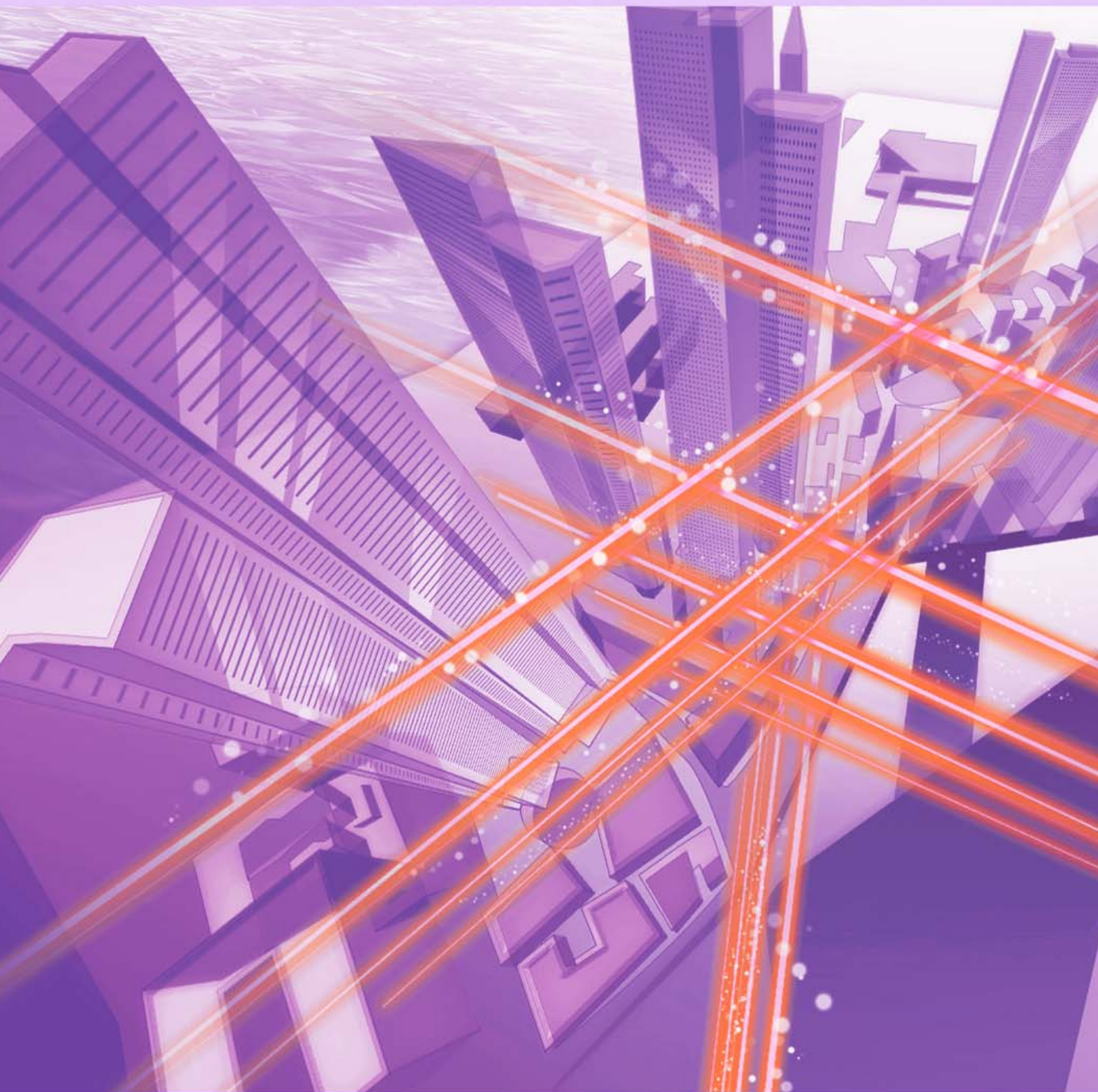


# NTT Technical Review

8

2024



August 2024 Vol. 22 No. 8

## **NTT Technical Review**

### **August 2024 Vol. 22 No. 8**

#### **View from the Top**

- Shinichi Yokohama, Chief Executive Officer, NTT Security Holdings

#### **Front-line Researchers**

- Kyosuke Nishida, Senior Distinguished Researcher, NTT Human Informatics Laboratories

#### **Rising Researchers**

- Mitumasa Nakajima, Distinguished Researcher, NTT Device Technology Laboratories

#### **Feature Articles: R&D and Commercialization of NTT's Large Language Model "tsuzumi"**

- NTT's LLM "tsuzumi"
- NTT's LLM "tsuzumi": Capable of Comprehending Graphical Documents
- Commercialization of NTT's LLM "tsuzumi"

#### **Global Standardization Activities**

- Report on ITU-T Industry Engagement Workshop

#### **Practical Field Information about Telecommunication Technologies**

- Extending Usability of Copper Cables Installed Underground by Using a Cable-retracting System

#### **External Awards**

## When You Think You Are About to Hesitate, Step Forward. Increase Freedom and Trust People in the Field



***Shinichi Yokohama***  
***Chief Executive Officer,***  
***NTT Security Holdings***

### **Abstract**

In today's connected society, ensuring safety and security in cyberspace is an extremely tough challenge. NTT Security Holdings provides security services to clients on the basis of its proprietary cyber intelligence and threat-detection-and-response capabilities. We asked Shinichi Yokohama, chief executive officer of NTT Security Holdings and chief information security officer of the NTT Group, about the company's mission and strategy as well as his attitude as a top management.

*Keywords: cybersecurity, security operations center, trust*

### **The strengths of NTT Security Holdings—NTT's only company specializing in security**

*—Could you tell us how NTT Security Holdings was established and what are its features?*

NTT Security Holdings is the only company in the NTT Group that specializes in security. Our mission is twofold: first, to maintain the security of the NTT Group's companies and support their activities and second, to provide security services to our clients.

NTT Security Corporation was established in 2016 and was temporarily a division of NTT Ltd. In 2022, on the recognition that security will become even more important to the business strategy of the entire NTT Group, it was reorganized as a direct subsidiary of the holding company and renamed NTT Security Holdings.

I believe that NTT Security Holdings has five unique strengths. The first strength is NTT's scale. Since NTT operates a huge information and commu-

nications infrastructure, it is the target of numerous cyberattacks at any moment. Therefore, we are able to learn about the world's most-cutting-edge cyberattack methods.

The second strength is our advanced technologies for promptly detecting cyberattacks and rapidly responding to and recovering from them. Cyberattacks have become so complex and sophisticated that 100% protection is no longer possible. Accordingly, we assume that an intrusion (cyberattack) from outside will occur, and we are required to detect it quickly, respond to and recover from it, and minimize the damage it causes. To meet those requirements, the NTT Group operates a security information and event management (SIEM) engine—which also uses artificial intelligence—to collect activity and log data from security devices and automatically detect, visualize, and notify threats. To make this SIEM engine function even more effectively, we have a global partnership for collecting and sharing threat information with highly skilled analysis engineers present at





our globally linked security operations center (SOC). Therefore, a virtuous cycle is created in which the input threat information is analyzed by our talented engineers and knowledge is accumulated, which allows us to provide fast and effective services.

*—The above-described environment based on a huge global information technology (IT) infrastructure and excellent technical capabilities alone are of competitive advantage.*

Our third strength is an abundance of human resources. Employees of the domestic NTT Group companies are required to take cybersecurity-related training. Approximately 3% of its employees in Japan, or about 4500 people, have obtained the certificate of “intermediate level,” which recognizes them as people who can play an active role in the cybersecurity field. NTT also employs approximately 100 experts having industry-leading achievements recognized by outside experts. Since NTT operates a huge information and communications infrastructure, we can develop those talent in our own cybersecurity field.

Our fourth strength is expertise backed by a wealth of experience. As a partner of the international sporting event held in Tokyo in 2021, we were able to play

a part in the defense against cyberattacks and other threats to that event. We have also accumulated experience in handling major international events such as the G7 and G20 Summits. Global events are held every year in Japan, and every time a VIP from around the world visits Japan, such as when Ukraine’s President Zelensky visited the G7 Hiroshima Summit, it attracts the world’s attention. Such a stage is ideal for those who want to launch cyberattacks. I believe that our strength lies in the expertise we have acquired from our experience for cyber defense in these situations.

#### **Teaming up with clients and partners to build a more secure society**

*—With both the organization and knowledge in place, your company has a solid foundation. You have recently authored a book, correct?*

Yes, the book is called “Towards World-Class Cybersecurity Practice” and its publication is an example of thought leadership, which is our fifth strength. Efforts in cybersecurity are generally not disclosed because of concerns that doing so could provide attackers with hints on how to attack. In my role as the chief information security officer (CISO)



of the NTT Group, I exchange opinions on cybersecurity and other topics with various companies and organizations, and I meet many people who have problems in handling cybersecurity issues. Having the desire to help them, I decided to disseminate information about our cybersecurity efforts. We have a team dedicated to publicizing cybersecurity information. We have also set up a presentation room where our clients and partners can see the operational state of our SOC.

When we first started to disseminate cybersecurity information, we heard concerns that doing so might be handing an advantage to attackers; however, we are already being targeted enough by attackers. Since NTT aspires to become a truly globalized digital company, hiding is not an option for us. We changed our way of thinking and decided to team up with clients and partners instead of hiding. We live in the era of Internet of Things, where everything is digitally connected. I believe that it is important to protect not only our own company but also the entire supply chain, including electric-power companies and other companies on which NTT relies, as well as our clients in a manner that makes society as a whole more secure.



*—Many cybersecurity-related incidents have recently been reported. How do you view this situation?*

I am actually very positive about the cybersecurity being the focus of attention. When I first started working in the cybersecurity field ten years ago, I felt that awareness of cybersecurity in society as a whole was low. For example, the cybersecurity incident at a large international media company 10 years ago was covered by the media, but for many people and companies, it may have been perceived as “someone else’s problem.” However, it is now widely known that cybersecurity incidents occur frequently both domestically and internationally and that responding to them requires much effort and expense. I think there is a growing awareness that cybersecurity incidents are no longer “someone else’s problem” but a “crisis happening right now” and “a threat to our company that could occur at any time.” In other words, society is becoming more aware of the importance of cybersecurity. There is also a growing awareness that cybersecurity is not an IT issue but a management issue.

Even if companies recognize that cybersecurity is a management issue, many do not know what they should do and to what extent they should take measures. In many situations, therefore, a security officer is appointed and security matters are left to them, and those security officers are troubled about what they should do and to what extent they should take measures.

The NTT Group consists of both large companies and small companies, such as startups with around ten employees. Such diversity of companies is what characterizes the NTT Group. We have a wealth of expertise in cybersecurity in regard to both large and small companies, and we want to offer that expertise to our clients.

The NTT Group also experienced a major data breach due to internal fraud at NTT WEST. We profoundly regret that the incident lasted 10 years. To regain the public’s trust, we intend to raise the level of our internal security.

External services and internal defense for the NTT Group are two wheels of the same cart for fulfilling our responsibilities. Regarding external services, our diverse clients range from domestic companies to multinational corporations of all sizes. For all of these clients, we leverage our five above-mentioned strengths to provide consulting, knowledge provision, system integration, and other security-related services that are tailored to their needs and environments.



In regard to internal defense for the NTT Group, we will ensure that the so-called Three Lines of Defense Model of governance and risk management (defined by the Institute of Internal Auditors) is fully functional, and the incident that we experienced is an opportunity for us to go back to basics and reaffirm our commitment to security. While it is fundamental that each operating company takes the initiative, NTT Security Holdings, under the direction of the NTT holding company, will promote complementary initiatives to enhance the collective strength of the NTT Group. For example, since security is not only a technical issue but also a management issue, NTT Security Holdings will provide training to the presidents of domestic NTT Group companies.

#### **Trust each other and become a reliable presence**

*—The pressure of being the head of security for the NTT Group seems considerable. What attitude do you have towards your daily work?*

I sleep with my smartphone next to my pillow, and since I receive reports of security incidents that occur for the NTT Group both in Japan and abroad and respond to them 24 hours a day, I would be lying if I said I feel no pressure. When an incident occurs, the

most-important task is to minimize the damage it causes. To minimize the damage, it is important to give freedom to the people in the field who understand the situation better than anyone else, including the incident that has occurred.

It doesn't matter if there is a slight delay in reporting a security incident to me, the CISO of the NTT Group. To be honest, it can be frustrating if the situation is not reported to me immediately, but I try to think that it is because the people in the field are busy doing everything in their power to minimize the damage. I think it is important to increase the freedom of those in the field and trust them.

I joined NTT after working at the Ministry of International Trade and Industry (now the Ministry of Economy, Trade and Industry) and at a consulting firm. When I worked as a consultant for a major beverage company, I would sometimes ride with the company's delivery truck drivers as they drove around town to think about measures to improve the sales capabilities of the company. One time, a driver told me about the things you won't get to hear unless you are in the field, including a delicious local Chinese restaurant at which drivers would eat between shifts. Listening to their voices, I became convinced that leveraging their strengths, namely, knowing the overall situation and environment in the field, would lead

to good work. This experience was the starting point of my desire to trust the people in the field.

It is also important to communicate my belief to the people in the field. Therefore, at my first meeting as the chief executive officer (CEO) of NTT Security Holdings, I told the participants that my belief in my work is “When you think you are about to hesitate, step forward.” I’m not sure if this belief always produces the right results, but if we proceed and find that we are wrong, we can correct mistakes immediately.

*—Employees must be reassured to know that the CEO has confidence in their work. What are your future aspirations and what would you say to readers?*

A good corporate culture cannot be created overnight, but I want to make it a priority to maintain unwavering trust in those in the field. To that end, I focus on the things that only I can do, and I want each and every employee to focus on the things that only they can do. This way of working is important not only for gaining my trust but also for employees to gain the trust of each other in a manner that enables us to all to be a reliable presence and say, “I can leave it to them.”

Although NTT’s security business is still in the development stage, we aim to be a world-class security company. Achieving this aim does not mean being the best in the world; it means being proud to sit at the same table as top-class companies. I have said that external services and internal defense for the NTT Group are two wheels of the same cart, but it is

important that NTT Security Holdings becomes the “security hub” of the entire NTT Group, and I believe that if NTT can pool all its strengths, we will be able to win trust externally.

I want all of you involved in research and development to approach your work with pride. Expectations from around the world concerning the security capability of Japan and NTT are growing. In that sense, NTT has an enormous opportunity to excel. I want you to conduct research and development with pride in building security practices in your field. To our clients and partners. We will disclose as much of our knowledge and expertise as possible, so please join us in building a more secure society.

### Interviewee profile

#### ■ Career highlights

Shinichi Yokohama joined the Ministry of International Trade and Industry (now the Ministry of Economy, Trade and Industry) in 1984 and McKinsey & Company, Inc. in 1992. He became the head of the Business Integration Department at the Global Business Headquarters of NTT DATA in 2011 and a special advisor to NTT, responsible for increasing the external presence of cybersecurity, in 2014. He has been the CISO of the NTT Group since 2018 and CEO of NTT Security Holdings since June 2023.



## NTT's Large Language Model “tsuzumi”: A High-performance and Low-energy-consumption Large Language Model with Expertise in Specific Fields

***Kyosuke Nishida***  
***Senior Distinguished Researcher,***  
***NTT Human Informatics Laboratories***

### **Abstract**

Since ChatGPT, OpenAI's generative artificial intelligence (AI) chatbot, was released in November 2022, large language models (LLMs) have rapidly drawn worldwide attention, and new LLMs have been announced one after another. While the “bright side” of such LLMs, namely, high performance, has been brought to attention, their “dark side,” namely, increased energy consumption, has also been pointed out. On November 1, 2023, NTT's LLM called “tsuzumi” was announced. It was developed with a particular vision in mind, creating a future in which many and small LLMs having different characteristics are combined, rather than creating one huge and monolithic LLM, to achieve social well-being through AI-human collaboration. Thus, tsuzumi has made it possible to achieve the bright side and address the dark side of LLMs simultaneously, and its commercial service was launched on March 25, 2024. We interviewed Kyosuke Nishida, a senior distinguished researcher at NTT Human Informatics Laboratories, about how the tsuzumi project was launched, features of tsuzumi, trends in LLMs, and his thoughts on the challenges of competing in a highly competitive market.

*Keywords: large language model, tsuzumi, natural language processing*



## NTT's large language model "tsuzumi" was developed—from announcement to commercial launch—in five months

*—Could you tell us about the research you are currently conducting?*

I am researching large language models (LLMs), including NTT's tsuzumi, which was announced in November 2023. Since joining NTT in 2009, including a stint at an operating company from 2013 to 2015, I have been involved in various research and development activities. In the field of natural language processing, which I began researching around 2017, I have been focusing on themes such as machine reading of natural language text by artificial intelligence (AI) models and visual machine reading, which combines visual information with textual information to help AI models understand documents in the same way humans do when reading documents. My research colleagues and I received Paper Awards from 2018 to 2024 (Best Paper Awards in 2018 and 2021) at the Annual Meeting of the Association of Natural Language Processing and were the runner-up in the Document Visual Question Answering competition at the International Conference on Document Analysis and Recognition in 2021.

After Google announced their Bidirectional Encoder Representations from Transformers (BERT) language model in 2018 and NTT began constructing an NTT version of a Japanese BERT model, I started researching language models. In parallel with the release of ChatGPT, OpenAI's generative-AI chatbot, in November 2022 and LLMs gaining worldwide attention, I sensed that a global paradigm shift toward the emergence of early versions of general-purpose AI that had never existed was underway. I therefore emphasized the necessity and importance of the development of LLMs at NTT and launched a project on LLMs in February 2023. Since then, I have been investigating LLMs as my main theme.

In this project, I procured computing resources, collected the data necessary for training, and built an LLM with team members. The development of our LLM was completed in a very short time; pre-training of the LLM started around June 2023, tsuzumi was announced on November 1, 2023, and the start of its commercial service was announced on March 25, 2024.

*—Could you explain what an LLM is?*

A language model is a machine-learning model that models the likelihood of generating a sequence of words (tokens) and predicts the probability of future (or missing) tokens. An LLM is based on a deep-learning model in the form of a neural network called a transformer and enables a variety of language-processing tasks such as highly accurate information retrieval and generation and modification of computer programs. An LLM's capability of advanced language processing is primarily achieved through two processes: pre-training and instruction tuning. Pre-training is the process of pre-training a language model, i.e., the transformer is pre-trained with a large corpus of text to create a pre-trained language model. Since its emergence in 2018, BERT has been gaining attention as an effective language model for natural language processing. Since the release of GPT-3 in 2020, the capability of pre-trained language models has advanced to the point that these models can handle any task to some extent—without fine-tuning—by having them generate text following task-defining prompts.

