

Искусственный интеллект в звуковом сопровождении виртуальных обучающих сред: современные подходы и тенденции

Урынбаева Азада Бахадировна, Кокандский Государственный университет
urinbaevaazada@gmail.com

ABSTRACT

В статье представлен обзор современных технологий искусственного интеллекта, применяемых в звуковом сопровождении виртуальных обучающих сред. Рассматриваются основные направления: синтез речи, голосовые ассистенты, адаптивное озвучивание, генеративная музыка и мультязычная озвучка. Подчеркивается роль ИИ в повышении доступности и интерактивности учебного процесса, а также улучшении эмоционального восприятия материалов. Представлен анализ ключевых технологий и их образовательного значения.

This article presents an overview of modern artificial intelligence technologies applied in audio support for virtual learning environments. The main directions covered include speech synthesis, voice assistants, adaptive audio feedback, generative music, and multilingual voice-over. The role of AI in enhancing accessibility, interactivity, and emotional perception of educational content is emphasized. A detailed analysis of key technologies and their educational significance is provided.

Ushbu maqolada virtual ta'lim muhitlarida ovozli qo'llab-quvvatlash uchun sun'iy intellektning zamonaviy texnologiyalari ko'rib chiqilgan. Asosiy yo'nalishlar qatoriga nutqni sintez qilish, ovozli yordamchilar, moslashuvchan ovozli javoblar, generativ musiqa va ko'p tilli ovoz berish kiradi. Sun'iy intellektning ta'lim jarayonining mavjudligini oshirish, interaktivligini kuchaytirish va materiallarning hissiy qabul qilinishini yaxshilashdagi roli ta'kidlangan. Muhim texnologiyalar va ularning ta'limdagi ahamiyati tahlil qilingan.

ARTICLE INFO

Received: 30th April 2025

Accepted: 26th May 2025

KEYWORDS:

искусственный интеллект, виртуальное обучение, синтез речи, голосовые ассистенты, адаптивное озвучивание, генеративная музыка, мультязычная озвучка. artificial intelligence, virtual learning, text-to-speech, voice assistants, adaptive audio feedback, generative music, multilingual voice-over. sun'iy intellekt, virtual ta'lim, nutqni sintez qilish, ovozli yordamchilar, moslashuvchan ovozli javoblar, generativ musiqa, ko'p tilli ovoz berish.

Введение. Современное образование все активнее интегрирует виртуальные обучающие среды (ВУС), позволяя создавать интерактивные, доступные и персонализированные курсы. Звуковое сопровождение является важным компонентом таких сред, повышая эффективность усвоения материала и вовлеченность пользователей. Искусственный интеллект (ИИ) значительно расширил возможности создания адаптивного, естественного и мультимодального звукового контента, что актуализирует необходимость систематического обзора современных подходов.

Методология. Для написания статьи был проведен анализ научных публикаций, технической документации и отраслевых отчетов, опубликованных в период 2016–2023 годов. Основное внимание уделялось исследованиям и продуктам, связанным с применением ИИ в синтезе речи, голосовых

ассистентах, адаптивном озвучивании, генеративной музыке и мультязычной озвучке. Результаты систематизированы с учетом образовательных приложений.

Результаты

Современные технологии ИИ в звуковом сопровождении виртуальных обучающих сред включают пять основных элементов:

Синтез речи (Text-to-Speech, TTS). Современные TTS-системы на базе нейросетевых моделей, таких как Tacotron 2 и WaveNet, обеспечивают высокое качество и естественность речи. Они поддерживают разнообразие голосов и эмоций, что улучшает восприятие материала, помогает слабовидящим и способствует изучению языков.

Голосовые ассистенты. Голосовые ассистенты обеспечивают интерактивный диалог с пользователем. Используя модели NLP (GPT, BERT), они способны распознавать голосовые команды, поддерживать диалог с учетом контекста и адаптировать подачу материала, повышая мотивацию и вовлеченность.

