

Making the World Smart and Technology Natural

Katsuhiko Kawazoe

Senior Vice President, Head of Research and Development Planning, NTT

Overview

This article introduces NTT's latest research and development activities based on a lecture presented by Katsuhiko Kawazoe, NTT Senior Vice President, Head of Research and Development Planning, at NTT R&D Forum 2018 Autumn held November 29–30, 2018.

Keywords: smart world, highly realistic communication, artificial intelligence



1. Transformation of the world with B2B2X

Until the present time, the NTT laboratories have engaged in research and development (R&D) mainly to support NTT's services and systems. NTT has shifted its business strategy to the business-to-business-to-X (B2B2X) model, thus placing greater emphasis on R&D for value creation with partners. We have been able to develop groundbreaking co-innovations in collaboration with numerous parties.

For example, in 2014, Mitsubishi Heavy Industries, Ltd. and NTT agreed to work together on R&D for *social infrastructure × ICT*, which refers to the integration of social infrastructure and information and communication technology (ICT). The two companies are engaged in different business fields, but one technology linked them, and this connection led to a discovery that transformed the manufacturing industry. From the wide-ranging research results produced by NTT, Mitsubishi Heavy Industries set its eye on one technology, one that NTT had never imagined would catch their attention. They took an interest in optical fiber technology.

The photonic crystal optical fiber, which was an NTT first, was characterized by having air holes inside the fiber, and light propagated through these

holes while being confined in them (**Fig. 1**). By changing the hole diameter and the spacing between the holes, we can finely control the refraction index of the light in the fiber and achieve optical transmission characterized by extremely high output and unparalleled quality.

Application of this optical fiber technology for communication to laser beam machining is the key objective of the cooperation between the two companies. However, because the light energy level required for laser beam machining was more than 10,000 times higher than that required for communications, we faced a number of new challenges. We solved these problems, and optical fiber technology for communications found a new application in laser beam machines for cutting and welding.

It had not previously been possible to transmit a high-power single-mode laser beam for machining over more than a few meters. The application of NTT's technology made it possible to extend the distance by several dozen times. The people at Mitsubishi Heavy Industries discovered new value in our technology that those at NTT could never have envisaged on their own. This is a prime example of co-innovation through B2B2X. The case involving Mitsubishi Heavy Industries originated from our attempt to

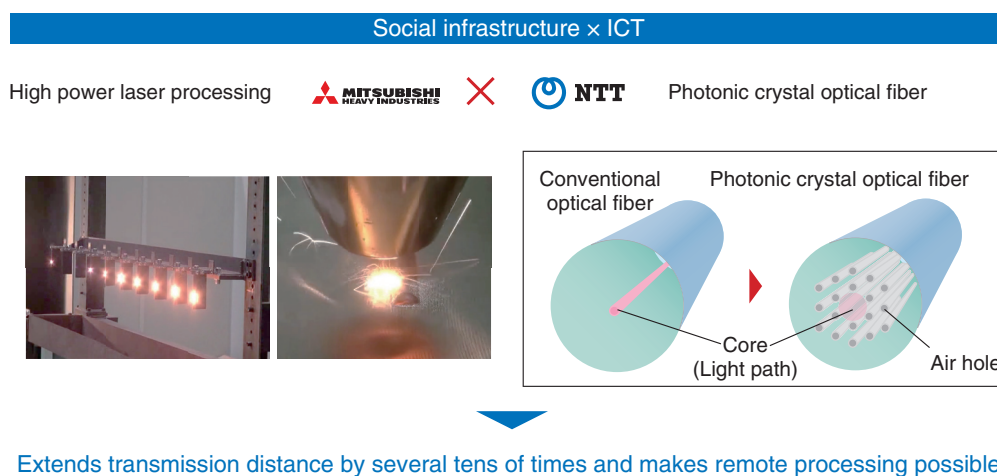


Fig. 1. Technology transforming the manufacturing industry.