Instruction tuning is the process of explicitly learning pairs of instructions and responses for various tasks to further develop the ability acquired through pre-training. In conventional tuning using task data, for example, two sentences are input to BERT as tokens called "separators" to produce a score with three labels, "0" (yes), "1" (no), or "2" (possible), which is used to determine the relationship between the two input tokens. This process is data-driven learning, i.e., the language model is trained in a data-driven manner with no knowledge of what tasks it is given. In instruction tuning, however, a task is defined in language. Specifically, the task that the LLM is to solve or the output the LLM should produce in response to the input text (instruction) is defined in language so that the LLM can respond to those instructions (**Fig. 1**). This process—instruction tuning—increases the versatility of an LLM and enables it to respond to unknown tasks without learning.

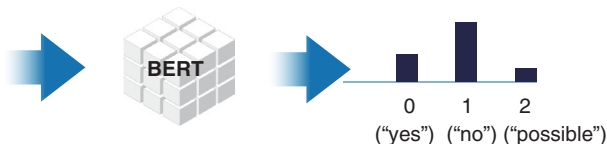
Current LLMs are becoming larger, and following scaling laws\*, language models with a large model size (number of parameters) are being created; in fact, models with over a trillion (1000 billion) parameters

\* Scaling laws: Empirical laws stating that as model size (number of parameters), dataset size, and the amount of computation used to train a natural language-processing model increase, loss decreases according to a power law.

- **Explicitly learn** pairs of instructions and responses concerning a variety of tasks to further develop the ability acquired through pre-training.
- Instruction tuning increases the versatility of an LLM and enables it to respond to unknown tasks without learning.

#### Image of conventional tuning using task data

[CLS] Seafood is served on a plate.  
[SEP] Many types of vegetables are served on a plate.

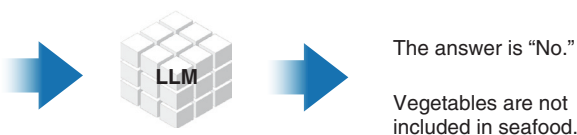


#### Instruction-tuned model

Can an implication relationship between the premise and hypothesis sentences be established?

Premise: Seafood is served on a plate.  
Hypothesis: Many types of vegetables are served on a plate.

Output: Choose from "Yes", "No", or "Neither" and provide a reason for your answer.



CLS: classification

SEP: separator

Fig. 1. Instruction tuning.

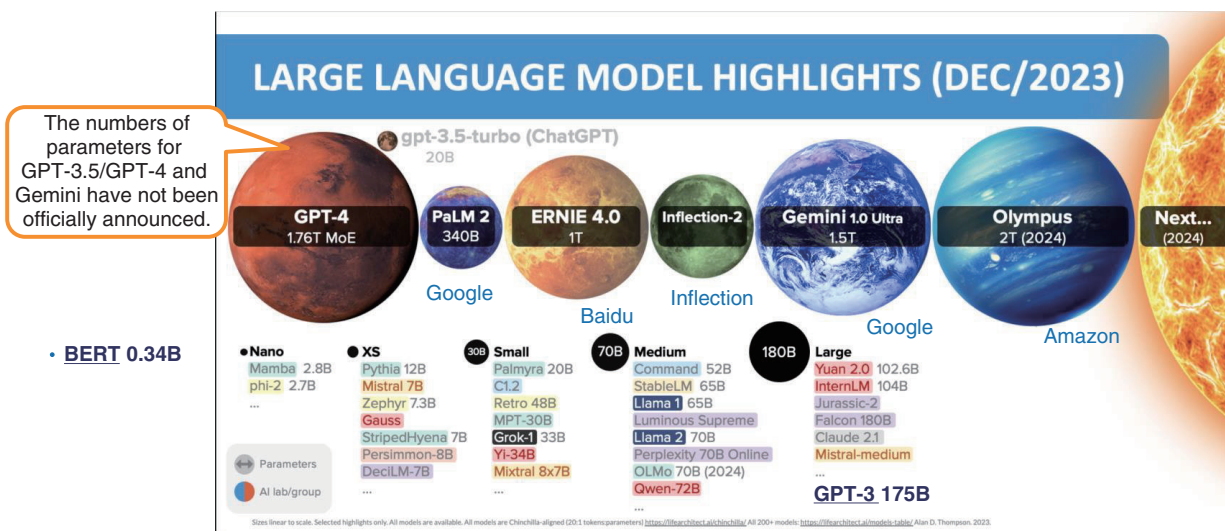


Fig. 2. Increasing the scale of LLMs.

are under development (Fig. 2). For example, GPT-4 is said to have 1.76 trillion parameters.

As this trend progresses, however, energy consumption for using and training LLMs increases exponentially, and costs increase accordingly. It has been estimated that training an LLM such as GPT-3, which is said to have 175 billion parameters, uses

about 1300 MW/h of electricity per training session, which is equivalent to an hour's worth of electricity produced by one nuclear power plant.



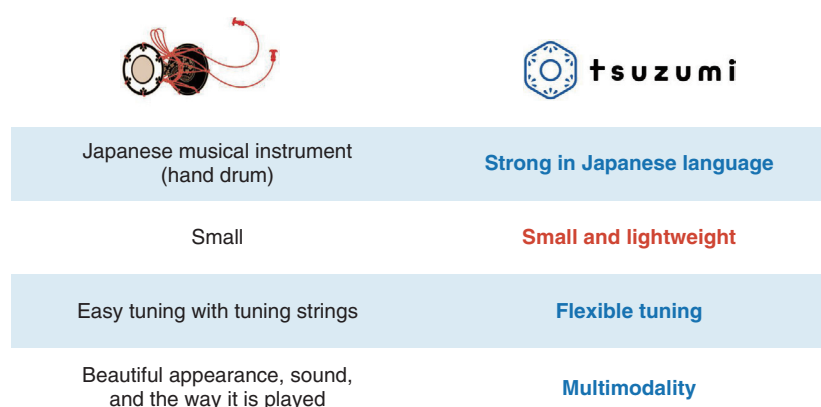


Fig. 3. Japanese hand drum and tsuzumi.

**Japanese hand drum and tsuzumi are both associated with the keywords “Japanese language,” “compact and lightweight,” “flexible tuning,” and “multimodality”**

—*What kind of LLM is tsuzumi?*

While the global trend is toward larger-scale, general-purpose LLMs that can do anything, their power consumption is becoming problematic. We therefore adopted the vision of creating a future in which many and small LLMs having different characteristics are combined, rather than creating one large-scale and monolithic LLM, to achieve social well-being through AI-human collaboration. With that vision in mind, we decided to develop a high-performance LLM with specialized knowledge in a specific field by improving the quality and quantity of language training data, rather than competing on model size. We also intended to pursue the concept of “AI constellation,” in which multiple small and high-performing LLMs, such as LLMs that are strong in the medical and educational fields, are combined to form a superior LLM. We named our LLM “tsuzumi” after a traditional Japanese hand drum (Fig. 3), because our LLM has the following features: strength in the Japanese language, high performance despite small and lightweight, ease and flexibility of tuning, extendibility to multimodal applications (multimodality).

I will present the evaluation results of tsuzumi based on the Rakuda benchmark, which compares the output of two models by using an evaluation set of 40 questions covering 4 categories, i.e., Japanese geography, politics, history, and society. This benchmark

uses GPT-4 as the evaluator instead of a human evaluator. In the evaluation example shown in Fig. 4, GPT-4 compares the output of tsuzumi with that of GPT-3.5 and rated tsuzumi highly in terms of specificity and detail. According to the overall evaluation using this benchmark, tsuzumi beat GPT-3.5 and other Japanese LLMs at an overwhelmingly high win rate, demonstrating tsuzumi’s high-level Japanese language comprehension and generation abilities (Fig. 5).

For pre-training tsuzumi, we prepared a total of over 1000 billion tokens in a wide range of domains from technical documents to entertainment. We will continue to improve the quality and quantity of training data. It is also important to consider the ratio of languages contained in training data during pre-training, and it is not yet clear what kind of language balance should be taken in pre-training to build an LLM that is strong in the Japanese language with a limited corpus of Japanese. Another company has recently adopted a method of additionally training non-Japanese language LLMs on Japanese language and performed well in Japanese language processing. We pre-trained tsuzumi by using in-house data from scratch and added codes and multilingual data, mainly in Japanese and English. We are continuing to verify this approach from various perspectives, particularly the effect of adding multiple languages to pre-training corpus.

When creating a pre-training corpus, we also focus on tokenization, which breaks down sentences into tokens, and preprocessing. Tokenization is done by taking into consideration Japanese words. If word constraints are not taken into account, unnatural and redundant tokens can easily be generated due to bias

## Question

Please give examples of the social and political changes that occurred during the transition from the Heisei era to the Reiwa era and explain their significance.

## Answer by tsuzumi-7B

The transition from the Heisei era to the Reiwa era has seen many social and political changes. Below are some of the main changes.

1. Spread of the Internet and digital technology: During this period, information and communication devices such as smartphones and personal computers spread rapidly, which has greatly affected people's lifestyles. It is especially noteworthy that the number of users of social networking services (SNSs) has increased, which has made it easy to communicate among individuals. Moreover, cloud computing and AI have advanced and they are now being used in the business world (...)

## Answer by GPT-3.5

Several social and political changes occurred during the transition between the Heisei and Reiwa eras. Some of them are described below.

(1) Political change: Compared to the Heisei era, the Reiwa era has seen continued political stability, and the current administration is supported by the younger generation. This state of affairs indicates that political values are changing from generation to generation. Compared to the Heisei era, the Reiwa era has seen an increasing population inflow from overseas, and policies to accept foreign human resources are being promoted. This trend is important in terms of helping to cope with globalization and Japan's declining population. (...)

## Judgment by GPT-4: usefulness, relevance, accuracy, detail, Japanese-language ability

The answer of Assistant 1 (tsuzumi) provides a detailed explanation of five specific social and political changes that occurred during the transition from the Heisei to Reiwa eras. The significance and impact of each change is specified, so the information presented is highly useful and relevant. In addition, the Japanese expressions are appropriate, and the text is easy to understand. (...) Both tsuzumi-7B and GPT-3.5 understood the intent of the question and provided appropriate answers; however, I judge Assistant 1's answer to be superior because it **provided more specific and detailed information**. My evaluation is therefore higher for Assistant 1.

Fig. 4. Example of judgment based on the Rakuda benchmark: tsuzumi-7B vs. GPT-3.5.

### Improved world-class Japanese language ability of tsuzumi and increased win rate against GPT-3.5 by about 30%

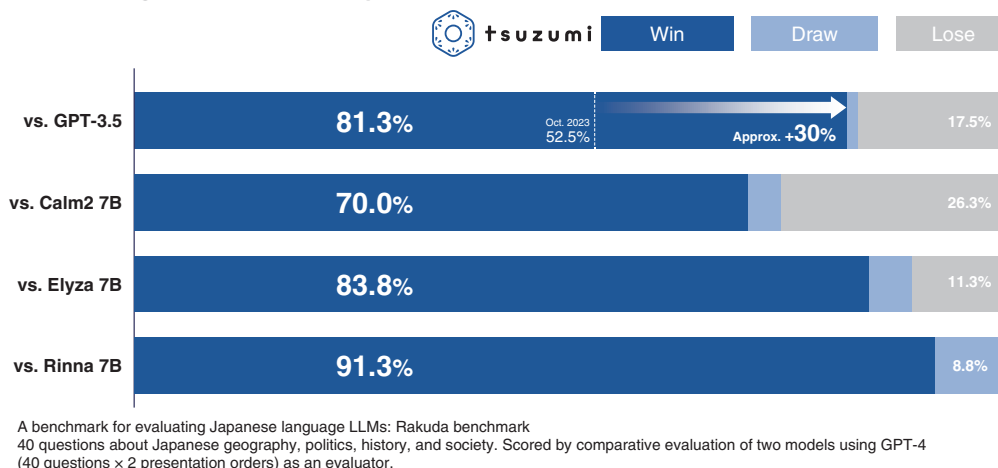


Fig. 5. Comparison of Japanese language LLMs based on the Rakuda benchmark.

in the training corpus. Our tokenizer incorporates the results of NTT's research on word segmentation that has been conducted for many years, achieving the word segmentation that is close to that of natural language. To improve the quality of the text that is gen-

erated as much as possible, we use various methods to remove noisy text. Regarding instruction tuning, we have compiled our own instruction-tuning data covering a wide range of categories of questions and instructions to use in training and are currently

maintaining and expanding the data from the perspective of both usefulness and safety.

*—Now that tsuzumi has been launched as a commercial service, what are your plans for the future?*

Conventional generative AI models may not perform well in fields where technical terms and industry-specific expressions are often used, such as medical care and software development. Since tsuzumi can be customized for industry-specific data, it can expand the application areas of AI. To improve customer experience, it is essential to enable AI to read and comprehend visual information such as charts and graphs in manuals, etc. and personalize AI services by updating customer information. By providing world-class Japanese-language processing capability as well as visual information reading comprehension capability, tsuzumi will support the evolution of customer-support areas such as contact centers and consultation chatbots.

In the medical field, although electronic medical records have been introduced in Japan, it has been difficult to collect and analyze medical-record data because the way of writing medical records depends on the hospital and doctor even for the same symptoms. By ultra-lightweight, flexible, and secure processing of medical-record data, tsuzumi reads and comprehends medical-record data recorded by doctors, arranges the data sets in a common format with appropriate expressions, and puts them in a state ready for analysis.

NTT DOCOMO's contact centers receive more than 40 million inquiries a year from customers. These centers currently use communicators (human operators) and chatbots to answer these inquiries; however, many types of manuals are used and new information is frequently added to them, so it takes much time and effort to create appropriate responses and lists of questions and answers for the chatbots. I believe that introducing tsuzumi into this process will enable AI to understand the content of calls and manuals with visual information, understand customer requests, and provide the most-appropriate answers.

We announced the commercial launch of tsuzumi on March 25, 2024, but we had just created a basic model in the research phase. We are currently working with engineers across the NTT Group to refine that model to provide value to our customers. While supporting that work, I want to improve the versatility of tsuzumi to achieve AI constellation—a next-

generation AI architecture in which multiple, small, and specialized LLMs are combined, rather than creating one huge, monolithic LLM—and advance research toward a world in which humans and AI can coexist naturally. Although LLMs have made remarkable progress, they still have many shortcomings that must be addressed before they can be implemented in society and assist people in all kinds of environments. It is particularly important to ensure that AI has an input-output interface comparable to that of humans and can tackle tasks that require multimodality and physicality. Our main focus is currently on the combined understanding of vision and language; however, I want to investigate how to link hearing, force, touch, and biological signals, such as brain waves, to language models.

### **Taking on the challenge with the awareness of what is important now and what should be done naturally now**

*—What do you keep in mind as a researcher?*

When pursuing my research, I want to do something that involves both academia and business, something important that everyone naturally thinks they should do at that moment in time. I often talk about this to my team members. Around 2017, I worked on the theme of “reading tables,” which involves computers automatically reading tabular data as humans can, and that technology was used in a solution for contact centers. Then, around 2018, I worked on machine reading and summarization to understand and generate natural language, and those technologies were offered as part of an AI service provided by NTT Communications called COTO-HA®. Since then, I have moved in the direction of enabling computers to understand human language under conditions more similar to those of humans, such as comprehension of visual documents, and have approached themes that can create value according to the technological level of the time. LLMs are a highly competitive field and a great challenge for me, but I felt that NTT should definitely develop an LLM, so I launched a project in February 2023.

While I was investigating machine reading comprehension as my research theme, I decided to participate in the MS MARCO challenge, an international competition for machine reading comprehension hosted by Microsoft. Until then, I had no experience competing with many other researchers on a major topic, but our model initially took first place, and the



letters “NTT” appeared on top of the leaderboard. Although I was delighted, I was also impressed that the people around me were also very happy. This experience made us believe that we could compete on a global stage, and we went on to participate in several competitions and ranked top of their leaderboards. It took courage to take on LLMs as a research theme because they are attracting much attention and evolving fast, but I feel that the successful experience at MS MARCO has pushed me forward. The business world has high expectations for LLMs, and as an academic topic, LLMs have much potential, from analyzing the mechanisms of LLMs to proposing new models, so I find them very rewarding to investigate.

—*What is your message to junior researchers?*

Mainstream research themes are “red ocean” fields in which many researchers are competing aggressively, and some people may avoid those fields because it is difficult to achieve results in those competitive environments. However, it can also be said that LLMs is the field that many researchers focus on because it is currently the most-important research theme. When you research LLMs, you can make new discoveries and achieve results, and I think there are many areas in which you can compete, so I encourage young researchers to take on LLMs. When taking on a new theme that is not mainstream, it is a good idea to be aware of whether the theme is valuable enough for other researchers to follow you. For example, when you are young, you might work on themes that you think they are important, even if they are somewhat incremental, then once you have gained more research ability, you might jump into mainstream

themes in red oceans. In any case, I believe it is important to take on the challenge of important and valuable themes.

Many researchers are investigating LLMs, making it truly a red-ocean field. However, it is a new and rapidly changing field, so it is not impossible for us to keep up with it, and I believe we have ample opportunity to compete at the forefront. We are working on the LLM project in cooperation with many people, and we expect to achieve good results.

I think it takes courage to change your research field or theme, but from my own experience of working in various fields and on themes other than natural language processing, I am glad that I was flexible and tried different things. If you challenge yourself to aim at what is important now and what you should naturally do now, you will feel relatively little resistance in changing fields and themes.

#### ■ Interviewee profile

Kyosuke Nishida received a B.E., M.I.S., and Ph.D. in information science and technology from Hokkaido University in 2004, 2006, and 2008. He joined NTT in 2009. His current interests include natural language processing, vision-and-language, and AI. He is currently serving as the technical lead for NTT’s LLMs. He received the Paper Award from the Association of Natural Language Processing from 2018 to 2024. He is a member of the Association for Computing Machinery, Association of Natural Language Processing, Information Processing Society of Japan, and the Database Society of Japan.

## Machine Learning Using an On-chip Photonic Neural Network—Changing the Future through Photonic Technology

*Mitsumasa Nakajima*

*Distinguished Researcher, NTT Device Technology Laboratories*

### Abstract

Artificial intelligence (AI) is now entering its fourth wave, and this time, it is penetrating everyone's daily life and bringing many surprises. However, training an advanced AI model requires a huge amount of resources such as time in units of several months and power equivalent to all of the power consumed by a medium-sized city. To solve this problem, studies are being conducted on AI-model computing using optical circuits that are fundamentally different in principle from past technology. In this article, we talked with NTT Distinguished Researcher Mitsumasa Nakajima about the potential of future photonic technologies and his research on “machine learning using an on-chip photonic neural network” that can greatly reduce the power consumed by machine learning through the use of photonics.

*Keywords: machine learning, optical circuit, optical computing*



### Revolutionizing next-generation computing technology through the properties of light

*—Dr. Nakajima, exactly what kind of technology is “machine learning using an on-chip photonic neural network”?*

In short, “machine learning using an on-chip photonic neural network” is technology that can significantly reduce the amount of power consumed by machine learning through the use of photonics. The background to the research of this technology is the remarkable progress in artificial intelligence (AI) in

recent years. The lives of many people in society are already benefiting from AI, but on the other hand, there are still many problems that need to be solved from the viewpoints of calculation speed and power consumption. For example, training an advanced AI model based on an electronic computer currently requires time in units of several months and power on the level of several hundred megawatts (equivalent to the power consumed by a medium-sized city). Since a huge amount of resources are already needed, it has been pointed out that future advances in machine learning will eventually hit a hardware limit by simply extending current technology (**Fig. 1**).

