Current and Future R&D at NTT DOCOMO for the 5G Era Current State and Progress in Each Area Trends in AI and Big Data Analysis

Service Innovation Department Masayuki Tsuda Research Laboratories Wataru Takita Innovation Management Department Tomovoshi Oono

In recent years, the attention given to AI has increased dramatically, and NTT DOCOMO is continuing its various initiatives to create value and solve social issues with AI. This article describes the significance of AI in DOCOMO's thinking, and DOCOMO's AI core technology developments and AI platform strategies. Also, this article introduces "AI agent," "AI taxi," "AI stamp rally" and "disease prevention and early discovery" as specific examples of DOCOMO's AI initiatives. Finally, this article describes the future outlook for the "Beyond 2020" vision.

1. Introduction

We are now in the third Artificial Intelligence (AI) boom after the two previous booms since AI first appeared in 1956. This is because it's now possible to handle big data at practical costs thanks to advances in computing power and algorithms.

There are three specific reasons why the boom in AI has occurred, as follows: (1) Increases in the amount of available data

Data now exists at an effective level for machine learning due to the increase in the amount of data available with the spread of digital devices such as smartphones, cameras, and PCs, and information publishing services such as social networking services, User Generated Contents (UGC)*¹, and video sites.

(2) Improvement of computing power The spread and increase of hardware

*1 UGC: A generic name for contents created by users.

^{©2018} 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.

environments such as General Purpose computing on Graphics Processing Unit (GPGPU)*² that enable high-speed, large-scale floatingpoint computations, cloud computing environments that are easily scalable to suit different purposes, and high-performance memory, Solid-State Drives (SSD), and optical networks to handle high speeds with infrastructure have made it possible to use large-scale computing resources cheaply.

(3) Advances in learning algorithms and the spread of libraries^{*3}

In addition to the dramatically improved performance of learning through "deep learning^{*4}," one of the types of machine learning technologies, the open sourcing of software libraries such as TensorFlow^{®*5} and the spread of library environments that enable easy trial-and-error have accelerated the machine learning Plan, Do, Check, Act (PDCA) cycle^{*6}.

With these factors, the focus on AI has increased in academia, business and the mass media, and a large movement of investment and human resources into these fields from all over the world has occurred, resulting in the wide spread of AI.

Generally, AI applications are split into two types - "general-purpose" and "specific-purpose."

Theoretically, general-purpose AI is not restricted to particular work or tasks, but has generalized capabilities similar to or greater than humans, and can apply its own capabilities to respond beyond the scope of its specifically programmed functions. A typical example is an autonomous robot with its own will that appears in animation and so forth.

*2 GPGPU: The use of GPUs ordinarily used in computers for rendering and other types of image processing for other types of applications. GPGPU excels at parallel distributed processing.

*3 Library: A collection of high-versatility programs which are reusable.

*4 Deep learning: Machine learning using a neural network with

However, there are no current prospects for the realization of such technology due to hurdles such as the difficulty of gathering correct training data over a wide area.

In contrast, specific-purpose AI is designed to perform a specific task, and is specialized with one function such as the games Go or Shogi or for tasks such as image recognition, machine translation or dialogue. A typical example of specific-purpose AI is Google's "AlphaGo*7" developed using deep learning, and renown for beating professional Go players. Currently, the AI that has spread in society and through various fields of business is this specific-purpose type. This does not simply entail the use of AI technology - more important is the way it is used in combination with domain knowledge*8. In addition, because better results can be gained efficiently by machine learning with large amounts of data in some cases rather than humans thinking up the procedures to solve a problem using their knowledge and experience, securing overwhelming amounts of data is becoming ever more crucial.

NTT DOCOMO has also built a specific-purpose AI (known as "narrow-AI") platform for analyzing and using data, and we are which it is continuing to apply it to various areas of business with the aim of creating value and solving social issues. This article describes the future direction of AI at NTT DOCOMO.

2. NTT DOCOMO Initiatives

2.1 The Significance of AI Technology in NTT DOCOMO's Thinking

With AI technology, we aim to achieve the ultimate personal agent, and solve social issues and

a multi-layered structure.

^{*5} TensorFlow[®]: A deep learning programming framework. A registered trademark of Google LLC.

make industry more efficient (Figure 1).

As part of our medium-term strategy to 2020, the "Declaration beyond," the former initiative aims to achieve an AI agent that will transform the lifestyles of our customers and become an indispensable item for "personal" lifestyles in various scenes. This agent will use AI technologies for naturallanguage dialogue and behavior prediction to accurately understand information that customers want and provide optimized suggestions with appropriate timing. An example of this is the "my daiz^{TM*9}" AI agent service.

