National Institute of Informatics News ISSN 1883-1966 (Print) ISSN 1884-0817 (Online)

NII Today

66
June. 2018

National Institute of Informatics News

NII Interview

Three-way conversation

What We Gain from the Digitalization of Medical Care

Makoto Suematsu [President, Japan Agency for Medical Research and Development]
Masaru Kitsuregawa[Director General, National Institute of Informatics]
Jun-ichi Taki[Senior Writer, NIKKEI INC.]

Dialog

Aim Behind the Succession of Medical Research Centers Being Built

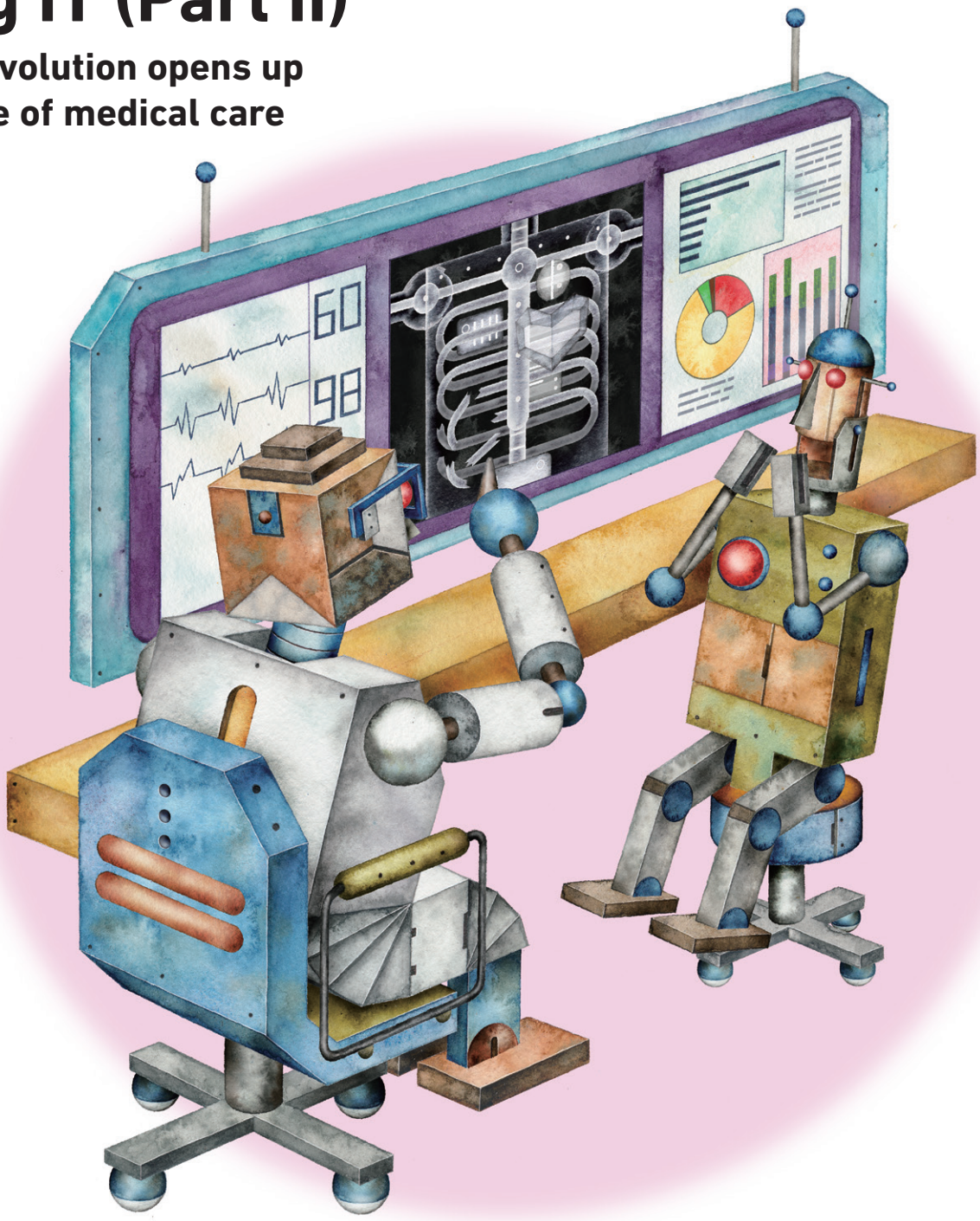
Tomoyuki Higuchi [Director General, The Institute of Statistical Mathematics]
Masaru Kitsuregawa[Director General, National Institute of Informatics]

Taking on Health, Medical and Social Challenges with Machine Learning

Feature

New Support for Medical Care Using IT (Part II)

Digital revolution opens up the future of medical care



This English language edition
NII Today corresponds to
No. 80 of the Japanese edition

What We Gain from the Digitalization of Medical Care

Putting IT to work in diagnostic screening and diagnostic support

Makoto Suematsu
President, Japan Agency for Medical Research and Development

Masaru Kitsuregawa
Director General, National Institute of Informatics

Interviewer: Jun-ichi Taki
NIKKEI INC. Senior Writer

Under the banner of a "Digital Revolution in Medical Care," the Japan Agency for Medical Research and Development (AMED) aims to raise the quality of medical services by changing clinical practice through the use of information technology. The National Institute of Informatics (NII) joined in this effort and newly established the Research Center for Medical Bigdata. AMED President Makoto Suematsu and NII Director General Masaru Kitsuregawa talked about the background and aims of the digital revolution.

Taki The foremost mission of AMED is to make it possible for the latest available research results in the medical field to be put to use in clinical practice at the very first opportunity. First, Mr. Suematsu, I'd like to ask you about the goal of achieving a digital revolution in medical care, which AMED is actively promoting. You're bringing together a number of individual projects under the name of "Achieving a Digital Revolution", and I'd like to ask you about the background behind these projects, as well as their aims.

Suematsu The foremost aim is to put information technology (IT), artificial intelligence (AI), and other fields to work in clinical practice, in order to improve the quality of medical care being delivered to the people of Japan. Let me give you one example that's pretty straightforward. One project that we are enthusiastic about, working with NII, is what we call "The Four Brothers of Imaging". What we're trying to do with this project is to create an infrastructure for utilizing big data from medical images. The "four brothers" involved in the project are the Japan Gastroenterological Endoscopy Society, the Japanese Society of Pathology, the Japan Radiological Society, and the Japanese Ophthalmological Society, and we call them "brothers" because they're working together to provide comprehensive cooperation in the project.

To describe the situation briefly, the workforce in medical care today is facing some severe working environment challenges, and physicians and medical staff are constantly constrained. At the same time, patients want to be diagnosed quickly and accurately, and to be able to choose an optimum treatment for their particular condition. We're looking at how we can make image-based diagnosis faster and more accurate, and how we can relieve the workforce from the burden. We believe that utilizing big data from medical images is one feasible approach.

