Project Humanity: Providing Intimate Support to Respect the Humanity of Individuals

Mariko Nakamura, Ryosuke Aoki, Asuka Ono, Iwaki Toshima, Kengo Okitsu, Atsushi Ando, and Takeshi Mori

Abstract

At NTT Human Informatics Laboratories, we are engaged in Project Humanity aiming at solving problems on the basis of the human-centric principle of respecting the humanity that each person values and in a way that does not burden the user. In this article, we introduce five case studies in our effort to achieve Project Humanity.

Keywords: communication, human understanding, motor reconstruction, neurodiversity, spinal cord injury

1. Enrichment of communication for people living with ALS

As amyotrophic lateral sclerosis (ALS) progresses, cognition remains normal while muscle strength throughout the body gradually loses function. People with ALS can live their lives to the full term by using a ventilator, but they lose their voice due to the tracheostomy surgery required to put on the ventilator. Thus, the choice to continue living comes with the trade-off of losing speech. Due to the double loss of communication through both spoken language and physical expression, many people fear disconnection from society and lose hope of living. Globally, more than 90% of ALS patients refuse to wear a ventilator.

We have been developing cross-lingual text-to-speech technology, a type of text-to-speech technology that reproduces a person's voice from a recorded voice. This technology makes it possible to communicate in multiple languages while maintaining the person's unique tone of voice. In 2022, a DJ artist who is unable to speak due to ALS, was able to share dialogue and a musical performance in English using his tone of voice. In 2023, he took on the challenge of communicating nonverbally by slightly moving his body.

For communicating his bodily expressions through an avatar in the metaverse, NTT Human Informatics Laboratories used motor-skill-transfer technology based on the artist's biosignal information. On stage at the Ars Electronica Festival in Linz, a festival of art, advanced technology, and culture, the artist with ALS performed via an avatar remotely manipulated from Tokyo by just slightly moving his body. He stirred up the crowd by reproducing his bodily expressions that excited the audience before the progression of ALS.

Surface electromyography (sEMG) can be measured with sEMG sensors when muscles are moved even slightly. When the artist with ALS moved the parts of his body that could be moved slightly, even a few millimeters, sEMG measurement could be taken (**Fig. 1**).

When a person intentionally moves a part of the body, muscles in other parts of the body may react. We observed the muscles of other parts of the body
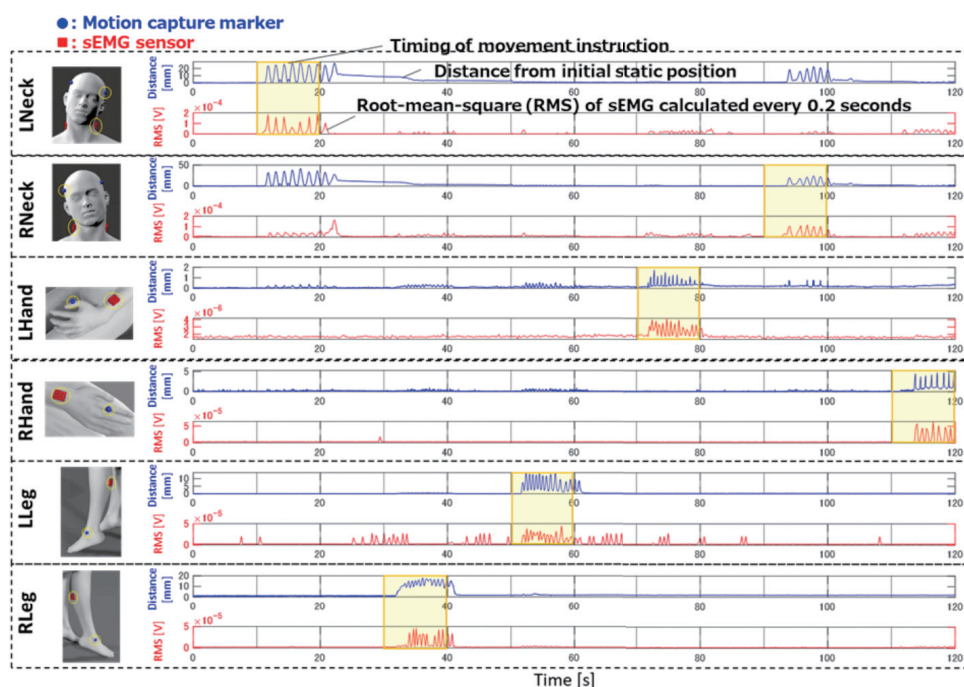
Fig. 1.   Time-series changes in the marker position at each body part as the distance from the initial static position (blue line) and the root mean square (RMS) calculated with the sEMG sensor (red line). The yellow highlights indicate the body parts instructed to move.

