NICT NEWS

National Institute of Information and Communications Technology

No.6
2022
Vol.496

FEATURE

Special Issue on AI

Foreword

TEZUKA Osamu and AI?

N

Appears in
NICT PR video entitled
"A Future with N"

START A SCOUTER

the AI robot character N
comming from
from the year 203X,
developed by NICT.

CONTENTS

NICT NEWS 2022 No.6 Vol.496
National Institute of Information and Communications Technology

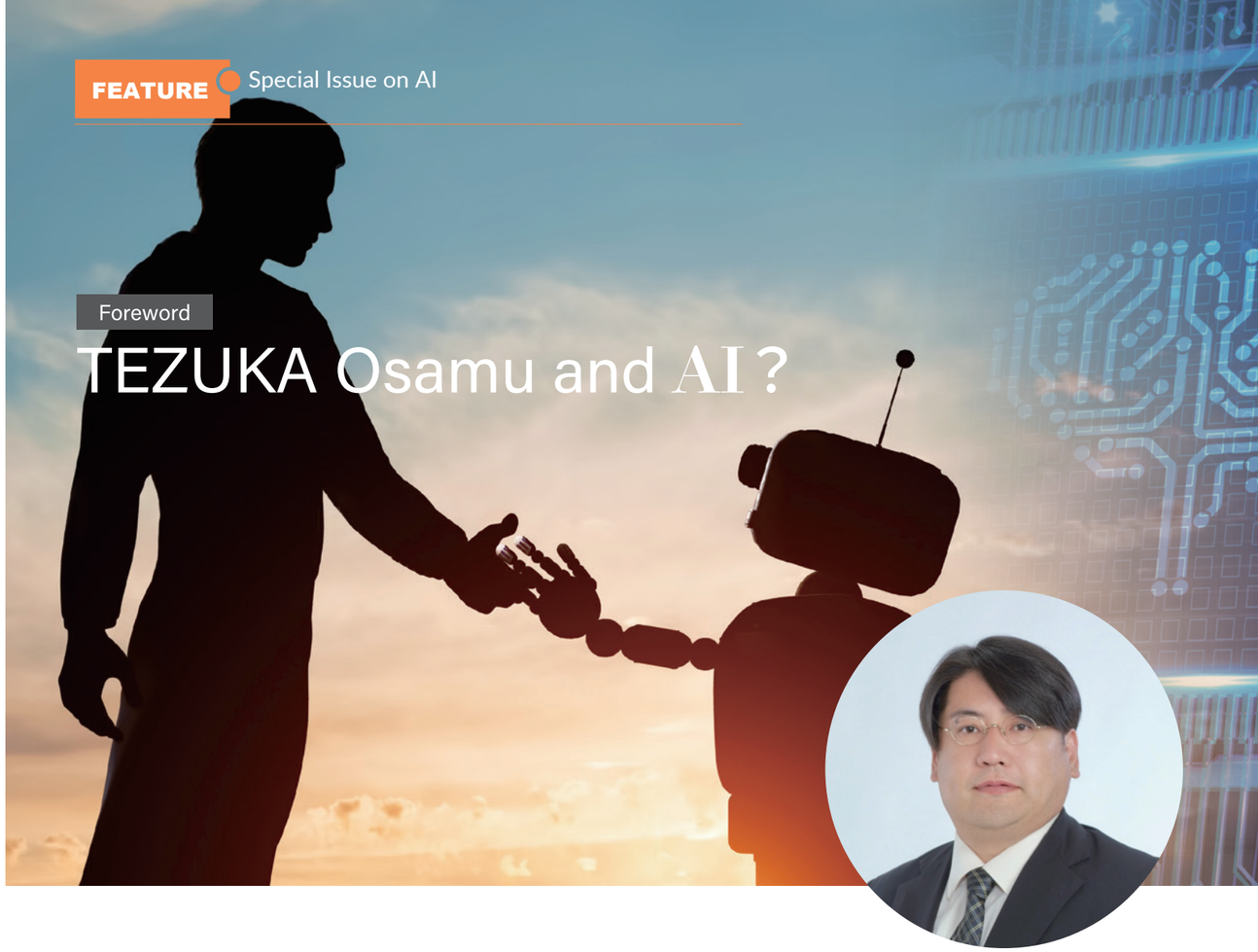
FEATURE

Special Issue on AI

Foreword

INFORMATION

Cover Photo:
Computer Lab of the Universal Communication Research Institute and "N," official mascot of NICT
NICT's basic research on deep learning using powerful computers has helped drive various other AI research projects.
N is an imaginary futuristic AI robot that serves as NICT's mascot.



Foreword

TEZUKA Osamu and AI?

TORISAWA Kentaro

NICT Fellow /
Associate Director General, Universal Communication Research Institute /
Distinguished Researcher, Data-driven Intelligent System Research Center (DIRECT), Universal Communication Research Institute

Kentaro Torisawa joined NICT in 2008. After serving as the director general of DIRECT, he became an NICT Fellow in 2020 and is currently serving a distinguished researcher in DIRECT. He has been studying natural language processing and has received many awards and grants including the JSPS prize and Twitter Data Grants.

Artificial intelligence (AI) has been attracting a great deal of attention in recent years. AI technologies are evolving rapidly, owing to newly developed techniques such as deep learning. This NICT News issue features various AI-related R&D at NICT and efforts to put R&D products into practical use. This includes AI-based forecasting of space weather to protect ICT infrastructures from solar flares and other cosmic events, AI network control, AI capable of addressing the sensitive social issue of handling personal information confidentially, the VoiceTra multilingual speech translation application which is already available to the public, and collaboration between brain science and AI.

I have been researching natural language processing—a subfield of AI—for nearly 30 years. As an AI expert, I'm deeply impressed by the recent explosive emergence of so many amazing AI techniques, which are beginning to be applied to various social issues.

Efforts to solve social issues using AI are often hindered by a wide range of new, distinct social issues. That is to say, AI technologies are advancing so rapidly that their capabilities are beginning to surpass social expectations and, in many cases, have begun to outstrip systematic efforts to properly promote and control them, including legislation. I think it is important for future AI researchers and developers to carefully consider potential social issues that may obstruct the development of new AI technologies. When I was a boy, I was captivated by the stories of Osamu Tezuka, one of the most influential Japanese manga artists, ranging of early work like Metropolis to his later work, such as Message to Adolf. Some of the issues I have encountered doing AI research remind me of issues raised in Tezuka comics like Astro Boy and Phoenix more than 50 years ago. In the following, I would like to describe a few of such issues that I encountered during the development of our spoken dialogue system.

Since 2018, DIRECT has been developing MICSUS—a multimodal dialog system for the elderly—in collaboration with KDDI Corporation, NEC Solution Innovators, Ltd. and the Japan Research Institute, Ltd. This project has been supported by the Cabinet Office's Cross-ministerial Strategic Innovation Promotion Program (SIP). MICSUS's front end device is a stuffed *shiba inu* dog (Figure) equipped with a microphone but

Foreword

TEZUKA Osamu and AI?

