Applied Science Innovation (ASI)
Published by College of Physical and Applied Sciences, MOUAU

- *Received:* 10 May 2026
- *Revised*: 18 May 2026
- *Accepted:* 11 June 2026
- *Published:* 19 June 2026

MODEL OF A GAN-GENERATED DIGITAL HUMAN FOR REPLACEMENT OF RETIRED UNIVERSITY ACADEMIC STAFF

Prisca I. Okochi*[1], Enefiok A. Etuk[2], and Ndidiamaka N. Eluwa[3]

[12]Department of Computer Science, Michael Okpara University of Agriculture, Umudike Abia State, Nigeria.
[3]Department of Geology, Michael Okpara University of Agriculture, Umudike Abia State, Nigeria.

***Corresponding author: okochi.prisca@mouau.edu.ng**

ABSTRACT

The retirement of experienced university academic staff often results in a significant loss of institutional knowledge, pedagogical expertise, and instructional continuity. This study proposes a comprehensive model for a Generative Adversarial Network (GAN)–generated digital human designed to serve as a functional replacement for retired university professors. The proposed system integrates speech recognition, large language models (LLMs), curriculum-driven knowledge retrieval, and GAN-based facial video synthesis to simulate human-like teaching interactions. The model captures the instructional behavior of retired academics through curated curriculum repositories and conversational intelligence, enabling real-time interaction with students via voice, text, and visual modalities. The architecture combines automatic speech recognition, semantic document embedding, retrieval-augmented generation, text-to-speech synthesis, and GAN-driven video generation to deliver a realistic digital lecturer. This research contributes a novel educational technology framework capable of preserving academic legacy while addressing staff shortages and scalability challenges in higher education.

Keywords: Digital Human, GAN, Artificial Lecturer, Educational AI, Large Language Models, Knowledge Preservation

1. Introduction

Universities worldwide face increasing challenges related to the retirement of senior academic staff, including knowledge discontinuity, instructional gaps, and reduced mentorship capacity. Retired professors often possess decades of accumulated subject mastery, teaching methodology, and contextual academic insight that are difficult to replace through conventional hiring alone.

Recent advances in artificial intelligence particularly Generative Adversarial Networks (GANs), (Goodfellow et al., 2014), large language models, and multimodal interaction systems present an opportunity to preserve and replicate academic expertise digitally (Vaswani et al., 2017). A GAN-generated

digital human can act as a virtual lecturer, capable of answering student questions, delivering lectures, and visually engaging learners in a human-like manner.

Digital humans have been explored in entertainment, healthcare, and customer service domains. Facial animation using GANs and 3D Morphable Models (3DMMs) has demonstrated significant realism in lip synchronization and expression generation. Meanwhile, conversational AI systems powered by transformer-based LLMs have shown strong performance in educational tutoring and dialogue-based learning environments **(Karras et al., 2020; Liu, 2023).**

However, existing systems typically operate in isolation either focusing on speech, text, or visuals. Few studies propose an integrated academic replacement model that combines curriculum-based knowledge retrieval, conversational intelligence, and realistic facial synthesis tailored for higher education.

This research presents a structured model for a GAN-generated digital human specifically designed to replace retired university academic staff while maintaining pedagogical integrity, curriculum alignment, and interactive realism. It bridges that gap by proposing a unified system architecture that supports real-time academic interaction using a GAN-generated digital professor.

2. System Architecture

The pipeline includes ASR (OpenAI, 2022), semantic retrieval (Reimers & Gurevych, 2019), FAISS indexing (Johnson et al., 2019), LLM reasoning, TTS, and GAN-based facial synthesis.