reacting during the interval (shown in yellow) in which the artist with ALS moved the indicated body part. Therefore, to obtain the sEMG of the muscles of the body part the person intends to move, it is necessary to calibrate the sEMG during the state of muscle enervation, such as at rest, as the reference values, and set a threshold for each body part to determine muscle contraction. The resting state of muscles immediately after the artist performed was set as the reference values, and the thresholds were set in accordance with his exertions during a performance.

Muscle fatigue is a point that should be considered when converting continuous sEMG into commands to operate an avatar in the metaverse. Muscle strength and endurance in individuals with severe physical disabilities deteriorate. It is thus necessary to enable intentional operating commands for the metaverse while conserving energy in the moving body parts. We thus adopted the policy of avoiding avatar commands that come from long, continuous muscle contractions or the degree of strength of muscle contractions. An avatar command is issued on the basis of the determination of muscle contraction of each body part. An avatar movement in response to an operation command is maintained for a certain period. If the

same bodily operation command is repeated during this time, the duration of the reflected avatar movement is extended. Such a policy may seem ineffective because it does not reflect the reality of body movements. However, the artist with ALS who experienced this technology felt the avatar was moving as he had intended. He felt the actual sense of operating the avatar through his own body movements.

2.   Development of information-prompt technology that supports the independence of people with dementia in their daily living

The decline in cognitive functions, such as memory and language, makes it challenging for individuals with dementia[*1] to perform tasks that were once routine, necessitating assistance in daily living. Even

---

*1   Dementia: "A syndrome due to disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgement. Consciousness is not clouded. The impairments of cognitive function are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behavior, or motivation." (From the International Classification of Diseases Version 10)

Feature 1. Presentation of information to stimulate start of task needed for daily living

1 "Let's have lunch!"
2 Heading to the kitchen.
3 When a PLWD loses track of their purpose …
4 Audiovisual clues are presented to them.
5 Recalling the purpose.

Feature 2. Presentation of information to stimulate continuation of task needed for daily living

9 Enjoying the sandwich.
8 The task is presented in a sequence of small steps.
7 When a PLWD loses track of their task …
6 "Let's make a sandwich!"

① Get two pieces of bread.
② Get the fillings from the refrigerator.
③ Butter one piece of bread.

