

ChatNekoHacker: Real-Time Fan Engagement with Conversational Agents

Anonymous Author(s)



Figure 1: ChatNekoHacker User Interface

Abstract

This paper introduces ChatNekoHacker, a real-time conversational agent system designed to enhance fan engagement within the music industry. By integrating Amazon Bedrock Agents for autonomous dialogue generation, Unity for immersive 3D broadcast environments, and VOICEVOX for high-quality Japanese text-to-speech synthesis, the system simulates interactive dialogues through character personas representing the music duo Neko Hacker. In a one-hour YouTube Live session with 30 participants, comprehensive experiments were conducted to evaluate the system's impact on user interest and engagement.

The experimental results indicate that interactions with the conversational agents significantly improve fan interest. Notably, regression analysis revealed that the primary contributor to this increased interest is the enjoyment—or fun—derived from the broadcast experience. Additional survey findings suggested that the engaging nature of the live interaction not only heightened curiosity

towards the artists but also positively influenced fans' willingness to listen to more music and participate in live events.

These findings underscore the importance of delivering an entertaining and interactive broadcast experience as a key factor in boosting fan engagement. The contributions of this work provide valuable insights into leveraging conversational agents in the entertainment sector, while also highlighting areas for further improvement, such as enhancing response diversity, reducing latency, and ensuring the accuracy of artist representation. On the other hand, concerns have been raised about the risk of the conversational agent providing misinformation to viewers. It is considered necessary to implement risk mitigation measures, such as improving accuracy and ensuring appropriate follow-up.

CCS Concepts

• Human-centered computing → Interactive systems and tools.

Keywords

Conversational Agents, Large Language Model, RAG

ACM Reference Format:

Anonymous Author(s). 2025. ChatNekoHacker: Real-Time Fan Engagement with Conversational Agents. In *Proceedings of Conference Title (Conference acronym 'XX')*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/XXXXXXX.XXXXXXX>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX', City, Country

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-XXXX-XXXX-X/20XX/XX

<https://doi.org/10.1145/XXXXXXX.XXXXXXX>

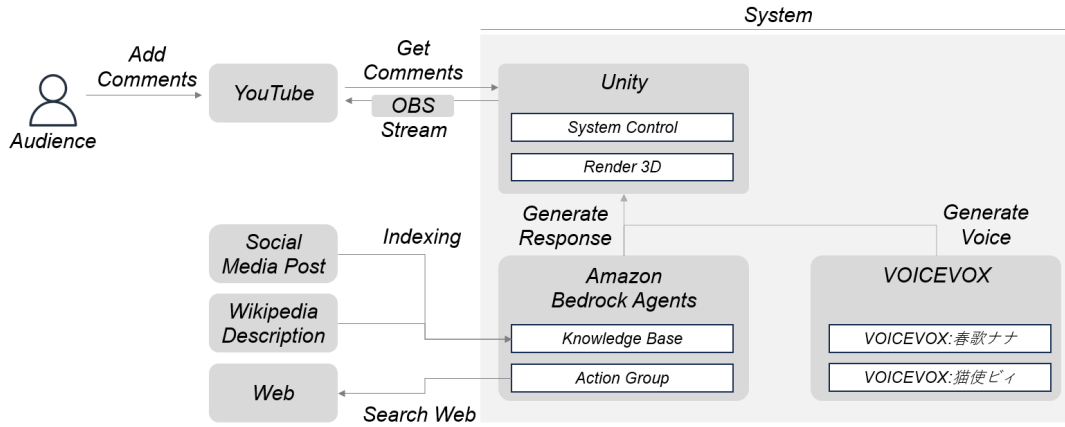


Figure 2: System Architecture

1 Introduction

In the entertainment domain—particularly within the music industry—the utilization of social media and dedicated applications for interactive engagement with fans has garnered significant attention as a key factor in enhancing brand value [2, 3]. Maintaining fan interest and motivation to act is therefore of paramount importance.

