Technology Reports (Special Articles)
DX | Natural Language Processing | Document Classification
Special Articles on AI—Expansion of AI Technologies to Diverse Industries and Basic Technologies Supporting AI Applications—

# Natural Language Processing for Realizing DX

Service Innovation Department Yutaro Shiramizu Keisuke Oka Shusuke Tatsumi

In recent years, there has been a great deal of interest in the development of technologies that use AI to help promote DX. In particular, attention has been focused on the field of natural language processing, which is an AI technology that can quickly and accurately process large amounts of text data accumulated by businesses. This article introduces the natural language processing AI being developed by NTT DOCOMO, its technical features, and the GUI tools that support the use of AI. It also discusses actual implementations of this technology, its synergistic effects with RPA tools, and its future prospects.

## 1. Introduction

In recent years, there has been a government-led push to reform the way people work, and there is a strong demand for the use of digital technologies such as Robotic Process Automation (RPA) tools*1 and chatbots*2 to speed up business operations, improve productivity, and reduce workloads. In particular, RPA tools can make significant improvements to the processing of large quantities of text data, including questionnaires and user inquiries, and are becoming increasingly popular, especially in front and back offices. However, with RPA tools alone, it is difficult to achieve the Digital Transformation (DX)*3 of work requiring intelligent judgments, such as text categorization and



*1 RPA tool: A software-based robot that records PC operations as scenarios and automates PC operations.

tagging. There is consequently also an increasing demand for AI-based automation in the field of natural language processing, which can replace human linguistic comprehension.

However, AI has to be diverse because the available computational resources and required performance differ from one project to the next in practical situations. People also have reservations about the security aspects of using AI to process a company's private internal text data in public external Software as a Service (SaaS)*4 applications. Furthermore, departments without system engineers can find it very burdensome to implement the introduction and operation of AI by themselves. Due to issues such as these, the introduction of AI is not always an easy task.

To address these issues, NTT DOCOMO has developed a high-performance natural language processing algorithm that makes full use of the computational resources, and a lightweight and high-speed algorithm that can run on low-spec PCs. These algorithms have made it possible to select and use flexible natural language processing AI technologies that meet the needs of users in diverse scenarios, including cloud and local processing environments, PCs and servers with diverse specifications, and various user security policies. To reduce the burden on operators, we have also developed a Graphical User Interface (GUI)*5 tool that makes these AIs easy to use and fine-tune.

This article describes the development of natural language processing AI and GUI technologies that meet diverse and widely varying needs in practical tasks such as text categorization and the concealment of personal information, as well as in internal and external applications.

## 2. Natural Language Processing AI to Help Drive DX

### 2.1 Overview

In recent years, natural language processing has been the subject of academic research all over the world, and many high-performance algorithms have been proposed in studies conducted under well-developed experimental settings. These algorithms are often released as Open Source Software (OSS)*6, so the latest algorithms are readily available for anyone to use. On the other hand, in the business world, there are many restrictions on the availability of computational resources and data, so it is not possible to satisfy the needs of users by applying these algorithms directly. At NTT DOCOMO, we have therefore been working on the development of additional functions and the creation of methods for deploying these functions to meet the diverse needs of real users while using the latest OSS.

### 2.2 Functions

1) Document Classification

A document classification function is a function that automatically assigns labels to documents (**Figure 1** (a)). NTT DOCOMO provides two types of classifier (lightweight and high-performance) to satisfy users with different needs.

(a) Lightweight classifier

The lightweight classifier meets the needs of local environments using ordinary business PCs for situations such as when the cloud cannot be used due to concerns about

*2 Chatbot: A program that automatically conducts interactive dialog with people via speech or text.
*3 DX: The use of IT technology to revolutionize services and business models, promote business, and change the lives of people for the better in diverse ways.
*4 SaaS: Software that is used remotely via the Internet or other networks.
*5 GUI: An interface that consists of a combination of buttons and icons that are clearly visible and can be operated intuitively.
*6 OSS: Software whose source code is released free of charge for anyone to reuse or modify.