Working with the four societies, we can compile a large number of images such as ultrasound scans, computerized tomography scans, and magnetic resonance images (MRIs). Using AI technologies, we can identify the particular site in whatever organ has been imaged, and, for example, if it's the gallbladder, the system can automatically distinguish whether it's looking at a polyp that needs to be carefully watched for a number of years or needs to initiate immediate treatment. Physicians who work in radiological diagnosis are extremely busy, so if we can use AI for whatever it's capable of, the doctors can be freed up to concentrate on more difficult diagnoses. This will undoubtedly provide a great benefit to patients.

Multiple modalities can provide more accurate diagnoses

Taki This is a question that comes up frequently in fields other than medicine as well, but do you see AI replacing some of the work done by specialists in the future?

Suematsu I don't think it will replace what they do. Rather, I think that some things can be done faster and more accurately using machines, but more subtle diagnoses are better made drawing on the implicit knowledge and experience that only physicians can provide. At the current stage, we're evaluating whether screening that requires careful diagnosis based on numerous images can be done using AI. Genomic medicine has made its debut, and molecularly targeted agents are becoming available, but that doesn't mean that breakthrough treatments are available for the entire spectrum of cancers. It's important to detect and diagnose cancer early through imaging.

Taki If we're talking about screening, it seems like it could be done with the technologies we have now.

Masaru Kitsuregawa

Suematsu Research in pathology has shown that sometimes pancreatic cysts appear to be benign at first, but after a number of years have passed, they turn malignant. If we could accumulate records of cases like these, we could know when we first see a cyst what kind of characteristics require particular caution, and we could use deep learning to be able to foresee what will happen in the future at the stage when the cyst is first diagnosed.

Taki That would require gathering image data from a large number of patients in a time series, wouldn't it?

Suematsu One other thing that's important is the modality, meaning devices for medical imaging diagnosis. For instance, something might look one way on an ultrasound but might look different on an x-ray CT. It's important to accumulate information about a single patient using multiple modalities. I've come up with a term that I call a "biomarker signature". What I mean by that is, say you took my name, Makoto, and you wanted to use that information to make a definite identification of a disease state. You would use all of the letters, and not just the first M and A. If you don't have one organization linking everything together, each society would compile information using its own format, and you could end up with each one developing its own AI for diagnostic purposes. Each society has its own issues and challenges; an ultrasound is an ultrasound, and pathology is pathology. If each one tries to address issues on its own, we would end up with each developing its own approach.

But if you look at it from the patient's point of view, after a number of years have passed, when we have a medical "My Number" personal identification system in place, patients will want to have definite diagnoses based on a combination of different modalities. If we turn around at that point and decide to integrate, I'm not sure it will go all that well.

Taki I'm sure it definitely would not go well.

Suematsu Well, we don't know. But if we're going to need to integrate anyway, it would be better to understand from the outset what each of us wants to do and then develop a common platform. We initiated this project because we think we can draw on NII's knowledge and wisdom to achieve that.

If you'll let me make just one more comment, people tend to think that if we just take the information that has been read from ultrasounds and CTs in general hospitals, and have AI learn it, then we'll be able to make diagnoses, but it's not actually that simple. Not everything is black or white; you can't always say that one particular case is tuberculosis and another isn't. There are many gray areas where it's not tuberculosis but it may not be normal, either. How do we take that kind of know-how, or the implicit knowledge that physicians have, and teach it to machines? How do we achieve deep learning of those gray areas? I'm hearing from people on the clinical side as well that they don't know how to go about this.

Putting medicine and engineering on an equal footing

Taki It's a difficult problem, isn't it? Dr. Kitsuregawa, from an NII standpoint, how do you view the project to digitalize medical care?

Kitsuregawa We've moved out of the era when we were conducting research in IT itself and we are now looking at how we can use IT to bring about innovative changes in society. There's a big shift underway from "of IT" to "by IT". As we focus seriously on the actual application of IT, new issues will emerge

Makoto Suematsu

Graduated from Keio University School of Medicine in 1983; after serving as an assistant in the Department of Internal Medicine at Keio University School of Medicine, became an exchange student at the Institute for Biomedical Engineering, the University of California San Diego, in 1991. Returned in 2001 to Keio University School of Medicine as a professor in the Department of Biochemistry and Integrative Medical Biology, and served from 2007 until March 2015 as the dean of the School of Medicine. Has served as the president of the Japan Agency for Medical Research and Development since April 2015.

that IT people will have to work on in earnest. Professor Takeo Kanade of Carnegie Mellon University in the US, who was awarded the Kyoto Prize in 2016, actually succeeded in having a car drive itself from the east coast to the west coast of the US back when the concept of self-driving cars didn't even exist. He understood that, by taking on the challenge of autonomous driving, which will absolutely be necessary in the future, IT people would need to identify and understand problems that will come up as a result, and he took on that challenge.

Right now, vast amounts of data exist in the clinical workplace. Rather than simply letting the data sit there, could we be actively utilizing the data to address issues that physicians are facing in the workplace? That's the challenge that this project is taking on. IT researchers are enthusiastically taking part in the project as well. I have to admit that, in the past, we had a distinction between medical and engineering personnel, where medical people had a superior role and engineering people worked in a supportive role. But in this project, we're putting medical people and engineering people on an equal footing, under equivalent conditions. Nowadays, IT is capable of looking at things at a level largely equivalent to that of physicians, and people in IT can speak up on an equal basis.

Taki I'd like to ask the same question that I asked Mr. Suematsu. Do you think that AI can replace specialists?

Kitsuregawa Mr. Suematsu has already clearly broken down the problems into pieces, but I think that they can also be broken down into a number of types. First of all, it's equally easy for physicians and IT to distinguish between cases that are obviously normal and those in which disease is obviously present. When it comes to the gray areas, though, there are many subtle problems. We don't know whether a specialist is diagnosing a patient as normal because they really know that he or she is normal or whether they are overlooking something pathological. Actually, in our project, there was a case in which a physician judged an image to be normal, but IT judged it to be abnormal. When we apologized to the physician and said that our technology was not yet sufficiently developed, the specialist went back to the image and looked at it carefully, and then came back and said, "Hold on. It's possible that we didn't look at the image in the right way." That put us on a friendship basis. The specialist had to make a very subtle judgment, and IT was able to give some advice and point out an area that should be examined more carefully. In that moment, we were able to take a step into a world in which AI was not threatening to take away people's jobs, but in which human beings and machines can work together.

Suematsu That was the moment at which physicians and IT researchers could start to trust each other, wasn't it? Being able to acknowledge each other means trust. If there is no relationship of trust, people simply keep information to themselves. What we really hope to see from the digitalization of medical care is that the medical workforce will, through technology, move in the direction of building trust across professions and organizations, and that more information will be shared.

