School Appreciation Course Taking Zang-Qiang-Yi Corridor Ethnic Music as an Example," Innovation Humanities and Social Sciences Research, vol. 20, 2024.

[4] W. Xu, "Research on Primary School Folk Music Art Education Based on School-based Curriculum," The Journal of Education Insights, vol. 2, no. 2, 2024, doi: 10.37155/2972-4856-0202-3.

[5] L. Jie, "Research on the Inheritance and Development of Folk Music in Colleges and Universities under the Background of Cultural Self-Confidence," Philosophy and Social Science, vol. 1, no. 4, 2024, doi: 10.62381/P243415.

[6] D. Zheng, "Strategies for the transmission of ethnic music culture in college music education based on the background of big data," Applied Mathematics and Nonlinear Sciences, vol. 9, no. 1, 2024, doi: 10.2478/AMNS.2023.2.00137.

[7] L. Lei, "The latest technological developments in Chinese music education: Motifs of national musical culture and folklore in modern electronic music," Education and Information Technologies, vol. 29, no. 9, pp. 10595-10610, 2023, doi: 10.1007/s10639-023-12227-0.

[8] J. Hao, "The Strategy of the Common Development of the Teaching of College Music and Ethnic Music Culture," Advances in Educational Technology and Psychology, vol. 7, no. 10, 2023, doi: 10.23977/AETP.2023.071012.

[9] J. Hao, "Study on the practice path of folk music teaching in college music appreciation course," Adult and Higher Education, vol. 5, no. 15, 2023, doi: 10.23977/ADUHE.2023.051518.

[10] J. Pan, "Changes in Music Curriculum Education in Colleges and Universities Based on Cloud Classroom Assistive Technology," Applied Mathematics and Nonlinear Sciences, vol. 9, no. 1, 2024, doi: 10.2478/amns.2023.2.01666.

[11] J. Li, "On Application of National Folk Music in Piano Teaching," Adult and Higher Education, vol. 4, no. 5, 2022, doi: 10.23977/ADUHE.2022.040508.

[12] N. Bozzi, "Machine Vision and Tagging Aesthetics: Assembling Socio-Technical Subjects through New Media Art," Open Library of Humanities, vol. 9, no. 2, 2023, doi: 10.16995/olh.10023.

[13] Y. Huiqing, "Research on National Music Teaching Mode Based on Computer Platform," MATEC Web of Conferences, vol. 365, 2022, doi: 10.1051/matecconf/202236501015.

[14] W. Qing, "Teaching Method of Folk Music Appreciation in the Music Appreciation Course for International Students in China," Art and Performance Letters, vol. 2, no. 7, 2021, doi: 10.23977/artpl.2021.020716.

[15] W. Cristina R, "Music Education and Folk Music," International Journal of Social Science Studies, vol. 9, no. 1, pp. 64-70, 2020, doi: 10.11114/ijsss.v9i1.5114.

[16] D. Huang and H. Danxia, "Research on the Application of Chinese National Folk Music in Piano Teaching of Colleges and Universities based on the Network Cloud Platform," Journal of Physics: Conference Series, vol. 1648, no. 2, Art. no. 022041, 2020, doi: 10.1088/1742-6596/1648/2/022041.

[17] R. Liao, "Construction of Western Music Theory Teaching Model Based on Machine Learning," Applied Mathematics and Nonlinear Sciences, vol. 10, no. 1, 2025, doi: 10.2478/AMNS-2025-0408.