

## NTT Research and Development for the Age of Transformation

**Hiromichi Shinohara**

***Senior Executive Vice President and  
Senior Vice President of Research and  
Development Planning, NTT***



### Overview

This article introduces NTT's research and development (R&D) activities designed to create new value by providing cutting-edge technologies and also through collaboration with various partners toward realizing a better society. This article is based on the keynote lecture presented by Hiromichi Shinohara, NTT Senior Executive Vice President and Senior Vice President of Research and Development Planning, at NTT R&D Forum 2016 held February 18–19, 2016.

*Keywords: R&D strategy, artificial intelligence, IoT*

### 1. Introduction

In response to increases in the number of security threats and in the volume of traffic on the network, the various research and development (R&D) laboratories at NTT (hereafter, NTT R&D) are carrying out R&D to address social issues, reinforce industrial strength, help revitalize local economies, and thereby build a better society by providing advanced technologies that enable information and communication technology (ICT) to further penetrate our lives.

Toward these ends, NTT R&D is focused on networking, cloud, security, and basic technologies that provide foundations for those technical areas. This article first addresses artificial intelligence (AI) and the Internet of Things (IoT), both of which have come under the spotlight in recent years.

### 2. Directions for AI technologies

AI has become a hot topic and aims at duplicating the intellectual faculties of humans using machines.

The workings of the human brain can be regarded as consisting of three processes: 'recognition/understanding of the external world,' 'inferring/judging,' and 'providing feedback to the external world.' Based on this understanding, AI is beginning to be used in a number of different fields. What are the intellectual faculties of humans? Human activity for perceiving/recognizing the external world means recognizing not only objects and people but also human emotions and nuances of human expression. Naturally, we make the most of all five senses. Human activity for *judging and generating answers* does not necessarily mean capturing and processing all available information but rather, concentrating on the information needed to facilitate judgment. An AI machine playing a game of *go* can determine its next move but cannot explain why it has selected that move. Human activity for *providing feedback to the external world* includes the ability to take actions that please others and to communicate with them while taking care to sense what is pleasing, which can vary from person to person and also depends on the historical background.

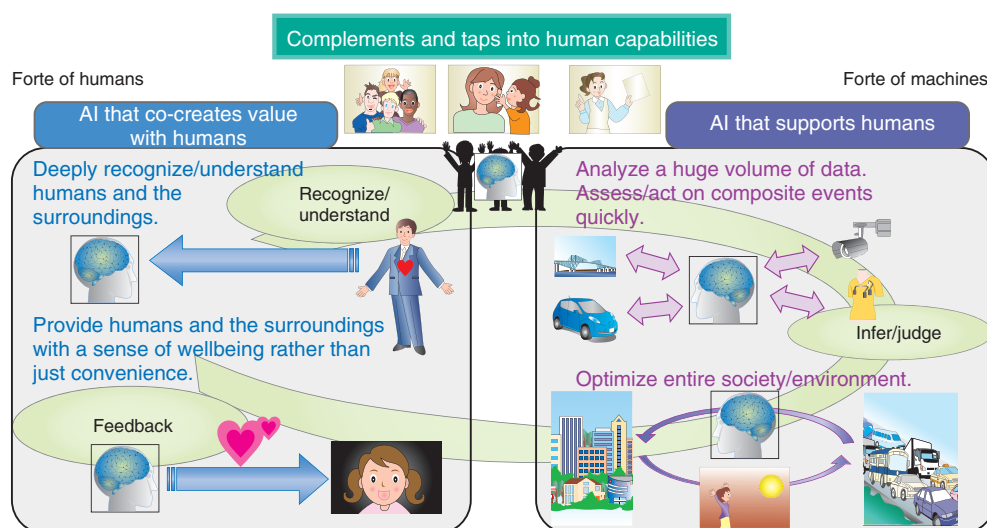


Fig. 1. AI targeted by NTT R&D.

When we humans touch a cup containing a piping hot beverage, then instantly let go of it to avoid being burned, we do so intuitively, without thinking. Similarly, there is a gap between the AI that is currently attracting interest and the intellectual faculties of humans.

With this gap in mind, NTT R&D believes that the role of AI is not to emulate the intellectual faculties and thinking of humans but, rather, to complement and tap into human faculties. In other words, our concept of AI's role is as a facility to substitute for or support some human activities in such ways that it can co-exist with humans or co-create value with humans to enrich their lives (Fig. 1). Naturally, there are areas in which machines are more effective and/or efficient and other areas in which humans are more adept. For example, machines are better at analyzing enormous volumes of data and rapidly assessing composite events. It is necessary to thoroughly polish this ability and incorporate it into NTT R&D's AI. On the other hand, humans are still way ahead of machines in recognizing and profoundly understanding other human beings and the environment, and in providing agreeable feedback to humans. NTT R&D will polish its technologies in order to bring its AI closer to those human faculties. Thus, our goal is to combine the AI that supports humans and the AI that co-creates value with humans in order to develop a type of AI that complements and taps into human faculties.

## 2.1 Four types of AI targeted by NTT

We have defined four specific types of AI—Agent-AI, Heart-Touching-AI, Ambient-AI, and Network-AI—and we are undertaking R&D in line with these categories (Fig. 2).

Agent-AI is the closest to the type of AI that is attracting interest today. It interprets information generated by humans, understands the surroundings, intentions, and emotions of humans, makes inferences based on enormous volumes of data, and enables robots to conduct a sophisticated dialogue with humans. A major feature of Agent-AI is that it expands the parameters of our daily lives, for example, substituting for or supporting the routine work done in a contact center, supporting intellectual activities such as medical diagnosis, and assisting humans in moving and communicating.