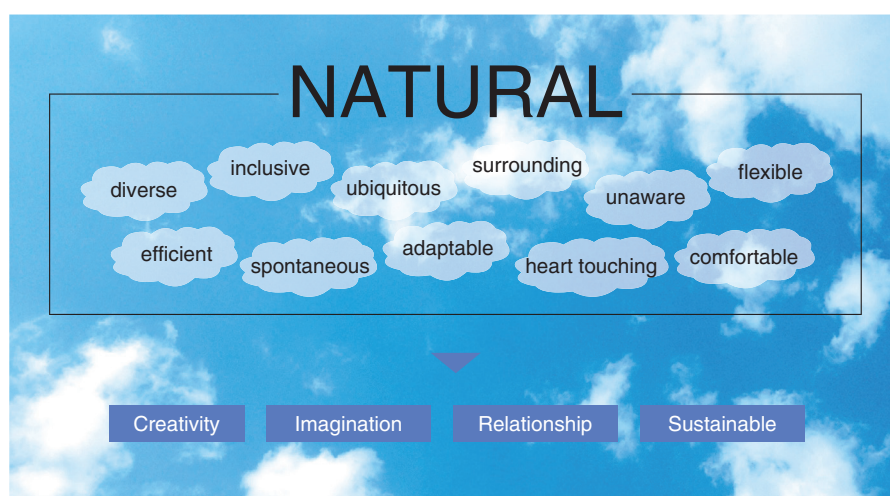


Fig. 2. The concept of *natural* in digital transformation.

solve a problem besetting their business. In the end, we have produced a result that will dramatically transform the manufacturing industry. Looking back, we realize that this constituted a transformation of the world.

2. Keyword for making smart world a reality: *natural*

NTT has innovative technologies that originated in its R&D—technologies that have been the best in the world, the first in the world, and that have amazed the world. With these technologies and through co-inno-

vation with a range of partners, we will push digital transformation in society and industry to solve the problems they face, and thereby make a smart world a reality. To do this, we will strengthen our involvement in a wide range of technical fields.

How should we advance these technologies and push digital transformation so that everyone regardless of nationality, age, or background can benefit from technology? We believe the keyword for this endeavor is *natural* (Fig. 2). We are aiming at the creation of a world in which technology keeps *caring eyes* on the lives of people from all walks of life without them being aware of it; a world in which technology

