

Technology Reports

AI

Multilingual

Inbound

Interactive and Multilingual AI-based “Oshaberi Annaiban”

Solution Service Department Shin Oguri

Communication Device Development Department Naofumi Hara Anna Shima

Japan's labor force is decreasing year by year and recruiting qualified personnel is becoming difficult in a variety of industries. As a result, foreign workers are being increasingly accepted to fill this void, but providing support for them as well as for foreign visitors to Japan whose numbers are expected to grow is now a major issue facing Japan.

To help mitigate this labor shortage and provide smooth support for foreign visitors, “Oshaberi Annaiban” (Multilingual Information Board with Touch and Voice) has been developed using NTT DOCOMO AI technology. This service consists of an AI information board supporting touch operations and voice functions. It features an interactive multilingual format and customization capabilities and the ability to be used as a source of marketing data. Given that the number of foreigners coming to Japan—including workers and immigrants in addition to inbound travelers—is expected to increase from here on, there are great expectations that this service will be a tool not only for providing sightseeing information but also for other purposes such as displaying disaster information and providing evacuation instructions.

1. Introduction

The number of foreign visitors to Japan is increasing steadily with expectations that it will reach

60 million people by 2030 [1]. Furthermore, due to a revised Immigration Law that took effect in April 2019, the acceptance of foreign workers will be expanded over the next five years [2] pointing

©2020 NTT DOCOMO, INC.

Copies of articles may be reproduced only for personal, noncommercial use, provided that the name NTT DOCOMO Technical Journal, the name(s) of the author(s), the title and date of the article appear in the copies.

to an increase in the number of foreigners living in Japan [2] (Figures 1 and 2). However, while environments that will help foreigners adapt to life in Japan are now being set up in large metropolitan areas, there are still many local governments and companies in regional areas that are not sufficiently prepared [3]. This is turning out to be a major

problem especially at times of an emergency such as a natural disaster since measures should be established for providing information and evacuation instructions to people living in an evacuation zone [4] [5]. On the other hand, Japan’s working population is on a downward trend, which means that there is a limit as to how much support can be provided by

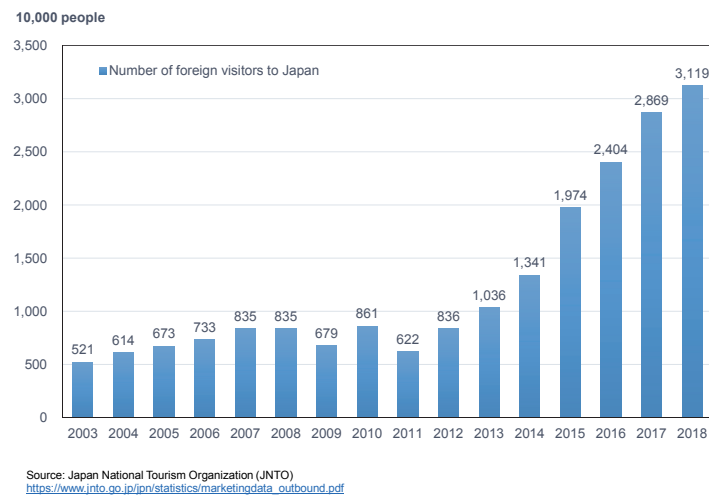


Figure 1 Number of foreign visitors to Japan

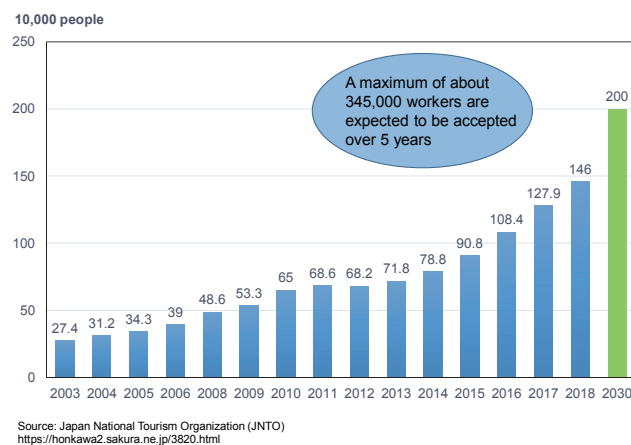


Figure 2 Change in number of foreign workers

people to foreigners who are increasing in number.

