Artificial Intelligence Research Activities and Directions in the NTT Group

Takeshi Yamada, Satoshi Takahashi, Futoshi Naya, Takashi Ikebe, and Shigeto Furukawa

Abstract

The research and development of artificial intelligence (AI) at NTT is advancing along four directions: (1) *Agent-AI* for analyzing information issued by people and understanding intentions and emotions in that information; (2) *Heart-Touching-AI* for analyzing unconscious and unnoticeable aspects of a person's mind and body and understanding deep psychological, intellectual, and instinctual states in that person; (3) *Ambient-AI* for analyzing and understanding just about anything in the world (objects, people, the environment) and instantaneously predicting and controlling those things; and (4) *Network-AI* for organically connecting and cultivating multiple types of AI and optimizing the entire social system. This article introduces the technologies supporting these four forms of AI and specific AI-related activities in the NTT Group.

Keywords: artificial intelligence, machine learning, brain science

1. Introduction

We are now experiencing another artificial intelligence (AI) boom as interest in its possibilities continues to grow in science, business, and the mass media. There are various reasons for this renewed interest in AI. First, research in deep learning, a type of machine learning technology, is progressing and coming to dominate a variety of information-communication fields such as image/voice recognition and understanding [1]. Also, when IBM's Watson question and answer (Q&A) system beat human champions in a television quiz show, it ignited a trend in the use of AI-related technology in the business domain. Finally, the commercialization of various types of humanoid robots has helped to make many people more comfortable about a society in which people and robots coexist. In addition, some researchers predict that a technological singularity in which AI comes to surpass all human ability will occur by 2045 [2]. This possibility is generating a great deal of interest throughout the world.

While it is important to get on board this AI trend without delay and take a lead in advancing the research on it, it is also important to analyze both the technological and business aspects of AI in a calm and thorough manner. For example, deep learning and Q&A systems constitute only a portion of technologies that will be needed in order to make people more comfortable with AI, while in the business domain, the first signs that AI may bring about a revolution in the industrial structure are starting to appear.

In these Feature Articles, we introduce some of the activities underway in the NTT Group that are expected to lead to a genuine implementation of AI. As an introduction to these articles, we explain here the four directions that NTT is taking in the research and development (R&D) of AI.
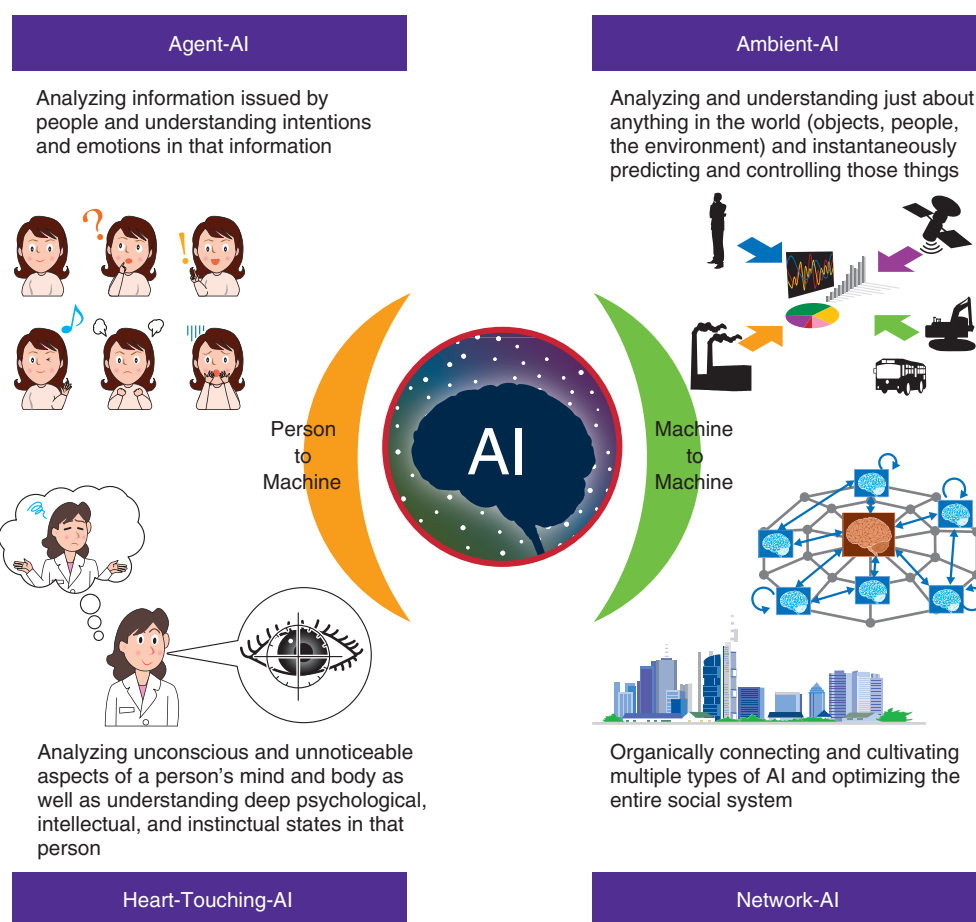
Fig. 1.   Four AI directions set by NTT.