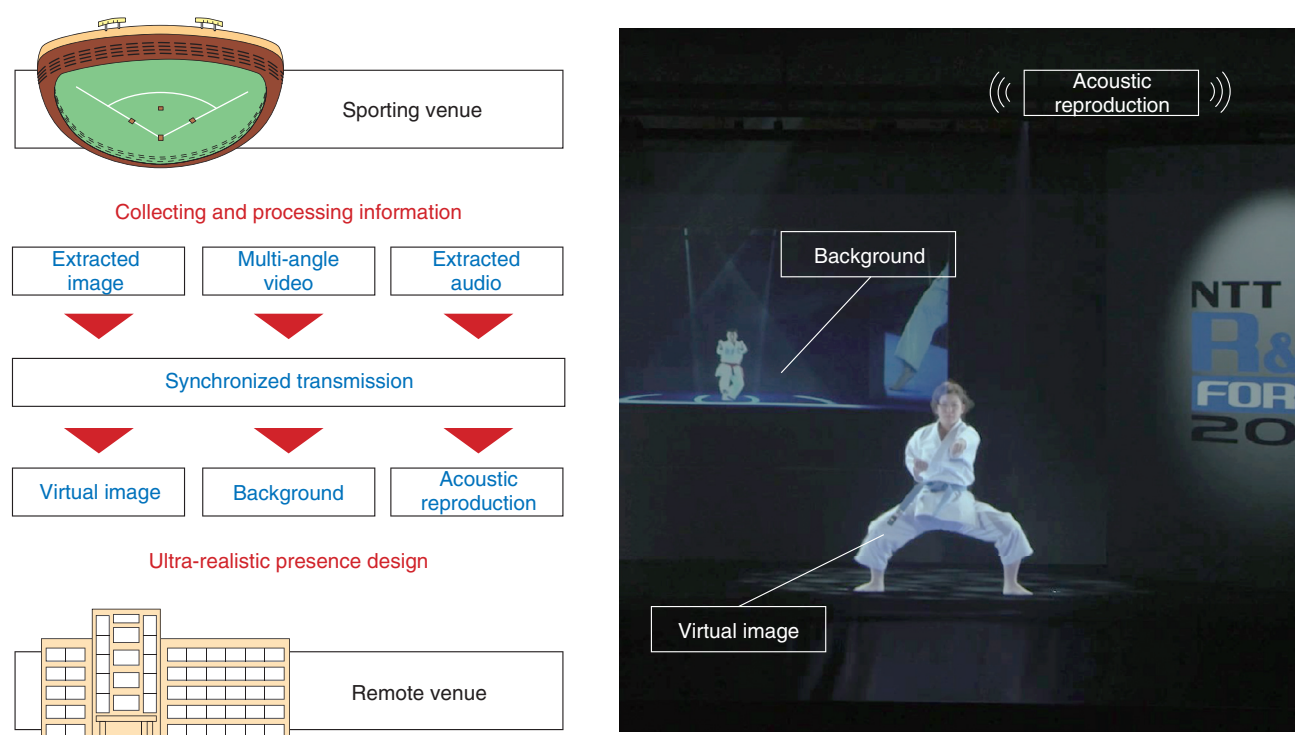


Fig. 3. Ultra-realistic communication that conveys excitement.

sometimes helps people do things more efficiently and appeals to their feelings; a world in which technology provides an environment that is friendly to both people and the earth. This initiative will lead to a world in which people can engage in human-centered activities. Let me introduce some of our activities that are aimed at making such a world come true.

3. Ultra-realistic communication that conveys excitement

The long-awaited regular broadcasting of 4K and 8K television (TV) started on December 1, 2018, in Japan. Its high definition of video images will offer a higher than ever sense of excitement. As the next step, we want to realize a world in which people transcend time and space and experience excitement as if they were at the site where an event was unfolding. This is the ultra-realistic communication that NTT is aiming at (Fig. 3).

For example, in the fashion show Tokyo Girls Collection 2018, which was held at the Yokohama Arena on March 31, 2018, scenes from the show were transmitted to a remote live viewing site in real time, enabling a large number of people to view the scenes

happening at the event site. Kirari!'s ultra-wide video synthesis technology stitched together a number of 4K videos in a natural manner in real time to produce a video that allowed the audience to enjoy a high sense of reality.

In addition, the videos shown on long vertical screens, three-dimensional (3D) and virtual reality videos, and multi-angle videos displayed on tablets were synchronized precisely using Kirari!'s advanced media streaming and synchronization technology called Advanced MMT*. Kirari! is an ultra-realistic communication technology that enables people at any location to enjoy the sensation of being at an event venue. It brings to remote sites the excitement and sensation of tension that usually only those at an event venue can enjoy. The ability to share in the excitement of a game or event is what we aim to achieve with Kirari!.

Kirari! has been continuously evolving since the

* Advanced MMT: Extended protocol of the MPEG Media Transport (MMT) undergoing standardization in ITU-T (International Telecommunication Union - Telecommunication Standardization Sector) Study Group 16 Immersive Live Experience. The MMT is an optimized protocol for synchronous data transmission developed by the Moving Picture Experts Group (MPEG).

concept behind it was announced in 2015. Initially, viewers' attention is captured by its quasi-3D image display, but what is more important is that Kirari! creates a natural and highly realistic space by decomposing a scene into elements such as the objects' video and audio streams, transmitting them separately, and recomposing them in a way that is best suited to the conditions at the viewing site.

I was once deeply moved by the performance of a figure skater in a top-level competition. Her entry into this event was to be the culmination of her long efforts, but she lagged well behind in the short program. Despite that, her performance in the free program the next day was outstanding, gaining her sixth place in the end. I believe that many people around the world were also moved when they witnessed her complete her performance, look up, and start to cry. Why were we so moved at that time? I believe that we were impressed not just by her performance but also because the images of the path she had followed from childhood, her aspirations for that contest, and her errors in the short program ran through our minds like fragments of a story and were multiplied by the video.

Kirari! will evolve further with the aim of not only transmitting the video and audio faithfully but also evoking deep emotions in the hearts and minds of viewers based on a relevant story, such as past experiences and knowledge. Ultra-realistic communication that conveys deeply felt emotions. That is the *naturalness* that we aim at.

4. Artificial intelligence (AI) that guesses what you think

Technologies for listening and speaking AIs are now being incorporated into home appliances and are commonly used in our everyday lives. If we are to convey human feelings in a natural manner without causing the person involved to be conscious of interacting with a machine, it is essential to make further advances in speech dialogue technology. For example, if AI can handle unstructured chatting, which makes up a good part of our conversation, it can communicate with people naturally. Therefore, research on chat technology is underway at the NTT laboratories.

