

Developing Artificial Intelligence Services that Satisfy Customer Demands: Moving forward with Social Implementation of corevo[®] Technologies

Hideaki Ozawa

Abstract

Artificial intelligence (AI) technologies are playing a steadily increasing role in supporting human work and social activities, such as in contact center operations and smart speaker implementations. This article introduces NTT's initiatives in developing AI technologies that are essential in providing services needed by customers in the real world, including the technologies needed for effectively utilizing AI in the real world, such as AI applications for household voice communication systems and medicine. Technologies for accelerating processing speed to collect more extensive data from the real world and improve the accuracy of AI are also introduced.

Keywords: corevo, AI, DNN

1. Introduction

The NTT Group is pursuing technology development, service proposition, and social implementation of its artificial intelligence (AI) technologies called corevo[®] based on the following four concepts:

- (1) Agent-AI: Supporting interaction between humans and computers such as speech recognition and natural language processing
- (2) Heart-Touching-AI: Generating added value by leveraging the internal human mental state such as emotions and the ways humans understand surrounding events and phenomena
- (3) Ambient-AI: Generating added value by collecting information from our social environment such as factories, vehicles, and roads, and analyzing that information, for example, by applying AI technologies that detect abnormalities in machinery and control the flow of people in crowded areas
- (4) Network-AI: Generating knowledge by connecting AI in various locations—similar to the concept of *two heads are better than one*—and enabling stable operations by finding failures in the network itself

The past year has seen rapid progress in AI initiatives among the NTT Group companies, with more than 60 press releases (in Japanese) related to corevo issued. In particular, new Agent-AI services have been developed based on natural language processing and speech recognition to generate new value by enabling computers to understand information produced by humans. For example, computers process information from users of machine translation systems to return translation results, analyze calls in a contact center system to rapidly comprehend the issue described by the customer, or act as reception robots that talk with customers and provide guidance information (**Fig. 1**).

