TourAi an NPC tour guide

Ferdinando Boccia*
Domenico D'Antuono
f.boccia28@studenti.unisa.it
d.dantuono7@studenti.unisa.it
University of Salerno (UNISA)
Fisciano, Salenro, Italy

ABSTRACT

Tourism plays a pivotal role in global economies, with Italy being a prime destination due to its rich history and culture. However, challenges persist in providing personalized and language-accessible experiences for diverse visitors. In response, TourAi introduces an innovative solution leveraging Natural Language Processing (NLP) and 3D graphics technology to create an interactive virtual assistant. Through a structured methodology encompassing avatar generation, knowledge base construction, and application development, this project aims to enhance the tourist experience by offering real-time localized information, recommendations, and navigation assistance.

Contents

*Both authors contributed equally to this research.

1 CONTEXT AND MOTIVATION

Tourism represents a crucial sector for many global economies, significantly contributing to economic growth and the well-being of local communities. In particular, Italy, with its rich history, culture, and scenic beauty, attracts millions of tourists from around the world every year.

However, despite its enormous tourism potential, Italy faces challenges in offering high-quality tourism experiences that meet the increasingly high expectations of international visitors. Among these challenges is the need to provide accurate and personalized local information, as well as the ability to overcome language barriers to ensure a seamless experience for tourists from diverse backgrounds.

The "TourAi: an NPC Tour Guide" project emerges precisely in this context, with the aim of addressing these emerging needs. Leveraging advancements in Natural Language Processing (NLP) and 3D graphics technology, the project aims to create a virtual assistant capable of acting as an interactive tour guide.

The significance of this project lies in its ability to enhance the overall tourist experience by providing detailed and localized information in an intuitive and accessible manner. The NPC tour guide will be able to provide recommendations on restaurants, attractions, and activities, as well as respond to tourists' questions in real-time, thereby making the journey smoother and more memorable.

Furthermore, the project fits into a broader context of technological innovation in the tourism sector, contributing to positioning Italy as a leading tourist destination and attracting a new generation of digital travelers.

2 CONVERSATIONAL AGENTS

Conversational agents, also known as chatbots or virtual assistants, are software programs powered by artificial intelligence (AI) that interact with users through natural language conversations. These agents are designed to simulate human-like conversations and provide assistance or information to users in a conversational manner. Chatbots are rapidly used in different sectors to enhance better customer experience and increase the financial performance of the industries[1].

The reasons why these systems are becoming increasingly popular and used are related to their features and the benefits they offer.

- **24/7 Availability**: Chatbots provide users with instant responses and can be accessed 24/7, offering a significant advantage over traditional services. Users can have an instant conversation in real-time instead of waiting in long queues for solving the query[3].
- **Reduction in Workforce**: By automating responses and services, chatbots can reduce the need for human intervention in certain tasks, leading to cost savings for businesses.
- **Integration with Emerging Technologies**: Chatbots are often integrated with technologies like artificial intelligence (AI), machine learning, big data, and the Internet of Things (IoT) to enhance their capabilities and improve user interactions[2].
- **Enhanced User Experience**: Chatbots aim to provide a seamless and convenient user experience by offering real-time responses and personalized interactions .
- **Applications Across Various Sectors**: Chatbots find applications in diverse sectors such as education, tourism, banking, and entertainment, showcasing their versatility and utility[1].
- **Advancements in Natural Language Processing**: Chatbots leverage advancements in natural language processing to understand and respond to user queries effectively, contributing to their efficiency .

Furthermore, findings from the study highlight The Open University, Pecking University, Politecnico di Milano, and Microsoft Corporation as leading contributors in annual document output within the chatbots and virtual assistants domain. Researchers interested in delving deeper into this area are advised to reference these institutions for comprehensive insights.