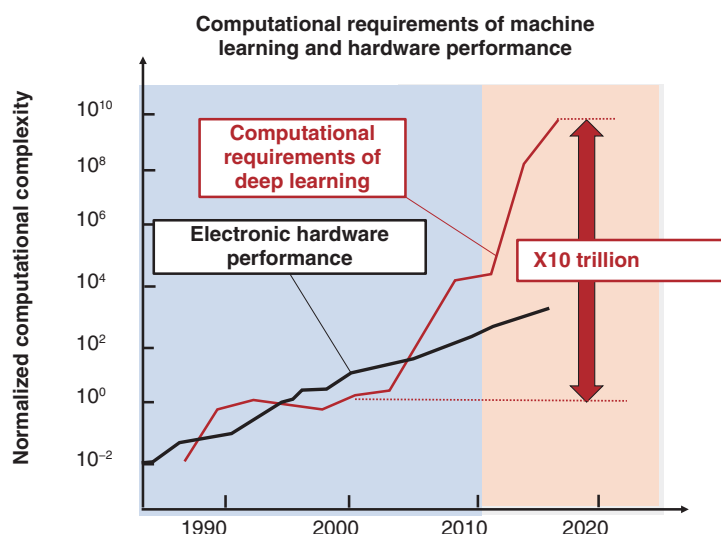


Fig. 1. Correlation between computational complexity of machine learning and performance [1].

To solve this problem, I am researching AI-model computing using optical circuits as hardware that is fundamentally different in principle from past technology. The reason as to why I began this research in the first place goes back to 2013 during the AI boom of that time. I was hardly aware of all the excitement since I had only overheard people in other departments saying things like “This is really amazing!” Eventually, however, I began to hear similar things from all sorts of people, and I too began to think, “This will no doubt become a major research theme that will somehow bring about a new era.” At that time, I was engaged in the research of optical waveguides and optical communication devices, and I came to focus on “analog matrix operations by light” as a contact point between machine learning and my own research. This was truly a calculation method at the heart of machine learning, so I got the idea that the technology developed for devices like switches incorporated in optical waveguides that I was researching could be used here as well. I then began my research.

*—What are the advantages of using optical circuits in AI computing?*

To be specific, making full use of the “space,” “wavelength,” and “time” features of optical operations makes it possible to execute matrix operations at the root of AI models at high speed with low power consumption (Fig. 2). It has been experimentally

shown that optical circuits that we have so far fabricated can achieve  $10^{16}$  calculations per second, which is about 1000 times faster than that of conventional general-purpose central processing units (CPUs), and can decrease power consumption per operation to the picojoule–femtojoule range, which is about one-tenth the conventional level (Fig. 3). Of course, it has been suggested for some time that using light has the possibility of enabling high-speed calculations, but announcing that we have actually demonstrated this to be true at academic conferences and elsewhere has been well received with many people expressing surprise. I have also received invitations like “Let’s try using the optical circuits that you fabricated in communication transmission experiments.” The wide network of fellow researchers that I have been able to cultivate through this research is a driving force behind my research to this day.

*—What difficulties have you encountered in your research?*

The most difficult thing that I had to face occurred when I first started out on this research, since I had no technical experience about machine learning. Fortunately, I was able to start this research with the support of a unique research fund in our laboratory. In this scheme, we can use part of our office hours and budget for our own research topic. Before I took on this research, I was engaged in a single research and development theme along with several tens of other



## Advantages of optical computing

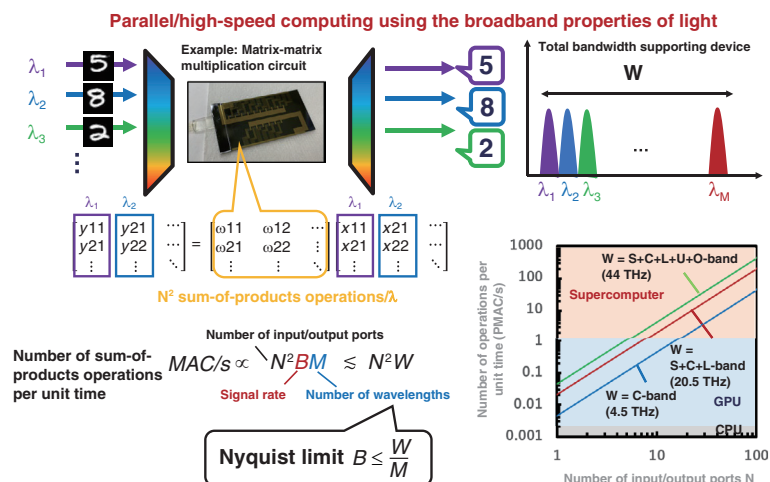


Fig. 2. High-speed computing exploiting the advantages of optical operations.

	Chip type	Fiber type	Our proposal (chip type)
Outline	 Nature com. 2014 IEEE JSTQE 2018	 Nature com. 2016 Optica 2016	 Commun. Phys. 2021
Multiplexing method	Space	Time	Space × Time × Wavelength
Number of nodes	≤ 32	≤ 1000	512
Size	cm <sup>2</sup>	(several tens of cm) <sup>3</sup>	cm <sup>2</sup>
Operation speed (operations/s)	≤ 500 giga	≤ 500 giga	10 tera/λ (parallel computation using wavelength division multiplexing is possible)
Function integration	Only reservoir layer is integrated on the optical circuit Not programmable 		Optical integration of input and reservoir layer Programmable 

Fig. 3. Comparison between conventional technology and this technology.

researchers. Once I became familiar with the work that I was placed in charge of, I would proceed with that work as my main line of research. However, since I myself had chosen this research related to machine learning, I had to do everything on my own. I, of course, had to grapple with device-related technology such as device design and evaluation for per-

forming optical computing, but in addition, I had to research machine-learning algorithms. Nowadays, there are plenty of books and know-how related to machine learning, but when I started this research, there was not a sufficient amount of well-organized information. To therefore study up on this field and accumulate knowledge, I began with a trial-and-error

process in which I would read journal papers and attempt to reproduce their results. Of course, I couldn't understand everything by doing only that, so I would sometimes take my questions to neighboring NTT laboratories or to university professors specialized in this field. Looking back, acquiring new knowledge and encountering fellow researchers in this way was a truly valuable experience.

*—Please tell us about your research vision going forward.*

In terms of future themes, I would first like to aim for a phase in which members of the computer industry come to use the results of this research. Current optical computing makes use of multiple units of large laboratory equipment, but in the future, I would like to enable optical computing to run at least on the cloud by making further advances in our research and refining the mechanisms involved.

Additionally, in terms of optical computing as part of NTT's Innovative Optical and Wireless Network (IOWN) vision, my aim is to convert communications between electronic computers and between CPUs, memory, etc. to photonics. Once we enter this era, I believe that optoelectronic integrated circuits having a mix of optical circuits and electronic circuits will become commonplace on computers. Then, as the next research stage, my goal will be to achieve practical device technologies that incorporate optical computing functions. Certainly, it's difficult to perform general-purpose calculations as commonly done on a CPU by optical means, but I think that it will be possible to perform those calculations not conducive to light by electronic means, and conversely, to perform those calculations not conducive to electronics by optical means. Dividing up calculations among electronics and optics in this way will enable general-purpose computational processing to be performed using light.

Another goal is to achieve an optical computer that improves power and calculation speed by several orders of magnitude compared with conventional computers. In this regard, I would like to contribute to the ongoing development of computations for ever-expanding AI and to the easing of environmental loads by achieving a dramatic reduction in power consumption. I also plan to search for applications that make good use of the unique features of photonic operations beyond computers. For example, signal processing in optical communications uses many matrix calculations, but it is thought that assist-

ing this processing by optical means can help reduce power consumption and lower delay in communications, so I am studying an expansion of my research into this area.

Through these efforts, I hope to contribute to performance improvements in optical computers, to the widespread use of optical computing, and to the creation of optical-computing applications. At present, the number of researchers engaged in the field of computational processing using light is increasing around the world, so I would like to enhance NTT's presence in this field while continuing to produce research results.

**Having a mindset that enjoys trial and error and is willing to say "Now I understand what won't work!"**

*—Please tell us about other research ideas that you would like to pursue.*

It has only been 80 years since the invention of the transistor, but the electronic computer based on that simple mechanism has come to enrich our lives beyond imagination. Yet, I don't think that this trend will necessarily continue in the 21st century and beyond. For example, it is thought that bits, neurons, and qubits will become important units of computing in the future. Bits are the units of von Neumann computers, neurons are the units of neural-network operations for AI, and qubits are the units of quantum computers. History has proven that digital electronic circuits are currently the best way to handle bits, but it is not yet known what types of circuits would be best for handling neurons and qubits. Additionally, the conventional idea was to do all bit operations on the same type of circuit (if the operation is electronic, do it all on electronic circuits). However, it is also possible to use different technologies as in using electronic circuits for gate operations and optical circuits for information transfer. In this way, when stepping out of the commonly accepted framework of conventional computers, optical computing becomes an attractive candidate. High-speed and low-power computing for AI using the parallel properties of optical wavelength, space, and time can be provided for neurons while high-speed, room-temperature quantum operations using the properties of photons can be provided for qubits. Going forward, I will conduct my daily research toward the spread of computing using light in these ways.

*—Please tell us about NTT Device Technology Laboratories that you belong to.*

Before entering NTT, I visited the company as part of my job search, and I remember how surprised I was on being introduced to an optical-fiber manufacturing method (vapor-phase axial deposition (VAD) method approved as an IEEE Milestone by the internationally respected IEEE) that supports current ultra-high-speed communications. This was because I was majoring in material physics at the time at university, and from the general knowledge that I had so far obtained, I thought that using this method to fabricate extremely transparent glass fiber as required in optical communications must be difficult. However, I heard that NTT researchers had identified the superior points of this technology and implemented and commercialized it as a foundation of today's society. In addition, this was around the time of the "Lehman shock" (the global financial crisis of 2008), and it was difficult for many companies to allocate resources to basic research that could produce innovative and foundational results as I just described. However, NTT, as always, placed great importance on basic research. Thinking that I would like to work in such a fortunate environment, I applied to NTT.

NTT Device Technology Laboratories that I belong to was formerly a research laboratory excelling in devices for communication purposes, so I believe that one of its key features is world-class device technology that was nurtured during that time. At the same time, its breadth of research is wide and its researchers are diverse. There are researchers who lead the world in communication devices as I just mentioned, researchers like me who put that technology to use in the search of new application areas, and researchers who seek to launch new technical fields and enter



cross-industry fields. Another of its features is the great diversity of human resources at other NTT laboratories at the same location including specialists in machine learning and quantum optics. In addition, it provides an environment that enables me to receive advice even on unusual research themes like mine and an environment that supports optical device prototyping and evaluation based on know-how accumulated over many years, all of which is a strongpoint of NTT. I believe NTT Device Technology Laboratories to be an open-minded organization that has many motivating policies including the Director's Fund, which provided me the opportunity to begin my research, and that respects the ideas and visions of researchers.

*—Dr. Nakajima, can you leave us with a message for researchers, students, and business partners?*

Yes, of course. Once, when I was asked to conduct an interview for an article as a member of the editorial committee of an academic journal, I received advice for young researchers from a certain NTT senior researcher who said, "Do original work different from others." At that moment, I took that to mean, "Learn to think outside the box," but today, looking back from a different perspective, I believe that advice holds a different meaning. In short, "different from others" means that there is currently no one in that field or line of research. In other words, you must take the lead in pursuing and demonstrating that possibility and getting other people involved. You will, of course, have to search out past knowledge in journal papers and elsewhere, but that has its limits. It is important that you think about many things on your own, and in any case, that you get to work on demonstrating the possibilities that you have found. Of course, jumping into the unknown is scary, and there are many things that can go wrong, but it's crucial that you have the willpower to try and try again while being optimistic like Thomas Edison who would often say, "Now I understand what won't work!" I also think that having users actually try out a constructed device is an important step in verifying its usability.

In reality, there are many hardships in such a process. However, it is exactly where no one is working on something new that there is much room for research—many worthwhile research topics are lying dormant. I feel that planting the seeds of new research and growing its branches and leaves is a blessing and an irreplaceable joy for a researcher. This kind of

undertaking, however, cannot be pursued alone. If any readers of this article are interested in my research, I would be happy to work with you in creating a new future of possibilities.

### Reference

- [1] N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso, "The Computational Limits of Deep Learning," 2020. <https://arxiv.org/abs/2007.05558>

### ■ Interviewee profile

Mitsumasa Nakajima received his M.E. and Ph.D. in Innovative and Engineered Materials from the Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology in 2010 and 2015, respectively. He entered Nippon Telegraph and Telephone Corporation (NTT) in 2010. He is engaged in the research of non-von-Neumann computing through optical operations. He received the Best Poster Paper Award, Nature Conference, Nature Publishing Group in 2019, the PN Research Award from the Institute of Electronics, Information and Communication Engineers (IEICE) Technical Committee on Photonic Network in 2017, and the IEICE Young Researcher's Award in 2013 among other awards.



## NTT's LLM "tsuzumi"

*Kentaro Shimizu, Kosuke Nishida, and Kyosuke Nishida*

### Abstract

In November 2023, NTT announced tsuzumi, a large language model (LLM) based on NTT laboratories' over-40 years of accumulated research in natural language processing. The tsuzumi LLM excels in Japanese-language processing ability, and its key features include being lightweight and providing multimodal support, which enables handling of non-text media. Its lightweight model reduces power consumption and makes on-premises use possible. Its multimodal feature enables comprehension of photos and diagrams, distinguishing tsuzumi from other publicly available LLMs. In this article, we give an overview of tsuzumi on the basis of these key features.

*Keywords: tsuzumi, lightweight, multimodal*

### 1. NTT's LLM "tsuzumi" expands the application domains of LLMs

A variety of large language models (LLMs), such as OpenAI's ChatGPT, have been released, attracting significant attention. To accumulate a greater amount of knowledge and provide natural answers, many LLMs have exploded in size. It has been reported that a model of the size of GPT-3 requires about 1300 MWh (equivalent to one hour of electricity generated by a nuclear power plant) for training [1]. Because of the prohibitive initial and operational costs, including the hardware required for large-scale training, many companies find it impractical to create their own language models. They thus turn to commercial LLMs provided as cloud services. However, these companies often handle data with personal and sensitive information, so simply uploading such data to the cloud is problematic. A company's data are considered a part of its assets. How to make effective use of such data with LLMs has thus become a challenge.

To address these issues, NTT Human Informatics Laboratories has been engaged in the research and development of tsuzumi, a lightweight LLM with superior Japanese-language processing ability, and released it to the public in November 2023. The significant advantages of tsuzumi's lightweight nature include reduction in hardware resources and power required for using and tuning the LLM. These benefits enable the use of tsuzumi by companies on their

premises, allowing them to apply and effectively use an LLM on data that would be problematic when stored on external cloud servers.

A decrease in the size of an LLM generally leads to a reduction in performance. However, NTT laboratories' over-40 years of experience in natural language processing has allowed tsuzumi to outperform publicly available LLMs with parameters of similar size in Japanese language processing. Furthermore, tsuzumi features adapter tuning to achieve flexible tuning and multimodal support, which enables comprehension of documents that include non-text content such as figures and diagrams. The key features of tsuzumi are explained in detail in the following sections.

### 2. Lightweight model

The size of an LLM is measured in terms of its parameter size. There are two versions of tsuzumi: a lightweight version with 7 billion (7B) parameters and an ultra-lightweight version with 600 million (0.6B) parameters.

The parameters of an LLM are variables used to store knowledge and skills acquired by the model during training. The greater a model's parameter size, the greater the model's capabilities tend to be in accumulating knowledge and responding to queries from humans. However, increasing parameter size also leads to an increase in the computational resources

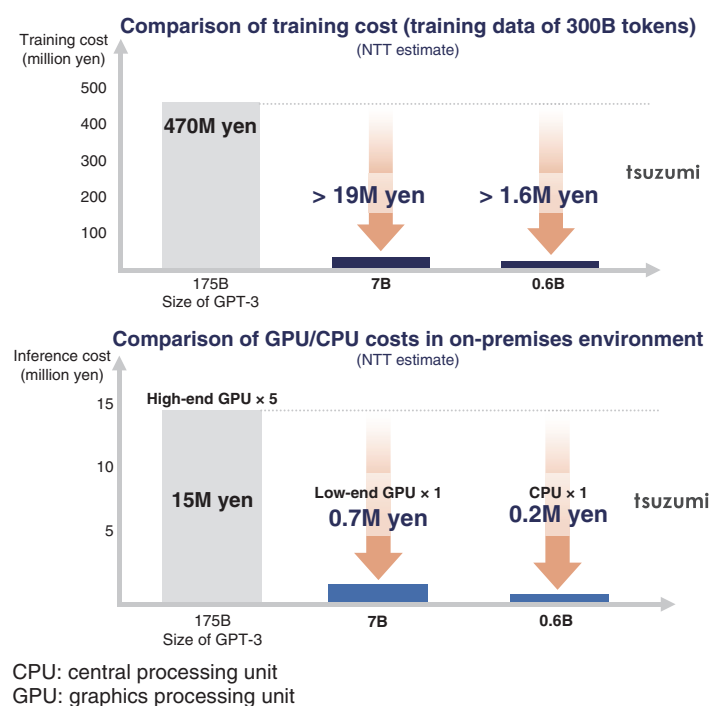


Fig. 1. Relationship between lightweight models and cost reduction.

and power consumption required for training and inference. The technological challenge is thus to reduce an LLM's parameter size while maintaining its performance.

The smaller an LLM's parameter size, the less knowledge it can generally store, leading to reduced performance. For *tsuzumi*, we have thus taken the approach of improving the quality of training data. For example, we seek to remove noise—information that should not be learned, such as redundant or erroneous information—and create training data with information from a wide range of fields instead of a few selective fields. For *tsuzumi*'s ultra-lightweight version, we succeeded in making the model even more lightweight while maintaining its performance by focusing the training on domains and tasks handled by the model.

With this approach, *tsuzumi*'s parameter size was trimmed to 1/25th to 1/300th the size of publicly available massive LLMs, resulting in significant cost reduction (Fig. 1).

### 3. Superior Japanese-language processing ability

When training an LLM, text included in the training

data is broken down (tokenized) into “tokens.” Tokens are the fundamental units processed by an LLM.

Tokenizers of recent LLMs determine the token set or vocabulary by learning from a collection of texts. LLMs developed outside Japan have a small amount of tokens for the Japanese language included in their vocabulary, and most break down Japanese text into tokens of a single character or byte. Therefore, a great number of tokens are needed to generate Japanese text, which is highly inefficient in terms of text-generation speed. We have thus developed a tokenizer with vocabulary that can efficiently generate Japanese text by training it based on Japanese text. However, the training algorithms of commonly used tokenizers do not take the structure of the Japanese language into account. Thus, there is the issue of the likelihood of redundant tokens, which often appear in the training corpus but not in other texts, being included in the vocabulary.

