

In its medium-term strategy to 2020, the "Declaration beyond", NTT DOCOMO will offer a new AI agent to revolutionize the lifestyles of its customers, and as the initial step, NTT DOCOMO announced the docomo AI Agent Open Partner Initiative, which entails co-creating voice dialogue assistance functions and hardware with partner companies. This article describes the target world of AI agent, and various related initiatives.

1. Introduction

With the spread of smartphones, and improvements to voice recognition and natural language processing technologies^{*1}, services that assist users via voice dialogue have increased rapidly. Common examples include Apple's Siri^{®*2} and DOCOMO's Shabette Concier.

Additionally, the appearance of new hardware called "smart speakers" has been gaining attention in recent years. For example, Amazon's "Amazon

©2019 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. Echo^{*3}" enables users to access a voice assistant service by talking to the "Alexa" AI system, which responds appropriately. Google also has released its "Google Home[™]" smart speaker that connects to the "Google Assistant^{™*4}" voice assist function provided by the company. LINE also released a smart speaker called "WAVE" that uses the company's "Clova^{*5}" cloud AI platform, and plans to release a smart display called "FACE".

Thus, as its medium-term "Declaration beyond" strategy to 2020, NTT DOCOMO will offer a new AI

† Currently, Nippon Telegraph and Telephone Corporation Media Intelligence Laboratories

*1 Natural language processing technology: Technology to process the language ordinarily used by humans (natural language) on a computer. agent to revolutionize the lifestyles of its customers. As the first step, NTT DOCOMO announced the docomo AI Agent Open Partner Initiative in June 2017 [1]. This initiative entails co-creating voice dialogue assistance functions and hardware with partner companies.

This article describes the features of AI agent and a platform for its achievement, and business developments and partnerships using the docomo AI Agent API.

2. AI agent World

Conventional voice assistance service design tends to include a single-personality character (hereinafter referred to as an "agent") that response to various user requests. Although such designs could be viewed as intended for agents with omniscient abilities, the functions and services that can be provided through voice assistance using the assets of only one company are limited.

Thus, initiatives to solve this issue through the provision of open platforms are ongoing.

With Alexa, it's possible to register individual functions and services called "skills". This allows developers to freely expand Alexa's functions. In other words, skills enable various companies to provide their services through Alexa.

Like Alexa, AI agent also has functions that enable other companies to provide services, but they are not exactly the same. With Alexa, other companies' services are provided through only one agent (Alexa). However, AI agent provides the freedom to individualize the characters of the agents that provide the services of each company. Thus, this enables the existence of many different agents to provide users with combined support. This is analogous to real human society in which services of companies and stores are provided by their respective staff members, and if something is unknown, then a person who knows can be asked.

AI agent is divided into two parts - main agents and expert agents.

• The main agent is positioned as the closest partner of the user, and is the first agent that responds when the user speaks to a device such as a smart speaker or smart appliance.

The main agent is not an agent provided by NTT DOCOMO, but one configured with unique characters for the devices of individual companies. These features are called "device open".

 In contrast, expert agents appear when they are called by a main agent. These are created as experts in certain fields, and could presumably be staff member agents of companies or shops for example. This function is called "service open", and enables users to call particular agents when required, providing users with the services of each company.

This variety of agents with different characters plays an important role in bringing about new value through co-creation with partner companies. In companies, and in-service provision, brand image is of paramount importance, and as a contact point, an agent has a direct connection with a brand image. Having partner companies use DOCOMO's AI agent asset and give it their own original character will enable them to provide services with the brands that they have fostered through time, without the NTT DOCOMO brand at the forefront. In

^{*2} Siri®: A trademark or registered trademark of Apple Inc. in the United States and other countries.

^{*3} Amazon Echo: Amazon, Echo and Alexa are registered trademarks of Amazon.com, Inc. and its related companies.

^{*4} Google Assistant™: Google Assistant and Google Home[™] are trademarks of Google LLC.

^{*5} Clova: Clova and WAVE are registered trademarks of LINE CORPORATION.

Japan, where the use of company mascots is particularly common, the technology will enable such companies to use their mascots as agents without having to change them.

3. docomo Al Agent API

In the docomo AI Agent Open Initiative, we are offering the docomo AI Agent API^{*6} as a central system to configure the AI agent world.

The docomo AI Agent API consists of a multipurpose dialogue engine, a proactive support engine and an IoT access control engine (Figure 1).

(1) The multipurpose dialogue engine is a voice

dialogue system that plays a central role in achieving the AI agent world. To create agents for voice dialogue, voice recognition, natural language understanding, and speech synthesis^{*7} are provided as a set. The system enables both voice and text input. A variety of models are registered to enable creation of voices suitable for the various agent characters.

(2) The proactive support engine is for collecting and analyzing diverse information, and analyzing and using various user profile^{*8} information. Collected information includes information related to user activities, and



Figure 1 AI Agent platform structure

*6	API: An interface that enables software functions to be used	*7	Speech synthesis: Technology for artificially creating speech
	by another program.		data from text and verbally reading out text.
		*8	Profile: Here, this refers to information about users place of

8 Profile: Here, this refers to information about users place of residence, tastes and interests etc. estimated from system operational history. information of a highly public nature such as weather, public transport operation and disaster information. Analysis of this information is aimed at providing the necessary information to those who need it in a timely manner.

(3) The IoT access control engine is for using IoT devices in the home via the Internet. We adopted "DeviceConnect" software to control IoT devices with different respective specifications in a unified format [2], to enable control of different devices, with the same interface specifications.