3 PURPOSE AND OBJECTIVES

The purpose of the "TourAi: an NPC Tour Guide" project is to develop a virtual assistant integrated with an NLP (Natural Language Processing) model tailored to the tourism domain. The primary objectives of the project include:

- **Creation of an NPC connected to an NLP model**: Develop a virtual NPC (Non-Player Character) capable of interacting with users based on an NLP model trained on tourist knowledge of select cities. This will enable the NPC to understand and respond to user queries in a conversational manner.
- **NPC speaking capability**: Implement speech synthesis functionality to enable the NPC to communicate verbally with users, providing a more immersive and natural interaction experience.
- **Tour guide functionality**: Equip the NPC with the ability to act as a tour guide, offering information about the history of the city, recommending restaurants, and suggesting places to visit based on user preferences and contextual cues.
- **Integration into a working application**: Develop a functional application that integrates the NPC tour guide seamlessly, providing users with an intuitive and user-friendly interface for interacting with the virtual assistant.
- **Voice and text communication**: Design the application to support both voice and text-based communication with the NPC, allowing users to interact with the assistant using their preferred mode of communication.
- **Multilingual support**: Ensure that the application supports input in both Italian and English languages, catering to the diverse linguistic preferences of users.
- **Navigation integration**: Implement a feature whereby the NPC can recommend places of interest, and the application provides users with a link for navigation to the suggested location, enhancing the practical utility of the virtual tour guide.

By achieving these objectives, the project aims to offer a comprehensive and user-centric solution for enhancing the tourist experience, leveraging advanced technologies in NLP and virtual assistant development.

4 METHODOLOGY

The process of developing the avatar and integrating it into the application follows a series of structured and methodical phases to ensure an optimal and functional outcome. Below are the main phases of the process:

(1) **Avatar Generation and Training with Convai**: To generate and train the avatar, we use the Convai web application [4]. Convai provides an intuitive interface for designing characters with multimodal perception abilities in both virtual and real-world environments. Initially, the avatar is trained with information related to one or more tourist cities to acquire a knowledge base. Subsequently, training may be expanded with additional data to enhance its responsiveness and provide increasingly detailed information.
(2) **Appearance Modification and Basic Information Setting**: The avatar's appearance is modified to make it visually appealing [fig: 1]. Subsequently,

basic information such as name, voice type, and background is set through Convai's Character Description section. Avatar appearance creation is managed through Ready Player Me: Ready Player Me [7] is a platform that allows users to create personalized 3D avatars that can be used in various virtual reality (VR), augmented reality (AR) applications, and in the metaverse. The platform is designed to be interoperable, meaning that avatars created with Ready Player Me can be used across a wide range of applications and virtual environments without the need to recreate them for each one.

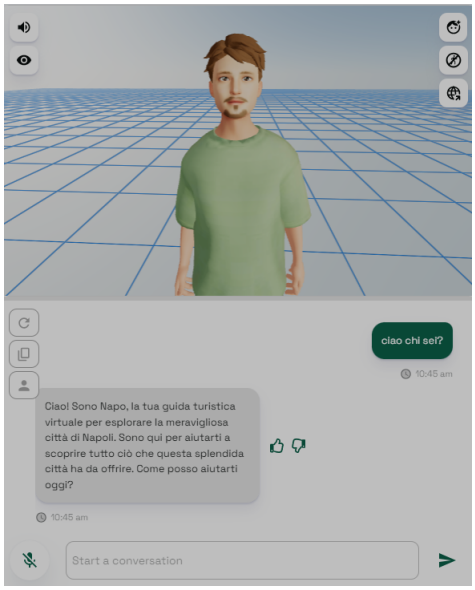

Figure 1: Avatar panel in Convai.

(3) **Construction of the Knowledge Bank**: A knowledge bank containing varied tourist information is created, divided into separate files for categories such as historical sites, tourist attractions, entertainment venues, restaurants, etc. This database will serve as the repository of knowledge from which the avatar can draw during interactions with users.
(4) **Avatar Testing and Evaluation**: The avatar undergoes testing to identify any issues and measure the accuracy of the responses provided. This validation process is crucial to ensure that the avatar can deliver correct and useful information to tourists.
(5) **Integration of the Avatar into Unity and Mobile Application Development**: Once satisfied with the results, the avatar model is imported into Unity[8] via Convai's APIs. From there, mobile application development begins, with particular attention to designing the user interface to ensure an intuitive and satisfying user experience.