Compartmentalization of researchers poses problems

Kitsuregawa As one other type of problem, let's think about metastatic cancer. If different organs harbor the cancer, then specialists in each of those organs should have evaluated organs that are not in their own field. Undoubtedly, their ability to read images involving those organs is going to be somewhat less than their ability to evaluate organs in their own field of specialization. We can anticipate that AI will do a little better than the physician in a case like that. In other words, AI would be able to help the physician.

Additionally, recent CTs, for instance, are able to produce vast amounts of data in just a few seconds. Who's going to be looking at those data? Measurement technology has advanced far beyond human capabilities. Actually, this applies to fields other than medicine as well, but we need to provide measuring instruments with functions that enable them to view images. The era when we simply used machines for measurement and left the rest to human beings is over. As this illustrates, it's feasible that doctors and IT will be able to coexist in a number of different forms.

Taki Has the imaging project already reached a stage where judgments made by humans and by machines can be compared?

Kitsuregawa In November 2017, we opened a cloud environment where image data provided by the various societies have been compiled, and we officially began analyzing the data. We've only just started. We've got the ecosystem in place now for handling the data, so I think you can look forward to some good developments.

Taki Does that mean that you've developed a platform capable of handling data from multiple modalities in a time series?

Kitsuregawa That's the primary core component where researchers who are in the mainstream of IT can really help us. What made NII really want to initiate this project was the realization that it would test the comprehensive strength of IT. The societies are taking data from hospitals, anonymizing the data, and then sending the datasets to NII's cloud server. Carrying out a series of steps involves communications technology, as well as security technology and cloud technology. We need to create a platform that can be accessed by anyone and will allow them to run an application for deep learning. Software engineering is a critical part of that. That's exactly where the comprehensive strength of IT will be needed.

Researchers in deep learning tend to say that if you hand them the data, then they will have no problem with doing the analysis, but they want the data to be properly organized. Actually, the analysis is less than 10% of the overall volume of work. The tough part is organizing the data, and generally, that accounts for more than 90% of the work. We can't create an ecosystem if we ignore that fact. It's for this point that the comprehensive strength of IT will be called into play, and NII has a diverse range of IT researchers working on that ecosystem.

Taki When do you think that will be ready for use in the actual clinical workplace?

Kitsuregawa It's hard to say because there are some circumstances that are rather touchy and complicated. What I mean is, radiation and ultrasound constitute the methodology for diagnosis, but at the same time, there are physicians working in hospitals who are specialists in certain areas of disease, such as digestive organ diseases or cancer. In many cases, the methodology doesn't directly connect between specialists and doctors in the various fields of disease. We've asked Mr. Suematsu to work with physicians who are specialists in various disease fields and those who are involved in measurement methodology, so that they get together with each other as a community.

Also, it probably differs depending on the hospital, but for example, to whom do x-ray images belong? Obviously they belong to the patient, but does the Radiological Diagnosis Department have the data or does the department handling the medical care for each particular organ have the data? That's something that's very unclear.

Suematsu It's an extremely fine point, but it's something that has to be dealt with. If you think in terms of the IT space in the hospital, I think that electronic charts are under the jurisdiction of the hospital director, or the chief information officer (CIO) to whom that responsibility has been delegated. Conversely, at institutions such as university hospitals, which use image data for which patients have given informed consent in research, each classroom has its own compartment, or "silo", and each laboratory has only the data that it needs for its own research. What AMED needs to do is function in the background to create a consensus through which image information is shared and utilized at core hospitals that play a central role.

Taki On that point, why did you choose the four societies that you selected?

Suematsu It was because they're all societies that can be linked cross-disciplinarily. It also makes sense in terms of the flow of tests that patients undergo; they have an ultrasound, and if something of concern is seen, then they have a CT, and after that a biopsy. We believed that it's important to share data in a form that incorporates all of these.

Also, the Japanese Ophthalmological Society is different from the other three in that they already have digital data about individual patients acquired through noninvasive means using multiple modalities. This field is promising in terms of digitalization. The other three societies started in fiscal 2015, and the Japanese Ophthalmological Society was added in 2017, so they were the last organization to join, but expectations are high.

Taki Are you going to be adding more in the future?

Suematsu The Japanese Dermatological Association and the Japan Society of Ultrasonics in Medicine will be joining us starting this fiscal year. It would be nice if our budget increased even more, though.

Hospitals need more professional data managers

Taki Are you thinking at all about putting the power of the private sector to work, for example in the area of funding?

Suematsu I think it's important to build a framework that would allow us to bring private corporations into the mix, not for the money they could provide, but because this is something that they themselves want to do. I'd like to create a big wave. We would start by building some success stories from data that we've already compiled, and then it would start growing here and there as people heard about what we were doing and wanted to get on board.

Kitsuregawa At the point when we're taking the first step, so that we all share data and are starting to produce benefits for patients, at the stage where we are verifying what we can do, I think that national research organizations acting in a neutral capacity could play a large role. What we want to do is, once we've produced a certain level of results through our efforts, to look at how we want to roll out the technology that we've developed to private industry in a way that will feel comfortable for the societies, hospital organizations, physicians, and patients. We'll need AMED to provide a venue where those issues can be discussed. It's not going to be an easy task, but Japan is the only country in the world that has data on this scale, and this is an issue that we are going to have to address ourselves.

Taki NII established its Research Center for Medical Bigdata in November 2017. That center will be creating the platform, won't it?

Kitsuregawa We have platform specialists and people who are specialists in medical image analysis at the Center. Unlike NII centers in the past, a succession of specialists from universities such as The University of Tokyo, Nagoya University, and Kyushu University have been joining us. As a rule, this is a nationwide endeavor. By becoming a member of the Center, they can look at the data that have been compiled. There's nowhere else where they could have the opportunity to look at such a vast compilation of images. We have one person who used to be in charge of medical devices at a corporation but who said that the Center was his idea of a dream, and he quit his company and joined us. There are also more and more medical fields that want to be added to the original "four brothers". I think that the number of core organizations aligned with us will continue to grow.

Taki My sense is that this not only will change the medical workplace but is also tying in to moves that redefine the ideal image of what Japanese society should be.

Suematsu AMED's task centers around medical care, but I believe that we need more people in hospitals who are able to use data skillfully and to be useful to society. The UK has built biomedical research centers in 30 domestic locations and has been cultivating human resources over the past decade. They have professional data managers who work in hospitals and who engage in what they call reverse TR (translational research), deciding what kind of basic research should be undertaken based on human data. In Japan, we invest money in research involving individual diseases, but we haven't invested much in human resource development. In parallel with our image project, I think that we have to initiate some innovative changes in the way that hospitals cultivate human resources.

Kitsuregawa When we do that, we need to make sure that the importance of data engineers is recognized along with that of data scientists. Also, the latter field, engineering, is going to be particularly important in terms of giving back the results of our work to society. That's because by finding inventive ways to compile, organize, and use data, we'll be able to bring out the real value of the data.