To address this issue, *tsuzumi* incorporates unique processing (lexical constraint) that takes Japanese words into account when tokenizing. The tokenizers of other companies may include redundant tokens, such as those that appear frequently in Wikipedia but not in other texts, in its segmentation results. In

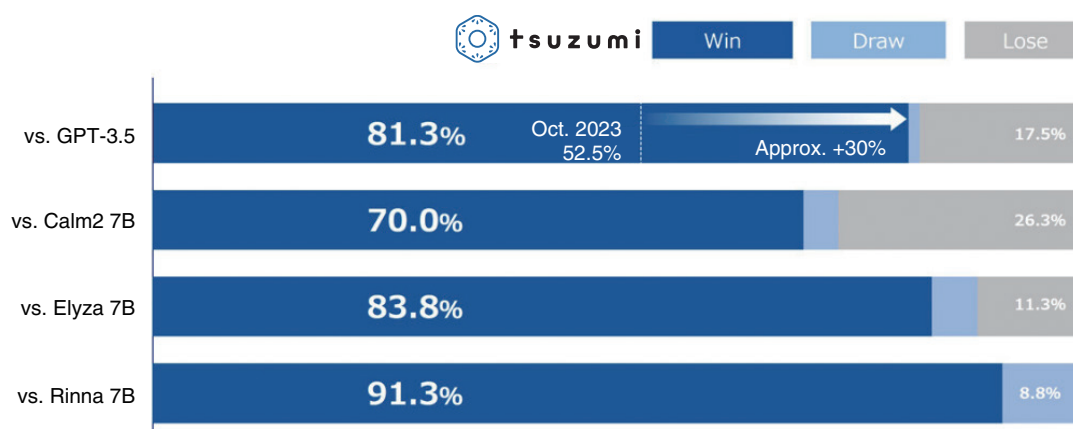


Fig. 2. Comparison of tsuzumi's Japanese-language processing ability against other commercial LLMs.

contrast, tsuzumi uses lexical constraint to achieve segmentation that strongly reflects the structure of the Japanese language. In deploying lexical constraint, NTT applied the results of its long endeavor in the research and development of morphological analysis tools and dictionaries.

Tokenization affects not only the speed of text generation but also the accuracy of text comprehension. To improve the quality of training data described above, tsuzumi's unique tokenization also results in high Japanese-language processing ability. **Figure 2** shows the results of tsuzumi's Japanese-language processing ability compared with other publicly available LLMs, as measured using the Rakuda benchmark. This benchmark consists of a series of questions testing knowledge related to Japan and is often used to compare the performances of Japanese-language LLMs. The same questions were given to LLMs for comparison, and GPT-4 was used to determine a better response of each question between two LLMs. The results indicate that tsuzumi has superior Japanese knowledge and Japanese-language processing ability, not only against LLMs of similar size but also against GPT-3.5.

#### 4. Adapter tuning

The tsuzumi LLM supports adapter tuning. This tuning adds (adapts) a small domain-specific model to the original model. **Figure 3** provides an image of adapter tuning.

Other tuning methods include prompt engineering, which provides relevant information at the same time as the query during inference, and full fine-tuning,

which updates the entire model. **Figure 4** provides a summary of each method's features. Each tuning method has its own characteristics in terms of cost and accuracy. Adapter tuning has the advantage of achieving both at a certain level. When considering use cases within different companies, the deployment of different adapters for each company makes it possible to easily create company-specific models.

We are working on providing support for multiple adapters, as shown in **Fig. 5**. This will allow the requirements of a variety of use cases to be considered by connecting multiple adapters on top of tsuzumi.

#### 5. Multimodal support

As its name implies, an LLM is a model for processing language, so it does not take graphical and audio data as its input or output. However, figures and tables are often included in general documents; such non-text content often has significant meaning. The tsuzumi LLM was designed to be a new type of LLM that comprehends figures and tables [2, 3] as well as speech and user situations (**Fig. 6**). Such use cases were exhibited at the 2023 NTT R&D Forum.

The feature article "NTT's LLM 'tsuzumi': Capable of Comprehending Graphical Documents" [4] in this issue explains tsuzumi's techniques for visual comprehension of graphical documents, one of its multimodal capabilities. Please read it in addition to another feature article "Commercialization of NTT's LLM 'tsuzumi'" [5], which describes the usage domain and usage applications, solution menu, and other features of tsuzumi commercial services

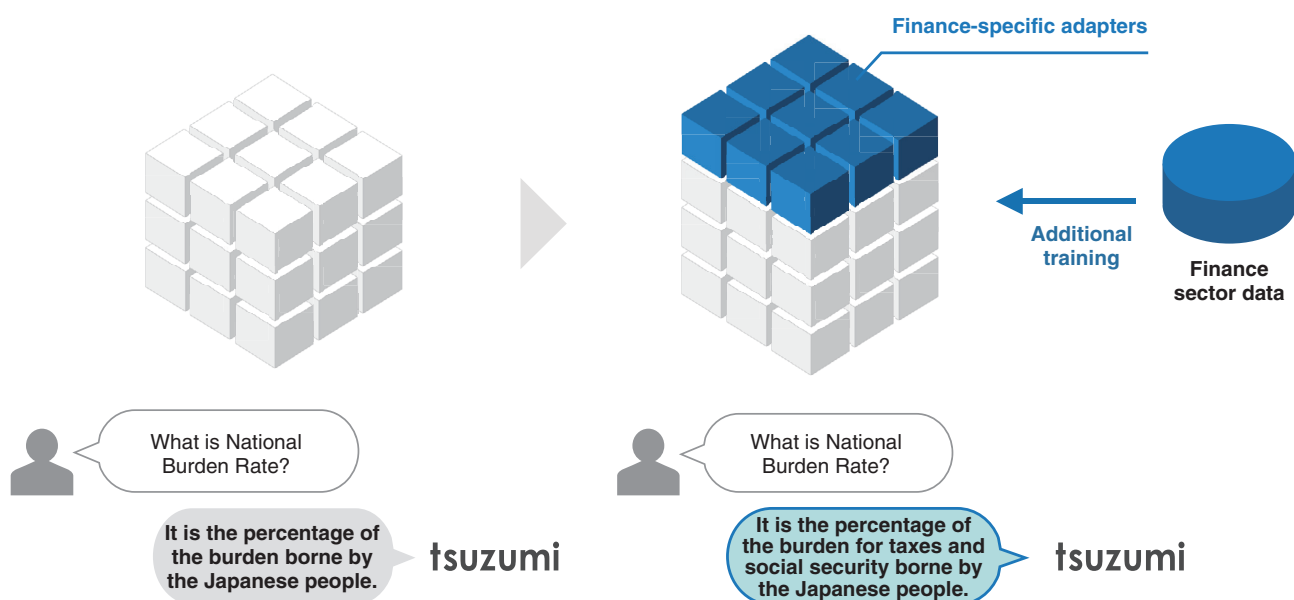


Fig. 3. Image of adapter tuning.

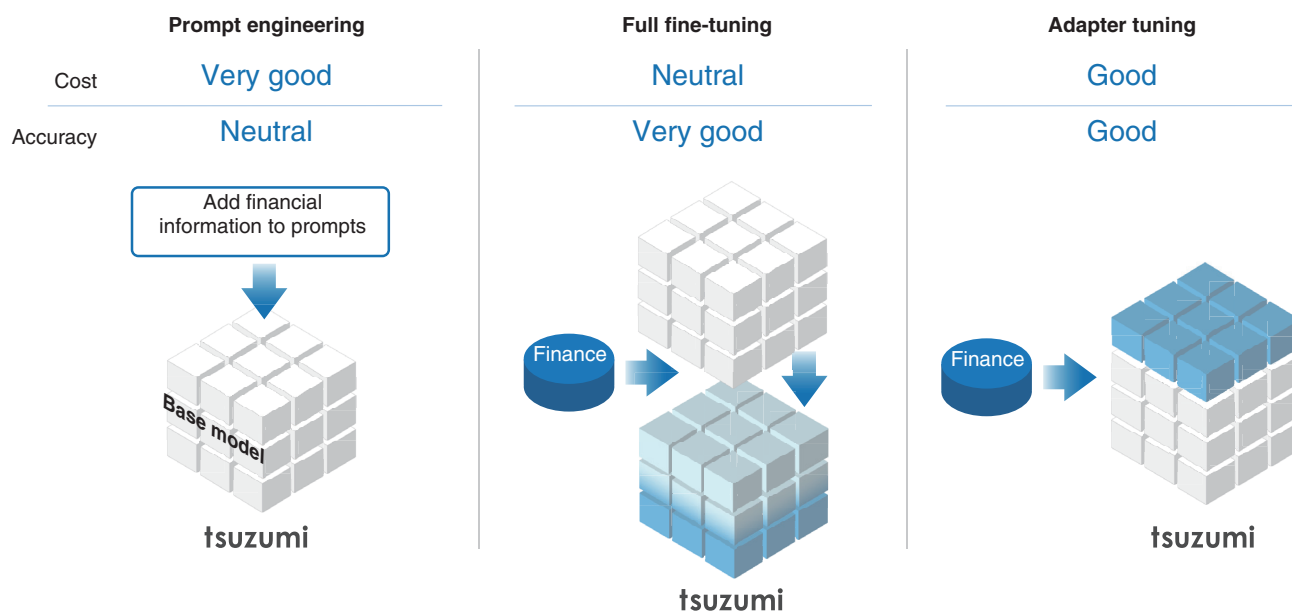


Fig. 4. Comparison of tuning methods.

launched in March 2024.

## 6. Future developments

As we have described, tsuzumi boasts key features, such as being lightweight while possessing superior

Japanese-language processing ability and flexibility in tuning, as well as multimodal support. We are continuing research and development of tsuzumi in the following areas:



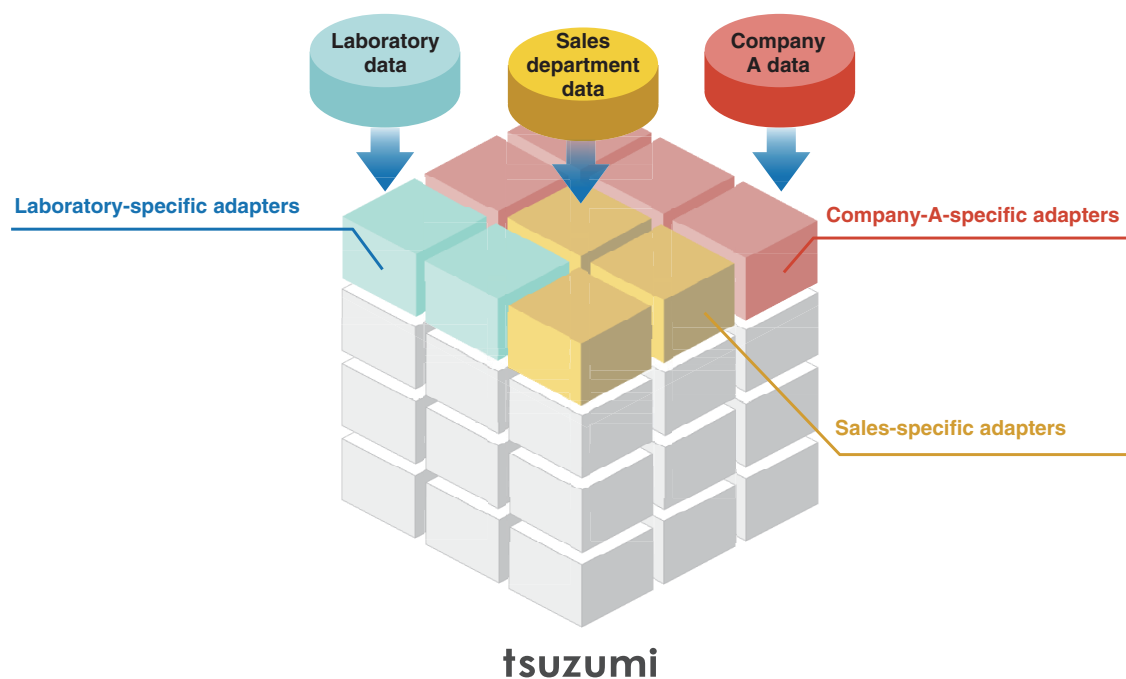
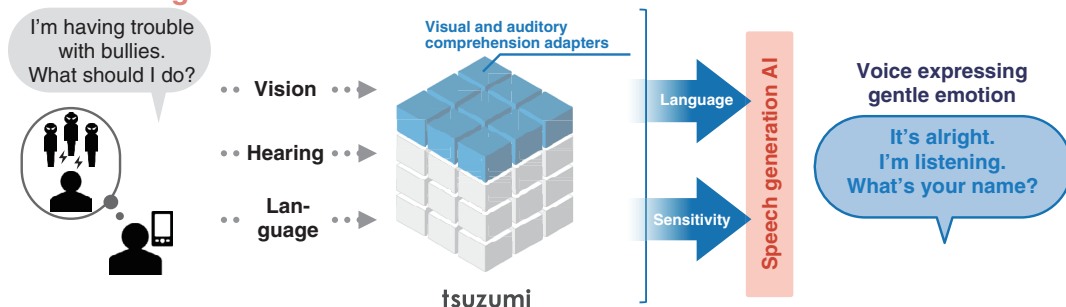


Fig. 5. Multiple adapters: Enable switching and combining adapters in accordance with the user or use case.

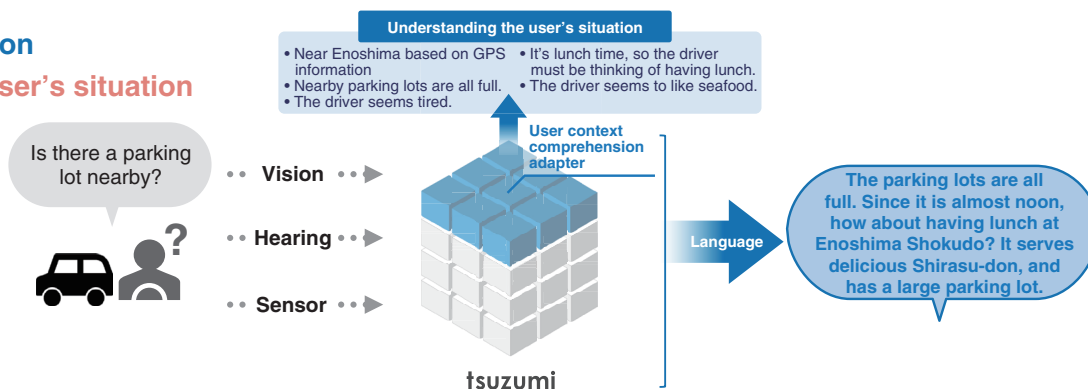
### Modal extension

#### Language + Vision + Hearing



### Modal extension

#### Language + User's situation



GPS: Global Positioning System  
AI: artificial intelligence

Fig. 6. Image of modal support.

### 6.1 Multilingual support

Currently, tsuzumi supports both English and Japanese. We seek to further improve its language processing of these two languages while also working to provide support for other languages (e.g., Chinese, Korean, French, and German) to expand its user base worldwide.

We are also working on adding support for not only human languages but also programming languages. For example, if tsuzumi can comprehend the content of a specification document and output source code in the specified programming language in accordance with the specification, it will contribute significantly to reducing software-development workload. Currently, tsuzumi can output source code to a certain extent. We are working to improve its performance in this area by collecting more training data and implementing more training.

### 6.2 Medium-sized version

In the announcement of tsuzumi's release in November 2023, it was revealed that in addition to ultra-lightweight and lightweight versions, a medium-sized version is also scheduled for release. This version will have 13B parameters, roughly double the parameter size of the lightweight version (7B parameters). In addition to multilingual support, it will be able to store a greater amount of knowledge and improve its performance as an LLM.

The medium-sized version's greater parameter size means more hardware resources are required compared with the lightweight version. However, we seek to reduce the required resources by investigating a technique (quantization) that enables LLMs to operate efficiently with limited hardware resources.

### 6.3 Support for safety and security

There is much debate on the safety and ethical

aspects of generative artificial intelligence, including LLMs. From the institutional viewpoint, lawsuits have been filed in other countries over the inclusion of copyrighted works and personal information in training data and the proper solutions. In Japan, these issues are being actively discussed, but there are still no clear conclusions. From the technological viewpoint, systems that robustly reject inappropriate questions, such as those that violate human rights, are needed.

To tackle these issues, we will not only prepare higher quality training data but also collaborate with other research institutes and organizations studying the institutional and technical aspects of these issues. At the same time, we will continue to investigate how to provide tsuzumi users with greater safety and security.

We will continue to expand the capabilities of tsuzumi. The best is yet to come.

## References

- [1] M. DeGeurin, "Move Aside, Crypto. AI Could Be the Next Climate Disaster," Gizmodo, Apr. 2023. <https://gizmodo.com/chatgpt-ai-openai-carbon-emissions-stanford-report-1850288635>
- [2] R. Tanaka, T. Iki, K. Nishida, K. Saito, and J. Suzuki, "InstructDoc: A Dataset for Zero-shot Generalization of Visual Document Understanding with Instructions," Proc. of the 38th Annual AAAI Conference on Artificial Intelligence (AAAI-24), Vancouver, Canada, Feb. 2024.
- [3] T. Hasegawa, K. Nishida, K. Maeda, and K. Saito, "DueT: Image-Text Contrastive Transfer Learning with Dual-adaptor Tuning," Proc. of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023), pp. 13607–13624, Singapore, Dec. 2023. <https://doi.org/10.18653/v1/2023.emnlp-main.839>
- [4] R. Tanaka, T. Iki, T. Hasegawa, and K. Nishida, "NTT's LLM 'tsuzumi': Capable of Comprehending Graphical Documents," NTT Technical Review, Vol. 22, No. 8, pp. 26–30, Aug. 2024. <https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202408fa2.html>
- [5] N. Shibuya, I. Kudo, M. Kaido, and S. Yamamoto, "Commercialization of NTT's LLM 'tsuzumi,'" NTT Technical Review, Vol. 22, No. 8, pp. 31–37, Aug. 2024. <https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202408fa3.html>



#### Kentaro Shimizu

Senior Research Engineer, Human Insight Laboratory, NTT Human Informatics Laboratories.

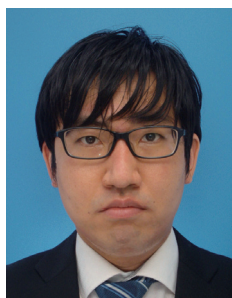
He received a B.E. and M.E. in engineering from Yokohama National University in 1997 and 1999 and Ph.D. in medicine from Tokyo Women's Medical University in 2009. He joined NTT Cyber Space Laboratories in 1999 and studied image processing. His current interests are LLMs and machine learning. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE), Information Processing Society of Japan (IPSJ), the Japanese Society for Artificial Intelligence (JSAI), and the Japan Society of Computer Aided Surgery (JSCAS).



#### Kyosuke Nishida

Senior Distinguished Researcher, Human Insight Laboratory, NTT Human Informatics Laboratories.

