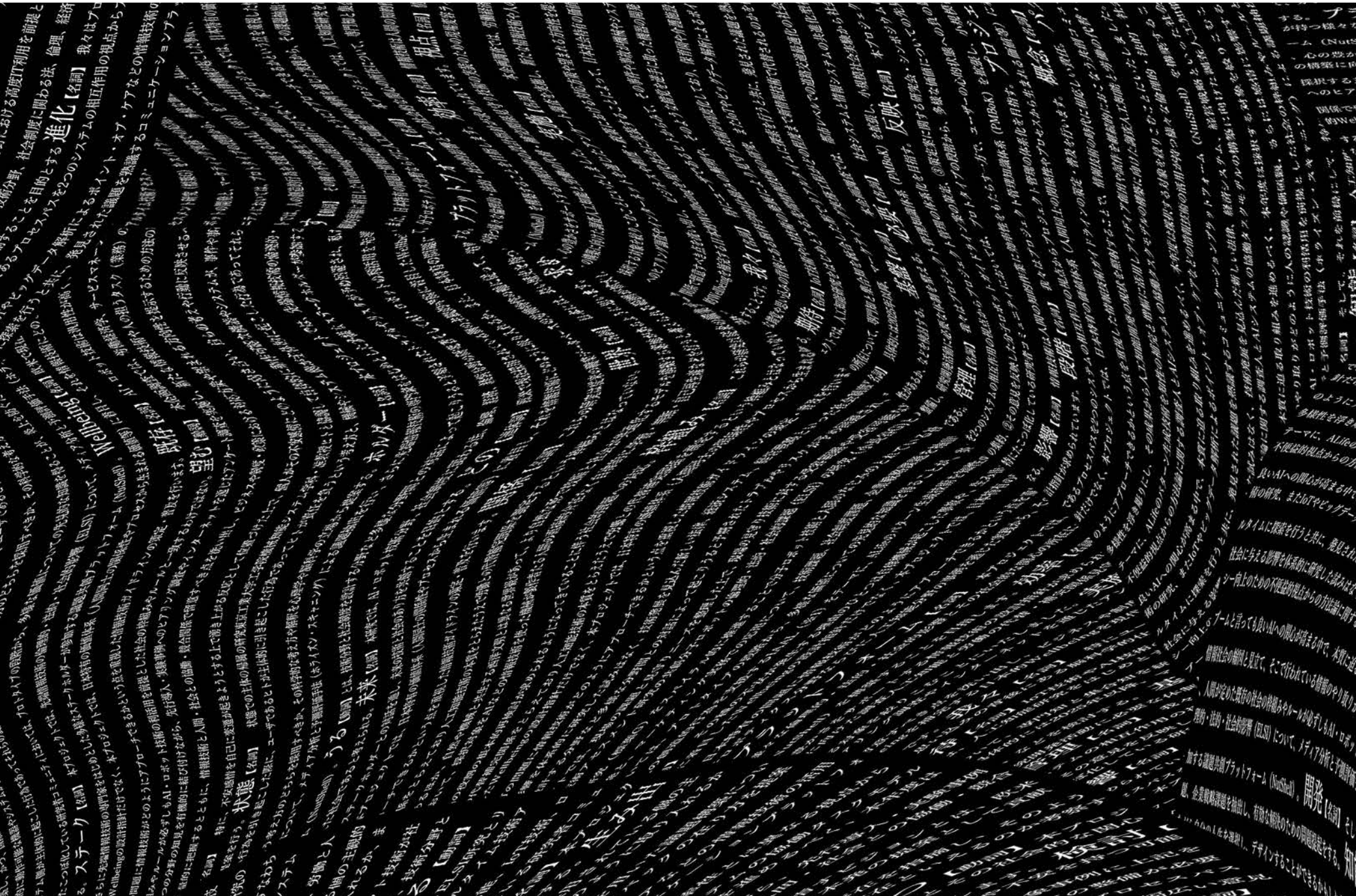
# Human-Information Technology Ecosystem

Vol.02

HITE
Human
Information
Technology
Ecosystem



## Message from the Program Supervisor

It is probably not productive to overestimate the power of AI and to exaggerate when speaking about its negative social impact. Issues such as the destruction of employment due to productivity improvements or automation have existed since before AI, even without mentioning the Luddite movement. We have been, or at least have been trying to be, wise enough to come up with social solutions to maximize benefits while minimizing the negatives. In addition, people are already living their everyday lives while entrusting judgments to various automatic control systems. Perhaps, unexpectedly, there are not many new problems that are unique to AI.

