National Institute of Information and Communications Technology

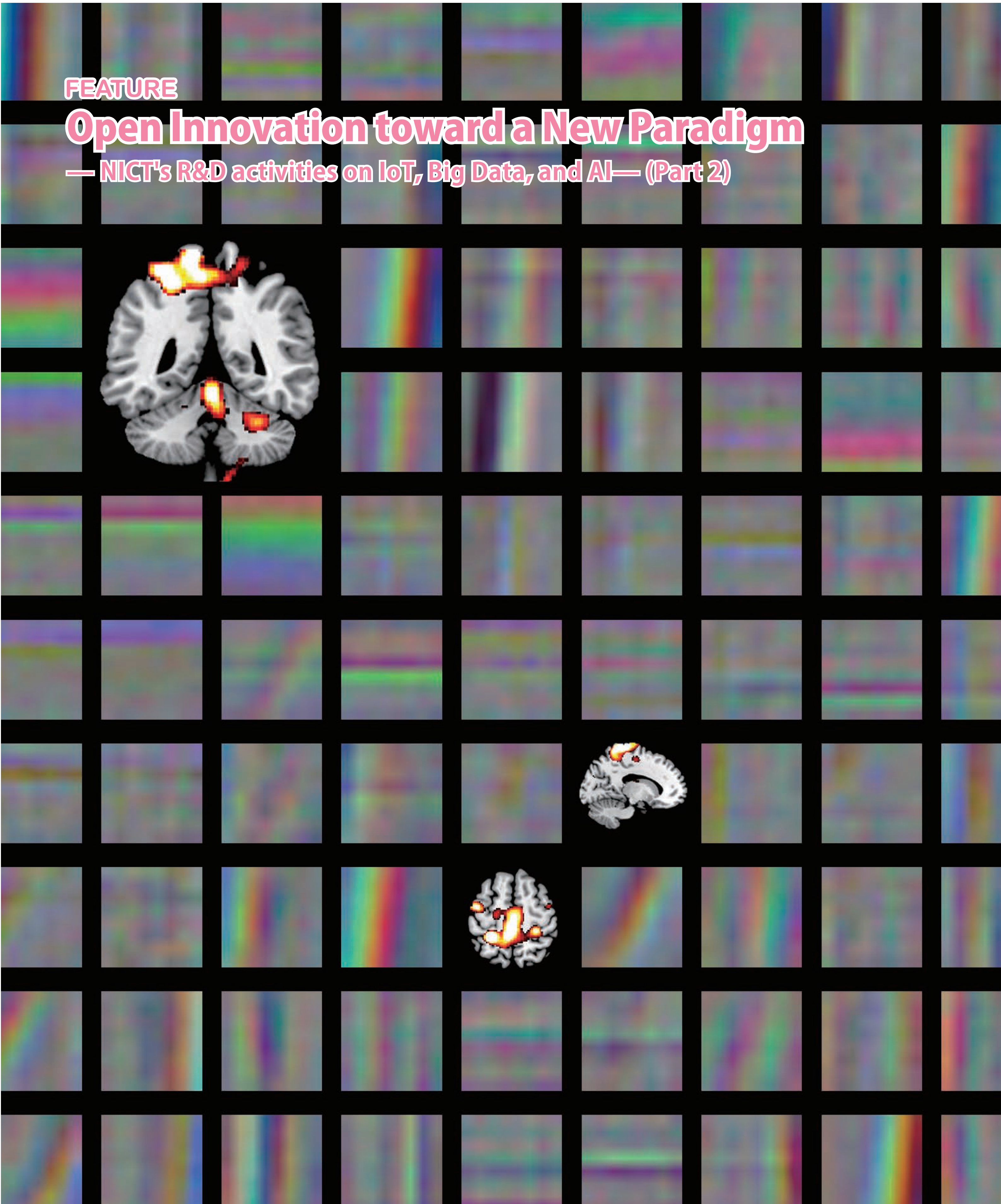
NICT NEWS

No.462 MAR 2017

FEATURE

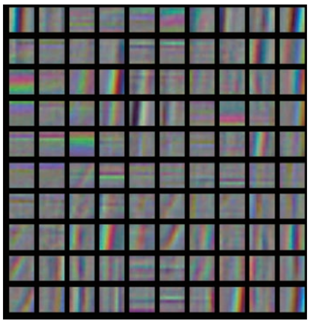
Open Innovation toward a New Paradigm

— NICT's R&D activities on IoT, Big Data, and AI— (Part 2)



NICT NEWS No.462 MAR 2017
National Institute of Information and Communications Technology

CONTENTS



Cover photo:
Each of the squares in the image is a basic component able to detect movement in video, and was obtained using a neural network modeled on the brain (developed at the Center for Information and Neural Networks (CiNet)). The cover was designed by overlaying functional Magnetic Resonance Imaging (fMRI) data, captured at CiNet, on a grid of the basic components described above (fMRI colors ranging from yellow to red indicate brain activity).

FEATURE

Open Innovation toward a New Paradigm
—NICT's R&D activities on IoT, Big Data, and AI— (Part 2)

A Special Feature on AI Research

Vice President Shinro MASHIKO

The emergence of AlphaGo in the world of Go in March last year was a great shock to those working toward innovation. An artificial intelligence (AI) able to defeat human Go players is an accomplishment that thrusts into the limelight, the speed with which AI R&D is progressing around the world. AI R&D is moving with unprecedented speed, with new technologies appearing on a two-year cycle, and being put immediately to practical use. Beyond the areas usually considered to be AI strengths—the ability to analyze data and store it efficiently as knowledge—these technologies are also able to make sensitive refinements in areas conventionally considered the realm of humans, so applications are extending into fields of art and culture. As such, to get a late start in AI R&D could present increasing risks of dangerous situations in national industry and security aspects.

For this reason, the government, and particularly the Cabinet Office, has established the Artificial Intelligence Technical Strategy Committee to promote more efficient research planning. Its goal is to bring to light important research areas and directions and create effective research strategies for this field, which does not have many researchers. This is accomplished through collaboration among three entities: the National Institute of Information and Communications Technology (NICT), RIKEN, and the National Institute of Advanced Industrial Science and Technology (AIST). As a research facility in the information and communications field, NICT recognized the importance of this field early and established ongoing R&D in various areas. As such, it has achieved more than other institutes. Specifically, we are working in technologies such as multi-lingual speech translation, social wisdom analysis, brain information processing, network control, cybersecurity, and IoT related technologies. We are creating an integrated research system as a base for all such research among industries and universities interested in NICT research results.

This special feature is the second in our series on open innovation toward new paradigm —NICT's R&D activities on IoT, Big Data, and AI—. It introduces some of the R&D on multi-lingual speech translation technology and social wisdom analysis technology being done at the Universal Communications Research Institute in Keihanna Science City, and on brain-inspired information processing technology being done at the Center for Information and Neural Networks in Suita City, Osaka. These belong to the "Creating Value" research area, as outlined in the Fourth Medium- to Long-Term Plan that NICT has been following since April of last year. Their objectives are to improve efficiency and optimize in areas of health, medicine, transportation, and living environments.

We appreciate your continued understanding, support, and cooperation in these endeavors.

Shinro MASHIKO
NICT Vice President, Member of the Board of Directors
Joined the Ministry of Posts and Telecommunications Communications Research Laboratory (currently NICT) in April 1989. Previously Executive Director of the Strategic Planning Department and Senior Executive Director of NICT. Appointed to his current position in April 2013. Ph.D. (Engineering).

