Event Report: NTT Communication Science Laboratories Open House 2019

Atsunori Ogawa, Xiaomeng Wu, Masaaki Nishino, Mathieu Blondel, and Takemi Mochida

Abstract

NTT Communication Science Laboratories Open House 2019 was held at Keihanna Science City, Kyoto, on May 30 and 31, 2019. Around 1500 visitors enjoyed 5 talks and 30 exhibits, which included our latest research efforts in the fields of information and human sciences.

Keywords: information science, human science, artificial intelligence

1. Overview

NTT Communication Science Laboratories (CS Labs) aims to establish technologies that enable *heart to heart* communication between people and people, and between people and computers. We are thus working on a fundamental theory that approaches the essence of human beings and information, as well as on innovative technologies that will transform society.

NTT CS Labs Open House has been held annually with the aim of introducing the results of the CS Labs' basic research and innovative leading-edge research to both NTT Group employees and visitors from business industries, universities, and research institutions who are engaged in research, development, business, and education.

Open House 2019 was held at the NTT Keihanna Building in Kyoto on May 30 and 31, and around 1500 visitors attended it over the two days. This year, we invited the former athlete, Deportare Partners Representative, Mr. Dai Tamesue, and held a special talk with NTT Fellow Dr. Makio Kashino. We also tried an outdoor demonstration exhibition for the first time. We prepared many hands-on exhibits to enable visitors to intuitively understand our latest research results and to share a vision of the future where new products based on the research results are widely used. This article summarizes the event's research talks and exhibits.

2. Keynote speech

The event started with a speech by the Vice President and head of NTT CS Labs, Dr. Takeshi Yamada, entitled "Processing like people, understanding people, helping people—Toward the future where humans and AI will cohabitate and co-create" (**Photo 1**).

Dr. Yamada pointed out that CS Labs places particular importance on basic research not only toward the development of technology that can approach human abilities but also technology that can be used to elucidate human functions and characteristics and to understand what it means to be human, and technology to help people in their daily lives. In Japan, 2019 has turned out to be both a year marking Reiwa, a new name in the traditional Japanese era system, and a year on the verge of holding the Olympic and Paralympic Games. Dr. Yamada introduced the latest

Photo 1.   Dr. Takeshi Yamada delivering keynote speech.

Photo 2.   Dr. Kunio Kashino giving research talk.

artificial intelligence (AI) technologies developed by CS Labs at this transition point between eras and declared that they will boldly and tenaciously undertake new challenges with a focus on technologies that carry out processing the way people do, and that also understand and help people.

3.   Research talks

Three research talks were given, as summarized below, which highlighted recent significant research results and high-profile research themes. Each presentation introduced some of the latest research results and provided some background and an overview of the research. All of the talks were very well received.

(1)   "See, hear, and learn to describe—Crossmodal information processing opens the way to smarter AI," by Dr. Kunio Kashino, Media Information Laboratory

Recent advances in AI and machine learning research are breaking down the barriers between modalities such as language, sounds, and images, which have been studied separately so far. There are various implications for this dramatic change. Most importantly, AI is now achieving a new breakthrough—its ability to learn concepts on its own based on multimodal inputs. The key is to acquire, analyze, and utilize common representations that can be shared among those multiple modalities. In this research talk, Dr. Kashino called this approach crossmodal information processing, introduced its concept, and discussed how it will help people in their

future lives (**Photo 2**).

(2)   "Measuring multiple visual abilities in daily circumstances—Towards establishment of daily selfcheck for eye health," by Dr. Kazushi Maruya, Human Information Science Laboratory

The human visual system has considerable interpersonal differences, and its ability varies with the context, task, and circumstances. To grasp the variability in visual ability in daily circumstances, a novel set of eye health-check tests are proposed to measure visual abilities. Each test can be finished in a short time (around 3 min), and some tests are gamified so that users can check their visual ability in an enjoyable way. In this research talk, Dr. Maruya explained the details of this self-check test battery and clarified problems that should be overcome in order to establish daily self-checks of eye health (**Photo 3**).

