JAPANESE CIVILIZATION INSTITUTE

日本文明研究所

24-5 Sakuragaoka-cho, Shibuya-ku, Tokyo 150-0031
Tel: 03-5456-8082 Fax: 03-5456-8388
Mail: info@japancivilization.org
http://www.japancivilization.org/

Newsletter Vol. 12 Summer 2018

On 23rd May 2018, a special lecture under the theme, "Will Artificial Intelligence (AI) surpass human intelligence?" took place for the 12th Symposium of the Japanese Civilization Institute. Mr. Yutaka Matsuo, associate professor at the University of Tokyo gave the lecture, while Naoki Inose, director of the Japanese Civilization Institute moderated the lecture. The lecture featured what we should keep in mind in a world that is drastically changing with the fast development of deep learning, and the current situation of research in artificial intelligence in Japan and in what way the Japanese industry can survive provided the circumstances. (Reprinted from "Shukan Dokushojin," 10th August, 2018, No. 3251 & Web Dokushojin)

Greetings

Masashi Takai
Board of Director of Japanese Civilization Institute, Chairman and President of Kinokuniya Company Ltd.

I have served as Chairman of this Japanese Civilization Institute ever since it was founded. The Chairman should be here today giving the speech, but I will take her place. Today, we have Mr. Matsuo, leading AI researcher, giving a lecture to us all, which I am looking forward to very much. Mr. Matsuo has written numerous books, ranging from academic books to manga, which are all selling extremely well at our bookstore. Our theme today is "Will Artificial Intelligence (AI) Surpass Human Intelligence?" I personally am looking forward to discovering whether business management will definitely be better if artificial intelligence took place of CEOs. I think this is a good opportunity for us all, so if you have any questions, don't hesitate to ask. I hope you enjoy the lecture to the very end. Thank you.

"Spirituality" required in the era of Artificial Intelligence

Toshio Goto
President of Japanese Civilization Institute

When we think about artificial intelligence, which we are going to learn about today, and relate it to Japanese Civilization, what comes to my mind is Daisetz Suzuki who spread zen to the world in English. Artificial intelligence is said to surpass people's intellectual, intelligence and whatever part our left-brain dominates. It's nearing the realm of creativity too, including knowledge, logic and reason, and research is pursued every day in order to go beyond that and reach the area of sensitivity. A couple years ago, a book written by artificial intelligence was chosen as a candidate for the "Nikkei Hoshi Shinichi Award" (sponsored by the Nihon Keizai Shimbun) and brought controversy.

Other words, there are possibilities for artificial intelligence to reach the standards of humans or even surpass humans in literature. Similarly, many music created by robots can be heard on the Internet today. Artificial intelligence may surpass both our left-brain, which dominates logic, and our right-brain, which dominates creativity. And in the end, will there be an area in which artificial intelligence cannot cover? While thinking about this, I came across the word "spirituality," which Daisetz Suzuki mentioned.

So, what exactly is spirituality? Daisetz Suzuki calls it something that lies in the middle of our spirit that cannot be completely embraced, and something that is related deeply to our religious consciousness. He did not refer to a particular religion, but instead he cited it to be an absolute religious existence. Concerning this, he stated;

"When the spirit confronts material substances, and thereby suffers in shaking off bonds, if an opportunity to touch their own spirituality is provided, the suffering of conflict will naturally fade away."

He said, spirituality will not come into existence unless humans go through suffering like any other humans do. Other words, babies will not know spirituality at all. In addition, Japan as a nation recognized spirituality for the first time, during the Kamakura Era (1185-1333) and amongst it was zen.

People think of Daisetz Suzuki as a researcher in zen, but when you look at his life carefully, you can see that he spent his whole life researching spirituality, particularly "Japanese spirituality." While artificial intelligence is nearing the realm of logic and creativity today, I think we should focus more on understanding humans. The more artificial intelligence becomes the heart of our lives, the more important civilization becomes. In addition, the soul, which exists in the realm of what Daisetz Suzuki calls spirituality—that lies somewhere deeper than knowledge or creativity—is put into question. I think if we listen to Professor Matsuo's lecture while keeping this notion of "Japanese spirituality" in mind, it might give us a clue as to how we should face the future to come.

Special Lecture

"Will Artificial Intelligence (AI) surpass human intelligence?"