Data-driven Intelligent System Research Center (DIRECT)

Kentaro TORISAWA
Director General
Data-driven Intelligent System Research Center

Kentaro TORISAWA became an assistant professor at the University of Tokyo before obtaining his Ph.D., and then was appointed associate professor at the Japan Advanced Institute of Science and Technology (JAIST). Since 2008, he has been working at the National Institute of Information and Communication Technology (NICT), currently serving as the director general of the Data-Driven Intelligent Research Center (DIRECT). Engaged in Natural language processing research and recipient of JSPS and other awards. Ph.D. (Science).

The Data-driven Intelligent System Research Center (DIRECT) is a new research center established in April 2016. It is evolved from the former research group, the Information Analysis Laboratory. R&D on intelligent systems that process natural languages such as Japanese is the Center's research mission.

State of natural language processing research

Natural language processing research began in the 1950s as a research area of artificial intelligence (AI) technology, and development of practical applications has accelerated since 2010. There are two major reasons for this. The first is that machine learning technology has matured and the second is the availability of large volumes of text data from sources such as the Web.

Before mature machine learning technology was available, people had to write programs for analyzing the meaning of language. It was extremely difficult to write programs to analyze natural languages, with their huge vocabularies and exception-riddled syntax.

However, by applying machine learning technology to natural language processing, it is possible to make computers learn how to analyze natural language automatically, by preparing concrete examples of text and the results of semantic analysis, and using them as learning data. In other words, the computer can be made to create a program that analyzes language automatically. This does not mean that by simply preparing the training data, the program is created automatically at the touch of a button. There are many possibilities with machine learning technology, and an infinite number of ways to apply it, so when implementing a particular type of analysis, deciding an appropriate machine learning technology and how to apply are often very difficult problems. This situation does not fundamentally change, even when using the “Deep Learning” technology that has been receiving attention lately. Creating the required training data also involves work, with humans writing concrete examples of analysis results, so making this process more efficient is another major issue. These issues have been a main focus of natural language processing research in recent years.

Today there are also billions of documents that can be obtained through the Web and other sources. By analyzing these using computers,

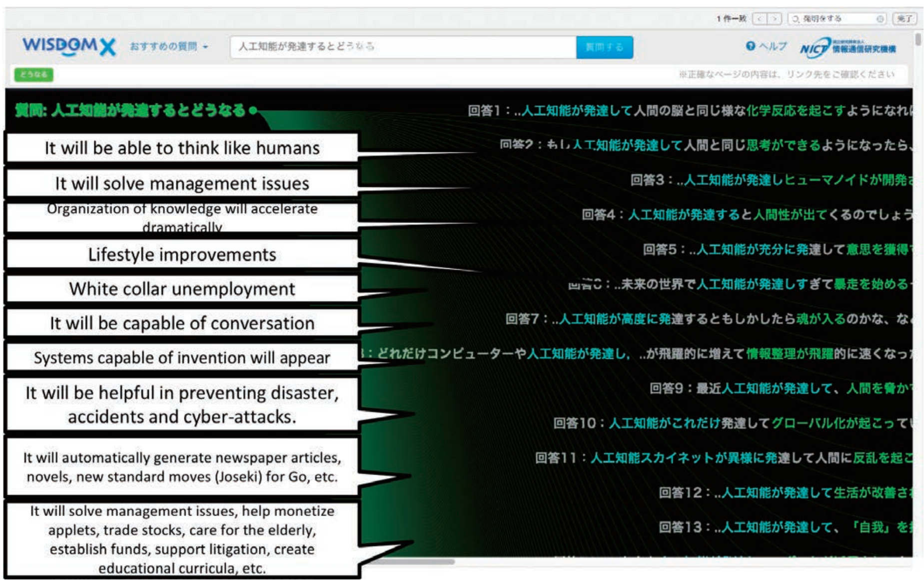


Figure 1　Answers to the question "What will happen if AI develops further?" from WISDOM X

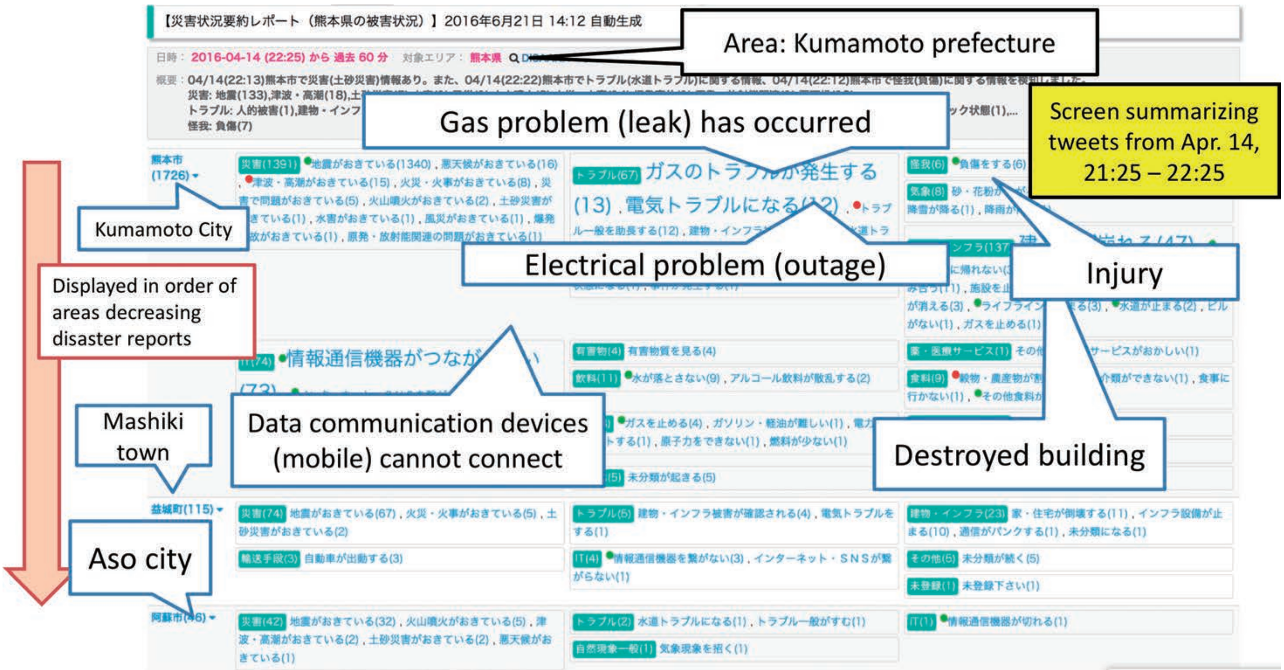


Figure 2　Results of using D-SUMM to summarize Twitter data in the one hour following the 2016 Kumamoto Earthquake

an amount of knowledge not comprehensible by any one person can be obtained automatically, enabling provision of knowledge that is useful to people in activities ranging from every-day to business. Thus, a very broad range of natural language processing applications can be realized, but there are still basic issues that have not been solved, such as exactly what knowledge or wisdom is useful to people, and how can this useful knowledge be presented in efficient and user-friendly ways? These are research themes that are now starting to attract much closer attention.