Heart-Touching-AI goes beyond the type of AI that is based on superficial knowledge. It interprets subconscious human mental and physical conditions and gains an understanding of the deep psyche, intellect, and instincts of humans in order to help build a society that provides a sense of well-being rather than merely convenience.

Ambient-AI seeks to give objects a kind of knowledge and make them behave in an organized manner. NTT R&D proposed the concept of *ambient intelligence* in 2006. In those days, computer power was lacking, and no satisfactory cloud environment was available. Now that these conditions have improved, we believe that Ambient-AI has become a realistic

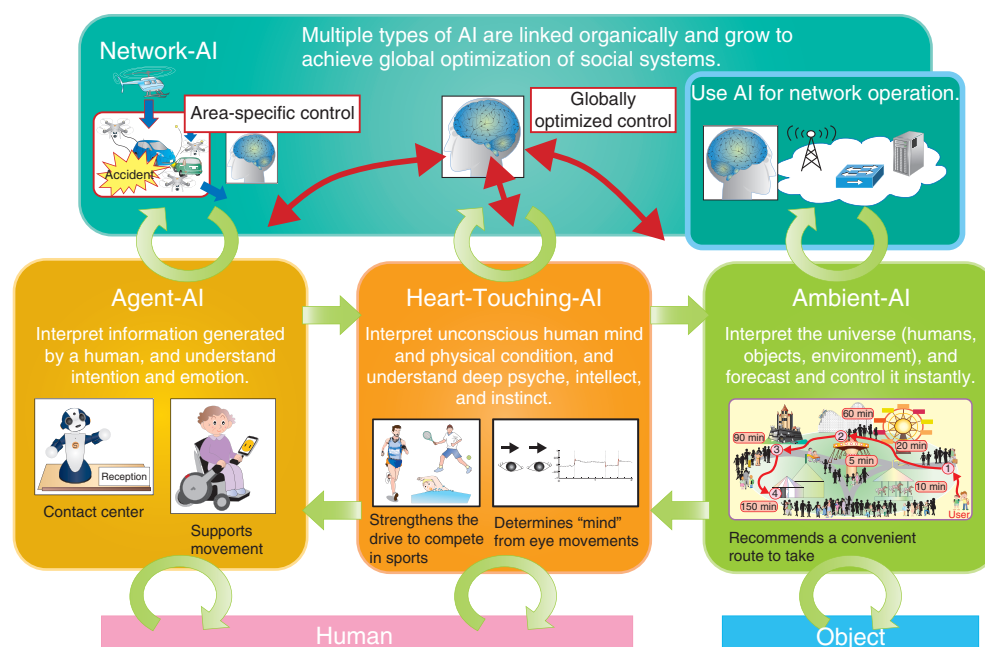


Fig. 2. Four types of AI in NTT R&D.

proposition as an evolved form of ambient intelligence. It encompasses IoT technology in that it interprets humans, objects, and the environment, and forecasts and manages them instantaneously.

Network-AI includes two concepts. The first is to apply AI to network operations so that, for example, cyber-attacks can be detected early and signs of impending failures in communication equipment can be discovered, resulting in stable network operation. The other is to interconnect a number of AIs and make them grow in such a way that they will optimize the overall social system. The AI being discussed today is designed to substitute for and tap into the thinking of just a single person. In contrast, Network-AI is aimed at having various AIs work together in order to acquire collective knowledge and thereby create new value. For example, in times of disaster, it is important to make decisions and act based on information in a small area, but it is also important to seek total optimization from a nationwide perspective. In short, Network-AI is aimed at emulating what is done by the human cerebellum and cerebrum.

Implementation of these four types of AI requires three technical elements. The first is to recognize and understand humans and objects by decoding, rather than simply measuring, data about them. The second is to infer and judge through exploration rather than simple analysis, which is the case with machine

learning. The third is to provide feedback in a pleasant manner by design rather than simply by control. We will build a circle of these technical elements on three foundations: data science, human science, and a mathematical base (**Fig. 3**).

## 2.2 Main AI technologies

Representative AI technologies being studied by NTT R&D are introduced below.

In the area of Agent-AI, NTT R&D's distortionless noise reduction technology has been combined with our deep-learning speech recognition technology. The resultant technology attained the world's top speech recognition level among 25 competitors in CHiME-3 (the 3rd CHiME Speech Separation and Recognition Challenge), an international technology evaluation event in which participating organizations competed in the ability to recognize English speech in a noisy environment (**Fig. 4**). Our intonation and accent conversion of speech can convert a speech in standard Japanese into one with the intonations of, say, an Osaka or Nagoya dialect without losing naturalness. It enables a robot to speak with pronunciation similar to that of a human.

