Interview: Artificial Intelligence

Professor Sakura Osamu (left) and Professor Nishigaki Toru

Professor Sakura Osamu (left) had an interview with Professor Nishigaki Toru about the artificial intelligence (AI) on September 2, 2016, at the Office of the Dean of the University of Tokyo Interfaculty Initiative in Information Studies.

Professor Sakura (hereinafter "Sakura"): First, I would like you to give a brief self-introduction. I have just read your book entitled *Big Data and Artificial Intelligence: Gain Insights into Their Possibilities and Traps*, which you published in July 2016 through Chuokoron-Shinsha Inc. In this book, you discuss the cultural and social background behind artificial intelligence (AI), with relation to recent big data and singularity. I found it very interesting. I am remembering the excellent impression I got from reading your book for general readers entitled *AI: The Concepts Behind Artificial Intelligence*, which you first published in 1988 through Kodansha Ltd. *Big Data and Artificial Intelligence* came out this year, 28 years later. I think that much of your way of thinking has remained unchanged at its root. We just have the concept of big data today, 30 years later. Please tell me a little bit about what has changed and what has not changed in that context, along with an outline of your career.

Professor Nishigaki (hereinafter "Nishigaki"): I published that book in 1988, around the end of the second AI boom. We have had three AI booms so far, with the first boom flourishing from the 1950s to the 1960s. There were pretty essential implications as a result of the keyword during that boom being "logic." AI is considered to be merely a field of computer science, but my opinion is that, fundamentally speaking, the invention of computers was intended to create AI from the beginning. Of course, initial applications included scientific/technological calculations and business data processing. But clearly, the pioneers of AI like Alan Turing and John von Neumann aimed to achieve thinking machines. Behind this was *logicism*. That is, logicism was based on the idea that you can obtain the truth if you put together a group of logical propositions which accurately describe the world according to a set of stylized formal rules.

The Dartmouth Artificial Intelligence Conference in 1956 is said to have raised public awareness of AI, but people had dreamed of achieving thinking machines long before that. Let me tell you the reason why the first AI boom ended in failure. Among other things, note that very few problems could be solved by logic alone, except for some puzzles and easy games. Difficult games, including chess and *shogi* (the Japanese version of chess), have so many permutations which could not be analyzed then. All they were able to handle was the so-called *toy-problems*. Actual problems existing in the world were more complicated. Nevertheless the researchers were even undertaking machine translation projects. The problems caused by words with multiple meanings and grammatical exceptions were an extremely intractable challenge.

Now, let me get back to the Japanese story. Few people had computers in the 1950s, and Japan did not see the arrival of an AI boom. It wasn't until the 1980s that an AI boom came to Japan—during the period of the second boom. At that time, I was a visiting scholar at Stanford University. In those days, Stanford had already been a center of AI studies along with Carnegie Mellon University and MIT (Massachusetts Institute of Technology). Edward Albert Feigenbaum, who had come up with the idea of the expert system, was the director of the Computer Science Department at Stanford, and he was hugely popular. I was also influenced by him. (Laughs)

Sakura: How long did you study at Stanford?

Nishigaki: From 1980 to 1981.

Sakura: That was in the middle of the period of the second AI boom, wasn't it?

Nishigaki: Sure it was. Professor Osuga Setsuo at the University of Tokyo, from whom I received guidance for my doctoral thesis, was also an AI researcher, which triggered my interest in AI during the second boom.

Sakura: Looking back at your career history, you joined Hitachi, Ltd. after graduating from the Department of Mathematical Engineering and Information Physics in the School of Engineering of the University of Tokyo. Did you go to Stanford University during your Hitachi days?