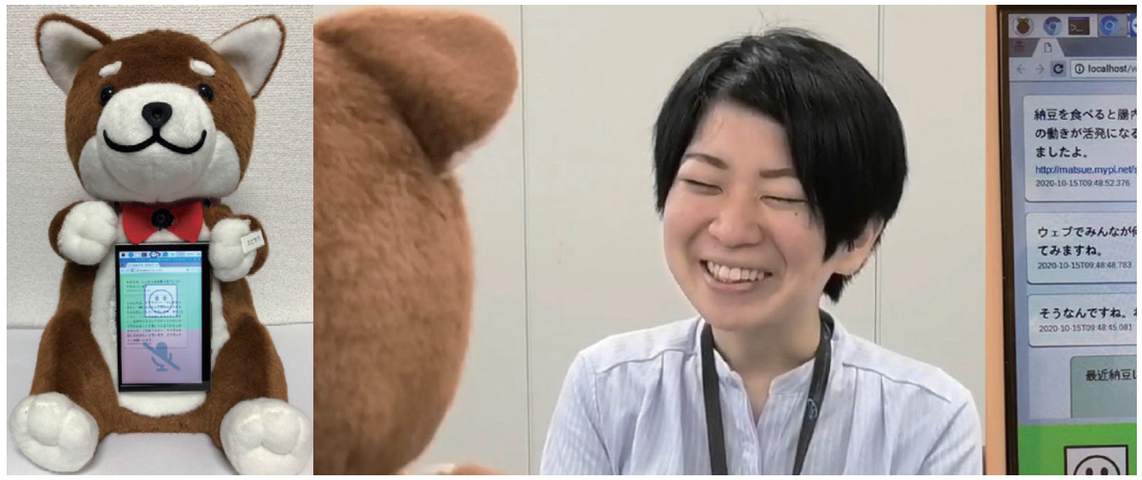


Figure Shiba inu dog-like MICSUS device engaging in conversation with a person (an NICT engineer). Dialogue records are displayed on the device's display (right).
This technology was developed jointly by NICT, KDDI Corporation, NEC Solution Innovators, Ltd. and the Japan Research Institute, Ltd. with the support of the Cabinet Office's Cross-ministerial Strategic Innovation Promotion Program (SIP).

also a speaker and a camera to enable the device to recognize emotions from the facial expressions of users. It is connected to an AI cloud containing various deep learning technologies.

MICSUS has two purposes. First, it collects information concerning the health of its elderly users through multi-modal dialogs. The collected information is used in making so-called care plans, which specify what kind of rehabilitation or nursing care should be administered to the users. This health monitoring, called "nursing care monitoring" has conventionally been performed monthly by a nursing care administrator called a care manager. Each care manager is tasked with the heavy burden of monitoring dozens of elderly people. MICSUS is designed to reduce their workload automating some of their monitoring responsibilities. Furthermore, MICSUS can monitor elderly users daily by engaging in conversation with them, potentially enhancing care quality. We are currently testing the usability of MICSUS by allowing elderly people to use it in their homes. We ran an experiment in which four participants communicated with MICSUS for 15 consecutive days. During this time period MICSUS asked approximately 400 questions about their health condition and received verbal responses. We found that, MICSUS equipped with our original deep learning technology, MICSUS was able to correctly interpret these responses with an accuracy of approximately 96%. Many of the erroneous interpretations were attributed to ambiguous answers from users, which were difficult to understand even for humans. These results indicate that MICSUS has achieved sufficient accuracy for practical use. An actual dialogue between an elderly user and MICSUS can be viewed on YouTube (https://www.youtube.com/watch?v=bRgo31EoIMg, at around 3 minutes and 39 seconds into the video).

The second purpose of MICSUS is to ease social isolation and resulting health deterioration among the elderly. Research has found that socially isolated elderly people face a 50% greater risk of developing dementia and requiring nursing care over a 10-year period than those who interact with others on a daily basis.

(Source: Masashige Saito, Katsunori Kondo, Toshiyuki Ojima and Hiroshi Hirai. Criteria for social isolation based on associations with health indicators among older people: a 10-year follow-up of the Aichi Gerontological Evaluation Study. Japanese Journal of Public Health. 2015. 62(3): 95–105. DOI: 10.11236/jph.62.3_95)

MICSUS is intended to compensate for elderly peoples' lack of social interaction by engaging in casual conversations in between nursing care monitoring sessions, using information from the internet.

During casual conversation with its users, MICSUS provides useful information from the internet about everyday life based on their conversations with MICSUS. When MICSUS determines that a user has said something that can be developed into a casual conversation, it automatically generates questions related to benefits or risks associated with what the user said. For example, if an elderly person says, "I like tomatoes," MICSUS may generate questions such as "What are the benefits of liking tomatoes?" or, "What are good ways of using tomatoes?" These questions are then forwarded to WISDOM X, system capable of answering questions using information from the internet and deep learning (accessible at https://www.wisdom-nict.jp/). WISDOM X finds answers to MICSUS-generated questions from the internet (e.g., heating tomatoes allows more efficient absorption of lycopene) and selects the best answer using deep learning. Finally, the device conveys this answer to the user after slightly modifying it to fit the context of the conversation.

Through these casual conversations, we hope to offer elderly people more frequent chances for communication, making their lives more stimulating and healthier. During the experimental use of MICSUS by elderly people mentioned above, it engaged in casual conversation approximately 160 times. Roughly 90% of the responses delivered by MICSUS during these conversations were judged appropriate. In addition, nearly half (47%) of the appropriate responses made the elderly person either smile or respond positively (e.g., by saying "I will try what you suggested"). Unlike similar existing systems that give simple preprogramed responses to users (e.g., by nodding or repeating the user's speech), we are attempting to create a system capable of providing them with new, appropriate information, although this is a great challenge. The fact that nearly half of MICSUS users smiled or responded positively indicates that MICSUS is able to engage in high-quality casual conversation with them.

Osamu Tezuka's comic, Phoenix, has a robotic character named Robita who has an electronic brain in which human memories have been implanted. Robita sometimes exhibits human-like behavior. For example, using the memories of a young child implanted in his brain, Robita teaches new games to young children, making it popular among them. He also makes goofy mistakes like humans. I had Robita's human-like qualities in mind when developing MICSUS. Its technology makes casual conversation using the internet—a massive repository of collective human memory. My motivation in developing MICSUS was to cheer up elderly people, especially those living alone, by providing them with new, life-enhancing information derived from human memories.

Tezuka seemed to favor technologies with human-like attributes, like Robita. In the real world, however, efforts to create interactive systems that emulates humans will require dealing with various social issues. The first is the need to comply with copyright laws related to gathering information from the internet. We resolved this issue by consulting with lawyers and making the casual conversation function of MICSUS as legitimate as existing search engines. We designed MICSUS to display its search results (i.e., the URLs of source websites and the extracted snippets*1 used in casual conversation) on the device's display in a clearly identifiable manner. In addition, MICSUS often cautions users about the risk unreliable information on the internet and avoids responding to user questions related to serious matters, such as medical issues.

Some MICSUS components were intentionally designed not to employ deep learning. The latest neural networks can hold human-like conversations even without using search engines or question answering systems connected to the internet. However, responses generated by such technologies sometimes contain groundless, inappropriate information. It is therefore still unclear whether full-scale implementation of such technologies will be widely accepted. Also, such technologies seem to be capable of only casual conversation without any underlying mission, like nursing care monitoring. According to these observations, we chose to create rather conventional but reliable dialog scenarios for the health monitoring, (i.e., part of the MICSUS's dialogs), while using deep learning technologies for semantic interpretation of users' speech. The scenarios were developed based on check lists for elderly people health check created by elderly care professionals.

Similar to neural networks, humans also often make meaningless and inappropriate remarks and fail to accomplish the objectives of dialogs. As Tezuka suggested through his Robita character, these errors can also be seen as human-like behaviors. Moreover, we often felt during our usability testing that the participating elderly people were looking for human-like qualities in MICSUS. Introducing AI technologies into nursing care—a sector requiring close human interaction—is indispensable amid the shrinking labor force caused by the declining birthrate. Future AI developers—particularly those pursuing the development of AI technologies with human-like attributes, like MICSUS—should seriously consider how user-friendly AI technology can balance attractive human-like qualities with the risk of errors associated with them.

Also, it may be necessary in the future for the public to change their perception of human-like AI, due to its potential to resolve many serious social issues. Despite continued technological evolution, the risk of AI technology making errors cannot be completely eliminated. When humans make mistakes that affect others, they apologize, admit responsibility, take recurrence prevention measures and resume their duties, if the mistakes are not too serious. On the other hand, apologies from AI technologies are practically meaningless. However, if AI acquires the ability to learn from its mistakes and improve itself, such apologies may potentially become socially acceptable and recurrence prevention measures may be seen as a positive solution. This approach to AI development may lead to the production of more properly behaved, socially acceptable AI technologies which can be used to resolve a wide range of social issues. I recall that Astro Boy also frequently made mistakes and apologized. I sometimes wonder whether the greatest obstacle to the development of social problem-solving AI technologies is the public perception that AI must be completely error-proof.