In the area of Heart-Touching-AI, the body- and mind-reading technology can infer the condition of a person—for example, how deeply he is concentrating or if he is sleepy—from tiny eye movements or pupil

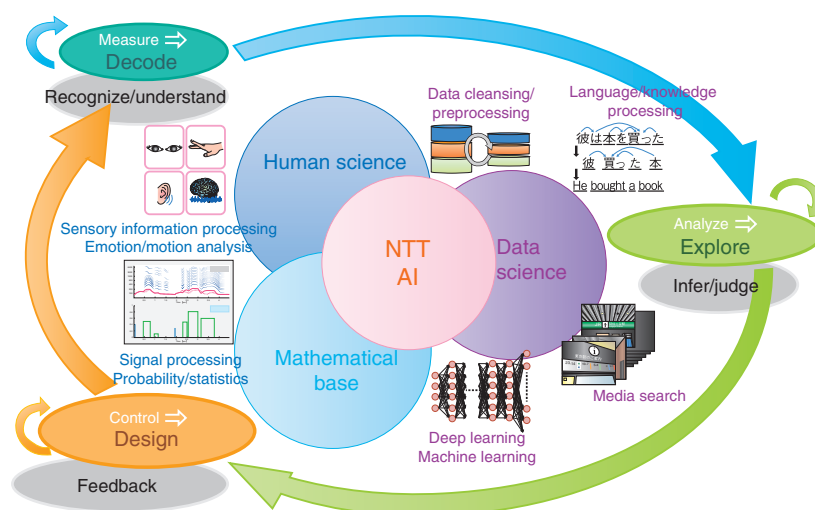
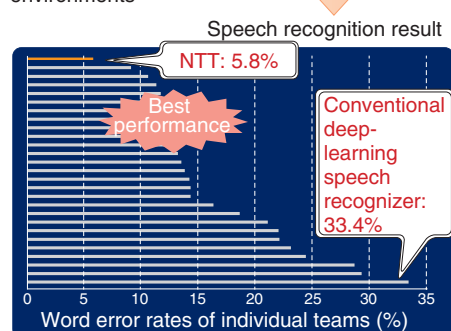


Fig. 3. Technical elements that support AI.

NTT attained the top level among 25 participating organizations in the CHiME-3 Challenge international speech recognition event.



Speech recorded in diverse environments



NiN: network-in-network  
RNN: recurrent neural network  
WFST: weighted finite-state transducer

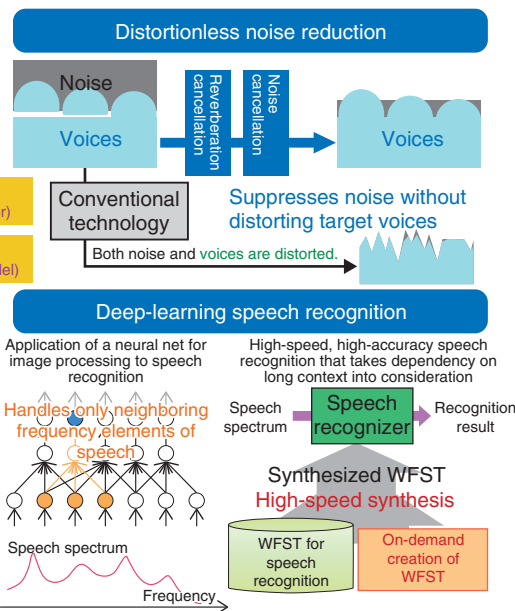


Fig. 4. Agent-AI technologies.

reactions. We are also studying how a person can train his or her brain to be able to achieve victory in a sports contest, based on brain science that involves analyzing the muscle activity patterns and heart rates of professional sports players. We are also consider-

ing use of tactile information presentation technology (Buru-Navi) to convey feedback not only through speech or video but also by appealing to the tactile sense with vibrations.

In the area of Ambient-AI, spatio-temporal analysis/

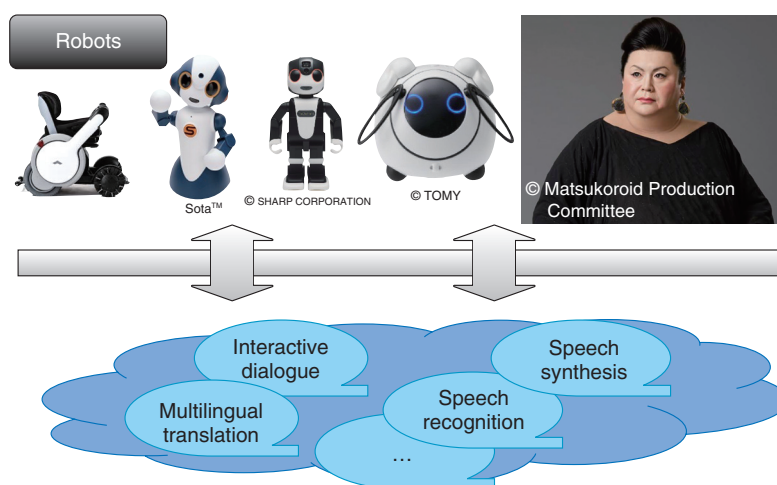


Fig. 5. Cooperation example of robot and AI.

mathematical optimization technology makes it possible to predict the near future of just a few minutes from now rather than predicting a vague future.

In the area of Network-AI, in which AI is applied to networks, resource and QoE (quality of experience) optimization technology is being studied with a view to achieving stable telecommunication service quality through proactive control of various traffic-fluctuating factors.

### 2.3 AI utilization cases

NTT R&D's activity related to the use of AI in robots is introduced here. Robot hardware will continue to evolve in different forms. Some robots will evolve to converse with people, and some will evolve to support construction work. We believe that focusing on a robot designed for a specific purpose and trying to make it smarter will not bring far-reaching change to society. Our approach is to apply our AI technologies to all kinds of robots. We will advance and enhance the capabilities of robots by providing common technologies such as speech synthesis, interactive dialogue, and multilingual translation to a variety of robots, as shown in Fig. 5.