Nishigaki: Yes. Precisely speaking, I specialized in operating systems while I worked for Hitachi, and I was involved in creating mathematical models of a computer system. This area was not so closely related to AI. At Stanford, however, I was free enough to attend AI classes and seminars. At the time, we had an amazing array of big-name AI researchers at Stanford besides Professor Feigenbaum. I thought about AI a lot in that environment, and I wrote the above-mentioned book back in Japan. When I wrote the book, I had already left Hitachi and was teaching the basics of computers to liberal arts students at a Japanese university. In this way, I shifted my specialty to social and cultural aspects of a computer society, and the core of these was AI. This is because the essence of AI is based on the assumption that computers have the ability to think like human beings. In this context, the fundamental question was asked, that is, whether it is possible for computers to do so or not.

The keyword of the second AI boom was "knowledge." A typical example of this was the expert system. This system was based on the idea that all you have to do is, to logically express the knowledge that experts have, and to put it together properly using a computer.

In the 1980s, a national project for developing the Fifth-generation computer was launched in Japan as a joint academic, industrial and government project. In fact, I participated in the project as a Hitachi engineer for a brief period of time shortly after returning to Japan from the United States. It was the largest project in Japan's computer history. And it was a famous project in that it was not able to be commercialized despite the injection of more than 50 billion yen. (Laughs)

Sakura: Ordinary people as well as you gave the project poor evaluations, didn't they? I would like to ask you a little bit more about the core concept of "knowledge" behind the second boom. The first AI boom was centered on logic, whereas the second AI boom was focused on knowledge. That is, the simple practice of logical reasoning only enabled you to solve problems with toys. But if you want to solve a broader range of everyday tasks, you have to make full use of expert knowledge. People tried to achieve this by computer.

Nishigaki: That was exactly what the Fifth-generation computer was intended to do. More specifically, you express knowledge in the form of logical propositions, and then you put it together through automatic logical reasoning using a computer. The Fifth-generation computer was intended to work as a high-speed parallel logical reasoning machine. It seemed a convenient tool, but it was hardly put to use. The second AI boom itself died down at the end of the 1980s. This was because human knowledge is not as logically strict as had been expected and unavoidably included ambiguous factors. For example, medical knowledge is something statistical to a great extent; a particular symptom merely suggests the strong probability of a particular disease. Accordingly, if you only put together logical propositions in a stylized formal way without examining the details, you may reach a very strange conclusion. This was where the fundamental challenge lay.

Sakura: People failed to grasp the nature of human knowledge, did they?

Nishigaki: Human knowledge involves ambiguous factors. Experts who are human beings do not simply carry out logical calculations according to local knowledge, but make a comprehensive judgment by making use of intuition. That is where the essential differences lie. For example, a famous expert system MYCIN was able to carry out an automatic diagnosis of bacterial blood diseases, but it could not be put to commercial use. This was probably because of a question as to who should be deemed to be responsible for a misdiagnosis.

You should think about what the third AI boom will be like with a focus on such history. The key phrase of the third boom is "statistics and learning." That is, it is based on the assumption that even if you make a mistake, it is no fatal problem; all you have to do is to try to be approximately correct and to continue to learn. Let me take machine translation as an example. Words with multiple meanings and grammatical exceptions make it very difficult to constantly create correct translations in a given context. But you have corpuses. You can retrieve and search a large amount of corpuses and pick up the translated statement with highest probability. The English word *dog* can mean not only a dog as a mammal, but also something insignificant and worthless. But if most corpuses say that the word is used to express a mammal, you just pick it up.

Now, remember that the keyword of the first boom was "logic." The greatest advantage of AI was that computers *never* make logical mistakes. Alan Turing and John von Neumann considered the computer to be a machine that could achieve correct human thinking. Today, however, we are shifting to the idea that slight mistakes are acceptable. You should not ignore this. Naturally the utilization of statistics leads to a significant expansion of the scope of AI applications, though.

Sakura: You mean that a priority should be how to use AI in real daily situations, not logical consistency, don't you?