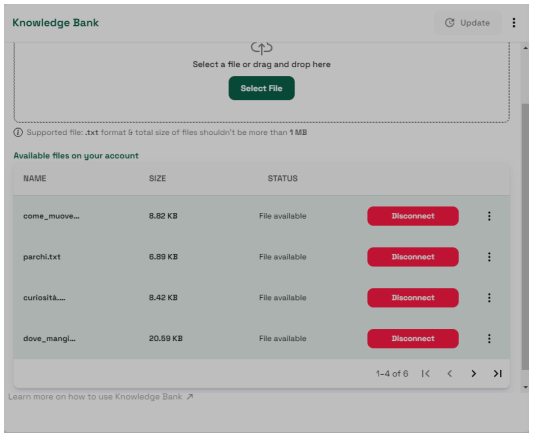

Figure 2: Knowledge Bank in Convai.

(6) **Implementation of Additional Features**: If necessary, additional scripts are developed to implement extra functionalities in the application, enriching the user experience and improving the avatar's capabilities as a tour guide.
A new feature has been developed: **augmented reality**. By launching the application and pressing the "Switch to AR" button, users can view the avatar in augmented reality. This action will change the scene, activate the camera (after granting necessary permissions), and load the avatar's 3D model. All standard features such as chat, animations, and lipsync are retained. The main difference is that users are no longer in a static scene but a dynamic one; to interact with the bot, they need to approach at an appropriate distance, at which point the chat panel and settings will appear.
(7) **Application Build, Testing, and Release**: Upon completion of development, the application is built, likely for the Android operating system, and tested on external devices or emulators. Any issues encountered are resolved before the final release version.

5 NON-FUNCTIONAL REQUIREMENTS

For the development of the tool, it is essential to consider certain requirements that must be/are guaranteed by the tool [5].

(1) **Accuracy**: The system must provide a high number of correct responses; otherwise, it would be of little use or even harmful to users.

(2) **Reusability**: The bot should be usable in future projects and different contexts, such as those related to the metaverse. This requirement is ensured by Convai, which allows for easy and quick implementation of avatars with various game engines or frameworks. Thus, the same avatar can be used in different ways.
(3) **Usability**: The ease of interaction with the assistant is key to making the system engaging for users. Ensuring users can ask questions easily and receive appropriate responses is paramount.
(4) **Consistency**: Requests from different users should yield similar results; otherwise, the system may be perceived as unreliable.
(5) **Retrainability**: Since the system will interact with a real-world context that is constantly changing, it is crucial to ensure it can maintain performance levels with context updates (e.g., new monuments, street name changes, restaurant closures).
This requirement is guaranteed by the system through the use of the knowledge bank.
(6) **Moderation**: The system must moderate and filter the responses provided to the user to protect them from harmful, offensive, or inappropriate content.
This requirement is ensured by Convai through the use of moderation classifiers [5].

5.1 Optimization

Since Convai does not allow direct interaction with the model but only with the data, many optimization operations are not possible. Considering the available tools, we have decided to improve only a subset of the requirements. These include:

- **Accuracy**: The Convai system allows for a specific feature to improve accuracy: users can rate the responses provided by the avatar with a "thumbs up" or "thumbs down." This feedback helps the model improve future responses [6].
- **Usability**: This requirement can be enhanced by modifying the user interface to make interaction with the avatar more pleasant and user-friendly.
- **Consistency**: This requirement can be optimized by using the knowledge bank. This tool not only helps to better characterize the avatar or add knowledge and information to the base model, but it also helps mitigate the 'hallucination of information' issue that plagues even the best large language models (LLMs).

6 DATA CLEANING AND PRE-PROCESSING

When training language models, the initial step requires pre-processing the data to ensure consistency, followed by implementing feature engineering techniques to optimize model outcomes.