(3)   "Like likes like strategy: search suitable for various viewpoints—Picture book search system 'Pitarie' with graph index based search," by Mr. Takashi Hattori, Innovative Communication Laboratory

A novel method is proposed for finding similar objects in a large-scale database, based on a graph index. Each vertex corresponds to an object, and two similar vertices are likely to be connected. The graph index, constructed using a *like likes like* strategy, shows small-world behavior: any two vertices can be connected by a small number of steps. A graph index search terminates quickly and is applicable to many types of media, including text, images, and audio. In this research talk, Mr. Hattori introduced Pitarie, an

Photo 3.   Dr. Kazushi Maruya giving research talk.

Photo 4.   Mr. Takashi Hattori giving research talk.

application of this proposed method that enables searching for similar picture books by both text and images (**Photo 4**).

4.   Research exhibits

The Open House featured 30 exhibits displaying NTT CS Labs' latest research results. We categorized them into four areas: Science of Machine Learning, Science of Communication and Computation, Science of Media Information, and Science of Human.

Each exhibit was housed in a booth and employed techniques such as slides presented on a large-screen monitor or hands-on demonstrations, with researchers explaining the latest results directly to visitors (**Photos 5** and **6**). The following list, taken from the Open House website [1, 2], summarizes the research exhibits in each category. (Abbreviations in the titles have been defined.)

4.1   Science of Machine Learning
- Learning and finding congestion-free routes—Online shortest path algorithm with binary decision diagrams
- Efficient and comfortable AC (air conditioning) control by AI—Environment reproduction and control optimization system
- Recover urban people flow from population data—People flow estimation from spatiotemporal population data
- Improving the accuracy of deep learning—Larger capacity output function for deep learning
- Which is cause? Which is effect? Learn from

data!—Causal inference in time series via supervised learning
- Forecasting future data for unobserved locations—Tensor factorization for spatio-temporal data analysis
- Search suitable for various viewpoints—"Pitarie": Picture book search with graph index based search

4.2   Science of Communication and Computation
- We can transmit messages to the efficiency limit—Error correcting code achieving the Shannon limit
- New secrets threaten past secrets—Vulnerability assessment of quantum secret sharing
- Analyzing the discourse structure behind the text—Hierarchical top-down RST (rhetorical structure theory) parsing based on neural networks
- When children begin to understand hiragana—Emergent literacy development in Japanese
- Measuring emotional response and emotion sharing—Quantitative assessment of empathic communication
- Touch, enhance, and measure the empathy in crowd—Towards tactile enhanced crowd empathetic communication
- Robot understands events in your story—Chat-oriented dialogue system based on event understanding

4.3   Science of Media Information
- Voice command and speech communication in

Photo 5.   Researcher explaining a demonstration.

Photo 6.   The latest research results were exhibited.

car—World's best voice capture and recognition technologies
- Learning speech recognition from small paired

data—Semi-supervised end-to-end training with text-to-speech
- Who spoke when & what? How many people

were there?—All-neural source separation, counting and diarization model
- Changing your voice and speaking style—Voice and prosody conversion with sequence-to-sequence model
- Face-to-voice conversion and voice-to-face conversion—Crossmodal voice conversion with deep generative models
- Learning unknown objects from speech and vision—Crossmodal audio-visual concept discovery
- Neural audio captioning—Generating text describing non-speech audio
- Recognizing types and shapes of objects from sound—Crossmodal audio-visual analysis for scene understanding