2.   NTT's four AI directions

The world of AI is quite diverse with many and varied supporting technologies, and people often view AI in different ways. For example, AI is commonly thought of as the pursuit of human-like behavior in a computer, so that the computer, for all practical purposes, appears to be human. However, in today's information and communication technology society, all sorts of devices and sensors are being connected to the Internet to form the Internet of Things (IoT), and the idea that these *things* can have some type of intelligence and behave in an organized manner can also be called AI.

With this in mind, NTT has established the following four concepts or directions reflecting different forms of AI that can be useful to society (**Fig. 1**).

(1) Agent-AI for analyzing information issued by people and understanding intentions and emotions in that information. If we consider Q&A-type AI to be a typical example of conventional AI, then Agent-AI would be located far out on the development line extending from that. Agent-AI can also be called AI in the pursuit of human-like behavior.

(2) Heart-Touching-AI for analyzing unconscious and unnoticeable aspects of a person's mind and body as well as understanding deep psychological, intellectual, and instinctual states in that person. It is thought that deepening our knowledge of the mechanisms underlying the human brain and mind will lead to a new world in which people and AI can interact in a more direct manner.

(3) Ambient-AI for analyzing and understanding just about anything in the world (objects, people, the environment) and instantaneously predicting and controlling those things. This type of AI would analyze information from things and the environment in addition to people and feed
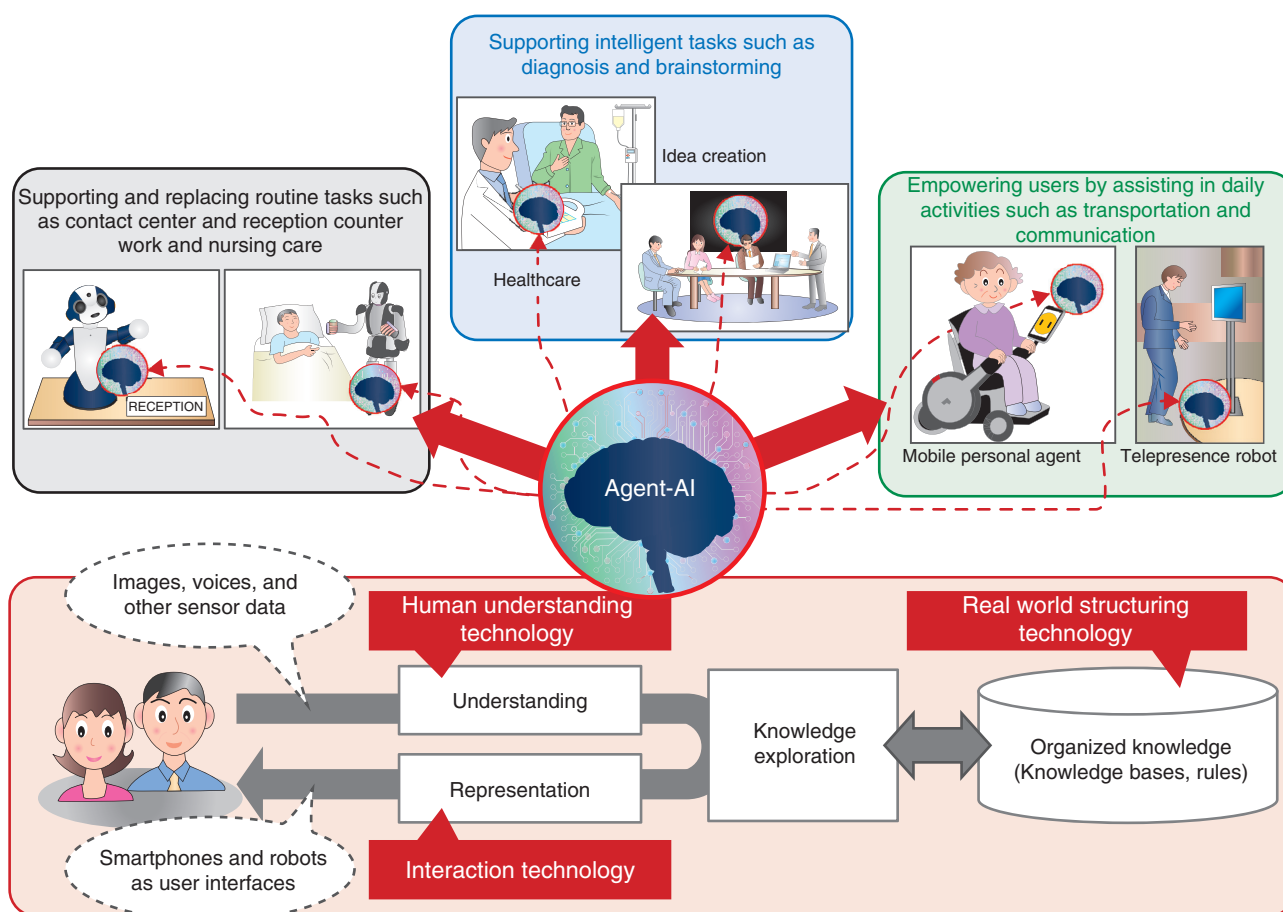
Fig. 2. Agent-AI.