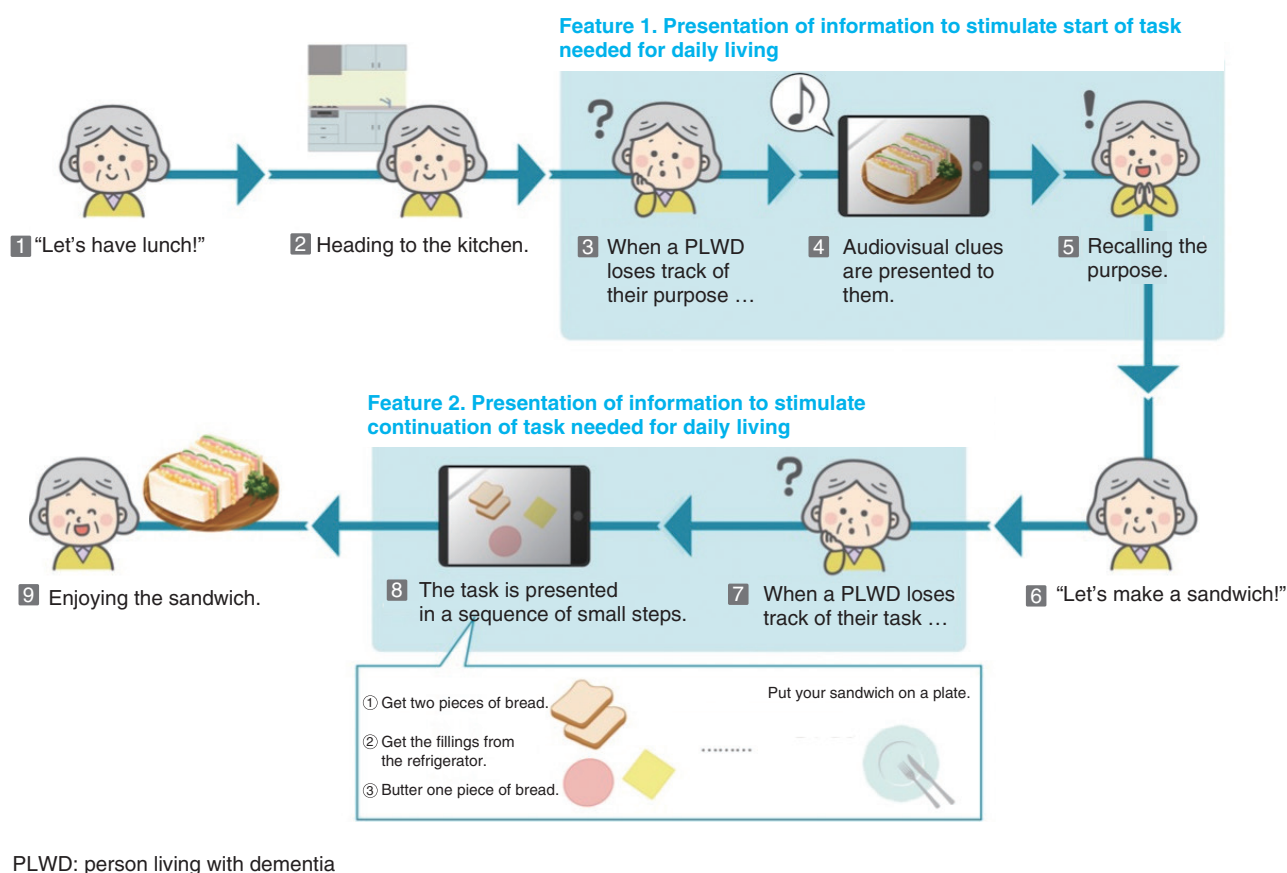Put your sandwich on a plate.

PLWD: person living with dementia

Fig. 2.   Features of information presentation technology to promote independence in people with dementia.

though the number of elderly people with dementia is increasing, there is a shortage of care professionals, and the difficulty of providing daily living support that meets the needs of each person has become a social issue. In terms of future societal trends, as a result of the enactment of the Basic Act on Dementia[*2] in June 2023 and the expected revision of the long-term care insurance system[*3] in FY2024, it is expected that the deployment of technology to address dementia will accelerate. While there are many technologies designed to watch over people with dementia, there are few technologies designed to assist them in their daily living. Developing such technologies in the field of dementia is thus highly important.

At NTT Human Informatics Laboratories, in collaboration with the University of Western Sydney and Deakin University in Australia, we have been conducting research and development (R&D) of information-prompt technology that encourages individuals with dementia to perform activities of daily living

(ADLs) independently. Specifically, we presented notifications and procedures related to tasks of ADLs to users with dementia through tablet devices and investigated how to effectively prompt the tasks using information such as images and sounds (**Fig. 2**).