Agent-AI technology is now being used in contact centers. AI assists operators not merely in comprehending what the customer says but, more importantly, in discriminating whether or not the customer is angry, which is considered to be particularly important for contact center operators. This may be peculiar to the Japanese, but we are working on understanding two distinct types of anger: hot anger, which is easy to recognize because the angry person shouts, and

cold anger, which is hard to recognize because the person pretends to be calm while silently seething. NTT R&D's AI makes it possible to recognize a caller's cold anger by analyzing the conversation flow and vocabulary. The NTT Group already offers systems that adopt this technology in contact centers and is convinced of its effectiveness.

### 3. Directions for IoT technologies

NTT R&D's activities related to IoT are described below from three aspects: sensing, security, and networking/clouds (Fig. 6).

We believe that penetration of a variety of sensors into our lives should not impose extra pressures on people or society. Sensors should merge with humans and society in a natural way. We express this requirement as *natural*. An example of a technology that reads information about a human is one that makes it possible to read the heart rate and muscle activity of a person, who needs only to wear a shirt or support garment made of the functional material called *hitoe*. One way to read information about an object is to install a sensor. It is also necessary to consider using a drone to photograph an object and analyzing the captured images. Given that it is not *natural* if an IoT device must be replaced each time its measurement or performance requirement changes, we are studying IoT device virtualization technology, which would allow an IoT device to be reprogrammed from a remote site.

When we provide added value using data from IoT devices, it is important to ensure data security. We



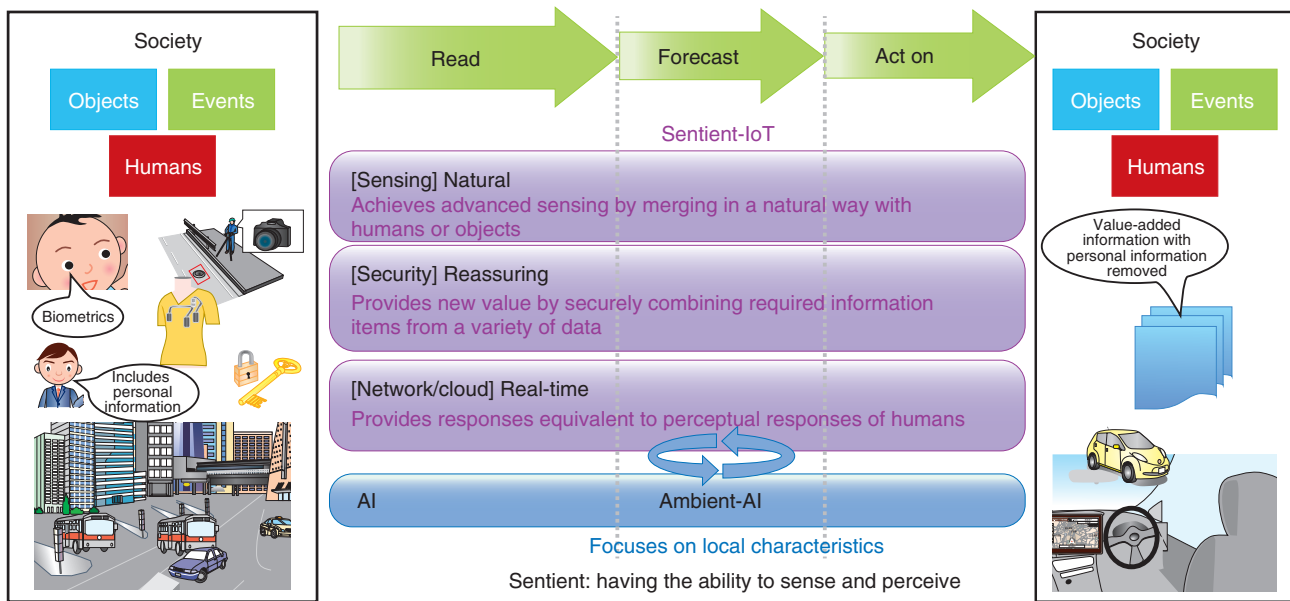


Fig. 6. IoT targeted by NTT R&amp;D.

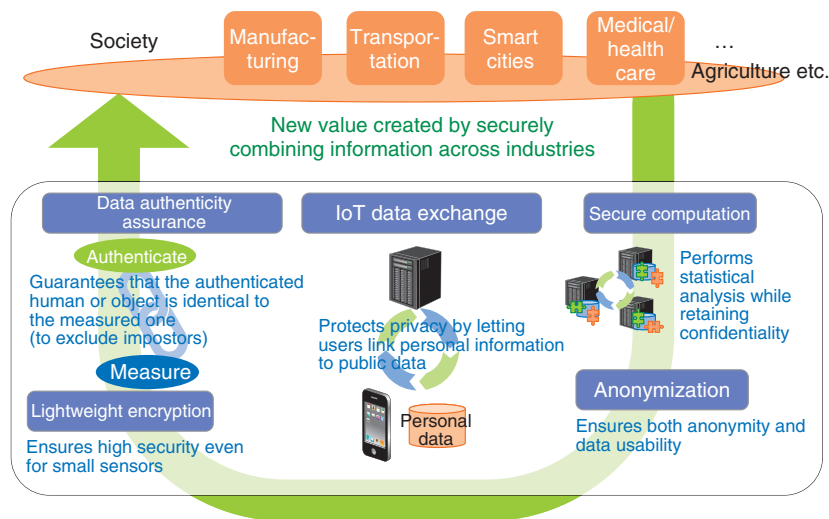


Fig. 7. IoT security: Reassuring.

express this requirement as *reassuring* (Fig. 7). Since sensors used in the IoT world do not have high computing power, data must be encrypted, and the authenticity of data must be guaranteed under severe functional constraints. Diverse types of data may be gathered from IoT devices across many industries, so in addition to our secure computation and data anonymization technologies, we are developing a technology that attempts to protect privacy by leaving it up to

users to link personal data to public data.

