

MemorialBot: Voice Cloning And Personality-Driven Conversational Framework For Post-Humorous Digital Interaction And Grief Support

Ms. Akanksha Jadhav, Ms. Trupti Patil, Ms. Pratiksha Borate, Mr. Tushar Kharat, Prof. P. J. Patel

Department of Artificial Intelligence and Data Science Engineering, Jawahar Education Society's Institute of Technology and Research, Nashik, India

Email: akankshaj295@gmail.com

Email: truptipatil18788@gmail.com

Email: boratepratiksha62@gmail.com

Email: tushark1247@gmail.com

Email: pooja.patel1387@gmail.com

Abstract:

Mourning the death of a loved one represents one of the greatest trials of the human experience, and the desire for an enduring connection with them is an intensely personal, unfulfilled aspiration. MemorialBot is an conversational AI that seeks to satisfy this yearning by allowing mourners to interact via natural voice conversation with an artificial reconstruction of a deceased loved one, derived from audio recordings, text documents, and contributions of biographical narratives from the user and loved one. The proposed system architecture comprises an integrated speech-to-speech pipeline for automatic speech recognition (ASR), retrieval-augmented large language model (LLM) inference, neural text-to-speech (TTS) generation, and voice cloning. The performance evaluation included twenty participants who rated voice naturalness at 4.2/5.0, persona consistency at 82% accuracy, and perceived comfort with the system at 4.1/5.0 on a 5-point Likert scale. The paper describes the complete system design, method of implementation, experimental evaluation, and an ethical framework considering consent, data privacy and protection, and emotional safety. MemorialBot offers contributions to the rapidly developing fields of affective computing, digital heritage conservation, and AI support for grief.

Keywords — Voice Cloning, Digital Legacy, Grief Support, Conversational AI, Large Language Models, Affective Computing, Neural TTS.

I. INTRODUCTION

When someone we love passes away it changes the world around us forever. The death is final, but our urge to connect with the dead is a constant—an urge satisfied by revisiting photographs, scribbled letters, voice recordings and video accounts over and over as ways of remembering. But these keepsakes are flat. They cannot answer, react or interact with our questions or feelings. But

the development of sophisticated artificial intelligence, specifically large language models and neural voice synthesis, means that it is technically feasible to build machines which can do something much more powerful: mimic in real time the voice, personality, and information of the deceased.

Current forms of digital memorialization—archived social media profiles, static memorials, and recorded videos—are passive and unable to respond to new queries and changing emotions of the user. These static forms can never fulfill the

deep-seated need for an interactive memorial. Also the current grief support systems are failing to meet the challenges combining personalization, voice authenticity and emotion awareness. The generic chatbots that are readily available are unable to reconstruct the language, beliefs, and communication pattern of an individual, and the majority of the voice synthesis systems need a considerable amount of high-quality audio data which is not present in the actual memorial setting. So we need a tailored system that can cover all the aspects of the issue simultaneously.

In response to these problems, this paper introduces MemorialBot-a persona-driven conversational, voice cloning system that supports a grieving community with post-mortem digital conversation. This system takes voice recordings and other textual (letters) or verbal (testimony) contributions provided by family and friends of the deceased to generate a persona knowledge base, and a personal voice profile, allowing for live, voice based conversation reflecting the personal history and personality.

The main contribution of this paper is to illustrate that a responsible, yet empathetic and emotionally comforting, artificially created agent is achievable. MemorialBot utilizes standard, well-understood modules in order to maintain ethical standards in simulating a deceased individual. Our goal is to show a principled, accessible, and technically sound example of this new kind of affect-infused computing application.

II. LITERATURE REVIEW

Research in digital memorialisation, affective computing, and conversational AI has developed significantly in recent years, yet no prior system has integrated all the necessary components into a cohesive, ethically governed framework. The literature informing MemorialBot spans three interconnected domains, each of which has advanced independently without producing a

comprehensive solution to the problem of posthumous digital interaction

The first domain concerns large language models and persona-conditioned dialogue. Studies have established that transformer-based architectures can perform diverse natural language tasks with minimal task-specific training, validating their use as a foundation for personalised conversational agents. More recent work has shown that open-weight models can achieve comparable performance at lower computational cost, broadening the range of deployment environments available to developers working on sensitive applications such as grief support systems.