the results of analysis back to the real world.

(4) Network-AI for organically connecting and cultivating multiple types of AI and optimizing the entire social system. Reassessing the network as an infrastructure from an AI perspective should make it possible to create a completely new social system.

2.1 Agent-AI

As the name implies, Agent-AI is aimed at achieving advanced interaction with human beings as a personified intelligent agent typified by a robot. This would be accomplished in diverse ways such as by understanding people and the conditions surrounding them through voice, language, images, and other types of media (human understanding technology), by performing multimodal interaction that includes human facial expressions, body language, and gestures (interaction technology), and by being conversant in multiple languages and performing compli-

cated reasoning based on a huge amount of knowledge (real world structuring technology) (**Fig. 2**).

One example of the above within the NTT Group is the development of an interactive system that is meant to achieve natural conversation with a computer. This is a joint project between NTT and NTT DOCOMO, the latter of which is already providing the *Shabette Concier* (Talking Concierge) voice-agent service for smartphones [3]. Further employing the above technologies and exploiting all relevant data such as the user's profile and vital signs data can enhance contact center operations and customer service operations by partially automating the customer reception process and by providing staff with useful information in a timely fashion according to the circumstances. The main point, however, would be to develop AI that could understand not only the user's intentions but also the user's emotions and that would always create a positive impression in the user. Similarly, in the healthcare field, AI technology could be
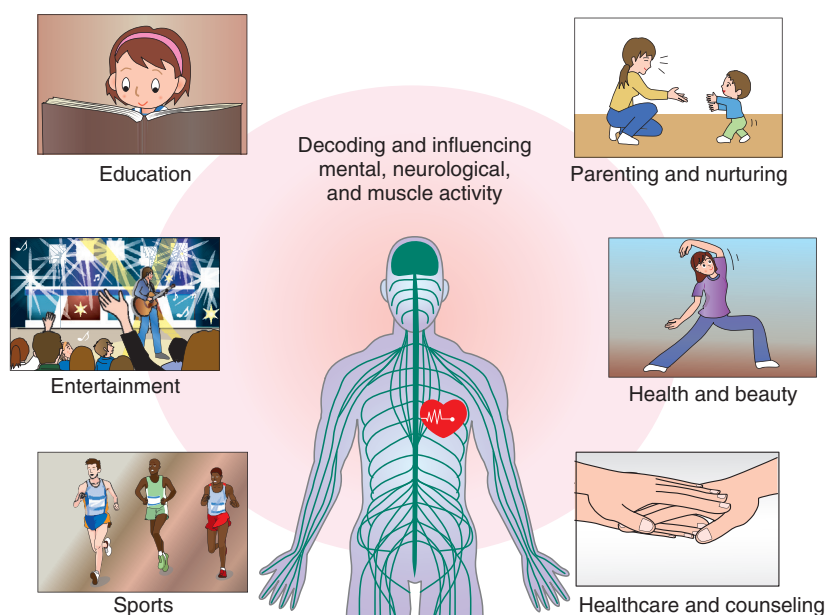
3