Data preprocessing entails cleaning and organizing raw text data for machine learning. Its aim is to standardize input data, simplifying the complexity for the model. Several techniques used in preprocessing data for training LLMs are:

Data cleaning, fundamental for LLM training, involves identifying and rectifying inaccuracies, inconsistencies, and irrelevant elements within the raw text data. Procedures include removing duplicates, handling missing values, and addressing formatting irregularities. Additionally, text-specific tasks like removing special characters, punctuation, and stop words streamline the input, significantly enhancing the quality and integrity of training data.

Normalization, crucial for standardizing textual data, ensures uniformity and consistency in language usage, minimizing complexity for NLP models. This involves converting text to a common case, usually lowercase, and standardizing numerical data and dates, simplifying the dataset for more effective training and improved generalization capabilities.

Tokenization, breaks down text into smaller units or tokens, like words, subwords, or characters, providing a structured input for the model. By analyzing and generating coherent sequences of words, tokenization helps the model gain a granular understanding of language usage and syntax.

Stemming involves deriving the root form of a word by removing prefixes or suffixes, though it may not always result in actual words.

Lemmatization, similar to stemming, also converts words to their root forms but ensures the root word belongs to the language, resulting in valid words.

To carry out these processes, we will use a data pipeline in Python leveraging certain libraries for natural language analysis such as **NLTK (Natural Language Toolkit)** or **Simplemma**.

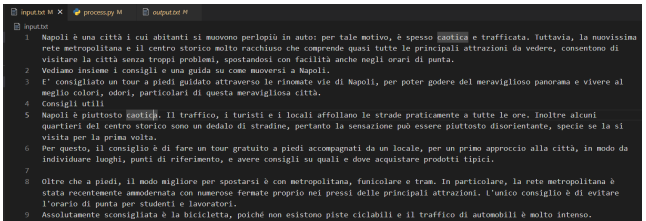

Figure 3: Input text before processing.

The data pre-processing and cleaning tool was developed using Python and has been uploaded to a GitHub repository.

Figure 4: Text after processing

For more information and to view the code, visit: https://github.com/Ferdi00/SE4AI_TourAI.

The tool is quite simple to use. To start it, just input the text you want to process into the input.txt file [fig: 3] and run the process.py Python script. Once the process is complete, the converted text will be available in the output.txt file [fig: 4].

7 EVALUATION AND VERIFICATION

To ensure the completeness of the project, it has been decided to verify the previously described non-functional requirements. Given that not all requirements can be directly verified, we will concentrate on a narrower **subset**. Consequently, tests will be conducted on the following requirements:

- **Accuracy**
- **Consistency**
- **Usability**
- **Moderation**

7.1 Note

Currently, Convai is working on a testing framework to support users in generating and evaluating test cases. The promised features include:

Dataset Creation and Management:

- **Seamless Integration**: Users can create test datasets directly through Convai's feedback system, using indicators such as thumbs up or down during interactions with their Characters.
- **Detailed Feedback**: Users can provide additional details on their evaluations to enhance the understanding of expectations and the context of the feedback.

Test Automation:

- **Dedicated "Testing" Tab**: A unified table that displays all interactions with detailed feedback, providing an intuitive and efficient way to manage test cases.
- **Test Filtering and Execution**: Users can filter test cases by tags and execute selected tests with a single click.
- **Instant Result Generation**: After executing the tests, the system immediately generates and displays new outputs, providing immediate insights into test outcomes.

Test Case Tagging:

- **Advanced Test Management**: Users can tag responses to better group and manage their test cases, facilitating filtering and organization.

Currently, these features are not yet available but will be accessible in the future.

8 METHODOLOGY FOR VERIFICATION

Due to the non-deterministic nature of Generative AI, automating testing is found to be not very advantageous: the bot could be responded to the same question in different ways. Therefore, a decision has been made to conduct manual testing. A series of questions will be formulated to be given as input to the bot, and the responses will be manually evaluated, with scores being assigned.