Incidentally, we're getting similar requests not just in the field of medicine but from an extremely large number of other fields as well. To cite just one example, in agriculture, crossing genetics with environmental factors has given rise to diverse phenotypes. We are moving in the direction now of reverse-analyzing this function. It doesn't take 30 years anymore to develop a new species of apple. We're shifting to a mindset in which all science and industry are based on data. Fields that are realizing that early on and making it a priority for investment are the ones that are getting steadily stronger on the global front. This is an issue that needs to be addressed by the nation as a whole.

(Photography by Yusuke Sato)

A Word from the Interviewer

The Editorial Department asked me to structure this as a three-way conversation, but there was no way I could keep up with the other two, with their wealth of insight and perception, and their excellent ability to express themselves. Ultimately, I took on the role of moderator. We didn't have a great deal of time, but the discussion was very thought-provoking, and there were some aspects of the topics that could not be included in the main article, because of space limitations, so I would like to mention them here.

The discussion had to do with professionals who are working with data. Mr. Kitsuregawa touched on the importance of people who are responsible for "annotation", in which they mark areas on images thought to be tumors by drawing a square around them. He said that the ecosystem won't function unless the contribution made by people doing annotation is recognized, and to that end, he said that he would like to note the names of people responsible for annotation in medical papers, although they may amount to several hundred people. Mr. Suematsu also said that starting from April, AMED's data management policy is always included in public recruitment projects, and the contribution made by researchers working with data is always emphasized. The discussion underscored the ways in which big data will change medical care and the world at large.

Medical care has a long history and has always called for diagnosis based on the five senses of the physician, treatment based on outstanding skills, and warmhearted communication. There will obviously be resistance to the shift from this analog world to digitalization. As referenced in today's discussion, however, rather than "substitution", the emphasis will be on seeking innovative change that will result in higher-quality services being provided.

Jun-ichi Taki

Senior Writer, NIKKEI INC.

Born in 1956. After graduating from the School of Political Science and Economics at Waseda University, joined Nikkei, Inc. After working in branch offices and covering corporate news, began covering science and technology, as well as environmental fields, starting from the mid-1980s. Authored "*Eco-Uma ni Nore!*" (Shogakukan) and co-authored "*Kansensho Retto*" (NIKKEI, Inc.), among others.

Aim Behind the Succession of Medical Research Centers Being Built

Providing support for medical care as a venue for system architecture and analysis, as well as for education

Tomoyuki Higuchi
Director General, The Institute of Statistical Mathematics

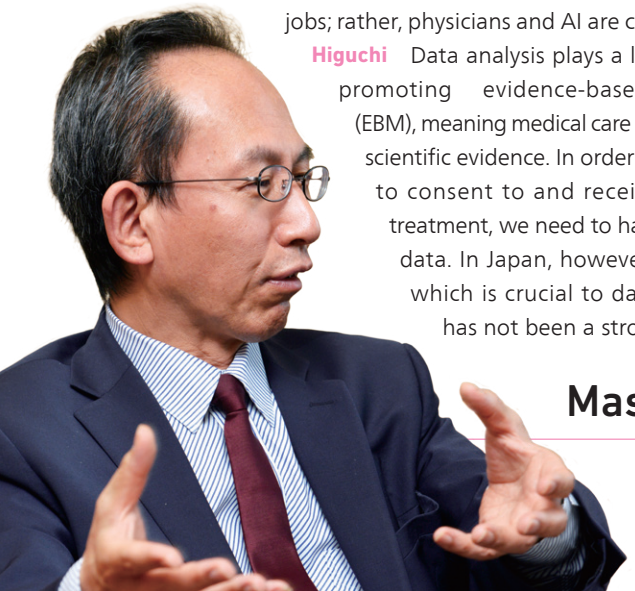
Masaru Kitsuregawa
Director General, National Institute of Informatics

Interviewer: Etsuko Nagayama Member of Opinion Group Editorial Committee, Mainichi Shimbun

A group of research centers involved in data science for medical care have been established in the laboratories of the Research Organization of Information and Systems. The National Institute of Informatics (NII) established its Research Center for Medical Bigdata in 2017, followed by the Research Center for Medical and Health Data Science that was opened at the Institute of Statistical Mathematics this April. We spoke with the heads of both organizations about the significance of their move into the field of medical care, and the direction that they plan to take in the future in terms of ties and connections.

Nagayama In medical research and the practical application of research in medical care, the use of big data and the appropriate analysis of the compiled data have become essential. I wonder whether expectations concerning data science in the medical field have also been growing.

Kitsuregawa In order for medical journal articles to be published, they have to be firmly grounded in acceptable data. For that reason, researchers in the field of medical care, where the link with IT has always been somewhat tenuous, are beginning to work proactively in data science. At the Research Center for Medical Bigdata opened by NII with strong support from the Japan Agency for Medical Research and Development (AMED), societies in medical fields are compiling large volumes of medical images from hospitals and other facilities, and are transferring them to a cloud built by NII, developing a system that will support diagnosis. Working with endoscopic images, we're already obtaining a high level of diagnostic accuracy with respect to a limited number of diseases. We've brought together experts in networks, cloud, security, and other IT fields who are building the system architecture. In April, I gave an overview of the present situation at the General Meeting of the Japanese Society of Gastroenterology that was very well received. Above all, I emphasized that AI will not be taking away physicians' jobs; rather, physicians and AI are coexisting.



Masaru Kitsuregawa

Higuchi Data analysis plays a large role in promoting evidence-based medicine (EBM), meaning medical care grounded in scientific evidence. In order for patients to consent to and receive medical treatment, we need to have accurate data. In Japan, however, statistics, which is crucial to data analysis, has not been a strong focus of attention. There have also been some very unfortunate events that happened in clinical research and other areas. We're standing on the brink of the "Age of Data". In particular, vast amounts of data are being accumulated in the fields of medicine and health science, and we built the Research Center for Medical and Health Data Science with the aim of promoting progressive research and education in data science in these fields.

Nagayama Why do you suppose that statistics hasn't ranked highly in the fields of medicine and health science?

Higuchi I think that the researchers themselves understood the importance of statistics, but that individual labs were left to their own self-help efforts when it came to human resource development. Maybe because of that, statistics was regarded as simply one component of medical research. Starting from this fiscal year, we initiated a new AMED project in which courses are being taught at the graduate schools of The University of Tokyo and Kyoto University aimed at promoting specialists in biostatistics who will analyze data in clinical studies and clinical trials, and that project has finally grown into a full-fledged, large-scale effort. That emphasizes just how vulnerable the system has been in terms of human resource development.

With that aim in mind, we're launching educational courses at our center on a systematic basis. The courses are targeting graduate students in doctoral programs, post-doctoral researchers, and young physicians in the workplace, and they're going to include the whole spectrum from lecture courses that teach basic knowledge to on-the-job training, where students will learn to address specific issues as they actually carry out the work.