Against the above background, NTT DOCOMO has developed “Oshaberi Annaiban” that is cheaper to operate and easier to customize than communication robots using NTT DOCOMO’s own AI technology. This is a multilingual information board that can be tailored to the location targeted for installation. As a service, it can provide disaster-related information as described above as well as helpful sightseeing information, event updates, etc. particular to that region. In short, it can provide real-time information that is needed “here and now.”

This article describes NTT DOCOMO’s “Oshaberi Annaiban” developed with the aim of achieving multilingual guidance using AI.

2. Overview of Oshaberi Annaiban

In contrast to conventional information boards or electronic message boards, Oshaberi Annaiban is an interactive, multilingual AI-based information board that supports touch operations and voice by

adopting a touch-panel display and voice-input function. It can customize characters and backgrounds to match the purpose of use and installation location and therefore provide a spatial presentation at that site. It can also perform intention interpretation^{*1} against utterances made by a user during conversation with a character appearing on the screen so as to provide the information needed in response to each utterance. It also provides an interface that induces the user to think “I would like to try talking to this board.” or “I would like to ask something else.”

2.1 Interactive Guidance Supporting Voice and Touch Operations

Oshaberi Annaiban features voice input and touch operations and provides multilingual support including Japanese, English, Chinese, and Korean. It can provide guidance to foreign visitors in place of staff at facilities having no interpreters or bilingual staff (**Figure 3**).

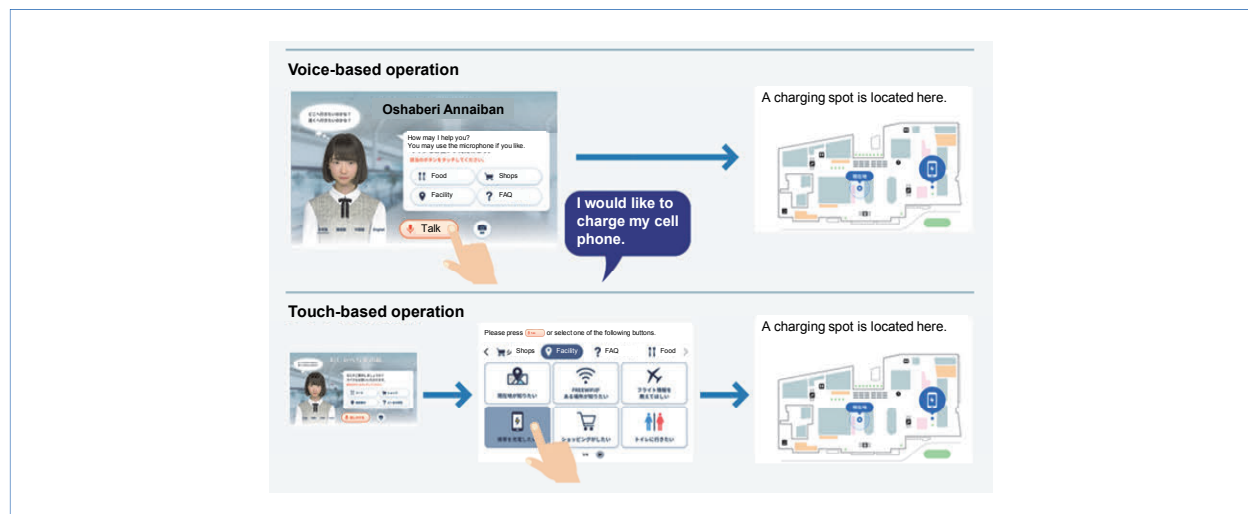


Figure 3 Communication by voice and touch operations

^{*1} Intention interpretation: Technology that uses machine learning and so forth to determine the user’s intention from the user’s utterances (natural language). User intentions are called “tasks.” For example, all the utterances “What’s tomorrow’s weather?,” “I wonder if tomorrow will be fine?,” and “Is it going to rain tomorrow?” are judged as weather tasks.

2.2 Friendly Communication

In conjunction with providing voice-based communication, Oshaberi Annaiban was designed to have an interface that would make it easy for the user to converse. This was accomplished through the following three features.

- To achieve a friendly tone, the content of the conversation is linked with the facial expressions of a three-Dimensional Computer-Generated (3DCG) character through intention interpretation using AI technology (Figure 4).
- To provide the user with optimal information, the conversation and guidance is made to match the attributes of the user during the conversation between the user and character.
- “Trigger” for expanding the conversation is created by displaying the character’s “inner thoughts” in a speech balloon (Figure 5).