Yutaka Matsuo vs. **Naoki Inose** (MC)

Big innovation in deep learning

Matsuo: 2 years ago, computer AI AlphaGo, developed by Google DeepMind, beat Le Se-dol, a South Korean Go master with 9 dan. The following year it won over Chinese Go master Ke Joe, who also has 9 dan, so there is practically no one who can beat the machine. AlphaGo learns the go players' playing record and continues to compete itself while becoming stronger. Then finally, in December last year, AlphaZero was developed and this didn't even need go data for learning. It became stronger by competing itself in areas not only in go, but also in chess and shogi. Meanwhile, elmo beat Ponanza at the World Computer Shogi Championships, but AlphaZero won over elmo in merely 2 hours, learning from scratch. These artificial intelligence devices only need 2 hours to accomplish what these masters build all their lives. The technology at the core of these artificial intelligence machines is called deep learning.

Today, artificial intelligence has become extremely popular, but in fact, artificial intelligence is in its third boom. Research on artificial intelligence started in 1956, so 62 years have passed since it was first realized. Since then, people have been reciting the fantasy that "someday, artificial intelligence will surpass human intelligence." However, technology never caught up and that fantasy turned into disappointment and dark years kept coming again and again. Amid the boom we have today, I think we need to know exactly what artificial intelligence can do and cannot do.

By the way, did you know that research on automotive driving goes back well over 10 years ago? Roughly speaking, the technology we have today is all the result of small improvements of what we accomplished years ago. Moreover, most of the things that we couldn't do back then, we still can't do today. Still, amongst all these technologies, the technology of deep learning has seen a great innovation.

So what exactly can be done with deep learning? It comes in 3 steps. The first thing it can do is to "recognize." Next, it can "learn movements," and lastly it can "understand language."

To begin with, let me go into detail about "recognition." People can distinguish a cat, a dog, and a wolf instantly when shown a photograph, right? But this is difficult for a computer to do.

Therefore, we provide it with a definition for classification like, "if its eyes are round, it's a cat." Next, we give it a criterion for judgement like "if its eyes are long and thin and its ears are loopy it's a dog, but if it's got pricked ears, it's a wolf." Seemingly, one would think that they can acknowledge under these rules, but there are in fact multiple kinds of dogs in the world. For instance, Siberian Husky have pricked ears but they are not wolves. It's in fact extremely difficult to define a dog and a wolf.

This notion of "looks like〜" is known as "feature amount." So long as humans are defining these feature

Yutaka Matsuo

amounts, the accuracy of image recognition will not increase. Humans are unconsciously learning this notion of "looks like〜." Computers too, need a system where they can learn "feature amount."

This became possible with deep learning. The "Google Cat" is famous. The system they developed goes like this: you simply show the computer a huge number of pictures that you find on YouTube and it will learn what a cat looks like.

The year 2012 was an epoch-making year for image recognition. All over the world, computers tried recognizing images using the same set of data and they competed in the percentage of accurate answers they gained. Surprisingly, it turned out that the winning team of 2012, suddenly produced an error ratio of 16%, which was 10 % lower than any previous record. This was possible due to the development of a system in which deep learning automatically learned "feature amounts" that were initially defined by humans. After this, the error ration of image recognition starts to drop fast. In 2013 it was 11 %, in 2014 it was 6.7 % and now it's 2.3 %. If a human goes through the same task, the error ration is said to be 5.1 %. This means that in the field of image recognition, computers have already surpassed humans. This kind of revolution has happened in merely 5 years.

Image recognition is done through various tasks, including detection of objects in the image and segmentation in which the domain of the object is cut out. For example, an object detection program called YOLOv2 can read individual objects in an image as fast as 47 flames in 1 seconds. This program is shared around the world, and it can easily be found on the Internet.

From recognition to learning movement and understanding the meaning of words

Furthermore, if you combine this image recognition with robot technology, "learning movement" becomes possible. When humans continue doing the same movement, they reinforce the movement they are good at, which makes them better at it even more. This is called "reinforcement." People are trying to develop a system so that computers can do the same thing. For example, various tasks such as combining bumps and dents of blocks, opening, and closing of lids, placing hangers onto poles of different heights are acknowledged by the eye and adjusted. Then a dozen computers photograph the movement sequences with a camera, share that experience and self-learn at high speed. Robot and machine technology up until now couldn't make adjustments to hold things of different shapes and sizes, because the accuracy of image recognition was bad and it couldn't see.

Humans, on the other hand, learn to hold things at different spots in no time, after repeating the movement of holding and dropping things ever since they are a baby. Robots are now learning to do the same.

The fact that computers are finding it difficult to do things that children can do is known as Moravec's paradox. Research on computers making a medical diagnosis, playing chess and proving mathematical theorem has shown improvement to some extent, but things like image recognition that 3-year-olds can do are lagging behind. This is because humans can't explain how they're doing it themselves.