Simultaneously, the rapid evolution of Large Language Models (LLMs) has introduced new possibilities in the design of conversational interfaces. Approaches employing LLM-based conversational agents—such as SimsChat [5] and Character-LLM [4]—have demonstrated that advanced fine-tuning can improve character consistency and factual accuracy. However, these methods typically require substantial computational resources and specialized expertise, presenting a significant barrier for adoption by artists and record labels. In this paper, we present the development of a conversational agent based on a persona constructed from various information sources concerning the duo known as Neko Hacker, utilizing Amazon Bedrock Agents. Additionally, we constructed a system that automatically collects viewer comments from YouTube Live broadcasts and delivers responses via speech synthesis.

Figure 1 shows a snapshot of the ChatNekoHacker system interface during a live broadcast.

During a one-hour live broadcast, surveys were collected from 30 viewers. The results were analyzed from the following two perspectives:

- The contribution to enhancing fan interest and motivation to act.
- The advantages and challenges faced by artists when implementing fan engagement via a conversational agent.

2 System

Figure 2 illustrates the overall system architecture.

2.1 Unity

To recreate an engaging live broadcast experience akin to traditional human-hosted sessions, a 3D character model and streaming environment were reproduced within Unity. Through Unity, the

system facilitates the collection of comments from YouTube Live, response generation via Amazon Bedrock Agents, and voice synthesis using VOICEVOX.

2.2 Amazon Bedrock Agents

Amazon Bedrock Agents provide functionalities for constructing autonomous agents. By leveraging a knowledge base, these agents can retrieve information from unstructured data sources and generate responses to user inputs. Moreover, through integrated action capabilities, they are able to execute functions such as web searches to obtain the latest information. In this study, by indexing summaries of group members’ social media posts and activity overviews from Wikipedia into the knowledge base, agents can generate response based on past member statements. Figure 3 illustrates the processing method for social media posts. For each member, an LLM was used to classify the Social Media Posts into 15 different categories. Subsequently, a summary was generated for each category, and the summary results were indexed into the Knowledge Base.

Prompt engineering was used to control the system to speak in Kansai dialect, which is the Japanese dialect spoken by the members. Additionally, by implementing web search functionality within the action group, the system could integrate up-to-date information on live events and releases. Given that Neko Hacker is a duo, two distinct conversational agents—Neko-Chan and Hacker-Chan—were developed.

2.3 VOICEVOX

VOICEVOX [1] is a text-to-speech and singing voice synthesis software compatible with Japanese. In the present study, the voice “VOICEVOX: Haruka Nana” was used for Neko-Chan, and “VOICEVOX: Neko-tsuka Bii” for Hacker-Chan.

3 Experiments

A one-hour live broadcast was conducted during which surveys were administered to 30 viewers; the results were subsequently analyzed. The survey items included the frequency with which respondents listened to Neko Hacker’s music, the number of times

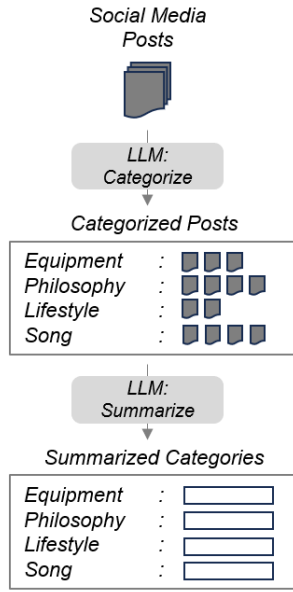


Figure 3: processing method for social media posts

they had attended Neko Hacker’s live events, and a series of questions concerning their viewing experience and changes in their motivation. Responses were recorded on a five-point Likert scale (Agree, Slightly Agree, Undecided, Slightly Disagree, Disagree). The questions are presented in Table 1