In agricultural monitoring, a small amount of IoT data can be processed without constraints on data processing time. In contrast, if IoT data are to be used to operate a machine in a factory or to control a vehicle, the delay time for data processing must be minimized to prevent accidents. We express this requirement as *real time* (Fig. 8). We are studying coordinated scheduling technology, which minimizes

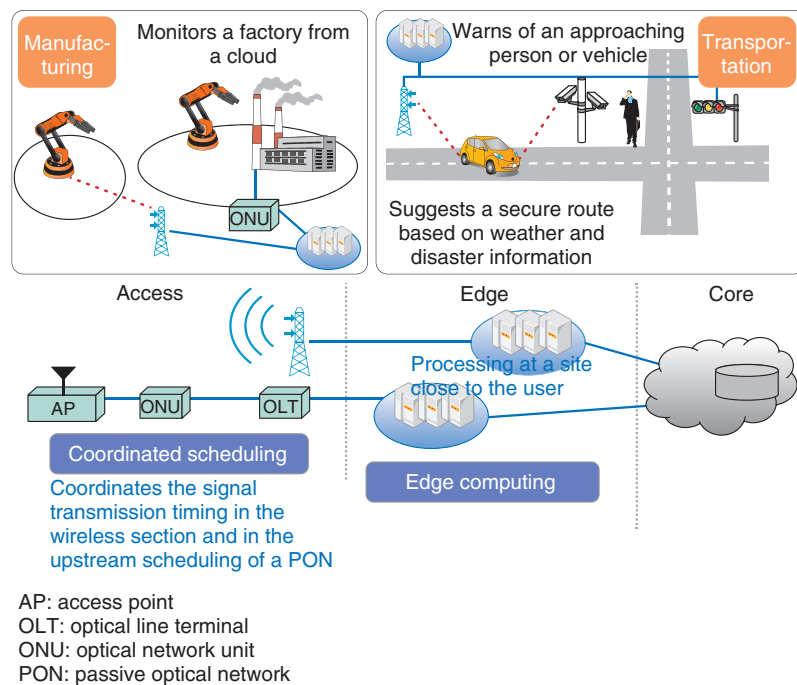


Fig. 8. IoT network/cloud: Real-time.

delay in access sections, in addition to edge computing, which processes data at points close to control terminals.

Like the Internet, IoT evokes a global perspective—being able to communicate worldwide. However, it is also important to build a network that focuses on a local perspective. For example, when controlling traffic lights in Tokyo, it is not necessary to consider the traffic situation in Osaka.

NTT R&D is aiming to develop an IoT that collaborates with the above-mentioned Ambient-AI technologies to create new value. We call this *sentient IoT*.

We believe that it will be important in coming years to have an environment that allows different types of services to be created flexibly by combining a variety of IoT devices, including robots. To build such an environment, it is necessary to have a general-purpose interface that allows these devices to be easily connected with the engines for tasks such as big data processing, image recognition, and speech recognition, and to have a programming language that can be used to simply define those connections. For this purpose, we have developed a technology called R-env<sup>TM</sup> that provides a mechanism enabling various parties to enter the world of IoT without difficulty (Fig. 9). For example, NTT DATA has collaborated with Resona

Bank, Ltd. to combine R-env with a robot and sensors so that the robot can interact with customers at the bank's reception desk. NTT DOCOMO has combined R-env with a power wheelchair from WHILL that provides greater maneuverability and encourages the user to move about. Daiichikosho Co., Ltd. has developed a robot that works with a karaoke machine to provide preventive care for seniors. We are also holding a hackathon to forge a community for R-env.

### 3.1 IoT utilization cases

NTT R&D's IoT technology can be used to detect abnormalities in equipment or to forecast an impending fault (Fig. 10). If sensor data are simply fed into Jubatus, which is a real-time, large-scale, distributed data analysis platform, Jubatus may make a false detection. Instead, sensor data are appropriately cleansed and preprocessed. As a result, Jubatus can detect abnormal behavior in real time and thus prevent faults from occurring. In a joint research project with the Japan Agency for Marine-Earth Science and Technology (JAMSTEC), we have begun using edge computing for advanced weather forecasting. The wide-area simulation on JAMSTEC's supercomputer is combined with area-specific processing on NTT R&D's edge server in order to enhance the level of forecasting accuracy.

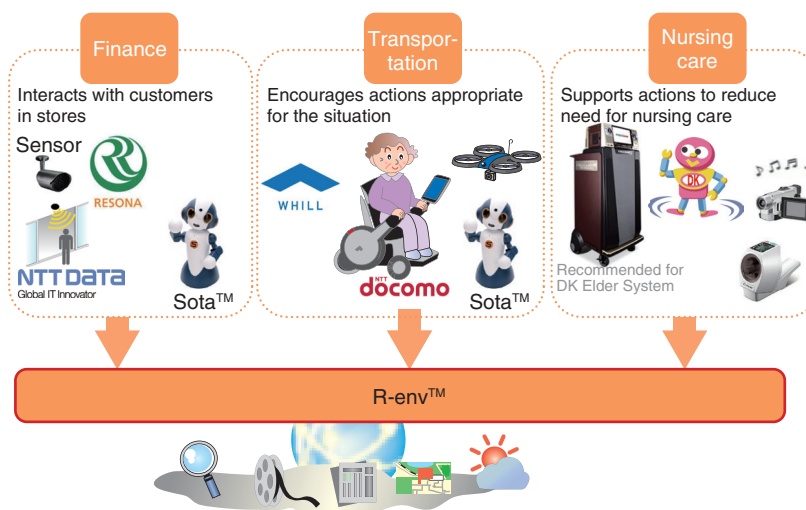


Fig. 9. R-env™: Interaction control.