We are developing an AI system that can converse more naturally by giving it a distinct character and, consequently, friendliness. In order to initiate and sustain a pleasant conversation, we have developed a function that responds with not just words but also through gestures such as nodding in agreement, a

verbal function whereby appropriate responses are made that indicate agreement with what you say or think, and a function whereby a follow-up question or a question relating to what you have said is asked. We have incorporated these into Totto, the android of Ms. Tetsuko Kuroyanagi, Japanese actress and TV presenter [1].

Viewing AI that recognizes objects in the surroundings and the current situation will also become more natural. NTT has been researching robust image recognition technologies. They include, for example, robust media search technology that requires only a few reference images in order to determine that an object in question is identical to the object in the reference images, and change point detection technology that instantly locates change points in observational photos taken from a satellite.

In addition, we have developed angle-free rigid and non-rigid object recognition technology that is expected to advance image recognition technology dramatically (Fig. 4). Things that AI needs to recognize are not necessarily limited to rigid objects. This new recognition technology can tell that an object with a shape that can change, such as a product in a bag, is identical to the object in the reference image. This ability is expected to significantly expand the application areas of image recognition AI.

The performance of AIs when listening, speaking, and viewing things is improving. The time will come when it is more than adequate. We believe that at some stage in the future, *thinking AI* that supports human thought processes and leads to co-creation by people and AI will become more important than ever (Fig. 5). The current technology for Totto is still at a stage where the distinct character of Ms. Kuroyanagi has been incorporated into the robot. In the future, AI needs to incorporate her values and personality. The AI we are aiming for at NTT is one that incorporates various values and personality traits and helps people to consider complicated problems for which there is no single answer.

In the future, we will aim for AI that, while conforming to commonly accepted rules and morals, absorbs all of the perceptions of value that may arise out of geographical regions or customs. Such AI could be called *generous AI* or *sincere AI*.

5. Stress-free device

When we think of the next personal communication device that will go beyond the evolution of cell phones and smartphones, we think that what is needed

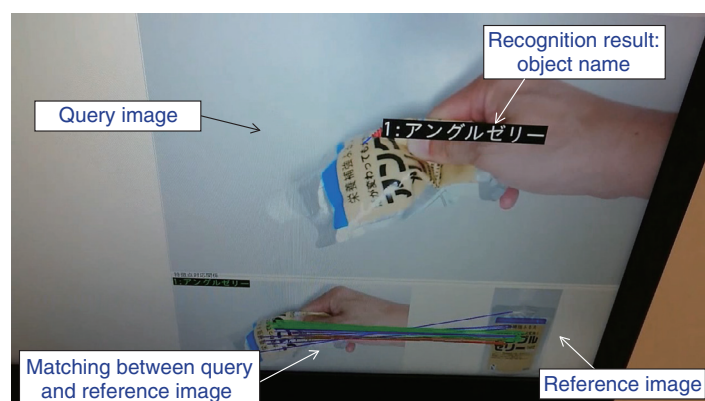


Fig. 4. Angle-free rigid and non-rigid object recognition technology.

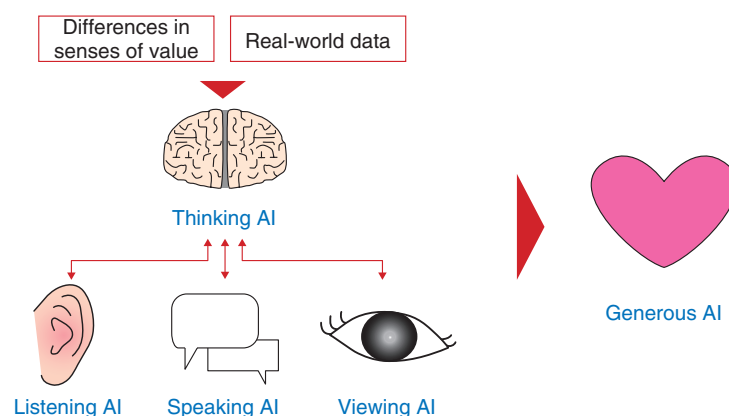


Fig. 5. AI that guesses what you think.

is a terminal or device that frees people from being dominated by applications and settings.