He received a B.E., M.I.S., and Ph.D. in information science and technology from Hokkaido University in 2004, 2006, and 2008. He joined NTT in 2009. His current interests include natural language processing, vision-and-language, and artificial intelligence. He is currently serving as the technical lead for NTT's LLMs. He received the Paper Award from the Association of Natural Language Processing (NLP) from 2018 to 2024. He is a member of the Association for Computing Machinery (ACM), NLP, IPSJ, and the Database Society of Japan (DBSJ).



#### Kosuke Nishida

Associate Distinguished Researcher, Human Insight Laboratory, NTT Human Informatics Laboratories.

He received a B.E. in engineering and M.E. and Ph.D. in information science and engineering from the University of Tokyo in 2015, 2017, and 2024. He joined NTT Media Intelligence Laboratories in 2017 and studied natural language processing. His current research interests are especially mathematical understandings of models and training algorithms for natural language processing. He is a member of the Association for Computational Linguistics (ACL).

## NTT's LLM “tsuzumi”: Capable of Comprehending Graphical Documents

*Ryota Tanaka, Taichi Iki, Taku Hasegawa, and Kyosuke Nishida*

### Abstract

Large language models (LLMs) are being applied to fields such as healthcare, customer support, and office digital transformation. Information handled in such fields includes not only text but also a variety of visual content such as figures and diagrams. To develop LLMs as the core of artificial intelligence, their capabilities must be expanded so they can comprehend visual information. NTT Human Informatics Laboratories has been researching and developing NTT's LLM called “tsuzumi.” In this article, we discuss our efforts related to tsuzumi's visual machine reading comprehension technology for comprehending the content of a document from visual information.

*Keywords: tsuzumi, LLM, visual machine reading comprehension*

### 1. Visual machine reading comprehension technology for visual document understanding

Besides text, the documents we handle every day include visual elements (e.g., icons, figures, and diagrams). Such information exists in a variety of forms and layouts. Developing technology to read and comprehend real-world documents is one of the most urgent challenges in the field of artificial intelligence (AI). Many AI models, including large language models (LLMs), which have the ability to achieve general-purpose language understanding and generation, have been developed. Although AI capabilities have expanded vastly, surpassing, for example, humans' reading ability, AI models still face the limitation of being able to understand only textual information in documents. To address this issue, NTT has proposed visual machine reading comprehension technology (**Fig. 1**). This technology enables comprehension of documents from visual information in a manner similar to how humans understand information.

We constructed datasets, such as VisualMRC [1]

and SlideVQA [2], to make this technology possible. These datasets contain question-answering pairs on single and multiple document images, such as screenshots of webpages and presentation materials. Comprehending document images requires comprehension of not only linguistic information but also visual information such as the size and color of characters, figures and diagrams, graphs, and layout. We proposed LayoutT5 [1], a visual machine reading comprehension model that integrates two sets of inputs. It first applies object-recognition technology to extract regions in a document (titles, paragraphs, images, captions, lists, etc.) and applies text-recognition technology to extract the position and appearance information of text as additional input. We also proposed M3D [2], which understands the relationships between multiple document images. These models, which take into account visual information, perform better than models that only handle text, confirming the effectiveness of this technology inspired by human information processing.

Using the knowledge we obtained from constructing the datasets and models, we participated in the



### Text-based machine reading (conventional approach)

Cannot read visual information such as figures, tables, graphs, text appearance, and layout.

Text extracted from the document

Light Plan usage fee Monthly fee (tax included) 5,800 yen/month Monthly usage fee Basic fee (tax included) 2,500 yen/month The usage fee increases by the amount you use 2 levels Flat rate plan 200MB 1,000MB 200MB Standard of 1,000MB Website browsing and email are mainly used once every 2 days for about 30 minutes Website browsing and email, plus video browsing once every 2 days for about 1 hour Internet usage

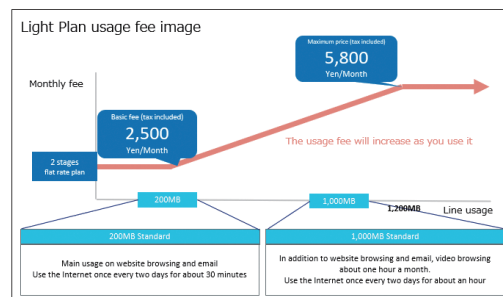
Q: If my monthly usage is 2000MB, what is the basic monthly fee for the Light Plan?

A: ????

### Visual machine reading

Can understand documents based on the visual information and handle a variety of document formats.

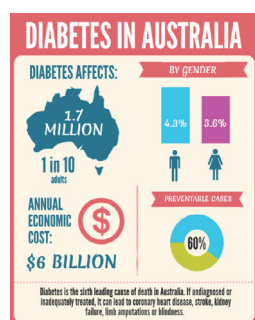
e.g., Documents in HTML or PDF format



Q: If my monthly usage is 2000MB, what is the basic monthly fee for the Light Plan?

A: 5800 yen

Fig. 1. Overview of visual machine reading comprehension technology.



Q1: How many females are affected by diabetes?

A: 3.6%



Q2: What total percent of B2B and B2C markets use Google+?

A: 79% (40% + 39%)

Fig. 2. Example of questions and answers in InfographicVQA.

Document Visual Question Answering (DocVQA) competition at the International Conference on Document Analysis and Recognition in 2021 (ICDAR 2021). This challenge tests a model's ability to answer questions presented in infographics containing visual representations of information, data, and knowledge. Examples of questions given at this competition and their answers are shown in **Fig. 2**. To answer Q1 in the figure, the model must understand that the icon shown in the center right of the document represents women. To answer Q2, the model must be able to extract the numerical values from the document and calculate " $40\% + 39\% = 79\%$ ".

Addressing these challenging questions requires a wide range of capabilities: understanding both textual content and visual information (e.g., icons, figures, and diagrams), comprehending the spatial relationships between text and visual elements, and executing arithmetic operations. For this competition, we thus proposed a model for answering infographic questions called IG-BERT [3]. We introduced a new method for learning layout relationships between text and visual objects in document images and a data-augmentation method for generating reasoning processes. IG-BERT achieved the highest performance among models of similar size while curbing the amount of pre-training data needed to 1/22 that of conventional models. It won second place out of 337 models submitted by 18 teams.

## 2. Issues with conventional visual machine reading comprehension technology

Conventional techniques in visual machine reading comprehension struggled with handling diverse tasks, such as extracting information from invoices. Typically, achieving high performance on specific tasks required extensive training on task-specific datasets, resulting in high data creation and training costs. This approach created barriers to developing models that could effectively meet user needs across different applications. We thus sought to develop a visual machine reading comprehension model that is effective in following instructions by using an LLM,

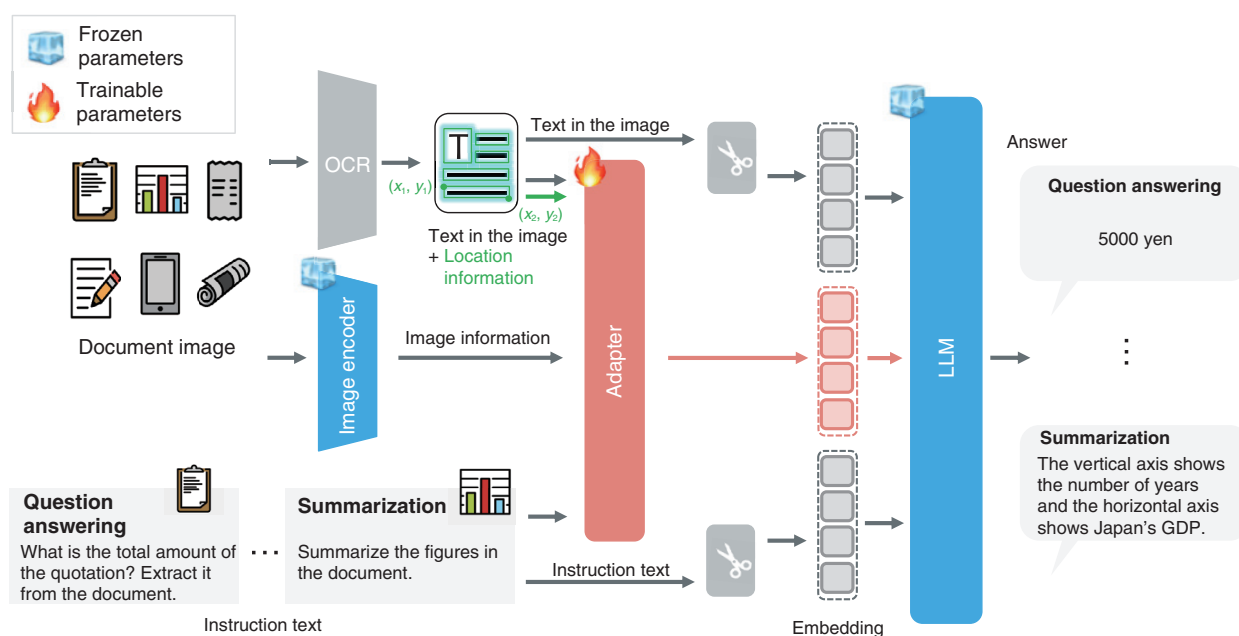


Fig. 3. Overview of LLM-based visual machine reading comprehension technology.

which is endowed with the general-purpose ability to understand and generate language. Our model would be able to respond to questions without training on each comprehension task. Specifically, our developmental challenge was to determine how to integrate visual information such as figures and diagrams contained in document images with text so the data can be understood by the LLM, which can only understand text information, without degrading the LLM's reasoning capability.

### 3. LLM-based visual machine reading comprehension technology

To enable an LLM to comprehend visual information such as figures and diagrams included in document images, visual information represented as a collection of pixels must be converted into a form that the LLM can easily process. Regarding visual document understanding of NTT's LLM "tsuzumi," we achieved the addition of visual understanding while maintaining text comprehension skills by combining an image encoder and lightweight adapter, as shown in Fig. 3 [4]. The image encoder maps a set of pixels to textual meaning. The adapter transforms the meaning so it can be processed by tsuzumi.

#### 3.1 Image encoder capable of understanding Japanese images characteristics

An image encoder processes the visual information of what appears in an image. We prepared an image encoder that converts images to vectors as well as a text encoder that converts text sequences to vectors. We train the image encoder so that the distance between an image vector and vector of the text representing the content of the image is close, and the distance is far when the image and text have no relationship. Therefore, visual information obtained by the image encoder can be connected to textual information. To train the image encoder, we constructed a dataset of several hundred million pairs of text and images. We collected not only general images and English captions but also images particular to Japan and their Japanese captions. The Japan-specific images contain, for example, Japanese writing and landscapes found only in Japan. By purposefully including Japanese captions during training, we enable the model to learn expressions particular to the Japanese language, such as "*aoi shingo* (blue traffic light)" and "*makkana taiyo* (bright red sun)." We are also engaged in constructing models that robustly support both English and Japanese by developing techniques that allow encoders trained on English text and images to adapt to the Japanese language [5].

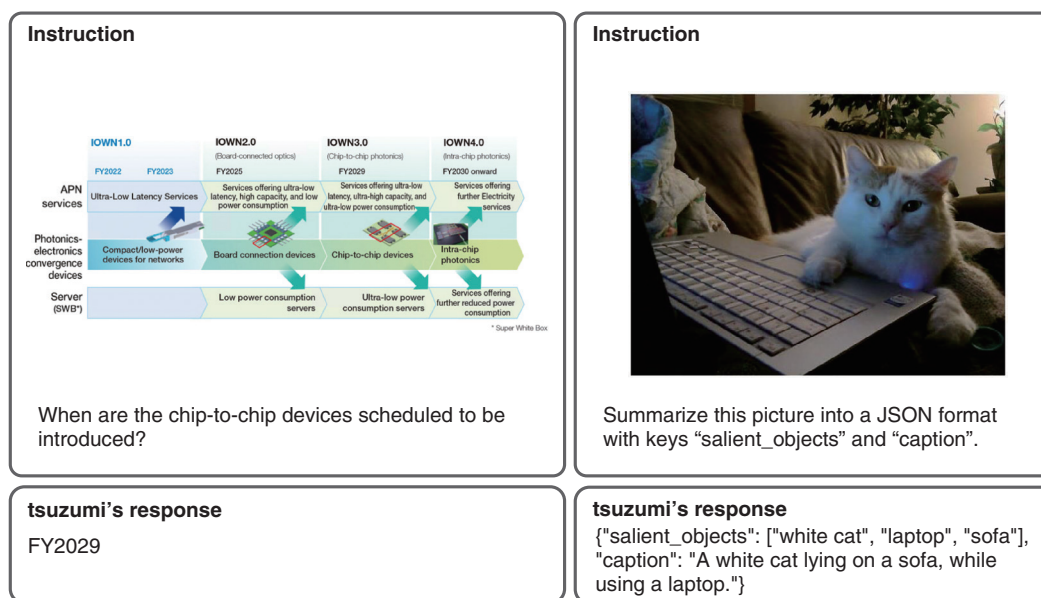


Fig. 4. Example of visual machine reading comprehension by tsuzumi.

### 3.2 Training adapter for understanding visual documents

An adapter has the role of translating a document representation into the “words” understood by the LLM. An LLM splits a text sequence into several strings called tokens. To input tokens into the neural network, they are converted to vectors. The vector corresponding to the token, called an embedding, is the “word” received by the LLM. The adapter communicates the image to the LLM by converting the output of the image encoder to an embedding.

Because the adapter is a neural network with a small number of parameters, training is necessary. In tsuzumi’s visual machine reading comprehension, we maintain tsuzumi’s reasoning capability by fixing its parameters and those of the image encoder and training only the parameters of the adapter. An adapter suited for images is also achieved by carrying out multi-stage training. Initially, large-scale image-caption pairs are used to train tsuzumi to recognize general visual concepts such as objects, scenery, and locations. Subsequently, tsuzumi is further trained using optical character recognition (OCR) from image and textual instructions as inputs, with answers as outputs. To this end, we created the dataset InstructDoc, which is a large-scale collection of 30 publicly available visual document understanding datasets, each with diverse instructions in a unified format [4]. For example, it includes tasks such as

information extraction, wherein given a question like “Tell the total amount” and a invoice image, the system provides answers such as “5000 yen.” We experimentally verified that an LLM trained on InstructDoc can achieve a high rate of success for unseen tasks (tasks not included in the training data).

### 3.3 Output examples

An example of tsuzumi’s visual machine reading comprehension is shown in **Fig. 4**. The left side shows a question-answering task on the figure. The figure depicts the Innovative Optical and Wireless Network (IOWN) roadmap and the question “When are the chip-to-chip devices scheduled to be introduced?” is posed. Our model responds with “FY2029,” the correct answer. This demonstrates that it is capable of reading the visual structure of the diagram, in which the roadmap is divided into columns showing fiscal years. Because the training dataset contains images of various figures and diagrams, our model understands standard visual layouts, thus could answer the question. The right side of the figure shows a photo-recognition task. A photo of a cat is shown, and the instruction “Summarize this picture into a JSON format with the keys ‘salient\_objects’ and ‘caption’” is given. JSON is a standard text-based format for representing structured data. Our model responded with “{“salient\_objects”: [“white cat”, “laptop”, “sofa”], “caption”: “A white cat lying

on a sofa, while using a laptop."}. The model could not only output text in JSON format with the given keys, it could also extract content from the image that matches the meaning of each key. Controlling output format based on text-image understanding can be used for diverse applications such as image tagging. Therefore, the flexibility of tsuzumi's visual machine reading comprehension connects text and image understanding to accomplish tasks that meet the needs of users.

#### 4. Future goals

We seek to further expand the capabilities of the current document comprehension model. We will also expand the application range of tsuzumi by connecting it with other modalities besides the visual modality, with the goal of advancing research and development and commercialization to ultimately achieve a society where humans and AI coexist.

#### References

- [1] R. Tanaka, K. Nishida, and S. Yoshida, "VisualMRC: Machine Reading Comprehension on Document Images," Proc. of the 35th Annual AAAI Conference on Artificial Intelligence (AAAI 2021), pp. 13878–13888, Feb. 2021.
- [2] R. Tanaka, K. Nishida, K. Nishida, T. Hasegawa, I. Saito, and K. Saito, "SlideVQA: A Dataset for Document Visual Question Answering on Multiple Images," Proc. of the 37th Annual AAAI Conference on Artificial Intelligence (AAAI 2023), pp. 13636–13645, Washington, D.C., USA, Feb. 2023.
- [3] R. Tanaka, K. Nishida, J. Xu, and S. Nishioka, "Infographic Question Answering Based on Integrated Understanding of Text and Visual Information," Proc. of the 28th Annual Meeting of the Association for Natural Language Processing (NLP 2022), pp. 52–57, Mar. 2022 (in Japanese).
- [4] R. Tanaka, T. Iki, K. Nishida, K. Saito, and J. Suzuki, "InstructDoc: A Dataset for Zero-shot Generalization of Visual Document Understanding with Instructions," Proc. of the 38th Annual AAAI Conference on Artificial Intelligence (AAAI 2024), pp. 19071–19079, Vancouver, Canada, Feb. 2024.
- [5] T. Hasegawa, K. Nishida, K. Maeda, and K. Saito, "DueT: Image-Text Contrastive Transfer Learning with Dual-adaptor Tuning," Proc. of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023), pp. 13607–13624, Singapore, Dec. 2023.



**Ryota Tanaka**

Researcher, Human Insight Laboratory, NTT Human Informatics Laboratories.

He received a B.E. and M.E. in computer science and engineering from Nagoya Institute of Technology, Aichi, in 2018 and 2020. He has also been pursuing a Ph.D. at Tohoku University, Miyagi, since 2023. He joined NTT Media Intelligence Laboratories in 2020. His research topics lie in multimodal machine learning, in particular, visual document understanding. He received the Best Paper Award at the 27th Annual Meeting of the Association of Natural Language Processing (NLP2021) and the Outstanding Paper Award at NLP2023 and 2024. He was a runner-up at the ICDAR 2021 DocumentVQA competition. He is a member of NLP.



**Taichi Iki**

Research Scientist, Human Insight Laboratory, NTT Human Informatics Laboratories.

He received a B.S. and M.S. in physics from the University of Tokyo in 2013 and 2015. After working as an AI engineer in research and development of natural language processing and dialogue systems, he entered a doctoral program. He received a Ph.D. in informatics from the Graduate University for Advanced Studies, SOKEND-AI, Kanagawa, in 2022. He joined NTT in 2022. His research interests include vision-language foundation models and AI. He is a member of NLP and the Japanese Society for Artificial Intelligence (JSAI).



**Taku Hasegawa**

Research Scientist, Human Insight Laboratory, NTT Human Informatics Laboratories.

He skipped undergraduate studies and received an M.S. and Ph.D. in computer science and intelligent systems from Osaka Prefecture University in 2016 and 2018. He joined NTT in 2019. His research interests include vision-language foundation models and AI. He received the Paper Award from the Japanese EC Society in 2017, the Paper Award from JSAI in 2020, and an award from NLP in 2023. He is a member of NLP and JSAI.