Likewise, new Ambient-AI services have also been

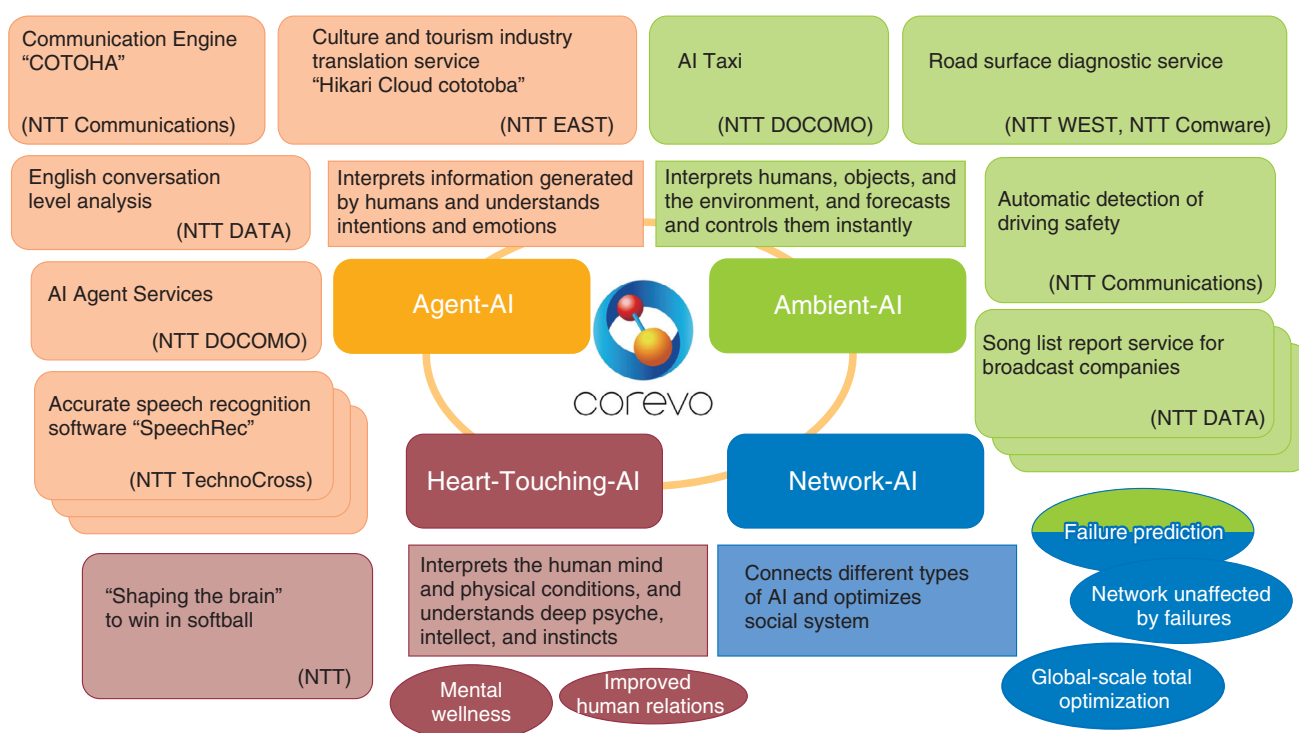


Fig. 1. AI technologies and services provided by the NTT Group.

developed to generate added value through the observation of various objects and phenomena in the social environment. These include the prediction of factory and plant failures by analyzing images from cameras and sounds collected by sensors or microphones, discovery of mechanical failures by detecting abnormal sounds produced by machinery, and measurement of holes and cracks in roads and the level of wear of manholes.

AI performance has been reported to surpass that of humans in the world of games, such as in *go* and *shogi* (Japanese chess). However, the NTT Group is aiming for co-creation between humans and AI of services and applications and has therefore endeavored to develop AI technologies that support human social activities in the real world, as AI that extends—rather than competes with—human capabilities. These initiatives have also revealed certain issues that need to be addressed in applying AI technologies in the real world.

2. Shift to AI that takes an active role in the real world from AI that merely processes information

AI technologies developed recently are capable of

collecting a vast amount of data and classifying the data based on internal data features through a method called machine learning. They also widely use deep neural networks (DNNs) and other deep learning methods for determining input-output correlations to produce the most probable output answer for newly input data. This is done through mechanisms for automatically learning relationships between input data and output data of correct answers provided by humans. These methods select the best solution from one type of combinatorial problem, making them very effective for selecting the best solution from among an extremely large number of combinations based on a fixed set of rules, such as in *go* or *shogi*.

Applying AI to the real world to address various problems that can potentially be solved using AI by simply collecting a large amount of data and finding the best state or solution is sometimes inefficient, though, because it requires a tremendously long computation time. For example, speech recognition technologies collect a large amount of human voice data and use DNN-based learning systems to correlate voice data with text transcribed from the voice data. This can be approached as an optimization problem for correlating human voice output with transcribed