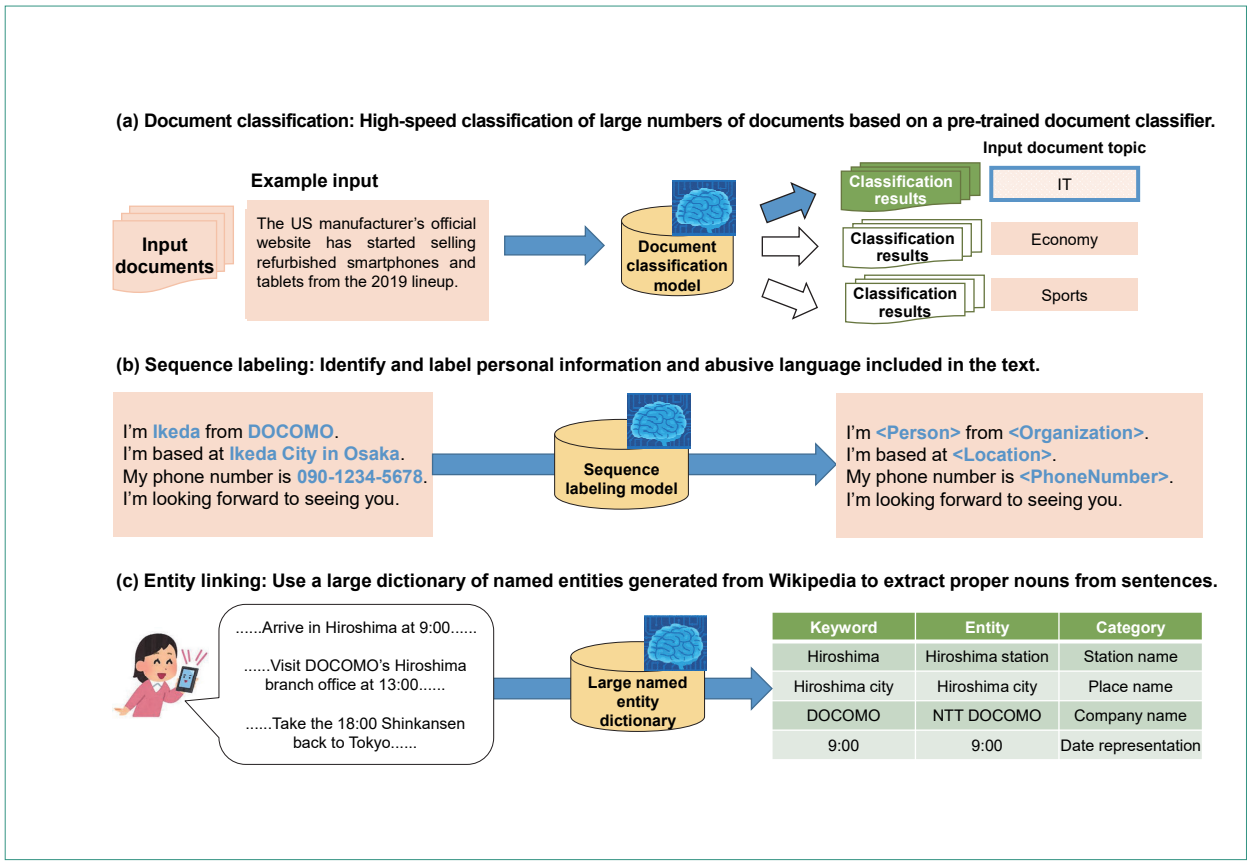


Figure 1 List of natural language processing AI features

cost and security. Its algorithm uses a multilayer perceptron*7 that is simple to implement, computationally inexpensive, and capable of running on a low-spec PC. The features*8 used in this algorithm consist not only of morphemes*9, which are commonly used for this purpose, but also the actual characters used in the text and the appearance of character N-grams*10 within it. This makes it possible to achieve high accuracy even with a small corpus*11, and enables the construction of classifiers that are robust against misspellings and omissions. **Figure 2** shows an example of FAQ classification to illustrate how these features can be used effectively. In FAQ classification, the text of a question entered by a user is automatically classified to match a prepared FAQ item. Here, if the input text contains misspellings, it cannot be classified correctly when using only morphemes. On the other hand, if characters and character N-grams are used, a question can be correctly classified by inferring substrings (like "b, l, u, e, t, o, ..." in the sample shown in the figure) as hints.

