Event Report: NTT Communication Science Laboratories Open House 2021

Kenichi Arai, Go Kato, Kenta Niwa, Hsin-I Liao, Daiki Nasu, and Tomoki Ookuni

Abstract

NTT Communication Science Laboratories Open House 2021 was held online, the content of which was published on the Open House 2021 web page at noon on June 3rd, 2021. The videos of 5 talks and 29 exhibits presented our latest research efforts in information and human sciences and were played more than 6000 times within the first month after the event.

Keywords: information science, human science, artificial intelligence

1. Overview

NTT Communication Science Laboratories (NTT CS Labs) celebrated its 30th anniversary this year. NTT CS Labs has aimed to establish cutting-edge 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 humans and information science, as well as on innovative technologies that will transform society. NTT CS Labs' Open House is held annually to introduce the results of our basic research and innovative leading-edge research with many hands-on intuitive exhibits to those who are engaged in research, development, business, and education.

Open House 2021 was held virtually due to measures to prevent the further spread of COVID-19. The latest research results were published with recorded lecture videos on the Open House 2021 web page at noon on June 3rd [1]. The content attracted many views within a month not only from NTT Group employees but also from businesses, universities, and research institutes. The event content is still available.

This article summarizes the event's research talks and exhibits.

2. Keynote speech

Dr. Takeshi Yamada, vice president and head of NTT CS Labs, presented a speech entitled "No matter how far apart, my heart will always be with you – Exploring the essence of communication that will create a spiritually rich society –," in which he looked back upon the history of NTT and establishment of NTT CS Labs then introduced present and future cutting-edge basic research and technologies (**Photo 1**).

For the past 30 years since its founding on July 4, 1991, NTT CS Labs has been developing innovative technologies such as media processing and machine learning that approach and even surpass human capabilities and seeking out fundamental principles in fields such as cognitive neuroscience and brain science that afford a deeper understanding of humans. After introducing the latest research results of NTT CS Labs from the three perspectives of "how to communicate accurately and efficiently" (information),

Photo 1. Keynote speech (Dr. Takeshi Yamada). "No matter how far apart, my heart will always be with you – Exploring the essence of communication that will create a spiritually rich society –."

"how to communicate by what means" (media), and "what is being indicated" (meaning), Dr. Yamada declared that NTT CS Labs will boldly and persistently tackle new challenges for the next 30 years to enable communication that can be conveyed to the heart through research that only NTT CS Labs can conduct.

3. Research talks

The following three research talks highlighted recent significant research results and high-profile research themes. Each talk introduced some of the latest research results and provided background and an overview of the research. After each talk, we received many questions from participants in real time for the question-and-answer (Q&A) session, confirming that the viewers had a high interest in our research.

(1) "The coming ages when AI becomes our conversation partners – Cutting edge of conversational systems with large-scale neural network models –," Dr. Hiroaki Sugiyama, Innovative Communication Laboratory

Dr. Hiroaki Sugiyama explained chatting dialogue systems for satisfying people's desire for dialogue through natural chatting with people. NTT CS Labs has been researching various dialogue systems for a long time, including research on giving systems personalities and improving the smoothness of dialogue by organically linking multiple dialogue robots. With the recent rapid development in deep learning, English social dialogue systems using large-scale deep learning have been proposed. He introduced details of NTT CS Labs' latest deep learning-based Japanese social dialogue system and its achievements and remaining issues (**Photo 2**).

(2) "Looking more, acting better – New concept of eye-hand coordination for skilled action –," Dr. Naotoshi Abekawa, Human and Information Science Laboratory

Dr. Naotoshi Abekawa explained skillful motor control using the mechanism of coordinated eye and arm motion as an example. One key issue in developing user-friendly information and communication technology is to understand "behavior" or "action," that is, to understand the essence of "motion." Although we seem to control motions easily, motor controls are achieved by sophisticated brain mechanisms, including the control of eye movements to obtain target information and the generation of limb movements. In his talk, he focused on the question, "Why is the eye important for skilled motor actions?", introduced explanations from the literature, and proposed an interpretation based on his findings (**Photo 3**).