Even maintaining such composure, however, one cannot assert that all of the concerns regarding modern information technology as exemplified by AI are without basis. These include the concern that the black box-like nature of judgments could increase and humans would no longer be able to control the technology; the concern that humankind could be threatened if such technology were used maliciously; the concern that the gap between rich and poor could be expanded based on the destruction of middle class employment or the ownership of systems (including intellectual property); and the concern that profiling technology could advance, we could move toward a society of mutual surveillance, and human rights could be violated. In proportion to the possibilities offered by AI, we wish to gain the wisdom to cultivate and accept the AI in order to bring about a more positive result for humankind.

In terms of an action in the near future that could facilitate this, the creation of some form of guidelines for technical development comes to mind. However, this is no simple task. This is because we cannot completely predict the direction in which technology will progress. Further, we know of historical instances in which the side of society accepting technology has developed it in unforeseen ways. People's values, which form the foundation for rules, change as well. The aspect of how commercialization will be undertaken also exerts a significant impact on both technology and society. In that context, the "Human-Information Technology Ecosystem" area of RISTEX perceives technology, social systems, and people as things that evolve together, and aims to create a platform in which people and information technology can coexist (adapt) through understanding and dialog. Toward those goals, we are performing activities with the aim of creating fundamental concepts, building a community of researchers across multiple fields, creating a platform in which stakeholders can give mutual feedback, and so forth.

Program Supervisor
Jiro Kokuryo

Jiro Kokuryo

Program Supervisor of "Human-Information Technology Ecosystem"
Professor of the Faculty of Policy Management, Keio University

Jiro Kokuryo graduated from the Faculty of Economics at the University of Tokyo in 1982, and then joined the Nippon Telegraph and Telephone Public Corporation. He obtained a Doctor of Business Administration at Harvard Business School in 1992. From 2003, he served in positions such as being a professor in the Faculty of Environment and Information Studies at Keio University, and in 2009, he became Executive Director of the Faculty of Policy Management. From 2005 to 2009, he also served as the Chief of the Keio Research Institute at SFC (Shonan-Fujisawa Campus). In 2013, he assumed the position of Vice-President at Keio University. His main authored works include Open Architecture Strategy(Diamond, Inc., 1999) and Business Strategy in an Onymous Economy (Nikkei Inc., 2013).

# Human-Information Technology Ecosystem

"Human-Information Technology Ecosystem" is an R&D area promoted by the Research Institute of Science and Technology for Society (RISTEX) of the Japan Science and Technology Agency (JST). This is the area's third year. In today's world, where AI, robots, IOT, and other types of information technology are progressing at an accelerating rate, this initiative is promoting R&D for the design of new institutions and technologies, while considering how such technology permeates society to become a familiar part of human life, and what sort of problems can arise in that process. The objective of this area is to realize a society that fosters more beneficial relationships between information technology and human beings.

**Website:** http://ristex.jst.go.jp/hite/

Special Tripartite Discussion

Date of interview: November 26, 2017

# What are the responsibilities and the subject of the artificial intelligence era?

**Kazuya MATSUURA**
philosophy

**Takako YOSHIDA**
psychology

**Tatsuhiko INATANI**
law





**If artificial intelligence causes an accident, who will take responsibility? In this era where information technology rapidly penetrates society, it is an important issue that has yet to be resolved. To challenge this contemporary problem, three researchers, each one having different expertise ——philosophy, psychology and law—— met to discuss what "responsibility" is and what "subject" in the age of artificial intelligence is.**

*First of all, please tell us about your research and the project contents being promoted by "Human-Information Technology Ecosystem" (hereinafter: HITE)*

**Kazuya MATSUURA (below: MATSUURA):** My specialty is Greek philosophy, mainly that of Aristotle. While I know some may feel Greek philosophy is too unwavering, I believe that Greek philosophers offer a timeless way of thinking which we share unconsciously, and their discussion and values serve as hints to comprehend the current society. At HITE we launched a project titled "Consideration and Suggestion on the Concept of Responsibility in the Sophisticated Information Society" last year, and starting this fiscal year we are carrying out a project "Consideration on the concept of 'responsibility' between autonomous machines and citizenries". These projects aim to clarify and reconsider the concept of "responsibly" in the coming sophisticated information society from the viewpoint of humanities.