Even with a design geared to responding efficiently and correctly to a user’s inquiry, the interface also incorporates measures for alleviating hesitancy in a user who might be thinking “I really don’t

know what to say ...” by becoming familiar with that person’s current situation and providing a warm communication experience.

Oshaberi Annaiban also features architecture that enables the character, tone of voice, etc. to be changed as desired, which means a format that can be customized according to the purpose of use and installation location. For example, collaboration between Oshaberi Annaiban and 3DCG character Saya^{*2}, who was born with the role of connecting people to people or people to place, enables the character’s facial expressions to be matched to the user’s utterances to create a friendly atmosphere. This function enables the provision of added value by combining Oshaberi Annaiban with any existing assets possessed by the board implementer such as a local mascot or corporate spokesmodel and enabling a person visiting the board to meet that character and receive local information.

2.3 Survey Function during Conversation

Up to now, most digital signage^{*3} or information boards have simply provided the information that the implementer wants to convey in a one-way



Figure 4 Communication linked with facial expressions of 3DCG character Saya [6]

^{*2} **Saya:** An original virtual human created by a husband and wife duo known professionally as TELYUKA. Using advanced computer graphics as an expressive medium, Saya was announced in 2015 becoming a worldwide sensation practically overnight. Born as a handmaid, Saya is an ongoing project showing evolution and growth as she searches out a new role (guide) with

a unique organic existence different from humans.
^{*3} **Digital signage:** Advertising media using digital technology. Using displays or projectors to change advertising content in response to time or location, this technology is gaining attention as an alternative to conventional advertising media such as posters etc.

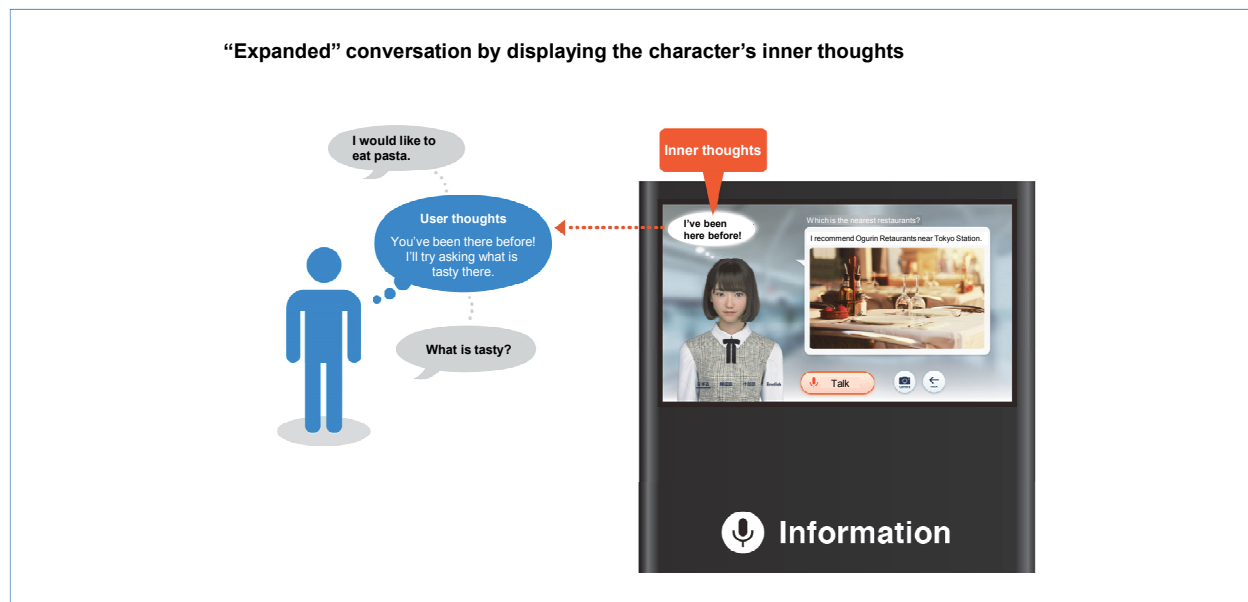


Figure 5 “Trigger” for expanding the conversation by presenting the character’s inner thoughts

manner to the user. For Oshaberi Annaiban, we have developed a survey function for asking the user questions from the character side during the course of a conversation. With this function, responses from users can be fed back to the installation company and used as marketing data for a variety of purposes such as measuring the impact of events, refining facility planning, and improving user satisfaction. By having the character ask questions in this way during a voice-based conversation, the user is able to complete the survey in a comfortable manner while enjoying the conversation (Figure 6).