Achievements

DIRECT and its predecessor, the Information Analysis Laboratory, have already developed a natural language processing technology using the vast amounts of text data on the Web as a starting point, large amounts of diverse training data, various machine learning technologies, and huge knowledge bases and dictionaries built using a combination of machine learning and manual checking. Results of this include WISDOM X (http://wisdomnict.jp/), a system that analyzes the large amount of information currently published on the Web, DISAANA (http://disaana.jp/), a system that analyzes SNS disaster-prevention information, and D-SUMM (http://disaana.jp/d-summ/), which can summarize disaster conditions.

WISDOM X: A question answering system with features that are unique in the world. It can answer questions in Japanese based on the massive information source of some four billion Web pages, can present hypotheses that were not in the original Web pages, and can also propose other questions that it can answer (Figure 1).

DISAANA: A system that can answer questions regarding disaster related information appearing on Twitter during a disaster. It was used by the Cabinet Secretariat to analyze needs in disaster-hit areas during the Kumamoto Earthquake in April 2016.

D-SUMM: A system published in October 2016 that is able to automatically summarize and present disaster related information originating from a specific local government that appears on Twitter. It holds promise for helping to understand an overall picture of a disaster quickly and efficiently (Figure 2).

We are also developing platform technologies such as the RaSC middleware (published at https://alaginrc.nict.go.jp/rasc/ja/), which enables these three systems to run easily and efficiently on clusters of several hundred machines.

Future prospects

In the future, we will deploy and improve these technologies, such as deep learning, in society through collaboration with private enterprise, including human resource sharing. We will also continue R&D to achieve advanced technologies, including technology to provide knowledge acquired from WISDOM X and other systems to users through a natural language conversation system, and technology to infer various useful hypothesis not in the input text.

Societal Implementation of Speech Technologies Leading up to 2020, and Advanced, Basic-Technology Research to Clear a Way to the Future

Hisashi KAWAI

Director of the Advanced Speech Technology Laboratory, Advanced Speech Translation Research and Development Promotion Center

After completing a doctoral course, he joined KDD R&D Laboratories (currently, KDDI Research) in 1989, where he was engaged in research on speech synthesis and recognition. He worked for ATR Spoken Language Translation Research Laboratories from 2000 to 2004 and for NICT from 2009 to 2011. He has been working on research and development of speech-to-speech translation at NICT since 2015. Ph.D. (Engineering)

The Advanced Speech Technology Laboratory conducts speech technology research to support communication using speech, which is a natural and ubiquitous medium for people, and to advance the exchange of information between people and machines. As an intermediate goal, we are conducting R&D on practical speech recognition and synthesis technology for ten languages (Japanese, English, Chinese, Korean, Thai, Indonesian, Vietnamese, Burmese, Spanish, and French), and a societal implementation of multi-lingual speech translation technology at the Tokyo 2020 Olympic and Paralympic Games.

Speech recognition and speech synthesis technology has made great advances recently through use of artificial intelligence (AI) techniques. Below, we introduce our speech recognition technology research for our societal implementation of multi-lingual speech translation. We also discuss research on topics that look at 2021 and beyond, including technology for speech recognition in real environments, which transcribes all kinds of speech content in our world into text, and a situated spoken dialogue technology, which enables effective dialog based on surrounding conditions and context.

Speech recognition using AI technology

Deep learning is an effective learning technology that uses neural networks having a deep structure (Deep Neural Networks, or DNNs) and has attracted much attention recently in research on AI technology. Great advances in speech recognition performance are being made using DNN and deep learning.

Generally, speech recognition is composed of an acoustic model, a pronunciation dictionary, and a language model, as shown in Figure 1. DNNs have been applied to the acoustic model, which converts speech feature parameters into phonemes (the smallest phonetic unit), and the language model, which predicts succeeding words from current words and converts to sentences.

We have placed particular emphasis on research to build a highly accurate acoustic model for speech recognition using DNN. Specifically, we are implementing practical speech recognition that is both computationally inexpensive and highly accurate, by incorporating simple, accurate methods for mapping feature parameters to phonemes, including chronological analysis with bidirectional recurrent neural networks and connectionist temporal classification. The acoustic model learning technology developed for speech recognition can also be applied in speech synthesis re-

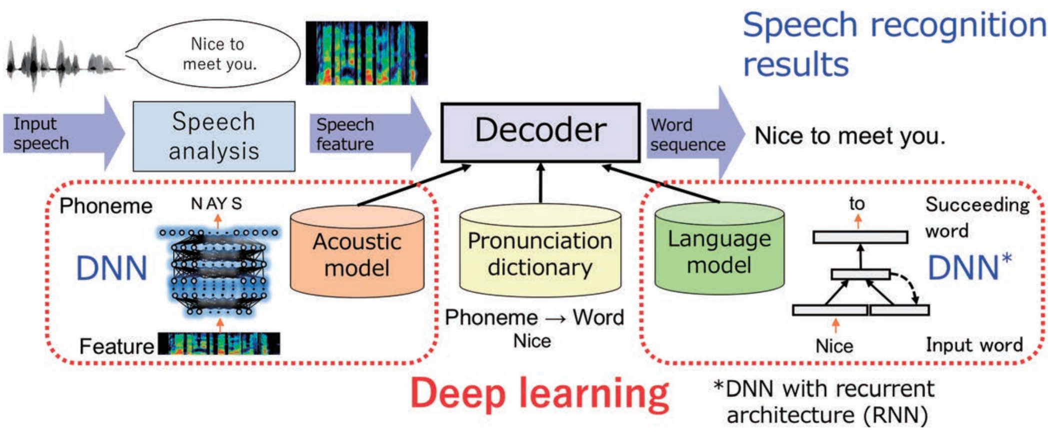


Figure 1 Speech recognition using neural networks with deep structure (DNNs)

search. We are using AI techniques in speech technology research, with speech recognition and speech synthesis augmenting each other, and advancing these technologies toward our societal implementation of multi-lingual speech translation.

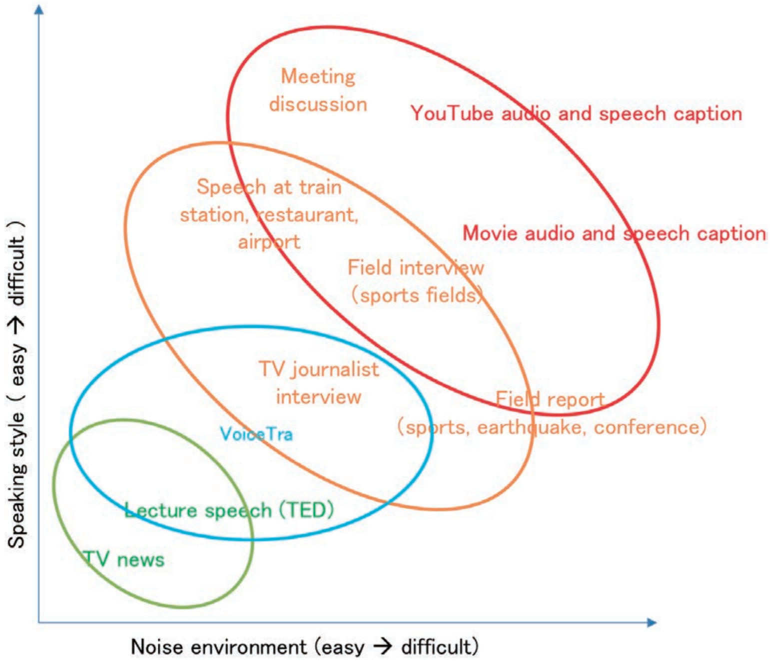


Figure 2　Main types of dialog in real environments and speech recognition difficulties