However, this situation changed with the third artificial intelligence boom. What's so brilliant about artificial intelligence of recent years is not how advanced it has become, but how it has come to be able to do easy things that 3-years-olds can do. We need to think about how the world will change with this. First comes recognition, then movement and next comes "words." This process is similar to what a child goes through; first, it becomes good at moving its body, then it starts touching various things and dropping them and then it starts to understand the concept. It is after this that they start to understand the meaning of language.

By the fall of 2014, it was predicted that artificial intelligence technology will develop through the following 6 steps; 1) increase in accuracy on image recognition 2) prediction of movement 3) learning of movement that isn't affected by environ-

mental changes 4)technology of movement while integrating smoothly with the surrounding environment 5) language understanding 6)consuming a wide range of knowledge. It will probably develop through these 6 steps and complete by roughly 2030. When artificial intelligence achieves recognition of steps 1) and 2), diagnosis of medical images, crime prevention, and surveillance will become possible. When learning of movement of stages 3) and 4) is achieved, automotive driving, watching over distribution and construction, automotive farming, nursing, cooking, and cleaning will become possible. Lastly, when AI reaches stages 5) and 6) and starts understanding the meaning of language, maybe translation and assistance in overall white-collar jobs will become possible. It's actually proceeding like this. But what we couldn't predict was that it's proceeding 3 to 5 times faster than expected.

At the current stage, computers are starting to understand the meaning of language. One is called image captioning, a technology to create a sentence after looking at an image. For example, when it takes in a picture of a man playing the guitar, it will display the sentence, "man in the black shirt is playing guitar." It is starting to learn to do vice versa too. When you put the sentence, "A very large commercial plane flying in blue skies," you get an image of a large plane flying against a blue sky.

Here, the computer is not searching for an image. It is instead creating an image based on the sentence it receives. We unconsciously create images in our head when we hear something. Artificial intelligence is doing exactly the same. From 2 years ago, Google translation has switched to deep learning and has increased its accuracy. For example, translation from English to Spanish and French to English stands pretty comparable to what humans translate. But translation from English to Chinese or Japanese is difficult for even humans to do, due to its difference in language structure, so there is still a difference in the quality of translation done by humans and computers. But even so, it's nearing that of humans.

Having said that, no matter how much computer increases its accuracy, it will never take place of the subtitle translator Natsuko Toda. Humans translate while taking into consideration people's feelings and its delicate changes as well as the cultural background.

Similarly, in the future, it will probably be difficult for automatic machines to create novels and haiku that moves our emotions. Novels and haiku are living expressions that express our emotions and instincts. Perhaps we can compare the relation-

ship of humans and artificial intelligence to that of birds and flying. Humans can understand the principle engineering of flying and create planes, but planes do not chirp in the morning. Similarly, artificial intelligence does not have vivid emotions and conflicts. Artificial intelligence is the result of humans trying to solve human intelligence with engineering.

Evolution of artificial intelligence, with the "discovering of vision"

There is a book by Andrew Parker titled "In the Blink of an Eye" that explains the Cambrian Explosion. The Cambrian Period is said to be a period in which all existing species came to present during the merely 10 million years in the long 4.6 billion years history of the earth. There are many theories as to why the Cambrian Period happened, but Parker stresses that it was because vision evolved, calling it the "Light Switch Theory." Until then, living things didn't have eyes, so they moved slowly using their sense of smell and touch, and ate when they bumped into something and ran away when something bumped into them. But then a trilobite with acute eyes appears for the first time. It obviously had advantages so it increased its number enormously. Soon, those that ran away started to grow eyes and began swimming fast or acquire mimicking abilities so it wouldn't get caught. By having eyes, strategies for surviving became diverse and with it, species became diverse too.

I think the same thing will happen in the world of robots. By discovering vision, tasks that can be accomplished and varieties in machines will increase enormously. Camera's image sensors are equivalent to the creature's retina. We see by processing signals, which we received from the retina, at the brain's visual area. This visual area is equivalent to deep learning. Our civilization up until now has been a history of how we should use machines that lack eyes. Products produced at factories needed to be the same in shape and material. Factory's pipeline too is systematized by machines with no eyes. Today, the area that isn't automated are the areas that machines with no eyes find it hard to handle. The prominent example of this would be farming, construction and various food processing industries.