(b) High-performance classifier

The high-performance classifier is designed to meet the needs of users who have either

*7 Multilayer perceptron: A type of neural network that implements machine learning algorithms.
*8 Feature: An amount (a numeric value) extracted from data to characterize that data.
*9 Morpheme: The smallest unit of a linguistic expression.
*10 N-gram: Any sequence of $n$ consecutive elements (e.g., words or letters).
*11 Corpus: A language resource consisting of a large volume of text and utterances, etc. collected and stored in a database.

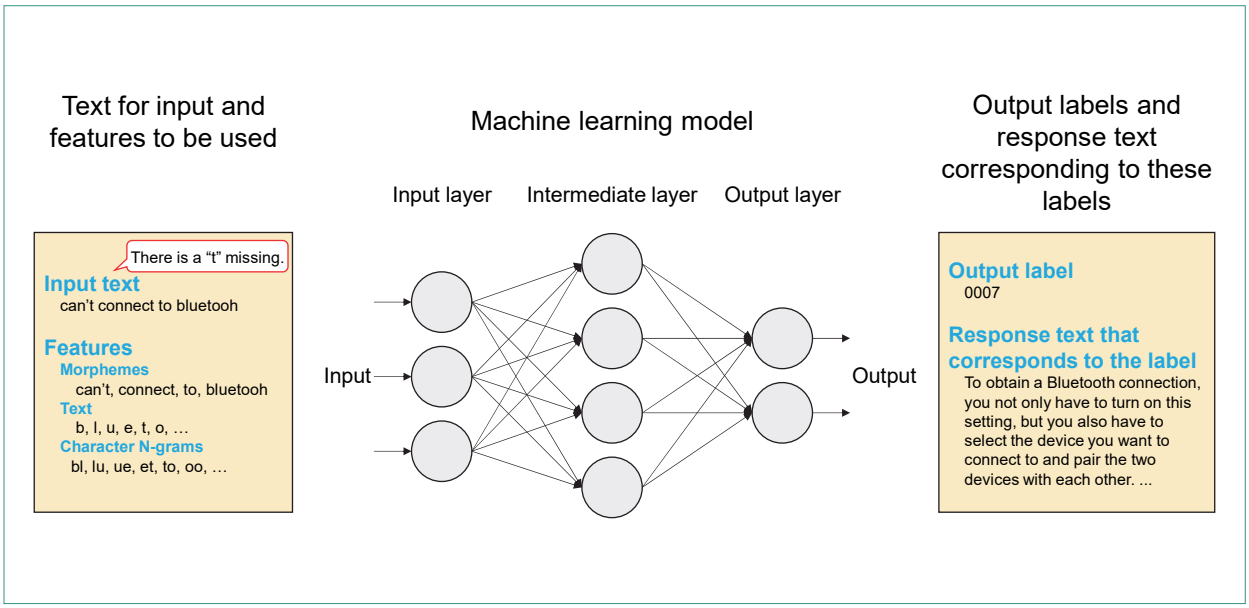


Figure 2 An example where the features of a lightweight classifier work effectively

a cloud server or an on-premises*12 server equipped with a Graphics Processing Unit (GPU)*13, and who want to achieve high performance by making the best use of abundant computational resources. It uses an algorithm called Bidirectional Encoder Representations from Transformers (BERT) [1], which has become the de facto standard for language processing in recent years. BERT requires large amounts of training data and computational resources, but performs very well. Its pre-learning model*14 is based on the NTT version of BERT*15, a technology developed by NTT Human Informatics Laboratories.

NTT DOCOMO has also provided its own support for multi-label classification*16. The NTT version of BERT only supported single-label classification*17, but there is a strong demand for multi-label classification in the business field. To meet this need, NTT DOCOMO altered the label output part of NTT’s BERT classifier to provide multi-label output.