Our results indicate the effectiveness of auditory prompts and the importance of dividing tasks appropriately on the basis of their complexity and familiarity [1]. The results also revealed user opinions highlighting the need to customize task division according

*2 Basic Act on Dementia: Abbreviated name of "The Basic Act on Dementia to Promote an Inclusive Society." "A law that establishes the basic principles related to dementia-related measures. It clarifies the responsibilities of the national government, local governments, etc., and sets forth the formulation of plans in relation to dementia measures, as well as the items that form the foundation of dementia measures." (translation by the author) (From "Basic Act on Dementia to Promote an Inclusive Society," Ministry of Health, Labour and Welfare of Japan (MHLW))

*3 Long-term care insurance system: "A system in Japan in which members of society as a whole support one another in caring for the elderly." (translation by the author) (From "Overview of the Long-term Care Insurance System," MHLW)

to each individual user's capabilities and preferences, and suggested that conventional support systems that provide uniform assistance to users are insufficient. We have also been studying the design of a notification system to facilitate attention management for completing tasks independently and studying an interface design for users to customize the system in accordance with their own lifestyle. Our results will be presented in future papers.

On the basis of the knowledge gained from the above collaborative research and other efforts, we have begun an investigation to expand this knowledge into an approach that takes into account "purpose in life" and "inclusive society," which lie beyond "independence," as our effort going forward. We are revising the concept of design with the aim of developing technologies that support communication for maintaining social connections with others in one's community and promoting psychological transformation to positively embrace dementia. Our efforts are currently at the stage of organizing issues, taking into account the perspectives of people with dementia by conducting surveys through questionnaires and interviews [2]. Collaboration with local high school students in the framework of an educational program is also underway to brainstorm ideas for addressing dementia-related challenges. By engaging in R&D while collaborating with individuals with dementia and high school students, we seek to create technology that is rooted in humanity living with dementia and in humanity shaping the future dementia society. In the future, we plan to apply this technology to care robots[*4] and information and communication technology that can comprehensively support living and life experiences of individuals with dementia from cognition to emotion.

3. Toward neurodiversity

Neurodiversity means perceiving differences between individuals in their brains and nervous system and various characteristics that arise from them as diversity, mutually respecting the differences, and making use of the diversity in society. For society, including communities such as workplaces, it means mutual respect and understanding and providing an environment in which it is comfortable for anyone to live. We expect that in communities where mutual respect and understanding is promoted, there is greater psychological safety and more active flow of information between members. As a result of such changes in the environment, greater creativity will be possible in companies, with diversity and psychological safety as the engine.

At NTT Human Informatics Laboratories, in collaboration with NTT Claruty Corporation, a special-purpose subsidiary of NTT, we discussed and agreed that elimination of miscommunication is most important in increasing psychological safety and allowing demonstration of creativity in a more diverse workplace. We believe that the largest cause of miscommunication lies in ambiguous expressions used in conversations during work. The workplace environment would greatly improve if the risk level of ambiguous expressions could be determined and pointed out during meetings. We are thus developing a prototype tool that notifies users ambiguous expressions. We have made improvements to the tool by using findings from its trial use. For example, a user requested that the tool should work well with other tools (web conferencing tools and various support tools). We thus learned the features needed for such a tool to satisfy a highly diverse workplace.

In developing this prototype, the following design concepts were implemented on the basis of discussions with NTT Claruty.

- Any person in any position should be able to look at the same screen and mutually confirm the notification of an ambiguous expression.
- During its use, a feedback mechanism should be provided for the notification of the ambiguous expression. The criteria for notification should be easily modifiable.
- In determining the risk level of an ambiguous expression, more consideration should be given to members whose communication errors could easily lead to serious issues such as leaving their jobs, for example, members on the autistic spectrum (members with autism spectrum disorder).
- Interference with current support software should be minimized.
- The tool should have a highly flexible configuration for conversation interventions that meet the characteristics of each individual.