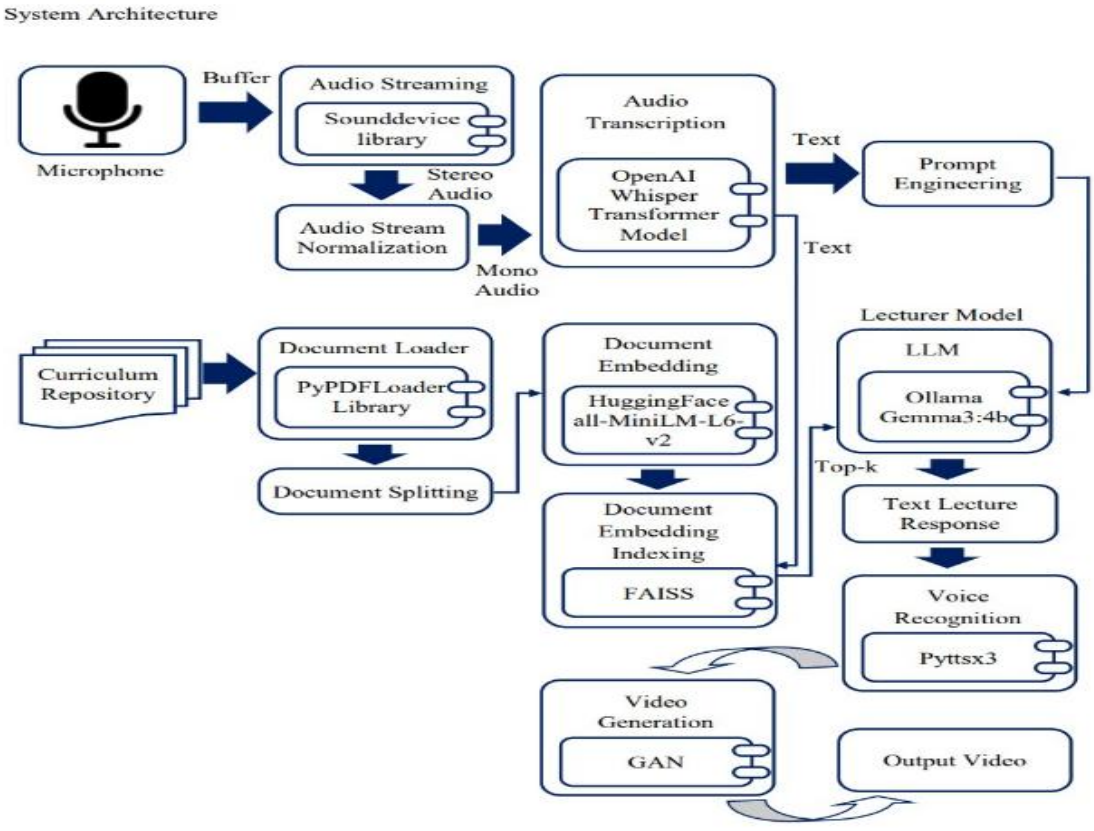


Fig. 1: System Architecture

Figure 1 illustrates the complete system architecture of the proposed GAN-generated digital human framework. The architecture integrates speech input, curriculum-based knowledge retrieval, conversational reasoning, and GAN-driven video synthesis to simulate the instructional behavior of retired academic staff.

The architecture consists of interconnected modules responsible for audio input capture, speech transcription, semantic document retrieval, language modeling, voice synthesis, and GAN-based facial video generation. The system ensures that student queries are processed through curriculum-aligned knowledge sources before generating both spoken and visual responses.