Nishigaki: That's right. The pattern recognition is a typical example. This field already existed when I was a student, but it was technically very difficult. The computer is poor at classifying patterns and understanding the meanings of images. This was said to be where the gaps with human beings lay. But it is the technology termed "deep learning," which is now attracting a great deal of attention, having brought about breakthroughs in this area through mechanical learning based on statistical processing. You could say the technology triggered the third AI boom.

Precisely because the keyword of the third boom is "statistics," you need a massive amount of data. I included the words *big data* in the title of my recent book. It was precisely the collection of a huge amount of data from the Internet that caused the success of deep learning. The basic idea of deep learning itself had existed for many years, but it was not achieved simply because there were not enough data, and they were also extremely hard to process. The advantage of deep learning as a pattern recognition technology is its automatic extraction of the characteristics of the patterns to be recognized from data. In conventional pattern recognitions, human beings must input the characteristics of patterns from outside, but for deep learning, the computer discovers the characteristics of patterns on its own without such processing. Deep learning is groundbreaking in this respect, but you need to process an enormous amount of data for actual execution. The third AI boom has been achieved by huge data processing ability.

Let me discuss the situation in Japan here. Japan is attempting to catch up with other countries in the third AI boom. Western countries are taking the lead in deep learning as well. But this is not the case in the area of robots. Japan has enjoyed the world's leading position in research on industrial robots. Germany is a major advocate of the Industry 4.0 project, through which the European country is planning to develop smart factories using AI. People usually mass-produce standard products in factories, but the German project is intended to facilitate the AI-based automatic production of a wide variety of products in small quantities in factories. As a result, it is described as industrial smartization. Because wages are higher in developed countries, including Germany and Japan, they are put at a disadvantage in terms of price competition against developing countries. However, they can restore the manufacturing industry by introducing smart factories. I speculate that Japan, which has high-level technologies for industrial robots, will succeed in that area. Namely, Japan is expected to make good use of AI/robot technologies for the efficient production of painstakingly manufactured goods.

Sakura: Fundamentally speaking, one of the major reasons for the Japanese automobile industry's success was its manufacturing of many slightly different types of assembly lines in accordance with highly intricate customer needs. This is exemplified by Toyota's "just-in-time" production system. Japanese people are dexterous and are good at perfecting details. The Japanese are characterized by this, I think. I am convinced that a good combination between this ethnic characteristic and AI will make the Japanese even more dexterous than the Germans. (Laughs)

Nishigaki: Exactly. But the Germans, who place a great deal of emphasis on logic, may readily accept smart factories, while Toyota engineers may raise objections, saying "It's not that easy" and "Human beings are trying to do something greater than AI." This is where my expectations lie. I think there is a radical difference between machines and human beings. Accordingly, a new wave of Industry 4.0 is expected to develop through the collaboration between human beings and AI at the Nissan, Toyota and Honda factories, etc.

Sakura: You discussed Industrie 4.0, but in Japan, we are now hearing the start of a discussion about Society 5.0 for the future. Society 5.0 is about human communication, nursing care and welfare. But your past arguments and writings suggests that you think this project will be rather difficult.

Nishigaki: Society 5.0 will include discussions about AI-based cultural and social activities. This reminds me of the concept of the ubiquitous society that was in fashion a few years ago. Ubiquitous technologies were successful to some extent on the supply side, such as factories and hospitals, but they were unable to make inroads into consumers' lives. This was quite a natural and logical conclusion. There is little demand for taking a hot bath and enjoying warm heat immediately after coming home. Ordinary people use inexpensive sensors and cannot do maintenance work at home. So their ubiquitous systems do not work long. The bath may be boiling when you get home. (Laughs) That is, engineers had an overoptimistic outlook for what to achieve for consumers. On the other hand, however, factory and medical workers who can do proper maintenance work for themselves can increase efficiency and enjoy the benefits by using ubiquitous technology skillfully. It is necessary to learn lessons from this.

Sakura: Instead of merely expanding the scope of areas for the application of AI, you should think harder about where to apply AI. Right?