The latter is the "+d^{*10}" (co-creation) initiative to generate new value by using the core technologies and big data required to achieve the aforementioned personal agent combined with the big data and domain knowledge of partner companies. An example of this is the "AI taxi" to predict future taxi demand using deep learning by combining NTT DOCOMO's population statistics data generated from cellular networks with the operational data of a taxi company [1], described later.

2.2 AI Core Technology Development

The main AI technology field is broadly divided into four categories:

- Recognition: Recognition of images, voice and language etc.
- Judgment/optimization: Demand prediction, error detection, optimal route discovery, judgment support in medicine/law etc.
- Generation: Generation of images, voice and text etc.
- Control: Control of autonomous driving and robots etc.

At NTT DOCOMO, we are developing competitive AI core technologies focusing on recognition, judgment/optimization using big data collected through NTT DOCOMO's customer base. These developments also include applying technologies to solve current issues, polishing technologies while





- *6 PDCA cycle: A method of ensuring smooth running of business. The PDCA cycle entails repeatedly and continually running through the four steps of (1) Plan (planning), (2) Do (performing), (3) Check (measuring results) and (4) Act (making improvements).
- *7 AlphaGo: A trademark or registered trademark of Google LLC.
- *8 Domain knowledge: Knowledge or information such as trends and so forth about a target business or industry.
- *9 my daiz™: my daiz and the my daiz logo is a trademark of NTT DOCOMO, INC.
- *10 +d: Name of the NTT DOCOMO initiative for creating new value together with partner companies.

acquiring feedback, and pursuing cutting-edge technologies through the use of NTT group technologies and partnerships with top class research institutions both in Japan and abroad, such as participation in the Stanford Data Science Initiative (SDSI)*¹¹ run by Stanford University [2].

2.3 AI Platform Strategies

Regarding the aforementioned AI core technologies, we are progressing with the creation of a platform in co-creation with partners to expand the scope of industries using the technology. This process consists of:

- "Point" stage: Generating specific cases for each core technology and accumulating knowhow as "points"
- (2) "Line" stage: Deployment and development like "lines" including know-how in target business areas of each core technology
- (3) "Surface" stage: Deployment and development like "surface" with a platform organically linking core technologies

As described above, we are aiming to create a platform with deployment and development from points to lines, and then lines to surfaces (Figure 2).

The platformization of natural-language dialogue technology is being driven through the process described above. This technology was used in 2012 with "Shabette Concier," after which NTT DOCOMO began offering a natural-language dialogue platform after improving performance and accumulating know-how using data collected from several hundred million instances of use. In 2015, we developed the communication tov "OHaNAS^{® * 12}" jointly with Tomy Company. Then in 2016, we began providing the "Oshaberi Robot for Biz," a customizable service to which companies can add their unique scenarios, and then in 2017, we jointly developed the "ATOM®*13" communications robot with Kodansha, Tezuka Productions, FUJI SOFT and VAIO. In this way, our developments went from "point" to "line." Then in 2018, our developments moved from "line" to "surface" with the deployment of the AI agent API*14, which is the core of the AI agent platform, and was achieved by organically linking



Figure 2 The platform promotion process

*11 SDSI: A joint industry-academic program to support collaboration between the Graduate School of Engineering at Stanford University and companies involved in data science and big data.
*12 OHaNAS[®]: A registered trademark of TOMY Company, Ltd.
*13 ATOM[®]: A registered trademark of Kodansha Ltd.
*14 API: A general-purpose interface for using functions and data.

the separate prediction engine and IoT access control engine AI core technologies.

3. Specific Examples of NTT DOCOMO AI

3.1 Al Agent

1) AI Agent API

As part of our medium-term strategy to 2020, the "Declaration beyond," we are developing and providing the "AI agent API" to form the core of the platform to achieve the new AI agent that will transform the lifestyles of our customers [3]. This API is the basic system for creating the new AI agent, and consists of the prediction engine, a multipurpose dialogue engine, and the IoT access control engine (**Table 1**). Part of the NTT group AI technology "corevo^{®*15}" is used for the prediction engine and the multipurpose dialogue engine.

Also, we have been promoting the "DOCOMO AI Agent Open Partner Initiative" as a joint development for a new "open for services and devices" service provision style based on the audio interface by making this API openly available. With this initiative, we are aiming for the following three points by making the services and the API for device providers that NTT DOCOMO has fostered to date openly available (**Figure 3**).

