NEC Innovation Day Q&A

Date/Time: December 15, 2023, 10:00-11:30 am JST
Location: NEC Tamagawa Renaissance City Hall Building
                Live online streaming
Presenters: Motoo Nishihara, Corporate EVP and CTO
                Toshifumi Yoshizaki, Corporate EVP and CDO
                Akio Yamada, Corporate SVP
                Shigeki Wada, Corporate SVP

Questioner A

Q:

What are your thoughts on monetizing technology?

A:

The most important thing is to "contribute to our business." Therefore, it is crucial to firmly implement the technology in the NEC Digital Platform (NDP), which plays a central role in DX, and to ensure that it is widely used. We have been developing and promoting "cotomi," our generative AI, in collaboration with CDO Toshifumi Yoshizaki's team, which is responsible for the business, by keeping business contribution in mind from the R&D stage.

In addition, to further monetize our technologies, we are also engaging in intellectual property business and implementing advanced technologies through startup companies such as NEC X to earn returns in the future. However, in working with startups, we need to take a long-term view of around 3 to 5 years.

The cumulative amount of income from intellectual property for the period covered by the Mid-term Management Plan (2021-2025) is increasing at more than double the pace of the previous five years (2016-2020). By making this the norm rather than a temporary trend, we will be able to build up a steady stream of profits for the years to come. Our intellectual properties are used in a variety of industries, including for networking and video-related technologies applications, and we expect our licensing revenues as a partner to increase in the future. We have received many requests from startup companies to use NEC's AI technology, and we expect to gain income in the form of licensing revenue through technology licensing in the future.

Q:

NEC has partnered with twelve companies and three universities in the Japanese-

language-specific generative AI business. What was the basis for choosing the partners?

A:

NEC has chosen partners from different industries for our generative AI "cotomi." It is important that we work together with different industry players in creating a framework for the future. To be able to actually use the system in a company, it is crucial for us to achieve extremely high accuracy, ensure fairness, etc. on the framework and to create a model that can be used quickly.

Questioner B

Q:

How large of a business scale do you aim to achieve with the generative AI "cotomi"? I think the profitability of Phase I is low since it is based on individual SI targeting specific customers. How do you aim to improve profitability in Phases II and III?

A:

We are aiming for sales of ¥50 billion in three years. Looking at the revenues for generative AI "cotomi" by phase, the majority, accounting for 40% to 50% of the total, will be in Phase II, where industry- and business-specific models will be developed and incorporated into solutions. Phase I, where individual SI will be the main focus, is expected to account for less than 20%, and Phase III, where partner companies will deploy the system, will account for 20% to 30%. In Phase II, we will link the system with NEC Digital Platform (NDP) to increase both sales and profits. We will transform our business model, including value pricing, by incorporating our industry know-how based on our generative AI and offering it as a service tailored to the Japanese market.

Questioner C

Q:

How are customers reacting to generative AI?

A:

Currently, most of our customers are positively considering how to use generative AI. In business, whether or not you use generative AI can make a big difference. It is important to find ways to reduce risks, especially how to ensure the accuracy and fairness of inputs and outputs. We are working with Robust Intelligence in assessing risks of LLMs based on global standards to provide safe and secure generative AI to our clients.

Q:

What are your thoughts on the expanding scale of LLMs ?

A:

NEC's stance is that LLMs are not about size competition. The strength of NEC's current LLMs is that they have very high performance and are strong in the Japanese language despite its small size. These advantages enable its use in portable terminals, for example. Currently, NEC is developing a new architecture that will enable linearly expanding the scale and intelligence by combining small LLMs and creating new AI by flexible linkage with various specialized AI such as face recognition in addition to LLMs.

Additionally, we are conducting multiple trials to find the best balance between size and data. As of July this year, we have started offering LLMs at a size of 13 billion parameters, which we believe is the best balance between load and performance for our customers. By combining these LLMs, we hope to create a scalable architecture that can cover both small and large size requirements. The larger the size, the more data and knowledge can be built in, and thus the potential to expand its capabilities will be greater. Usability, however, will diminish. We will therefore aim to provide the best solution for our customers' needs through trial and error to find out how to achieve the best balance going forward.

Questioner D

Q:

What fields and industries do you see as promising for the use of generative AI in the future? What business models do you have in mind?

A:

Areas such as medical care, local governments, and contact centers, where verifications are being conducted ahead of other areas, are very promising. We are also developing frameworks for the financial and manufacturing industries. Each industry has different priorities, use cases, and required functions, but we would like to establish several patterns, frameworks, and models, and eventually create a system for delivering value to various customers through value pricing.

Q:

What are your future investment plans for generative AI?

A:

We will continue to invest in AI supercomputers in order to develop a variety of advanced AI in the future. We have no specific plans that we can disclose at this time.

Q:

How do you plan to strengthen AI specialists in the future?

A:

The R&D department is bringing together members in charge of AI-related businesses into one organization to create a powerful unified workforce. Naturally, we are also

considering hiring new talents. We will strengthen the seamless linkage between R&D and business to accelerate the commercialization of research results.

Questioner E

Q:

What do you think about the possibility of offering services that incorporate NEC's generative AI to world-class IT service companies doing business in Japan? Have you already received inquiries?

A:

We have already received inquiries from several companies, and NEC's LLM is attracting attention also for its power-saving feature, not only for its Japanese language capability. We believe that rather than the difference in GPU performance, the algorithm is more important in determining LLM performance; thus, we are developing LLMs by focusing on the algorithm. Many global IT service companies have recognized the value of this approach and have asked us to work with them to develop LLMs for the Japanese market.

Questioner F

Q:

What specific research will be conducted at the Generative AI Center? How will collaboration be carried out with the NEC Generative AI Hub, which is responsible for the business aspect?

A:

In addition to the development of foundation models, the Generative AI Center will develop technologies for high-potential applications based on these models and for linking them with various AI models. The key point is to carry out development with a view of the applications while focusing on LLM as the foundation. We have excellent researchers in these areas in our research laboratories in Japan, Germany, and the U.S., and we are working closely with them. It is important to work together seamlessly with the NEC Generative AI Hub, which is responsible for both R&D and business.

The Generative AI Center is oriented toward agile research and development. This is an approach that is different from the traditional schedule-based flow of conducting research, development, and delivery in stages. Instead, we will conduct research concurrently with development, while incorporating information from our customers, to enable delivery in the shortest time possible.