The above three engines can be mutually interconnected, so that the multi-purpose dialogue engine can access the proactive support engine to reflect user information in conversation with an agent at the right time. Also, the multipurpose dialogue engine can access the IoT access control engine to enable operation of IoT devices with voice commands.

4. Business Developments and Partnerships Using the docomo AI Agent API.

The docomo AI Agent API is a platform to bring about new value through co-creation with partner companies, and has been released to enable anyone to freely use it. Since multiple partners can use the same platform to provide their respective services, NTT DOCOMO is engaging in mutual B2C^{*9} and B2B^{*10} business developments taking advantage of this characteristic.

4.1 my daiz - "By Your Side with Your Smartphone"

Released in May 2018, my daiz is a DOCOMO B2C service that uses the AI agent platform technology, and is a pivotal service in the docomo AI Agent Open Partner Initiative.

Aiming to build one-to-one relationships with customers, this service was developed under the concept of "by your side with your smartphone" to achieve the new AI agent to revolutionize customers' lifestyles, as stated in our medium-term strategy to 2020, the "Declaration beyond".

my daiz is characterized by the proactive support engine which is used to proactively pick up and deliver suitable information to the customer. The system displays on-screen information based on profile information estimated by the proactive support engine and updates it as necessary, and analyzes customer situations from registered schedules etc. to deliver suitable information with suitable timing.

Also, my daiz includes the concept of "members". "Members" are various services using the multipurpose dialogue engine, which can be added free-of-charge, and include many services offered by partner companies. NTT DOCOMO plans to expand these "members" going forward. When using members, agents created by other companies appear as contacts, and the multipurpose dialogue engine is used for this dialogue function. Also, my daiz enables IoT infrared remote control via the IoT access control engine.

4.2 Collaboration with Other Companies Using the docomo AI Agent API.

The docomo AI Agent API can be used to build

^{*9} B2C: Business-to-consumer transactions.

^{*10} B2B: Business-to-business transactions.

services or products of other companies. In the lead up to full commercial offering planned for spring 2019, discussions are ongoing towards co-creation and collaboration with a range of companies. Here, we introduce a few of these initiatives.

YKK AP Inc. announced its "UPDATE GATE" door of the future with a built-in AI and facial authentication system. The company has used the docomo AI Agent API for the door's AI dialogue system. The door has been developed with the concept of "Update with each daily pass" and aims to enrich the daily lives of residents by providing them with required information such as weather and transport information, etc., as shown in **Photo 1**.

Photo 2 shows NTT Resonant Inc.'s "Oshieru: romance consultation bot robot". Oshieru was originally a text input-type Web service, but the company held discussions with NTT DOCOMO and used the docomo AI Agent API to turn the system into a robot for exhibitions, etc.

In this way, as well as providing voice recognition and speech synthesis as a set, the multipurpose dialogue engine can cooperate with other companies' servers to add voice input/output functions to existing text input/output services.

Like the above examples, NTT DOCOMO has engaged in a range of consultations with various companies who plan release of diverse devices that should gain attention. Above, we introduced some examples of utilization of the multipurpose dialogue engine. Similarly, we had been taking inquiries on utilization of the proactive support engine and IoT access control engine, and we expect further usage of the docomo AI Agent API in various scenes in the future.



Photo 1 Usage image of the "UPDATE GATE" door of the future



Photo 2 "Oshieru" AI romance consultation robot

4.3 Fostering a Developer Community

For docomo AI Agent API to become a platform supporting a wide range of services, it's necessary for both partner companies providing services and developers building the systems for those services to understand the advantages of docomo AI Agent API and how to use it. For this reason, NTT DOCOMO is putting efforts into fostering a developer community.

.....

First of all, NTT DOCOMO has opened a trial site for developers [3]. This site provides information necessary for development such as instructions on how to use docomo AI Agent API and sample code information, and enables developers to actually try out the docomo AI Agent API.

NTT DOCOMO also has been holding study meetings and hackathons^{*11} around Japan that have been attended by more than 1,000 people to date.

Going forward, NTT DOCOMO plans to proactively disseminate information to developers around Japan as well as provide them with rich content including a wide range of development examples and sample code.

5. Conclusion

This article has described the features of AI agent and a platform for its achievement. The article has also introduced services provided by NTT DOCOMO using the docomo AI Agent API, products achieved through collaboration and cocreation with other companies, and initiatives to foster a developer community.

NTT DOCOMO has begun studies with partner companies on specific products using this API to

strengthen solutions. However, many issues with voice user interfaces (VUI) remain, such as the fact that design methods have not been established. Going forward, NTT DOCOMO will take initiatives to solve these issues to promote usage of the docomo AI Agent API.

Lastly, NTT DOCOMO planned the concept of this AI agent initiative around June 2016. At the time, this in-house project was called "Sebastien" and was developed by a mere seven people. However, having gained cooperation from many others, the project was included in the medium-term strategy and became a company-wide initiative in less than a year. Just as the name "Partner Initiative" suggests, in the future we expect these initiatives to bring about a wide range of solutions through cooperation with the people in a wide range of our partner companies.

REFERENCES

- NTT DOCOMO Press Release: "DOCOMO Announces docomo AI Agent Open Partner Initiative," Jun. 2017. https://www.nttdocomo.cojp/english/information/media_ center/pr/2017/0623_00.html
- [2] Device WebAPI Consortium Web site. https://en.device-webapi.org/
- [3] SEBASTIEN Web site (In Japanese). https://docs.sebastien.ai/