Fig. 3.   Heart-Touching-AI.

used to perform first-order processing of huge volumes of vital signs data, thereby reducing the doctor's diagnostics testing load. In addition, intelligent AI support could be used to ease the burden on nursing staff.

In this way, AI will come to replace or support certain human activities, but at the same time, AI and people will coexist and co-create sustainable societies by doing those tasks that each does best. This process will help to enrich people's daily lives. In the future, the aim will be to achieve an agent that can serve as a background assistant throughout a person's life. Such an agent, for example, could enhance the skills and know-how of young people in their 20s by offering specialized advice, enhance the creative skills of experienced staff in their 40s by supporting idea generation, and supplement the physical faculties and memory power of senior citizens to improve their labor force availability and achieve a sustainable workforce. The end result would be a brighter society for both younger and older generations of people.

NTT's current efforts in Agent-AI research and business development are described in separate articles in this issue [4–8].

2.2   Heart-Touching-AI

As described above, Agent-AI is a form of AI that supports people by providing them with knowledge and functions. However, no matter how advanced the knowledge possessed by AI and how complex the reasoning that it performs, there are still doubts as to whether AI can become human-like in character. Against this background, NTT aims to achieve what it calls Heart-Touching-AI (HT-AI) that further pursues human-like characteristics by understanding the mechanisms of human cognition through research in brain science (**Fig. 3**).

HT-AI does not restrict itself to intelligence. It is targeted at understanding the essential and fundamental components of a person—the intellect, instinct, and physical body—and to encourage that person and expand his or her abilities. The idea here is to break down emotional barriers between individuals and between an individual and society or the environment and to overcome physical barriers as well, and to create a comfortable human society. In other areas, too, such as sports and the arts, HT-AI is aimed at enabling anyone to experience the sensations and flashes of inspiration that only professional athletes or accomplished artists are thought to have.