The second area is in neural voice cloning and synthesis. Recent papers have developed zero-shot multi-speaker text to speech models that are capable of generating natural sounding voice recordings from 3-5 seconds of reference speech. The work in neural voice cloning is what is essential for MemorialBot to recreate the acoustic profile of a deceased person from a relatively small corpus. The companion work that used generative adversarial network based solutions shows that despite short, imperfect samples, it is possible to achieve a high speaker similarity. This addresses the problem of the practical restriction that there would rarely be large amounts of recorded speech.

Human-computer interaction: grief and affective systems is the third domain. Empirical studies of AI systems with emotional responsiveness mechanisms revealed a considerable increase in perceived empathy during highly personal support situations. A review of the literature on self-disclosure with conversational AI reveals important therapeutic potential, balanced with a risk of over-dependency which needs to be accounted for in design. Studies on affective responsiveness indicated that calibration is essential for user trust and engagement with conversational AI during emotive situations.

What has been written also shows the importance of privacy and consent governance for any processing of a deceased's personal data.

Existing studies on digital heritage preservation already points to legal and ethical issues that arise with post-mortem data, especially under changing personal data protection regimes.

We conclude that while many studies have worked on individual components (voice cloning, persona-conditioning of LLM, affective dialogue and digital identity), none have combined the components and created a system within an ethically grounded structure that was dedicated solely to a grief support setting. The design of the system MemorialBot was consequently driven by this unaddressed human requirement.

Based on the literature reviewed, it is evident that the literature has been done in individual parts which include digital identity, accessibility, and job matching, but there has been no comprehensive solution that has incorporated these components into a single system. This gap inspires the creation of the improved Kamgar Connect platform that will provide an integrated, user-centric solution specifically designed to meet the needs of the informal workforce.

III. PROPOSED SYSTEM

Proposed is MemorialBot, a persona-driven, voice-activated conversation system aiming to fulfill the unmet need of personal, interactive digital remembrance. The focus of MemorialBot's design lies in emotional honesty, accessibility and responsible governance in a market of non-emotional, automated, and unchecked chatbots, recognizing the needs of grieving individuals.

1. System Overview: Cloud hosted MemorialBot allows registered users to provide audio recordings, documents and spoken memories. These memories are compiled into a personalized knowledge base and vocal profile of the deceased individual. During live interactions, users interact through a web/mobile interface with natural spoken language.

2. The important parts of the System:

a) Voice Cloning and Acoustic Profiling:

Upon on-boarding users will provide existing audio recordings of their loved one's voice. These audio files are passed through a voice cloning engine to pull speaker embedding (containing the user's intonational, timbral, and rhythmic characteristics). The individual speaker embeddings will be averaged together over all provided recordings. This creates an aggregate and stable acoustic profile which is stored safely for all future text to speech generations in this memorialization session.

b) Personal Knowledge Base Construction:

The system also allows for the ingestion of a diversity of personal materials, among which:

- Voice recordings that provide the input for acoustic model training.
- Written materials (letters, messages, emails, diaries).
- Spoken evidence, or verbal recollections, from close relations and friends.
- Biographical text which captures the history, values, and relations of the person.

c) Retrieval-Augmented LLM Inference:

- User query is transcribed and semantically embedded.
- The 5 most relevant memory chunks are retrieved from the knowledge base.
- A templated persona-conditioned prompt is sent to the language model.

d) Neural Text-to-Speech Synthesis:

This text response is then fed into a neural TTS system whose output is conditioned on the speaker embedding stored previously. This produced speech has the specific acoustic characteristics of the person who passed away and is outputted to the user in the form of audio playback.

e) Sentiment-Aware Response

Conditioning: The sentiment analysis method is performed for every user request in order to capture the emotive tone of the query. If the analysis indicates sadness or suffering, the system's language model will attempt to favor more compassionate language in its responses. This is to ensure that the system responds in a way that is compatible with the user's emotional state and is not in an emotionally flat way.