**Kyosuke Nishida**

Senior Distinguished Researcher, Human Insight Laboratory, NTT Human Informatics Laboratories.

He received a B.E., M.I.S., and Ph.D. in information science and technology from Hokkaido University in 2004, 2006, and 2008. He joined NTT in 2009. His current interests include natural language processing, vision-and-language, and AI. He is currently serving as the technical lead for NTT's LLMs. He received the Paper Award from NLP from 2018 to 2024. He is a member of the Association for Computing Machinery (ACM), NLP, Information Processing Society of Japan (IPSJ), and the Database Society of Japan (DBSJ).



## Commercialization of NTT's LLM "tsuzumi"

*Naoyuki Shibuya, Ichiro Kudo, Masahiro Kaido, and Shunichiro Yamamoto*

### Abstract

NTT Human Informatics Laboratories has been researching and developing NTT's large language models called "tsuzumi," and on March 25, 2024, NTT held a press conference to announce the launch of tsuzumi commercial services. This article introduces the solution menu, partner program, members forums, and other features of tsuzumi as presented at the press conference.

*Keywords: LLM, tsuzumi, commercial use*

### 1. Introduction

NTT's large language models (LLM) "tsuzumi," which is now being researched and developed at NTT Human Informatics Laboratories, has been receiving good reviews since its announcement in November 2023 with more than 500 inquiries being received from a variety of companies and local governments. A press conference was later held on March 25, 2024 announcing the launch of commercial services. This article introduces the tsuzumi usage domain and usage applications, solution menu, partner program, members forums, and other features of tsuzumi as presented at the press conference.

### 2. Usage domain and usage applications

Among the inquiries that have been received about tsuzumi from a variety of companies and local governments, two-thirds have expressed an interest in training an LLM on in-house data and using a customized LLM. In other words, they would like to construct an LLM on their own premises (on-premise development) so that it can be trained in a closed and secure environment. Inquiries on introducing an LLM have also been received from a wide range of industries, many of which are from sectors such as manufacturing, local governments, and finance that handle highly confidential data.

We expect tsuzumi to be used in various types of applications. These include improving the customer experience (CX) at customer contact points, such as call centers, and improving the employee experience (EX) by streamlining in-house tasks such as automatically creating/summarizing the minutes of a meeting and creating a question and answer (Q&A) list from a business manual, as well as automating information technology (IT) operations and developing software.

### 3. Solution menu

Against the above background, tsuzumi commercial services that combine three usage environments and three types of solutions were announced on March 25, 2024.

#### 3.1 Usage environments

Since tsuzumi is a lightweight LLM model, it can be used on premise at the customer's office without having to construct a large-scale hardware environment or via a private cloud operated by an NTT Group datacenter. A customer may also select a public cloud as desired tailored to the way that tsuzumi is to be used. Using tsuzumi in an on-premise environment makes it possible to train the LLM safely on in-house data without having to send that data outside the company (Fig. 1).

# Usage environments

## Can be customized using in-house data without sending the data outside



Can be used in an on-premise environment or on an NTT Group private cloud, enabling safe handling of confidential or sensitive information

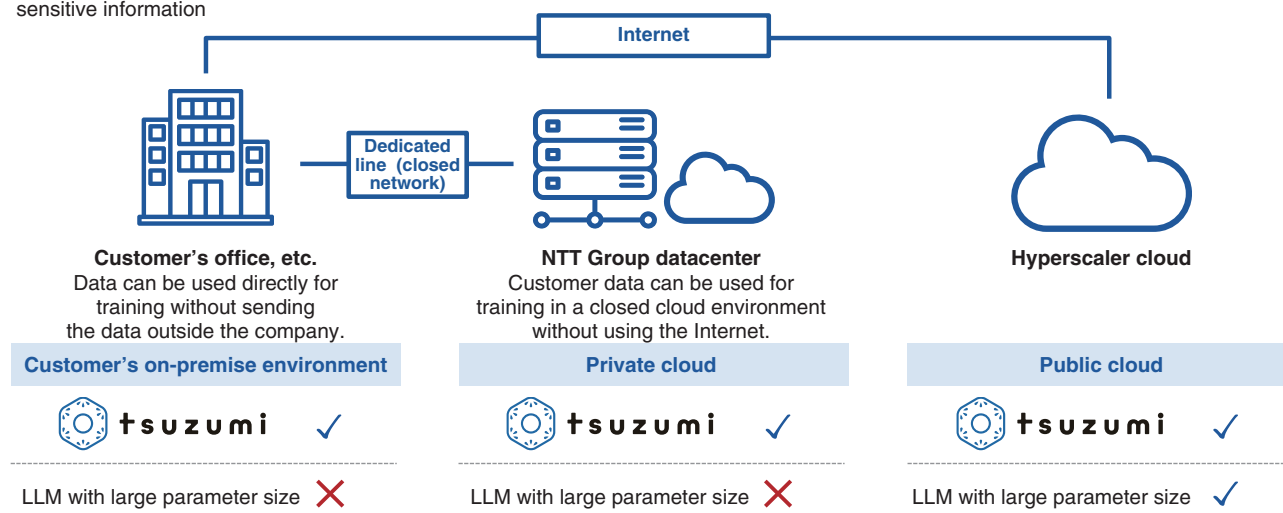


Fig. 1. Usage environments.

### 3.2 CX solutions

CX solutions aim to improve CX at all types of customer contact points such as contact centers, brick and mortar stores, and e-commerce sites. We introduce two specific CX solutions.

The first is operator support for contact centers (Fig. 2). After a customer has been helped at a contact center, there is a need for after-call work such as summarizing the help given and registering necessary information in the system. This CX solution makes it possible to automate the summarizing of a completed call, entering of knowledge for a frequently asked question (FAQ) page, etc., thus reducing the time taken up by after-call work. It can also help improve the quality of customer service by finding candidate answers for a variety of customer inquiries from knowledge data stored by tsuzumi and giving the operator advice in real time, etc.

Another CX solution is virtual concierge (Fig. 3). A virtual concierge is a digital human that provides customer service. A digital human expresses human-like behavior that combines the appearances of real persons or an appearance based on a model tailored to the usage scenario and by applying motion generative artificial intelligence (GenAI). Combining technolo-

gies such as advanced automated replies, speech synthesis, and facial-expression analysis through tsuzumi in this way can provide personalized attention during a consultation with a customer while observing the customer's information, facial expressions, etc. Thus, a virtual concierge shows promise for use in customer service at a variety of company counters and online contact points.

### 3.3 EX solutions

EX solutions are for automating and increasing the efficiency of common analog tasks in a variety of industries. These tasks include creating and summarizing the minutes of a meeting, searching through business manuals, creating a Q&A list, searching for products, supporting the preparation of reports, supporting the creation of application forms, checking internal rules and regulations, and conducting a legal check.

EX solutions provide employee support on the basis of industry knowledge and the customer's proprietary business knowledge by providing an employee-friendly user interface combined with a specialized LLM trained on industry knowledge and in-house documents. They help improve job satisfaction,

## Support for operator tasks at contact centers

Reduce customer-call wait time and improve the quality of customer service by executing real-time searches, automating summaries, and extracting information

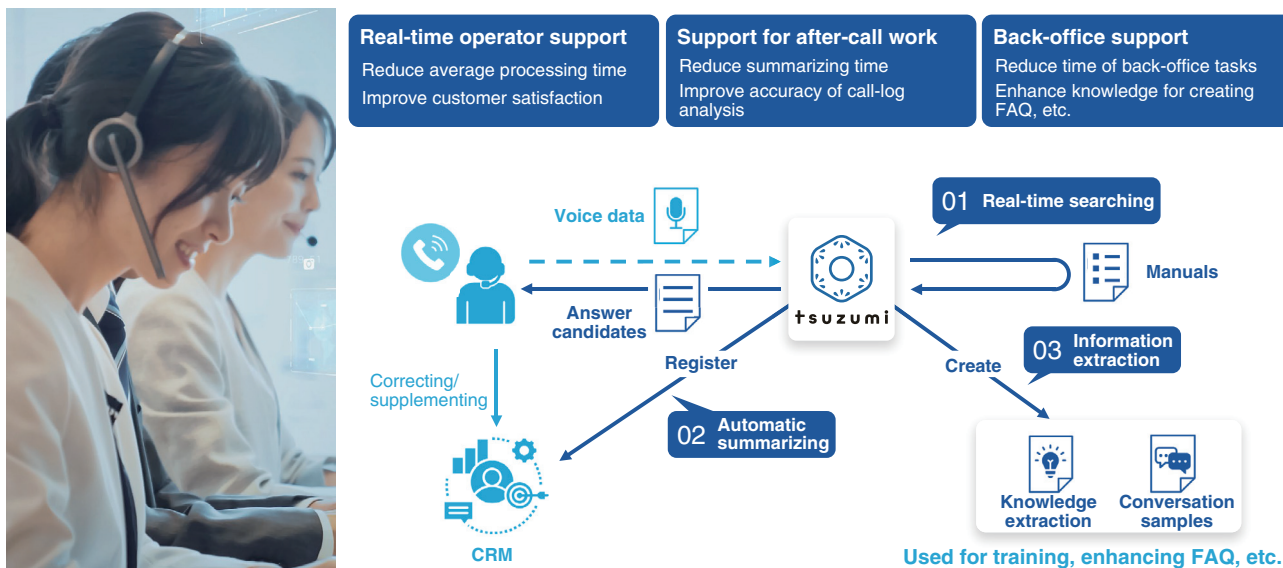


Fig. 2. Operator support at contact centers.

## Virtual concierge

Achieve new customer contact points by using digital humans created with proprietary technologies × GenAI

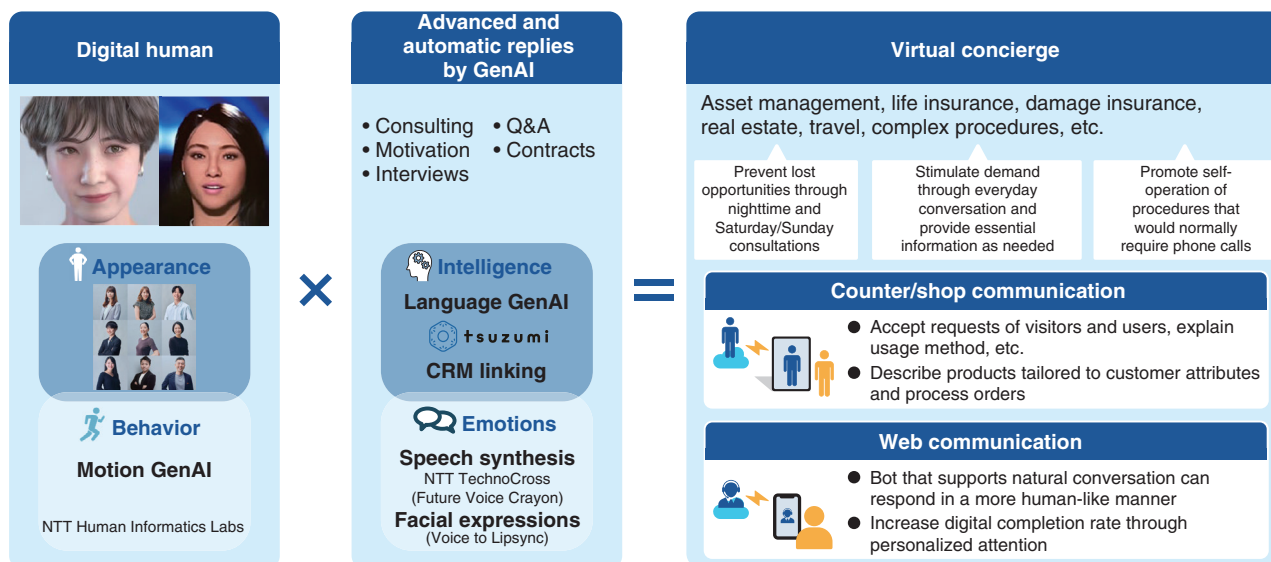


Fig. 3. Virtual concierge.

## IT-operations support solutions (security response)

### Support technologies that require highly skilled operations

- Reduce the load of a company's security operations through automation and simplification by using an LLM
- Resolve the contradiction between the long-standing problem of attacks by malware, etc. and shortage of security personnel

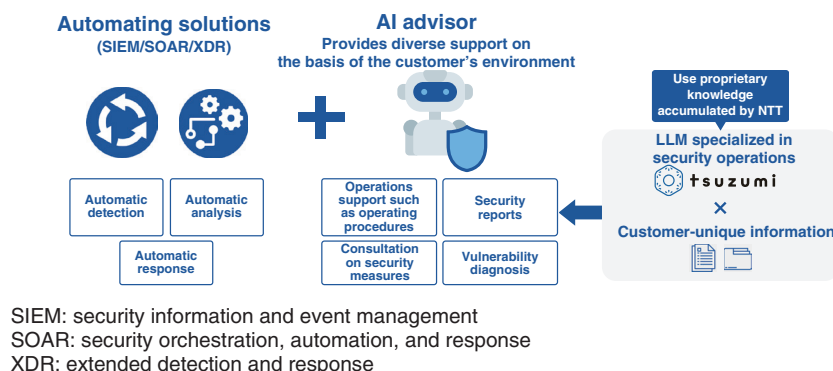


Fig. 4. Example of security response.

productivity, and EX at the target client. For example, an EX solution for local governments would enable personnel at an administrative counter to easily search for answers even for a wide range of inquiries without making residents wait. This would be accomplished by using a specialized LLM trained on government tasks in conjunction with the business processes carried out at that administrative counter.

Since tsuzumi can be directly constructed within a network storing highly confidential data inside a local government, it can be trained securely on diverse types of data unique to government services without having to send that data outside while generating answers tailored to the processes of an administrative counter.

### 3.4 IT-operations support solutions

To enable a variety of work styles, such as remote work and “workcations,” in a secure manner, the volume of work in IT operations, in-house information-system departments, and at help desks has recently been increasing. At the same time, the work is becoming more advanced and complicated, and the shortage of human resources is becoming increasingly serious.

IT-operations support solutions support the streamlining and automation of IT operations within a company and the strengthening of business continuity combined with the unique expertise of the NTT Group in providing IT services. They can be used, for

example, in security measures. The threats of cyber attacks, such as malware, are increasing yearly and the attacks are becoming increasingly sophisticated while the number of security personnel that can respond to these threats is decreasing.

The NTT Group has had a number of achievements in the field of security such as defending against cyber attacks directed at international events. In fact, by using these solutions that detect security threats and respond automatically using NTT proprietary knowledge, 95% of these achievements involved cases of automatically responding without human intervention by automatically detecting access-log records that suggest cyber attacks. The remaining 5% involved access-log records of new, previously unknown cyber attacks. These records had to be analyzed and dealt with from various angles by personnel specialized in security, which took time to carry out.

By having tsuzumi learn the cyber-defense expertise of the NTT Group and the customer's unique network environment then using it as an LLM specialized in security operations, tsuzumi can give advice on methods for dealing with new attacks that had been escaping from the filters of existing AI automated solutions. These IT-operations support solutions can both mitigate risks and reduce the labor of specialized personnel by reducing the range of the role played by people in security measures (Fig. 4).

## tsuzumi Launch of partner program

Provide some tsuzumi APIs at no charge and promote the use of tsuzumi in partner services, development of specific models, and integration (recruitment of partners began in May 2024)

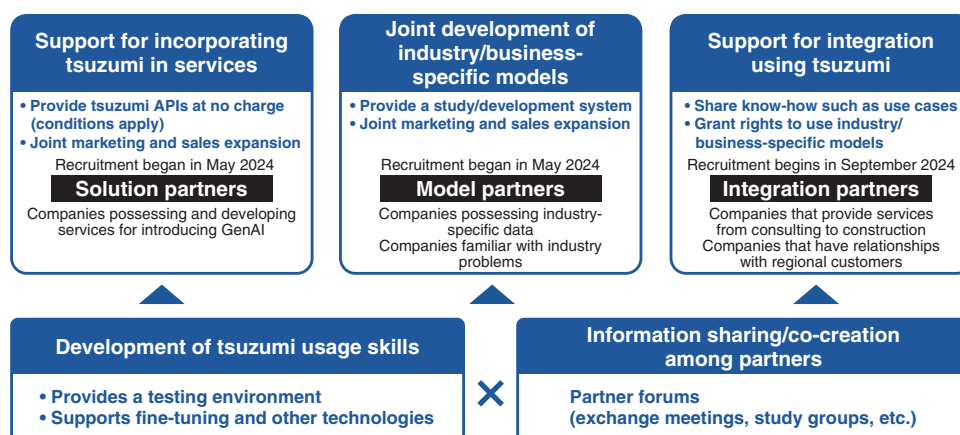


Fig. 5. Partner program.

### 3.5 Total support for introducing tsuzumi

For a company introducing tsuzumi solutions, the NTT Group can provide total support for a series of processes from initial studies to development, installation, and operation.

With GenAI, the desired response accuracy differs in accordance with the training data used, number of tunings executed, and target use case. The NTT Group supports agile development in accordance with the circumstances by revising training data and varying the tuning method while testing. It also supports customized maintenance toward further improvements in CX and EX after tsuzumi has been installed and put into use.

Before customized development of an LLM targeted for installation, an AI risk evaluation can be conducted using expertise from the AI governance initiative undertaken by the NTT Group from 2019. Customers can therefore use GenAI including tsuzumi in a safe and secure manner.

## 4. Partner program

Recruitment of three types of partners for different objectives began in May 2024 (Fig. 5).

The first type is a solution partner targeting companies having application services. This type of partner will be able to use some of tsuzumi's application programming interfaces (APIs) free of charge to

make it easy to study how tsuzumi can be incorporated behind the scenes in solution services. By incorporating tsuzumi functions in existing customer relationship management (CRM) services and business application services provided by a partner company, and by enabling startup companies that are studying new services using GenAI to try out tsuzumi, we hope to increase the added value of new services and existing solution services.

The second type is a model partner targeting companies having data specific to certain industries. The aim is to train tsuzumi on industry/business-specific data, promote the joint development of specific models for solving industrial problems, and enable joint marketing and expansion of sales.

The third type is an integration partner targeting companies such as system integrators engaged in system integration from consulting to construction to improve the digital transformation (DX) and CX of their clients. By becoming an integration partner, a company can obtain the right to use "industry/business-specific models" developed, for example, by a model partner (second type of partner), and by sharing expertise, such as use cases, partner companies can jointly deliver solutions using tsuzumi to even more customers.