2) Sequence Labeling

The purpose of the sequence labeling function is to automatically assign labels to sequential data*18. At NTT DOCOMO, we use this function to mask text including content such as personal information and abusive language (Fig. 1 (b)). In developing this function, we used an open-source natural language processing library called Flair [2], which has a large developer community and can handle lightweight algorithms. In the selection of OSS, the use of lightweight algorithms was an important consideration because users often have to avoid using the cloud for security reasons when processing highly confidential data, and therefore need a solution that can also operate in a local environment. A large developer community is also

*12 On-premises: Refers to an environment where the constituent hardware of a corporate system is owned, operated and maintained by the company itself.

*13 GPU: A processor unit that excels in parallel computing. It is ideal for deep learning processes that requires parallel computation.

*14 Pre-learning model: A model that has been trained unsupervised on a large corpus before supervised learning of the target task.

*15 NTT version of BERT: A BERT model that was pre-trained using data collected by NTT Human Informatics Laboratories.

important, because it makes it more likely that the software will be continuously updated with interface improvements and bug fixes based on feedback from diverse users. We avoided OSS from small developer communities because it tends to have meager documentation and less accumulated know-how, so that the development and operation of systems often depends on an individual developer.

3) Entity Linking

The entity linking function extracts keywords from sentences and then links these keywords to entities. For example, sentences may contain not only common nouns like "desk" and "train", but also proper nouns like "America", "USA", and "United States". Sometimes, different nouns may refer to the same concept (entity) such as the last three examples in the previous sentence, which all refer to the United States of America. Entity linking is a function that mechanically extracts keywords from a sentence (words that convey a unique meaning) and infers the entities to which these keywords refer (Fig. 1 (c)).

(a) Building a dictionary based on data extracted from Wikipedia

At NTT DOCOMO, we perform entity linking by using data extracted from Wikipedia [3]. This data – including each article's body text, anchor text, and number of page views – is statistically processed and used to weight the keywords and construct an entity-keyword dictionary. Using Wikipedia has several advantages: (a) its articles are frequently created and updated, and it responds quickly to new phenomena and popular words, (b) new database dumps*19 are published every day, making the data easy to use, and (c) the cost of maintaining the dictionary is low because the entire process from data extraction to dictionary construction is fully automated.

(b) Category assignment

NTT DOCOMO also assigns categories to entities as higher-level concepts based on an extended named entity hierarchy*20 [4], and classifies the extracted results into about 200 categories before they are used. For example, in a dialogue system, it would not be practical to prepare individual system responses for every single possible user utterance, such as "I want to eat noodles" and "I want to eat an apple." However, if we use the fact that the words "noodles" and "apple" both belong to a "food name" category, it is only necessary to prepare a system response for "I want to eat [food name]," which greatly reduces the burden of creating response scenarios.

(c) Use case

One possible use case of entity linking is the extraction of keywords from news. The use of Wikipedia statistics for entity linking is a good match with keyword extraction from news items, because Wikipedia updates its article content and increases the number of page views to reflect current trends. Furthermore, since it is possible to perform not only simple character string extraction but also to accommodate different ways of representing the same concept, this approach makes it easy to use extracted/linked entities or categories associated with entities as article tags.

*16 Multi-label classification: A classification method where multiple labels can be assigned to a single piece of data.
*17 Single-label classification: A classification method where a single label is assigned to a single piece of data.
*18 Sequential data: Data consisting of a series of elements, such as character strings, audio waveforms, and purchase histories.
*19 Database dump: A file containing the content of a database.

# 3. GUI Tools to Accelerate the Deployment of Natural Language Processing AI

## 3.1 Overview

When implementing and using natural language processing AI, it is important to tune the AI to maintain and improve its accuracy. Tuning refers to a series of tasks involved in the construction and continuous updating of AI models, such as verifying the AI accuracy and modifying the training data by performing annotation*21 work. It also involves visual work by operators, which incurs a certain level of human labor costs.