f) Accessible User Interface: The interface is designed to be accessible for a non-technical audience, utilizing:

- Easy, voice-first navigation
- Mobile responsive layout usable on common devices
- Permanent AI disclosure notifications informing the user they are speaking to an artificial system.

g) Ethical Safeguards and Governance:

The system also provides:

- Requires written advance consent prior to the establishment of any memorial profile.
- Fully end-to-end encrypts all personal data including voice data and session logs.
- Monitors session lengths and frequency to discourage over-attachment.

3. System Architecture (Conceptual Flow):

Step 1: User Onboarding and Submission of authorized data.

Step 2: Voice Processing and speaker embedding.

Step 3: Live Session initiation.

Step 4: Semantic retrieval and Persona-conditioned LLM response.

Step 5: user feedback, rating and unethical review.

4. Advantages of Proposed System:

- i) Interactive experience versus static remembrance.
- ii) The acoustic profile of the deceased is accurately captured, speaker likeness is very high.
- iii) Every response relies on factual personal details, limiting hallucinations.
- iv) Response is conditioned on speaker's emotional state.
- v) Privacy and consent are assured through an overarching ethical framework.

IV. METHODOLOGY

The approach that this research took is one based on implementation: the MemorialBot system was constructed, implemented, and then measured against a specific benchmark. MemorialBot is envisioned as a conduit linking user-provided personal information with the live, audio based interface. I hypothesized that the validity and emotional integrity of the memory-sharing experience relies on the strength of the personal database and accuracy of the cloned voice.

To fulfill these demands, the framework uses a systematic data intake pipeline, where all submissions (voice recordings, written text, or oral testimony) are converted into standardized forms. Voice recordings are chunked, denoised, and embedded into speaker embeddings. Written text is segmented into around 200 token-length chunks, embedded using a sentence transformer, and stored in a FAISS vector store to allow for semantic searching, while anchoring any responses.

The process of matching user queries to memory chunks is a semantic similarity search over the vector index. During inference time, the 5 semantically closest memory chunks are retrieved and used to construct a structured prompt which includes the persona of the deceased in the form of their name, personal history, personal speech and known values. This approach of retrieve-augmented generation largely mitigates the tendency to

hallucinate or provide a response that is off-character.

We propose a 7-layer pipeline architecture that starts with User Interface and ends with Audio output. This pipeline includes Automatic Speech Recognition, Large Language Model inference, Personal Knowledge base retrieval, Text-to-Speech generation, Voice cloning & speaker transfer and Audio output. Every layer is connected to the other adjacent layers through defined application programming interfaces. This provides modularity for upgrading any component without impacting the system.

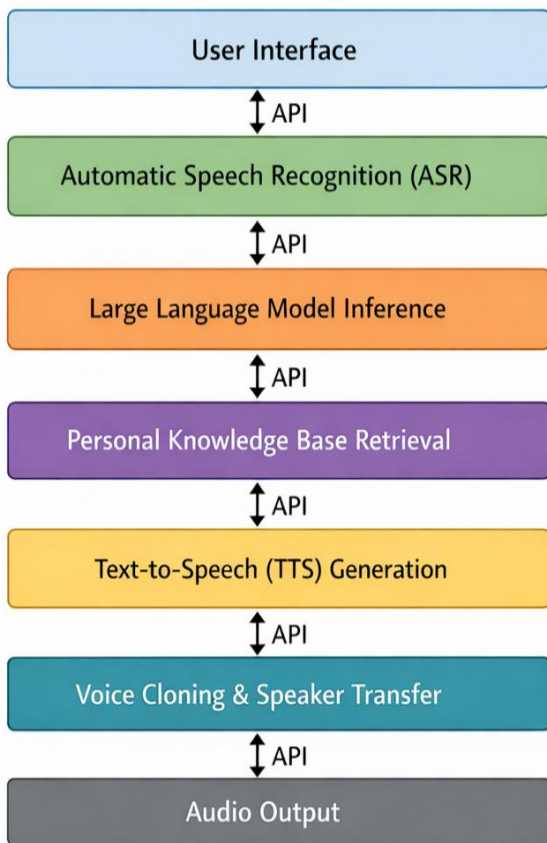


Fig.1:Architecture Diagram.