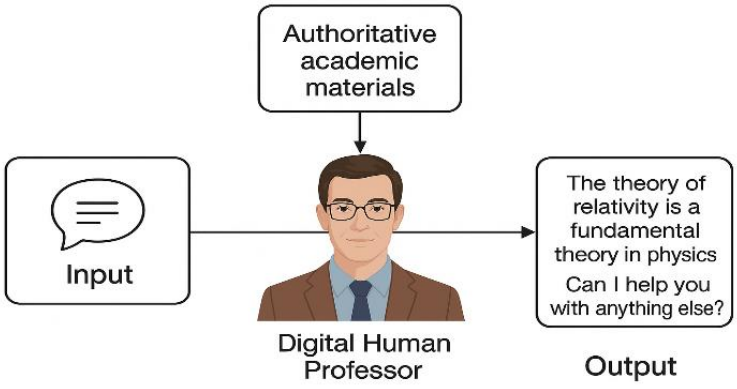


Fig. 2: The integration of semantic retrieval with LLM

Figure 2 illustrates the integration of semantic retrieval with the Large Language Model (LLM), forming a retrieval-augmented generation framework for academic response synthesis. The LLM receives both the student query and the retrieved academic content. This ensures that generated responses remain contextually accurate, pedagogically relevant, and aligned with the retired professor's expertise.

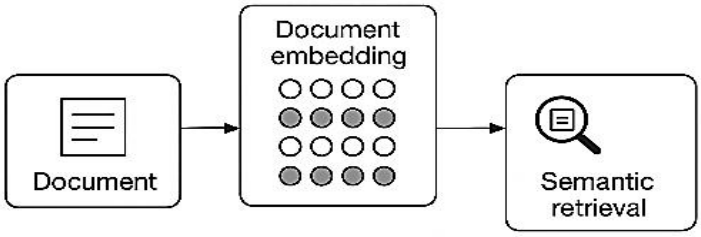


Figure 3: Document embedding and semantic retrieval

Figure 3 shows the document embedding and semantic retrieval mechanism used to ground the digital professor's responses in authoritative academic materials. Curriculum documents are converted into dense vector embeddings using a transformer-based embedding model. These vectors are stored in a FAISS index to allow efficient similarity search during real-time interactions.

3. System Overview

The model consists of six tightly coupled layers:

- User Interaction Layer
- Speech Processing Layer
- Knowledge Management Layer
- Conversational Intelligence Layer
- Speech and Facial Synthesis Layer
- Output Delivery Layer

Together, these layers enable a retired academic's knowledge to be digitally preserved and interactively delivered.

User Interaction Layer – Students interact with the system through a user interface supporting voice-based input. This interface mimics traditional classroom questioning, allowing natural spoken queries rather than typed commands.

Speech Processing Layer – Voice input is captured and processed using a speech recognition engine. Audio normalization ensures consistent quality, after which a transformer-based speech-to-text model converts speech into structured text. This step enables the system to understand student queries accurately, even in noisy environments.

Knowledge Management Layer – Academic knowledge is stored within a curriculum repository, containing lecture notes, textbooks, presentations, and research materials previously used by retired staff. These documents are segmented and embedded using semantic vector representations. A vector database enables fast retrieval of relevant content using similarity search techniques. This ensures responses are grounded in authoritative academic material rather than generic language generation.

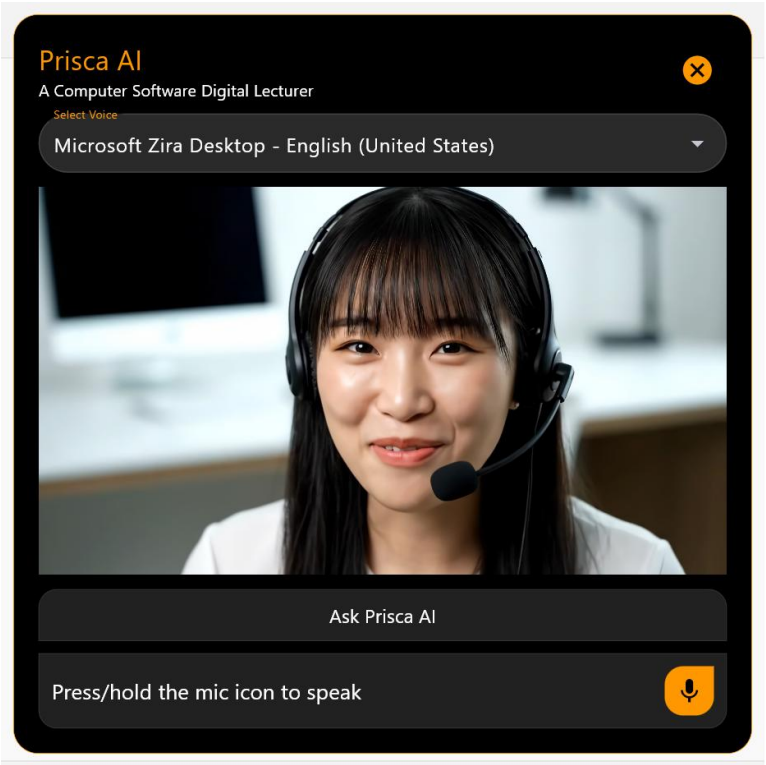
Conversational Intelligence Layer – The Large Language Model (LLM) functions as the digital lecturer's reasoning core. It combines:

