Social Well-being Achieved through Social Infrastructure That Is Friendly to People and the Planet

Sachiko Oonishi

Abstract

With the emergence of generative AI (artificial intelligence), the world is becoming more convenient; however, electricity consumption is increasing, and environmental problems are becoming more serious. By personalizing products and services to meet diversifying needs from a market-in perspective, while using energy-saving technologies and IOWN (Innovative Optical and Wireless Network) access networks as new human and Earth-friendly industrial and social infrastructure, we can achieve social well-being. This article is based on the keynote speech I presented at Tsukuba Forum 2024 held in May 2024.

Keywords: generative AI, IOWN, marketing

1. Delivering desired value: From information and communications infrastructure to industrial and social infrastructure

The global situation and social issues are becoming increasingly serious. Food loss is also a major issue; 2.5 billion tons of food, or 40% of the world's annual food production, is wasted, and 1.5 billion of the 2.9 billion garments produced in Japan each year are discarded. Environmental and energy issues are becoming increasingly severe on a global scale, and Japan's declining birthrate, aging population, and labor shortages are also becoming more serious. In this environment, people's values are diversifying. In the past, people sought an abundance of material goods, but this pursuit has gradually shifted to spiritual values, intangible qualities, and satisfying experiences. Put another way, people's present mindset is to minimize loss and possess only what is necessary rather than an abundance of goods.

In response to such diverse values, NTT has combined research and development with marketing, and in June 2023, we launched the Research and Development Market Strategy Division within the holding company. In addition to the Research and Development Planning Department, which manages the research laboratories, Market Planning and Analysis Department and Alliance Department were established in a manner that unified the three departments into a single headquarters. We want to continue to evolve and develop our research technologies while aligning with the market and considering research and development of technologies on the basis of requirements from a market-in perspective. In addition to research and development, we will pursue customer experience (CX) in all business processes, ranging from service development to operations and service quality. We will also continue to improve our services from the customer's perspective by taking into account how our corporate and individual customers feel. To achieve these goals, we are building a marketing infrastructure that will enable us to digitize and visualize customer feedback and other information and strengthen our system for promoting CX.

Our product-out research and development has evolved from "connect" and "communicate" technologies, starting with telephones, to optical-fiber technology and more recently, IOWN (Innovative Optical and Wireless Network). From a market-in perspective, we look at things from the viewpoint of people, society, and the Earth as if we were imagining them with our senses. Therefore, "product-out" is

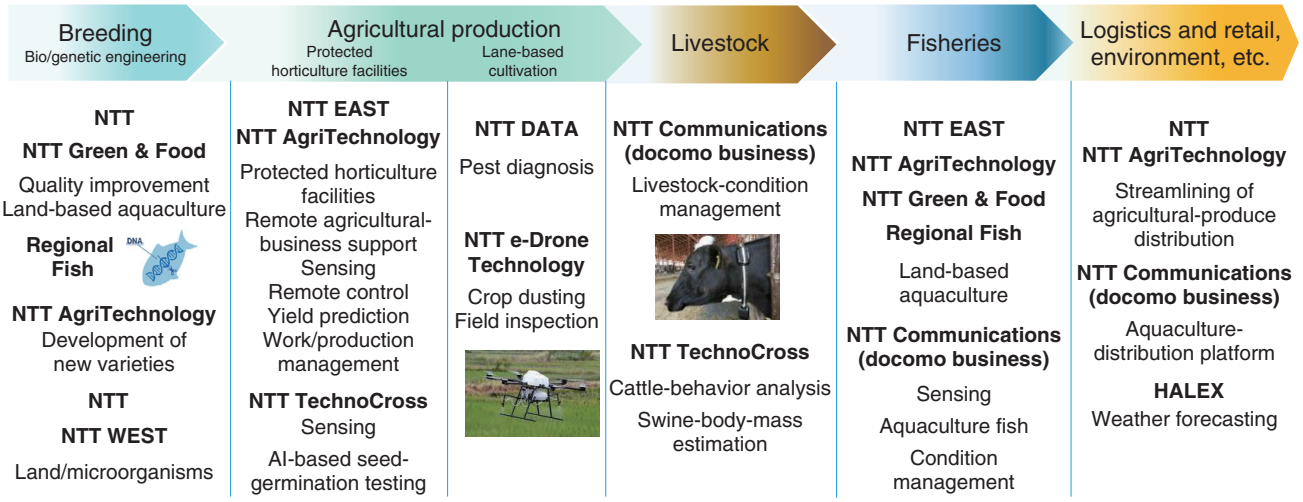


Fig. 1. How innovation is transforming food: from breeding, agricultural production, livestock and fisheries to distribution.