Table 1: Item and Description

Item	Description
Interest	Increased interest in the artist
Fun	Conversation was enjoyable
Useful	Gained meaningful information
Reality	Made statements resembling the artist’s
Unity	Felt a sense of unity during broadcast
Frequent	Frequency of listening to Neko Hacker’s music
Experience	Number of times attended Neko Hacker’s live events
ListenMore	Felt like listening to more Neko Hacker’s music
JoinMore	Felt like attending more Neko Hacker live events
Comments	Free-response comments

According to the survey results, 83% of respondents provided positive responses (Agree or Slightly Agree) to the item *Interest: Increased interest in the artist*. To assess how interactions with the conversational agent contributed to this enhanced interest, a regression analysis using the least squares method was performed based on four survey items. The results are presented in Table 2

With an adjusted R-squared of 0.56, an F-statistic of 10.27, and a p-value of 0.00005 for the F-statistic, the regression model is both valid and statistically significant. The analysis revealed that “Fun” was the only factor contributing significantly to “Interest,” while “Useful,” “Reality,” and “Unity” had minimal influence and were

Table 2: Statistical Analysis Results

Item	Positive Rate	Coefficient	p-value
Interest	83%	–	–
Fun	86%	0.59	0.01
Useful	67%	0.24	0.13
Reality	67%	0.00	1.0
Unity	67%	0.25	0.11

not statistically significant. In free-response comments, respondents noted positive aspects such as “The responses were natural” and “The conversation closely resembled everyday dialogue.” Conversely, comments highlighting areas for improvement—such as “The conversation lacks variety” and “I expected more interesting responses”—indicate that viewers primarily sought an enjoyable experience.

Next, the contribution to enhancing the willingness to act was evaluated. Respondents who indicated that they listened to the music “every day” were categorized as the *Frequent* group, while those who reported listening less than several times per week were categorized as the *Infrequent* group. The response tendencies for the item *ListenMore: Felt like listening to more songs* were compared between these groups. The results are presented in Figure 4. Furthermore, respondents were categorized into two groups based on their live event participation: those who had attended live events were labeled as *Experienced*, while those who had not were labeled as *Inexperienced*. The response tendencies for the item *JoinMore: Felt like participating in more live events* were then compared. A chi-square test conducted for the items *ListenMore* and *JoinMore* revealed no significant difference in the response distribution between the groups.

During the live broadcast, the following exchange took place. Although Japanese was used during the broadcast, an English translation is provided. “Spaceship” and “City Runner” are actual song titles, and the concept described by Hacker-Chan is based on the members’ past posts.

- Audience: Do you have a song or memory that has left the strongest impression on you during your production process?
- Hacker-Chan: That’s a tough question to answer. But the song ‘Spaceship’ is particularly memorable. I created it with the image of a spaceship in mind, and it turned into a sparkling Vocaloid track. I also put a lot of care into the guitar phrasing. Additionally, ‘City Runner’ is another memorable piece. I made it with the idea of running through a futuristic city, and it gave off a feeling of being chased. I really enjoyed the production process.

4 Discussion and Conclusion

This study demonstrates that conversational agents have the potential to enhance fan engagement for music artists. In particular, the findings indicate that the enjoyment derived from the conversation is a critical factor in increasing interest in the artist, suggesting that highly entertaining interactions provided by conversational agents are key to boosting fan engagement. Moreover, the agents