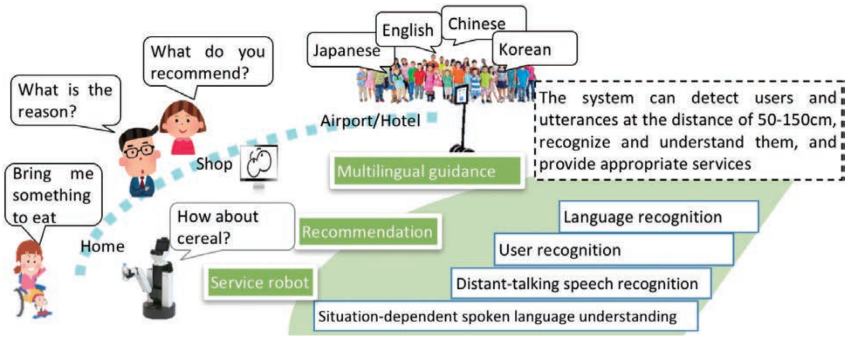


Figure 3　Service applying situated spoken dialogue technology and background technologies

Technology for speech recognition in real environments

As mentioned, development of deep learning has led to noticeable improvements in performance of speech recognition technology, and some research institutes have reported performance equal or better than a human transcriber. However, such speech recognition performance results are obtained under conditions where aspects such as noise and speaking style are controlled. Under uncontrolled situations, speech recognition performance deteriorates seriously. We have defined speech under these conditions as *real environment speech* and are conducting R&D to improve the accuracy of real environment speech recognition. Figure 2 shows the main types of real environment speech and the difficulty of speech recognition. For example, for interviews in live TV reporting, speech recognition can be difficult due to the adverse influence of background noise and distant speech, speaking style may not be clear or grammatical, and the form could change, with ambiguous statements and informal expressions.

This is the background for our research on speech recognition in real environments, examining how to transcribe all kinds of speech content in our world into text, and how to improve recognition accuracy for uncontrolled speech, which is usually difficult. To solve these problems, we are making basic improvements to advance the acoustic model using deep learning, and also conducting research to identify acoustic information other than speech, to understand conditions of the utterance and process accordingly, in what we call *situation-dependent speech recognition*.

Situated spoken dialogue technologies

We are also conducting R&D on more advanced speech technologies for future society. One of the main societal issues in the future will be an increasing need for lifestyle support in an aging society and there will be a shortage of workers to provide it. Service dogs are currently carrying some of this load, but there are issues that are fundamentally difficult to solve such as taking care of the dog and the cost of training. Recently, communicative service robots have begun to appear, but they currently only understand user's speech within predetermined scripts. This is not adequate to provide practical lifestyle support.

This has prompted the idea that a communicative service robot, able to provide support comparable to that of a service dog, could solve both the need for lifestyle support and the shortage of workers to provide it. Toward implementation of such a robot, we have built a research platform for robust spoken dialogues including advanced speech recognition and synthesis technologies. We have built a cloud robotics platform for spoken dialogues, rospeex, which has been used by 40,000 unique users for various spoken dialogue tasks. In the future, we will extend the work and conduct research to build technologies needed for the practical services shown in Figure 3. These include situation understanding technology, capable of estimating the positions and roles of speakers in multi-party conversations, and situated spoken dialogue technology, capable of spoken language understanding dependent on the speaker and real-world conditions, and recommendation dialogues based on user preferences.

Future prospects

Our laboratory will continue to focus our efforts on developing speech technologies for societal implementation in 2020, and also on nurturing seeds of long-term research for the world in 2021 and beyond.

Research and Development of AI-Based Automated Translation Technologies (Toward 2020 and Beyond)

Eiichiro SUMITA

NICT Fellow
Associate Director General
Advanced Speech Translation Research and Development Promotion Center (ASTREC) / Director of Advanced Translation Technology Laboratory, ASTREC

Eiichiro SUMITA received an M.S. in computer science in 1982 and a Ph.D. in engineering in 1999. Before joining NICT, he worked for ATR and IBM. His research interests include machine translation and e-Learning.
Ph.D. (Engineering).

The Advanced Translation Technology Laboratory is conducting R&D on automated translation technologies based on AI, to overcome the Japanese language as a barrier to communication with foreigners. This article introduces (1) the scope of our R&D toward a societal implementation by 2020, and longer term basic research on (2) the bilingual data gathering ecosystem and (3) simultaneous interpretation technologies.

Societal implementation of automated translation using AI technologies by 2020

Automated translation using AI technology went through a large paradigm shift to methods utilizing "Big" bilingual data*1 in the 1990s, before Big Data was a hot topic. Till the 1980s, rule-based translation methods were common, based on rules to convert between the different vocabulary and structure of the two languages, while preserving meaning. These methods reached a limit, impeded by the difficulty of controlling interactions among the rules.

In 1988, a method that infers translation from bilingual data, called *statistical translation*, was proposed, and has gradually developed. Collection of bilingual data, which is the basis of the method, has progressed and the accuracy of translations has increased dramatically.

Deep neural networks (DNN) are neural networks that have a deep hierarchical structure and are widely used in AI. In recent years, they have started to be used for automated translation through renewed use of distributed representation, which uses real vectors of high dimension to represent words. This is referred to as *neural translation*.

Both statistical translation and neural translation can be used to build automated translation systems from bilingual data and are suitable for extension to many languages. Our laboratory is collaborating with the Advanced Speech Technology Laboratory to fulfill the Global Communication Plan,*2 which is implementing multi-lingual speech translation technology in society by 2020. To do so, accurate, automated translation among ten languages*3 is being implemented using statistical and neural translation, is published as the VoiceTra*4 smartphone application, and is continuously being improved. Research is also advancing on adapting it to other fields such as medical care.

It should be noted that these methods have

English source	The actuator according to claim 1, wherein an even number of notches are formed in said body, and the displacement of said rod in the axial direction is extracted.	
Break into structural parts using English patterns	PREA	the actuator according to claim 1
	TRAP	wherein
	PURP	an even number of notches are formed in said body, and the displacement of said rod in the axial direction is extracted
Generate Japanese patterns corresponding to the English patterns	PURP	an even number of notches are formed in said body, and the displacement of said rod in the axial direction is extracted
	TRAP	wherein
	PREA	the actuator according to claim 1
Translate each part from English to Japanese to complete the translation	PURP	偶数個の切込みが形成されている前記本体であり、前記ロッドの変位には、軸方向 を抽出する
	TRAP	ことを特徴とする
	PREA	請求項 1 に記載のアクチュエータ

Figure 1　Rule-based translation for patent applications

*1 Also called a corpus.
*2 http://www.soumu.go.jp/main_content/000285578.pdf
*3 Japanese, English, Chinese, Korean, Thai, Indonesian, Vietnamese, Burmese, Spanish, and French.
*4 http://voicetra.nict.go.jp/

both advantages and disadvantages. For example, statistical translation is quite unstable for translating synonyms, but neural translation can handle them. Statistical translation can be built using ordinary CPUs, but neural translation requires special hardware called a GPU. Our laboratory considers a hybrid of these two methods as promising, and our effort is focused there.