• To create new service experiences for endusers through text and audio

Table 1 Characteristics of the three engines	
Prediction engine	Behavior analysis technology enables timely provision of information tailored to individuals.
Multipurpose dialogue engine	Natural language processing technology enables provision of services through natural dialogue.
loT access control engine	Device Web API technology enables control of IoT devices with different communication standards from a single application.

Device Web API: A mechanism that achieves linkage between various devices discussed and studied by the Device Web API Consortium, which consists of 120 participant companies (as of July 2018).



Figure 3 Generating new services by making AI agent API openly available

*15 corevo[®]: A registered trademark of Nippon Telegraph and Telephone Corporation.

- To build win-win business relationships with partner companies
- To shorten the development period for AI agent services
- 2) The my daiz AI Agent Service

Using the aforementioned AI agent API etc., NTT DOCOMO began providing the "my daiz" AI agent service in the spring of 2018, which enables NTT DOCOMO or partner companies to provide optimized suggestions with appropriate timing by the AI agent understanding the information that customers seek.

This service makes possible to provide customers with optimized suggestions with appropriate timing by understanding each customer (Figure 4) such as:

• Sounding an alarm early, once weather and traffic information is understood if it looks like commuting to work or school might take time due to rain

- Purchasing insufficient food supplies through Internet shopping and having them delivered to the home
- Providing weather and peripheral information of destinations as well as travel bookings

3.2 AI Taxi

For taxi drivers to increase their income, it's important to find passengers efficiently in unfamiliar areas or timeslots to reduce the time the taxi is vacant and maximize the time it's engaged. To address this issue, NTT DOCOMO has begun commercially providing "AI taxi" developed technology to predict taxi passenger demand in particular areas using past taxi operational data and NTT DOCOMO's real time population statistics data generated from cellular networks. This enables taxi drivers to keep track of changing passenger demand in real time and operate their taxis more efficiently.

This technology uses data input of real time



Figure 4 my daiz

population statistics, weather forecasts, past taxi passenger rates, and uses deep learning to effectively extract latent feature values^{*16} and then perform regression. Then in combination with the Vector Auto Regressive model (VAR, also multivariate autoregression model), the technology estimates taxi demand in 500 m mesh^{*17} units 30 minutes in the future (the number of passengers predicted to be picked up in the mesh) (Figures 5, 6).

In field trials, we confirmed 93 to 95% accuracy of these predictions, and there was an approximately JPY 1,400 average sales increase per day per driver for drivers using the prediction system.

3.3 AI Stamp Rally

The way people acquire tourism information has



Figure 5 Demand prediction and taxi movement



Figure 6 Anticipated effects of AI taxi

*16 Feature values: Values extracted from data, and given to that data to give it features.

*17 Mesh: A grid dividing the country into square sections, along lines of latitude and longitude.

changed dramatically with the spread of smartphones and social networking services. Hence, disseminating information about sightseeing requires not only publication in conventional travel magazines or posting on Web sites, but also requires new and more efficient methods.

To provide value as a sightseeing information dissemination method using dialogue AI technology, NTT DOCOMO developed "AI Stamp Rally^{TM*18}" with partners, which disseminates sightseeing information in conversation with dialogue AI (chat bot) on social networking services with many users and achieves a stamp rally using image recognition AI, and performed trials of this service in Ishinomaki city in Miyagi Prefecture from March to August 2018 (**Figure 7**). This service is available on social networking services prior to sightseeing so that users can find out recommended information about sightseeing or events while enjoying conversation with a character (dialogue AI) before setting out. Users can tour around sightseeing spots while enjoying a bingo rally by taking photographs, and easily find out the closest dining and sights from location information. After sightseeing, it's also possible to find out about new sightseeing information such as events.

Main advantages compared to conventional sightseeing spot information dissemination are:

- Effective information dissemination: Information disseminated to individuals in conversation by dialogue AI
- (2) Attract tourists, generate migration: Attract tourists and generate migration to sightseeing spots with stamp rallies
- (3) Information acquisition: Acquire tourist attribute information and migration information



Figure 7 AI stamp rally service image

*18 AI Stamp Rally™: A trademark of NTT DOCOMO, INC.

through conversation with dialogue AI and stamp rallies

(4) Generate constant contact: Create contact through connection on social networking services to continually disseminate information

Figure 8 describes an image of the AI stamp rally Web API linkage. With the Web API, NTT DOCOMO's dialogue AI links with image recognition AI to recognize images of sightseeing spots with deep learning, and stamp rally systems developed by partners. According to the details of the conversation, the dialogue AI responds to the users, for example, by disseminating sightseeing information or providing stamp rally functions. In future, this service will make it possible to achieve more advanced and effective sightseeing information dissemination by linking existing sightseeing spot services and data using the Web API.