Naoki Inose

Humans are the ones controlling tractors and combine harvesters. Robots can be used for say rice and potato harvest, which you can harvest in chunks, but for harvesting vegetables and fruits like tangerines, apples, tomatoes and cucumbers, the human eye is essential. In fact, the human eye is necessary for almost every process of farming. That's why there is a shortage of hands in farming. Similarly, construction sites need the human eye too for labor like setting up reinforcing steel, welding and casting concrete. Food processing too has to have eyes or they wouldn't be able to cut the ingredients, place them on the frying pan and turn them over when they are cooked.

But with the technology of vision, there are possibilities that farming, construction sites, and food industries to become automatic. This will stir a huge impact. If it were farming, tomato harvest would be relatively suitable for becoming automotive with reasonable cost. If it were construction sites, welding is menial labor that lacks a working hand, so I think it might be feasible. Cooking requires extremely complex routines so it might be a difficult area for automation, but I think we can start by putting the dishes in the dishwasher. Like these examples, we can start automating things where we can, and gradually develop it from there.

In fact, Japan is strong when it comes to machines for farming, construction and food processing, and boasts a huge share in the world. I think Japanese companies can excel in developing automotive machines by combining hardware and deep learning. Japan has been economically suffering these 20 years, being unable to start a big business in the Internet field. However, I think it can do well against the world, in developing automatic machines with visual capabilities. During the first industrial revolution, the steam engine was invented, replacing what used to be done by human muscles or farm animals. Similarly, if eyes are added to machines, "recognition" will be taken away from humans and

be relocated where society needs it.

In reality, the medical field is one step ahead in diagnosis using X-ray, CT, MRI, dermatology, and endoscope. Automotive machines that surpass overall doctors, have already been developed. Just last month, The Food and Drug Administration (FDA) of the U.S. government authorized fundus testing equipment that uses deep learning. Recently, artificial intelligence not only supports doctor's jobs, but also are close to replacing doctors altogether.

The field that's progressing next is facial recognition. We all know that iPhone X has incorporated it, as well as China's Tencent Inc. for supervising their employees' entry and exit. There are over 100 million cameras across China, so deep learning is spreading fast in that country. There will no longer be the need for people to monitor rooms for surveillance. Moreover, it will only take 10 seconds to check which camera the person you want to find showed his or herself last. China wants to watch over its people so there is a great match here. There is a great concern as to how this kind of trend will evolve in the future. Also, the vision of artificial intelligence can be used for looking over patients 24 hours straight, at nursing homes and hospitals.

Moreover, presently, there are cleaning robots but there are no tidying robots. This is because tidying is an act of recognizing an object and placing it back to where it used to be. But with the technology we have today, there is a good chance that this tidying robot will become a product in the near future. If tidying robots become common, you can leave your home in the morning and come back at night with everything back in place, like a hotel. Our living environment will no doubt improve enormously.

Deep learning equals to the least-squares method

So let me explain to you in detail about the structure of deep learning. In one word, deep leaning is "least-squares method that uses the deep function." Least-squares method is a method used for statistics. For example, if you draw a scatter diagram with Excel, and add an approximation straight line, you have a line. The algorithm used to draw this line is a least-squares method.

For instance, say you drew a scatter diagram of a store, stating the relationship between the temperature and how many drinks you sell on that day. You draw a line by adding

an approximation straight line and you get the equation, <y'=a+by>. This y' expresses estimated value and the error is calculated by measuring how much gap there is compared to the actual value and how much each point is apart from this line. There are errors going in the plus direction and minus direction, so you square the error and add it up. All points of the square sum are more or less out of line, so when you decide on the value of a and b so it becomes minimum, you are able to find a line that rides on top of the points.

We just created a formula using the amount raised on beverages against one variable stating temperature, but let's make it two variables of temperature and humidity. Then, it becomes three dimensional and the line becomes a surface. But even so, the process is the same: you find out how much y' is apart from the actual value with the equation <y′=a+bX1 +cX2> that uses two variables, X1 and X2. You then square the error and add it up and figure out the value of a, b and c so that the square sum becomes minimum. I just converted one variable to two variables, but the method is the same even if it's 10,000 variables.

One wouldn't regularly come across a case that requires 10,000 variables but it's needed for people who want to find out if a cat is in an image of 100×100. Other words, in deep learning, you need to find a function that returns 1 for a cat and 0 if it wasn't a cat against 10,000 number of variables. Again, what you do is the same: you create the equation <y′=K0+K1X1+K2X2+ .....+K10000X10000> using 10,000 variables from X1 to X10000, take the square sum of the error compared to the actual value and decide on the value from K0 to K10000 that makes it minimum. Image recognition of a cat is acquired like this with minimum square sum.