The Claims section of a patent is the most important and can contain very long sentences that neither statistical nor neural translation can translate well. However, accurate translation can be implemented by including rule-based methods. In superficial terms, a sentence can be partitioned into three parts, the PREA, appearing before "wherein," the TRAP, which is the "wherein" itself, and the PURP, appearing after the "wherein," as in Figure 1. This is then rearranged into the Japanese order, of PURP, TRAP, and PREA. Finally, the individual parts are translated by the automated translation system and assembled to get the required Japanese translation. This system*5 has been released and widely distributed in private industry.

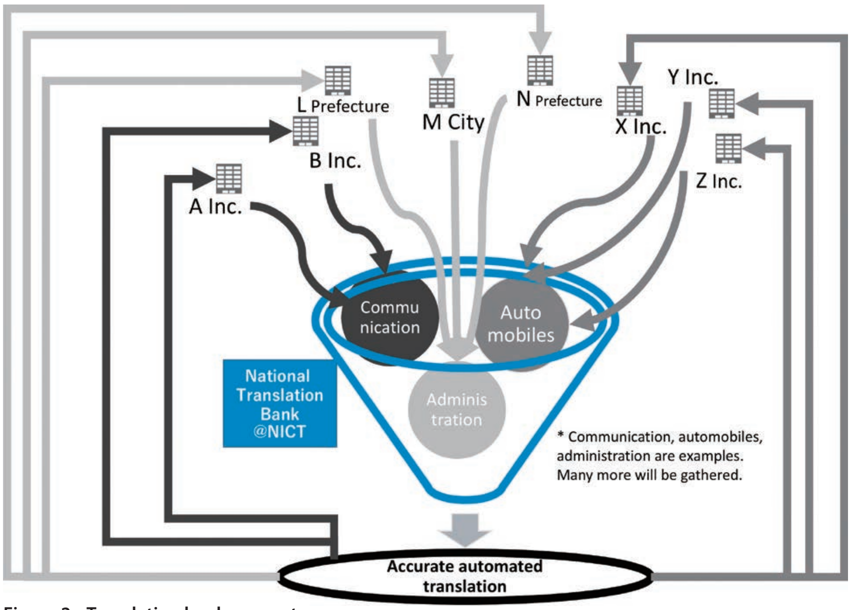


Figure 2 Translation bank concept

Figure 3 Simultaneous interpretation prototype screen shot

The bilingual data ecosystem

Statistical translation and neural translation both automatically improve in performance as the amount of bilingual data increases, so it is important to gather bilingual data efficiently. Our laboratory has built a large scale bilingual corpus using patent data provided by the Japan Patent Office (350 million Japanese-English sentences over 20 years) and realized accurate automated translation.

Translation data are scattered through various organizations, but if it could be gathered at NICT, a public institution, and used to build a translation system (referred to as "Translation Bank" below), we could achieve highly accurate translation systems for various fields, such as automotive manufacturing, telecommunications, and government administration, as shown in Figure 2.

If accurate automated translation is implemented, the cost of translation should drop and the amount of translation that can be added to the translation bank should increase, leading to a continuous positive cycle of improving translation accuracy.

The ultimate goal of simultaneous interpretation

Current speech translation technology processes speech after hearing entire sections, so for long inputs, there is a large delay before the translation reaches the user. To avoid this, translation must start before the end of the sentence is received, as is done by simultaneous interpreters.

Simultaneous interpretation from Japanese into English is particularly difficult. The basic word order in Japanese is subject, object, verb (SOV), with the critical V at the end. The basic word order for English is SVO, with the V after the subject. To output the English V, it is necessary to wait till the V at the end of the Japanese sentence. Sentences cannot be translated before hearing the end of the sentence, but simultaneous interpreters anticipate what the speaker is saying to decide their translation.

Our laboratory is tackling this problem, and has implemented a prototype that finds break-points in an uninterrupted input stream to begin translation. A screen shot is shown in Figure 3, with the recognized English word shown above in red, the English words up to a break-point shown below in red, and the result of automated translation, a string of Japanese words, shown below on a yellow background. It is only a first step, but it shows the significance of establishing simultaneous interpretation technology.

Our laboratory will continue to focus our efforts on advancing automated translation technology over the medium and long term, as discussed above.

*5 https://mt-auto-minhon-mlt.ucri.jgn-x.jp/

Semi-Supervised Learning for the Next Generation of Human-Like Artificial Intelligence

Takashi SHINOZAKI
Researcher
Brain Networks and Communication Laboratory
Center for Information and Neural Networks

After completing his doctoral course, he became a Special Postdoctoral Researcher at Riken BSI and a postdoctoral researcher in New York University, before taking his current position in 2010. He is engaged in computational neuroscience, EEG, MEG, and artificial intelligence. Ph.D. (Science).

Deep learning is a neural information processing technique that recently has developed rapidly, become a core technology in AI innovation, and is being applied in new fields that conventionally could only be handled by humans. However, unlike the brain, deep learning requires large amounts of training data with the associated correct results. This can be a significant barrier for applications in fields such as medical care. The Center for Information and Neural Networks (CiNet) is conducting research to solve this problem by implementing learning that is closer to how the brain learns.

Background

Deep learning is a machine learning technique that uses neural networks with many layers, simulating human sensory information processing. It achieves dramatically better performance than earlier AI techniques, so applications in various fields have advanced rapidly in recent years. Versatile applications such as voice search systems on smartphones, keyword search of photographs uploaded to the cloud, and even multi-lingual translation are bringing a new ICT innovation into our daily lives.

Current deep learning techniques use what is called *supervised learning* in the machine learning field. Correspondence relations are learned from "labeled data" consisting of, for example, a photograph of an apple associated with the label, "apple," that has been checked for correctness. However, a problem with supervised learning is that a very large amount of data are required. For example, to learn general image recognition, more than a million labeled data items are used. This amount of data is not necessarily available in all fields and it can be very difficult to increase the number of data items available, particularly in fields such as medical care, so the technique is difficult to apply.

Flexible learning in the brain

How then, does the brain learn? To know what something is when we look at it, we have not learned by looking at more than a million images, checking what they are one-by-one (and for deep learning, this must be repeated dozens of times!). Learning in the brain rarely makes use of such correct data, and most of the time it learns by just looking. In the machine learning field, this is called *unsupervised learning*. It is equivalent to what is called clustering in data analysis, in which data are classified based on their statistical properties. For this type of statistical analysis, techniques such as Independent Component Analysis (ICA) are often used, but controlling the resulting classification can be difficult, resulting in difficulties. On the contrary, brain learning uses unsupervised learning combined with a minimal amount of supervised learning to subtly control the various data classifications and achieve efficient learning. In this way, most basic things are learned automatically from the environment, and everything else is learned with a minimal number of questions, resulting in very effective learning (Figure).