Nagayama It sounds like what's being done by NII and the Institute of Statistical Mathematics is really a huge boon for the medical world.

Kitsuregawa My position is that we're all working on it together. For example, putting together a support system for imaging

diagnosis requires being able to point out abnormal sites such as cancer on each and every image. That can only be done by medical practitioners. It takes a huge amount of work and effort on the part of those in the medical world, and we're not aiming to become a main player and take over the stage; rather, this is something we can only accomplish by all working together.

Higuchi Last fall, before we opened the Center, we created what we called a Health Science Research Network (we later changed the name to the Research Network for Medical and Health Data Science), and invited participation. As a result, almost before we knew it, 66 organizations such as universities and pharmaceutical companies nationwide had signed on. It really made us realize how much interest there is in what educational centers are doing, as well as in research and development projects involving artificial intelligence (AI), big data analysis, clinical research, and other areas. Through the activities being carried out by our new center, we hope to change the way that statistics is regarded in medicine and health science.

Nagayama How have you established links and connections between your centers in the course of promoting data science in the medical care field?

Kitsuregawa There was a time when data mining was the big thing. If you take the total amount of time required as 100, normally about 90% of the time was spent preparing data. The other 10% was used for data mining, or analysis. That hasn't changed all that much. The work of compiling and organizing data and building a system isn't glamorous, and is a lot of work, but data can't be analyzed without that preparation. It seems to me, however, that that 90% of the work involves engineering, rather than data science. In other words, data engineering is also important. The University of Tokyo had information science and information engineering. Often, it's hard to identify the delineation between science and engineering, but obviously, ties with the Institute of Statistical Mathematics started long before the term "data science" first came into use, and of course will continue from here on out as well.

Higuchi We can't do our work without data, so we really depend on NII. At the same time, there are many commitments in place worldwide in the fields of medical research and in clinical studies and clinical trials, and data have to be accurately analyzed in line with what they require. Without professionals in data analysis who do their work properly, those results can't be written up in journal articles, and we won't be able to get research results out into the world. I see it as our role to lay the foundation for doing that.

Nagayama Data science and the medical field really are working closely together, just like a pair of wheels, aren't they?

Higuchi Going forward, the academic world and industry are going to need to create a matrix-based strategy. That will mean driving methodology, which forms the foundation, together with application. Unless we do that, we won't be able to keep up with the speed at which things are currently changing. Similarly, human resource development will need to take a matrix-style approach. Students have an obligation to obtain the fundamental knowledge and skills at universities, after which they join corporations and research organizations and further strengthen their abilities while gaining experience in solving real problems. Our center was established based on that concept.

Kitsuregawa If NII is going to support one side or the other, it would probably be the comprehensive strength of IT. The primary focus of the Institute of Statistical Mathematics is mathematics. I served as the president of the Information Processing Society of Japan, which is the largest IT-related society in Japan, and it encompasses about 40 fields. AI is only one of those. Developing systems requires a broad range of IT, and NII covers many diverse IT technologies.

Nagayama What kinds of advantages do you think patients can look forward to with research now underway at these two centers?

Kitsuregawa I've been told that no other country is on a par with Japan when it comes to medical organizations meticulously obtaining and accumulating test data. If we have an environment in place that makes large amounts of high-quality data available, it should help to reduce the risk of medical problems being overlooked when images are used for diagnosis and will relieve the burden on physicians. In other words, physicians will no longer need to spend as much time on simple diagnoses and will be able to spend more time on diagnoses in which judgments are difficult. Overall, I believe that this will help boost the quality of medical care.

Higuchi The aim of the network that I talked about earlier is to cover the entire nation of Japan. First of all, the network will help to spread awareness of the importance of data science. Also, our modern society is constantly and dynamically changing, and new technologies and new types of data, among them big data, are emerging in rapid succession. By covering all of Japan, I believe that we can help to research what modern clinical trials need to be in order to respond to the expectations of patients, and we can also contribute to building new systems.

(Photography by Yusuke Sato)

A Word from the Interviewer

I remember that, while covering a certain study in epidemiological research, a specialist in statistics said in exasperation, "I really can't understand what led to results like these." There were also some questions about how the data being analyzed had been collected. It is a serious issue because the outcome would affect qualities and quantities in medical care and policy decisions. I really wonder how medical care in Japan will change as a result of the efforts of both of these research institutions, which are providing powerful backup in the form of data science. Interest is growing in terms of what developments we will see in the future.

Etsuko Nagayama Graduated from the Faculty of Law at Keio University in 1991 and joined the staff of Mainichi Shimbun. After covering the Department of Science and the Environment, the Department of Medical Welfare and other departments, became a member of the Opinion Group Editorial Committee in 2017 and continues to serve on the committee.

Tomoyuki Higuchi

In 1989, Doctorate of Science completed through the Department of Geophysics, Faculty of Science at the University of Tokyo and joined the Institute of Statistical Mathematics of the Ministry of Education, Culture, Sports, Science and Technology. Since 2011, has served as the Executive Director of the Research Organization of Information and Systems, and as the Director-General of the Institute of Statistical Mathematics. Field of specialization: Bayesian Modeling. Member of the mathematical science and informatics fields of the Science Council of Japan.

Taking on Health, Medical, and Social Challenges with Machine Learning

What emerges as correlations are sought and quantified

Youichirou Ninomiya
Researcher by Special Appointment Research Center for Medical Bigdata(NII)

The primary objectives of the Research Center for Medical Bigdata that opened on November 1, 2017 are developing AI for use in medical image analysis and developing a cloud infrastructure. In April 2018, Youichirou Ninomiya transferred from the Cognitive Innovation Center (CIC) to the Research Center for Medical Bigdata (RCMB) and is now working on the development of AI that will assist in medical diagnoses by examining medical image data to differentiate between pathological and normal conditions. We spoke with Dr. Ninomiya about the Center and his work.

The aim: Quantification of phenotypes

In the world of biology, the volume of information available has been growing explosively ever since the genome became a target of research, and using computers and other forms of machine learning to interpret data is becoming essential.

Youichirou Ninomiya started out research on jaw development. As a jawbone develops, growth factors are secreted and intercellular and intracellular signal transductions and other processes take place. If some signals are excess or below certain thresholds in the biological pathway, then the jawbone could end up with dysplasia or malformation. Specifically, phenotypic abnormalities, which are abnormalities in observable morphology, can occur. Dr. Ninomiya's research has involved, in particular, using machine learning to analyze correlations between phenotypes and genotypes, which are composed of the genes that make up a biological organism.

Youichirou Ninomiya

"For example, if we look at an x-ray of a jaw, we can see a number of parameters, such as the angle, length, and position of the bone, and the center point of the joint. There are various kinds of diseases that can occur in the jaw, and in order to understand what kind of condition is present at what location, we make measurements of characteristics in order to quantify the phenotype. At the same time, we do a genome analysis to identify the patient's genes. Then we look at the correlation between the genotype and phenotype, to know what kind of genetic mutation is causing the disease."

