Sound Strategies
SEARCH ENGINE OPTIMIZATION · WEB DESIGN · WEB HOSTING

Eternos.Life – AI Chatbot – Human Emulation with Voice Synthesis – Case Study

Prepared by Sound Strategies, Inc. – Petr Voinov, Sergey Gavrik, Alex Popovich, Tarin LoCascio
Last updated April 13, 2024

Abstract

Sound Strategies designed and built the Eternos AI application and web presence. This document provides the details of the first successful human emulation using voice synthesis and specifically targeting the lifetime memories of a single individual; Michael Bommer. The press release and testimonial can be found here: https://eternos.life/2024/04/michael-bommer-becomes-the-first-eternos-ai/

Index

Sound Strategies, Inc. Manhasset, NY USA
Andy LoCascio – Phone: (516) 978-1634 – AndyL@Soundst.com
Tarin LoCascio – Phone: (702) 499-6988 – TarinK@Soundst.com

1. Project Overview

The Eternos project aimed to develop an AI chatbot capable of emulating the voice and behavior of any one individual. The first case was German born Michael Bommer. The chatbot's purpose was to provide detailed information in first-person view sharing all the details of Michael Bommer's life as well as providing guidance and companionship for his spouse, family, and descendants. Access to the final AI chatbot is currently controlled by Michael Bommer and upon his death, by the appointed administrator for his AI via codicil. The lifetime of the AI is indefinite and the intention is that it will be available for many future generations.

2. Objectives
The following is a partial list of the project objectives:

- Develop a natural language processing (NLP) model capable of understanding and generating responses in an individual person's style using data specifically provided by that person.
- Develop a training approach and methodology to capture information primarily from the person, and additionally from his or her family members, and limited internet sources.
- Create and deeply a speech synthesis system that accurately replicates the individual's voice.
- Implement machine learning algorithms to continuously improve the chatbot's ability to emulate the individuals' behavior, accurately retrieve historical data and apply guidance using the samples provided by the individual.
- Ensure the chatbot's responses are contextually appropriate and aligned with individual's principles and demeanor.
- Test the chatbot's effectiveness in engaging users and providing valuable insights and guidance.

3. Implementation

- Data Collection – Created a carefully customized set of over 200 questions regarding all aspects of Michael Bommer's life and his opinions and advice on multiple topics.
- Audio Capture - Captured over 8 hours of audio responses to the questions that were provided.
- Audio Conversion - Converted speech to text, corrected grammar and spelling, and added paragraph breaks.
- LLM Creation – Built a fine-tuned Chat-GPT model using the relevant prompts and responses for the entire set of paragraphs based on an algorithm specifically for first-person view responses.
- Context Retrieval - Created a custom content retrieval engine specifically for life history recall and guidance.
- Voice Synthesis - Created a Microsoft Azure custom neural speech voice model utilizing over 400 carefully chosen utterances.

4. Integration

The Eternos team Integrated the LLM fine-tuned model, voice synthesis model, and the context retrieval engine into a unified chatbot framework accessible via web or mobile interfaces.

Sound Strategies, Inc. Manhasset, NY USA
Andy LoCascio – Phone: (516) 978-1634 – AndyL@Soundst.com
Tarin LoCascio – Phone: (702) 499-6988 – TarinK@Soundst.com

5. Fidelity Testing and Iterations

Fidelity Test 1

One hundred carefully chosen questions based on the data provided by Michael Bommer were used to test the fidelity and performance of the chatbot using an automated tool. The initial test score was 79/100. The 21 failed responses included 12 that were due to missing, inconsistent, and incomplete datasets. There were 5 failed responses that were due to missing context elements. The causes for the remaining 4 were not identified.

Iteration A

The context engine was altered to provide notification to the visitor when data was missing. Additional data provided by Michael was added to the model and several datasets were manually edited. The fine-tuned model was rebuilt using an updated algorithm for prompt and response creation. The context generation algorithm and system role were also altered.

Fidelity Test 2

The same set of questions was used to test the fidelity and performance of the chatbot. The test score improved to 97/100. The 3 failed responses were due to missing datasets.

Fidelity Test 3

Four 50 question real-world usage tests (200 total questions) were performed specifically with the intent to generate failed responses. The test score was 170/200. There were 13 failed responses due to missing, inconsistent, and incomplete datasets. There were 11 failed responses due to missing context elements. The causes of the remaining 4 were not identified.

Iteration B

The fine-tuned model was rebuilt using an updated algorithm for prompt and response creation. The context generation algorithm and system role were also altered. A critical feedback function was added to notify the administrator of null responses (previously identified as missing content).

Fidelity Test 4

The initial set of 100 questions were used to retest the fidelity of the model. The test score remained at 97/100. The 3 failed responses were due to missing datasets to be addressed via the feedback function.

Fidelity Test 5

Four 50 question real-world usage tests (200 total questions) were performed specifically with the intent to generate failed responses. The test score was 182/200. There were 12 failed responses due to missing, inconsistent, and incomplete datasets to be addressed via the feedback function. The causes of the remaining 6 were not identified.

Iteration C

An additional feature was enabled that allowed the real-world test group to request additional responses for the same prompt and then identify them as the preferred response. Data from those interactions was then used to add rules to the context generation algorithm and to enhance the fine-tuned model. The same four 50 question real-world usage tests (200 total questions) from fidelity test 5 were performed again with the testers providing addition tuning data.

Sound Strategies, Inc. Manhasset, NY USA
Andy LoCascio – Phone: (516) 978-1634 – AndyL@Soundst.com
Tarin LoCascio – Phone: (702) 499-6988 – TarinK@Soundst.com

Fidelity Test 6

The same four 50 question real-world usage tests (200 total questions) from fidelity test 5 were performed again. The test score was 197/200. The causes of the 3 failed responses were not identified. **Final fidelity success rate 98.5%.**

Final Fidelity Testing and Iteration Results

- Accurate Voice and Behavior - The AI chatbot successfully replicated Michael Bommer's voice and behavior, providing users with an authentic conversational experience.
- Engagement and Satisfaction - Users reported high levels of engagement and satisfaction with the chatbot, praising its ability to deliver history, relevant advice, and insights reminiscent as expected.
- Versatility - The chatbot demonstrated versatility in addressing a wide range of topics, including early life, family, career, education, and guidance.
- Continuous Improvement -Through the visitor feedback function, the chatbot continuously improved its emulation of behavior, ensuring relevance and accuracy over time.

---

6. Conclusion

The successful development and deployment of an AI chatbot capable of emulating Michael Bommer's voice and behavior represents a significant achievement in the field of conversational AI. By leveraging fine-tuned models, voice synthesis, content retrieval, and visitor feedback, the chatbot not only provided users with valuable insights and guidance but also preserves the essence of Michael Bommer's legacy for future generations. As AI technology continues to evolve, such projects serve as a testament to the potential for AI to enhance human communication and preserve the voices and wisdom of prior generations.