Semi-supervised learning

In the field of machine learning, combining supervised and unsupervised learning in this way is called *semi-supervised learning*. It has been researched in various ways for a long time, and is of course, being studied around the world with respect to deep learning as well. However, the unsupervised and supervised learning are done in completely separate learning processes in much of this research, and it is difficult to achieve the type of learning in the brain, where knowledge is accumulated and revised at the same time. We have developed a new technique able to apply unsupervised and supervised learning seamlessly to exactly the same network. This technique is called a Convolutional Neural Network (CNN). When implemented on the most widespread deep learning network and a handwritten digit

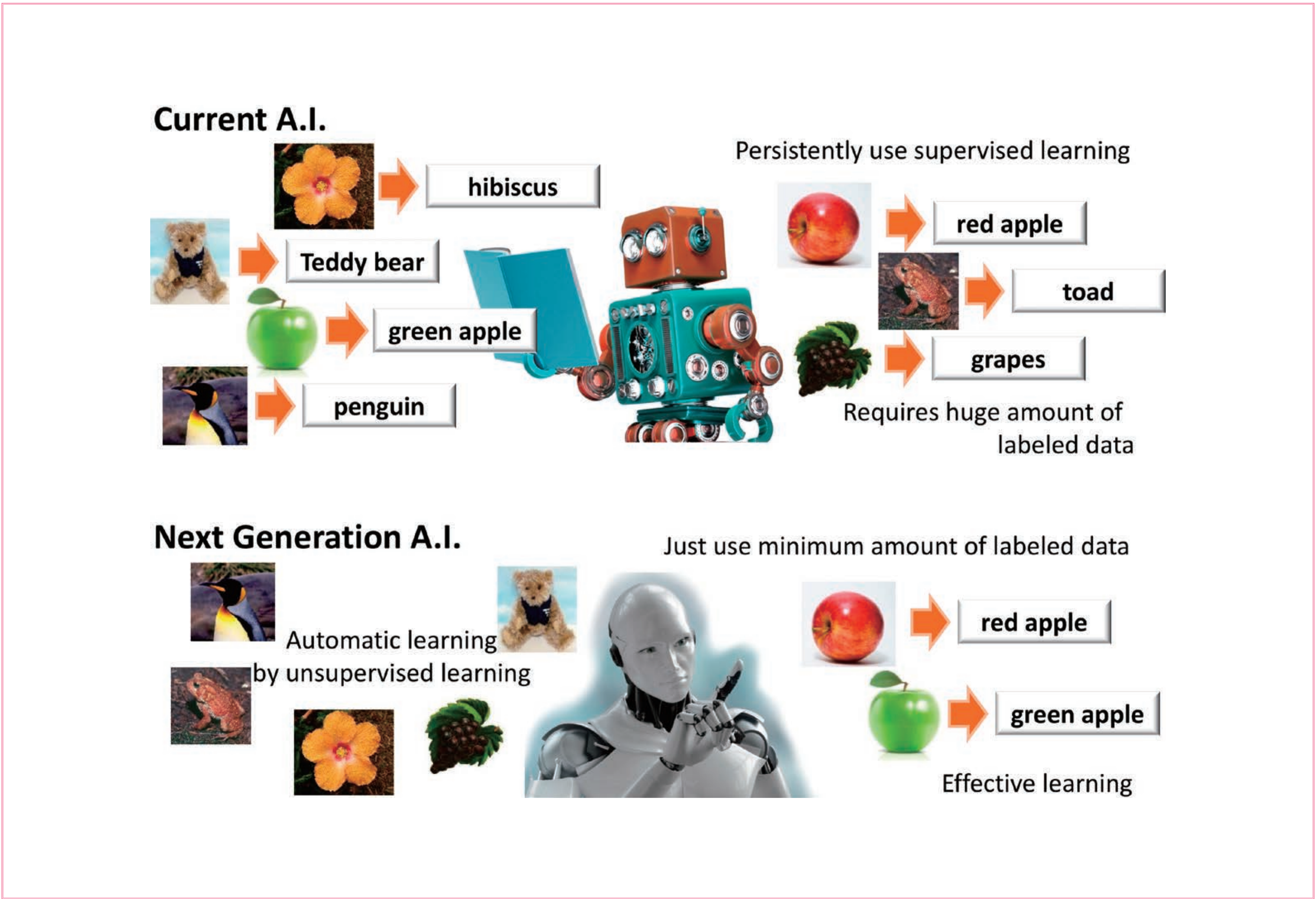


Figure: Differences between current and next generation AI

recognition task using the MNIST data set* performed, our method achieved learning to a level equivalent to earlier methods using only approximately one quarter the amount of labeled data.

Future prospects

Our method has currently only be tested using a limited data set, so we intend to test it further using natural images and other generalized data. This will make it possible to use deep learning in areas that were previously difficult, such as with medical or IoT data, so we anticipate a range of new applications. On the other hand, the semi-supervised learning that we have implemented still only achieves a fraction of the learning functionality of the brain. As such, we will continue to utilize the latest neuroscientific knowledge from CiNet and integrate techniques such as the powerful self-organizing capabilities of competitive learning, which is used in the Neocognitron from Professor Kunihiko Fukushima, formerly of Osaka University. In doing so, we work to develop next-generation AI that is capable of learning more like the brain.

＊ http://yann.lecun.com/exdb/mnist/

Understanding and Realizing Intelligence by Learning Brain Function

Toward Next Generation AI Systems

Michiaki IWAZUME
Research Executive Director, Center for Information and Neural Networks
Director of Preparation Office for Integrated Promotion of AI Research Strategic Program Produce Office, Social Innovation Unit, Open Innovation Promotion Headquarters

After completing a doctoral course, he held positions as a JSPS post-doctoral researcher, an assistant professor at Kinki University and research scientist at the Brain Science Institute, RIKEN. He joined NICT in 2005. After a temporary transfer to the Cabinet Office, he was appointed to his current position in 2011. He has been studying knowledge-based systems and web information processing. Ph.D. (Engineering).

Recently, society's interest and expectations of AI technology have increased and we are experiencing a third "AI boom." A direct cause of this is development of a type of machine learning called deep learning, which is learning by neural networks having deep structure that mimics structures in the human brain. Deep learning exhibits performance dramatically better than prior machine learning techniques for problems including image and speech recognition, and it has sent shock waves through society by defeating world-class players in the game of Go, a feat which experts expected would take another ten years.

As an indirect cause, improvements in high-performance computing technology, such as the GPGPU, have enabled an IoT environment able to generate and process Big Data in real time on smartphones and also network connected sensors and other devices.

In Japan, the government has also established an Artificial Intelligence Technical Strategy Committee with collaboration among the Ministry of Internal Affairs and Communications (MIC), the Ministry of Education, Culture, Sports, Science and Technology (MEXT), and the Ministry of Economy, Trade and Industry (METI), to promote research, development, and implementation in society, of AI technologies. The Center for Information and Neural Networks (CiNet), together with the Universal Communication Research Institute, its Data-driven Intelligent System Research Center and the Advanced Speech Translation Research and development promotion Center (ASTREC), are participating in this government initiative and are raising expectations for the next generation of AI technology, largely inspired by the brain.

Limitations of current AI technology and brain inspired approaches

As mentioned above, deep learning is responsible for much of the current AI boom, but it still is not a panacea, as it assumes that a large amount of good quality training data (that have been labeled correctly beforehand) is available. Since these data require labeling by people, it is not practical, in terms of cost, to prepare such data for all application areas. Then, after a model has completed deep learning, considerable computing resources, a data center to support them, power and other infrastructure on a large scale are needed to put it to practical use. This limits the range of possible applications.