Inside the human brain, the neocortex is the outermost region and the most recent to appear in the evolutionary process. This region is mostly responsible for rational and analytical thought and language functions, which is generally the pursuit of Agent-AI. Underneath the neocortex, there are the cerebral limbic system, brain stem, and cerebellum, which are older parts of the brain. These parts cannot be ignored
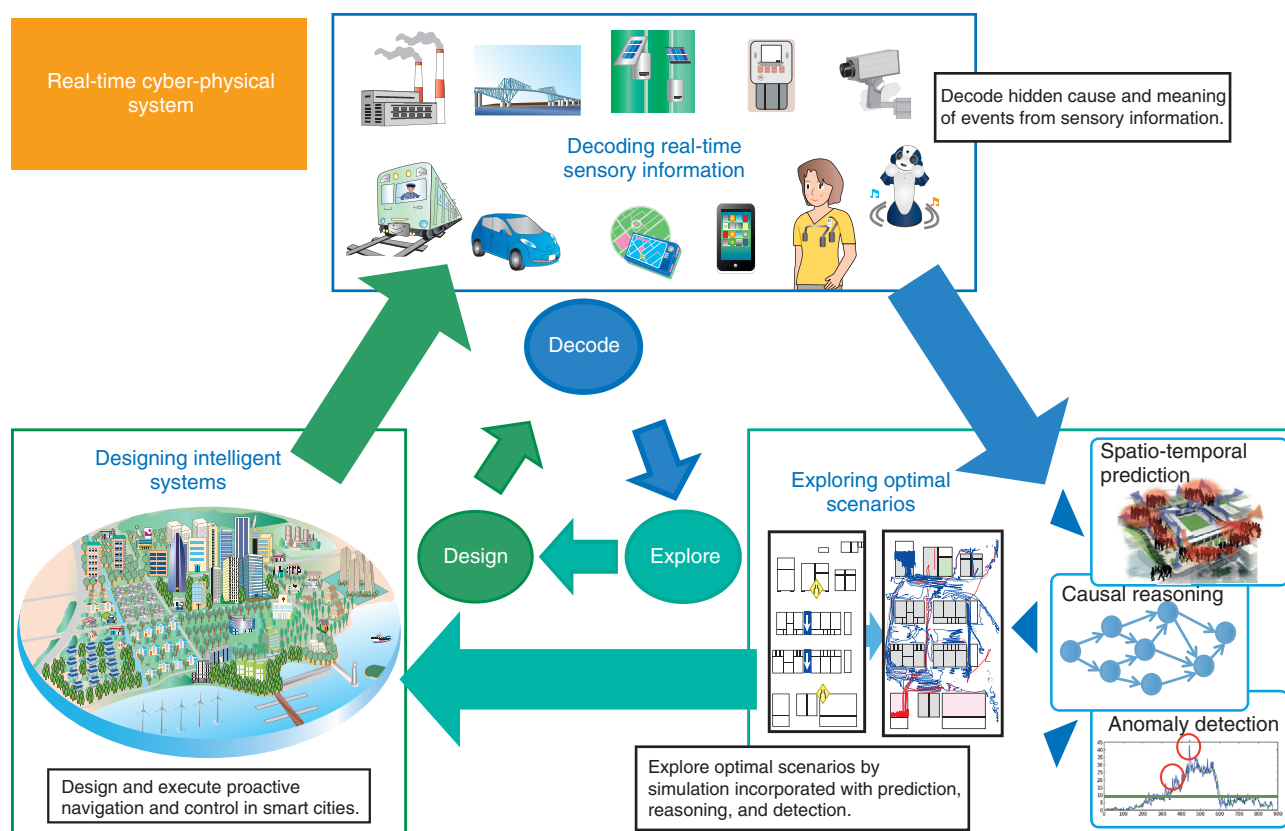
Fig. 4.   Ambient-AI.

if we are to understand or influence unconscious cognitive processing, physical movements, and human characteristics such as trust, loyalty, likes/dislikes, amiability, and motivation. These functions and mechanisms, however, are still not sufficiently understood, and measurement and evaluation techniques are still being developed.

HT-AI is therefore a form of AI that aims to decode or influence the state of human brain activity. As the name implies, HT-AI attempts to *touch the heart*. It aims to provide the individual with a comfortable existence rather than simply being useful. The road to achieving HT-AI and enabling people to enjoy its benefits is long, but NTT has already begun activities toward its realization. These efforts are described in a separate article in this issue [9].

2.3   Ambient-AI

Ambient-AI pursues *things* that have some level of intelligence and that behave in an organized manner. The word ambient is similar in meaning to atmosphere and environment. For example, ambient music

is meant to create an atmosphere in the listener's mind. In Ambient-AI, things that exist just about everywhere in the environment—that is, devices having sensors and actuators—possess intelligence. Interconnecting them in a network so that they can communicate and interface with each other will enable them to make decisions and act autonomously and to support people in decision-making (**Fig. 4**). In 2006, NTT proposed the concept of *ambient intelligence*, in which the existence of a thing with intelligence can be likened to a fairy or guardian angel "who normally remains hidden while quietly watching over a person but comes to the rescue whenever a difficulty or emergency arises in that person's life" [10]. In the ten years or so that have passed since then, the cloud, big data, and mobile communications have made great inroads in society, and significant progress has been made in IoT and machine learning. It can also be said that the elements essential to achieving ambient intelligence are finally being readied.