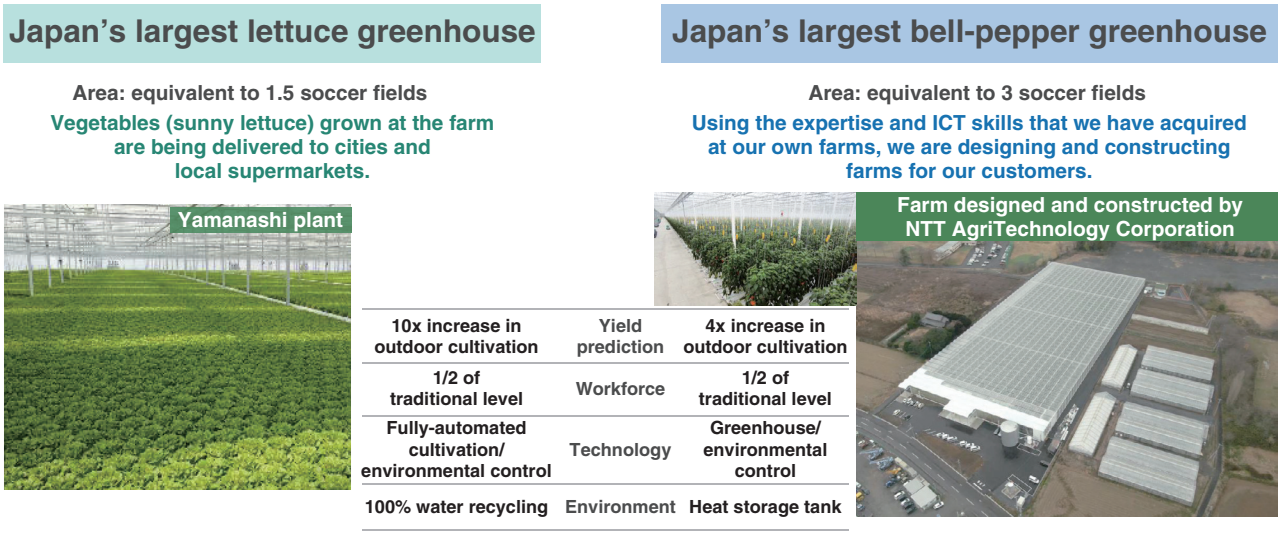
more of a functional approach that focuses on connection, digitalization, data generation, and artificial intelligence (AI) analysis. In contrast, "market-in" is more of a value approach that involves using a variety of technologies to visualize, optimize, improve efficiency, save energy, and personalize. By combining these two approaches, we can respond to diverse values and work towards social well-being. The products that we have been developing through our traditional product-out research and development focused on connecting and communicating have played a role as an information and communications infrastructure for corporate and individual customers; however, from now onwards, we hope our products and services will play a role as an industrial and social infrastructure in relation to people's values, which cover food, clothing, shelter, lifestyle, education, healthcare, energy, and entertainment.

2. Food and healthcare play a role as an industrial and social infrastructure

How can we make food more abundant and sustainable? The livestock industry, fisheries, forestry, and the recycling of waste produced by these industries are all related to food production but have not been directly related to the industry supporting information and telecommunications infrastructure. I will give examples of how we are using our technology for sustainable food production. As we all know, the risks to a stable food supply are severe. In 2024, food costs have increased by about 1.5 times compared with those in 2023, and this trend is starting to affect eating habits. Japan's food self-sufficiency rate is 38%, which is the 12th lowest among the 13 major developed countries [1]. The agricultural workforce in Japan has decreased by 70% compared with that in 2000, and the average age of an agricultural worker has increased to 68 [2]. Thus, the amount of abandoned farmland has increased by 70% compared with that in 1995 [3]. The Netherlands and the UK also faced similar challenges, but through innovation, they now have food self-sufficiency rates of over 60%. To achieve such innovation in Japan, many NTT Group companies are working in various areas such as breeding, agricultural production, livestock, fisheries, and distribution (**Fig. 1**). I will introduce one of these areas, greenhouse horticulture.