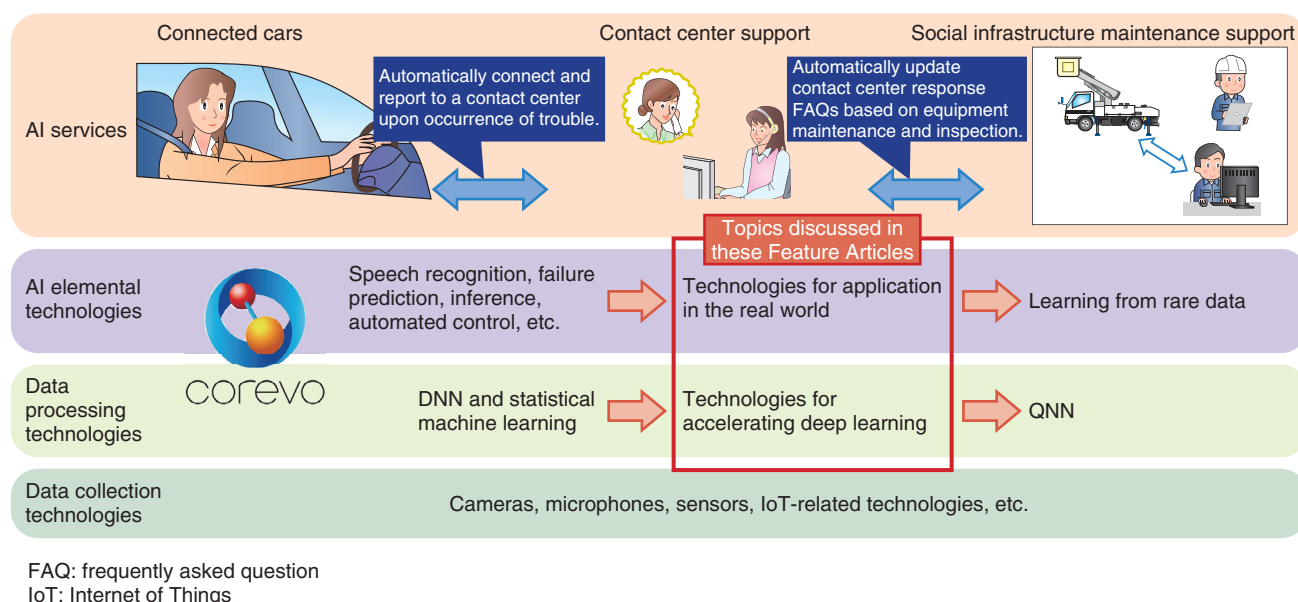


Fig. 2. Architecture of AI technologies as conceived by the NTT Group.

text such as by using AI to learn human voice data collected in a low-noise environment and identifying variations in sound between male and female voices or finding differences in sounds that are easy to pronounce for certain people using speech recognition in a low-noise environment.

The real-world environment where speech recognition is used, however, requires distinguishing the human voice amidst various types of noise such as sounds from the television or stereo when using robots in living rooms, and from stereo music, the sound of air conditioning, and road noise when using speech recognition inside vehicles. Since the types and levels of noise vary depending on the environment where speech recognition is used, simply collecting data and analyzing combinations would require an enormous amount of data and computing power. It is therefore necessary to perform learning by taking into account the possible sources of noise in the particular environment where speech recognition is used such as by controlling road noise or other sound disturbances. This will effectively increase the accuracy of speech recognition.

In addition, learning more data from the real world within a short period will make it possible to deliver AI systems that will become increasingly more useful for customers every day. The NTT Group, therefore, is also focusing on developing methods to accelerate the process of learning massive amounts of data. In

addition to developing technologies to optimize learning through DNN, which is currently one of the main AI technologies, we are also developing quantum neural network (QNN) technologies for solving combinatorial optimization problems at speeds faster than that of DNN methods.

To offer the best services to users of AI services in the real world, we will continue to strive to make AI wiser by creating a mechanism for accurately finding relevant learning data and rapidly processing a larger amount of information. The Feature Articles in this issue introduce examples of AI technologies being developed by the NTT Group for solving customer problems in the real world and also present the current status of our initiatives to accelerate learning using DNN methods (Fig. 2).

3. AI elemental technologies for the consumer sector

Smart speakers and other agent devices are now becoming particularly popular in the consumer sector. Users can listen to music, turn on electronic appliances, and enjoy various other services by giving voice instructions to an agent device. The Feature Article in this issue entitled "Efforts to Enhance Far-field Speech Recognition" [1] introduces a speech recognition technology that accurately captures the user's voice and is a key technology for agent devices.

Healthcare is another promising sector for targeting consumers. The NTT laboratories are engaged in research to better understand daily physical activity states and health conditions based on sensor information and medical and health data, as well as research to support disease prevention and health promotion based on the analysis of physical activity and health conditions. Another article in this issue, “Biosignal Processing Methods Targeting Healthcare Support Services,” [2] introduces a technology for accurately estimating vital data from biosignals obtained from sensors, as a way to more thoroughly understand health conditions. The article entitled “Artificial Intelligence-based Health Management System: Unequally Spaced Medical Data Analysis” [3] introduces AI technologies that support blood sugar management based on clinical data from a university hospital, as part of research to support disease prevention and health promotion.

4. AI elemental technologies for the business sector

The first thing that comes to mind in applying AI for business is optimization of operations. NTT is engaged in research on complementing and drawing out human capabilities using AI.

For example, maintenance inspection of communication facilities and other infrastructures owned by the NTT Group is very labor-intensive and therefore urgently requires optimization. To address this issue, the NTT laboratories are pursuing research and development of ways to optimize maintenance and inspection operations. As an example, the article entitled “Automatic Degradation Estimation of Manhole Covers for Efficient Inspection via Vehicle-mounted Cameras” [4] introduces a technology to estimate the extent of manhole degradation based on images from vehicle-mounted cameras.