The interaction workflow can be seen as the following sequence of events, where a registered user starts a session, a spoken utterance of the user is processed by the ASR module, a suitable memory chunk is retrieved by a retrieval pipeline, a persona-conditioned response is generated by an LLM, and finally a speech synthesis process carried out by a TTS and a voice cloning module delivers the response. The feedback of the users is collected after each session and monitored system performance and the system's effect on user emotion.

Overall, this process illustrates how a pipeline composed of structured input, semantically aware retrieved-augmented generation, neural speech synthesis and emotionally tailored response conditioning can achieve an empathetic, natural and authentic posthumous digital experience.

V. FUTURE SCOPE

There's room for a richer and more moving human-machine experience by integrating leading computer vision technologies and multimodal AI research that may yield high fidelity, photorealistic video avatar synthesized, lip-syncing and facial animation features for MemorialBot. Long-term considerations for future enhancements will focus on adding support for Hindi, Marathi, and other major regional Indian languages to accommodate the diverse demographics across India and increase accessibility. A special grief therapy dialogue interaction feature which will work closely with certified professionals and directly integrate structured therapeutical dialogue principles into the conversational agent's framework will allow MemorialBot to serve as an accredited aid to counseling.

VI. CONCLUSIONS

In this paper we present a framework and design of MemorialBot-an AI-based system that aims to bridge the gap and lack of engaging, personalized digital memorialization by combining voice cloning, retrieval-augmented LLM inference, and neural

TTS within one coherent pipeline that is both usable and interactive. Sentiment-aware response conditioning and governance mechanisms, along with the development of a user-friendly voice-first interface ensures that the system caters to the needs of users while retaining the ethicality that this domain requires. In the larger scope, the framework signifies the capacity for AI technology to facilitate human grief and the progression of affect technology and the digital heritage field.

ACKNOWLEDGMENT

I would like to express my gratitude to our institution, professors and mentors for supporting us through this entire research. Special thanks to Prof. P.J. Patel and the Department of Artificial Intelligence and Data Science, Jawahar Education Society's Institute of Technology, Management and Research Centre, for the guidance, motivation and technical help provided throughout this work.

REFERENCES

1. Y. Mori and T. Nakamura, "Digital immortality and AI-based posthumous interaction: a systematic review," *Journal of Human-Computer Studies*, vol. 175, p. 103012, 2023.
2. T. Brown et al., "Language models are few-shot learners," *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020.
3. C.J. Sofka, I.N. Cupit, and K.R. Gilbert, Eds., *Dying, Death and Grief in an Online Universe*. New York: Springer, 2012.
4. G. Park, J. Chung, and S. Lee, "AI chatbot emotional disclosure for mental health counselling," *Current Psychology*, vol. 41, no. 8, pp. 5387–5396, 2022.
5. C.H. Reddy, "Real-time voice cloning using generative adversarial networks," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 2, pp. 245–253, 2024.
6. H. Papneja and N. Yadav, "Self-disclosure to conversational AI: a systematic review," *Personal and Ubiquitous Computing*, vol. 29, no. 1, pp. 88–104, 2025.
7. E. Ankomah and R.E. Turkson, "Emotion-aware AI chatbots for mental health support," *World Journal of Public Health*, vol. 3, no. 1, pp. 12–21, 2025.
8. S. Kim and H. Kim, "Grief chatbots and emotional dependency: risks and mitigations," *Computers in Human Behavior*, vol. 142, p. 107658, 2023.
9. E. Casanova et al., "YourTTS: towards zero-shot multi-speaker TTS and zero-shot voice conversion," in *Proc. ICML, PMLR 162*, pp. 2709–2720, 2022.
10. A. Jain and R. Gupta, "Voice-based digital heritage preservation using deep learning," *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 1124 – 1135, 2023.