- Transcribed student queries
- Retrieved curriculum content
- Prompt engineering rules that preserve academic tone and instructional style

This retrieval-augmented generation approach ensures responses remain factual, contextual, and pedagogically appropriate.

Speech and Facial Synthesis Layer – Generated text responses are converted into natural speech using text-to-speech synthesis. The speech signal then drives a GAN-based video generation module that produces synchronized facial expressions and lip movements. The GAN learns realistic facial dynamics from training data, enabling the digital human to visually resemble a professor delivering a lecture. This multimodal output significantly enhances student engagement.

Output Delivery Layer – The final output is a video stream of the digital academic speaking in real time. Students perceive the system not merely as a chatbot, but as a responsive virtual lecturer capable of delivering explanations, answering questions, and sustaining dialogue.



4. Application - Use Case Scenario

A student asks a question during an online class session. The system:

- Captures the voice input
- Transcribes the speech
- Retrieves relevant academic material
- Generates an instructional response
- Synthesizes speech and facial animation
- Presents a video of the digital professor answering the question

This process occurs in near real-time, closely resembling a live academic interaction.

4.1. Evaluation Criteria

The effectiveness of the proposed model may be evaluated using:

- Response accuracy and curriculum relevance
- Speech recognition accuracy
- Facial realism and lip-sync precision
- Student satisfaction and engagement metrics
- System latency and scalability

5. Discussion and Conclusion

The proposed model demonstrates how retired academic expertise can be preserved digitally without compromising educational quality. Unlike static recorded lectures, the system enables dynamic interaction, adaptive explanations, and continuous availability. Ethical considerations include transparency, data consent from retired staff, and clear distinction between human and digital instructors. When responsibly deployed, the system complements not replaces human educators.

In conclusion, this research presented a comprehensive model for a GAN-generated digital human designed to replace retired university academic staff. By integrating speech recognition, curriculum-aware conversational AI, and GAN-based facial synthesis, the system delivers an interactive, realistic, and scalable academic solution. The model offers a viable pathway for preserving institutional knowledge, enhancing access to education, and addressing staffing challenges in modern universities.

Acknowledgements

We acknowledge the services of Dr. C. E. Osodeke and Mrs Chioma for ensuring that all necessary grammatical corrections were made.

Conflict of Interest

The authors declare no conflicts of interest.

References

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial nets. Advances in Neural Information Processing Systems, 27*, 2672–2680.

Johnson, J., Douze, M., & Jégou, H. (2019). *Billion-scale similarity search with GPUs. IEEE Transactions on Big Data, 7*(3), 535–547.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). *Analyzing and improving the image quality of StyleGAN. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8110–8119.

Liu, Q., Geertshuis, S., & Grainger, R. (2023). *Digital knowledge repositories and lifelong learning in higher education. Education and Information Technologies, 28*(2), 1925–1943.

OpenAI. (2022). *Whisper: Robust speech recognition via large-scale weak supervision. arXiv preprint arXiv:2212.04356.*

Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence embeddings using Siamese BERT-networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 3982–3992.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need. Advances in Neural Information Processing Systems, 30*, 5998–6008.