

# Advanced Initiatives for Contact Center AI

*Kimihito Tanaka, Takashi Yagi, and Tetsuya Iizuka*

## Abstract

Contact centers are becoming increasingly important as a point of contact where feedback from many customers can be obtained. NTT Media Intelligence Laboratories is carrying out research and development of the application of artificial intelligence technology in contact centers. This article introduces some of the latest technologies for solving various issues at contact centers using the speech and natural language processing technologies that we have cultivated over many years.

*Keywords: artificial intelligence, contact center, digital transformation*

## 1. Introduction

It has been more than three years since the third artificial intelligence (AI) boom began to take hold in industry, and practical use of AI has begun in various fields. The NTT Group announced its corevo® AI brand in June 2016 and has been advancing various initiatives [1] since then. There is no doubt that the main technology accompanying the third AI boom has been machine learning, and deep learning in particular. Deep learning can be applied as a basic technology in various areas, but in most cases, simple application of the technology will not achieve the best performance when used in actual business. To achieve performance adequate for practical use, a network model specialized for the application must be created and trained using suitable data.

In the early days of AI, there was discussion regarding general AI as opposed to narrow AI. General AI refers to artificial intelligence that is able to solve any type of problem, similar to the human brain. We are still far from achieving this. In contrast, narrow AI can solve particular types of problems, which can be said of all AI available today. By specializing on a particular type of problem, narrow AI is often able to achieve performance equal to or better than humans. As such, an application area must be defined and the AI must be applied to the specific problem to use AI as currently available.

Contact centers are one area where NTT Media Intelligence Laboratories is focusing research and development (R&D) to apply AI technology. Contact centers respond to many telephone calls and chats daily and also conduct many knowledge searches and a lot of call analysis. Thus, it is a field where the speech and natural language processing technologies that we have cultivated for many years can be utilized. There are many contact centers within the NTT Group, and they can provide a lot of data across many fields that can be used in refining our technologies. Our objective is to use these environments and this store of data to create technologies that can be applied in real business scenarios.

## 2. Contact center issues

Contact center departments specialize in dealing with customers through various channels such as telephone, email, and chat. They were originally positioned as administrative centers for accepting applications or providing customer support, but with recent changes in business environments, they are becoming increasingly important as customer contact points that gather feedback from many customers. The contact center environment in Japan is beset with an increasingly serious shortage of personnel, and hiring and retaining operators is a serious challenge [2]. At the same time, service quality, or customer

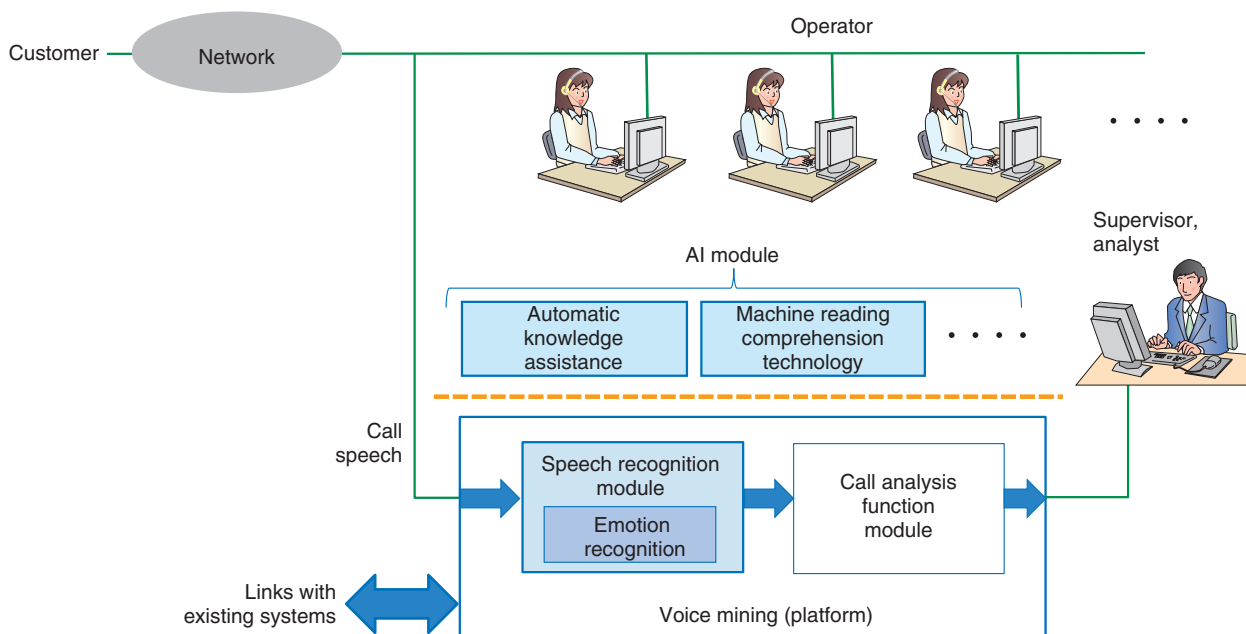


Fig. 1. Example deployment of voice mining technology.