*1 **Snippet**: the original meaning of this word is a fragment of text. In the context of this report, it means short passages extracted from webpages using a search engine. MICSUS displays both the snippets extracted and the URLs of the source webpages.

*2 **Neural network**: a computing system composed of numerous interconnected artificial neurons that mimic human brain cells. Although individual artificial neurons can perform only simple calculations, when a large number are connected in a precise manner, they can handle complex processing after being trained with data. By drastically increasing the number of interconnected artificial neurons in a neural network, today's deep learning technology is capable of sophisticated, high-precision processing needed for such tasks as image recognition and semantic interpretation of natural languages, which were previously achievable only by humans.

The Future of Space Weather Forecasts

World's First Operational Solar Flare AI Forecasting Model

NISHIZUKA Naoto
Senior Researcher,
Space Environment Laboratory,
Radio Propagation Research Center,
Radio Research Institute

After completing graduate school, he worked at JAXA, NAOJ and UCL/MSSL. He joined NICT in 2014. He has been engaged in R&D related to solar flare prediction models using machine learning. He holds a Ph.D. in Science.

A recent report issued by the Ministry of Internal Affairs and Communications warns that solar flares can disrupt Earth's magnetic field, interrupting mobile phone services intermittently for up to two weeks. Although the sun sustains life on Earth, solar flares—intense localized eruptions of electromagnetic radiation on its surface—sometimes produce powerful storms in the heliosphere with the potential to damage infrastructure vital to communications and GPS navigation on Earth. At a meeting on space weather forecasting hosted by the Ministry in 2022, a panel of experts from the industrial, public and academic sectors discussed responses to solar flare-caused disasters. The next solar activity peak is predicted to occur around 2025, making the creation of more prompt and accurate solar flare forecasting systems an urgent issue.

Social impact of Space Weather

X-rays, ultraviolet emissions, high energy particles and hot coronal gas released by solar flares sometimes affect our lives in various ways by disrupting radio communications and broadcasting, compromising the accuracy of our car navigation systems and map applications and causing satellite failures and blackouts (Figure 1). To minimize these impacts, NICT has been forecasting space weather daily and publicly releasing its forecasts, including levels of solar flare activity and geomagnetic and ionospheric storms. In addition, since 2019, NICT has been working 24/7 to provide up-to-date space weather information to the International Civil Aviation Organization (ICAO).

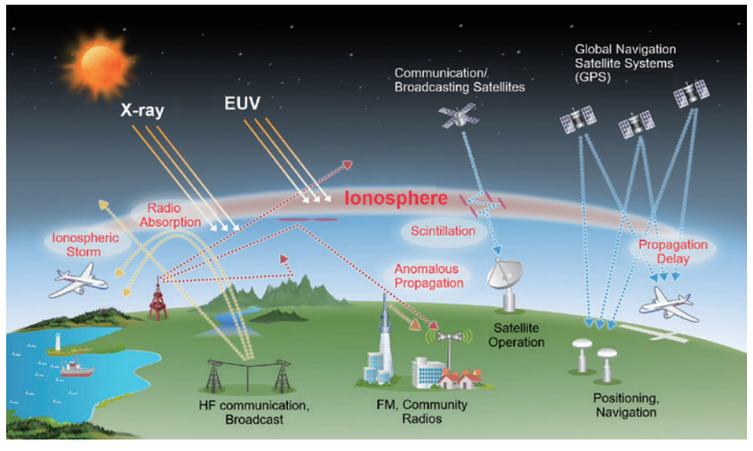


Figure 1 Social impact of space weather (solar flares)

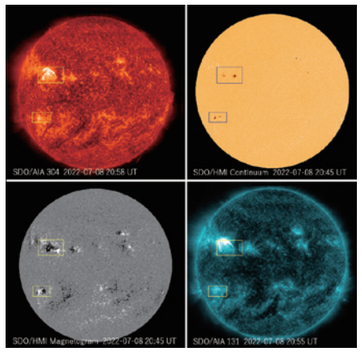


Figure 2 Solar observation images captured at different wavelengths: (upper left) 304 Å extreme ultraviolet , (upper right) white light, (lower left) magnetogram and (lower right) 131 Å extreme ultraviolet. Images taken by NASA's SDO Satellite.

Mechanisms and Prediction of Solar Flare Occurrence

Solar flare occurrence mechanisms—a long-standing mystery—have been researched through satellite and ground-based observations (Figure 2). Sunspots are formed when the sun's magnetic field rises from its inner depths to the surface. Magnetic shear energy builds up around sunspots which is then rapidly released during solar flares. Because of this mechanism, close monitoring of sunspots is important in solar flare forecasting. Our understanding of the relationship between sunspot growth and the occurrence of solar flares has significantly advanced since solar observations from space—unaffected by weather or time of day—became feasible. Solar flare forecasting had traditionally been carried out based on manual analysis of solar flare images. To improve the accuracy of NICT's solar flare forecasting and deepen our understanding of how solar flares are produced, we introduced AI techniques into our forecasting.

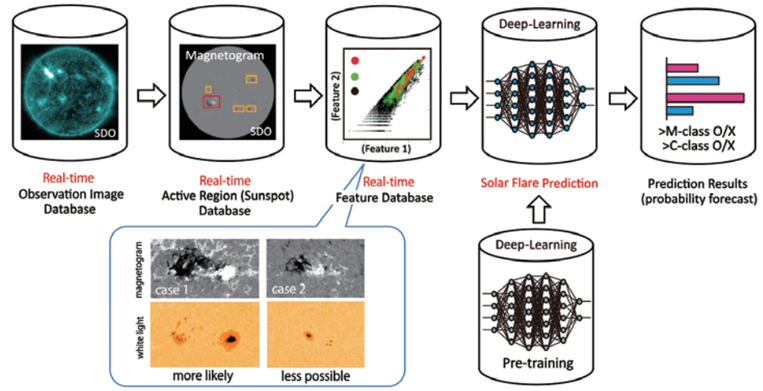


Figure 3 Outline of the Deep Flare Net forecasting model

Deep Learning-based Prediction Model

The use of machine and deep learning and other AI techniques in solar flare forecasting is being actively pursued around the world fueled by an abundance of solar observation image data. A joint team of solar scientists and AI researchers formed by NICT developed and publicized the world's first solar flare prediction AI model. This model (1) is fed satellite observation data obtained from NASA's Solar Dynamics Observatory website, (2) uses magnetograms to detect active solar regions, including sunspots, (3) extracts 79 sunspot features and (4) analyzes the features using deep learning and forecasts the probability of solar flares occurring within the next 24 hours (Figure 3).

This model takes into account features that were found to be important in previous solar research and in the experience of solar flare forecasters (e.g., complexity of sunspots' shapes, strength of sunspots' magnetic fields and small-scale brightening). We trained the model on these features using about 300,000 solar images. As a result, forecast accuracy improved from the 30–50% achieved by conventional forecasting based on manual image analysis to more than 80% via the model. To transform untrained, imperfect AI models into effective and accurate tools, they need to be trained in a manner similar to human students—by providing them with teaching material. We went through a great deal of trial and error in training the model before it finally achieved adequate prediction accuracy. In addition, the model is now able to carry out comprehensive sunspot feature analysis using databases, providing us with a new approach to investigating solar flare occurrence mechanisms.