If physicians know what kind of change occurred, at what point in the development process, and to what extent, then it is easier for them to put together a treatment plan. Furthermore, this knowledge could also be connected to discovering new treatment methods.

The basic approach is quantification of phenotypes. Having come from both backgrounds in bioinformatics and experimental embryology, Dr. Ninomiya always felt that genotype quantification was advancing far more than was the quantification of phenotypes. He joined the research group at CIC of NII in 2016 focusing on the latter. The work he was doing then has many commonalities with his current work at RCMB, namely, the development of AI that will assist in medical diagnoses.

"What I will pursue is clarification of what machines see and how they see it. It's not simply a question of differentiating between something that's normal and something that isn't; I would like to understand what the machine focuses on and, if it's quantified, what that variable is. In other words, I want to be able to quantify phenotypes. I hope that I can delve deeper into this research at RCMB."

Because genes are composed with four different letters (base

pairs), genotypes are easy to quantify as digital data. Phenotypes, however, which are the outcomes of genetic make-ups, are all different, just like no two faces are alike. Phenotypic quantification is difficult in a universe that involves such vast diversity.

"However, we can identify individual differences in people's faces simply by glancing at the person, and at the same time we can recognize someone as the same person even when they are displaying different facial expressions, which is something difficult for machines to perform. I want to address what features people are looking at and to clarify what the differences are."

Predicting diseases before symptoms appear

During the three-year term at CIC specified in the agreement with IBM Japan, the researchers could access new cognitive computing technologies, including IBM's Watson and Bluemix, to address societal issues such as aging, and working and living environments. A significance of the project was that around 20 of Japan's private leading enterprises in different business areas worked together with CIC to carry out research using real enterprise data.

Research at CIC embraces five topics, among them support in health and child raising, and research pertaining to consumption behavior. For his part, Dr. Ninomiya was involved in research aimed at predicting diseases before symptoms appeared, using health-check records and personnel data. In the initial fiscal year of the project, researchers examined data from 20,000 people spanning a five-year period, provided by a participating enterprise, but unfortunately there were significant differences in the data that prevented the research from advancing as hoped.

During the second year of the project, the researchers focused on clarifying health disparities among workers and the mechanisms behind those disparities by using the data from a multi-purpose panel study called J-HOPE, with public health classes in the Kitasato University School of Medicine at the center of the effort. The J-HOPE study provided data from 10,000 cohorts (epidemiological observational research conducted on groups) gathered over a five-year period and identified parameters that enabled measurement of workplace environments and health-check records, classified by workplace.

In the J-HOPE project, researchers looked at the relationships between environmental problems in the workplace and health problems of individuals, and used machine learning to analyze correlations, focusing particularly on high blood pressure, diabetes, and depression, and examined which kinds of factors were involved. In the course of the study, they found evidence indicating that the way a person works can increase the risk of developing physical disease.

Based on those results, the researchers embarked on new research in 2018, obtaining temporal data once again from participating enterprises, and looked at connections between the way people work and their workplace environment, and worsening or improvement of disease. The study targeted companies that advocate health management, but the researchers found that in some workplaces, there was localized cases worsening of diseases.

The data that NII receives have been coded so that researchers do not know specifically which workplaces are seeing a downturn in workers' health. Dr. Ninomiya says, "I think that from this point on, companies will work with industrial physicians and look at workplaces rather than individuals."

For the study, researchers used conventional machine learning rather than deep learning. The reasoning was that if the study involved data from 10,000 or 20,000 people, better results could be obtained using existing, traditional machine learning. If it becomes possible to compile larger-scale data in the future in fields such as genetics and lifestyle habits, it may be necessary to use a different approach more suited to that scale.

One other research project that Dr. Ninomiya was engaged in during his time at CIC involved the way that skills are passed on in manufacturing companies. The aim of the project was to quantify tacit knowledge and implicit knowledge that skilled workers have, and identify key points in their knowledge. Specifically, the researchers videotaped operations such as visual inspection of given parts and used statistical processing including dimensionality reduction and clustering (grouping) to identify key processes in which skilled workers used a series of elemental steps to get the work done. The aim was to define the procedures that should be carried out during the task, based on the results of the study, and design a system that will allow machines to instruct not-skilled workers during the job training. This research is still ongoing at CIC.

Traditional machine learning has its advantages

One model case of technology designed to assist in imaging diagnosis being developed at RCMB targets stomach cancer. Working with the Japan Gastroenterological Endoscopy Society, the Japanese Society of Pathology, the Japan Radiological Society, and the Japanese Ophthalmological Society as partners, NII has collected over 100,000 medical images from all over Japan. For example, a single endoscopy examination generally produces around 40 images. A diagnosis then takes place and standardized text is used to create a report that is then used as training data, as well as annotated endoscopic images. With assistance of the AI, endoscopic diagnosis should be able to see dramatic improvement in terms of both efficacy and efficiency.

In recent years, deep learning has been used in much of the technology designed to assist in medical diagnosis using AI that has been developed by various organizations. Most of the image analysis conducted by RCMB has also been done using deep learning.

However, Dr. Ninomiya indicates that there are problems, saying, "Deep learning is highly capable when it comes to classification, but the process remains a black box, and we have no idea what the machine is looking at in making its judgments." Because of that, he hopes to use traditional machine learning to understand what machines are looking at, and how. In other words, quantification is the aim here as well, and the hope is that being able to quantify data will contribute to improved accuracy of imaging diagnosis. Dr. Ninomiya's ultimate goal remains the quantification of phenotypes.

(Interview/Report by Kazumichi Moriyama
Photography by Yusuke Sato)

NEWS 1

Research aims to address societal issues by utilizing LINE

Partnership agreement concluded among Hyogo Prefecture, Amagasaki City, Tamba City, LINE, and the Graduate School of Informatics at Kyoto University

On April 12, the National Institute of Informatics concluded a partnership agreement with Hyogo Prefecture, Amagasaki City, Tamba City, LINE Corporation, and the Graduate School of Informatics at Kyoto University in an effort to address societal issues by utilizing LINE, a communication app.

Prior to concluding the agreement, NII and LINE Corporation established a collaborative research department on April 1, 2018, for the purpose of conducting research on a robust knowledge base aimed at addressing societal issues and focusing primarily on the areas of robust intelligence and social technology. The Center for Robust Intelligence and Social Technology (CRIS) was established as a research base for the undertaking. CRIS is headed by Director General Masaru Kitsuregawa of the National Institute of Informatics, with Professor Sadao Kurohashi, Department of Intelligence Science and Technology, Graduate School of Informatics at Kyoto University and affiliate professor at NII, serving as the deputy director.