To reduce these labor costs, NTT DOCOMO has also developed GUI tools to support AI tuning. With these tools, it is possible to use simple screen operations to rapidly perform AI training and evaluation, which requires repetitive work such as construction and tuning. In addition, the prediction results of AI trained by users can be used for annotations to adjust the amount of manual effort and prioritize tasks so that work can be performed efficiently. Furthermore, since the constructed AI and data are managed by the system in units of revisions, it is no longer necessary for the user to be aware of complicated information management such as which data was used to construct which AI and when, and how accurate it was. As regards the introduction of these tools, their back-end*22 engine runs on container virtualization technology*23 and thus does not depend on work environment attributes such as a local/cloud server or host OS. These GUI tools are also containerized so that the back-end tools and front-end*24 GUI can both be introduced easily.

In this way, by using GUI tools that are easy to introduce, users can easily build, tune and manage AI systems by performing simple screen operations, which is expected to reduce the hurdles that have to be negotiated in order to set up and use natural language processing AI.

## 3.2 GUI Tools and Their Functions

The GUI tools provide a suite of functions that are needed for the implementation of AI, including uploading datasets, creating annotation data, learning, and performing evaluations. Everything from the initial construction of AI systems to the maintenance of these systems can be done on-screen via a GUI (**Figure 3**).

The details of the main functions are outlined below.

1) Learning/evaluation

Users can create their own AI systems by using the GUI tools to upload pre-prepared annotated text data to the backend server and perform learning. To judge the accuracy of labels automatically assigned to the input, the user can upload annotated data prepared as correct answers to the GUI tool, whereby the AI can then automatically calculate the correct answer rate and evaluate the accuracy.

In this way, users can easily construct their own AI systems and evaluate their accuracy simply by preparing annotated text data.

2) Tuning

With this function, users can check and correct labels by using trained AI to automatically assign labels to text data. In general, AI can reproduce human judgment by learning from texts labeled by humans. The more texts it has to learn from,

*20 Extended named entity hierarchy: A semantic classification structure wherein words are classified into about 200 categories such as “names of people”, “names of cities and towns”, and “names of countries”. These categories have a structure of up to three levels, such as “place name > name of astronomical body > planet name”.

*21 Annotation: Manually annotating data such as text and images

*22 Back-end: The part of a system that runs the GUI. It mainly consists of a processing engine and a system part that connects the engine with the GUI.