Our project emphasizes human history and culture. When some sort of "autonomous machine," such as cars with AI, spread throughout our society, the point of argumentation would not be how the machines actually move or behave, but how ordinary people could accept the machines. Therefore, we put less weight on seeking the definition of autonomous machines, but we ask this question instead. Namely, "what abilities are needed for machines to be regarded as the equivalent of human beings?" This question, of course, just leads to the more universal question of "What a human being is." If we want the system of society co-existing with autonomous machines to be more ideal, it is necessary to advert not only to the social models offered by modern Western philosophy and political theory, but models in other periods and eras, such as ancient Greece, the Edo period of Japan, ancient India, etc.

**Takako YOSHIDA (hereinafter: YOSHIDA):** In our laboratory, we are working on machines and systems that work in accordance to the actions of the human body. At HITE, our theme is to observe the human user's psychological state to answer the question of "Which can be the subjective and objective subject of a specific action when human beings and artificial systems are integrated into one and working together, the human or artificial system?" Our interest is especially on the user's subjective feeling of affinity of machines and the human body from the approach of brain science. For example, when using a wearable power support robot attached to the human body and working with it, when a certain condition is satisfied, the human user gradually feels that the machine system can be part of his/her own body. In the end, he/she feels strongly that he/she, himself/herself is the only subject to control his/her own body, and loses the feeling that the machine may be also controlling his/her body. My question is, "Who is responsible when some socially unwanted event happens in that state?" From the perspective of the person himself/herself, all actions feel as if they were made on his/her own solo intention, which then leads to them feeling responsible for himself/herself even if the problem is physically on the machine side. On the other hand, despite the result of an action caused by himself/herself, he/she declares that it is a malfunction of the machine and passes it to the manufacturer's side in some cases. Who judges whether the person himself/herself is really thinking this or just lying? Furthermore, he/she can also declare that the machine took over their own body actions. This can be a tough situation because from a third person's point of view, as the machine and the person are working together, it is difficult to distinguish which controls which, the machine or human user. Under such circumstances the boundary between such machines and humans becomes ambiguous. I think it is very important to figure out "subject" and "responsibility" of who performed a specific act in this type of human and artificial system co-operation.

**Tatsuhiko INATANI (hereinafter: INATANI):** My specialty is criminal justice and criminology. Specifically, I'm conducting research of substantive and procedural law in the criminal justice and legislative theory, while applying the theory of other areas adjacent to law such as philosophy, economics, sociology and cognitive science. Nowadays, with the progress of globalization and technology, fluctuations have occurred in the idea that "A human being is a rational 'subject' with 'free will'", which has been regarded as a solid premise in the modern legal world up to now. So while making full use of contemporary philosophy, cognitive psychology, behavioral economics and other knowledge, I critically examined the current criminal justice system and explored a new criminal justice system. In "Legal being: electronic personhoods of artificial intelligence and robots in NAJIMI society, based on a reconsideration of the concept of autonomy" project I participated in with HITE, we have dealt with recent hot legal issues such as whether a legal personality should be given to artificial intelligence. We are also developing discussions on what is legal responsibility concerning the development and use of artificial intelligence, especially on what punishment system is necessary for artificial intelligence to exist in harmony with human beings.

## Is there no "free will" for humans?

*What is the most important point concerning "responsibility and subject" in the artificial intelligence era?*

**INATANI:** As I mentioned earlier, in modern criminal law based on modern philosophy of the West, it is a basic premise that humans with free will can control objects without being influenced by the external environment. However, artificial intelligence such as deep learning is something that continues to develop, so complete control of it cannot be assumed. On the other hand, according to recent neuroscience and brain science, human beings always exist under the influence of the external environment, so it is also revealing that the existence of a firm free will is doubtful in the first place. Then, it can be said that the premise of the argument that "if a harmful event occurs, the person who could control the danger should take responsibility" is now questionable.