Адаптивное озвучивание. ИИ анализирует поведение и эмоциональное состояние пользователя, подбирая оптимальное звуковое сопровождение: одобрение, подсказки или поддержку. Такой подход помогает удерживать внимание и снижает когнитивную нагрузку.

Генеративная музыка. ИИ-композиторы создают уникальную музыку, адаптированную под учебный процесс. Фоновая музыка улучшает концентрацию, снижает тревожность и создает нужную атмосферу, способствуя лучшему усвоению знаний.

Мультязычная озвучка. ИИ-системы обеспечивают автоматический перевод и озвучку на множество языков с сохранением интонаций и эмоций. Это расширяет доступ к образовательным ресурсам для международной аудитории и помогает в изучении языков.

Интеграция ИИ в звуковое сопровождение обучающих сред открывает новые горизонты для персонализации и доступности образования. Однако вопросы этики, качества перевода и эмоциональной адекватности остаются предметом дальнейших исследований. Развитие мультимодальных систем с комбинированием аудио и визуальных элементов будет способствовать более глубокому вовлечению обучающихся.

ИИ-технологии существенно трансформируют звуковое сопровождение виртуальных обучающих сред, делая обучение более интерактивным, доступным и эмоционально насыщенным. Адаптивное озвучивание и генеративная музыка создают уникальные образовательные условия, а мультязычные решения способствуют глобализации образования.

Список литературы

1. Wang, Y., Skerry-Ryan, R., Stanton, D., Wu, Y., Weiss, R. J., Jaitly, N., ... & Saurous, R. A. (2017). Tacotron: Towards end-to-end speech synthesis. *arXiv preprint arXiv:1703.10135*.
2. van den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., ... & Kavukcuoglu, K. (2016). WaveNet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*.
3. Microsoft Azure Cognitive Services. (2023). Speech Service Documentation. Retrieved from <https://learn.microsoft.com/en-us/azure/cognitive-services/speech-service>
4. Amazon Web Services. (2023). Amazon Polly Developer Guide. Retrieved from <https://docs.aws.amazon.com/polly>
5. Bocklisch, T., Faulkner, J., Pawlowski, N., & Nichol, A. (2017). Rasa: Open source conversational AI. *arXiv preprint arXiv:1712.05181*.
6. Huang, Z., Kim, H., & Smaragdis, P. (2019). Music source separation with deep learning: A review. *IEEE Journal of Selected Topics in Signal Processing, 13*(6), 2063-2079. <https://doi.org/10.1109/JSTSP.2019.2922027>
7. Briot, J.-P., Hadjeres, G., & Pachet, F. (2020). Deep learning techniques for music generation – A survey. *ACM Computing Surveys, 54*(5), Article 93. <https://doi.org/10.1145/3397191>
8. OpenAI. (2020). Jukebox: A generative model for music. Retrieved from <https://openai.com/research/jukebox>

9. UNESCO. (2023). Artificial intelligence and the future of teaching and learning: Policies and practices for sustainable development. Paris: UNESCO Publishing.
10. Duolingo Research. (2022). Personalization in language learning using AI. Retrieved from <https://research.duolingo.com>
11. Koryakovskiy, A., & Timofeev, A. (2022). Adaptive audio systems in e-learning: A review of approaches and technologies. *Journal of Educational Technology & Society*, 25*(2), 45-57.
12. Lee, J., & Oh, S. (2021). Multilingual text-to-speech synthesis using neural networks. *IEEE Transactions on Audio, Speech, and Language Processing*, 29*, 1586-1598. <https://doi.org/10.1109/TASLP.2021.3052347>
13. Zhang, Y., & Chen, C. (2020). Emotion recognition in speech using deep learning for adaptive feedback in education. *Computers & Education*, 147*, 103789. <https://doi.org/10.1016/j.compedu.2019.103789>
14. Park, J., & Lim, S. (2019). Generative models for personalized music creation in learning environments. In *Proceedings of the 27th ACM International Conference on Multimedia** (pp. 266-274). ACM. <https://doi.org/10.1145/3343031.3350905>