(3) "Developing AI that pays attention to who you want to listen to – Deep learning based selective hearing with SpeakerBeam –," Dr. Marc Delcroix, Media Information Laboratory

Dr. Marc Delcroix discussed approaches to achieve computational selective hearing. Selective hearing refers to the ability to focus on listening to a desired speaker, even in a noisy environment such as a cocktail party. He first introduced SpeakerBeam, which is

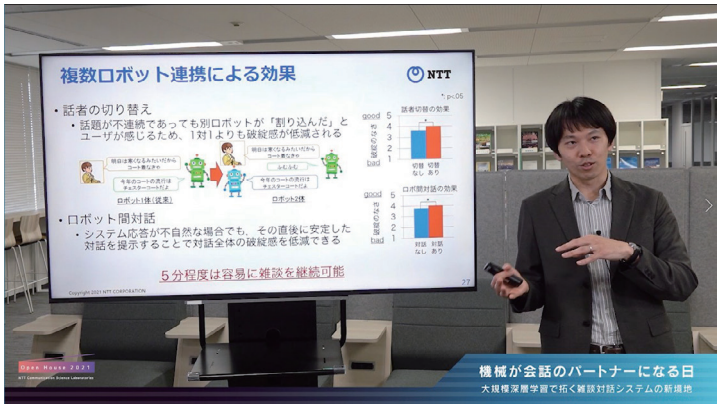

Photo 2. Research talk (Dr. Hiroaki Sugiyama). “The coming ages when AI becomes our conversation partners – Cutting edge of conversational systems with large-scale neural network models –.”

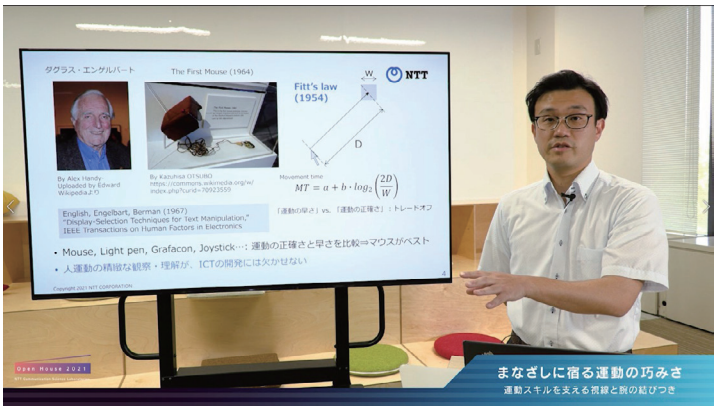

Photo 3. Research talk (Dr. Naotoshi Abekawa). “Looking more, acting better – New concept of eye-hand coordination for skilled action –.”

a deep-learning-based method he and his co-researchers proposed to extract speech of a desired target speaker in a mixture of several speakers by using a few seconds of pre-recorded audio data of the target speaker. He then presented recent research on (1) extension to multi-modal processing, where he and his co-researchers use video of the lip movement of the target speaker in addition to audio pre-recording, (2) integration with automatic speech recognition, and (3) the generalization to the extraction of arbitrary sounds (**Photo 4**).

4. Research exhibition

The Open House featured 29 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 Humans*. Each exhibit prepared videos explaining the latest results and were published on the event web page (**Photo 5**). Several provided online demonstrations or demo videos to make them closer to direct demonstrations. We also introduced a Q&A system in which page visitors can freely post questions and comments and CS Labs’ researcher can answer them. There were more than 130 public questions, and some expressed the enthusiasm of the visitors, such as long comments and professional questions.

The following list, taken from the Open House website, summarizes the research exhibits in each category.

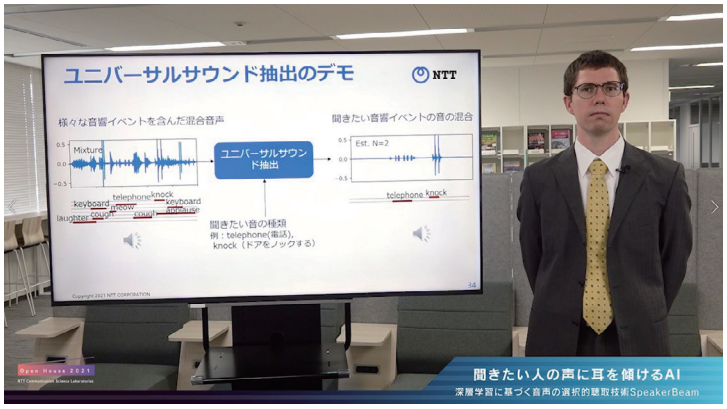

Photo 4. Research talk (Dr. Marc Delcroix). "Developing AI that pays attention to who you want to listen to – Deep learning based selective hearing with SpeakerBeam –."