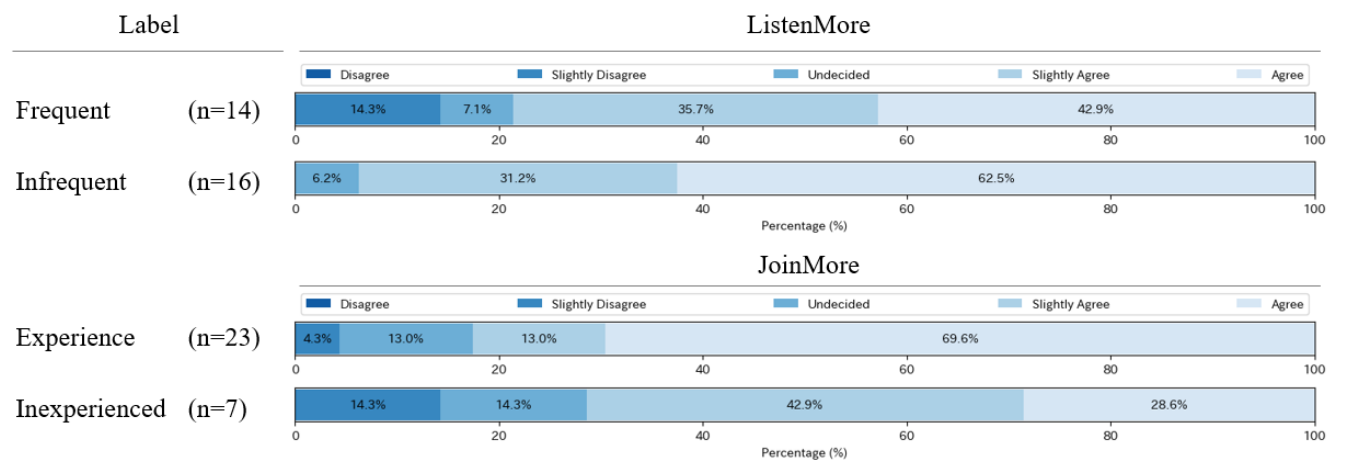


Figure 4: Results for “ListenMore: Felt like Listening to More Songs”, “JoinMore: Felt like Participating in More Live Events”

were found to also increase interest in both songs and live events, thereby potentially enhancing the willingness to act. An interesting observation was made during the live broadcast: a viewer commented regarding an out-of-stock item from an online store, and following the subsequent restocking, the item was purchased. This example indicates that real-time conversational agents may foster purchasing intent and support consumer behavior.

However, this study is subject to several limitations. Due to the small sample size of 30 respondents, caution must be exercised when generalizing the results. Future work should address issues highlighted in the free-response comments, such as enhancing the variety and interest of the responses, better incorporating conversational context during broadcasts, improving the interplay between the two conversational agents, and reducing the latency from comment collection to vocalization. Furthermore, some responses generated by the conversational agents included inaccuracies or chronologically inconsistent information, posing a risk of disseminating misinformation to viewers.

Looking ahead, there is potential for conversational agents to evolve into tools that more accurately reflect an artist’s personality and support deeper fan relationships. By more precisely incorporating an artist’s past statements, works, and musical style, it may be possible to facilitate conversations that are more authentically representative of the artist. Additionally, such agents could learn individual fan preferences and interests to provide tailored conversational experiences. With continued technological advancements, conversational agents hold the promise to transform fan engagement within the entertainment industry by establishing a novel communication channel between artists and their fans.

References

[1] [n.d.]. VOICEVOX. <https://voicevox.hiroshiba.jp/> Accessed: 2025/2/3.

[2] Jessica Edlom and Jenni Karlsson. 2021. Hang with Me—Exploring Fandom, Brandom, and the Experiences and Motivations for Value Co-Creation in a Music Fan Community. *International Journal of Music Business Research* 10 (2021), 17–31.

[3] Sihan Li. 2024. Behavior and Identity Construction of Online Fan Community—Taking Pocket 48 as an Example. *Interdisciplinary Humanities and Communication Studies* (2024).

[4] Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. 2023. Character-LLM: A Trainable Agent for Role-Playing. *ArXiv abs/2310.10158* (2023).

[5] Bohao Yang, Dong Liu, Chen Tang, Chenghao Xiao, Kun Zhao, Chao Li, Lin Yuan, Guang Yang, Lanxiao Huang, and Chenghua Lin. 2024. SimsChat: A Customisable Persona-Driven Role-Playing Agent. *ArXiv abs/2406.17962* (2024).