To assess the previously described requirements, specific criteria must be specified:

8.1 Accuracy

- **Definition**: Measuring the model's ability to produce correct responses to prompts
- **Testing Methodology**:
  (1) Formulate a set of benchmark questions covering various aspects and complexities of the expected input.
  (2) Compare the bot's responses to a predefined set of correct or expected answers.
- **Scale**: 1-5 scale for each response, where:
  (1) Completely incorrect
  (2) Mostly incorrect with some correct elements
  (3) Partially correct but incomplete or partially incorrect
  (4) Mostly correct with minor inaccuracies
  (5) Completely correct
- **Metrics**: The accuracy percentage will be evaluated based on all the questions provided. It will be calculated using the following formula:

$$\text{Accuracy Percentage} = \left(\frac{\sum \text{score per response}}{\text{total number of responses} \times 5} \right)$$

8.2 Consistency

- **Definition**: Consistency refers to the AI's ability to provide stable and coherent responses across similar queries over time.
- **Testing Methodology**:
  (1) Input similar questions or paraphrased versions of the same questions to the bot from different users
  (2) Evaluate the similarity and coherence of the responses.
- **Scale**: 1-5 scale for each response, where:
  (1) Completely inconsistent and illogical
  (2) Mostly inconsistent with minor similarities
  (3) Partially consistent with some coherence
  (4) Mostly consistent with minor variations
  (5) Completely consistent and coherent
- **Metrics**: The Consistency percentage will be evaluated based on all the questions provided. It will be calculated using the following formula:

$$\text{Consistency Percentage} = \left(\frac{\sum \text{score per response}}{\text{total number of responses} \times 5} \right)$$

8.3 Moderation

- **Definition**: AI's ability to provide appropriate responses without generating harmful, offensive, or inappropriate content.
- **Testing Methodology**:
  (1) Input a set of sensitive or potentially provocative questions.
  (2) Manually review the responses to assess their appropriateness.
- **Scale**: 1-5 scale for each response, where:
  (1) Highly inappropriate and offensive
  (2) Somewhat inappropriate with offensive elements
  (3) Neutral but could be improved for sensitivity
  (4) Appropriate with minor issues
  (5) Completely appropriate and sensitive
- **Metrics**: The Consistency percentage will be evaluated based on all the questions provided. It will be calculated using the following formula:

$$\text{Moderation Percentage} = \left(\frac{\sum \text{score per response}}{\text{total number of responses} \times 5} \right)$$

8.4 Usability

To evaluate this requirement, a **user-based testing** of the application will be carried out. Users will be involved to try out the application and will be asked questions related to the user experience, simplicity, satisfaction level, etc. For each criterion, users will be asked to assign a score. A good average of scores across all criteria will determine the completeness of the project.

8.4.1 Usability Evaluation Questionnaire.

GENERAL INFORMATION

(1) Name (optional):
(2) Age:
(3) Level of experience with technology (1 - Beginner, 5 - Expert):

INTERACTION EXPERIENCE

(1) **How easy was it to start using the application?**
  - 1: Very difficult
  - 2: Difficult
  - 3: Neutral
  - 4: Easy
  - 5: Very easy

(2) **Were the system's responses clear and understandable?**
  - 1: Not at all clear
  - 2: Slightly clear
  - 3: Neutral
  - 4: Clear
  - 5: Very clear

(3) **Were the responses provided by the system useful for completing your tasks?**
  - 1: Not at all useful
  - 2: Slightly useful
  - 3: Neutral
  - 4: Useful
  - 5: Very useful

(4) **How intuitive was the use of the application?**
  - 1: Not at all intuitive
  - 2: Slightly intuitive
  - 3: Neutral
  - 4: Intuitive
  - 5: Very intuitive

(5) **Did you encounter any technical difficulties while using the system?**
  - 1: Many difficulties
  - 2: Some difficulties
  - 3: Neutral
  - 4: Few difficulties
  - 5: No difficulties