(From left) Kyoto University Professor Sadao Kurohashi, LINE Corporation CEO Takeshi Idezawa, NII Director General Masaru Kitsuregawa, Hyogo Prefecture Governor Toshizo Ido, Amagaseki City Mayor Kazumi Inamura, and Tamba City Mayor Shinichi Taniguchi at the ceremony celebrating the signing of the partnership agreement

In the research project, a system is being developed on the basis of information from city websites and other existing services in which artificial intelligence will be used to promptly respond to inquiries received from city residents via LINE regarding information about child raising, disaster prevention, city government, and other areas, in an interactive format. Additionally, by analyzing the kinds of inquiries that tend to be received and ascertaining potential needs, the project will be playing a useful role in the creation of new services that will contribute to improved conditions for residents and revitalization of the community.

NEWS 2

Contributing to addressing societal issues

Graduate University for Advanced Studies Department of Informatics holds commemorative medal ceremony for graduates, presents awards to outstanding students

On March 22, the National Institute of Informatics held a ceremony to award commemorative medals to graduates who had completed degrees with majors in the Department of Informatics, School of Multidisciplinary Sciences in the Graduate University for Advanced Studies (Sokendai) and also presented awards to the most outstanding students in the program (photo). NII is a participating organization in Sokendai and offers graduate education in the form of a five-year doctoral program in which students spend all five years at Sokendai, as well as a doctoral program for third-year transfer students. There were three graduates from the department who were awarded degrees in the spring term of 2017: FENG Jingyun, Susumu Kono, and TRUONG Thao Nguyen, all of whom had studied at NII. At the ceremony, each of the students' supervisors described their accomplishments, after which Mr. Masaru Kitsuregawa, Director General of the National Institute of Informatics, presented medals to each of the students commemorating their graduation. Mr. Kitsuregawa also presented commemorative shields to Takuma Ebisu and NGUYEN Phi Le, who had won outstanding student awards.

In offering his congratulations, Mr. Kitsuregawa spoke about the contributions of data science in the field of undiagnosed diseases, saying that it is a difficult challenge to address issues such as assuring privacy when data are collected, but it is even more difficult to address the various factors involved in those issues, and that this is the current state of affairs regarding societal issues. His closing message to the graduates was, "Of course you have to keep researching your own topic. However, I sincerely hope that you will use your 10-20% energy to think about how you can solve these real societal problems". With those words, he expressed his hopes that the graduates would make a contribution to addressing societal issues in the future.

SNS

"Hey, this is great!" Hottest articles on Facebook and Twitter (March 2018-May 2018)

National Institute of Informatics, NII (official) Facebook
www.facebook.com/jouhouken/

Detailed information about INFOMANIA, a 3D puzzle action game that lets players experience mathematics, can be found at the following URL:https://bigdata.nii.ac.jp/wp/pr/infomania/ (2018/4/16)

National Institute of Informatics, NII (official) Twitter
@jouhouken

Research Introduction 2018
"Matching Algorithms and Combinatorial Optimization" Yu Yokoi # NII YouTube
(2018/3/29)

Bit on Twitter! Twitter
@NII_Bit

Bit comes to the award ceremony for the Commendation for Science and Technology by Minister of Education, Culture, Sports, Science and Technology in the field of science and technology. Bit covers the Commendation for Science and Technology recipient Professor Yamaji. (2018/4/17)

* Some text edited/omitted.

NEWS 3

Professor Kazutsuna Yamaji awarded Science and Technology Prize

2018 Commendation for Science and Technology by Ministry of Education, Culture, Sports, Science and Technology in field of science and technology

On April 17, the Ministry of Education, Culture, Sports, Science and Technology held its annual commendation ceremony to award the 2018 Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology that the Ministry presents every year to a person who has achieved noteworthy outcomes in research and development, public awareness, and other achievements involving science and technology. The ceremony was held at the auditorium in the Ministry. Kazutsuna Yamaji from NII was awarded the Science and Technology Prize (Development Division) for his contributions to cloud service architecture for the Institutional Repository (IR) [1]. Professor Yamaji is a Digital Content and Media Sciences Research Division professor of NII, as well as the center director of the Research Center for Open Science and Data Platform. Photo: Professor Yamaji receiving the award.

Professor Yamaji developed a working model of an institutional repository system using CMS [Content Management System] that is provided to universities and research institutions as "JAIRO Cloud", a cloud-type service. This allows universities and research institutions to easily create an architecture for their own unique IR, and puts Japan at the forefront worldwide in the number of IR architectures that have been created. The award was presented to Professor Yamaji in recognition of his development of a new academic information communication infrastructure that will contribute to the further development of open science in Japan.

Professor Yamaji expressed his pleasure at receiving the award, saying, "I never even dreamed that I would be receiving this award for my development of a repository system. It makes me very happy that the importance of institutional repositories is being acknowledged. Starting from fiscal 2017, as Center Director of the Research Center for Open Science and Data Platform, I've been working with other members of the Center on the research and development of a new research data infrastructure that includes a repository system. Japan still has a long way to go in open science. Receiving this outstanding award will motivate me to keep working even harder in the future."

[1] Institutional Repository (IR): A group of services that universities and research institutions provide to members making up their community that allow digital documents created by the universities, research institutions, and their constituent members to be managed and transmitted.

NEWS 4

Holding a competition to discover graphs showing efficient network topology

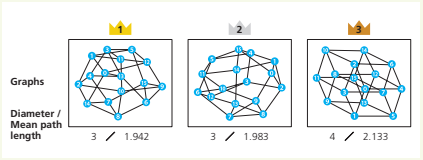
How would you connect CPUs in a supercomputer?

NII is holding a competition called "Graph Golf" to discover a way to produce a straightforward graph that could tie in to an efficient design for networks inside and between CPU chips, and would replace the complicated networks used in devices such as supercomputers with a simple graph [1].

Computers are embracing larger scales and have become more complex in recent years, with supercomputers having millions of processor cores all interconnected. The design of the network topology, which connects these vast numbers of cores to each other as efficiently as possible, hugely impacts the processing capability of the supercomputer. In NII's competition, contestants will create models of network topology represented as graphs in which the cores are represented as vertices and the cabling connecting one core to another is represented as edges. The maximum number of hops between any vertex in the graph and the vertex most distant from that vertex (that is, total number of the vertices that are passed through along the way and the terminal vertex) is called the "diameter", and the mean number of hops between pairs of vertices is called the "mean path length". The challenge is to come up with a graph that has the smallest diameter and mean path length under the specified conditions.

Applications are being accepted through a dedicated website (http://research.nii.ac.jp/graphgolf) until October 14. Contestants who come up with outstanding graphs will receive commendations at CANDAR2018, an international symposium involving computer systems and network technology being held in Takayama City, Gifu Prefecture in November.

[1] Graph: A model comprising an aggregate of vertices and edges, which show the connections between the vertices.