CUzo is a device that can be operated very naturally just by pointing it at something such as a person or a landmark. You can get information about something just by pointing the device at it. For example, CUzo will present information about a tourist spot in the user's own language on its transparent display, provide easy-to-understand navigational advice in an unfamiliar street, or enable people from different countries to have a face-to-face conversation by looking at translations that appear on the display (**Fig. 6**). We want to enable people to have a natural experience by doing away with annoying tasks such as activating and operating applications, and having the device provide services in response to prompts conveyed by an individual's natural actions.

We have developed this system in collaboration

with Panasonic Corporation. It uses device function virtualization technology developed by NTT, which virtualizes the processing functions conventionally performed within a device such as a smartphone, and places them in a cloud or on an edge computer. This technology makes it possible to provide advanced services even with simple devices.

After we have made it possible to use devices without the conscious effort of setting up and operating applications, what will come next? We foresee a future in which we no longer need to be conscious of visible devices, and various things around us look after our lives. For example, several ICT devices in a room work together and provide an illusion that a raincoat hung on the wall appears to be trembling and the floor looks wet, thereby letting the resident know in this natural manner that it is going to rain today (**Fig. 7**). This is one of the manifestations of naturalness.

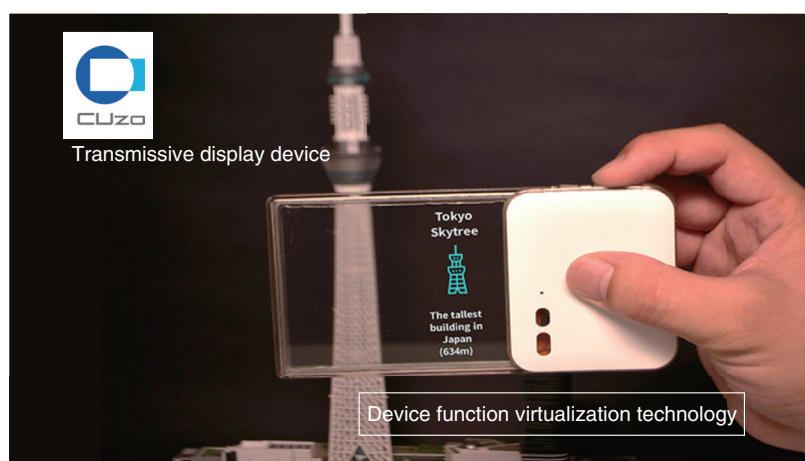


Fig. 6. Stress-free device.



Fig. 7. A world where we no longer need to be conscious of visible devices.

To pursue this concept, we plan to launch a new project called Point of Atmosphere. Please follow what we do in this field in the coming years.

6. Future network that blends seamlessly into society

Networks support various services. They will also evolve to become more natural.

A next-generation network will integrate various functions and roles across different layers. It will understand social order and priority and operate in a natural manner without humans being required to make conscious efforts to select the optimal choice. A mechanism we are working on for this purpose is the Cognitive Foundation®. It manages and operates various resources in different layers in an integrated manner. In the public safety solution in the City of Las Vegas, the Cognitive Foundation manages various ICT resources such as cameras, sensors, edge computers, networks, and clouds, as necessary in an optimal and integrated manner in order to flexibly adapt video monitoring resources to what is needed at a particular time.

At the current stage, the Cognitive Foundation only performs basic operations automatically, but our aim is that as the network AI continues to learn, it will execute globally optimal control in real time in a constantly changing environment and achieve advanced coordinated operations. The Cognitive Foundation will expand its coverage area to people, towns, transport and energy, work with enterprises involved in various layers of society and applications, and

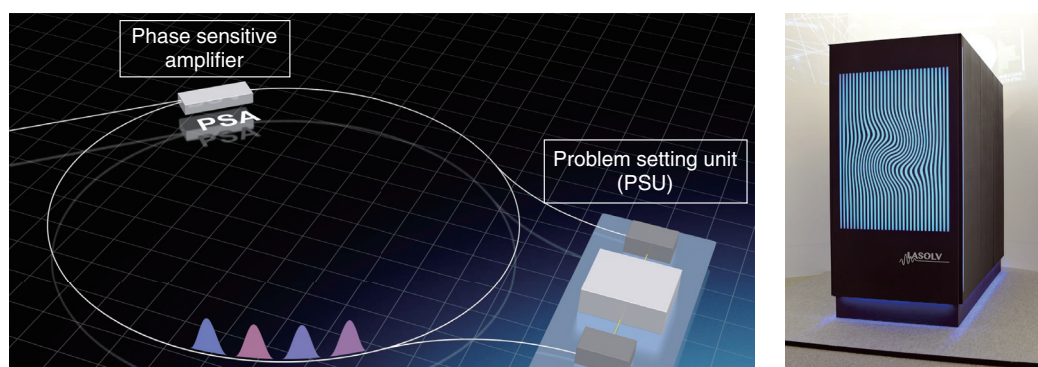


Fig. 8. Computer that solves challenging problems with light.