AI-based Solar Flare Forecast Website

We named our prediction model Deep Flare Net and made its solar flare forecasts viewable on the Deep Flare Net website (Figure 4). The website's 24-hour flare forecast is updated every few hours. To help the general public understand the forecast, we adopted a display format similar to that used in weather forecasting. For example, we classified flare activity levels into the "dangerous," "warning" and "quiet" categories in a manner similar to precipitation intensity categorization. In addition, graphs indicating the probability of flares occurring in different active regions on the sun were designed to be similar to precipitation probability graphs. Since Deep Flare Net—the world's first solar flare prediction AI model—came online in 2019, its forecasts have been used as reference materials in the space weather forecasting meetings held daily at NICT. We are continuously improving the Deep Flare Net website based on feedback from its users. The fact that large solar flares have been observed more frequently since 2021 has increased the value of the website for general users who use solar flare information in their activities.

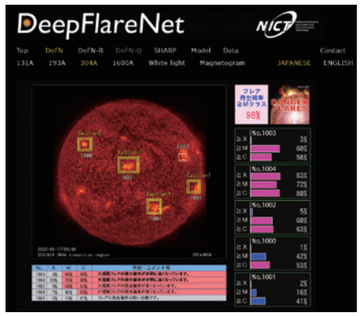


Figure 4 The Deep Flare Net website (https://defn.nict.go.jp/), which publicizes AI-based solar flare forecasts

Future Prospects

Although we have already met our earlier goal of establishing the Deep Flare Net website, we plan to further improve the reliability of our model, enable it to perform longer-term forecasting and offer viewers physical explanations of flare occurrence mechanisms by applying explainable AI. Moreover, we envision expanding the model by taking geomagnetic storms and Earth's atmosphere into account, thereby enabling it to forecast the impact of solar flares on satellite operations, communications and GPS systems. Commercial space weather businesses have been launched in the United States. NICT would also like to promote commercial use of space weather in Japan through continued space weather research and services.

Technology Enabling AI Collaboration for the Automated Management of Large-scale Networks

Pedro Martinez-Julia
Researcher
Network Architecture Laboratory, Network Research Institute
He joined NICT in 2016, working on the automation of network management and control by AI and the standardization of the network architecture.
Ph.D. (Computer Science).

Ved Kafle (left)
Research Manager
Network Architecture Laboratory, Network Research Institute
He joined NICT in 2006, working in current position since 2018. Engaged in research, development, and standardization of new generation of network architecture and technologies.
Ph.D. (Informatics).

ASAEDA Hitoshi (right)
Director of Network Architecture Laboratory, Network Research Institute
He joined NICT in 2012 after working at IBM Japan, Ltd. INRIA (France), and Keio University. Working on research and development of network protocols.
Ph.D. (Media and Governance).

The increasing size and complexity of current networks makes them impossible to be managed by humans alone to ensure the networks operate as needed. Therefore, networks may break or experience performance degradation at anytime. Automating the control and management functions of networks is essential to overcome human limitations by using Artificial Intelligence (AI). However, network administrators (humans) must still be involved in the operation to decide about the actions to be applied to the network because there are situations where AI alone cannot decide, such as when it finds different actions even with the same input. Therefore, a new technology is needed to allow the seamless collaboration between AI and humans, as well as between different AI components and between various management functions.

Managing Large networks

The key problem behind the management of large-scale networks is twofold. On the one hand, the management system must tackle with a huge number of network elements, which implement network functions, such as switches, routers, firewalls, and caches, either in hardware physical equipment or in logical software components such as virtual or cloud-native network function (VNF/CNF). On the other hand, the management system must tackle with the huge amount of monitoring and control information associated to these elements. Every element of the network must be closely monitored to know its state and, in general, to find out if there is any undesirable situation occurring in the network that must be corrected. This means that a massive amount of information must be collected by the network elements and retrieved by different AI components involved in the management by a technology called "network telemetry".

In addition, since different AI components/methods may produce different kinds of results of the analysis, they must interoperate to each other in order to agree on the result and/or exchange information that are relevant to each other. The flow of information and decisions among AI components must involve human administrators at some points through interfaces for human interventions.

The management system requires new methods for collecting and processing network telemetry, while efficiently handling the huge size and diversity of control information, as discussed in next sections.

Optimizing Network Telemetry

In order to find the state of the network and detect and correct any undesired situations, a huge amount of information must be collected by the network elements that implement the network functions and communicated to the management system and their AI components through the control network, which is a special type of (physical or logical) network used for control and management tasks only. The control network is separated from the data plane network through which user application data is transmitted.

The monitoring information is obtained by keeping the log of operational status of network elements, such as the quantity of data they transmit or amount of time they take to process or handle data. The amount of monitoring information that can be transmitted without overloading the control system is limited because of limitations in the available processing and bandwidth capacity of the control network. Consequently, the granularity of control information received by the management system and AI components would be low, resulting in the lower accuracy of decisions carried out by the AI components.

To address the above issue, we have designed a new telemetry mechanism based on data compression techniques that minimizes the amount of control data to be transmitted through the control network and that enables the network management system and AI components to retrieve the enough amount of information.

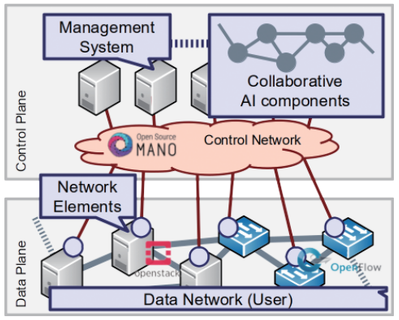


Figure 1 System Overview

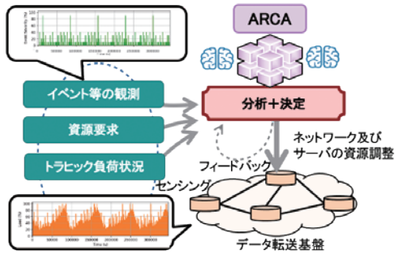


Figure 2 ARCA workflow

Figure 3 Platform Computers running ARCA

Enabling AI to Collaborate

The AI components analyse control data to know the state of the network, identify any potentially undesirable situations, and perform appropriate actions to resolve them. However, it is difficult to identify undesirable solutions in a complex situation by a single AI component. Therefore, we propose a technology that enables seamless communication among multiple AI components involved in the management. AI components are enabled to be connected forming a pipeline, so the output of an AI component (e.g., control data analyser) can be fed to other AI components (e.g., action decider).

Sometimes the AI components can produce multiple possible actions to resolve an undesirable situation, or an action with some parameter that is hard to determine. To overcome these situations, AI components are enabled to share their outputs, so that they all get a better view of the collective findings and decisions and can choose the most appropriate decision that will be finally enforced. The selection process may prioritize the findings and decisions reached by the majority of AI components. Generally, the more AI components interact to one another, the better the resulting AI capability will be and, thus, the more accurate results they will produce collectively. To enable efficient communication among the AI components, the data transmission in the control network has to be agnostic of any particular underlying network structure as much as possible.

Based on the above discussed concept, we have developed a network management architecture, called the Autonomic Resource Control Architecture (ARCA), which enables the collaboration among AI components. ARCA provides a full system for management automation to perform the elastic adaptation of resources assigned to networks, enable them to meet the dynamically changing demands of network service users. Decisions for appropriate actions on resource adaptation are made and enforced on the basis of the collective outputs produced by multiple AI methods.

Promoting Network Automation

We have been aiming to integrate our optimised network telemetry and AI-based network control and management technology into Open-Source MANO (OSM) framework, which is a widely used management and orchestration architecture developed as an open-source software by many companies under the umbrella of ETSI (European Telecommunications Standards Institute). By integrating in the widely used network management system, our technology can be useful for the network industry and, thus, society.

The position of our technology is shown in Figure 1. On the one hand, we propose to extend OSM by introducing the AI components on top, and connecting them to the management system, so they can collaborate as discussed above. On the other hand, we propose to implement our technology to provide the efficient monitoring of the network elements of data networks and users. With OSM, this system can be based on various cloud and software-defined network platforms, such as OpenStack and OpenFlow, which are widely used in the network industry. Figure 2 shows ARCA workflow, which consists on retrieving event notifications and performance measurements, analysing them using intelligence methods to estimate the amount of resources that must be enforced in the network, and finally retrieving feedback from the network to assess that changes are taking the desired effect.