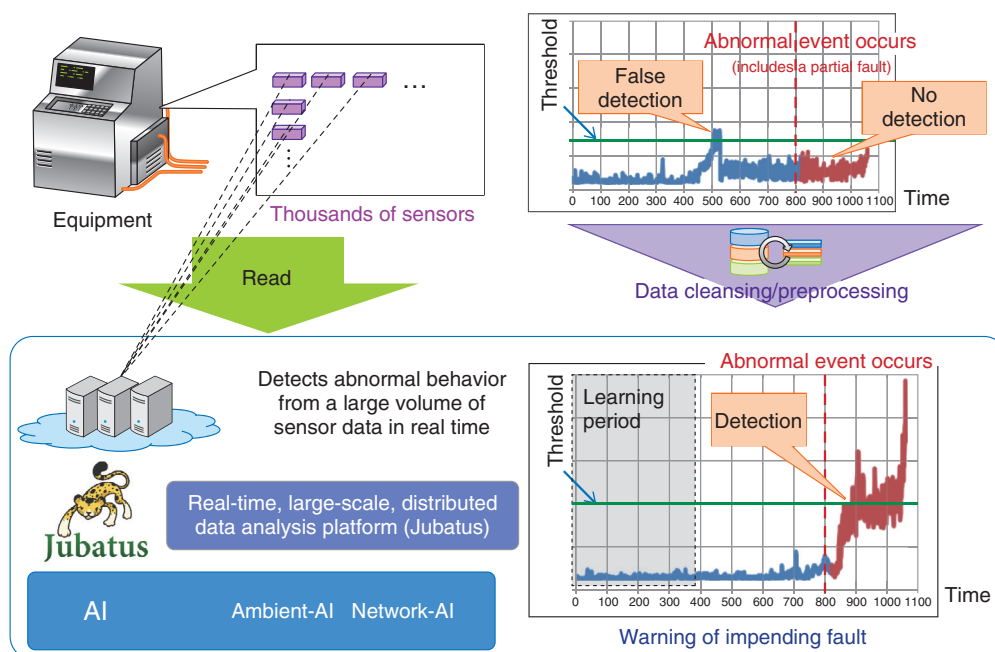


Fig. 10. Detection of abnormality in a machine.

### 3.2 Security technology

In promoting AI and IoT, it is important to continue R&D of security technology. DDoS (distributed denial of service) attacks are growing in scale and sophistication in Japan. We urgently need to be able to implement prompt and effective measures against such attacks. NTT R&D is working on security

orchestration that will enable us to respond to attacks from a network-wide perspective. The security orchestrator takes the specific network condition into consideration. It blocks attacks at the optimal point, recovers service through automatic control of routing and routing policy, or dynamically controls traffic in a manner that is appropriate for the attack.



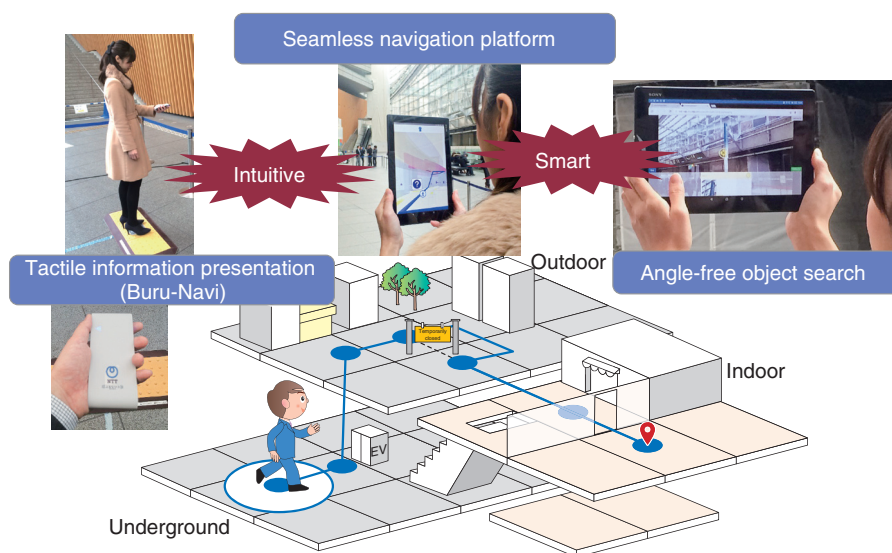


Fig. 11. Trial near Tokyo Station.

As we approach the year 2020, it is necessary for not just NTT but all major infrastructure companies to share information about and develop human resources for security. No matter how skilled it may be, a single company cannot ensure strong cybersecurity. All concerned parties need to work together.