Example of graphs in which there are 16 vertices and 4 edges coming from each vertex. The graph at the far left, which has the shortest diameter and the shortest mean path length, is the best of the three.

Agenda

2018 public lectures to start with "At the Forefront of Informatics" in July

The schedule and programs have been decided for the 2018 series of public lectures for the general public titled "At the Forefront of Informatics", which are presented by researchers at the National Institute of Informatics and cover various topics involving informatics. For detailed information and to apply to attend, please refer to the NII public website below.
https://www.nii.ac.jp/event/shimin/
(Support is provided by Chiyoda Ward.)

1st lecture: Tuesday, July 10
Instructor: Hiromichi Hashizume (Professor, Architecture Science Research Division)
Topic: Indoor positioning and navigation technology
— Getting directions even inside buildings beyond the reach of GPS signals —

2nd lecture: Friday, August 24
Instructor: Takeaki Uno (Professor, Principles of Informatics Research Division)
Topic: Understanding and discovering data mining
— AI does not do everything for us —

3rd lecture: Thursday, September 13
Instructor: Ryoichi Ando (Assistant Professor, Digital Content and Media Sciences Research Division)
Topic: World of digital art drawn using fluid mechanics
— Computer graphics in scenes that bring luck: Does beauty come down to mathematics? —

4th lecture: Wednesday, October 24
Instructor: Megumi Kaneko (Associate Professor, Architecture Science Research Division)
Topic: Wireless access networks of the future
— Existing frequencies will run out if we do not take action! --

5th lecture: Tuesday, November 20
Instructor: Keizo Oyama (Professor, Digital Content and Media Sciences Research Division / Director, R&D Center for Dataset Sharing)
Topic: Shared use of real data
— Your information goes to academic research!? But don't worry —

6th lecture: Tuesday, December 11
Instructor: Yoichi Iwata (Assistant Professor, Principles of Informatics Research Division)
Topic: Logic and reality of calculation
— Calculations that should be hard are actually very (?) simple —

7th lecture: Wednesday, January 23, 2019
Instructor: Kazuhiko Hirakawa (Professor, Institute of Industrial Science, The University of Tokyo / Researcher, The Global Research Center for Quantum Information Science at the National Institute of Informatics, new academic field of Hybrid Quantum Science)
Topic: New developments in terahertz electromagnetic waves
— Far infrared rays do not just roast coffee beans —

Essay

Thinking of Innovation

Mitsuru Ishizuka
Project Professor, National Institute of Informatics /
Director, Cognitive Innovation Center

As the belief has been growing in recent years that we need innovation as a motive force to drive society, the word "innovation" is being incorporated more and more often in department names, not only in the national government but also at universities and corporations. That increase has been noticeable recently and expectations regarding innovation are growing amidst an era of change. This directly affects me on a personal level, as in February 2016, I became the director of the Cognitive Innovation Center (CIC) established with funding provided by IBM Japan.

In Japan, AI research centers have been established at the National Institute of Advanced Industrial Science and Technology and at RIKEN. Because it is not a good strategy for us to compete in terms of the number of researchers, our center aims to provide Japanese companies with innovation involving cognitive functions and AI, with cooperation from IBM Japan, by addressing real issues that directly affect Japanese companies. That was one reason that we used the word "innovation" in the name of our center, but when we chose the name, we did not actually put that much deep thought into it.

In research work, novelty is an important factor, along with its usefulness, and becomes one key criterion for evaluating academic papers. Novelty involves invention, but novelty alone is still far from innovation. Innovation comes into play in bringing revolutionary changes to society, people's lives, and industry, and has a strong social impact. In a happy-case scenario, invention leads to innovation, but the reality is that such cases are far from commonplace. In many cases, innovation is created when real-world issues, including near-future issues, are confronted and ways to address them are found.

The difficulty here is that technical novelty is not necessarily a requirement for innovation, and in many cases, the challenge of innovation involves work that calls for modalities different from those of research. Because I was involved in research at a university for a long time, I was not used to work that was oriented toward innovation, and have experienced struggling with this issue. It is very difficult for one individual to take on work with two different directional characteristics, and the individual needs to be careful that the work is not done in a half-baked manner and to make sure that it carries sufficient impact. Of course, an orientation toward innovation that strives for technical novelty has value as well (because it is sometimes also used by others for innovation), and I am certainly not criticizing it. I am simply pointing out the need to recognize that the two have different orientations. In cases where it is difficult for one person to aim for both, a teamwork-based strategy is a good approach.

When we talk about innovation today, the book "*The Innovator's Dilemma*" by Clayton M. Christensen comes to mind. If we trace back along this concept of innovation, however, we arrive at the thinking of the economist Joseph Schumpeter. Schumpeter saw innovation as a fundamental function, as well as the most important function, driving capitalist economies, and called innovation a "new combination", a view which is suggestive, even today. Keynes, a great economist of the 20th century, was also born in 1883, the same year as Schumpeter, and the two were rivals. While Keynesian economics, which was the mainstream of economics for many years, addressed static conditions, Schumpeter's approach can be defined as dynamic economics. Given the recent era of change in which we find ourselves, the approach to innovation advocated by Schumpeter, as well as his concept, may well be worth rethinking.

Future Schedule

July 21 | 2nd 2018 Karuizawa Saturday Forum: "150 Years since the Start of the Meiji Era and 30 Years since the Start of the Heisei Era: Looking Back at Modern Japan in the Context of Era Names". (Speaker: Takashi Mikuriya, Professor Emeritus, The University of Tokyo; Professor Emeritus, Tokyo Metropolitan University)

September 8 | 3rd 2018 Karuizawa Saturday Forum: "Animal Communication and the Origin of Language". (Speaker: Kazuo Okanoya, Professor, University of Tokyo Graduate School of Arts and Sciences)

For detailed information, please refer to the public website of NII at the following URL:https://www.nii.ac.jp/event/karuizawa/

Notes on cover illustration The cover shows a robot acting as a patient and being surprised at the checkup results being conveyed by the doctor. The display in the background shows a photograph of an x-ray of the affected area.

Weaving Information into Knowledge
NII

National Institute of Informatics News [NII Today]
No. 66 June 2018
[This English language edition NII Today corresponds to No. 80 of the Japanese edition.]
Published by National Institute of Informatics, Research Organization of Information and Systems
Address | National Center of Sciences 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430
Publisher | Masaru Kitsuregawa **Editorial Supervisor** | Ichiro Satoh
Cover illustration | Toshiya Shirotani **Copy Editor** | Madoka Tainaka
Production | MATZDA OFFICE CO., LTD., Sci-Tech Communications Inc.
Contact | Publicity Team, Planning Division, General Affairs Department
TEL | +81-3-4212-2028 **FAX** | +81-3-4212-2150 **E-mail** | kouhou@nii.ac.jp

Bit (NII Character)

https://www.nii.ac.jp/en/about/publications/today/