NTT AgriTechnology is building one of Japan's largest lettuce greenhouses in Yamanashi prefecture (**Fig. 2**). The area enclosed by the greenhouse is 1.5 times the size of a soccer field. In this greenhouse, natural light is used instead of artificial light, and the spraying of water and nutrients is done automatically. With only half the workforce required, the greenhouse has achieved a 10-times-higher yield than that of conventional greenhouses. As Japan's population of agricultural workers declines and the general population ages, we are making good use of a combination of manual labor and technology. Thus, we were commissioned to design and build a bell-pepper greenhouse (right-side image in Fig. 2), which has also been a great success.

Another food-related area of interest is fisheries.



*Due to the need to prevent pests, requests to conduct tours of the facilities are not currently being accepted.

Fig. 2. Reduced environmental footprint due to 100% water recycling.

Japan's fisheries industry is in a situation in which it is no longer possible to catch as many fish as before. Increased amount of atmospheric carbon dioxide (CO_2) is warming the oceans and reducing the regions of the oceans that can be inhabited by fish. Although ocean acidification has led to a decrease in phytoplankton and zooplankton, which fish feed on, global production of marine products, including aquaculture, is increasing. Japan, however, is the only country where seafood production is declining. Japan's fishing industry was ranked number one in the world in 1980, but it has now fallen to seventh [4, 5]. In light of this situation, we decided to start land-based aquaculture, and in collaboration with NTT and a startup from Kyoto University, called Regional Fish, we established NTT Green & Food to promote land-based aquaculture. Although fish change gradually in the natural world, by scientifically modifying them at an earlier stage and making them meatier, we have succeeded in breeding fish with 60% more edible parts. The algae that fish feed on are modified to absorb as much CO_2 as possible from the ocean, and that CO_2 is thus adsorbed into the bones of the fish that eat the algae. The NTT Group is building land-based aquaculture plants for flounder, white-leg shrimp, and salmon. The first flounder farmed at the plant located in Kyushu, Japan is now available for shipment under the name "Hirameki Hikari."

We are also engaged in various initiatives in the fields of health, healthcare, and medicine. In one such initiative, we are investigating which foods, for example, a banana or cookie, cause a greater rise in blood-glucose level after eating. The glycemic index (GI) determines how much a person's blood-glucose level rises after eating. According to the GI [6, 7], cookies raise blood-glucose level more. However, the level varies from person to person, even if they eat the same food [8]. When we asked 1089 people at risk of developing diabetes to eat a banana and a cookie then measured their blood-glucose level after eating each one, 445 of them experienced an increase in blood-glucose level after eating the banana, but their blood-glucose level did not change after eating the cookie. For the remaining 644 people, their blood-glucose level increased when they ate a cookie but remained unchanged when they ate a banana (**Fig. 3**). As biometric devices in the form of wristwatches and pedometers become available to measure blood-glucose level, people will be able to eat whatever food they like while checking whether it raises their blood-glucose level. NTT's biometric wristwatch device irradiates radio waves onto the skin and measures glucose concentration by analyzing the reflected signal (**Fig. 4**). We will continue investigating this type of personalized support for promoting health.

3. Entry of generative AI

The next area concerning NTT Group companies is generative AI. Hardly a day goes by without someone