#### 4. Activities looking towards 2020

By 2020, NTT R&D should have fulfilled three major missions: providing high-quality and stable network services, implementing reliable network security measures, and providing hospitality that leaves a deep and positive impression. Our activities related to the third mission are described below. In last year's NTT R&D Forum, I introduced the technologies that NTT R&D had conceived for the third mission as a proof of concept. As the next step, we have initiated some pilot trials.

A pilot trial being conducted in the international terminal of Haneda Airport is aimed at enabling visitors to Japan who cannot read Japanese or who are handicapped in some way to move around without barriers. One service already provided at Haneda for visually impaired persons is an intelligent sound sign service, which provides information using speech technology. The terminal building is full of noise from various sources. The conventional measure for coping with such noise is to increase the speech volume. Instead, NTT slightly modifies the frequency

spectrum of the voice so that the sound will carry well. Neither the volume nor the type of voice is changed. The user can hear sound signs clearly, even amid a cacophony of background noise. Our angle-free object search technology enables visitors to understand information on a direction board by simply pointing their smartphones toward the board; the information is translated and displayed on their phone screens.

A pilot trial is also being carried out near Tokyo Station (Fig. 11). The station is very complex, with a number of passageways and floors, including several underground floors. Currently available maps of the station are not efficiently linked to one another. The main aim of the service developed by NTT R&D is to connect the maps seamlessly so that people can move about intuitively and smartly.

A pilot trial being operated at a stadium uses a technology that integrates a video generated using immersive telepresence (Kirari!), a real video, and computer graphics. This technology has been developed for use in sports activities. For example, in a public viewing, the life-size hologram of a player at a remote stadium is projected on a super-wide screen in synchronization with a video that shows the background. Or, a user wearing a head-mounted display stands in a batting box and experiences the sensation of a ball thrown by a pitcher coming toward him or her.

In the 2020 Showcase at R&D Forum 2016, a screen in the exhibition hall displayed a three-dimensional

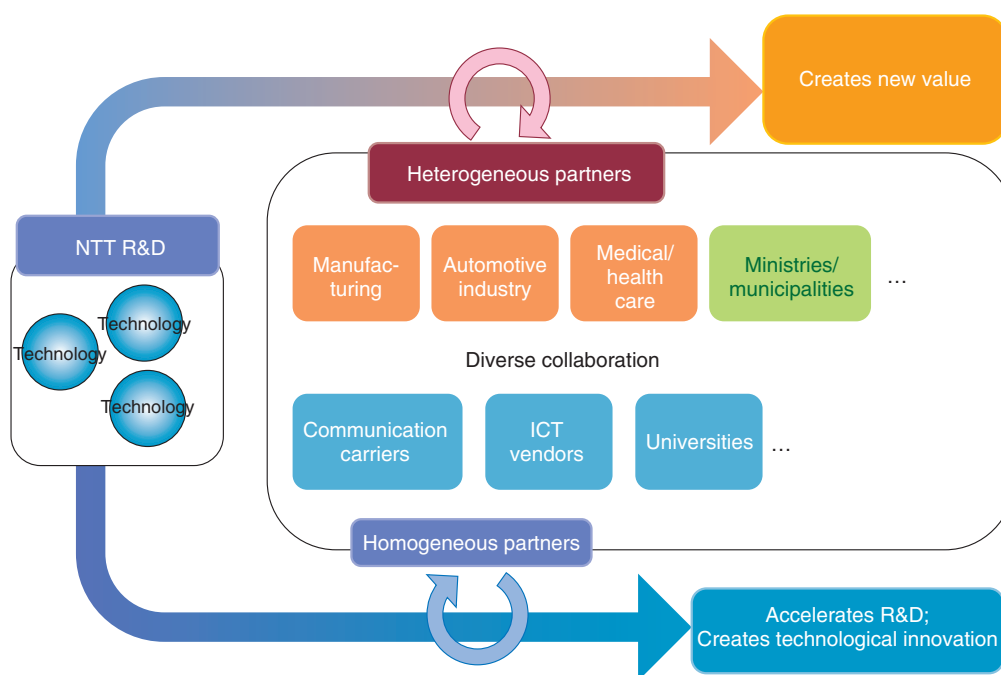


Fig. 12. Objectives of R&D collaborations.

video of the lecture being given in the main hall. The video, generated by Kirari!, provided the sensation of the speaker talking in the exhibition hall. The conventional way of achieving this has been to extract a segment from a recorded video and project it, but on this occasion the video was displayed in close to real time. In addition, angle-free object search technology was used to enable any user who pointed his or her smartphone at an exhibit panel to obtain detailed information displayed on the screen.

## 5. R&D collaborations

We have been collaborating with several partners on various projects. Our collaborations over the past year are described below.

In the area of networking, last year we announced NetroSphere, an initiative that indicates the direction of networks by NTT R&D. The key concepts of NetroSphere are three types of separation: separation between optical and electrical, separation between functionality and hardware, and separation between functions. To be able to implement these separations and combine the disparate parts to build a scalable and flexible network, we need to cooperate with multiple ICT vendors and providers. Over the last 12 months, we have identified a number of parties that

support these concepts and have initiated joint R&D programs with them.

In the area of security, we are developing a common, sharable interface for the security orchestrator so that various combinations of security appliances and switches can be controlled in a timely manner. We will collaborate with other telecommunication providers to address cyber-attacks because it is not sufficient for us to be concerned solely with the NTT Group's networks.

