

NEC Innovation Day 2024 Q&A

Date/Time: Wednesday, November 27, 2024, 10:30-11:30 am JST

Format: Live online streaming at NEC Tamagawa Office

Presenters: Motoo Nishihara, Executive Officer, Corporate EVP, and CTO

Akio Yamada, Corporate SVP, Managing Director, AI Technology Service Business Division, and AI Research Officer

Shigeki Wada, Corporate SVP, Managing Director, Corporate Business Development Division

Questioner A

Q :

Since the release of NEC's generative AI "cotomi" version 1 (v1), which was announced in the previous fiscal year, what challenges have you overcome, and how has it evolved into the current "cotomi" v2? In particular, could you please discuss the feedback received from customers over the past year and any internal insights gained within NEC?

A :

LLMs (Large Language Models) are built to meet general standards so that they can provide general responses to general queries. However, we have learned that when customers actually use LLMs, the level of specialization, or the ability to offer specialized and more accurate answers to specific queries, is crucial. Meanwhile, there remains a certain degree of expectation for LLMs to handle general queries effectively. Therefore, we are making improvements, with a necessary condition of ensuring performance that meets or exceeds that of ChatGPT-4. That said, speed is more important. When using an AI agent, the agent exchanges data with LLMs multiple times behind the scenes. As a result, small delays in each response can accumulate, affecting the customer's user experience. Over the past year, we have learned that fast responses, which feel like natural conversations, are critical. Therefore, while establishing speed as an absolute condition, we are targeting performance that meets the level I just described.

Questioner B

Q :

NEC is targeting the creation of ¥300 billion in business value from growth businesses in FY2025. Could you please provide a breakdown of this business value? Also, I believe that your best example of a successful source of intellectual property revenue is dotData. What is your assessment of dotData's business value?

A :

Just over half of the business value of ¥300 billion will come from the Healthcare and Life Sciences business, including AI drug discovery. The remaining portion, just under half, will come from other areas such as data-driven DX/dotData and AI agritech. Additionally, there is revenue from intellectual property licensing. This component is not included in business value, so we believe that it has growth potential.

Among the growth businesses presented today, the data-driven DX business is growing significantly. In the data-driven DX business, along with dotData products, we are expanding business in areas such as relevant NEC consulting services and the supply of platforms. This structure will enable growth in both sectors. As we grow under this structure, we expect business value to increase through future actions, including IPOs and M&As.

Questioner C

Q :

With the AI agents based on "cotomi," I get the sense that in the future NEC will focus more on automating human work rather than competing in the learning race. It seems that NEC is moving away from the LLM learning race, and instead settling down into conventional AI models, including the application integration that lies beyond learning. Is my understanding correct?

A :

Various companies are offering many different versions of LLMs. We believe that the question is not about which will win or lose. Each LLM's individuality stems from its learning methods and how it uses data, so we envision a society in which a variety of LLMs operate alongside one another. That is why the approach of LLM orchestration, which we discussed as a conceptual framework last year, will be

crucial. We will continue to develop an LLM with the level of specialization that NEC considers essential. Based on this, the value we offer customers will come from the applications that use LLMs, rather than the LLMs themselves. NEC is particularly focused on automating business processes.

Q :

To significantly boost productivity, AI will be used across companies and will be integrated with one another across divisions. In some cases, decision-making may also be automated. I believe this kind of world will soon be a reality. Do you think it might take a little longer to arrive?

A :

That is exactly what we are working on right now. If AI is only used for questions and answers, it remains just a supporter of tasks. However, AI agents can be implemented autonomously, including performing tasks, and can automatically design and implement coordination between multiple tasks. As a result, AI agents target entire processes, not just individual tasks. The question of how to optimize the efficiency of the entire process lies at the heart of digital transformation (DX).

Questioner D

Q :

What is your outlook for the future of AI agents? Do you expect their use will expand beyond simple text input and response, to include controlling movements in hardware, such as physical AI, and eventually be fully integrated into hardware?

A :

Yes, we do. While NEC is also working on robotic control using AI, today's presentation focused more on fields closer to IT than robotics. Beyond helping individuals as a supportive partner, we aim to create a world where AI can comprehensively manage all of an organization's work processes, with AIs collaborating with each other much like different parts of an organization work together. Although integration with the physical world will eventually occur as a matter of course, we believe that the progress on actuators will be made in cooperation with our partners.