Deep leaning is close to multiple regression analysis and multivariate analysis of economics and marketing. However, about 30 to 40 variables are used for multiple regression analysis and multivariate analysis at the most, but deep learning requires tens of thousands of variables to hundreds of millions of variables. You need tons of data and in order to estimate the parameters, you need a huge amount of power from calculators. Other words, deep learning is like a monster with the minimum square sum. Moreover, I have somewhat simplified the explanation up until now, so there are parts where it's a little different from reality. Equations used for deep learning is a deep equation, not a linear equation. Rather than defying functions directly, you defy a different, intermediate function of f1 or f2.....f100 out of variables X1 or X10000, and use this function to create a cat function. The equation would be <f cat(X) =K0+ K1X1+K2X2+ .....K10000X10000> and the more you increase the number of functions, the more 2 layers, 3 layers, 4 layers, 5 layers—layers will increase. But what you do, again is the same: you will be able to find the parameter that deeply defies the cat function through minimum square sum. No matter how much you defy intermediate function to a linear equation, it remains linear, but if you add a non-linear element to this, it tries to imitate the function of neuron—a nerve cell that ignites when stimulation exceeds to a certain degree. It creates a sigmoid function (σ) in which 1 is selected when the value of X is a plus and 0 is selected when the value of X is a minus. By placing (σ) before the head of each function, you get <f1 (X) = σ (K01 + K11X1 + K21X2 + ... + K100000, 1X100000)>. Non-linearity of each intermediate function becomes strong and the overall cat function turns into a non-linear function. But at the bottom of this parameter too, we're again using minimum square sum.The reason we deepen function is that the world we live in is hierarchical, so we need functions that are hierarchical too. Also, when we think about space that is structured by different kinds of images, the area occupied by a "cat" is a subspace, but that space is twisted and bent, forming a complex shape. In order to cut off such a shape, we need to build a function that is high in the power of expression. Normally, to increase the power of expression, we need a huge number of parameters, but in this case, we need a lot of data. When we think of a way to keep the number of parameters low, while heightening the power of expression, the most effective way is to build layers of simple functions.

Building functions with deep layers and predicting data out of parameters, is something researchers have been trying to do from decades ago, but with little success. The reason they couldn't do it was because they lacked power and data. Today, that problem has been answered and minimum square sums with massive parameters—like a monster, can be solved.

From attractive Sigmoid function to realistic ReLU

Today, people have started shifting from a Sigmoid function to an activation function called ReLU (Rectified Linear Unit). The equation of ReLU is expressed as <max (0, x)> and it changes to 0 when in the area of a minus and to X when in the area of a plus. The reason why ReLU is better than a Sigmoid function is that with Sigmoid function, the more input value X becomes bigger, the more output value Y nears 1, and sticks to it and stops moving. It tries to differentiate and search for an optimal value but most differentiation end up to be 0, so it can't get close to an optimal solution. However, the value of ReLU is X, so a constant

term always remains when it is differentiated, so it ends up in an optimal solution.

They started using ReLU around 2012. Ever since researchers started studying artificial intelligence neural networks decades ago, they were using the mostly Sigmoid function, because its equation < σ(X) = 1/(1 + e)> was attractive. Compared to this, the equation of ReLU; <max (0, x) > was not. At the time, deep learning was unrealistic, so if they can't make anything out of it, they thought they might as well use the one that was theoretically beautiful.

Lastly, let me just refer to you another method called Batch Normalization. When you optimize deep function with a neural network, there is a method called error back propagation, in which you return errors due to how much you made mistakes. Up until now, the ratio of change was converged to 0 or released endlessly, and thereby deep function was said to be hard to optimize. But with Batch Normalization, they can adjust each value's output to an average of 0 and dispersal of 1. Other words, they decided to create a deviation rate out of the value output of each neuron. It's very simple, but this made it work. Differentiation was no longer converged to 0 or released endlessly, and each neuron released some kind of deviation rate—a fixed value. As a result, the quality of optimization improved immensely.

This was suggested in 2015. They were doing research on this for decades, but nobody attempted to normalize it. People involved in statistics go through these kinds of procedures quite naturally. Everyone thought it wasn't possible, so it lagged behind for decades, but once that barrier was taken away, it evolved at a high speed, resulting in the rise of accuracy we have today.

Is deep learning, a technology suitable for the Japanese people?

In short, deep learning is a "least-shares method" in which you can predict the value of Y if you enter the value of X. By using 10,000 to 100 million variables, and creating a deep, non-linear equation by piling up simple intermediate value, it imitates the process of neuron and makes deep learning possible.