Nishigaki: That's right. I am opposed to the idea of implementing an all-round policy based on the principle of pleasing everyone under the banner of "society." I think it is a good policy to focus on industries. There are gaps in academic arguments about "society" between researchers of humanities and those of technologies.

Sakura: Certainly, because factory and hospital workers have limited requirements, they can control the situation more easily and with fewer contexts than in scenarios of society in general. That enabled the expert system to be successful to some extent. But for daily lives in society in general, it has a wide range of contexts, and as you mentioned, people can just use economical items. In addition, it is fairly easy to boil water and turn on a heater. The automation of such devices does not provide you with a significant benefit. (Laughs) Machines are difficult to use for handling various daily scenarios, even with work that can easily be handled by human beings. That is, the more daily work there is, the more difficult it is to handle it by machine. I cannot help feeling that engineers just have a poor imagination in terms of this aspect.

Nishigaki: I have a feeling that if you are too technology-minded, you tend to have poor insight into human beings and society.

Sakura: This means that it is important to nurture specialists who can develop such insight. Now, let me move on to another point of view. What struck me about your 1988 book entitled *AI: The Concepts Behind Artificial Intelligence* was that behind AI or computer science were very profound ways of thinking based on Judaism and Christianity—monotheism. You argue very convincingly that such ways of thinking are closely associated with the stylized logical patterns you already mentioned—this can be clearly observed in Gottfried Wilhelm Leibniz, I think. When I read the book, I was a graduate school student. Until that time, I had thought that science, if not technology, was so universal that I vaguely assumed that it would barely be influenced socially and culturally. But your book showed me the very fresh recognition that behind AI, or rather computer science, lay deep-rooted historical and cultural backgrounds. Of course, the current third AI boom also shares the same traditions. But as you pointed out, computer engineers sometimes are too indifferent and insensitive to such aspects. Please tell me a little bit more about these historical and cultural elements behind the third AI boom and singularity theories.

Nishigaki: As for singularity hypothesis, futurist Ray Kurzweil expressed his forecast that AI exceeding human intelligence would have emerged by 2045. In 2005, when he published his book *The Singularity is Near*, his forecast did not attract much attention. But now that the third AI boom is occurring because of deep learning, his theory is in the limelight. His forecast is based on the idea of the Law of Accelerating Returns (LOAR). This is a kind of generalization of Moore's law, according to which the total number of transistors included in a single chip of an integrated circuit doubles every 18 months. While computer performance grows exponentially, the number of brain neurons is limited, and their reaction speed is slower than that of computers. This suggests that it is quite natural that the human brain will be surpassed by the progress of computers. The book is a funny one in that it also describes future and current technologies in a mixed and confusing way. (Laughs)

I have at least two questions about the singularity hypothesis. The first question is whether or not it is appropriate to evaluate human intelligent ability using single-dimensional standards. Advocates of this hypothesis start discussions from a comparison between the brain and the computer, but this is a misleading point. The brain is important, but it is a part of the body. The human being is a multicellular organism, where groups of cells in the body interact with the ecological environment in a complex way. Therefore it seems to be inadequate to merely focus on the brain.

The second question about the singularity hypothesis is that it does not at all consider the differences between living creatures and machines. Machines can process *past* data at such a rapid pace that they can achieve great efficiency under certain fixed environmental conditions. On the other hand, however, creatures live by adjusting themselves to the *present* situation in real time, although they have habitual behaviors based on memory. Creatures' activities constantly emerge anew and flexibly. This is why creatures manage to survive in a rapidly changing environment. Certainly, human beings have biological features to live efficiently depending on logical models, but this is not all the aspects of human beings. Basically, human beings live for the present moment, flexibly in real time. In contrast, machines process past data strictly based on formal rules and cannot cope with totally unknown new situations. Machines may take completely erroneous and pointless actions because they are simply subject to past data. Since creatures and machines differ in terms of qualitative ability like this, I am against the idea that AI's ability will "exceed" that of human beings.