Humans, on the other hand, can learn from

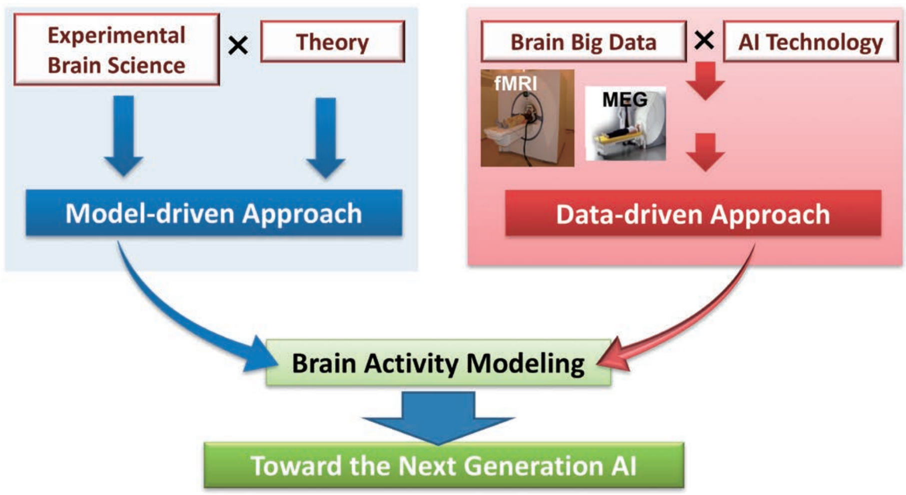


Figure 1 Complementary Integration of Model-driven and Data-driven Approaches

just a few seconds of experience and can acquire knowledge, such as languages, without explicit training. They can also apply knowledge from one field to other fields, and the brain, as a computing resource, consumes very little power. Such an advanced, flexible intelligence is still difficult to realize with machine learning and other current AI technologies. It is CiNet's goal to realize even more advanced AI technologies using approaches inspired by the human brain, which is the only embodiment of such an intelligence that we have.

Deciphering the brain using AI technology

To learn from the brain, we must decipher its functionality from large amounts of brain-activity data, create an engineering model of that functionality, reproduce it on computing devices, and analyze its characteristics. For this sort of data-driven R&D, building a large-scale database of brain activity is imperative (Figure 1). One would expect that with equipment to measure brain activity, such as functional Magnetic Resonance Imaging (fMRI) and magnetoencephalography (MEG), we could build brain Big Data that would be useful for next-generation AI research. This is not the case, however. CiNet and research institutes around the world have been collecting brain activity data, but most of these data are extremely inconsistent, with biases in dimensionality, measurement time and number of samples in the data, and wide-ranging or even lost attributes such as test conditions, age and gender of subjects.

Considering the issues with current AI technology identified earlier, it is a bit of a paradox, but AI technology, and machine learning in particular, could be a useful tool for deciphering this extremely complex brain-activity data. To do so will require methods to derive and model relationships among this very-high dimensional, heterogeneous brain-activity data (with differing or lost attributes), stimulus data such as images, and movement data.

Machine learning technology could also infer associations within data measured and collected under differing conditions, from subjects having neural and psychological ailments, and in many different medical facilities. This could then be used to support tasks such as early detection of diseases or selecting optimal treatment plans in a wide range of fields, beyond R&D on next-generation AI technology.

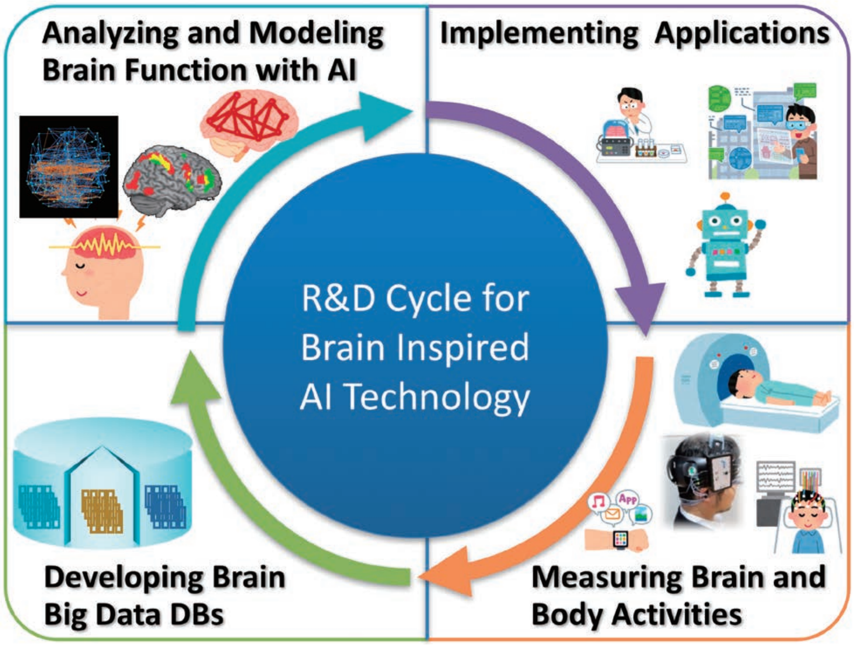


Figure 2 R&D cycle for brain inspired AI technology

till 2020	till 2030	till 2040
Developing Brain Big Data		
[Phase 1] AI learns brain (Brain function analysis by AI)	[Phase 2] AI imitates brain high-order functions	[Phase 3] AI imitates whole-brain (AI system for whole-brain architecture)
[Expected outcomes] ・Brain decoding technology and its applications ・Brain biomarker ・BMI etc.	[Expected outcomes] ・Innovative machine learning (Post deep learning) ・Neuromorphic devices ・AI based mobile-edge computing etc.	[Expected outcomes] ・Whole-brain neuromorphic system ・AI-driven autonomous vehicles, robots, IoT and ICT Infrastructures

Figure 3 R&D Roadmap

This sort of R&D cycle—measuring brain and body activity, accumulating brain Big Data, analyzing and modeling it using AI, developing application systems, and deploying them in society—is essential for developing and enhancing brain-inspired AI technology (Figure 2).

Future prospects

As the IoT spreads, it is expected that *edge computing*, with processing done closer to users at the "edges" of the network, will become increasingly important (Figure 3), in contrast to how processing is currently being centralized in the cloud. Edge computing has the benefits of reducing network communication costs and improving real-time response of ICT services, but is disadvantageous in conserving computing resources and power, with extra specifications such as user friendliness and flexibility also needed. For these types of requirements, next-generation brain-like AI technology, capable of low-power, highly flexible intelligent processing, is very attractive.

The future is very promising for innovation with AI technology, not only in robots and self-driving cars, but in devices with various actuators, built into services, and utilizing Japan's strengths in manufacturing. Continuing to accumulate knowledge about the brain from data, steadily and from a broad perspective, will be essential for R&D on the next generation of brain-inspired AI.

TOPICS
NICT Intellectual Property -Series 2-

A new method to correct chromatic aberration
—Observation of structures in living cells—

With the improvement of microscope performance, we developed a technology to observe living cells stained with various colors. The technology we developed is the 'chromatic aberration correction method.' Specifically, high resolution was realized by using this original technology to correct the chromatic aberration of image data obtained with an optical microscope. If we can more accurately obtain information on the relative positions of intracellular structures, researchers in a wide range of fields will benefit and may be able to make remarkable discoveries.

■Outline of the technology and its fields of application

The image taken by the camera contains many colors (wavelengths), and the differences between these colors shift the focus of the image. This phenomenon is called chromatic aberration. The technology we developed uses image processing to correct the shifts in focus caused by the differences between colors, but without any artificial fiducial markers for calibration. To correct these shifts, we used the leakage light from the color with the shortest wavelength and photographed a reference image in all other color channels. The reference image thus looked exactly the same in all color channels. From this image we sought the optimum solution for several variables, such as translation on each axis, the rotation angle, and the magnification of each axis. The resulting parameters are applied to a target image that is acquired separately for each color by using the corresponding excitation light. The resolution of a super-resolution microscopy is on the order of 30 to 120 nm (nanometers). By the above techniques, we are able to correct the chromatic aberration of three-dimensional images within a few nm. Of note, we can obtain even higher resolution by electron microscopy; however, since the object to be measured must be exposed to a vacuum state, the cell will die and cannot be observed alive.