At any rate, it's least-shares method, so it's all about thinking what to put in X and Y. If you put an image in X and a class, a cat or a dog in Y, image recognition becomes possible. Similarly, if you set it up so that it outputs an image in X and a class and bounding box (a rectangular grid-line that surrounds the image) in Y, you get an algorithm of object detection and if you put English sentences in X and Japanese sentences in Y, you are provided with a translation. If you think of ways to connect that with neural network, it can basically learn anything. The basic principle is simple.

I think the fact that the basic principle of deep learning is simple and that it requires numerous small tuning, makes designing deep learning a suitable job for Japanese people to tackle. The more they revise small circuits, the more tuning of parameters is done in detail and the more data is collected, the more accuracy is increased. Other words, the more effort you put in, the more results you get. Moreover, you can gain huge merit in the way you combine the design of deep learning and hardware. That's also something the Japanese people are good at doing.

However, Japan is way behind when it comes to technology in deep learning. When you look at 1500 academic essays related to deep learning and see how many times they have been quoted, the most quoted essay has been quoted 17000 times. It's a legendary essay of 2012 about image recognition written by professor Geoffrey Hinton and others.

Moreover, if you rank these researchers on the number of times they have been quoted, the top few are occupied by leaders of the deep learning field such as Hinton, Yoshua Bengio, and Yan LeCun. These are all exceptional people who continued to research during the dark years and succeeded in breaking the barriers. But apart from these people, these deep learning researchers are all in the late twenties or early thirties. Also, most of these essays were written after 2012. When you look at academic essays of other areas and line them up in order of how many times they have been quoted, you realize there are very few young people in the higher ranks. This just shows you how much deep learning is new in the field, something completely different from conventional technology. Unfortunately, there are no Japanese people in these ranks. The highest would be the 400th place. Regrettably, in Japan, countries, and companies invest in artificial intelligence, which comes before deep learning. And this competition is won by how much you can gain good personnel. Excellent personnel who are brilliant at deep learning are mostly young people. People ranking in the top 300 of the list I mentioned earlier, receive a starting salary of over 50 million yen a year. One more 0 is added to those who rank in the top 30. There are people who receive a couple billion yen a year, similar to that of a top soccer athlete. There are over 100 brilliant brains such as these in Deep Mind. When it comes to Google and Facebook, there are a couple hundred. Similarly, companies like China's Alibaba and Tencent are gathering talent fast. Meanwhile,

Japan is not joining the competition at all.

To begin with, the problem with Japan is that it's old. It can't keep up with the fast pace of deep learning, which keeps evolving. Of course, there are young Japanese people who are enthusiastically tackling deep learning, just like when the Internet appeared. But the older generation, who have the authority to manage personnel and operate the company, needs to understand its value, or else they will keep investing in fields of the past.

The second problem is that it's slow. There are big chances for manufacturing industries, but most Japanese corporations are bad at decision making. They cannot make decisions so they spread their targets out to everything sparingly. I call this phenomenon; "passive sparing in every area." Even though there are departments of big corporations that deal in artificial intelligence, they are just there. They don't operate at all.

The third problem, which is the most critical, is that they don't invest in people. Even though the government creates a budget for this area, what they choose to do in the end is buy a supercomputer with it. This is because, with the current Japanese system of personnel, they can't afford to pay tens of millions of yen to a young employee.

From 2015, I have a research group at University of Tokyo where I teach, that specializes in deep learning. 70 people gathered in the first year, but the number is growing. 190 students joined in the second year, while 360 students joined in the third. The year 2015 is the same year that top universities of the world started teaching deep learning, so I have confidence that we're no way behind. If you acquire the skills, chances will emerge and several people can launch startup companies. 2 companies have been listed from Matsuo Research Group so far.

Building structure of the industry and cultivating human resources

What we're dealing in is extremely innovative technology, so I figured we need to build an industry structure. That's why I launched the Japan Deep Learning Association in June last year. One of its aims is to increase engineers who can handle deep learning. We can only produce a few from Matsuo Research Group, so I wanted to provide an intensive learning opportunity to people from other universities too, together with people who are in the business of various fields, people who are working and students.

I also prepared a certificate exam because artificial intelligence is a field in which it's hard to tell a fake from real. One is an engineering certificate (E Certificate), given out to those who have a certain level of skills for deep learning. The other is G Certificate, which examines how much the person understands the basics of deep learning, and how that person can use it in business.