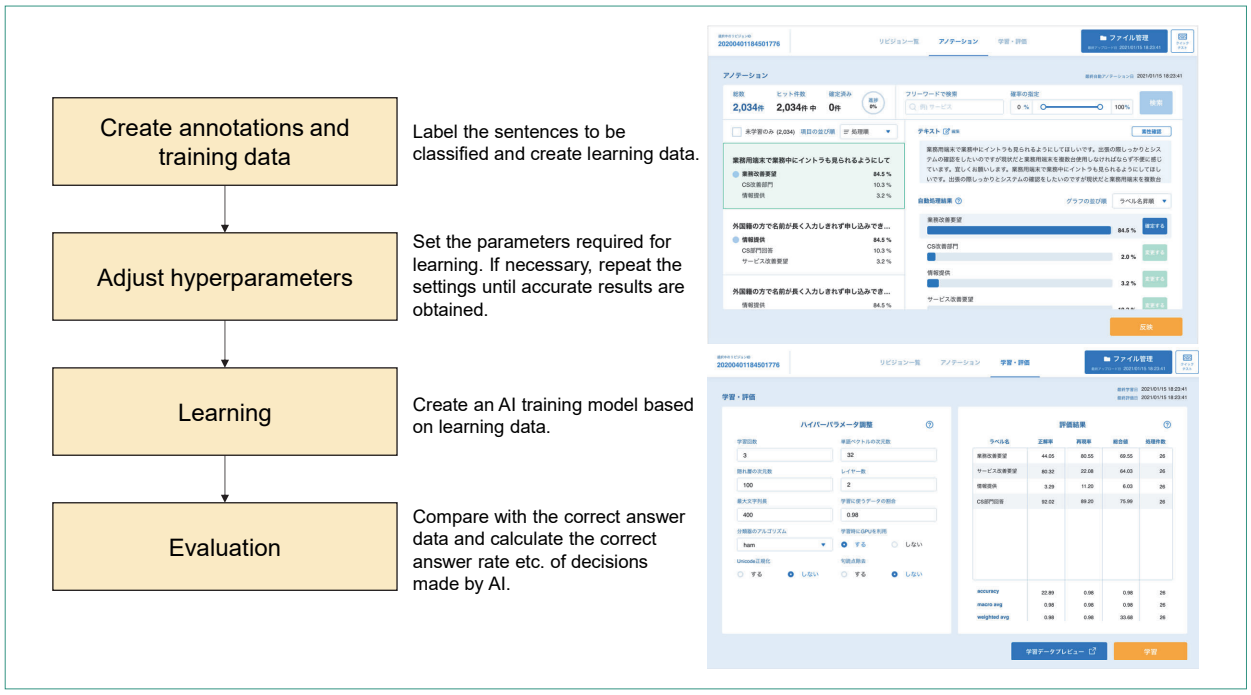


Figure 3 Tuning flow and screenshot of the GUI tools

the better its accuracy.

Once the AI has been trained to some extent, the annotation process can be made more efficient by allowing it to start annotating text, and using human judgment to correct its mistakes. The corrected annotations can then be used to improve the AI accuracy by further training.

As a result, this function makes it possible to improve the accuracy of AI while reducing the cost of creating labeled text data.

It also provides a screen where users can perform actions such as executing learning, evaluating the accuracy of learning, and changing hyperparameters*25 during learning. This makes it easy to adjust parameters*26 during tuning.

3) Revision Control

When tuning an AI system, the complexity of managing previous datasets and evaluation results can often become an operational issue. This tool provides functions whereby trained AI systems and the datasets and evaluation summaries used in learning can be managed in work units called "revisions." In this way, users do not need to be aware of the abovementioned complex management. Instead, the previous datasets and evaluation summaries can be individually managed on the system side, allowing the effects of tuning and differences of data to be analyzed for each dataset.

## 4. Use cases

### 4.1 Overview

At NTT DOCOMO, we are not only developing natural language processing AI and GUI tools, but we are also promoting DX using natural language processing AI both inside and outside the

*23 Container virtualization technology: A technology in which an application itself and all its necessary files are packaged together as a "container" and run on a process called a container engine.

*24 Front-end: The part of the system that the user sees and interacts with. This mostly corresponds to the GUI.

*25 Hyperparameter: A value that is set during training. Since the performance varies depending on the setting values, they should be optimized for the best performance.

*26 Parameter adjustment: The selection of relevant setting values for the best performance.

company.

Internally, we are cooperating with the Customer Satisfaction (CS) department, which collects text-based feedback from users and employees. We are working to improve their operations and help them achieve greater user satisfaction through measures such as constructing and supporting the introduction of natural language processing AI, and sharing the accumulated results of manual classification.

Outside the company, we are also working with the corporate sales division to develop RPA-AI collaborative solutions, and we are conducting a wide range of demonstration experiments using not only natural language processing but also Optical Character Recognition (OCR) and speech recognition[*27] technologies.

## 4.2 Working with the In-house CS Department

NTT DOCOMO collects feedback from users and employees at its shops and call centers. After it has been anonymized so that specific individuals cannot be identified, we use this data to create better services and operations that lead to increased user and employee satisfaction. For example, questionnaires submitted by users are read by specialist staff who compile them into units such as "requests" and "plaudits" before they are shared with related departments. However, a large number of questionnaires are received every day, and it takes a lot of time to read them all properly. Since this process relies on human judgment, it is inevitable that the accuracy will decrease due to oversights or misclassifications. Another issue is that it is difficult to equalize the quality of this work because the classification results are dependent on the knowledge and experience of staff.

NTT DOCOMO has therefore been working with the in-house CS department to promote the automatic classification of questionnaires by introducing natural language processing AI with the aim of improving operational efficiency by promptly and appropriately responding to user feedback. Since the collected questionnaires were used and analyzed within a closed in-house network, it would have been difficult to have them automatically processed by an external high-spec server. We therefore developed the abovementioned lightweight classification algorithm and made it available as AI for use in classifying questionnaire responses. Since data from previous questionnaires had already accumulated by manual classification, the CS department used it as AI learning data.