Figure 3 shows the computers that execute ARCA software. The computers run a virtualization platform based on OpenStack and software-defined networking platform based on OpenFlow, as well as the OSM management system. They are high-performance machines, so they can execute ARCA and its AI-based functions in a very short time and provide timely decisions to control the data plane network.

Our future work will focus on improving the scope and accuracy of automation level. In the new level of automation, not only the existing network services are managed by AI, but also new network services are designed and configured based on service requirements. New technology is required for this to become a reality, especially enabling the autonomic management system to decide by itself about the appropriate configuration parameters for creating a new network service.

Privacy Protection Technology using AI-based Federated Learning

LE TRIEU PHONG
Senior Researcher,
Security Fundamentals Laboratory,
Cybersecurity Research Institute
Joined NICT in 2015. Senior Researcher. Engaged in research of cryptography and privacy-enhancing technology. Ph.D.

Lihua WANG
Senior Researcher,
Security Fundamentals Laboratory,
Cybersecurity Research Institute
Joined NICT in 2006 first time, worked as an Associate professor at Kobe University in 2018, and joined back NICT worked as a senior researcher in 2019. Engaged in research of cryptography and its applications on privacy-preserving machine learning. Ph.D. (Engineering).

Privacy-preserving federated learning (PPFL) is an important method in AI, allowing multiple entities to perform machine learning over the combined dataset of all, without sharing the data.

In the sections below, we introduce PPFL systems: one based on neural networks-based deep learning (DL) called DeepProtect, and the other based on gradient boosting decision trees (GBDT) called FL-XGBoost and eFL-Boost.

Two Frameworks of the Systems

In figure 1, there are N distributed learning organizations and a central parameter server. The organizations hold and train their data locally by executing DL or GBDT, and also interactively communicate with the server via secure communication channels. Two frameworks: homomorphic encryption-based one (Figure1 left) and symmetric encryption-based one (Figure1 right) are used to securely share the information for update via the server while no information being disclosed to the server.

DeepProtect: Privacy-Preserving Federated Learning using Deep Neural Networks

DeepProtect is a set of schemes for privacy-preserving federated learning using deep neural networks proposed in [1,2].

The overview is given in Figure 1. The organizations execute replicas of a deep neural network, and interactively communicate with the server. The initial weight vector of the neural network is initialized by an organization, e.g. Organization 1. In DeepProtect using homomorphic encryption described in Figure 1 (left), that organization sends the homomorphic encryption $\mathbf{Enc}(W)$ to the server. The encryption is to protect the secrecy of the weight vector against the server, which is assumed honest in operation, but curious in extracting any information on the organization data. After the above initialization, and at any round of weight update, the organizations obtain the latest encrypted $\mathbf{Enc}(W)$ from the server, and perform decryption to have W. Then, a gradient vector $\frac{\delta J(W,data)}{\delta W}$ is computed using W and a local data batch at the organizations. The encrypted gradient vectors are then sent to the server.

The server task is to recursively update the encrypted weight parameters, utilizing the additively homomorphic property of encryption **Enc**. In particular, the server computes

$$\mathbf{Enc}(W) - \alpha \cdot \mathbf{Enc}\left(\frac{\delta J(W,data)}{\delta W}\right) = \mathbf{Enc}\left(W - \alpha \cdot \frac{\delta J(W,data)}{\delta W}\right)$$

owing to the homomorphic property of **Enc**. Therefore, the weight vector, while encrypted, is still updated by the same way as standard stochastic gradient descent algorithm.

DeepProtect with secret-key encryption is described in Figure 1 (right). In this case, the central parameter server only stores and shares the lasted encrypted weight vector. An organization starting the training process uploads the encrypted weight to the server. Other organization follows the process sequentially by downloading the latest encrypted weight vector from the server, and decrypt it using a secret key shared by the organizations. Then, the weight vector is updated using a data batch from the organization. The updated weight vector is encrypted and uploaded to the server for storing and sharing with others who continues the training process. It is showed in [2] that the weight is protected from the server, and its privacy can also be ensured in relatively involved scenarios such as dishonest organizations and/or the dishonest collusion of organizations and the server.

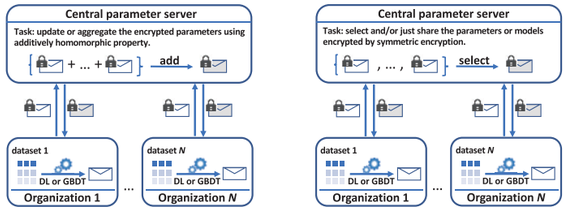


Figure 1 Overview of proposed systems using homomorphic encryption (left) and symmetric encryption (right).

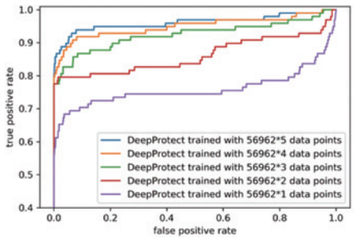


Figure 2 Performance of DeepProtect on the Credit Card Fraud Detection dataset.

In Figure 2, we demonstrate the performance of DeepProtect on the Credit Card Fraud Detection dataset [3]. In the experiment, DeepProtect is trained over increasing number of 56962*1, … , 56962*5 data points. The receiver operating characteristic curves are depicted in Figure 2, showing that the results become better when the data point numbers increase. Therefore, it is expected that DeepProtect can improve the learning outcome when the number of organizations increases.

Privacy-Preserving Federated Learning using GBDT

In collaboration with Kobe University, we have devised inter-organizational ensemble decision tree federated learning schemes, FL-XGBoost [4] and eFL-Boost [5], that can be used by multiple organizations without mutual data disclosure.

In FL-XGBoost [4] (See the overview given in Figure1 (right)), each organization executes XGBoost, a specific GBDT, using their own data to obtain current local model – decision tree T_i then sends the encrypted local model $\mathbf{Enc}(T_i)$ and average gradient $\mathrm{G}_{\mathrm{ave}}^{(i)}$ to the central server, where $\mathrm{G}_{\mathrm{ave}}^{(i)}$ means the estimation error in the global model before updating, and is used as a selection index for the global model. The server task is only to select the encrypted model with maximum selection index (i.e.,$T^* = argmax_{\{T_i\}}\left\{\mathbf{G}_{\mathbf{ave}}^{(1)}, \ldots, \mathbf{G}_{\mathbf{ave}}^{(N)}\right\}$) as a new tree being added to the global model $\{\mathbf{Enc}(T_1), \ldots, \mathbf{Enc}(T_N)\} \rightarrow \mathbf{Enc}(T^*)$, and then send to all organizations. Thereafter, each organization decrypts $\mathbf{Enc}(T^*)$ using a secret key pre-shared among organizations and adds the resulting T^* to the global model. The global model is continuously updated by executing the above process sequentially.

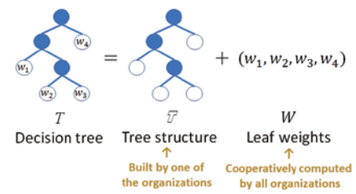
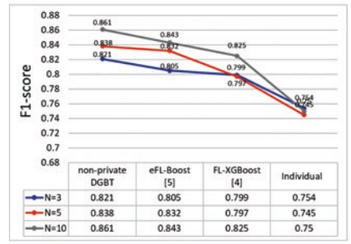


Figure 3 Decision tree model consists of tree structure and leaf weights.