One of the reasons why the Japanese IT industry failed to keep up in world competition, is because deep learning literacy of the person who orders and the users, including the country, the local government, medical and financial institutions, were very low. Therefore, the price is still paid according to workload. If this were a restaurant, it's the same as saying; "This dish is 10,000 yen, so it's delicious." Normally, customers try and find delicious restaurants that are also cheap, and this ultimately increases competition and raises the quality of the industry.

If we want to change the industry with deep learning, we need to increase the level of those who order. This is why I created G Certificate. I hope it will be used by presidents of companies who are thinking of business using artificial intelligence, people who are starting a new business, people in human resources and people who are engaged in marketing. It's a certificate, which you can acquire in 1 to 2 weeks of studying. If you gain it, you will understand what you should value when you talk to IT vendors. The first G Certificate took place in December last year, and half of approximately 1500 people who took the exam passed. We intend to hold it 3 times a year.

I think 'monozukuri' (making things) using deep learning is a big chance for Japan. There are high possibilities that the technology of the "eye" will solve the issue of lack of labor brought upon by the decreasing birth rate and aging population. But if we continue at this speed, the U.S. and China will win the competition, and there will be no place for Japan. If we want to catch up from here, we need to seriously think about how we're going to invest in deep learning and how we're going to do everything with more speed.

Creating a Comprehensive Deep Learning Research Institution in Japan

Inose: So, you're saying heightening the technology of deep learning is suitable for the Japanese people. But the Japanese system of promotion by seniority, lifetime employment results in people in their 50s and 60s making decisions. Almost every engineer in the world who are leading the field of deep learning are in their 20s and 30s, but Japanese companies can't pay young people tens of millions of yen for their start-

ing salary, because if they do so, their employment system will collapse. The only way they can survive is to create their own company, or maybe companies can form a separate company.

Matsuo: Yes, there are many companies that are forming separate companies. But I think it's best if they change the existing human resource system altogether, if they can. Companies are all in need of people who can handle deep learning. But their salaries are kept low. There are no changes there. If they don't change the basis of how they think, nothing will evolve. I'm sure that once they set an appropriate salary, people will come flooding in.

Inose: You're saying it's an area that is suitable for the Japanese people, but the system in which it is developed, and the old way Japanese companies are managed is resulting in Japan to get left behind from Silicon Valley and China's Shenzhen. Although, I think Matsuo Research Group is standing up like a Silicon Valley on its own (laugh). In that kind of situation, it's meaningful that G Certificate was established and acknowledged by society. By the way, is the Ministry of Economy, Trade, and Industry of any help?

Matsuo: They do offer support... The truth is, I think qualification systems such as these should be established by the government. I wish they could quickly start expanding education too. There are no chances for Japan to win in the academic research world of deep learning, but that doesn't matter. Why? Because Japanese automotive and electronic companies never really invented the engine or the transistor. Instead, they added originality to what was invented overseas and opened a new market. We should do the same with deep learning too.

Inose: How many researchers are there in Matsuo Research Group?

Matsuo: Officially there are less than 30 people. But there are 60 to 70 people involved in the project. We need to spread our knowledge not only inside our group, but also to other universities and people living in local areas. So, we have taken multiple approaches like opening online courses. We hope we can have people from technical college tackle deep learning too.

Inose: The faster you acquire the skills the better, right?

Matsuo: Yes, that's right. There are graduates of a technical college in Matsuo Research Group too, but

they move their hands than complain. In addition, there are many good hardware companies in local areas that are in need of people with deep learning skills. I think we could raise personnel in local technical colleges and have them contribute to these local companies.

Also, times are changing drastically, so I recommend that people in their 40s and 50s start studying fast. I'm quite sure they'll profit if they do.

Inose: I wrote the book "Defeated in War in the Summer of 1941," but Japan seems to be not at all different from those times, which is sad.

Matsuo: I think we need a to create something like a Comprehensive Deep Learning Research Institution, where we can objectively analyze the situation of artificial intelligence in Japan and examine where we're losing and where we have a chance to win.

Inose: Yes, we surely need a Comprehensive Deep Learning Research Institution in order to excel. The nation as a whole, should strategically tackle this issue.