experience (CX), must be improved. Finding ways to improve CX is another major issue for contact center management, especially in centers with high staff turnover and limited operational resources, where often there are personnel shortages and a number of inexperienced staff.

### 3. AI in contact centers

Various information technology systems have been introduced into contact centers, for example, those for interactive voice response, computer telephony integration, knowledge systems, and customer relationship management. AI is being used to make these systems more sophisticated, to provide new functionality, and to improve operational efficiency and CX.

A typical AI technology used in contact centers is voice mining. Voice mining has achieved dramatic increases in speech recognition accuracy through the use of deep learning, and it is now one of the most widely used AI technologies in contact centers. Through our subsidiary, NTT TechnoCross, NTT is developing and marketing a commercial voice mining technology called ForeSight Voice Mining [3]. It converts call speech to text using speech recognition and applies statistical analysis and visualization from a large volume of calls to obtain hints regarding issues with contact center operation, and how to solve

them [4].

Statistical analysis can be used, for example, to extract know-how from skilled operators and deploy it horizontally within the contact center, thereby increasing the overall skill level, or to compile statistical information for the entire center and quickly identify anomalies (such as rapid increases in a particular type of inquiry). Since the speech has been converted to text, it is also easier to review details after a call has been completed, increasing speed and accuracy when entering the call response history. These features are already being used in contact centers within and outside the NTT Group and are achieving results such as improved efficiency and increased sales.

An example of a voice mining technology configuration is shown in **Fig. 1**. The voice mining technology performs speech recognition on the call speech data, and the results are then analyzed by a call analyst. Voice mining has a role as an AI platform in the contact center, and other AI modules can be added to it to use speech recognition results in more sophisticated ways.

The NTT laboratories are also conducting R&D to extract information other than text, for example, emotional content, using AI. Such AI can be used to extend the speech recognition component and expand the scope of analysis. We are also advancing R&D on

AI that creates and presents knowledge based on the results of speech recognition in order to support operators in dealing with inquiries. These types of AI are configured so that they can be built using the voice mining technology as a platform. This enables them to be deployed quickly in real contact center environments.

#### 4. Initiatives for further advancement

---

Several new issues that need to be addressed have become apparent through commercial deployment of voice mining technology.

- (1) Further improvement of customer speech recognition accuracy
- (2) Creation of still more value using speech recognition results
- (3) Reduction of the cost and time involved in tuning the AI upon introduction

Currently, the accuracy of operator speech recognition in contact centers exceeds 90%. However, accuracy of customer speech recognition is generally ten percentage points lower. The customer speech can include more background noise, depending on their location, and if they are using a mobile phone, the sound quality can be degraded due to encoding and other factors. The speech may also be less stable and less formal than that of the operator, so there are several conditions that can make speech recognition more difficult.

The level of customer speech recognition accuracy is currently adequate for a person to understand what was said from the results and to use the results for statistical analysis at the individual word level. However, higher accuracy is required in order to use results in more sophisticated ways such as for knowledge creation or search. The article, “Evolution of Speech Recognition System—VoiceRex” [5] in the Feature Articles in this issue, introduces some of the latest technologies for resolving issue (1) above.

There is increasing demand when introducing voice mining technology to use speech recognition results in more sophisticated ways, to provide more advanced support for operators, and to reduce costs. Regarding issue (2), the NTT laboratories are focusing their efforts on R&D on knowledge support, which will facilitate acquisition of the knowledge needed for operators to handle inquiries.

The article “Toward Natural Language Understanding by Machine Reading Comprehension” [6] introduces technology that derives answers to questions from manuals and other documents. Knowledge used

in contact centers includes materials such as frequently asked questions (FAQs), manuals, and user agreements, but there are many contact centers that have not adequately organized their FAQs.