■Search for uses, applications, and research collaborators

Microscopy with which we can keep living cells alive and observe them would lead to significant discoveries for a wide range of fields in life science, and it would advance new research fields. We have been working with a corporation that manufactures microscopes. First of all, it is essential that we promote practical applications of the microscope, being aware of the users who would most want to use this microscope in their research. For that reason, we must have universities, which we anticipate will be users of this microscope, participate in the joint research. If we can imagine areas where this technology's superiority could be harnessed or would be crucial—the kind of fields it would be useful in and the purposes for which it could be used these fields—then the research partners we should work with or the licensing candidates we should look for will become clear. At the Intellectual Property Promotion Office we intend to advance the practical application of these research results in cooperation with researchers.

If you are interested in this technology, please contact us below. Software lending is also possible.

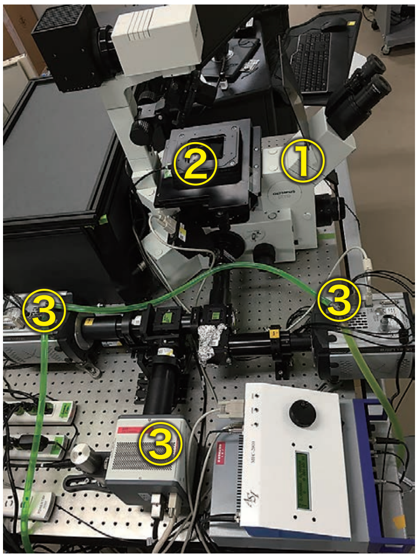
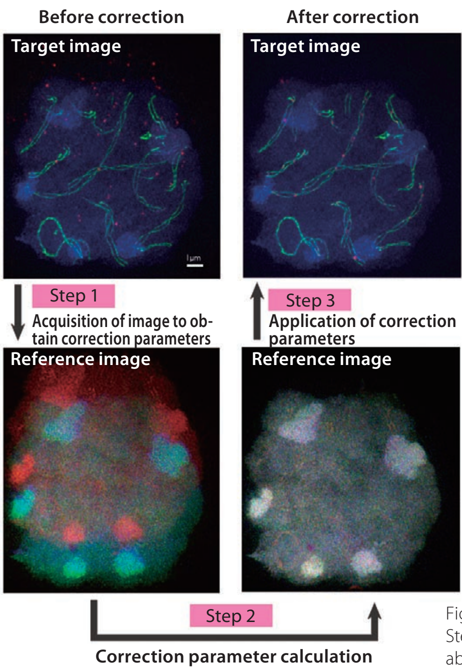


Figure 1
Microscope setup for chromatic aberration correction
1) Microscope
2) Sample
3) Camera



In the image before correction, the positions of the green line and the red dot are shifted, and this is corrected.

Figure 2
Steps in chromatic aberration correction

<Patent Information>
Publication No.: Patent Pub 2015-216495
Name of invention: A new method to correct chromatic aberration

<Contact (Inquiries,etc.)>
Intellectual Property Promotion Office, Innovation Promotion Department
E-mail: ippo@ml.nict.go.jp
TEL: +81-42-327-6950 FAX: +81-42-327-6659

Development of and technical transfer of a next-generation human musculoskeletal model 'Def Muscle'

Masaya HIRASHIMA
Senior Researcher, Brain Networks and Communication Laboratory,
Center for Information and Neural Networks

Development of a versatile, PC-operated musculoskeletal model

There are hundreds of muscles in our body, and the brain controls them well. To clarify the mechanism by which the brain moves the body, we must know the positions of these muscles in relation to each other. However, there is no human body model that can accurately express the positions of the muscles. In conventional human body models, the muscles are simplified as a straight line or a polygonal line without volume. So the muscles are buried in the bones, and muscles that should originally make up the surface layer are buried in the deep muscles. Thus unnatural situations sometimes occur, creating a major barrier in all areas dealing with human movements.

To fundamentally solve this problem, I have been working on development of a new type of musculoskeletal model. This model takes into account deformation due to muscle volume (size and shape) and interference (collision). Calculating the volume deformation is quite expensive. And so, we adopted the GPU parallel computing method, which has developed rapidly in recent years, and tried to solve this expensive calculation problem. We had GPU installed on a personal computer and succeeded in operating a musculoskeletal model with the PC. Without this approach, a costly supercomputer would have been necessary to create such a model (See Figure).

The GPU-based computing platform that we developed is a versatile technology that can be used for applications such as human motion simulation, motion load analysis, animation, etc. Currently, this platform can be used only around the shoulder at the present time, but if the muscle shape can be created and set up, it will also be possible for it to handle the movement of the lower extremities and the trunk, and thus the motion of other animals.

Its fields of application are quite wide, ranging from not only neurology but to biomechanics areas such as to orthopedics, rehabilitation, and sports, and further to education and entertainment. With the cooperation of the Intellectual Property Promotion Office, this technology will be transferred to 3D Co., Ltd., a company that specializes in 3DCG (three dimensional computer graphics). The technology will be sold as "Def Muscle— Next Generation Musculoskeletal Model Development Tool."

As a result of the NICT press release in September 2016, we had inquiries from organizations such as automobile companies and orthopedic hospitals, and we now go out more often with 3D Co. Ltd. staff members to explain this model to organizations and the public. We are also participating in academic conferences. In 2016 we exhibited 'Def Muscle' at the Japanese Society for Motor Control and at the System Integration Division of Society of Instrument and Control Engineers. From March 20-24, 2017, as an IT business exhibit, we plan to join CeBIT2017, the largest scale technology exhibition in the world, at Hanover in Germany. In the future, while getting useful information from users at exhibitions, we will work to improve this model's performance, and we intend to make efforts to establish it as a standard in this field.

Figure: A shoulder area model created with Def Muscle

Def Muscle is a group of function libraries that simulate three-dimensional changes in muscle. Users by themselves must prepare the basic shape data of the muscles underlying the simulation. This figure is the result of a simulation with Def Muscle based on human body shape data provided by BodyParts3D*. The musculoskeletal model created in this way can be used for simulating the kind of skeletal motion that occurs when a motion command is sent from the brain to muscle (forward dynamics calculation). From skeleton motion data measured by motion capture, this model can also be used to estimate the muscle shape and strength at that time (through inverse dynamics calculation).

* http://lifesciencedb.jp/bp3d/info/

Masaya HIRASHIMA
After finishing his graduate school doctoral course and earning his Ph.D., he was an Assistant Professor of the Graduate School of Education at The University of Tokyo. He joined NICT in 2014. He is engaged in research on motion control / motor learning and biomechanics.

NICT NEWS No.462 MAR 2017
Published by
Public Relations Department,
National Institute of Information and Communications Technology
<NICT NEWS URL> http://www.nict.go.jp/en/data/nict-news/

4-2-1 Nukui-Kitamachi, Koganei, Tokyo
184-8795, Japan
TEL: +81-42-327-5392 FAX: +81-42-327-7587
E-mail: publicity@nict.go.jp
URL: http://www.nict.go.jp/en/
Twitter: @NICT_Publicity

ISSN 2187-4050 (Online)