The screen of this prototype for ambiguous expression detection is shown in **Fig. 3**. All speech is transcribed using speech-recognition software. When an ambiguous expression is discovered, users are notified through a pop-up window in which the word is highlighted by changing its color, font, etc., in line

---

*4   Care robot: "Care equipment in which robot technology is applied to support patients in being independent and reduce the burden on care partners." (translation by the author) (From "Development and Dissemination of Care Robots," MHLW)
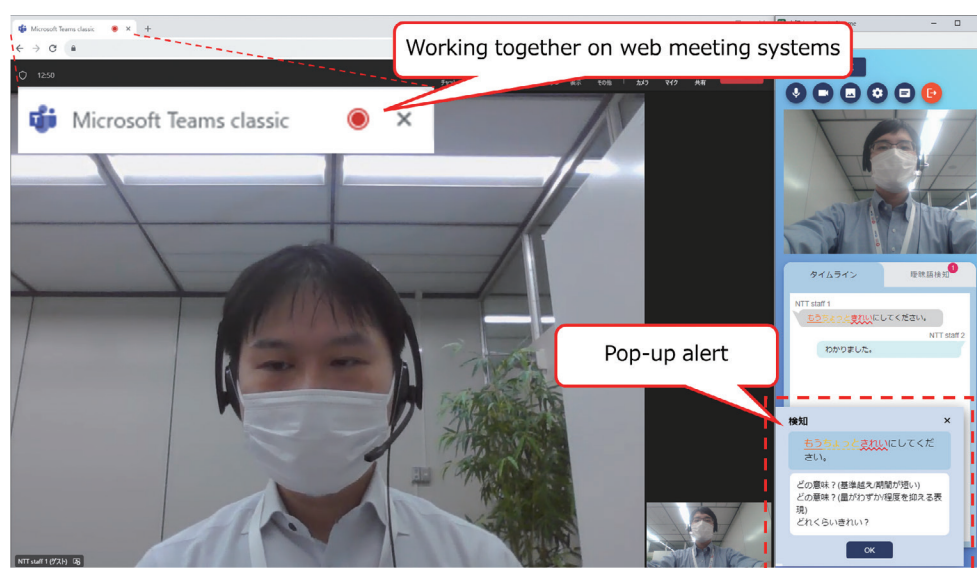
Fig. 3.   Prototype of a communication support tool based on ambiguity detection.

with the level of assessed risk of miscommunication. (The risk level is a numeric value that expresses the predicted severity if miscommunication due to the ambiguous expression occurs. Its initial value is determined on the basis of the opinions of specialists in the field of employment support for people with disabilities and adjusted on the basis of feedback by trial users of this tool.) The tool is used to promote more detailed discussion of ambiguous expressions. According to a survey after the trial use, users reported that this tool increased their sense of self-efficacy. They gave comments such as "This tool made it easier to form clear agreements" and "It made it easier to understand instructions" (**Fig. 4**).

Going forward, we plan to use an improved version of the tool and confirm its performance in terms of user experience and effectiveness. We will study adding functions that enable understanding of the internal state of the user by continuing discussions with stakeholders. Such functions will enable understanding of the user's level of understanding of conversations, their familiarity with work, and their degree of concentration.

4.   Movement support for people with spinal-cord injuries

The number of people with quadriplegia due to spinal-cord injury and who have difficulties with daily activities is increasing annually. In Japan, it is reported that about 6000 people per year acquire such injuries [3]. Movement support for people with spinal cord injuries has thus become an important issue. As an effort worldwide to address this issue, invasive brain-computer interfaces (BCI) are being developed mainly by research medical institutes to restore the patient's ability to move. To restore movements with invasive BCI, electrodes are surgically implanted in the brain and muscles (or nerves), and electrical stimulation is applied to the muscles on the basis of the results of analysis of brain activity information.

At NTT Human Informatics Laboratories, while respecting individuals' diverse lifestyles, we believe that movements on one's initiative using one's body is one of the components of well-being in daily living and in society. We are engaged in the development of invasive BCI technology to restore arm movements, which are essential for daily activities such as eating. Specifically, we are studying technology that outputs electrical muscle stimulation from brain activity to restore muscle-coordination movements. Our goal is to restore not only relatively simple movements such as wrist flexion, which has been achieved in previous research, but also daily-life movements that require more complex muscle coordination such as drinking water from a cup.