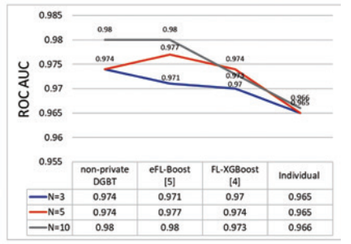


Figure 4 Experimental results for the three patterns of N=3, 5, and 10 when each organization was assigned 22784 training data.

eFL-Boost [5] is a further development of FL-XGBoost, which improves prediction accuracy, reduces communication costs and information disclosure to other institutions. In this protocol, decision tree construction is divided into two stages (Figure 3):

(1) Tree structure determination, which is performed by a participating organization that has been chosen to play a role called Builder. This process is done by a different Builder each time, in order to suppress the biases introduced when updating the global model. The builder constructs only the tree structure $\mathbb{T}$ of the latest local model (i.e., decision tree T) using his own dataset and the global model before updating and the general decision tree construction algorithm in GBDT, and shares with other participating organizations.

(2) Leaf weight calculation, which is cooperatively performed by all participant organizations for each update. For $d\,(d = 1, \ldots, N)$ organization d based on its own dataset D computes $G_d = \sum_{i \in D} g_i$ and $H_d = \sum_{i \in D} h_i$ the sums of the gradients $g_i = \partial_{\hat{y}_t} L(y_i, \hat{y}_t)$ and $h_i = \partial^2_{\hat{y}_t} L(y_i, \hat{y}_t)$ corresponding to each leaf of $\mathbb{T}$, encrypts them using additively homomorphic encryption, and sends the ciphertext $\mathbf{Enc}(G_d)$ and $\mathbf{Enc}(H_d)$ to the server (Figure 1 (left)). The server task is to aggregate the encrypted gradients, utilizing the additively homomorphic property. In particular, the server computes $\mathbf{Enc}(G) = \sum_{d=1}^{N} \mathbf{Enc}(G_d)$ and $\mathbf{Enc}(H) = \sum_{d=1}^{N} \mathbf{Enc}(H_d)$, and send them to all participant organizations. Organizations decrypt $\mathbf{Enc}(G)$ and $\mathbf{Enc}(H)$ to obtain H and G, calculate the global leaf weights $W = -\frac{G}{H+\lambda}$, where λ is a hyperparameter. Then, the global model is updated by adding the local model Twhich is completed as a decision tree by adding the leaf weight W to the tree structure $\mathbb{T}$, to the global model.

It is showed in [4, 5] that the trained models are protected from the server, and its privacy can also be ensured in some scenarios that dishonest organizations without collusion are involved.

Through performance evaluation on the Credit Card Fraud Detection dataset [3], eFL-Boost outperforms existing schemes that incur low communication costs and was comparable to a scheme that offers no privacy protection (Figure 4). For example, when the number of participant organizations N=5, the privacy-preserving system eFL-Boost can obtain the F1-score of 0.832 and ROC AUC of 0.977, while that of the original non-private GBDT are 0.838 and 0.974, respectively.

References

[1] Le Trieu Phong, Yoshinori Aono, Takuya Hayashi, Lihua Wang, Shiho Moriai: "Privacy-Preserving Deep Learning via Additively Homomorphic Encryption," IEEE Trans. Inf. Forensics Secur. vol.13, no.5, pp.1333-1345 , 2018
[2] Le Trieu Phong, Tran Thi Phuong: "Privacy-Preserving Deep Learning via Weight Transmission," IEEE Trans. Inf. Forensics Secur. vol.14, no.11, pp.3003-3015 ,2019
[3] Credit Card Fraud Detection dataset, https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
[4] Fuki Yamamoto, Lihua Wang, Seiichi Ozawa: "New Approaches to Federated XGBoost Learning for Privacy-Preserving Data Analysis," ICONIP2020-27th International Conference on Neural Information Processing, LNCS 12533, pp.558-569, Springer ,2020
[5] Fuki Yamamoto, Seiichi Ozawa, Lihua Wang: "eFL-Boost: Efficient Federated Learning for Gradient Boosting Decision Trees," IEEE Access, vol.10, pp.43954-43963, 2022

Hon'yaku (Translation) Bank

A Large-scale, Public Accumulation Base of Translation Data that to Serve as a Basis for High-precision Automatic Translation

SUMITA Eiichiro

NICT Fellow

After working for IBM Japan and ATR, he joined NICT in 2007. He is also the president of Asia-Pacific Association for Machine Translation[*1] (AAMT). He launched the speech-to-speech translation service VoiceTra in 2010 and the text-to-text translation service TexTra in 2014. He has been running "Hon'yaku Bank" from 2017. He also engages in promoting "Global Communication Plan 2025."[*2] He published *AI Hon'yaku Kakumei* (AI Translation Revolution) in August 2022.[*3] Ph.D. (Engineering).

With the emergence of deep learning, the practical application of artificial intelligence (AI) has made progress in a variety of fields. Automatic translation by computer is a typical example of successful AI application. In addition to general-purpose automatic translation that handles any text safely, there is also high-precision automatic translation for specific fields. Various services have been launched by various companies and their spread is accelerating. This article introduces the Hon'yaku (Translation) Bank, which is an unprecedented mechanism in the world for accumulating translation data that serves as the basis for higher accuracy.

Background

In recent years, automatic translation systems based on deep learning have been released by various companies in and outside Japan. These systems provide remarkably better translation accuracy than conventional systems, and have made automatic translation usable in a wide range of fields for the first time in the history of R&D concerning automatic translation.

The translation accuracy of automatic translation is expected to improve still further due to the reasons below.

(i) Regarding deep learning algorithms, researchers are competing to publish new ideas in papers, and at the same time, their implementations as programs are being made available to all people as open-source software. Therefore, the best programs are constantly being shared around the world.

(ii) Meanwhile, regarding the translation data to be fed to deep learning, tireless efforts are being made to pursue an increase in the amount accumulated and an improvement in quality, which are also vital for improving translation accuracy.

This article looks into the latter factor in detail.

Reality of Data Being *Low-quality and Large-amount or High-quality and Small-amount*

Various experiments conducted by different institutions have demonstrated that the larger the amount of translation data, the higher the translation accuracy will be. Accordingly, institutions developing automatic translation systems are fiercely competing to attain large-scale accumulation of translation data, and the accumulation amount has continued to increase.

If we input a simple Japanese sentence, for example, "toire ga nagaremasen" (which means that the toilet doesn't flush), to different, most-advanced automatic translation systems, we do not necessarily obtain the same translation output. One system may return a mistranslation, "The toilet doesn't flow," while another system may return a correct translation, "The toilet doesn't flush." This is presumed to be attributable to the quality of the training data.

As an AI system imitates its training data, the quality of the training data affects the accuracy of the AI system. The pursuit of the quality of training data is recently drawing attention in the AI research community. Also in the case of automatic translation, the data quality relates to the accuracy of the automatic translation system. Here, the *quality* of translation data means whether a sentence is translated *well* or *poorly*. Moreover, mistranslation with missing information or extra information as compared to the original sentence could be regarded as *low-quality* translation. It goes without saying that *high-quality* and large-amount translation data is favorable, whereas *low-quality and small-amount* data is the most unfavorable. However, examining and screening the quality of *large-amount* data, and preferably correcting the data manually would require enormous costs and be unrealistic.

Generally, either *low-quality and large-amount* or *high-quality and small-amount* data is available. Many institutions have demonstrated through different experiments that the amount of data and the translation accuracy are correlated under such an environment, and that the larger the amount of translation data, the higher the translation accuracy will be. Therefore, as a realistic solution, it is primarily important to accumulate *low-quality and large-amount* data, and further, to effectively mix *high-quality and small-amount* data into it.