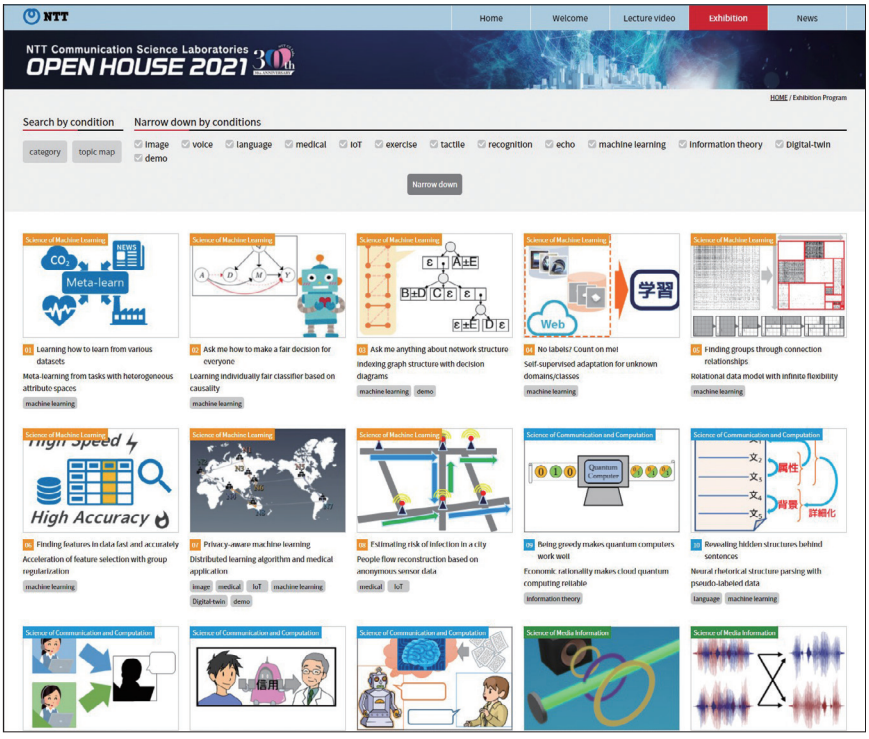

Photo 5. Exhibition web page.

4.1 Science of Machine Learning

- Learning how to learn from various datasets
  – Meta-learning from tasks with heterogeneous attribute spaces –
- Ask me how to make a fair decision for everyone
  – Learning individually fair classifier based on causality –
- Ask me anything about network structure
  – Indexing graph structure with decision diagrams –
- No labels? Count on me!
  – Self-supervised adaptation for unknown domains/classes –
- Finding groups through connection relationships
  – Relational data model with infinite flexibility –

- Finding features in data fast and accurately
  – Acceleration of feature selection with group regularization –
- Privacy-aware machine learning
  – Distributed learning algorithm and medical application –
- Estimating risk of infection in a city
  – People flow reconstruction based on anonymous sensor data –

4.2 Science of Communication and Computation

- Being greedy makes quantum computers work well
  – Economic rationality makes cloud quantum computing reliable –
- Revealing hidden structures behind sentences
  – Neural rhetorical structure parsing with pseudo-labeled data –
- Two experts, one result
  – Fusing two experts for enhancing their specialties –
- Can a chatbot mediate trust between humans?
  – Bridging doctor-patient rapport through a chatbot –
- Recipes for enjoy-talking conversational systems
  – Development of transformer-based conversational systems –

4.3 Science of Media Information

- Detecting faint sound by light
  – Non-contact sound measurement by precision interferometry –
- Extracting voices out of noise & reverberation
  – Joint signal separation, dereverberation, and noise reduction –
- AI that acquires knowledge just by watching TV
  – Crossmodal learning for concept acquisition of human movements –
- Real-time speech emotion controller using face
  – Emotional voice conversion via facial expression recognition –
- Visualizing touched places for Corona prevention
  – Touched places detection using heat traces by thermography –
- Telestethoscope: Looking into body by listening
  – Biomedical sound analysis utilizing physical characteristics –

4.4 Science of Humans

- Is falling birthrate related to population density?
  – Prediction and demonstration via life history theory –
- Move of magnetic fields, move by magnetic fields
  – Magnetact: A magnetic force-based tactile technology –
- Make hard objects soft, make rough objects smooth
  – Simple method for modulating tactile texture –
- Sense of touch connects our hearts beyond distance
  – Empathetic telecommunication by vibrotactile transmission –
- Why do mothers approach to infants' "crying"?
  – Oxytocin as a neural regulator of maternal implicit approach –
- Auditory attention that appears in the eye
  – Relation between microsaccades and auditory spatial attention –
- Blink pattern of elite car race drivers
  – Professional drivers always blink at the same time over laps –
- Essence of "keeping your eye on the ball"
  – Eye-hand coordination in motor learning –
- How do you define your dominant hand?
  – Quantifying motor-skill performance using a smartphone –
- Brain functions to recognize and hit a fastball
  – Brain mechanisms for quick judgment and motor control –

5. Special lecture

We asked Professor Asa Ito, professor of Tokyo Institute of Technology, to give a special lecture entitled "Bodies in the remote era," have a talk with NTT Fellow Makio Kashino, and review the research exhibitions. In the lecture, based on her research on haptic communication with the visually impaired, she talked about the problems of remote communication and the potential of the body (haptics). As an example of how trust was created through haptic communication, she shared an episode in which she simulated the experience of a visually impaired person being accompanied by a runner, and the rope connected to the runner quickly transmitted subtle information such as emotions and sensations, resulting in a surprising level of trust. She mentioned, however, that remote communication tends to be a one-way transmission of information, making it difficult to build a relationship of trust. She also introduced her research on watching sports with the visually impaired, remotely operated alter-ego robot communication, and memory of sensation and haptic perception. She mentioned that in haptic remote communication, we

can sometimes feel the existence of a person because we cannot touch their actual body, and that she has sensed the possibility for a new form of communication.