Questioner E

Q :

Could you please discuss any issues that may become obstacles from legal and technical perspectives as you expand the AI agent business?

A :

We believe the biggest issue is the perspective of social acceptance. Since AI operates based on probability theory, errors will inevitably occur with a certain likelihood. While errors and hallucinations are generally seen as negative, they actually indicate that AI has logically processed information and generated an answer, but the result was incorrect. These instances should be viewed as just one possible line of reasoning. For example, in daily business tasks performed by a person, a colleague might check the person's work before it is implemented. Likewise, AI could be viewed as a "colleague" of sorts and incorporated as a checking mechanism, enabling business tasks to be carried out efficiently. If social acceptance of this approach emerges, we believe that society will make progress toward overcoming the state of not being unable to introduce AI because it is not perfect. In the U.S. and other countries, many people say they are open to trying AI even if it makes mistakes. We believe this reflects differences in the level of AI adoption between Japan and Europe and the U.S.

Q :

You explained that NEC will add generative AI-related skills to its AI talent development. When do you plan to expand these efforts? Also, could you please discuss any future plans to increase staffing to address specific job categories where you face personnel shortages?

A :

We currently face personnel shortages across every field. While many people can explain generative AI, the key feature of the program we discussed today is to develop talent capable of independently discovering solutions through discussions with customers.

Questioner F

Q :

As various forms of AI emerge, I am concerned about whether companies will be able to truly use them effectively. I believe functions such as AI governance will

be necessary within companies. In my view, this role should be fulfilled by IT service vendors. What are NEC's thoughts on this?

A :

We strongly believe that governance will be key to advancing business. When organizations and individuals adopt AI, it is crucial to clearly define the roles people will fulfill. NEC is currently conducting demonstration tests on in-house AI use, positioning itself as client zero. Throughout this process, a variety of opinions will be expressed. Our basic approach is to advance AI while sharing the insights we gain with our customers. Based on this, we believe the role of an IT vendor is to develop AI while exploring these issues together with our business partners.

Unlike IT systems, the role of AI technology is not pre-determined, and it is highly flexible. AI should be seen as an AI worker that generates probabilistic answers, and businesses need to consider how these AI workers will collaborate with their human counterparts. This should be considered by businesses in their respective industries within the context of their overall governance and audit structures. Under current rules and guidelines, the focus is primarily on how to safely use AI. I believe this will be taken a step further, and new concepts, such as how to audit environments where AI workers and humans coexist, will likely emerge.

Questioner G

Q :

Could you please provide more details on the reduction of power consumption in generative AI, particularly regarding your future activities and direction?

A :

Power consumption by AI has been rapidly increasing. Until last year, the power consumed during the learning phase was the issue. Currently, the power consumed in the inference process has become a major concern. NEC is focused on reducing power consumption for each inference and will thoroughly consider reduction measures.

NEC has successfully doubled energy efficiency for the same generative AI performance by improving GPU usage efficiency. As AI agents evolve, the number of queries to LLMs will increase significantly, leading to further rises in power consumption. That is when reducing power consumption per transaction with LLMs will become essential. Additionally, since AI agents greatly increase the

number of transactions with LLMs, slow response times will severely reduce system performance. Balancing these two priorities will be crucial. We believe that NEC can reduce power consumption while improving performance by leveraging the middleware format, including resources like its existing HPC technologies.

Questioner H

Q :

How do you expect AI agents to be used? Will a company initially use them to develop strategies based on internal data? Or are you also contemplating the use of AI for automated negotiations with parties outside a company?

A :

NEC's AI agents will execute business tasks not only through access within a company, but also by accessing external parties, if necessary. We have developed a function that enables automated inter-AI coordination when coordination with a third party is needed in this process. NEC's inter-AI automated negotiation technology we previously presented is a type of AI agent. However, we believe this technology will be introduced to the market in the next step beyond the current implementation.

The direction we foresee for AI agents is their evolution into multi-AI agents. Inter-AI integration is important when AI agents with different roles are connected. We envision an architecture that connects people with a system that integrates numerous AI agents. However, humans will still bear responsibility in this situation. Therefore, we must consider governance and audit structures as a unified whole to fulfill this responsibility. We can describe the current phase as a period for constructing an architecture together with the people who will serve as operators of society.