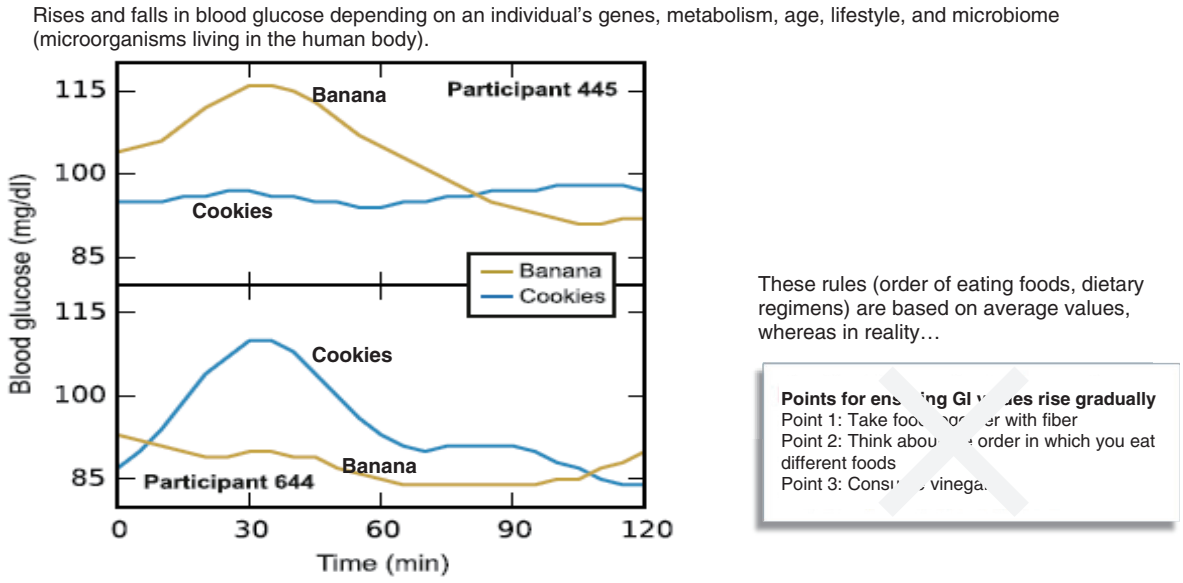


Fig. 3. Changes in blood glucose among individuals after eating the same foods.

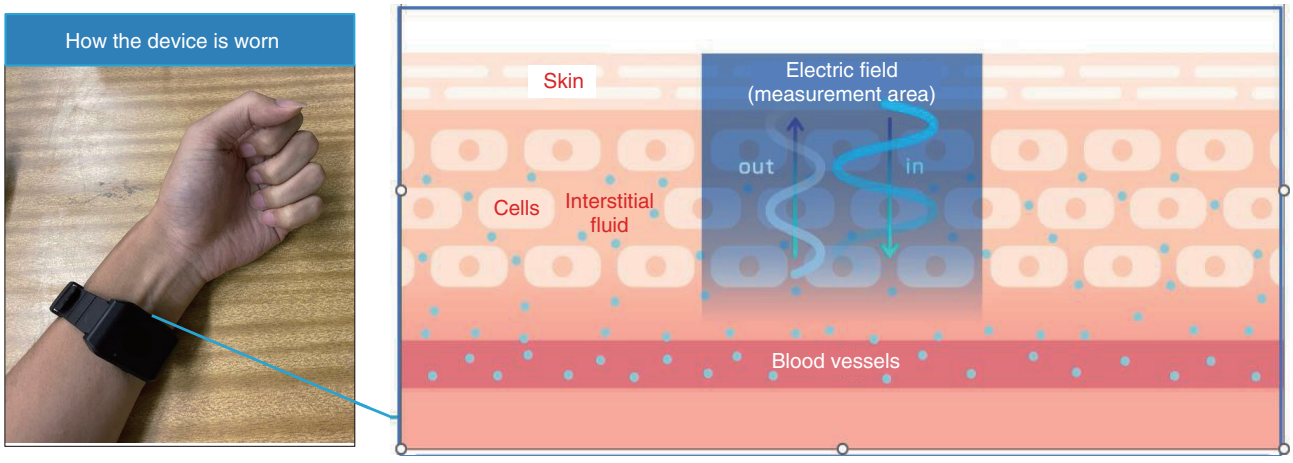


Fig. 4. Analysis of signal of reflected electrical wave passing through the skin.

talking about generative AI [9, 10]. Let's think about the size of generative AI. The size is determined by multiplying two factors: the amount of text data to train generative AI and the size of the neural network that connects the data. As the size increases, the number of graphics processing units (GPUs) used for training the AI increases, and more GPUs means increased electricity consumption, increased processing time, more operations, and higher cost. It is thus essential to determine the appropriate neural network and suitable data. OpenAI's ChatGPT-3.5 has 570 GB of training data, and the number of parameters for the size of the network that connects these data is said to be 175 billion. This amount of data is roughly the amount of words spoken by an adult male per day, which is said to be 7,000 words (20,000 words for women), and when converted to annual character count and bytes, it comes to 5.11 million bytes. In other words, the training data for ChatGPT-3.5 are equivalent to the number of words uttered by 110,000 adult men in a year. Whether this amount feels like a lot or a little depends on the individual, but it is the

Fig. 5. Features of tsuzumi.

actual amount of data that ChatGPT-3.5 learns from.