6. Concluding remarks

Open House 2021 was held as an event to present our latest results on a website. The lecture videos were viewed more than 6000 times in June by various segments of users. Participants provided many valuable opinions that have encouraged us to pursue further research activities through interactive means of communication with CS Labs researchers such as the Q&A system and Q&A session during the lectures. In closing, we would like to offer our sincere thanks to all the participants of this online event.

Reference

[1] Website of NTT Communication Science Laboratories Open House 2021, http://www.kecl.ntt.co.jp/openhouse/2021/index_en.html

Kenichi Arai
Senior Research Scientist, Signal Processing Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received a B.S., M.S., and Ph.D. from Waseda University, Tokyo, in 1991, 1993, and 2003. He joined NTT Communication Science Laboratories in 1993 and has been engaged in research on the application of nonlinear dynamics and stochastic systems to signal processing. His current research interests include physical random number generators based on a chaotic laser and the prediction of intelligibility using an automatic speech recognition system. He is a member of the Institute of Electronics, Information and Communication Engineers, the Physical Society of Japan (JPS), and the Acoustical Society of Japan (ASJ).

Go Kato
Senior Research Scientist, Computing Theory Research Group, Media Information Laboratory, NTT Communication Science Laboratories.
He received a doctor of science in physics from The University of Tokyo in 2004. He joined NTT Communication Science Laboratories the same year and has been engaged in the theoretical investigation of quantum information. He is especially interested in mathematical structures emerging in the field of quantum information. He is a member of JPS.

Kenta Niwa
Distinguished Researcher, Learning and Intelligent System Research Group, Innovative Communication Laboratory, NTT Communication Science Laboratories.
He received a B.E., M.E., and Ph.D. in information science from Nagoya University in 2006, 2008, and 2014. Since joining NTT in 2008, he has been engaged in research on microphone array signal processing. From 2017 to 2018, he was a visiting researcher at Victoria University of Wellington, New Zealand and involved with research on distributed machine learning and mathematical optimization. He is especially interested in distributed system optimization, such as model training on edge computing. He also belongs to NTT Computer and Data Science Laboratories. He was awarded the Awaya Prize by ASJ in 2010. He is a member of the Institute of Electrical and Electronics Engineers (IEEE).

Hsin-I Liao
Senior Research Scientist, Human Information Science Laboratory, NTT Communication Science Laboratories.
She received a Ph.D. in psychology from the Department of Psychology, National Taiwan University, in 2009. She joined NTT Communication Science Laboratories in 2012 and has been studying auditory salience, music preference, and preference of visual images. She has also explored the use of pupillary response recording to correlate human cognitive functions such as auditory salience and preference decision. From 2007 to 2008, she was a visiting student at California Institute of Technology, USA, where she studied visual preference using recorded eye movements and visual awareness using transcranial magnetic stimulation. She received a Best Student Poster Prize of the Asia-Pacific Conference on Vision in 2008, a Travel Award of the Association for the Scientific Study of Consciousness in 2011, and a Registration Fee Exemption Award of the International Multisensory Research Forum in 2011. She is a member of the Vision Sciences Society, Association for Research in Otolaryngology, and Japan Neuroscience Society.

Daiki Nasu
Research Scientist, Kashino Diverse Brain Research Laboratory, NTT Communication Science Laboratories.
He received a Ph.D. in medicine from the Graduate School of Medicine, Osaka University, in 2014. He joined NTT Communication Science Laboratories in 2016 and has been engaged in research on human motor control and biomechanics in sports. He has recently been focusing on the relationship between sports performance skill and brain function in athletes. He is a member of the Society for Neuroscience, the Japan Society of Physical Education, Health and Sport Sciences, the Japanese Society of Biomechanics.

Tomoki Ookuni
Senior Research Scientist, Research and Planning Section, NTT Communication Science Laboratories.
He received a B.A. in economics from Kyoto University in 1990 and M.A. in marketing science from Osaka Prefecture University in 2004. He joined NTT in 1990 and has been engaged in managing research and development at both NTT WEST and NTT since 1998. He joined NTT Communication Science Laboratories in 2019.