To sum up, my argument is based on two points of view. First, I am skeptical about the brain model ignoring other parts of the body. The brain only started to emerge relatively recently in the long history of evolution. In particular, there are few creatures with developed brain parts for logical thinking. For about three billion years in the four-billion-year evolutional history of creatures, there were many unicellular organisms whose main activities were metabolic activities. Even human beings, which are multicellular creatures, are basically made up of such cells. Next, machines simply process past data, but the cells of creatures live flexibly for the present. My critical views contend that the singularity hypothesis is too wild and irrational, because it ignores these two points.

Sakura: I think that various efforts have been made to devise learning and introduce physicality for machines as well. Fundamentally speaking, deep learning studies triggered the current AI boom, originating from neural network studies in the 1980s. In addition, in the area of robot studies, it has long been thought that it is essential to consider physicality for the movement of robots in real spaces. In fact, robots play soccer in the Robot World Cup (RoboCup) games. Currently, this may still be at a simple, plain level, but I speculate that they believe that if the learning ability of machines continues to grow at its present pace, someday they will be able to go beyond the limits you just mentioned.

Nishigaki: I guess they think so.

Sakura: Is it absolutely impossible?

Nishigaki: The reason why it is impossible is that the machine is basically an *open* and *heteronomous* system, and is devoid of genuine autonomy. The logic within AI is so complicated that it looks difficult to keep track of the processing details of deep learning, but if you try to, you can meticulously keep track of all those processes. This is because machines definitely depend on rules (programs) provided by human beings from outside. Learning machines often utilize different rules, but these differences themselves are subject to meta-rules provided by human beings in advance. This is not true of creatures, however. As "Autopoiesis Theory" indicates, the living creature is a *closed* and *autonomous* system, whose activities can never be foreseen completely. That is, creatures make processing rules on their own within themselves, which can hardly be analyzed from outside. At most, you can just make a guess. For example, suppose you give a signal to both a creature and a machine. In such a case, the creature has options (freedom) for interpreting the meaning of the signal because it is an autonomous system. On the other hand, however, the machine that receives the signal simply takes a particular predetermined action. This applies to all types of machines, including learning machines. Even if some learning AI may take actions that are hard to understand, it is wrong to conclude that the machine has the same characteristics as the creature.

Sakura: The reason why that difference is "underestimated," if I may borrow your word, is that people think, albeit unconsciously, that the Creator produced human beings.

Nishigaki: That's right. There are too many scientists and engineers who think according to so-called *naive realism*, the secularized version of the Creator theory. Even great scientists, such as James Watson and Francis Crick, think of creatures from a totally mechanical perspective—someone built a logical world and supervises/observes the world. In accordance with this line of reasoning, they claim, "Creatures have rules of their own, just like machines, within themselves. You may say that the rules within creatures cannot be known, but we scientists work hard to unravel the mystery by examining the brain, for example." They raise the criticism that the opposition to this point of view is old-fashioned mysticism. But their opinion is also a modern myth, because they look at everything from the unique (God like) viewpoint.

Fundamental Informatics, in which I specialize, offers a totally different view. The essential difference between Fundamental Informatics and other information theories is whether or not the observer's viewpoint is relative. An attempt to describe an objective world from a privileged fixed perspective of observation often results in naive realism. This approach to the world is too one-sided a view, in my opinion. The world that actually exists consists of, a world I see through my lens, a world you see through your lens, and more broadly, diverse subjective worlds that individual creatures build up. But these are not self-righteous. As they communicate with one another, they create a kind of commonality. Note that this way of thinking is not my original one. Roughly speaking, it was originated by Immanuel Kant, who argued that what human beings see is not the things themselves but just phenomena. This approach was followed by Edmund Husserl, and Martin Heidegger. The emergence of the world from human daily lives was discussed by Heidegger. Like this, despite the tradition of the Creator theory, excellent Western philosophers have given profound thought to the perspectives of the world. Accordingly, a relative viewpoint of the world has become quite common in the disciplines of humanities. This causes a great gap between humanities approaches and scientific approaches in terms of information. I think this is where the most serious problem lies.