The ambient intelligence mentioned above includes elements of Agent-AI as a personified agent. However,
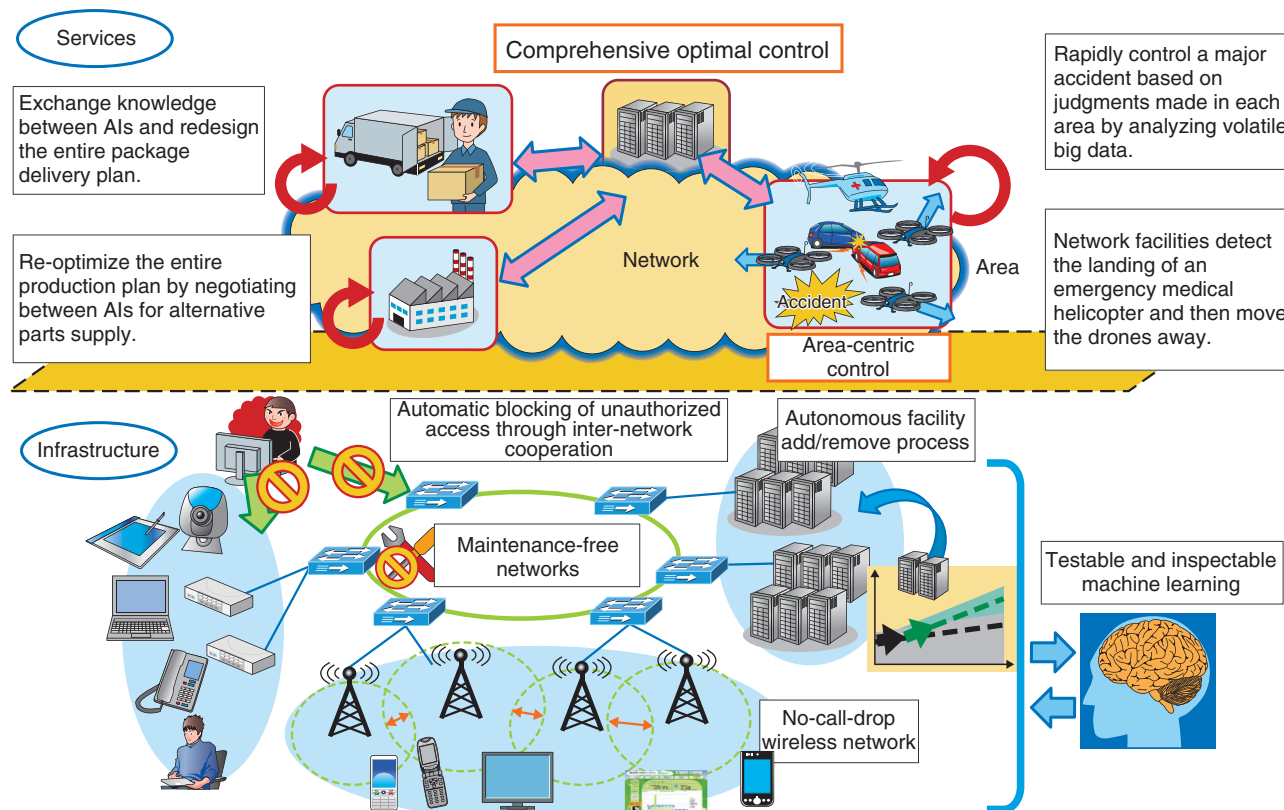
Fig. 5.   Network-AI.

the emphasis in Ambient-AI is a real-time cyber-physical system that continuously repeats the three processes of decoding hidden causes of events in the real world through intelligent real-time sensing, exploring optimal scenarios by performing prediction, reasoning, and detection using the cloud environment, and providing feedback based on an appropriate system design. In this way, Ambient-AI will be able to measure the real world (the physical component) in real time through devices and sensors situated ubiquitously in the environment. It will also be able to predict when, where, and what will occur in the near future by combining the above measurements with information on the Internet (the cyber component), to infer cause-and-effect relationships, to detect hidden signs with high accuracy, and based on simulations that incorporate the above, to search out and establish optimal scenarios and perform proactive control. NTT is researching and developing spatio-temporal multidimensional data analysis technology that models space-time relationships and predicts where and when a phenomenon will occur [11].

2.4   Network-AI

If Ambient-AI evolves even further, it can be envisioned that devices, people, and services connected to a low-latency, high-throughput network will become connected and linked as needed to all sorts of resources, thereby bringing about Network-AI, in which the network itself becomes AI. As a result, large volumes of time and space specific data (volatile big data) will be continuously generated at network terminal points and all areas covered by the network. It is important that these data be used effectively (**Fig. 5**).