To implement this technology, we are developing artificial intelligence (AI) technology that extracts from brain activity what kind of muscle activity is intended on the basis of the pattern of muscle synergy,
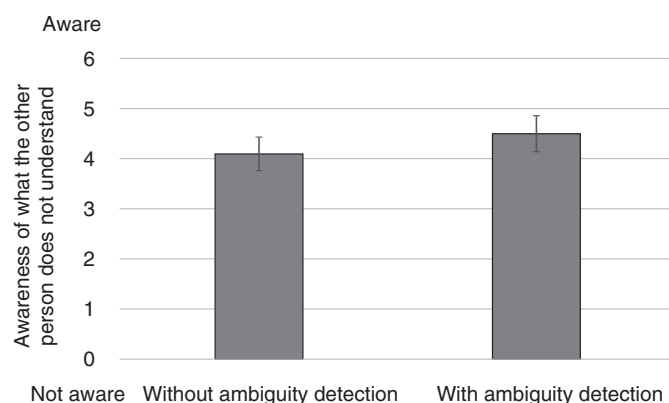
Fig. 4.   Increased frequency of awareness.

which is a mechanism of muscle coordination that the spinal cord possesses. Muscle synergy refers to the pattern of coordination of multiple muscle activities that appears during a body movement. It is considered to be the result of the brain's control of nerves in the spinal cord that are supplied to multiple muscles. On the basis of this muscle synergy pattern, we constructed a model that converts brain activity to muscle activity on the basis of biological data obtained from people with healthy arms. By incorporating the estimated muscle synergy in the BCI system for individuals with spinal-cord injury, we expect it can compensate for the muscle coordination impaired due to the injury and restore movements that require complex muscle coordination.

Conventionally, the estimation of muscle synergy has been conducted solely on the basis of muscle-activity data, resulting in estimations of how muscles have coordinated. We believe that muscle synergy that reflects nerve connections in brain-spinal cord-muscles are easier to control from the brain. We have thus developed a muscle-synergy-estimation method that takes into account brain-activity data acquired at the same time as muscle-activity data. Specifically, our method uses a deep learning model that features layers to simulate muscle synergy and estimate muscle synergy in the process of training muscle activity output from brain activity input. Comparing the conventional method with our method by using electrocorticography of the monkey motor cortex, the region of the brain involved in muscle activity, and muscle-activity data, we confirmed that our method can estimate muscle synergy with greater accuracy.

Going forward, we will investigate the method of electrical stimulation of multiple muscles to achieve actual muscle activity from the estimated muscle activity. We also plan to develop and test a BCI system in combination with the muscle-synergy model and the method of electrical muscle stimulation to confirm the effectiveness of the system in restoring muscle-coordination movements.

5.   Depression-symptom-detection technology

Support for mental health has become a major social topic. Depression (major depressive disorder), while familiar, is known to severely impair daily living. It has been reported that about 6 out of every 100 Japanese have experienced depression, and survey results indicate that Japanese society loses about 2 trillion yen annually due to depression in its members. The number of mental disorder cases has been on the rise as telework spreads and more people live alone. Reports state that the number of patients with depression more than doubled between 2013 and 2020. Because the more severe depression becomes, the harder it is to treat; thus, early detection and treatment are crucial. However, people around a person with depression and the afflicted person often do not recognize depression symptoms, and the case worsens [4, 5].