Discussions about Society 5.0 involve this issue. Many researchers in the humanities think in the way of Heidegger philosophy. On the other hand, however,-the thoughts of computer researchers are mostly based on the Claude Shannon's Information Theory. Shannon's theory leaves no room for considering the meaning interpretation of information. This question was already raised in the 1980s by Terry Winograd, a famous AI researcher at Stanford University. I quoted Winograd's argument in my book *AI: The Concepts Behind Artificial Intelligence*, which I published in 1988. The same question has been raised repeatedly, but is left unsolved.

Sakura: An internal theory focusing on a definitely classified scope of application as a small phenomenon is easy to understand, isn't it? If you plainly declare, "This will enable a telephone to be clearly heard," it is easy to understand. And it is actually effective and beneficial. The independence and autonomy of computers on the Internet is based on the separation between mechanical and human systems, which is also easy to understand. On the other hand, however, if you say, "That's only part of it. The whole of it is even more complicated," it will be rather hard to understand in terms of ease of understanding and ease of application to actual issues, even if what you say is correct. Your story about phenomenology reminds me of Hubert Dreyfus, an American phenomenologist who published *What Computers Can't Do: The Limits of Artificial Intelligence* in 1972 (revised in 1979 and 1992) and introduced Martin Heidegger to the English-speaking world. The fact that he criticizes the current state of AI studies is a reflection of your argument. It is difficult to artificially produce human intelligence through materialistic and monistic approaches alone. In fact, however, because materialistic and monistic approaches are easy to understand and work well for some areas, they are convincing to some extent. I think it is necessary to stop there for a while and fundamentally reconsider singularity theories. Meanwhile, nowadays the environment surrounding humanities studies is becoming tough all over the world, including Japan. There is even essential skepticism from outside that humanities studies are unnecessary. Judging from what you have just argued, I consider that the present situation is a crisis.

Nishigaki: Exactly. I have also paid attention to the argument of Dreyfus. Speaking of more specific topics, as I already indicated, Japan has strength in the area of robots. Today, the mass media often use the phrase *AI robots*, which suggests the integration of AI into robots. In fact, however, these two are significantly different in terms of technological development history. AI has been a consistent subject of research in Western countries. I assume that a monotheistic way of thinking leads to the conviction that super AI will appear someday. On the other hand, the superiority of Japanese robots mostly consists in fine mechanisms. Japan has a very high level of technologies for producing industrial robots that can work accurately without being disturbed by noise. Regarding this, a sort of *animism,* that a spirit dwells in everything, is closely related.

In Western countries, people traditionally believe that the correct human thinking is free of contradictions. In contrast, Japanese people place emphasis on intuitive awakening and enlightenment caused by contradictions. The idea that a spirit dwells in everything is more familiar and friendly to the Japanese than the idea that God created everything based on logical rules. As a result, the Japanese readily accept the idea that robots behave and think as if they were human beings. I think this is the essential point which we should now contemplate.

You can make a simulating robot pretending to have emotions as you like, so some people say, "This robot appears to have emotions." But as you know, the emotions emerged in biological evolution. Emotions such as fear and love are part of living. "Fear" is what creatures feel when they are escaping from the enemy, and "love" when raising children. I am convinced that robots, being indifferent to the value of "living," have no genuine emotions.

The story of emotions is closely associated with concepts such as independent *identity* and *free will*. Today, discussions about robotic weapons are being held. We have already seen the practice of using unmanned robots and drones to attack enemies. Such use of robotic weapons can involve cases of mistakenly bombing the wrong target, which inevitably causes objections and protests. Therefore some propose that robots should be made to have a sense of empathy and morals. But I am against such a proposal because a moral sense is inseparable from free will. It is difficult to define the concept of free will, but you cannot place the blame on people, at least unless they have freedom of choice. Remember that a robot merely acts according to given orders. The machine is simply a *heteronomous* system which operates by human-made programs, and has nothing to do with autonomy-based issues. Nevertheless people tend to make such animism-based irrational arguments as "You do not know what will happen. As long as robots make decisions on their own, they are to blame" and "Although they are cute, you cannot tell what they will do if they become unwilling to do what they have been told to do."