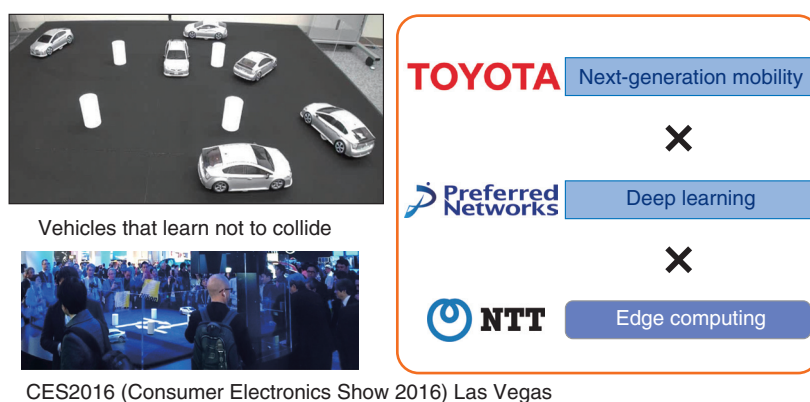
There can be two types of partners for R&D collaboration: homogeneous and heterogeneous (**Fig. 12**). In the areas of networking and security, collaborations with relatively homogeneous partners are expected to produce massive technical innovation and facilitate rapid development of new products. Our experience with collaborations with heterogeneous partners in the last couple of years has convinced me that this approach is conducive to generation of hitherto undreamed of new value. Some of our collaborations with heterogeneous partners are as follows:

### (1) Panasonic

We are collaborating with Panasonic Corporation to enable users with only very simple terminals to receive services equivalent to those that are normally available only to smartphone users (**Fig. 13**). Panasonic



Fig. 13. Panasonic × NTT.



CES2016 (Consumer Electronics Show 2016) Las Vegas

Fig. 14. Toyota × PFN × NTT.

has developed a prototype of a transparent see-through device with minimal functions. By providing it with additional capabilities via a cloud using NTT R&D's device virtualization function, we will be able to create new service possibilities.

## (2) Mitsubishi Heavy Industries

Our collaboration with Mitsubishi Heavy Industries, Ltd. (MHI) in the environmental area has led to the world's first successful laboratory test of online measurement of gas concentrations. By combining MHI's gas analysis technology with NTT R&D's laser technology for optical communication, it has become possible to measure gas concentrations within several minutes, as against the one to two days required for conventional chemical analysis based on sampling.

## (3) Toyota and PFN

We are seeking to implement the concept of vehicles that learn not to collide. A prototype was exhibited at CES2016 (Consumer Electronics Show 2016), which was held in Las Vegas in January 2016 (**Fig. 14**). By combining Toyota's next-generation mobility technology, PFN's deep learning technology, and NTT R&D's edge computing technology, it has become possible for vehicles to share information about their statuses and to learn how to avoid collisions so that they can run autonomously without colliding with other vehicles.

## (4) Dwango

We provided visitors to Dwango's Game Party Japan 2016, held in January 2016, with a smartphone application that displayed the current congestion state

and also forecast future congestion. The application assisted some 5000 visitors in moving around the exhibition hall. In addition, we trialed live video delivery of the event using an HEVC (High Efficiency Video Coding) encoder for a *niconico* live broadcast. A high-definition video was broadcast using half the amount of data that would be required with the conventional method.

(5) Daiichikosho

We are conducting feasibility tests with Daiichikosho Co., Ltd. on the use of karaoke not just for entertainment but also for providing preventive care to elderly persons. One feasibility test uses NTT R&D's noise removal technology to clearly pick up a speaker's voice, even in the middle of a loud karaoke performance, so that the user can search for a song by voice input. Another feasibility study is being conducted at a nursing care center. R-env, mentioned earlier, is used to interconnect a robot, a karaoke machine, and biometric sensors. The robot talks with the elderly persons so that they can better enjoy karaoke sessions.

(6) Shochiku

We are collaborating with Shochiku Co., Ltd. in the use of ICT to create a new form of Kabuki, which will be performed during the Japan KABUKI Festival in Las Vegas, in May 2016.

(7) Various uses of *hitoe*

In the area of materials technology, we are conducting feasibility tests on *hitoe*, which is capable of measuring two biometric signals: cardiac electrical activity and heart rate. In collaboration with Obayashi Corporation, Japan Airlines Co., Ltd., and NTT Communications, we are testing how wearing a shirt

incorporating *hitoe* can improve the work safety of construction workers and airport personnel in the field. To expand its application area, *hitoe* is now being given the ability to measure myoelectrical (muscle) activity. We are testing measurement of biometrics using a suit equipped with an anti-G-device in collaboration with the Acquisition, Technology and Logistics Agency of the Ministry of Defense, and we are measuring the physical stress on drivers in the IndyCar Series in collaboration with NTT DATA. At the Casio World Open golf tournament, we measured a player's muscle activity during a swing. We are hoping that this technology will be able to support athletes as they prepare for the 2020 Olympic Games. Together with NTT DOCOMO, we are providing a software development kit that facilitates use of biometric information obtained by *hitoe* so that other parties will be able to develop applications that utilize biometrics.

## 6. Conclusion

NTT R&D will continue to expand the scope of its collaboration with various partners. Meanwhile, we will endeavor to pursue groundbreaking R&D and enhance our technical capabilities so that we will continue to be selected as a value partner. Concerning the question of whether we should seek exclusive or non-exclusive collaboration, we believe that non-exclusive is the way to go because, in the world of AI, a single robot becoming smart will not change the world and because, in the world of IoT, the degree of added value created by a single industry through gathering information is not high enough.