USER SATISFACTION

(1) **Are you satisfied with the system's response speed?**
  - 1: Very dissatisfied
  - 2: Dissatisfied
  - 3: Neutral
  - 4: Satisfied
  - 5: Very satisfied

(2) **To what extent did the system meet your expectations?**
- 1: Not at all
- 2: Slightly
- 3: Neutral
- 4: Very much
- 5: Completely

(3) **Would you recommend this App to others?**
- 1: Absolutely not
- 2: No
- 3: Maybe
- 4: Yes
- 5: Absolutely yes

ADDITIONAL FEEDBACK

(1) **What were the best aspects of the application?**
(2) **What improvements would you suggest to make the system more usable?**

USAGE TIME

Did you find the time spent adequate for completing your tasks?

- 1: Not at all adequate
- 2: Slightly adequate
- 3: Neutral
- 4: Adequate
- 5: Very adequate

9 RESULTS

After defining the criteria and methodologies for the testing phase, it is important to analyze the collected data. To this end, a Google Sheets file was created where all questions of interest were gathered and categorized. A total of 90 questions were collected, divided as follows:

- 70 questions to evaluate accuracy
- 10 questions for consistency
- 10 questions for moderation

Following the assignment of scores, the following results were achieved:

- **Accuracy Percentage** = $\mathbf{281/350 \approx 80\%}$
- **Consistency Percentage** = $\mathbf{41/50 \approx 82\%}$
- **Moderation Percentage** = $\mathbf{50/50 \approx 100\%}$

Finally, a subset of questions in the "accuracy" file, containing both questions with inaccurate answers and questions with accurate answers, was evaluated. The accuracy post feedback was assessed. An increase in the accuracy percentage was observed from:

- **PRE**: $34/70 \approx 49\%$
- **POST**: $56/70 \approx 80\%$

To view all the collected questions and their respective evaluations, please refer to the following url:

https://docs.google.com/spreadsheets/d/19LLyov0YqCWSFVbPGVtNXtZWlpeKYPs9WUoUPCaBiVA/edit?usp=sharing

User-based testing was implemented by transposing the test questions into a questionnaire using Google Forms. Subsequently, all responses were entered into a Google Sheet for analysis. It was found that most responses were positive regarding the system's usability and user satisfaction level. The main strengths of the system, as reported by users, were the speed and efficiency of responses, as well as the system's quality and performance. Regarding suggestions, it was recommended to revisit the graphic design and enhance the user interface.

To view all the collected questions and their respective evaluations, please refer to the following url:

https://docs.google.com/spreadsheets/d/1NjvqWt_TnO25m7cck0Ckdc9Lgo1ISXbUEhCumW67gSs/edit?resourcekey=&gid=725793396#gid=725793396

REFERENCES

[1] Shubhangi Agarwal, Bhawna Agarwal, and Ruchika Gupta. 2022. Chatbots and virtual assistants: a bibliometric analysis. *Library Hi Tech* 40, 4 (2022), 1013–1030.

[2] Dr. Victor Chang, Dr. Dickson K.W. Chiu, Dr. Muthu Ramachandran, and Dr. Chung-Sheng Li. 2018. Internet of Things, Big Data and Complex Information Systems: Challenges, solutions and outputs from IoTBD 2016, COMPLEXIS 2016 and CLOSER 2016 selected papers and CLOSER 2015 keynote. *Future Generation Computer Systems* 79 (2018), 973–974. https://doi.org/10.1016/j.future.2017.09.013

[3] Dana Doherty and Kevin Curran. 2019. Chatbots for Online Banking Services. (2019).

[4] Convai Technologies Inc. 2024. Convai. https://www.convai.com Available on line.

[5] Convai Technologies Inc. 2024. Convai. https://convai.com/blog/introducing-convai Available on line.

[6] Convai Technologies Inc. 2024. Convai/testing. https://convai.com/blog/testing-framework Available on line.

[7] Ready Player Me. [n. d.]. Ready Player Me. https://readyplayer.me/it

[8] Unity Technologies. 2024. Unity. https://unity.com Available on line.