The other factor is the size of the neural network that connects data. A neural network is structured in a similar manner to the human brain, and humans recognize things by connecting over 100 billion nerve cells and 100 trillion synapses from input to output. The size of the ChatGPT-3.5 network, estimated to be 175 billion parameters, means that the network connects up to 1/600th the functions of the human brain. In other words, ChatGPT-3.5 must learn a lot of data, and that learning incurs high costs. Given these issues, we released NTT's version of a large language model (LLM) called "tsuzumi."

The tsuzumi LLM has four key features (**Fig. 5**). The first feature is it being lightweight. Both the amount of training data and number of parameters connecting the data are small. ChatGPT-3.5 has 175 billion parameters, but tsuzumi has only 7 billion, which is 25 times fewer. Thus, tsuzumi is easy to train on closed company data because of its low number of parameters. The second feature is its high language performance. NTT has been researching natural-language processing (NLP) for over 40 years. Therefore, we can effectively connect various NLP parameters and neurons for training LLMs, making it possible to attain high language performance. Despite being lightweight, tsuzumi has been evaluated as having language performance exceeding that of ChatGPT-3.5. The third feature is customizability. Training ChatGPT-3.5 on closed data from companies, etc. requires a huge number of parameters and GPUs, which requires high power consumption. However, tsuzumi can be easily customized with a small number of GPUs, so it can be tuned with low power consumption and at low cost. The fourth feature is multimodality. That is, tsuzumi can be trained on not only text data but also various figures and tables. "Lightweight = high cost performance" means that tsuzumi can be trained in 1/25th of the time it takes to train ChatGPT-3.5, which results in lower hardware, electricity, and operating costs. Despite it being lightweight, tsuzumi gives better answers than the top-ranked AI in Japan and ChatGPT-3.5, so it is currently ahead of its rivals.

4. Response to tsuzumi

Since the launch of tsuzumi in November 2023, we have received inquiries from more than 500 customers, mainly from corporate customers. Taking advantage of tsuzumi's features, 60% of customers want to train tsuzumi with their personal information and highly confidential information in their environments in a closed and secure manner, and about half the

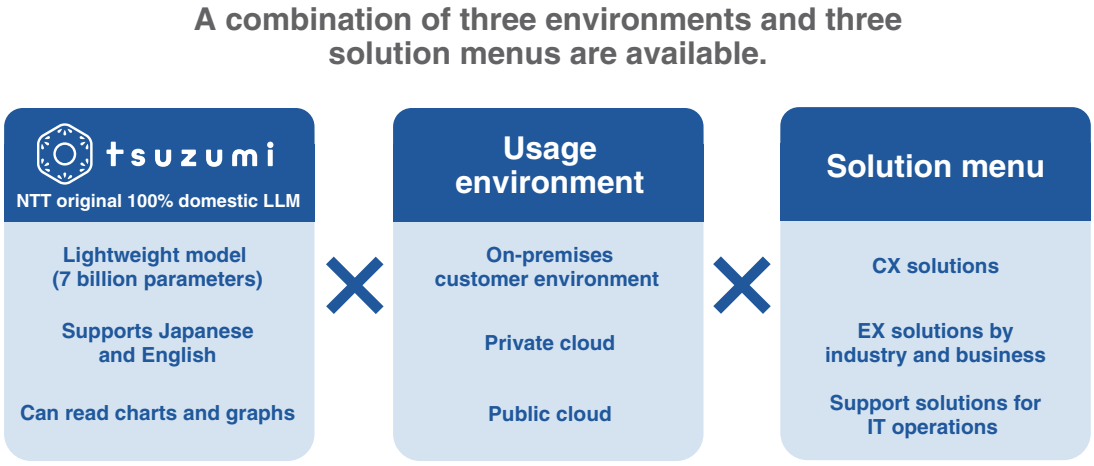


Fig. 6. Commercialization of tsuzumi.

customers are in the manufacturing, municipal, financial, and information technology (IT) industries. Many customers want to use tsuzumi for improving CX and customer response, improving employee experience (EX), and IT and automating operations. Given these customer requests, we started providing three commercial "solution menus" from March 2024 (**Fig. 6**).