In certain previous trials, the collection rate of paper-based surveys was a problem, and it was said that even a rate of 30% was good. In contrast, the survey collection rate by Oshaberi Annaiban was high at about 60%. There is also the capability of outputting tabulation results in the form of a log making it unnecessary to manually collect survey

forms. This service improves the survey collection rate through conversation and simplifies the tabulations of results, so its use in marketing tasks is anticipated.

2.4 Display of Disaster Information

In addition to the display of information at normal times, we also developed a function for receiving and displaying disaster information^{*4} at times of emergency. This information can be obtained from a local public agency corresponding to the terminal’s installation location, from the Japan Meteorological Agency, etc.

3. System Architecture

Oshaberi Annaiban is centered about AndroidTM^{*5} applications that store content such as images and video separately and that link up with a variety of Application Programming Interfaces (APIs)^{*6}

^{*4} Disaster information: Refers to earthquake bulletins, evacuation recommendations/instructions, civil protection information, evacuation-center opening information, weather alerts, tsunami alerts, etc. Disaster information is displayed using an information delivery platform provided by NTT Resonant Inc.

^{*5} AndroidTM: A software platform for smartphones and tablets consisting of an operating system, middleware and major applications. A trademark or registered trademark of Google Inc., in the United States.

^{*6} API: An interface that enables software functions to be used by another program.

(Figure 7). These constituent elements make for easy and flexible customization.

3.1 Free Dialogue through Natural Language Processing

DOCOMO “AI Agent API”^{*7} is a service that

can embed a dialogue engine in a variety of devices to provide information matching the user’s needs during free conversation using AI technology. It allows the API user to create an original agent^{*8}, which is a distinctive feature that makes for a high degree of freedom in customizing speech

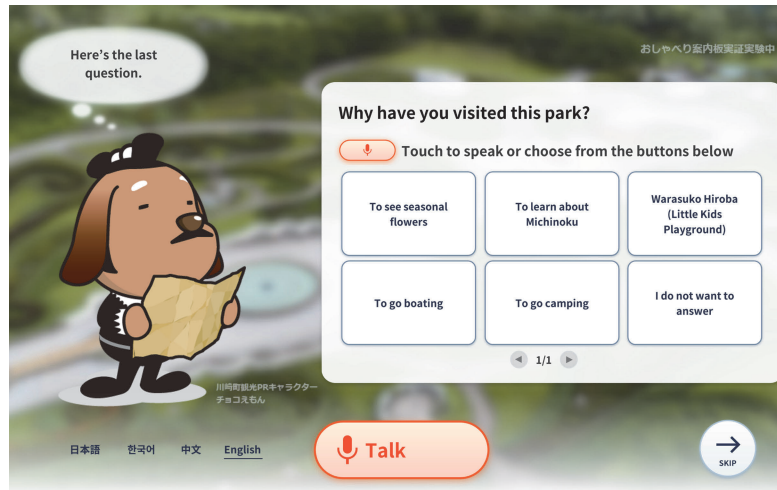


Figure 6 Survey function at Michinoku Park [7]