NTT Human Informatics Laboratories is developing AI technology that can detect symptoms of depression in daily life and promote early treatment. As a part of this effort, we have been working to create technology that can easily detect depressed mood, a major symptom of depression, by applying media-processing technologies cultivated at NTT laboratories over a long time. The technology detects depression from the responses of the user's voice and their
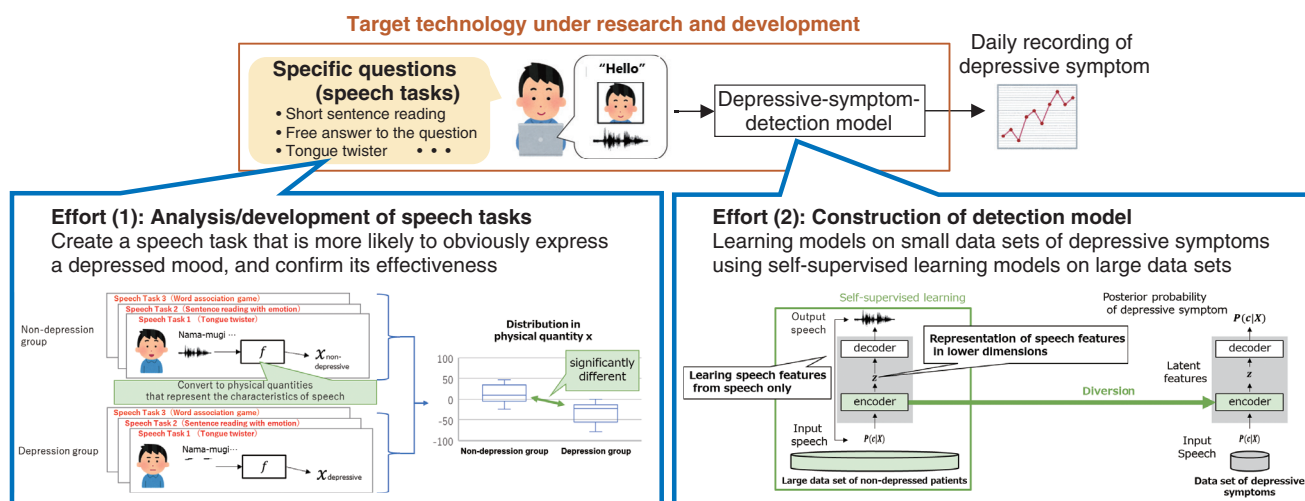
Fig. 5.   Efforts to create depressed-mood-detection technology.

expressions in response to specific questions (**Fig. 5**).

We collected and analyzed voice and video data of people with depressed mood. To date, we have collected voice data from more than 100 people, including those diagnosed by physicians as having a depressed mood. Comparison and analysis of the voice features of people with depressed mood and those without revealed the expressions of features such as the following in those with depressed mood: (1) lack of emotional expression (little difference between normal voice and voice when expressing happiness or anger), (2) lowered cognitive function (delay in answering questions, fewer words spoken), and (3) decline in articulatory function (difficulty speaking quickly) [6]. On the basis of these findings, we created multiple questions (speech tasks) suitable for detecting depressed mood and confirmed their effectiveness (Effort (1)). We are also constructing an AI model that can detect depressed mood (Effort (2)). This model is being developed on the basis of self-supervised learning[*5] to detect depressed mood with high accuracy from a small dataset of voice data belonging to people with depression symptoms. The AI model is trained in advance on the features of general speaking voices by using a large dataset of the voices of people without depression symptoms. It facilitates learning by the AI model of voice features of people with depressive symptoms.

We plan to develop services that prevent mental disorders and contribute to early recovery by using the above depression-symptom-detection technology to stimulate early mental care while monitoring daily living conditions. For the mental health management of employees in a company, for example, we plan to develop an AI agent service that can detect employees having depression symptoms from their daily communication and connect them to early treatment. Therefore, this technology can contribute to promoting health management in companies by implementing early prevention of mental disorders, reducing the need to take leave or retire due to depression.

References

[1] M. J. Radnan, R. Brookman, S. Parker, L. Hoon, H. Kawata, A. Ono, and C. Harris, "User Interface Preferences of Assistive Technology Device for People Living with Dementia at Home," MARCS Conference Series: Ageing, Westmead, NSW, Australia, Aug. 2022.

[2] A. Ono, A. Yamanaka, S. Seko, and H. Matsukawa, "Communication Problems for People Living with Dementia," Society for Cocreationology 7th Annual Conference, pp. 66–69, 2023 (in Japanese).