4.4 Science of Human
- Speech of chirping birds, music of bubbling water—Sound texture conversion with an auditory model
- Danswing papers—An illusion to give motion impressions to papers
- Measuring visual abilities in a delightful manner—Self eye-check system using video games and tablet PCs (personal computers)
- How do winners control their mental states?—Physiological states and sports performance in real games
- Split-second brain function at baseball hitting—Instantaneous cooperation between vision and action
- Designing technologies for mindful inclusion—How sharing caregiving data affects family communication
- Real-world motion that the body sees—Distinct visuomotor control revealed by natural statistics
- Creating a walking sensation for the seated—A sensation of pseudo-walking expands peripersonal space

5. Invited talk

This year, NTT Fellow Dr. Makio Kashino invited the former athlete, Deportare Partners Representative, Mr. Dai Tamesue and held a special talk entitled "Sports in the future and human potentiality." A wide range of topics was addressed, including *mental* control, record-breaking discontinuities, collaborative behavior, limit of growth due to premature optimization, language and sports skill communication, and balance of consciousness and unconsciousness. As an athlete who represented Japan in three Olympic Games and as a scientist who is studying brain functions underlying the amazing skills of athletes, respectively, Mr. Tamesue and Dr. Kashino talked about the essence of human nature revealed by the athletic practice and scientific research of sports, and shared their predictions of how science and technology will change sports and humans in the future. The audience engagement was high thanks to the vivid discussion that included a lot of physical gestures.

6. Concluding remarks

Just like last year, many visitors came to NTT CS Labs Open House 2019 and engaged in lively discussions on the research talks and exhibits and provided many valuable opinions on the presented results. In closing, we would like to offer our sincere thanks to all of the visitors and participants who attended this event.

References

[1] Website of NTT Communication Science Laboratories Open House 2019 (in Japanese).
http://www.kecl.ntt.co.jp/openhouse/2019/index.html
[2] Website of NTT Communication Science Laboratories Open House 2019 (in English).
http://www.kecl.ntt.co.jp/openhouse/2019/index_en.html

Atsunori Ogawa

Senior Research Scientist, Signal Processing Research Group, Media Information Laboratory, NTT Communication Science Laboratories.

He received a B.E. and M.E. in information engineering and a Ph.D. in information science from Nagoya University, Aichi, in 1996, 1998, and 2008. He joined NTT in 1998. He has been engaged in research on speech recognition and speech enhancement at NTT Cyber Space Laboratories (now, NTT Media Intelligence Laboratories) and NTT Communication Science Laboratories. His research interests include speech recognition, speech enhancement, and spoken language processing.

Xiaomeng Wu

Senior Research Scientist, Recognition Research Group, Media Information Laboratory, NTT Communication Science Laboratories.

He received a B.S. in energy and power engineering from the University of Shanghai for Science and Technology, China, in 2001, and an M.S. and Ph.D. in information science and technology from the University of Tokyo in 2004 and 2007. He joined NTT Communication Science Laboratories in 2013. His research interests include image processing, image retrieval, and pattern recognition.

Masaaki Nishino

Distinguished Researcher, Linguistic Intelligence Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.

He received a B.E., M.E., and Ph.D. in informatics from Kyoto University in 2006, 2008, and 2014. He joined NTT in 2008. His current research interests include data structures, natural language processing, and combinatorial optimization.

Mathieu Blondel

Distinguished Research Scientist, Ueda Research Group, NTT Communication Science Laboratories.

He received an engineering diploma from Telecom Lille, France, in 2008 and a Ph.D. in engineering from Kobe University, Hyogo, in 2013. He joined NTT Communication Science Laboratories in 2013. His current research interests include machine learning, mathematical optimization, the design of efficient machine learning software, and the application of these areas to real-world applications.

Takemi Mochida

Senior Research Scientist, Kashino Diverse Brain Research Laboratory, NTT Communication Science Laboratories.

He received a B.S and M.S. in engineering from Waseda University, Tokyo, in 1992 and 1994 and a Ph.D. in systems information science from Future University Hakodate in 2011. He joined NTT in 1994. His research interests include sensorimotor mechanisms in skilled human behavior.