Machine reading comprehension promises to reduce the cost of organizing such knowledge by eliminating the need to prepare question and answer pairs ahead of time, as is done for FAQ search. In interviews, operators have also expressed opinions such as, “The manuals used in training are usually more familiar than the FAQs and are easier to use,” and, “When we have an inquiry that is not in the FAQs, we need to look at the manuals anyway,” so there is also increasing anticipation among operators for the convenience such technology will provide.

Another article in this issue, “Automatic Knowledge Assistance System Supporting Operator Responses” [7], introduces an automated knowledge assistance technology that links speech recognition with FAQ search, automatically presenting relevant FAQ entries at the appropriate time to operators while they are handling inquiries. Automatic knowledge assistance eliminates the time required for the operator to search for knowledge. The article also introduces efforts to increase the efficiency of organizing FAQ information using the speech recognition results.

The accuracy of AI technologies such as speech recognition can be increased by tuning them for each application area, so it is important to reduce the cost and time required for tuning, especially when deploying it in small-scale contact centers. For issue (3), the NTT laboratories are working on developing new technologies that improve accuracy and also working on reducing the tuning required for each installation by strengthening the underlying AI models. We have been collecting large amounts of training data from voice calls and other sources and creating industry-specific base models for speech recognition and knowledge search. In the major industries for which we have built industry-specific base models, we are now able to implement highly accurate systems, quickly and at low cost, by simply performing additional training with a small amount of training data.

#### 5. Expanding scenarios for using AI

---

The accuracy of speech recognition and natural language processing has increased dramatically due to the emergence of deep learning, increases in computer performance, and training with large amounts of data. These technologies are becoming common in

real business environments. In the future, we intend to pursue successful business projects with various partners, using the media processing technologies introduced here and others. We will also continue promoting R&D on innovative technologies to expand the range of applications for AI.

## References

- [1] Corevo, <http://www.ntt.co.jp/corevo/e/index.html>
- [2] Monthly Call Center Japan Editorial Dept., “Call Center White Paper 2018,” RIC Telecom, 2018 (in Japanese).
- [3] ForeSight Voice Mining, [https://www.ntt-tx.co.jp/eng/products/fore-sight\\_vm/](https://www.ntt-tx.co.jp/eng/products/fore-sight_vm/)
- [4] S. Kawamura, K. Machida, K. Matsui, D. Sakamoto, and M. Ishii, “Utilization of Artificial Intelligence in Call Centers,” NTT Technical Review, Vol. 14, No. 5, 2016.
- <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201605fa7.html>
- [5] T. Oba, T. Tanaka, and R. Masumura, “Evolution of Speech Recognition System—VoiceRex,” NTT Technical Review, Vol. 17, No. 9, pp. 5–8, 2019.
- <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa2.html>
- [6] K. Nishida, I. Saito, A. Otsuka, K. Nishida, N. Nomoto, and H. Asano, “Toward Natural Language Understanding by Machine Reading Comprehension,” NTT Technical Review, Vol. 17, No. 9, pp. 9–14, 2019.
- <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa3.html>
- [7] T. Hasegawa, Y. Sekiguchi, S. Yamada, and M. Tamoto, “Automatic Knowledge Assistance System Supporting Operator Responses,” NTT Technical Review, Vol. 17, No. 9, pp. 15–18, 2019.
- <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa4.html>



### Kimihito Tanaka

Senior Research Engineer, Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.S. and M.E. from the University of Tsukuba, Ibaraki, and an MBA from the University of Manchester, UK. He joined NTT Human Interface Laboratories in 1995 and studied speech synthesis technologies. His current research interests are innovation management and AI technologies.



### Tetsuya Iizuka

Vice President, Head of NTT Media Intelligence Laboratories.

He received a B.E. and M.E. in information engineering from Gunma University in 1989 and 1991. He joined NTT in 1991 and conducted research on database systems and data mining. He was also involved in the development of the SIP (session initiation protocol) server of Hikari Phone, development and support of open source software, personnel management, and branch office management.



### Takashi Yagi

Senior Research Engineer, Supervisor, Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.E. in electrical engineering in 1990 and an M.E. in computer science in 1992 from Keio University, Tokyo. He joined NTT Human Interface Laboratories in 1992. His research interests include human-computer interaction, computer-mediated communication, and AI. He is a member of ACM (Association for Computing Machinery), IEICE (Institute of Electronics, Information and Communication Engineers), IPSJ (Information Processing Society of Japan), and VRSJ (Virtual Reality Society of Japan).