3.4 Disease Prevention and Early Discovery

Many diseases, including lifestyle-related diseases, are caused by a complex interaction between congenital genetic factors and acquired environmental factors such as lifestyles and living environments. To prevent the onset or progression of diseases, it is desirable to predict the risk of future onset of diseases by evaluating the current state of the individual's health from periodic monitoring for any signs or abnormalities of physiological data in which genetic and environmental factors are reflected.

Pregnancy complications^{*19} including preterm birth, gestational diabetes and hypertensive disorders of pregnancy are social issues that affect one in five pregnant women in Japan, but so far no fundamental preventive or therapeutic methods have been discovered. Taking the initiative to address this issue, NTT DOCOMO has embarked on "Maternity Log Study" collaborating with Tohoku University Tohoku Medical Megabank Organization to investigate the differences between pregnant women who develop pregnancy complications and those who don't to exhaustively analyze big data on both genetic and environmental factors [4] [5]. This initiative entails working on identifying risk factors for pregnancy complications



Figure 8 AI stamp rally Web API linkage image

*19 Pregnancy complication: Refers to diseases associated with pregnancy, and can include diseases that existed prior to a person becoming pregnant, or diseases that occur during pregnancy. and constructing prediction models for disease onset by integrated analysis of a wide variety of data such as the DeoxyriboNucleic Acid (DNA)^{*20} and RiboNucleic Acid (RNA)^{*21}, metabolites in urine and plasma^{*22}, oral cavity bacterial flora^{*23}, life log including activity and sleep etc., and diagnostic information such as medical records of pregnant women (**Figure 9**). In the future, we would like to contribute not only to the improvement of the health of pregnant women but also to the extension of the healthy life expectancy^{*24} of more people by using the knowledge obtained from this research to prevent the onset of disease in both non-pregnant women and men.

4. Future Outlook

Toward the Beyond 2020, we envisage the arrival

of a convenient and comfortable society in which AI is integrated throughout society in various ways (Figure 10). Thus, we will work with our partners to develop core technologies and services so that NTT DOCOMO can play a driving role in the realization and growth of this coming society.

REFERENCES

- S. Kawasaki et al.: "AI Taxi Passenger Demand Prediction Technology to Optimize Transportation Operations," NTT DOOCOMO Technical Journal, Vol.20, No.2, pp.14-21, Nov. 2018.
- [2] NTT DOCOMO Press Release: "DOCOMO Joins Data Science Affiliates Program of Leading U.S. University
 Will collaborate with big-data and AI faculty -," Jun. 2018.

https://www.nttdocomo.co.jp/english/info/media_center/ pr/2018/0608_00.html

[3] NTT DOCOMO Press Release: "DOCOMO Announces docomo AI Agent Open Partner Initiative - New "AI



Figure 9 "Maternity Log Study" overview

*20 DNA: A substance carrying genetic information in an organism and consisting of deoxyribose and phosphoric acid, and four types of nucleobases: adenine, guanine, cytosine, and thymine.
*21 RNA: A substance transcribed using DNA as a template, and composed of ribose and phosphoric acid, and four types of nucleobase a



Figure 10 A convenient and comfortable society in which AI is integrated throughout society in various ways

Agent API" will help accelerate increasing collaboration -," Jun. 2017.

https://www.nttdocomo.co.jp/english/info/media_center/ pr/2017/0623_00.html

- [4] D. Ochi et al.: "Initiative toward Prevention and Early Detection of Disease Using Healthcare Data and Genome Analysis," NTT DOCOMO Technical Journal, Vol.17, No.3, pp.23-29, Jan. 2016.
- [5] NTT DOCOMO Press Release: "Participant recruitment for research on prevention of pregnancy complications complete - integrated analysis of the world's largest amount of life log and physiological data begins," Nov. 2016 (In Japanese).

https://www.nttdocomo.co.jp/info/news_release/notice/2016/11/15_00.html

*23 Bacterial flora: A collection of bacteria growing in a certain

- environment.
- *24 Healthy life expectancy: The expected period of good health in daily life.