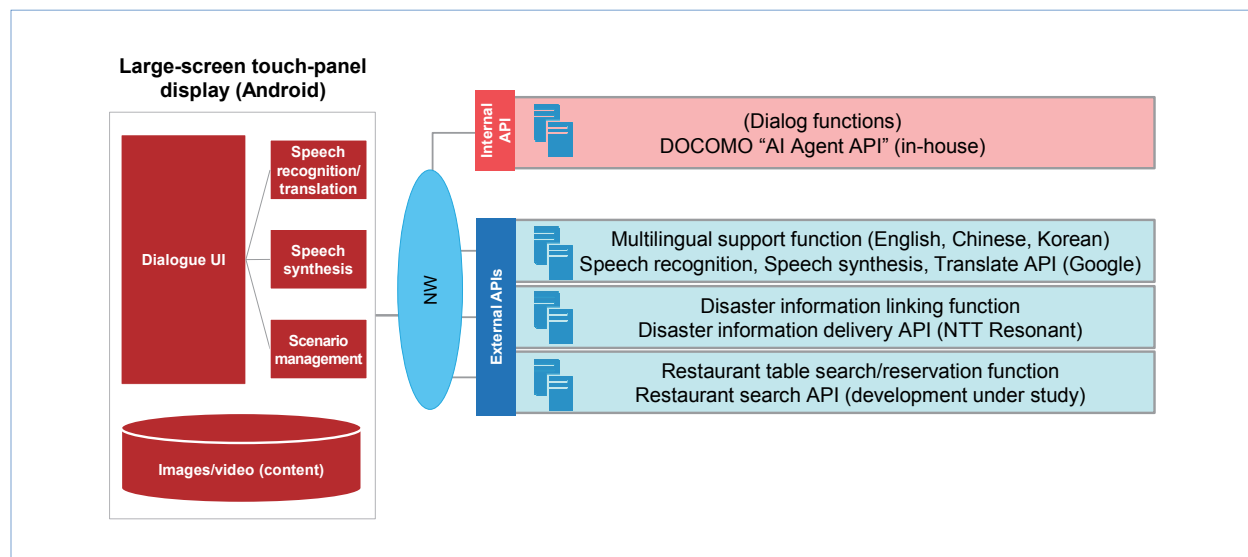


Figure 7 System configuration diagram

^{*7} DOCOMO “AI Agent API”: As part of NTT Group’s AI technologies under the “corevo®” brand name, this is an interactive AI-based Application Service Provider (ASP) service that packages a voice-based user interface. It can be used to achieve complex interactive scenarios by creating simple dialogue on a GUI and using Artificial Intelligence Markup Language (AIML).

It can also be used to respond to common inquiries as a chatbot for FAQ use that automatically generates interactive scenarios from question Q&A lists. “corevo” is a registered trademark of Nippon Telegraph and Telephone Corporation. DOCOMO “AI Agent API” is a registered trademark of NTT DOCOMO.

synthesis and using external content tailored to the environment where the API is being used, to user objectives and needs, etc.

Thanks to natural language processing in DOCOMO “AI Agent API”, Oshaberi Annaiban can now search for and provide information in a format that advances conversation with a 3D character, an actual image of a person, etc. displayed on the screen of a tablet or digital signage.

In addition, Oshaberi Annaiban links with Translate API* from Google to achieve a multilingual support function. This feature enables development of flexible multilingual support in a relatively short period of time.

3.2 Customizable Character Settings

Oshaberi Annaiban enables the information-providing character or screen background to be customized. It treats characters and tones of voice as content that can be replaced so that an implementer can easily customize the target installation

by deciding on the character to be used and setting tone of voice. In addition, the system is configured so that the selected character can operate in conjunction with DOCOMO “AI Agent API” or other third-party APIs from Google and other companies. This configuration makes it possible to display different types of interfaces for a variety of characters. It also enables flexible spatial presentation by tailoring the background image to the atmosphere of the installation location (**Figure 8**).

4. Trial Report

Since launching the project in August 2018, NTT DOCOMO has been conducting trials in cooperation with corporate-sales managers at NTT DOCOMO branch offices and stores to assess the usefulness of Oshaberi Annaiban for inbound travelers and to test the business model in terms of sales system, price setting, etc. Functional improvements have been made as a result of analyzing the needs of