However, sharing all of these data in the network—even in a high-performance network—is not efficient. It is therefore necessary to process the data first in an area-centric manner and to perform rapid execution and control based on judgments made in each area, even if it is not a 100% perfect solution (area-centric control). For example, in the event of a major disaster, we can envision how network facilities in charge of an affected area could detect the upcoming landing of an emergency medical helicopter and then activate the emergency control of drones already active in that area to rapidly move them away to avoid potential

collisions.

Next, after such processing in local areas has been completed, essential portions of that processing could be hierarchically shared over the network, enabling comprehensive optimal control. For example, we can envision how disaster information generated in certain regions could be used to redesign plans for global delivery of relief supplies. In this way, it is important that a good balance be established between area-centric control, which corresponds to processing in the cerebellum of the human brain, and comprehensive optimal control, which corresponds to the cerebrum.

In addition, it is important that a high-quality, safe, and secure network be provided in both normal and abnormal times by applying AI technology to the network. NTT is already applying machine-learning technology for early detection of signs of network failure and for operation automation and traffic prediction [12]. Future applications of AI may include automatic blocking of unauthorized access through inter-network cooperation, construction of maintenance-free networks, provision of a no-call-drop wireless network, and implementation of an autonomous facility add/remove process. New machine-learning techniques will be required to ensure the safety of these autonomous operations, in which people can test and inspect what is being learned and acquired.

3. Future outlook

This article introduced NTT's approach to AI in four key directions. These four forms of AI are in no way independent of each other but simply reflect different aspects of the broad world of AI. In any service, each of these four forms of AI will appear in one way or another, and the technologies common to them, such as machine learning, pattern recognition, and optimization, are not few in number. In addition, R&D of Agent-AI, for example, is making good progress, while that of the other forms of AI, for instance, Network-AI, is still in its early stages. In any case, NTT plans to further promote R&D in each of these four AI directions going forward. The goal is to achieve an enriching and comfortable society in which AI supports people in their daily lives and AI and people co-create sustainable societies by doing

what each does best.

References

[1] S. Araki, M. Fujimoto, T. Yoshioka, M. Delcroix, M. Espi, and T. Nakatani, "Deep Learning Based Distant-talking Speech Processing in Real-world Sound Environments," NTT Technical Review, Vol. 13, No. 11, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201511fa4.html

[2] R. Kurzweil, "The Singularity Is Near: When Humans Transcend Biology," Penguin Books, 2006.

[3] K. Onishi and T. Yoshimura, "Casual Conversation Technology Achieving Natural Dialog with Computers," NTT DOCOMO Technical Journal, Vol. 15, No. 4, pp. 16–21, 2014.
https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol15_4/vol15_4_016en.pdf

[4] Y. Matsuo, R. Higashinaka, H. Asano, and T. Makino, "Natural Language Processing Supporting Artificial Intelligence," NTT Technical Review, Vol. 14, No. 5, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa2.html

[5] T. Yamada and H. Yoshikawa, "Cloud-based Interaction Control Technologies for Robotics Integrated Development Environment (R-env™)," NTT Technical Review, Vol. 14, No. 5, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa3.html

[6] K. Ito, S. Nishido, and T. Yamazaki, "Business Transformation Using Artificial Intelligence at NTT Communications," NTT Technical Review, Vol. 14, No. 5, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa5.html

[7] O. Shirotsuka, "Artificial Intelligence Technology Development and Its Practical Use at NTT DATA," NTT Technical Review, Vol. 14, No. 5, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa6.html

[8] S. Kawamura, K. Machida, K. Matsui, D. Sakamoto, and M. Ishii, "Utilization of Artificial Intelligence in Call Centers," NTT Technical Review, Vol. 14, No. 5, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa7.html

[9] S. Furukawa, M. Yoneya, H.-I. Liao, and M. Kashino, "The Eyes as an Indicator of the Mind—A Key Element of Heart-Touching-AI," NTT Technical Review, Vol. 14, No. 5, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa4.html

[10] E. Maeda and Y. Minami, "Steps towards Ambient Intelligence," NTT Technical Review, Vol. 4, No. 1, 2006.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200601050.pdf

[11] F. Naya and H. Sawada, "From Multidimensional Mixture Data Analysis to Spatio-temporal Multidimensional Collective Data Analysis," NTT Technical Review, Vol. 14, No. 2, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201602fa2.html

[12] K. Ishibashi, T. Hayashi, and K. Shiomoto, "Improving Network Management and Operation with Machine Learning and Data Analytics," NTT Technical Review, Vol. 14, No. 2, 2016.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201602fa5.html

Takeshi Yamada
Executive Research Scientist, Research Planning Section, Machine Learning and Data Science Center, NTT Communication Science Laboratories.