orchestrate everything from services to devices, thereby achieving optimization of the entire society in a natural manner. For this purpose, we will work on scalable data processing infrastructure and super-secure authentication infrastructure technologies.

7. Computer that solves challenging problems with light

If we are to enable people to lead a better life in a natural manner, we must boost the power of computers. To overcome the limitations of conventional von Neumann computers, the NTT laboratories are developing LASOLV, a computer that solves challenging problems with light (Fig. 8). Conventional digital computers solve a problem by turning it into a mathematical problem. LASOLV solves a problem not as a mathematical problem but as a physical problem. This has made it possible to solve hard problems that could not be cracked by conventional digital computers.

LASOLV is implemented using leading-edge optical communication devices and optical parametric oscillator pulses generated by phase sensitive amplifiers, both of which have been developed by the NTT laboratories over many years. A series of optical pulses flow in a 1-km-long optical fiber ring. LASOLV expresses inter-pulse interactions using an optical parametric oscillator and a problem-setting unit. This has made it possible to solve the problems associated with optimizing large-scale combinatorial models.

To date, LASOLV has been able to solve relatively simple grouping problems. It has now become possible for LASOLV to expand its targets to a wider variety of problems such as the Japanese map coloring problem and scheduling problems. For example,

assume that the possibility for a town consisting of residential, commercial, industrial, and green sections to grow to become a better town is predicated on the condition that sections belonging to the same category should not exist side by side. Let us consider a simulation of town development that meets this condition. The beauty of LASOLV is that it can solve a problem incredibly quickly that conventional digital computers cannot solve in a short time.

We are now considerably expanding the libraries of LASOLV software and advancing its software development environment so that not only experts but also general programmers can make use of it.

In addition, we will increase the number of bits in hardware from 2048 to 100,000 so that LASOLV can be applied to a wider variety of fields. One example is drug development. At the basic research stage of searching for chemical compounds that can be used therapeutically, LASOLV has the capacity to quickly identify good combinations of chemical compounds so that new drugs can be developed in a short time. To solve traffic jams and city planning, LASOLV will suggest a better route by calculating a recommended route from one point to another or the most efficient route when it is necessary to visit several places. In the field of AI, LASOLV will have the ability to find from a data set to be used in machine learning, a sample that is the closest to the data being sought. This will lead to more robust learning.

8. Activities to accelerate innovation

I have so far introduced our R&D results and the direction we are taking. Let me reveal the details of our new measures designed to accelerate further innovation. The underlying policy is globalization of