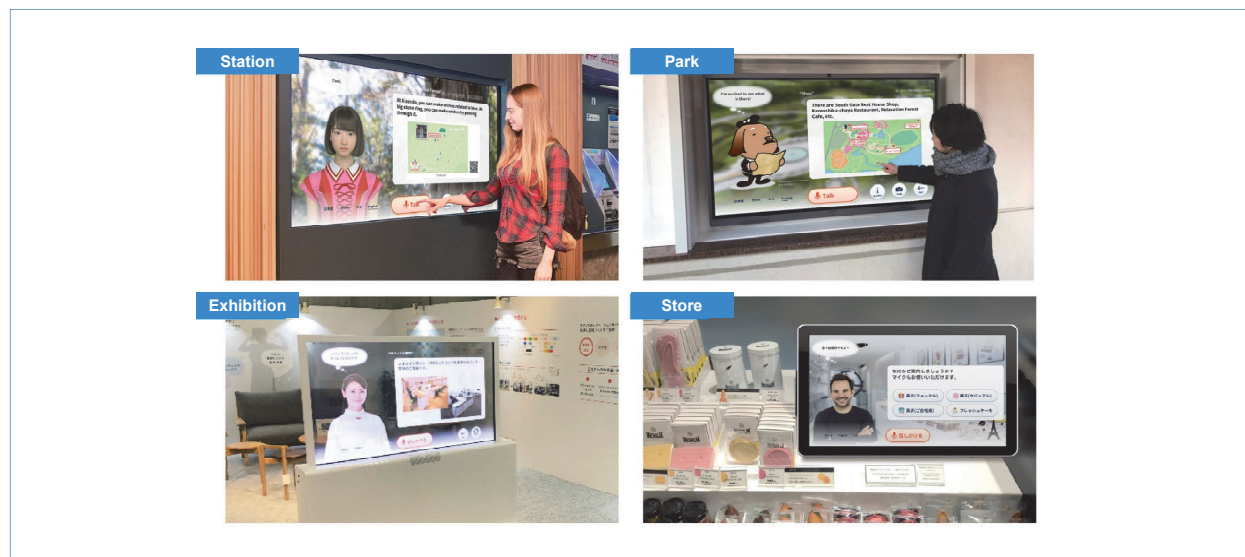


Figure 8 Character customization

*8 Agent: Software that acts as a representative of a user or other system, that functions as an intermediary among multiple elements. DOCOMO “AI Agent API” creates agents using a FAQ bot function.

* links with Translate API: At present, disaster information is displayed only in Japanese—it is not a target of multilingual translation.

implementers and of analyzing logs recording user interaction when customizing Oshaberi Annaiban. The following describes in particular a trial that NTT DOCOMO performed in Okinawa at Naha Bus Terminal.

4.1 Trial Description

Naha Bus Terminal is a transportation hub for traveling to tourist attractions within the city of Naha and throughout Okinawa Prefecture. As part of a remodeling initiative completed in October 2018, a “concierge counter” was set up to provide guidance on bus routes. However, at times of congestion, it could be seen that the counter would fill up with customers and that foreign travelers who do not speak Japanese would hesitate to talk to counter staff.

Naha Bus Terminal also welcomes many tourists

from the Asia region, but the difficulty of recruiting personnel that can provide multilingual guidance in Chinese, Korean, and other Asian languages has presented a problem. In response to the above issues, we conducted a trial to see if the installation of Oshaberi Annaiban could be effective in helping inbound travelers and reducing the workload of the concierge staff.

Oshaberi Annaiban used in this trial provided route information for four bus companies that depart from Naha Bus Terminal. The user could receive bus route information by speaking to the board and indicating what tourist site or attraction he/she would like to go to. We developed this bus-route searching function so that it could provide bus-route information by linking Oshaberi Annaiban to the Naha bus navigation site via an API (Figure 9).