By establishing partner forums, we will create opportunities for these three types of partners to develop skills in using tsuzumi and share knowledge



## tsuzumi Members forums

Members forums for customers using AI including tsuzumi will be held

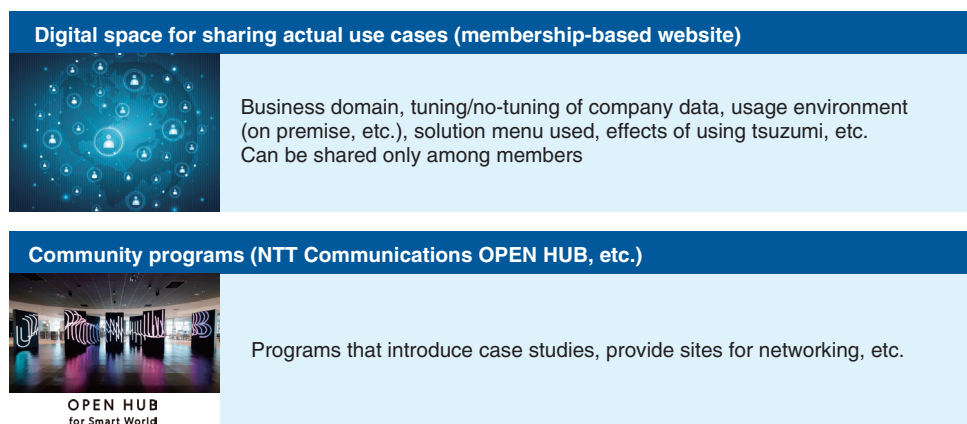


Fig. 6. Members forums.

with each other with the aim of creating new value together. The development of tsuzumi usage skills will not be limited to urban environments—we will also develop programs for supporting re-skilling and DX in non-urban regions by matching urban resources and non-urban work. DOCOMO gacco, Inc., which provides one of the largest online video-training services in Japan, plans to launch a program promoting improvement in DX skills including the use of tsuzumi in the summer of 2024 while providing specific support for the Noto Peninsula, which was struck by a strong earthquake in January 2024. In short, by participating in this program, participants can become directly involved in supporting the Noto Peninsula while developing tsuzumi usage skills.

### 5. Members forums

Members forums targeting customers who are using tsuzumi will also be held (Fig. 6). They will take place in both digital space and real space and enable customers who are using tsuzumi to share use cases and knowledge and ask questions such as “What kind of business are you using tsuzumi for?” “What kind of company data are you training tsuzumi

on?” and “What kind of effects have been obtained?” Targeting customers who are using tsuzumi, forums in digital space will enable use cases and other types of information to be shared via a membership website that only registered members can enter. There are also plans for programs to be held at NTT Group co-creation spaces where case studies can be presented to other members in person and members can network with each other.

### 6. Global expansion

The tsuzumi LLM is already attracting interest not only from customers in Japan but also from global customers. To promote the use of GenAI from Japan on a global basis, we are preparing a GenAI global platform on which customers will be able to use tsuzumi safely and securely. Specifically, we are incorporating tsuzumi in solutions that are being provided globally by NTT DATA and providing these solutions to customers via datacenters that are being rolled out globally. Our aim is to provide a multilingual environment in which global customers in a variety of industries can use tsuzumi in a safe and secure manner.



**Naoyuki Shibuya**

Director, Global Marketing Strategy Group, Market Planning and Analysis Department, Research and Development Market Strategy Division, NTT Corporation.

He received a B.A. in economics from Keio University, Tokyo in 2007 and MBA from GLOBIS University, Tokyo in 2020. He joined NTT DOCOMO in 2007.



**Masahiro Kaido**

Senior Manager, Global Marketing Strategy Group, Market Planning and Analysis Department, Research and Development Market Strategy Division, NTT Corporation.

He received a B.A. in economics from Aoyama Gakuin University, Tokyo in 2000 and joined NTT WEST the same year.



**Ichiro Kudo**

Director, Global Marketing Strategy Group, Market Planning and Analysis Department, Research and Development Market Strategy Division, NTT Corporation.

He received a B.E. in electrical engineering in 1998 and M.E. in informatics in 2000 from Kyoto University. He joined Business Communications Headquarters of NTT WEST in 2000. He joined NTT in 2014.



**Shunichiro Yamamoto**

Senior Manager, Global Marketing Strategy Group, Market Planning and Analysis Department, Research and Development Market Strategy Division, NTT Corporation.

He received a B.E. in information systems engineering from Osaka University in 1997 and M.E. in information science from Nara Institute of Science and Technology in 1999. He joined NTT in 1999.

---

## Report on ITU-T Industry Engagement Workshop

*Hiroshi Yamamoto*

### Abstract

Strengthening the relationship between the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) and industry is one of the top priority pledges of Seizo Onoe, Director of the ITU Telecommunication Standardization Bureau. The first Industry Engagement Workshop was held on April 19, 2024 in Geneva, Switzerland. The author was involved in the workshop as a member of its steering committee. The workshop details are introduced.

*Keywords: ITU-T, Industry Engagement Workshop, industry engagement metrics*

### 1. Background of the workshop

Strengthening the relationship between the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) and industry is one of the top priority pledges of Seizo Onoe, Director of the ITU Telecommunication Standardization Bureau, and the first Industry Engagement Workshop was held on April 19, 2024 in Geneva, Switzerland as an in-person event to allow for frank discussions.

At the Telecommunication Standardization Advisory Group (TSAG) meeting held in June 2023, a workshop was approved to be held, and a steering committee consisting of the leader Didier Berthomieux (Nokia, Finland) and two members appointed from each region was organized to set the goals of the workshop and consider a specific program. The author and Dao Tian (ZTE, China) were selected from the Asia-Pacific region to participate in the committee.

The goals of the workshop, as agreed in the Terms of Reference, were to

- attract relevant industry decision makers in regard to standardization where the ITU-T can provide value,
- contribute to the dialogue between all the parties,
- provide valuable feedback on the industry engagement action plan, and
- inform the World Telecommunication Standard-

ization Assembly 2024 (WTSA-24) preparations.

TSAG also approved an industry engagement action plan to attract intensive industry participation to take account of the latest technical trends and market needs.

The steering committee considered the above goals of the workshop, proceeded with the work of listing experts from each industry in each topic field, and finalized the presentation program shown in **Table 1**.

### 2. Workshop details

The keynote addresses and Session 2 (Metrics and Industry Engagement), which the author chaired, are only described below due to space constraints.

#### 2.1 Keynote addresses

##### 2.1.1 Seizo Onoe, Director, ITU Telecommunication Standardization Bureau

Mr. Onoe highlighted the significant gap of up to 10 years in deployment of mobile networks between developed and developing countries. He stated that ITU-T is the best place to help bridge this gap and mentioned a series of webinars with chief technology officers (CTOs), CxO meetings, and this Industry Engagement Workshop. He introduced the new Bridging the Standardization Gap program (**Fig. 1**).

Table 1. Workshop program.

<b>Welcome</b> <ul style="list-style-type: none"> <li>Glenn Parsons, ITU-T Rapporteur on Industry Engagement and Metrics; Principal Standards Advisor, Ericsson Canada: <i>Opportunity Statement and Goals of the Workshop</i></li> </ul>
<b>Keynote Address</b> <ul style="list-style-type: none"> <li>Seizo Onoe, Director, ITU Telecommunication Standardization Bureau</li> </ul>
<b>Keynote Address</b> <ul style="list-style-type: none"> <li>Ulrich Dropmann, Head of Standardization and Industry Environment, Nokia, Finland</li> </ul>
<b>Current Perspectives on Industry Engagement</b> <p>The purpose of this session was to provide the background necessary to understand the scope of the opportunity available to increase industry engagement in the ITU-T.</p> <p>Session Chair: Didier Berthoumieux, Chair of Workshop Steering Committee; Head of Standardization Office, Nokia, Finland</p> <p>Presentations:</p> <ul style="list-style-type: none"> <li>Bruce Gracie, Chair, WTSA-20: <i>Scene Setter on How Industry Can Engage in ITU-T</i></li> <li>Xin Chang, VP Standardization and Industry Engagement, Huawei, China</li> <li>Stephen Palm, Senior Technical Director, Broadcom, United States*</li> <li>Eliot Lear, Principal Engineer, Cisco Systems, United States</li> <li>Jun Shan Wey, Distinguished Engineer, Verizon Communications, United States</li> </ul> <p>Followed by a panel discussion with Uwe Baeder, Director Public Affairs, Rohde &amp; Schwarz, Germany and the presenters (on site).</p>
<b>Metrics and Industry Engagement</b> <p>The purpose of this session was to understand how to measure the impact of industry engagement.</p> <p>Session Chair: Hiroshi Yamamoto, Director, Head of Standardization Office, NTT Corporation, Japan</p> <p>Presentations:</p> <ul style="list-style-type: none"> <li>Christopher Clark, Head, Marketing and Partner Relations Division, ITU</li> <li>Hideyuki Iwata, President, Telecommunication Technology Committee (TTC), Japan</li> <li>David Law, Chair, Standards Board, IEEE Standards Association, United States; Distinguished Technologist, Hewlett Packard Enterprise, United States</li> </ul> <p>Followed by a panel discussion.</p>
<b>Value Proposition for Industry Engagement</b> <p>This session reviewed the current ITU-T value proposition and explored ideas to enhance the ITU-T brand to attract Industry to lead/participate/grow standardization work in the ITU-T.</p> <p>Session Chair: Arnaud Taddei (Global Security Strategist, Broadcom Europe, United Kingdom)</p> <p>Presentations:</p> <ul style="list-style-type: none"> <li>Bilel Jamoussi, Deputy to the Director, ITU Telecommunication Standardization Bureau</li> <li>Per Beming, Head of Standards &amp; Industry Initiatives, Ericsson, Sweden</li> <li>Graeme Burns, Partner - Tech, Media &amp; Telecom (TMT) and Digital, Boston Consulting Group, Switzerland</li> <li>Judy Zhu, VP Standardization, Alibaba, China*</li> <li>Xiaoja Song, Researcher, Artificial Intelligence and Intelligent Operations Center, China Mobile Research Institute, China</li> <li>Debora Comparin, Standardization Expert, Thales Group, France</li> </ul> <p>Followed by a panel discussion with Ulrich Dropmann (Head of Standardization and Industry Environment, Nokia, Finland), Bret Jordan (Chief Security Strategist, Afero, United States) and the presenters (on site).</p>
<b>Standardization Process and Industry Engagement</b> <p>This session was focused on the processes used to start, progress, approve, and publish Recommendations.</p> <p>Session Chair: Scott Mansfield, Standards Researcher, Ericsson Canada</p> <p>Presentations:</p> <ul style="list-style-type: none"> <li>Malcolm Betts, Consultant, ZTE Corporation, China</li> <li>Kam Lam, Senior Director, China Information Communication Technologies Group (CICT)</li> <li>Paul Doolan, Infinera Corporation, United States</li> </ul> <p>Followed by a panel discussion.</p>
<b>Workshop Summary, Outcomes and Next Steps</b> <p>Workshop Session Chairs summarized the discussions and provided outcomes of their sessions. Highlights of the draft workshop report were reviewed and discussed. An outlook on next steps was provided.</p> <p>Session Chair: Glenn Parsons, ITU-T Rapporteur on Industry Engagement and Metrics; Principal Standards Advisor, Ericsson Canada</p>
<b>Close of workshop</b>



Fig. 1. Keynote address: Seizo Onoe.

### 2.1.2 Ulrich Dropmann, Head of Standardization and Industry Environment, Nokia

Dr. Dropmann highlighted the importance of having global standards amid geopolitical tensions and the United Nations organization ITU as one of the best platforms to keep global coordination of standardization. ITU-T is only one among many standards organizations, and we need to better define its role in this landscape. The market has clearly recognized the excellence of ITU-T in the domains of transport and access and video codec but observed a limited market adoption and limited regional diversity in the work on cloud, protocols, and security. The importance of FRAND (fair, reasonable, and non-discriminatory) intellectual property rights rules is key to fuel a virtuous cycle of innovation. ITU should coordinate and leverage technical work in other specialized forums, maintain and improve the technical excellence acquired in certain domains, address issues in domains with limited success, and improve the decision process and metrics to assess efficiency.

## 2.2 Session 2 on Metrics and Industry Engagement

This session was moderated by the author.

### 2.2.1 Christopher Clark, Head, Marketing and Partner Relations Division, ITU

“Industry Participation Trends and Internal Tracking”

This presentation first provided an overview of the initiative of ITU-T’s internal tracking/metrics for

engagement. This initiative calculates an engagement index on the basis of information such as attributes of participants and participation statistics in recent events and makes tracking possible by creating a dashboard. The presentation then introduced the following member participation trends. 1) ITU-T industry participation is the highest it has ever been. 2) More diversity in membership regarding small and medium enterprises and regional and developing/developed nation balance is being achieved. Participation and contributions are also rising overall, especially in Asia/Pacific.

### 2.2.2 Hideyuki Iwata, President, Telecommunication Technology Committee (TTC), Japan

“Proposing Metrics and Industry Engagement through Standardization Activities at TTC”

This presentation provided the ITU standards localization efforts and its utilization status at TTC, a Japanese standards developing organization (SDO). Quantitative information was provided on how many localized standards have been downloaded and proposals up-streamed to ITU-T. Through these analyses, it was shown that localization trends among domestic organizations and the number of downloads can be effective measures of understanding the interest of the industry. It was also mentioned that TTC is holding various workshops and the participation trends in these workshops can be an effective means of learning about industry interests.



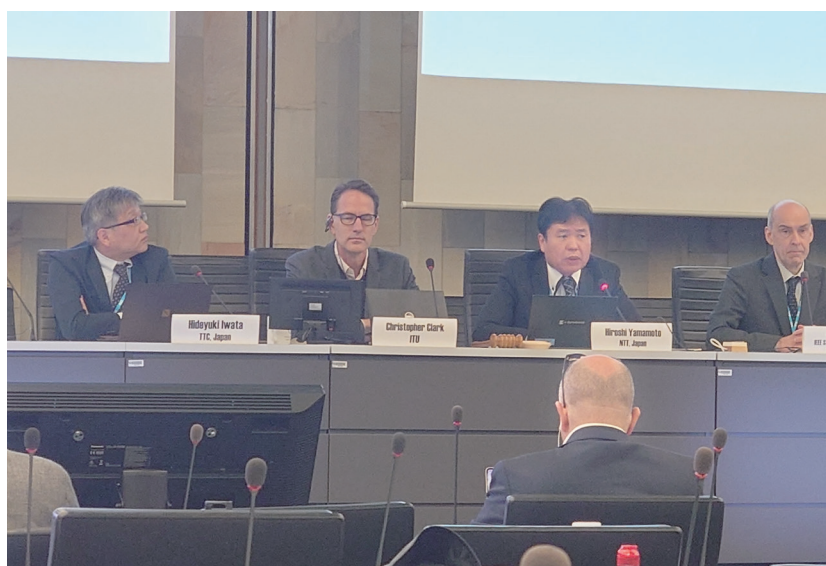


Fig. 2. Session 2.

### 2.2.3 David Law, Chair, Standards Board, IEEE Standards Association, United States; Distinguished Technologist, Hewlett Packard Enterprise, United States “Industry Metrics”

This presentation provided the latest trends in standardization activities at the Institute of Electrical and Electronics Engineers (IEEE) and considerations based on case studies regarding industry engagement metrics. Examples of recent activities are the IEEE 7000 series, which is an artificial intelligence (AI) series, addressing ways to protect personal data and safety in AI systems and IEEE 11073 for health informatics, enabling communications between medical devices and computer systems to achieve automatic capture of vital-sign information. In the second half, a discussion on industrial engagement and its metrics was introduced from the perspective of the IEEE Standards Association board, and it was pointed out that the number of recommendation downloads alone may underestimate or overestimate the needs and true value of industry, and the importance and difficulty of measuring engagement in industry was introduced.

### 2.2.4 Summary of Session 2

Three representatives from industry-recognized SDOs presented their metrics-related initiatives. While quantification and tracking of metrics for mea-

suring industry engagement are underway, the following future issues were also identified.

- Lack of indicators to quantify cross-industry collaboration and collaboration between SDOs, which is becoming increasingly important.
- Lack of means to know where and to what extent the established standards are being used and what impact they have on business.
- While the means of obtaining metrics-related data are improving, restrictions due to privacy concerns are increasing.
- The need for further development of statistical processing that takes into account the objectives of various standardization activities.

Further discussion and research is needed to identify the effectiveness of metrics and create criteria to compare one metric to another (**Fig. 2**).

## 3. Summary of the workshop

Many delegates expressed that this workshop was a success and provided a useful exchange of views from industry. It was also suggested that TSAG should hold such a workshop again, or even regularly, either separately or jointly with the CxO meetings. The action plan will be updated on the basis of the outcome of this workshop.



**Hiroshi Yamamoto**

Director, Standardization Office, Research and Development Planning Department, Research and Development Market Strategy Division, NTT Corporation.

He received a B.S. and M.S. in information and computer science from Waseda University, Tokyo, in 1999 and 2001. In 2001, he joined NTT Service Integration Laboratories, where he was engaged in the performance evaluation of Internet protocol (IP) networks, web applications, and video delivery. In 2006, he joined NTT Communications, where he engaged in the development of voice-over-IP systems. In 2010, he joined NTT Network Technology Laboratories, where he engaged in research and development of a quality-of-experience control video-delivery mechanism. In 2015, he was assigned as the primary USA liaison based in Washington, D.C. and was engaged in enhancing collaborations with US academia and industry. In 2020, he was assigned as a senior research engineer, supervisor at NTT Network Technology Labs and engaged in the research and development of future network architecture. He is currently the head of Standardization Office and oversees the NTT Group's standardization activities. He is currently the chair of the Asia-Pacific Telecommunity (APT) Preparatory Group for WTSA-24 Working Group 1 (Working Methods).

---

## Extending Usability of Copper Cables Installed Underground by Using a Cable-retracting System

*Technical Assistance and Support Center, NTT EAST*

### Abstract

The Technical Assistance and Support Center (TASC), NTT EAST developed a cable-retracting system for pulling underground copper cables that have moved back into their proper position. This article introduces this system and describes efforts to extend the usability of installed copper cables by using it. This is the eighty-third article in a series on telecommunication technologies.

*Keywords: copper cable, cable-retracting system, underground conduit*

### 1. Introduction

Cables for providing telecommunication services are installed in underground service tunnels, conduits, maintenance holes, and on utility poles above ground. Some copper cables installed in underground conduits may spontaneously move from the installed position over time due to vibration and other effects from vehicles traveling on the road directly above. This outcome is particularly pronounced around roads located on soft ground and on which many heavy vehicles travel. If a cable has moved significantly from its original position, it has been necessary to remove the cable and install a new one.

To solve such a problem, the Technical Assistance and Support Center (TASC), NTT EAST developed a cable-retracting system for pulling moved underground copper cables back into their proper position without having to replace them, thus extending the usability of the cables. This article introduces this system and describes efforts to extend the usability of installed copper cables using it while responding to various situations involving cable movement across Japan.

### 2. Effects of cable movement on surrounding facilities

When a cable installed in an underground conduit spontaneously moves from its proper position, tension in the cable due to pulling or overlength of the cable due to pushing occurs in the maintenance holes on each side of the conduit housing the cable (**Fig. 1**). Specifically, spontaneous cable pulling and pushing can cause the following problems in maintenance holes: a closure (i.e., cable-connection point) falls off the cable bracket due to the cable being pulled (**Fig. 1(a)**), a closure contacts the wall of the maintenance hole due to the cable being pulled toward the conduit (**Fig. 1(b)**), the position of a closure changes due to the overlength of the cable being pushed (**Fig. 1(c)**), and a cable buckles due to the cable being significantly pushed (**Fig. 1(d)**).