Also, the recent increase in inbound visitors to Japan has led to a surge in demand for translation of tourist and other information. The article entitled “Efforts toward Service Creation with Neural Machine Translation” [5] introduces initiatives to improve translation accuracy through actual field applications of AI-based automatic translation technologies.

Another highly promising field for AI business applications is data analytics. The NTT laboratories are working towards developing technologies for predicting and detecting phenomena that cannot be observed in advance, based on imperfect data. The

article entitled “People Flow Prediction Technology for Crowd Navigation” [6] introduces a means of spatio-temporal prediction for simultaneously analyzing time and space phenomena.

5. Platform technologies underpinning AI

AI elemental technologies have become possible through complex and massive computational processing. The NTT laboratories are also developing the technologies underpinning these processing operations. Platform technologies are divided into the frameworks related to learning and analysis, hardware, and the technologies for improving the operability and convenience of the frameworks and hardware. The Feature Articles in this issue introduce technologies related to AI learning.

Although learning based on massive data is necessary for AI to derive the right answer, the process entails a number of problems such as the extremely high computation cost of learning and the lack of guarantee of convergence of learning results. The article entitled “Advanced Learning Technologies for Deep Learning” [7] introduces initiatives aimed at resolving these issues inherent to DNN learning processes.

6. Future directions

To further enhance the lifestyles and business operations of many customers through corevo technologies, the NTT Group will continue to work with different partners in pursuing the applications of corevo technologies in the real world.

References

- [1] K. Matsui, T. Moriya, H. Itou, T. Fukutomi, S. Saito, Y. Shinohara, S. Kobashikawa, Y. Yamaguchi, and N. Harada, “Efforts to Enhance Far-field Speech Recognition,” NTT Technical Review, Vol. 16, No. 8, pp. 12–17, 2018.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa2.html>
- [2] K. Eguchi, R. Aoki, S. Shimauchi, A. Chiba, and N. Asanoma, “Biosignal Processing Methods Targeting Healthcare Support Services,” NTT Technical Review, Vol. 16, No. 8, pp. 29–36, 2018.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa5.html>
- [3] H. Kurasawa, A. Fujino, and K. Hayashi, “Artificial Intelligence-based Health Management System: Unequally Spaced Medical Data Analysis,” NTT Technical Review, Vol. 16, No. 8, pp. 24–28, 2018.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa4.html>
- [4] K. Murasaki, S. Ando, and T. Kinebuchi, “Automatic Degradation Estimation of Manhole Covers for Efficient Inspection via Vehicle-mounted Cameras,” NTT Technical Review, Vol. 16, No. 8, pp. 18–23, 2018.

- <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa3.html>
- [5] S. Yoshida, M. Mizushima, and K. Tanaka, "Efforts toward Service Creation with Neural Machine Translation," NTT Technical Review, Vol. 16, No. 8, pp. 42–46, 2018.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa7.html>
- [6] D. Sato, H. Shiohara, M. Miyamoto, and N. Ueda, "People Flow Prediction Technology for Crowd Navigation," NTT Technical Review, Vol. 16, No. 8, pp. 47–52, 2018.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa8.html>
- [7] Y. Ida, S. Kanai, Y. Fujiwara, S. Yagi, and Y. Iida, "Advanced Learning Technologies for Deep Learning," NTT Technical Review, Vol. 16, No. 8, pp. 37–41, 2018.
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201808fa6.html>



Hideaki Ozawa

Vice President, NTT Media Intelligence Laboratories.

He received a Ph.D. in engineering from the Graduate School of Science and Technology, Keio University, Kanagawa, in 1992. After joining NTT in 1991, he engaged in research and practical application of multimedia processing technologies at NTT Human Interface Laboratories and Cyber Solution Laboratories, and provision of local multimedia information at Walkerplus, Inc. He also worked on the development and management of "goo" internet services including the search engine and healthcare services and was involved in establishing new services at NTT Resonant Inc. In addition to his current position at NTT Media Intelligence Laboratories, he has also served since May 2013 as president of NTT Resonant Technology, Inc., which develops software for Android/iOS.