The next section introduces NICT's approach aimed at obtaining *high-quality and large-amount* data that overcomes the reality of available data being either *low-quality and large-amount* or *high-quality and small-amount.*

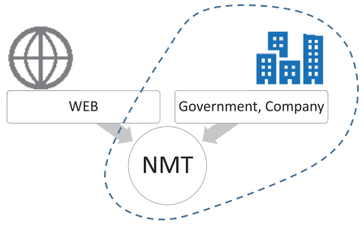


Figure 1 Two kinds of translation data

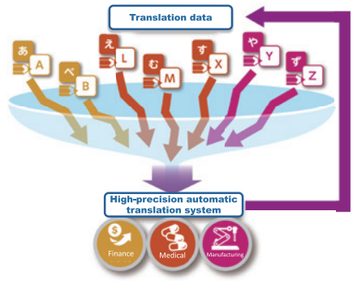


Figure 2 Idea of Hon'yaku Bank

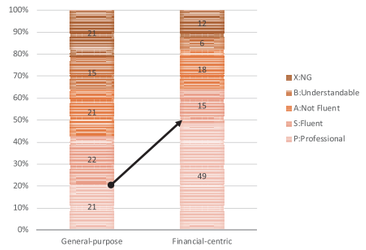


Figure 3 Comparison of translation quality between general-purpose and financial-centric translators

A large-scale, Public Accumulation Base of Translation Data

Institutions engaged in R&D concerning automatic translation are competing to collect translation data. As indicated earlier, the accuracy of the released systems is substantially affected by the translation field or language. The absence of a single system that realizes high-accuracy translation in all fields and languages implies that the current data collection method is insufficient. How can high-quality data be collected in a large amount (Figure 1)?

(i) There is a large amount of translation data on the web, but high-quality data and low-quality data are mixed together, and it is not easy to realize a technology for automatically judging the data quality.

(ii) On the other hand, translation data that has been proofread with sufficient care is scattered in individual organizations owning translation data, such as companies as well as local and national government offices. A promising means will be to collect reliable, quality-controlled data from public and private organizations.

As shown in Figure 2, the translation accuracy of automatic translation for specialized fields will dramatically improve if the scattered bilingual data is accumulated in a single organization, and is fed to deep learning for automatic translation. This approach is also expected to improve the coverage rate of general-purpose automatic translation.

The Ministry of Internal Affairs and Communications and NICT promote this scheme under the name "Hon'yaku Bank."[*4] The provided translation data has been utilized in deep learning, and the technology of the created high-precision automatic translation system has been transferred for the purpose of making it widely available (the translation accuracy of the latest version can be experienced with the speech translation app VoiceTra® and text translation system TexTra®, which are showcases of NICT's research results[*5,*6]). Private companies would likely have difficulty coordinating interests if they were to implement this initiative, but public institutions have little difficulty in that respect, and in fact have been able to make successful collaborations with various industries.

Hon'yaku Bank embodies a vision to build an automatic translation system of the people, by the people, for the people. Furthermore, if the scheme of the large-scale, public accumulation base can be spread to the entire world, unprecedentedly high accuracy would be achievable by a system built by accumulating bilingual data of previous translations in all fields in all languages.

Here is one of the latest case examples. The Financial Services Agency (FSA) collected, and provided to Hon'yaku Bank, a large number of translation documents owned by the FSA and financial industry groups. NICT converted the translation documents into data suitable for deep learning, and further automatically refined the data, to develop a high-precision automatic translation system.

NICT started technology transfer of this system to private companies on March 1, 2022. The system has been utilized in the FSA on a daily basis, and is highly commended by officials using the system.

As a result of evaluating the translation quality of this system, the percentage of translations that reached the highest-quality, financial-centric translator level (P: professional level) increased from the conventional 20% to 50% (Figure 3), and it became possible to dramatically improve the efficiency of the translation work of documents in the financial field.

Future Prospects

Hon'yaku Bank is expected to accelerate the realization of high-precision automatic translation in all kinds of fields in all kinds of languages. In addition, based on this achievement, simultaneous translation services are likely to become available everywhere within a few years by also taking advantage of the progress in speech recognition technology.

Furthermore, as a result of automatic translation becoming based on deep learning, a future is drawing near where issues that have been considered to be difficult or impossible to solve in the past, including utilization of context and non-text information, such as charts and graphs, will be solved in an integrated manner in the same paradigm.

The technological progress of automatic translation supported by Hon'yaku Bank and its peripheral technology is worthy of attention.

---

*1 https://www.aamt.info/
*2 https://www.soumu.go.jp/menu_news/s-news/01tsushin03_02000298.html (in Japanese)
*3 ISBN-13: 978-4023322639
*4 https://h-bank.nict.go.jp/
*5 https://voicetra.nict.go.jp/
*6 https://mt-auto-minhon-mlt.ucri.jgn-x.jp/

Integrating Features of the Brain into AI

Pursuing the Creation of an Artificial Brain

NISHIDA Satoshi

Senior Researcher, Neural Information Engineering Laboratory, Center for Information and Neural Networks, Advanced ICT Research Institute

He joined NICT in 2014 after receiving his Ph.D. in medicine and engaging in short-term postdoctoral work. He is also a Guest Associate Professor at Osaka University. His current research topics include cognitive computational neuroscience, neural information engineering and artificial intelligence.

We developed an AI which integrates features of the human brain (brain feature-integrated AI; BFI-AI). This BFI-AI is capable of simulating human brain information processing by predicting brain activity induced by audiovisual input. The BFI-AI was found to be able to estimate audiovisual input-induced human cognitive and behavioral patterns more accurately than existing AIs. Moreover, we confirmed that the BFI-AI is able to discern individual differences in brain information processing. This technology may be used to significantly advance the development of artificial brains capable of computationally simulating the idiosyncratic brain information processing of human individuals.

Background

AI technologies have evolved rapidly in recent years and are becoming more widely used in our daily lives. I envisage a society in which humans and AIs will coexist in the relatively near future. Developing fully socially acceptable AIs capable of behaving like humans will require a deeper understanding of humans. One potentially effective approach to developing such AIs is to enable them to simulate human brain information processing. In line with this approach, our research group developed an AI which integrates features of the human brain (BFI-AI). This BFI-AI is capable of simulating brain information processing using a prediction AI model and is able to accurately predict brain activity induced by audiovisual input from video and audio sources.

BFI-AI Capable of Simulating the Brain's Information Processing

The BFI-AI is composed of a prediction AI model—which predicts brain activities induced by audiovisual input from videos and other sources—and a decoding AI model—which translates the predicted brain activities into cognitive and behavioral responses associated with the audiovisual input (Figure 1). To construct this prediction model, we measured the brain activity of people viewing videos using a functional magnetic resonance imaging (fMRI) technique. A deep neural network (DNN)—a type of leading-edge AI technology—was then fed audiovisual data from the same videos to allow it to predict human brain activity induced by viewing them. The accuracy of the prediction model was optimized by adjusting the DNN-predicted brain activities to actual brain activity measurements. To construct the decoding model, human cognitive (e.g., preferences and impressions) and behavioral (e.g., actual time spent viewing the videos and the purchase of items displayed in the videos) responses associated with viewing the videos were measured. The translation capability of the decoding model was optimized by first allowing an AI technology to translate the brain activities estimated by the prediction model into cognitive and behavioral responses and then adjusting these results to match actual human cognitive and behavioral responses.

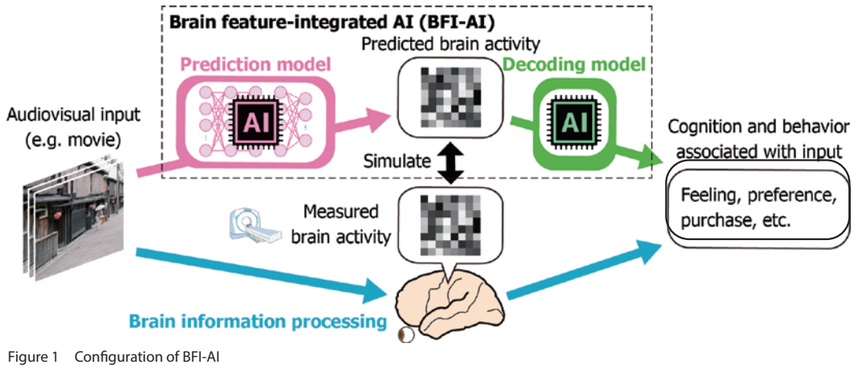


Figure 1 Configuration of BFI-AI

The human brain processes audiovisual input and then produces cognitive and behavioral responses through a series of information processing steps. This BFI-AI is composed of prediction and decoding models designed to simulate these information processing activities based on actual brain activity measurements taken from individuals.