If these problems concerning underground copper cables occur, they will lead to leakage of the dry air (hereinafter referred to as gas) sealed inside the cables and closures for the purpose of maintenance<sup>\*1</sup>; thus, immediate measures need to be taken.

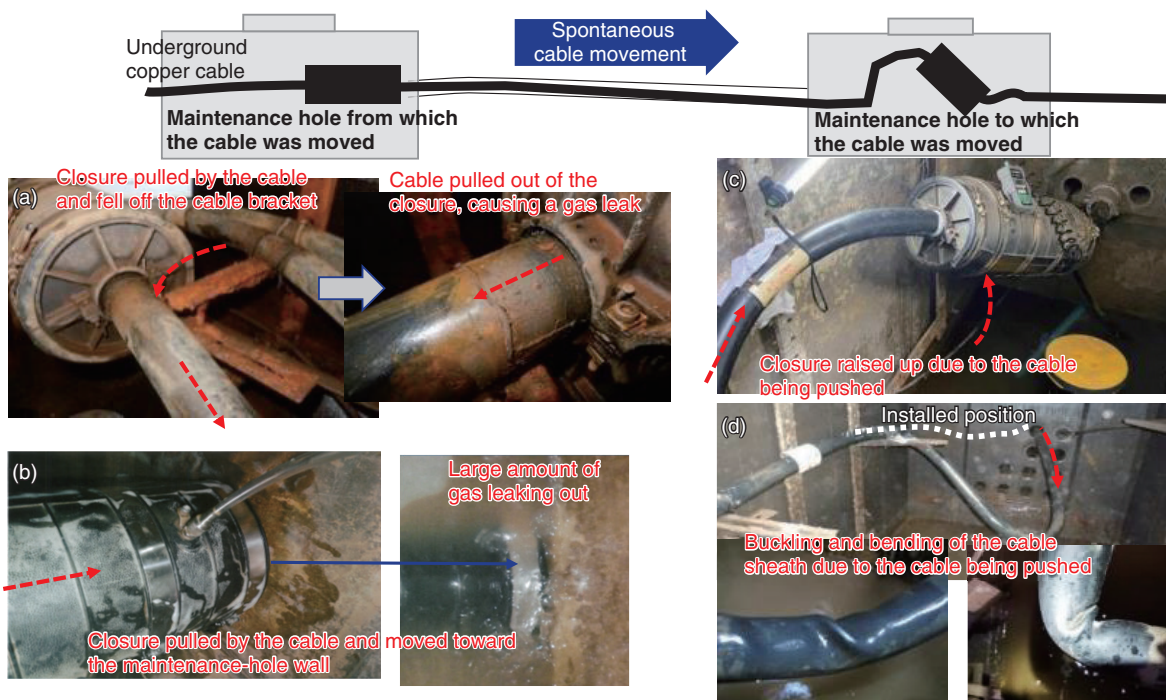


Fig. 1. Effects of cable movement on surrounding facilities.

### 3. Purpose of developing cable-retracting system

The conventional method of removing a spontaneously moved cable and installing a new one is extremely costly. To avoid such cable reinstalling and enable the continued use of cables, we developed a cable-retracting system for pulling moved cables back into their original position. We also conducted trials in the field where cable movement has occurred across Japan. The results of those trials have confirmed the effectiveness of the system.

### 4. Overview of cable-retracting system

The cable-retracting system consists of a main unit (**Fig. 2(a)**)—which is composed of a fixture for preventing spontaneous cable movement that holds the cable for retracting (hereafter cable-holding fixture) and two hydraulic cylinders—as well as a pump-connection hose, pressure meter, and manual pump (**Fig. 2(b)**). The hydraulic cylinders are bolted to the side of the cable-holding fixture and connected to the manual pump via the pump-connection hose. The mechanism for retracting the cable involves the following steps: when the manual pump is operated, pressure is applied to the hydraulic cylinders, and the

internal rod of each cylinder extends and pushes the cable-holding fixture away from the maintenance-hole wall. Thus, the cable is pulled.

During a trial conducted at an actual facility, we found that simply pulling the cable from the maintenance hole from which it was moved (pulling side) is not enough, because it only stretched the cable inside the conduit and did not move the cable. In addition to the pulling force, we thus added a force to “push” the cable into the conduit from the maintenance hole to which the cable was moved (pushing side). This combination of pulling and pushing forces enabled retracting long, heavy cables while simultaneously reducing the physical load on the cable by making it possible to retract the cable with less pulling force (**Fig. 3**).

An illustration of installation when a pulling device

\*1 In underground copper cables, high-pressure gas is continuously injected into the cable's interior to prevent water from entering the cable and closure in the event a flaw or pinhole is created in the cable sheath. If the internal pressure drops due to gas leakage, insulation failure or other malfunctions may occur due to water intrusion. The internal pressure is therefore constantly monitored, and when a pressure-drop alarm is issued, the cause of the gas leakage is investigated, damage is repaired, and gas cylinders are installed near the point where the pressure drops to supply the gas.

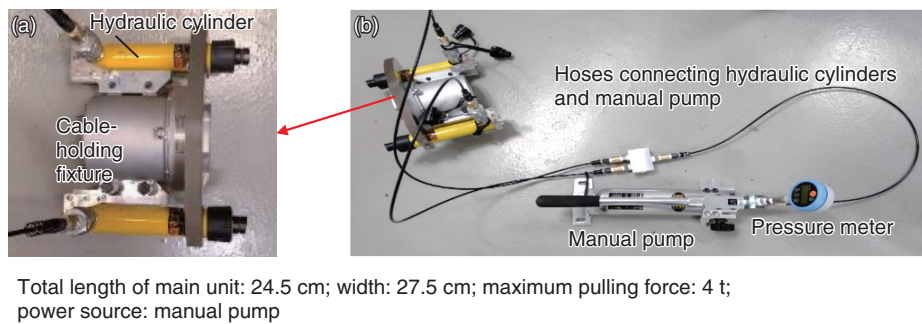


Fig. 2. Cable-retracting system.

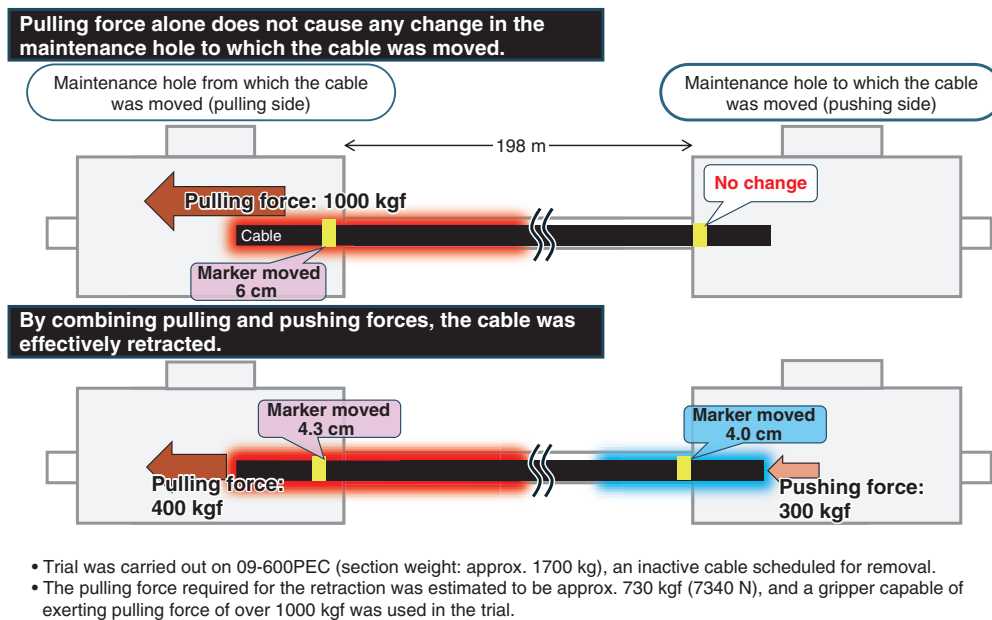


Fig. 3. Trial of retracting a cable by pulling and pushing forces.

is combined with a pushing device is shown in Fig. 4.

To push the cable from the maintenance hole to which it was moved, a strong pushing plate is installed behind the pushing device attached to the cable. The mechanism for pushing the cable is the same as that for pulling it; namely, the rod of each hydraulic cylinder extends and pushes the cable-holding fixture away from the plate. This action pushes the cable back into the conduit. The pushing plate is secured to the eye bolts on the four corners of the maintenance-hole wall by chains.

## 5. Operation of retracting a cable

Retracting a cable is done by the pulling side, assisted by the pushing side (ratio of pulling force to pushing force: 2:1 to 1.5). As shown in Fig. 5, the worker on the pulling side and worker on the pushing side communicate by mobile phone and cooperate in gradually retracting the cable.

The cable-holding fixture exerts a gripping force of up to 6000 N (611.8 kgf), and we confirmed that when a force exceeding that is applied, the gripping part of the fixture begins to slip so that the cable will not be damaged due to the pulling force exceeding the allowable tension of the cable.



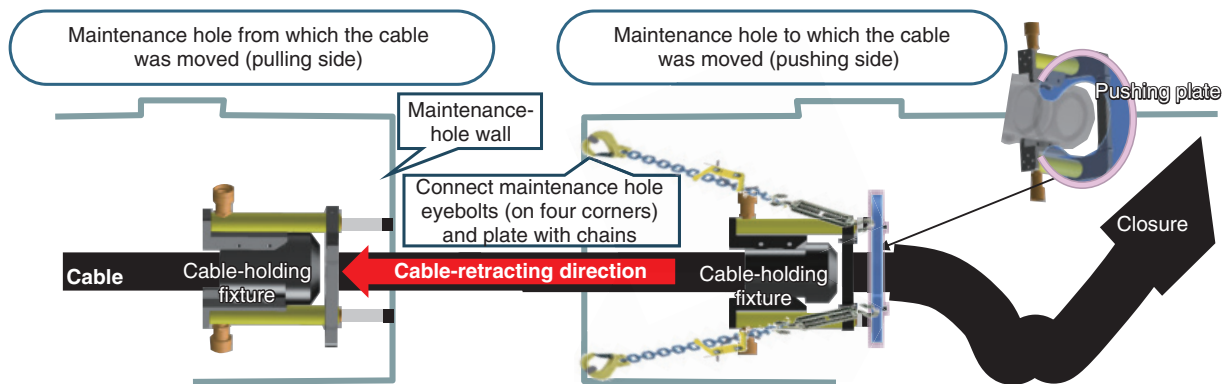


Fig. 4. Installation of pulling and pushing devices.

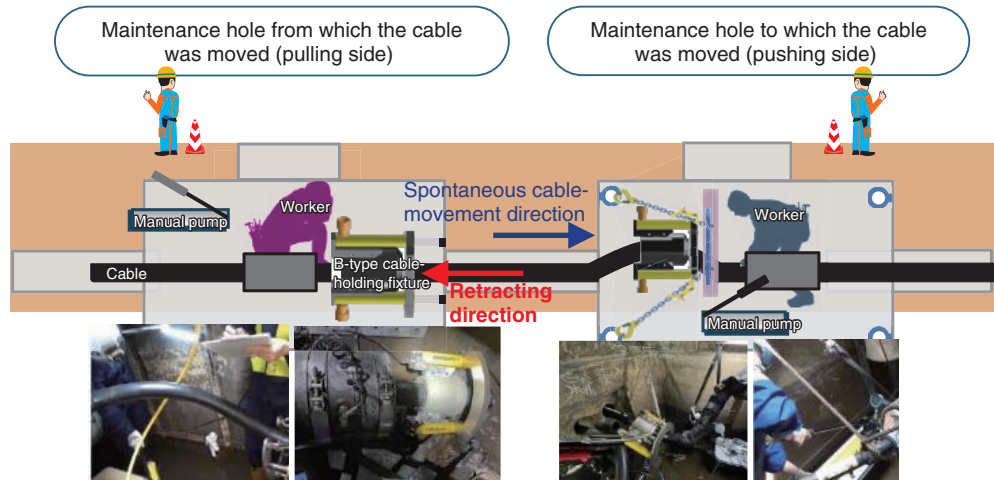


Fig. 5. Operation of retracting a cable.

## 6. Other components for assisting in retracting a cable

### (1) Rod spacer

The length of the rods extending from the hydraulic cylinders of the cable-retracting system is a maximum of 12.6 cm. When it became necessary to retract the cable further than this length, the pulling and pushing devices were detached from the cable, the devices were reattached, and retracting was resumed.

To streamline this process, we developed a dedicated rod spacer (Fig. 6). Attaching the rod spacer between the rod and maintenance-hole wall makes it possible to retract the cable longer than 12.6 cm, shortening the time it takes to detach and reattach the pulling and pushing devices.

### (2) Cable warmer

Cables that have been deformed over a long period due to cable movement and moved closures must be positioned in the proper location where they will not interfere with maintenance work. When a cable that was severely bent (buckled) is retracted, the buckled part of the cable needs to be straightened and returned to its original shape in accordance with the amount of retracting. However, this straightening cannot be done by human force, and the strong pulling force of the cable-retracting system could damage the cable sheath. A cable warmer is therefore wrapped around the buckled part to locally heat and soften it so that it can be returned to its original shape and repaired (Fig. 7).

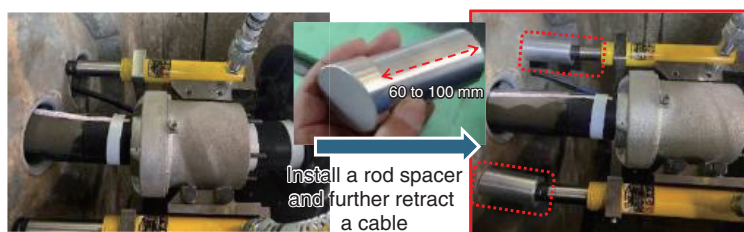


Fig. 6. Rod spacer.

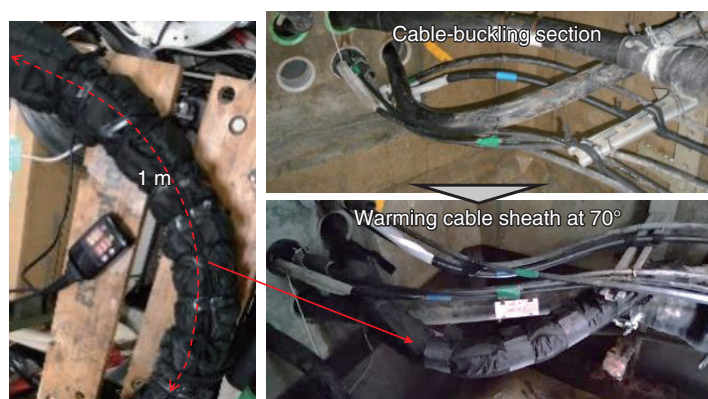


Fig. 7. Cable warmer.

## 7. Support for retracting cables nationwide

By using the cable-retracting system, we repaired moved cables in 14 locations and 26 sections\*2 nationwide, thus succeeding in extending the usability of the copper cables installed underground.

The results of our nationwide survey indicate cables have moved in many locations; thus, it is necessary to establish the conditions and mechanisms to enable on-site maintenance personnel to restore those cables. We are currently creating an operation manual and arranging the rental of the cable-retracting system.

We are also conducting a trial to verify the removal of copper cables, including lead-covered cables (which are difficult to remove), from underground sections by using this system.

## 8. Conclusion

We introduced a cable-retracting system we developed to solve the problem of moved underground copper cables. We also described efforts to extend the

usability of installed copper cables by pulling and pushing a moved cable at the maintenance holes on each side of the conduit housing it by using this system.

At TASC, we will use our accumulated knowledge, experience, and new technologies to provide technical support to resolve difficult-to-solve failures of access equipment and facilities that occur in the field across Japan. We will also continue our efforts to prevent gas leaks caused by movement of underground copper cables, the only solution to which had been replacing the cable.

## Reference

- [1] "Efforts to Pull Back Moved Underground Copper Cables," NTT Technical Journal, Vol. 33, No. 10, pp. 88–91, Oct. 2021 (in Japanese).

\*2 Section: A section of conduit connecting one maintenance hole to an adjacent maintenance hole.

# External Awards

## IEEE Haptics Symposium 2024 Outstanding Reviewer

**Winner:** Yusuke Ujitoko, NTT Communication Science Laboratories

**Date:** April 11, 2024

**Organization:** IEEE Haptics Symposium 2024

## The Laser Society of Japan Encouragement Award

**Winner:** Shota Kita, NTT Basic Research Laboratories

**Date:** May 31, 2024

**Organization:** The Laser Society of Japan

For “Ultralow Latency Operations Based on Linear Photonics Toward Photo-Electronic Converged Data Processing Infrastructure.”

**Published as:** S. Kita, K. Nozaki, K. Takata, G. Cong, Y. Maegami, M. Ohno, N. Yamamoto, A. Shinya, K. Yamada, and M. Notomi, “Ultralow Latency Operations Based on Linear Photonics Toward Photo-Electronic Converged Data Processing Infrastructure,” *The Review of Laser Engineering*, Vol. 50, No. 5, 254, May 2022 (in Japanese).

## Achievement Award

**Winners:** Shohei Matsuo, NTT Human Informatics Laboratories; Yukihiro Bandoh, NTT Computer and Data Science Laboratories (Shimonoseki City University since April 2024); Seishi Takamura, Hosei University

**Date:** June 6, 2024

**Organization:** The Institute of Electronics, Information and Communication Engineers (IEICE)

For research and development of video-coding technology and promotion of MPEG international standards.

## MWPTHz Young Scientist Paper Award

**Winner:** Katsumasa Yoshioka, NTT Basic Research Laboratories

**Date:** June 7, 2024

**Organization:** IEICE Technical Committee on Microwave Photonics and Terahertz Photonic-Electronics Technologies (MWPTHz)

For “On-chip Readout of Ultrafast Charge Dynamics in Graphene Using Terahertz Electronics.”

**Published as:** K. Yoshioka, “On-chip Readout of Ultrafast Charge Dynamics in Graphene Using Terahertz Electronics,” *IEICE Tech. Rep.*, Vol.123, No. 313, pp. 36–39, Dec. 2024.

## ICSS Research Award

**Winners:** Reika Arakawa, NTT Social Informatics Laboratories; Yo Kanemoto, NTT Social Informatics Laboratories; Eitaro Shioji, NTT Social Informatics Laboratories; Mitsuaki Akiyama, NTT Social Informatics Laboratories

**Date:** June 18, 2024

**Organization:** IEICE Technical Committee on Information and Communication System Security (ICSS)

For “A Precise Approach of Software Vulnerability Detection Using Set Similarity Calculation Algorithm.”

**Published as:** R. Arakawa, Y. Kanemoto, E. Shioji, and M. Akiyama, “A Precise Approach of Software Vulnerability Detection Using Set Similarity Calculation Algorithm,” *IEICE Tech. Rep.*, Vol. 123, No. 86, ICSS2023-7, pp. 32–39, June 2023 (in Japanese).