(Reprinted from "*Shukan Dokushojin*," 10th August 2018 Issue, No. 3251 & WEB Dokushojin)

The panelists

Yutaka Matsuo

Associate professor of the University of Tokyo School of Engineering, Institute of Engineering Innovation, Center for Knowledge Structuring and Department of Technology Management for Innovation. Matsuo is a leading researcher in artificial intelligence in Japan.
Born in Kagawa Prefecture in 1975. He graduated University of Tokyo, School of Engineering in 1997. In 2002, he received a Ph.D. degree from the University of Tokyo in 2002 and is a doctor of engineering. From 2002, he served as a researcher at the National Institute of Advanced Industrial Science and Technology and from 2005 he was a visiting scholar at Stanford University. He has been serving as an associate professor of University of Tokyo since 2007. He received the Japanese Society for Artificial Intelligence Best Paper Award in 2002, and the Information Processing Society of Japan Nagao Special Researcher Award in 2007. His books include "Is Artificial Intelligence Exceeding man? Deep learning and beyond," and "Why Artificial Intelligence will change the future."

Naoki Inose

Author. Born in 1946. In 1986, he received the Souichi Ooya Nonfiction Award for his book *"Mikado no Shozo"* (Portrait of the Emperor)." In 1996, he received the Bungeishunju Readers' Award for his book "*Nipponkoku no Kenkyu*" (A Report on Japan)." In June 2002, Prime Minister Junichiro Koizumi appointed him to the Promotion Committee for the Privatization of the Four Highway-Related Public Corporations. He served as Tokyo Governor from December 2012 to December 2013. In December 2015, he became Special Advisor to Osaka City. His books include *"Show 16-nen no Haisen"* (The Defeat in Showa 16), "*Persona—Mishima Yukio Den*" (Persona: The Story of Mishima Yukio) and *"Picaresque—Dazai Osamu Den."* (Picaresque: The Story of Dazai Osamu). Upcoming books include "*Kyushutsu*" (Rescue), *"Senso, Tenno, Kokka"* (War, Emperor, State), *"Seigi ni tsuite Kangaeyo"* (Let's Think about Justice), *"Minkei"* (Posse Man), *"Tokyo no Teki"* (Tokyo's Enemy) and in collaboration with Lully Miura, *"Kokumin Kokka no Riarizumu"* (Realism and the Nation-state).

Translation: Ayako Karino

Japanese Civilization Institute 2018 13th Symposium

As human beings, the last stronghold against artificial intelligence would probably be our 5 senses. However, in the modern world, we have come to lose our 5 senses. Our life used to be filled with scent and sound.
Kodo (traditional incense-smelling ceremony) which Shino Soshin first established in the Muromachi Era (1336-1573) in Kyoto has passed down its tradition, without breaking off once. On this occasion, we will have Hachiya Souhitsu, young *Soshu* who carries on the tradition of Shino School of Kodo, lecture us on *Kodo* with actual demonstrations.
They say, incense is something that you listen to. How did the *Gion Shoja* bells of "The Tale of the Heiko" sound like? Crystal ball player Ms. Asami Ishizuka will give a performance, while we immerse ourselves in ancient sounds and monko (listening to the incense).

▶ **Panel discussion:**

Making full use of the five senses

—What is traditional Japanese scent and sound?

Hachiya Ittshiken Souhitsu
21st successor and
head of Shino School of Kodo

Asami Ishizuka, Crystalist Asami
Representative director of the Crystal Bowl Academy Japan Foundation and
CEO of Crystal Bowl Academy Japan.

Naoki Inose
Author, director of the Japanese Civilization Institute

Date: 28th August (Tuesday), 2018, 7 p.m.-9 p.m. (doors scheduled to open at 6:30 p.m.)
Venue: Japan University of Economics, Tokyo Shibuya Campus Hall (the hall seats approximately 100 people)
Address: 25-17, Sakuragaoka-cho, Shibuya-ku, Tokyo, 151-0031

Admission: 2,000 yen (please pay at the door on the day)
How to attend: apply through the site below:
http://www.japancivilization.org/
Contact: Japanese Civilization Institute
03-5456-8082

JAPANESE CIVILIZATION INSTITUTE

Traditional Crafts Online Site

Worldwide：https://shops.japancivilization.org
Japanese：http://japancivil.shop9.makeshop.jp

Messages from Director

Japanese Civilization Institute has been introducing and selling traditional crafts of Japan. Although it is said that there is approximately 1,200 types of traditional crafts in Japan, its production is declining. As part of our activities, we hope to discover valuable traditional crafts spread across the nation, enjoy Japanese craftsmanship and its beauty inherited over the centuries, and develop it with you.


Tsumami Kanzashi (Crepe)


Folding Screen (half size)


Folding Screen (full size)


KYO-YAKI Earthenware


Japanese Bamboo Basket Tokyotrad


Odoshi - Samurai Armor

Contact：Japanese Civilization Institute 24-5 Sakuragaoka-cho, Shibuya-ku, Tokyo 150-0031 Phone : 03-5456-8082 Mail : info@japancivilization.org