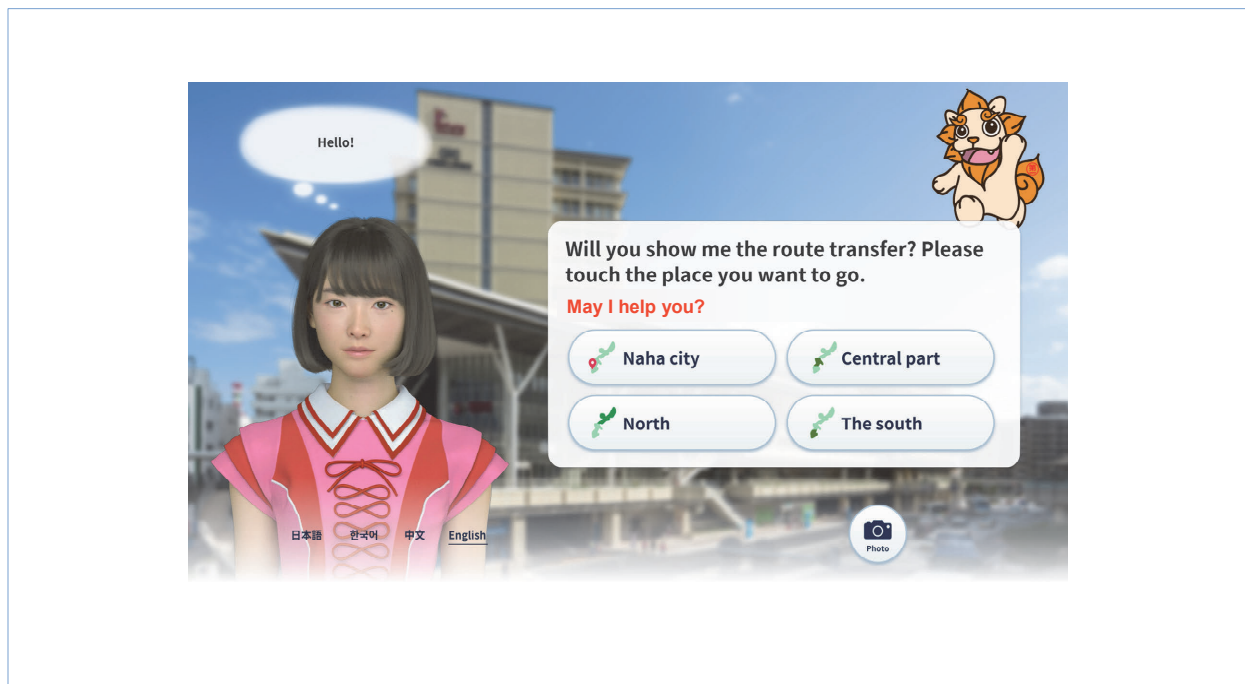


Figure 9 Application example at Naha Bus Terminal

4.2 Trial Results

The trial revealed that Oshaberi Annaiban was used when the concierge counter was crowded and that foreign visitors would ask counter staff questions to obtain more detailed information after first conducting a search on the board with its multilingual support. It was found through interviews that this coordination between Oshaberi Annaiban and counter staff helped to raise customer satisfaction. Management of Naha Bus Terminal gave this trial and Oshaberi Annaiban high praise, saying “The appropriate dissemination of information is essential to getting foreign visitors to Japan, whose numbers are increasing sharply every year, and local residents as well to use the bus more often. This trial revealed that Oshaberi Annaiban is highly effective in holding down labor expenses and making business operations more efficient.”

This trial involved the development of a new function for API-based linking with other companies’ sites, and though the scale of software development was large, the trial showed that the function could be customized for each implementer.

It was also found that achieving API-based linking with other companies’ sites in Oshaberi Annaiban could lead to the provision of information of even greater value to users through system linking with assets possessed by NTT DOCOMO business partners. In other words, the trial showed that API-based linking opens up new possibilities in customization.

On analyzing the effectiveness of support for inbound travelers using actual data obtained from this trial, we found that Oshaberi Annaiban was used in Japanese about 65% of the time and in a foreign language (English, Chinese, or Korean) about

35% of the time. As for use in a foreign language, this was a high result compared with the usage rate of around 10% obtained in trials of the same service in other regions (**Table 1**). These results showed that Oshaberi Annaiban was effective to a certain degree in providing support when guidance from the concierge counter was not available and in providing multilingual support for inbound travelers.

Next, on analyzing the change in the number of times Oshaberi Annaiban was used in one day during the trial period, it was found that a certain amount of usage could be obtained. Nevertheless, we will continue to conduct tests and make improvements to ensure that Oshaberi Annaiban comes to be used on an ongoing basis (**Figure 10**). In particular, we plan to work on increasing usage by encouraging users by some means to search out Oshaberi Annaiban at various types of locations and to use it by speaking to it and touching it.

5. Conclusion

This article provided an overview of NTT DOCOMO’s interactive and multilingual AI-based “Oshaberi Annaiban,” explained its system architecture, and

Table 1 Usage and language breakdown of Oshaberi Annaiban at Naha Bus Terminal

Language	No. of Times (Percentage)
Japanese	1,870 times (64.75%)
English	377 times (13.05%)
Korean	460 times (15.93%)
Chinese (simplified)	181 times (6.27%)
Total	2,888 times