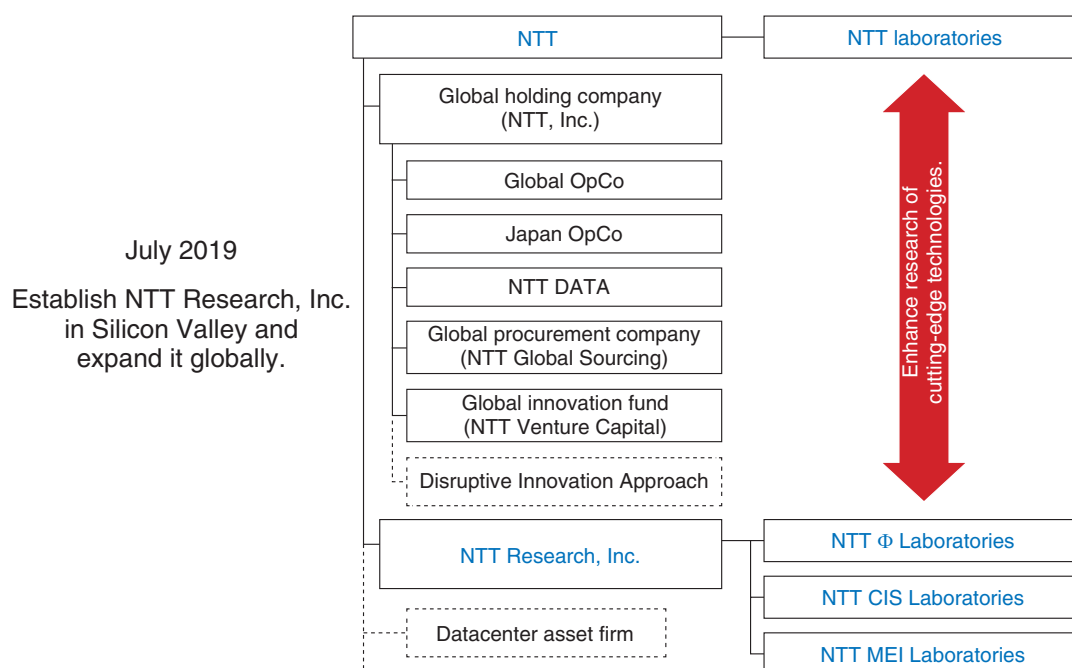


Fig. 9. Establishment of research organizations outside Japan.

our R&D. We are taking the following three measures to achieve this:

The first is global utilization of R&D results. We will take the R&D results of the NTT laboratories outside Japan and deploy them in ways adapted to individual regions.

The second is globalization of research targets. We will strengthen R&D that is adapted to global needs.

The third is establishment of research organizations outside Japan (**Fig. 9**). There is apprehension today that basic science and technology, the very source of innovation, are declining in Japan. We believe that basic research that serves as the foundation supporting the activities I have outlined above is the key that will allow us to take the next leap forward. To expand and strengthen basic research, we will establish a basic research organization overseas. It will be called NTT Research, Inc.

Three laboratories will be established within NTT Research, Inc. The first is NTT Φ Laboratories. This facility's mission is to make new discoveries and develop new technologies that will dramatically transform the world in the area of quantum science and computing, which is a co-creation area encompassing both physics and informatics. We came up with the idea of using PHI or Φ from the Greek alphabet—as an acronym of physics and informatics—for

the name of the laboratories. The focus at Φ Labs will be on basic research in the field of quantum theory, which will lead to future quantum computing technology and the formulation of completely new theories, including the application of quantum theory to information processing. We will invite Professor Yoshihisa Yamamoto, professor emeritus of the National Institute of Informatics and of Stanford University and the project manager of ImpACT (Impulsing Paradigm Change through Disruptive Technologies Program), which has collaborated with us in the development of LASOLV, to become the director of Φ Labs.

The second is NTT CIS Laboratories. This organization will work on advanced cryptographic theory and encryption, and information theory, which is the basic theory underlying secure information exchange in a complicated distributed environment. The director of CIS Labs will be Dr. Tatsuaki Okamoto, NTT Fellow and last year's recipient of the RSA Conference Award, which is one of the most prestigious awards in the area of cryptography.

The third is NTT MEI Laboratories. They will be engaged in medical and health information processing, which will assist the establishment of a natural relationship between people and ICT. To head MEI Labs, we will invite Dr. Hitonobu Tomoike, medical doctor

and advisor to the Sakakibara Heart Institute. He has considerable influence both in Japan and abroad as a physician expert in the cardiocirculatory system and is also well-versed in ICT.

The key aspect of these appointments is that all three individuals are active globally in their respective fields of expertise and thus have wide-ranging international human networks.

We will first establish NTT Research, Inc. in Silicon Valley, California, where research in these fields is advancing, and then extend its geographical presence to other parts of the world. Through R&D at global sites focusing on basic research, and in collaboration with universities, both at home and abroad, and with other partners, we will aim for the summit of basic research and produce unparalleled research results that will bring about game-changing innovations.

9. Future outlook

The NTT laboratories have shown the potential of new technologies well in advance of their rivals. We will aim to achieve impactful technical innovation, keep abreast of the changing times, and open up the next frontier.

NTT R&D will look ahead and paint a picture of what lies beyond the near future. With a view to bringing about a future characterized by naturalness and to realizing a smart world, we will continue to pursue research that will transform the world.

Reference

[1] Totto's room (in Japanese), <http://totto-android.com/>

Trademark notes

All brand names, product names, and company/organization names that appear in this article are trademarks or registered trademarks of their respective owners.