I will now give examples of using tsuzumi in the healthcare field. Does drinking coffee rejuvenate blood vessels or constrict them? In fact, it depends on the genes that each person possesses: it will have a positive effect on some people and a negative effect on others [11]. Similarly, the effectiveness of the drug "warfarin" for dissolving blood clots varies by up to 20 times depending on an individual's constitution and medical history [11]. In other words, for some people, one tablet is effective, and for others, 20 tablets are required. The current average dosage is considered 10 tablets. Consequently, some people may end up taking too much, but others may find that the drug has no effect at all. Given these issues, we are using tsuzumi to structure and analyze electronic medical records (information) in a way that makes it possible to personalize such medications. Electronic medical records have been introduced in over 90% of large hospitals in Japan [12]; however, they are written qualitatively, so they cannot be collected and analyzed. Although we have a system in place for structuring and formatting the data (records), we want to be able to transcribe the data automatically (since they are currently transcribed manually). For example, it takes one person a day to input various test data into a medical record; however, by using tsuzumi, it is possible to transcribe electronic-medical-record data and analyze them automatically in a manner that personalizes healthcare.

5. Energy-saving technology that supports the benefits of AI

Although using AI will make many things more convenient, the total amount of data generated and used in the world will increase. The electricity consumed by one training session for ChatGPT is greater than the amount of electricity generated by a single nuclear power plant in one hour; put another way, the more convenient it becomes to train it, the closer we get to the limit of the Earth's energy capacity. The power consumption of datacenters is also increasing rapidly, and building more datacenters poses the risk of power shortages for the general public. With that power issue in mind, some countries, such as the Netherlands and Singapore, have temporarily halted construction of datacenters.

IOWN is a technology that uses the power of light to reduce electricity consumption. Looking for ways to compensate for the shortage of datacenters, we are using optical technology—from transmission to data processing—and incorporating the features of IOWN, namely, low power consumption, large capacity, and low latency, into datacenters and networks. We launched the All-Photonics Network (APN) IOWN1.0 in March 2023 and the APN leased-line plan powered by IOWN in March 2024.

Building many large datacenters would significantly

impact electricity use in everyday life. We thus believe that this impact can be lessened by (i) building smaller, distributed datacenters or datacenters in locations where natural renewable energy can be used and (ii) connecting them with the APN to enable them to be used without delay as if they were a single datacenter. In fact, although the previously mentioned training of tsuzumi is being researched at our Yokosuka R&D Center, because space to install a GPU was not available there, we trained tsuzumi by connecting the Yokosuka R&D Center to a GPU installed at the Musashino R&D Center in Musashino City, 100 km away, via the APN. We are thus able to use the training data as if they were in Yokosuka.

In our overseas datacenters, we have experimentally connected datacenters 100 km apart via the APN. NTT currently possesses 98 global datacenters and 148 buildings. NTT's datacenters are ranked as the third most abundant in the world. We have also achieved a leading position in a market evaluation by IDC [13]. Demand for datacenters is forecast to grow at an average annual growth rate of 13.5%. With both hyperscalers and enterprises growing at the same rate, adding expected demand for generative AI will add a further 20% growth in demand for datacenters. Due to the above-mentioned issue of power consumption, supply is not enough to meet such high demand, so we hope to use the APN globally to resolve the supply shortage. We plan to double the current total power supply to each of our global datacenter locations (1100 MW) by FY2026 [14].

We are also considering using the APN for dynamic control of mobile base stations, that is, reducing power consumption by turning antennas on and off in accordance with the number of people covered by each antenna. In collaboration with Sony, we are also promoting remote production. We have been using temporary lines to stadiums each time a game is held, and people have been going to the stadiums to edit video of the game on the spot. However, we are considering a set up by which APN lines are connected to all the stadiums from the beginning, and the game video can be edited remotely simply by using those lines.

The construction industry now has a serious labor shortage, and to cope with this situation, we developed a system that allows construction equipment to be controlled remotely. At EXPO 2025 Osaka, Kansai, Japan, we will connect the NTT Pavilion and NTT's datacenter via the APN, which will make it possible to transmit the "Expo space" to the datacenter and analyze in real time what kind of people are visiting the NTT Pavilion and the emotions on their faces. From the analysis results, we will create a "living pavilion" that reacts to visitors' reactions in real time.