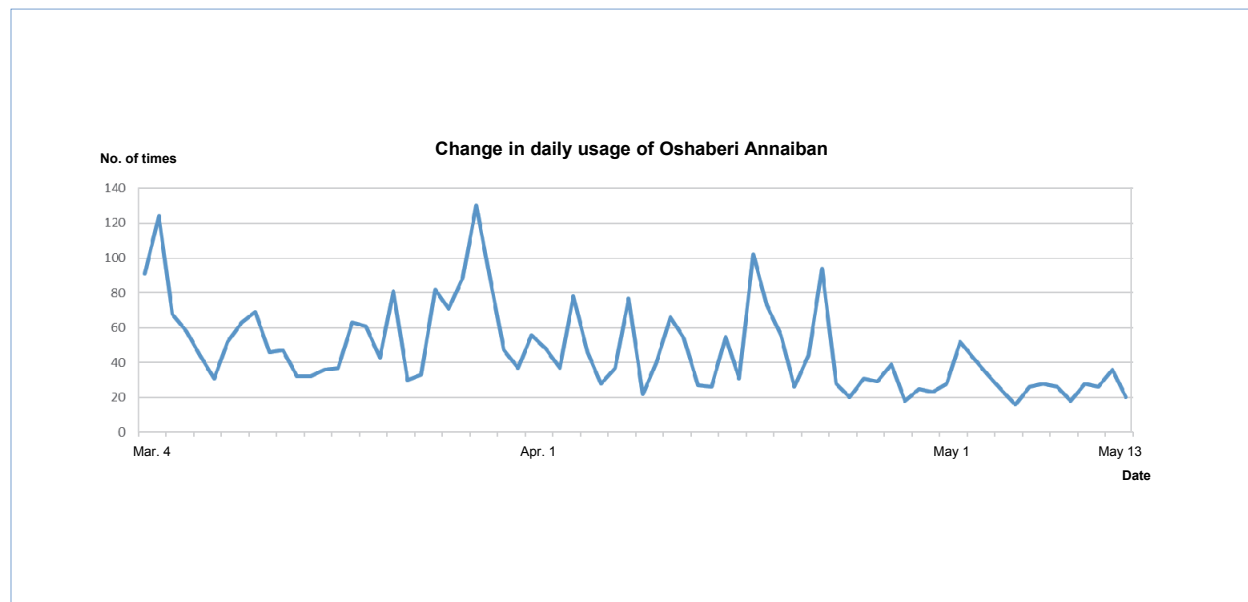


Figure 10 Daily usage of Oshaberi Annaiban at Naha Bus Terminal

described a trial held to assess its effectiveness.

The above trial was followed by the launch of Oshaberi Annaiban as a commercial service on July 8, 2019. Going forward, we plan to work together with corporate-sales managers at NTT DOCOMO branch offices and stores in proposing that Oshaberi Annaiban be applied to diverse usage scenarios such as stations, halls, and other public facilities throughout Japan, tourist attractions like hotels and leisure facilities, and commercial facilities and offices. We are committed to obtaining feedback from implementers and users to enhance the interactive and customization capabilities of this service. Finally, with the aim of providing users with highly convenient information, we will concentrate on information with high real-time properties reflecting the “here and now” at that location and will study the development of functions that can provide information on the availability of cooperating facilities, traffic congestion, etc.

REFERENCES

- [1] Japan Tourism Agency: “Tourism Nation Promotion Basic Plan and Vision Targets,” Oct. 2016 (In Japanese). <http://www.mlit.go.jp/common/001150395.pdf>
- [2] Nikkei: “Acceptance of Foreign Workers Expanded, Revised Immigration Law Takes Effect, Immigration Services Agency of Japan Established,” Apr. 2019 (In Japanese). <https://www.nikkei.com/article/DGXMZO43163430R00C19A4EAF000/>
- [3] Asahi Shimbun Company: “Acceptance of Foreign Workers Expanded, Premature Start, Shortage of Reception Desks,” Apr. 2019 (In Japanese). <https://www.asahi.com/articles/ASM3053T3M30UTIL01C.html>
- [4] Mainichi Shimbun: “Supporting Foreign Visitors During a Disaster Presents a Major Problem to a Tourism Nation,” Oct. 2018 (In Japanese). <https://mainichi.jp/articles/20181007/ddm/005/070/146000c>
- [5] Nishinippon Shimbun: “Support of Foreign Disaster Victims, Patrol Evacuation Centers, Multilingual Announcements, Kumamoto Example, Regional Organization,” May 2019. (In Japanese) <https://www.nishinippon.co.jp/item/n/509394/>
- [6] TELYUKA website.

- <https://www.telyuka.com/>
- [7] NTT DOCOMO Press Release: “Trial Begins of AI Information Board using “Oshaberi Annaiban” at Michinoku Lakewood National Government Park—Testing Effectiveness of Board in Supporting Foreign Visitors to Japan with Multilingual Information, Event Management using Visitor Surveys, etc.—,” Jan. 2019 (In Japanese).
https://www.nttdocomo.co.jp/info/notice/tohoku/page/2019/190123_01.html