He received a B.S. in mathematics from the University of Tokyo in 1988 and a Ph.D. in informatics from Kyoto University in 2003. He joined the Electrical Communication Laboratories at NTT in 1988. He was a visiting researcher at the School of Mathematical and Information Sciences, Coventry University, UK, from 1996 to 1997. He was a group leader of the Emergent Learning and Systems Research Group from 2006 to 2009 and an executive manager of Innovative Communication Laboratory from 2012 to 2013 at NTT Communication Science Laboratories. His research interests include data mining, statistical machine learning, graph visualization, metaheuristics, and combinatorial optimization. He is a senior member of the Institute of Electrical and Electronics Engineers (IEEE) and the Institute of Electronics, Information and Communication Engineers (IEICE), and a member of the Association for Computing Machinery and the Information Processing Society of Japan (IPSJ).

Satoshi Takahashi
Executive Manager, Executive Research Engineer, Supervisor, Audio, Speech and Language Media Project, NTT Media Intelligence Laboratories.
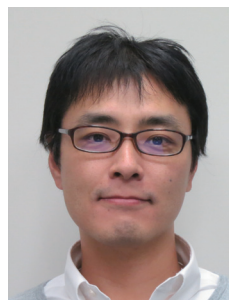
He received his B.E., M.E., and Ph.D. in information science from Waseda University, Tokyo, in 1987, 1989, and 2002. Since joining NTT in 1989, he has been engaged in research on speech recognition, spoken dialog systems, and pattern recognition. He is a member of the Acoustical Society of Japan (ASJ) and IEICE.

Futoshi Naya
Senior Research Scientist, Supervisor, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received a B.E. in electrical engineering, an M.S. in computer science, and a Ph.D. in engineering from Keio University, Kanagawa, in 1992, 1994, and 2010. He joined NTT Communication Science Laboratories in 1994. From 2003 to 2009, he was with Intelligent Robotics and Communication Laboratories, Advanced Telecommunications Research Institute International (ATR). His research interests include communication robots, sensor networks, pattern recognition, and data mining in cyber physical systems. He is a member of IEEE, the Robotics Society of Japan, the Society of Instrument and Control Engineers, and IEICE.

Takashi Ikebe
Senior researcher, NTT Network Service Systems Laboratories.

He received his B.E., M.E., and Ph.D. in engineering from the University of Electro-Communications, Tokyo, in 2000, 2002, and 2008. He joined NTT Network Service Systems Laboratories in 2002 and studied call control software, middleware, operating systems, and deployment scenarios. During 2006–2007, he was active in developing the Carrier Grade Linux specifications at OSDL (now part of Linux Foundation). He has extensive research and product experience in Linux and software-based call control systems. He received the 2006 OSDL Contribution Award. His recent research topics include IoT, device computing, and service-enable network architectures. He is a member of IEICE.

Shigeto Furukawa
Senior Research Scientist, Supervisor, Group Leader of Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received a B.E. and M.E. in environmental and sanitary engineering from Kyoto University in 1991 and 1993, and a Ph.D. in auditory perception from University of Cambridge, UK, in 1996. He conducted postdoctoral studies in the USA between 1996 and 2001. As a postdoctoral associate at Kresge Hearing Research Institute at the University of Michigan, USA, he conducted electrophysiological studies on sound localization, specifically the representation of auditory space in the auditory cortex. He joined NTT Communication Science Laboratories in 2001. Since then, he has been involved in studies on auditory-space representation in the brainstem, assessing basic hearing functions, and the salience of auditory objects or events. In addition, as the group leader of the Sensory Resonance Research Group, he is managing various projects exploring mechanisms that underlie explicit and implicit communication between individuals. He is a member of the Acoustical Society of America, ASJ (member of the Executive Council), the Association for Research in Otolaryngology, and the Japan Neuroscience Society.