Sakura: You mean, we must not do such a thing. Right?

Nishigaki: Exactly. Some experts argue that robots should be equipped with a moral sense, but I am against it. My point is that you should make robots do what people order them to do accurately and efficiently, and that if they make a mistake, the people who gave the order are to blame. This may be a mediocre opinion, but it is a better idea for human survival, because machines are what humans make. If AI makes AI one after another, its activity will become unpredictable and things will be massively confusing. I want to repeat that the real reason for the failure of the Fifth-generation computer was that people neglected to address head-on the difficult issue of logical ambiguities underlying human thinking. They merely pursued high-speed logical reasoning technology. Their development efforts were just outside the scope of humanities studies related to communication, such as linguistics, semiology and philosophy. It is absolutely essential to properly integrate the thought of humanities with technological discussions for effective AI developments.

Sakura: I agree with the argument you just made about the significant problem lying in the gaps among different academic categories. But your argument has caused me to think about another issue: In Japan, there are disconnections between academic research and real society and daily lives. This is probably true for researchers in philosophy, the humanities and social sciences as well as natural science researchers. I think that European philosophers, including Husserl and Heidegger, thought that their theories were linked with their own lives. This is one of the reasons why European philosophy had an enormous social impact. On the other hand, however, Japan imported Western philosophy in modern times and philosophical studies are mainly evaluated by how accurately researchers approached the text. It appears that the connection between academic studies and daily lives in Japan and their reality in society has grown thin.

Nishigaki: I think so too. It is important to study academic achievements made by foreign predecessors, but it is more important to formulate a hypothesis based on those preceding achievements on your own and applying the hypothesis to real society. That can result in the genuine development of academic studies. In Japan, however, it is considered important to build up an accumulation of achievements in accordance with authoritative paradigms as a member of a certain school. This approach is insufficient for achieving significant results in interdisciplinary studies like AI. Japanese academia should place greater importance on the types of people who can think profoundly and broadly on their own.

Sakura: I believe that it is the epitome of the entire Japanese society, rather than a matter of academia only. Today, poor emphasis on the humanities is a global issue, not an exclusively Japanese one. Your argument has made me think that it is our duty to correct the situation whereby in Japan, which has a different culture and traditions from Judaism and Christianity, the humanities cannot play an original and potential role in academia and among the general public.

Nishigaki: I think it is a globally common trend. But now that Japan is no longer a less advanced country and has become an advanced economic power, I hope young people study more and dynamically challenge interdisciplinary problems.

Sakura: You think that Japan is shifting in that direction, don't you?

Nishigaki: I think that the future will be a lot more hopeful than before. Let me summarize my views on the possibility of near-future AI in Japan. I would like to reconfirm that Japanese robotic technologies with painstakingly developed ingenuities are at the highest level in the world. Currently, however, there are still significant gaps between the philosophy behind AI and technological ingenuities. Accordingly, I repeat that it is better to stop working on the idea of creating robots equipped with emotions and making them take responsibility for their behaviors. You should focus on finely developed ingenuities just as tools, at least for the time being. For example, you can expect that easy-to-use nursing care robots that are truly friendly to human beings will be produced. In addition, self-driving cars using AI will also be very important in an aging society. Elderly people cannot do without cars in the local regions. But you should note that you must not place the legal responsibility on AI robots by over-simplistically combining animism with a monotheism-derived mechanistic view of human beings. I believe that it is much wiser to develop original Japanese AI studies based on a deep understanding of the background differences in thought.