6. Aiming for social well-being

We eventually hope to make even the wiring inside a chip optical, thereby lowering power consumption from transmission to processing. I believe that changing only a part of the social infrastructure will be ineffective, that is, the entire society must change to create a low-consumption infrastructure. We are looking forward to working with you to achieve this change. OpenAI has created very large LLMs. We are considering an "AI constellation" that combines many lightweight, small, expert LLMs connected by IOWN. In this AI constellation, multiple LLMs—each with a different specialty—are connected in the form of a constellation so that they can "discuss" questions with each other and come up with answers. Those answers are then compiled to elicit better answers that are deeper and more diverse than those that can be elicited from larger LLMs. We plan to achieve sustainable transformation (SX) by using optical transmission and optical transmission processing and dramatically reduce the power consumption of technology by improving the power efficiency, processing efficiency, and training efficiency of AI models. To achieve this, we hope to work with you to build a social infrastructure that will eventually enable personalization of not only food and healthcare but also entertainment. Through SX by using AI and IOWN, we aim to achieve "human and Earth-friendly social well-being" that allows a variety of people to live healthy, happy lives.

References

[1] Ministry of Agriculture, Forestry and Fisheries (MAFF), "Food Self-sufficiency Rate," aff, Feb. 2023 (in Japanese). https://www.maff.go.jp/j/pr/aff/2302/spe1_02.html

[2] MAFF, "Statistics on Agricultural Labor Force (2022)," 2022 (in Japanese).

[3] MAFF, "2020 Agriculture and Forestry Census," 2020 (in Japanese).

[4] MAFF, "White Paper on Fisheries," 2022 (in Japanese).

[5] Food and Agriculture Organization of the United Nations (FAO), "The State of World Fisheries and Aquaculture Production," 2022.

[6] GI value of banana (in Japanese), https://www.gi-gl.com/kudamono-50/

[7] GI value of cookies (in Japanese), https://www.gi-gl.com/keisyoku-118/

[8] D. Zeevi, T. Korem, N. Zmora, D. Israeli, D. Rothschild, A. Weinberger, O. Ben-Yacov, D. Lador, T. Avnit-Sagi, M. Lotan-Pompan, J. Suez, J. Ali Mahdi, E. Matot, G. Malka, N. Kosower, M. Rein, G. Zilberman-Schapira, L. Dohnalova, M. Pevsner-Fischer, R. Bikovsky,

Z. Halpern, E. Elinav, and E. Segal, “Personalized Nutrition by Prediction of Glycemic Responses,” Cell, Vol. 163, No. 5, pp. 1079–1094, 2015.
[9] Nikkei XTECH article, Jan. 10, 2024 (in Japanese). https://xtech.nikkei.com/atcl/nxt/column/18/02683/011000015/
[10] Nikkei XTECH article, Feb. 2, 2024 (in Japanese). https://xtech.nikkei.com/atcl/nxt/news/24/00170/
[11] NHK, “Special Program on The Human Body (Jintai),” May 2021.
[12] Ministry of Health, Labour and Welfare, “Trends in the Spread of Electronic Medical Record Systems, etc.,” 2020 (in Japanese).
[13] IDC, “IDC MarketScape: Worldwide Datacenter Colocation and Interconnection Services 2021 Vendor Assessment,” June 2021. https://info.equinix.com/rs/180-SLL-021/images/ar_idc_datacenter_and_colocation_vendor_assessment.pdf
[14] NTT, “Growth Strategy of NTT Global Data Centers,” 2023. https://group.ntt/en/ir/library/presentation/2023/pdf/231003_1.pdf

Sachiko Oonishi
Executive Vice President, Head of Research and Development Market Strategy Division, NTT Corporation.
She joined NTT in 1989 and became vice president of Regional Economy Vitalization, Strategy Business Development Division in 2016. In her career at NTT Communications, she became a member of the board, head of the Third Business Solutions at the Business Solution Division in 2020 and senior vice president, head of the Third Business Solutions at the Business Solution Division in 2021. She has been in her current position since 2023. She has also been an executive member of the board, chief customer experience officer, and co-chief artificial intelligence officer, NTT Corporation since June 2024.