BFI-AI Demonstrates an Improved Ability to Estimate Human Cognitive and Behavioral Responses

Standard AI technologies can also estimate human cognitive and behavioral responses by directly analyzing audiovisual input. By comparison, the BFI-AI uses an additional computational step to predict brain activity induced by audiovisual input. Our studies confirmed that this additional step made the AI behave more similarly to a human brain with an improved ability to estimate certain types of cognitive and behavioral responses (Figure 2). In these studies, we used internet and TV advertisement videos as audiovisual input sources and compared the ability of standard AIs and BFI-AIs to estimate cognitive and behavioral response measures, such as the proportion of viewers who watched an entire video without skipping any portion of it and the level of favorability which viewers demonstrated for the TV advertisements. As a result, the BFI-AI was found to be superior to standard AIs in estimating these measures.

Another interesting finding from the various trials we conducted is that the performance of the BFI-AI increases as the accuracy of its decoding model (which translates brain activity measurements into cognitive and behavioral responses) increases. This result indicates that this AI would be particularly effective in estimating cognitive and behavioral responses closely associated with brain activity. This is because the decoding model is likely to be more accurate in estimating responses strongly linked to brain information. Thus, it is potentially feasible to improve the AI's ability to faithfully simulate human brain behavior.

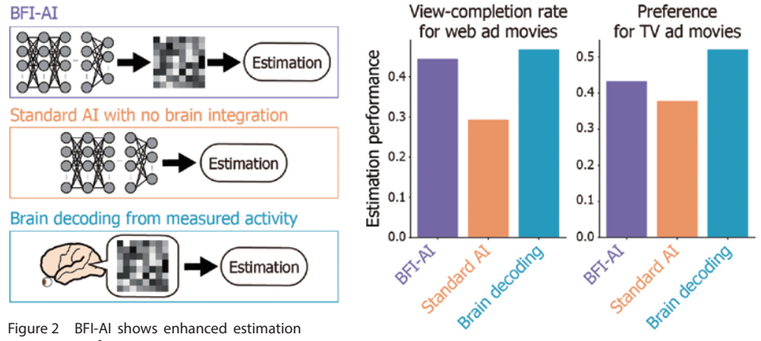


Figure 2 BFI-AI shows enhanced estimation performance

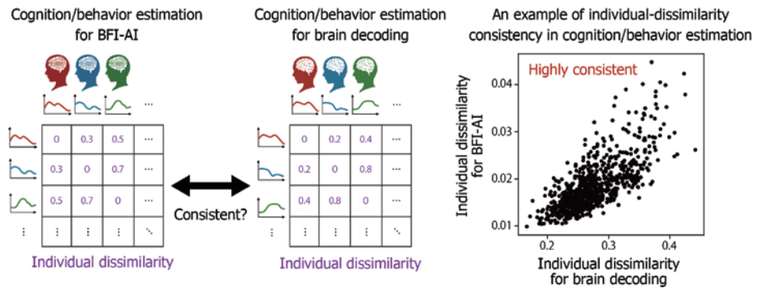


Figure 3 Individual differences in cognitive and behavioral responses discerned by the BFI-AI and the brain decoding model

BFI-AI Discerns individual Differences in Brain Information Processing

The BFI-AI is even able to simulate the brain activity of different individuals. Our studies found that the BFI-AI was able to simulate an individual's brain activity, including information processing properties unique to him/her (Figure 3). The BFI-AI estimates cognitive and behavioral responses in chronological order based on the chronological order of different audiovisual inputs. The BFI-AI then measures the degree of individual differences in cognitive and behavioral responses using an individual dissimilarity index. We hypothesized that if the BFI-AI accurately discerns individual differences in brain information processing, the individual dissimilarity in cognitive and behavioral responses discerned by the BFI-AI and the dissimilarity discerned by the brain decoding model (which translates actual brain activity measurements into cognitive and behavioral responses) would be similar. We tested this hypothesis by carrying out a correlation analysis and found that individual dissimilarities for 81 of 87 types of cognitive and behavioral responses analyzed correlated significantly between the AI and the brain decoding model.

Future Prospects

One advantage of a BFI-AI is its ability to adjust the information processing patterns it simulates to match those of the human brain. Through further advancement, this technology may potentially make a significant contribution to the development of AIs capable of behaving like and coexisting with humans. Another advantage of this AI is its ability to computationally simulate individual-specific brain information processing. This ability could potentially be used to develop artificial brains capable of serving as next-generation brain information technologies for various purposes, including human digital twins and digital archiving of brain activity.

The Meritorious Award on Radio is bestowed annually on individuals or groups who make significant contributions to improving the effectiveness and appropriateness of radio frequency utilization.

* Affiliations and positions are as of the time of the awards.

The 33th Meritorious Award on Radio

MATSUMURA Takeshi

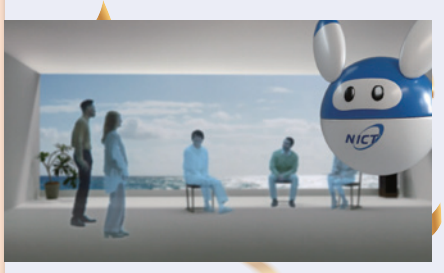
Director of the Wireless System Laboratory, Wireless Network Research Center, Network Research Institute

●Date: June 28, 2022

●Co-recipients: FUJII Takeo (The University of Electro-Communication), HARADA Hiroshi (Kyoto University), HAYASHI Takahiro (KDDI Research, Inc.), SHIMOMURA Masahiko (Mitsubishi Research Institute, Inc.), SAWAI Ryo (Sony Group Corporation), YOKOYAMA Hitoshi (IBM Japan, Ltd.)

●Abstract: These award winners developed a frequency sharing technology designed to enable different communications systems to dynamically share 2.3 GHz band frequencies. To put this technology into practical use, they conducted R&D and tested its usability while taking into account geographical and temporal factors affecting the operation of target mobile and existing radio communications systems. In addition, they coordinated the interests of stakeholders and actively promoted the practical use of the technology. These efforts were recognized as a significant contribution to making radio frequency utilization more effective.

●Recipients' comment: Dynamic frequency sharing technologies are expected to become a very important tool in making radio frequency utilization more effective in the future. We were able to actually integrate this technology into a system, which we believe is a major step forward in improving radio communications. We are very honored to have participated in this national research project as members of the National Institute of Information and Communications Technology. We thank the many other researchers who carried out various planning and data collection/entry work at the actual research sites. Because of their hard work, we were able to meet our goals by the end of this short-term, two-year project.



NICT video clip reaches a million views!

NICT PR MOVIE

A Future with N —NICT's AI Robot Mascot

We posted a NICT PR video entitled "A Future with N—NICT's AI robot mascot" on YouTube in June 2022. This video has been viewed more than a million times. In the video, the AI robot character "N" (shown on the front cover of this NICT News issue) from the year 203X explains NICT's future R&D vision in an easy-to-understand manner.

Mone Kamishiraishi—a Japanese actress—narrated the Japanese version of the video. We hope you enjoy it!

https://www.youtube.com/watch?v=GHtjA9LvgR4 (Japanese version)
https://www.youtube.com/watch?v=YOenAWHhRIA&list=PLBwwDuSrrNU1GuFPOguE5WrZvKzje1eZN (English version)

NICT NEWS 2022 No.6 Vol.496
Published by Public Relations Department, National Institute of Information and Communications Technology
Issue date: Nov. 2022 (bimonthly)

4-2-1 Nukui-Kitamachi, Koganei, Tokyo
184-8795, Japan
TEL: +81-42-327-5392 FAX: +81-42-327-7587
URL: https://www.nict.go.jp/en/
@NICT_Publicity
#NICT

ISSN 2187-4050 (Online)