By introducing natural language processing AI into this work, we were able to automate most of the work that had previously been done manually for a long time. As a result, the questionnaire processing workload was greatly reduced. In the future, in addition to the automatic classification of questionnaires, we also aim to achieve greater user satisfaction by introducing AI that automatically conceals personal information such as names and addresses that users have mistakenly entered.

## 4.3 Solutions for Linking RPA Tools and AI

In recent years, solutions to automate more complex tasks (Cognitive Automation[*28]) have been explored by combining RPA tools with "brains" implemented using AI for natural language processing or OCR. As part of the introduction and expansion

*27 Speech recognition: Technology that analyzes human speech in order to convert it into text or infer people's emotions.

*28 Cognitive Automation: A technology that combines RPA tools with AI tools such as natural language processing, image recognition and voice recognition to enable the automation of judgment-based work that is performed by humans.

of our WinActor RPA tool and our delivery of RPA-AI collaborative solutions, NTT DOCOMO has also been working with the Nara Prefectural General Medical Center on demonstrations of an experimental system for automatically inputting electronic medical records into existing systems.

Although humans are required to do some of the work in the medical field, such as performing medical examinations and tests, there are also many tasks that can be made more efficient by IT, such as creating medical records for each patient and inputting data into electronic medical record systems. However, especially when inputting data to the system, it is necessary to have some level of understanding of the text data before it can be transcribed to the proper location, so it is not easy to perform this task with RPA tools alone. In addition, since this process involves handling personal information, it is preferably performed within the confines of a local network between work PCs rather than by a SaaS solution that sends the data over an external network. In addition, hospitals often already have databases and systems up and running, and these have to be used as they are without any modification.

To accommodate these requirements, we constructed a lightweight natural language processing AI tool that can run locally and an RPA tool that requires only a PC connected to the hospital network and does not require an external server. We linked these systems together to automate the work of transcribing records into the system that had previously been performed by humans. By doing so, we were able to eliminate excess work by reducing the workload of medical personnel without making major changes to the later stages of the workflow. We also introduced new voice recognition software with the aim of improving the efficiency of data creation by making it possible to convert printed patient records into electronic data simply by reading them out aloud.

In this demonstration experiment, we evaluated not only quantitative effects such as the reduction of the time required to enter data into the electronic medical record system, but also the qualitative effects mentioned by people responsible for entering this data, who said that it reduced the incidence of data entry errors and gave them more time to spend with patients. In this way, the introduction of the RPA-AI collaborative solution made it possible for hospital staff to concentrate on their core business, and produced results leading to the creation of new value for customers. These results strongly suggest that DX is not only able to deliver cost reductions, but can also contribute to improving the quality of work itself.

## 5. Conclusion

We have discussed the development of natural language processing algorithms for delivering DX and GUI tools for improved operability and convenience, and the deployment and future prospects of these tools both within NTT DOCOMO and in other businesses. By developing a variety of natural language processing AI tools supporting diverse methods and algorithms, we have made it possible for users to select the appropriate AI according to their applications, requirements, and execution environments. At the same time, we sought to make operating AI in the field less burdensome (including making it easier to train systems, evaluate

their performance, and improve their accuracy) by developing GUI tools that incorporate functional requests from the operating department. We have also shown that the natural language processing AI we developed is not only able to quantitatively reduce the workload of operators, but also has a strong affinity with RPA tools and the improvement of work quality due to improvement of the workflow.

In the future, in parallel with the expansion of back-end functions such as the addition of new algorithms, we hope to promote front-end GUI tools for creating use cases inside and outside the company and increasing the number of organizations that use natural language processing AI.

## REFERENCES

[1] J. Devlin, M.-W. Chang, K. L. and K. Toutanova: "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proc. of NAACL, May 2019.
[2] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter and R. Vollgraf: "FLAIR: An easy-to-use framework for state-of-the-art NLP," Proc. of NAACL, Jun. 2019.
[3] Wikipedia home page. https://en.wikipedia.org/wiki/Main_Page
[4] Shinra Project: "Extended named entity." http://ene-project.info/?lang=en