[3] N. Miyakoshi, K. Suda, D. Kudo, H. Sakai, Y. Nakagawa, Y. Mikami, S. Suzuki, T. Tokioka, A. Tokuhiro, H. Takei, S. Katoh, and Y. Shimada, "A Nationwide Survey on the Incidence and Characteristics of Traumatic Spinal Cord Injury in Japan in 2018," Spinal Cord, Vol. 59, pp. 626–634, 2021. https://doi.org/10.1038/s41393-020-00533-0

[4] OECD, "Tackling the Mental Health Impact of the COVID-19 Crisis: An Integrated, Whole-of-society Response," May 2021. https://oecd.org/coronavirus/policy-responses/tackling-the-mental-health-impact-of-the-covid-19-crisis-an-integrated-whole-of-society-response-0ccafa0b/

[5] M. Sado, "What Are Society's Losses As a Result of Depression?—Cost-of-illness Study of Depression," Psychiatria et Neurologia

---

*5 Self-supervised learning: A machine learning method that uses a large amount of unlabeled data to acquire in advance features from the data for the model. After a model is generated through self-supervised learning, it is possible to build an accurate inference model by using a re-training method such as fine-tuning, using a small amount of labelled data.

Japonica, Vol. 116, No. 2, pp. 107–115, 2014 (in Japanese).
[6] Y. Murata and A. Ando, "Speech Analysis of Depression Patients for Selection of Speech Tasks to Detect Depression Symptoms," Acoustical Society of Japan 2023 Spring Research Presentation Conference, 3-4P-4, 2023 (in Japanese).

Mariko Nakamura
Senior Research Engineer, NTT Human Informatics Laboratories.
After joining NTT, she researched environmental impact assessment of telecommunication services. She then become interested in human-centered design in communication while studying risk-communication regarding chemical substances, disasters, and food. She is currently promoting projects in collaboration with different companies in various industries trying to solve communication issues with new technologies from NTT R&D.

Ryosuke Aoki
Senior Research Engineer, NTT Human Informatics Laboratories.
He received an M.S. and Ph.D. in information sciences from Tohoku University, Miyagi, in 2007 and 2014 and joined NTT in 2007. His current research interests include human-computer interaction, interaction design, research through design, medical engineering, human augmentation, and wearable/ubiquitous computing.

Asuka Ono
Researcher, NTT Human Informatics Laboratories.
She received a B.S. and M.S. in design from Kyushu University, Fukuoka, in 2016 and 2018. After joining NTT in 2018, she focused on hearing assistive technology with tactile speech enhancement from 2019 to 2020. From 2021 to 2022, she worked on prompting technology to support ADL for individuals with dementia. She is currently investigating communication challenges faced by people with dementia. She envisions contributing to the empowerment of individuals with diverse abilities through the development of inclusive technological solutions rooted in cognitive science. In addition to her research, she actively engages in science programs for high school students.

Iwaki Toshima
Senior Research Engineer, NTT Digital Twin Computing Research Center.
He received an M.E. from the University of Tokyo in 2002 and joined NTT the same year. He received a Ph.D. in engineering from Tokyo Institute of Technology in 2008. His research interests include human-computer interaction, robot audition, and digital twin computing.

Kengo Okitsu
Senior Research Engineer, Cybernetics Laboratory, NTT Human Informatics Laboratories.
He received a B.E. and M.E. from Tokyo Institute of Technology in 2008 and 2010. His research interests include brain machine interfaces.

Atsushi Ando
Senior Research Engineer, Human Activity Understanding Research Group, Human Insight Laboratory, NTT Human Informatics Laboratories.
He received a B.E., M.E., and D.E. from Nagoya University, Aichi, in 2011, 2013, and 2021. Since joining NTT in 2013, he has been engaged in research on speech recognition and non-/paralinguistic information processing. He received the Awaya Prize Young Researcher Award from the Acoustical Society of Japan (ASJ) in 2019. He is a member of ASJ, the Institute of Electronics, Information and Communication Engineers (IEICE), the Information Processing Society of Japan (IPSJ), and the International Speech Communication Association (ISCA).

Takeshi Mori
Senior Research Engineer, Human Activity Understanding Research Group, Human Insight Laboratory, NTT Human Informatics Laboratories.
He received a B.E. and M.E. from Tokyo Institute of Technology in 1994 and 1996 and D.E. from University of Tsukuba in 2007. Since joining NTT in 1996, he has been engaged in research on speech and audio processing algorithms. He is a member of ASJ and Institute of Electrical and Electronics Engineers (IEEE).