The development of AI will lead to a question regarding the allocation of roles. What's at issue is new relationships between the human being and the machine. Human beings will not lose jobs completely, although the specific content of human work will change. There are areas where machines are inevitably weak and human beings must handle. For example, a machine is recently checking tickets at a station's ticket gate while the station staff responds to inquiries from customers. At stations, there are often foreign travelers who want to ask various questions about, say, the delays in train service schedules due to accidents. It is only human beings who can flexibly cope with such situations. If you try to make AI handle such delicate services, things will fall into a stalemate because of the so-called *frame problem*. Environmental conditions change so easily in a complex way that the machine can hardly cope with them.

For another thing, you will have an increasing amount of maintenance work for AI programs, as well as its development. AI does not work by itself, but human beings must maintain it. It is merely an illusion to make AI create AIs on a full scale and make AI carry out maintenance work as well. You need to have a comprehensive idea of what kind of AI should be developed, including the troublesome maintenance work. In other words, you should make appropriate decisions as to which work needs to be done by human beings and which work by machines. What I am afraid is that many people believe it is efficient to make computers do everything. As you know, it is incredibly hard work to maintain programs created by someone else in the past.

Sakura: Especially, when the program is written in an old language such as Fortran. (Laughs)

Nishigaki: What is important is how to handle the joint operation between the human being and the machine, including development and maintenance. If human beings can solve this problem adequately, you will be able to achieve genuine services, or "hospitality," in a broad sense. You must say farewell to the idea that it is progress to make machines do everything.

Sakura: So you mean that it is a significant challenge as to how to design the man-machine system for facilitating efficient collaboration between the human being and the machine instead of separating the human being from the machine, regardless of whether it is in the area of welfare, nursing care, judgment about management, self-driving technologies or military issues rather than a particular industrial field, don't you?

Nishigaki: Yes, that is what's most important. As a result, I strongly suggest that truly desirable man-machine interfaces be pursued, instead of being misguided by the unrealistic singularity hypothesis.

Sakura: You've got it. We should recognize that robots and AI will work for the sake of convenience and happiness for human beings. Japanese cultural tradition is likely to work very positively to this end.

Nishigaki: That's right. I feel that Japanese people can be expected to have a wise balancing sense.

Sakura: Thank you very much for your input.

Translated from an original interview article made for Discuss Japan. [September 2016]

NISHIGAKI Toru, Ph.D.
Professor of Tokyo Keizai University, Professor Emeritus of University of Tokyo

Born in 1948. Graduated from the University of Tokyo in 1972 and earned Doctor of Philosophy in Information Engineering from the University of Tokyo in 1982. He worked as a researcher for Hitachi Ltd. (1972-1986), during which he also served as a visiting scholar for Stanford University (1980-1981). Interested in social/philosophical aspects of information technology, he stayed at University of Reims, France (1994-1995) to study French modern thought, while teaching at Meiji University. Then he served as a professor of information studies for the University of Tokyo (1996-2013) and currently a professor at Tokyo Keizai University, being involved in the research of Fundamental Informatics. His publications include *Digital Narcissus*, 1991 (Suntory prize for Social Sciences and Humanities 1991), *Multi-media*, 1994 (Telecom Social Science prize 1995), *Fundamental Informatics*, 2004, etc.

SAKURA Osamu, Ph.D.
Dean and Professor at the Interfaculty Initiative in Information Studies, University of Tokyo

Born in 1960. Earned his Ph.D. in 1992 from Kyoto University. Recent activities include the popularization of science for those without background knowledge. Served as Assistant Professor at Yokohama University (1993–1994) and Visiting Researcher at Freiburg University (1995–96). Has served as Advisory Board member of the academic journal *Biology and Philosophy* since 1998. His publications include *Watashitachi wa dokokara kite dokoe yukunoka?* (Where did we come from and where are we going?) and *Shinkaron no chosen* (A challenge to evolutionary theory).