**YOSHIDA:** In the field of cognitive science and brain science as well, the possibility that human beings are not actively and subjectively controlling their own actions as much as they think by themselves has been discussed. With human beings, we can find varieties of phenomena that the human subjective feeling to control

themselves does not correspond to their actual action in the physical world. In addition, recently the possibility that you tend to argue that you are not responsible for the action coerced by other people and agents such as machines and AI is being discussed. Before debating the social question of "Who is responsible", it is necessary to carefully consider the characteristics of this type of human cognition and behavior.

**MATSUURA:** The very concepts of "responsibility" or "subject" are the product of modern Western philosophy. One background of these concepts may be "the principle of alternative possibility," which enables us to blame someone and assign guilt to him/her, claiming that he/she was able to choose another action. However, this principle cannot be agreed by everyone. For example, if we recall an ancient Indian idea, we could say that the accident was caused "because of his/her karma from the previous life." Or, if a slave committed a crime in ancient Rome or Greek times, it was the master who was accused and had to pay the indemnity because of the responsibility of management failure. So our project, referring to social models in the past, reconsiders if the society formed with the modern concept of "responsibility", that of "subject", or "the principle of an alternative possibility" can really lead us to a happy and prosperous future.

## Where is the machine from? Is it a human?

*When responsibility is asked from the artificial intelligence side, at what level can the machine be determined to be "autonomous"?*

**MATSUURA:** I think that your question will be more important when artificial intelligence penetrates the daily lives of people. However, rather than defining the level of autonomy of machines or "autonomy" itself, it is more efficient to argue the possibility that ordinary people, including me, would regard artificial intelligence as "having autonomy and intelligence like human beings", if we aim to adapt them as "culture" into society. Therefore, we have no choice but to answer the question, depending on how each culture sees the "autonomy" or "human beings" in question.

**YOSHIDA:** There is scientific research on "animacy" to reveal how and when humans "feel" intelligence and life in mechanical and computer graphics objects, and it seems to be easier to define and study compared to the precise scientific definition and study of what is "true" intelligence and life. This is also relevant to the Turing test, which may be "artificial unintelligence" rather than artificial intelligence. In other words, it can also be said that human beings can have the subjective feeling or illusion that AI is a living thing, even for things that may not have life. I have to think carefully whether we should discuss today's rapidly developing AI in the same way.

**INATANI:** In the framework of the mainstream criminal law argument, we may try to capture the autonomy from the viewpoint of the sanity that human beings should have to be legally responsible. However, as I mentioned earlier, the answer to the question, "What is a human being?" that can take legal responsibility, in itself is changing. It might be the time to rethink the premise of traditional framework based on modern philosophy, which essentializes the mode of existence of human beings and the sanctions against them.

**MATSUURA:** If we seriously pursue "autonomous human beings", it may mean the people who are not affected from the outside at all. Whether such people exist is doubtful, except for the great philosopher Immanuel Kant.

**YOSHIDA:** I don't know Kant in person. If there is a person or artificial system such as AI or a machine that lives totally independent from their surrounding environment, they may act completely independently from human social common sense. I wonder if they can live without any trouble in an ordinary human society.

## The gap between technical "safety" and psychological "security"

*Recently, there has been debate on whether the manufacturer or the driver is responsible for an accident involving an automatic driving car. What do you think?*

**INATANI:** Autonomous driving is defined by 0 to 5 technical levels according to the SAE, and conditions are totally different depending on each level, so it cannot be said unconditionally.

**YOSHIDA:** I am concerned that the design concept of semi-automatic operation of level 3 or so for automatic driving is not widespread throughout the world; (the machine is driving in a specific place such as a highway, the person is sitting in the driver's seat only for the purpose, according to the design philosophy, of being able to respond to an emergency). Compared to the scenario in which a human and artificial system are operating relatively independently, it seems possible that we have more accidents in a state where the machine and human beings are cooperating together. This may be what is known empirically in other automation technology fields such as aircraft autopilot incident case reports. There seems to be a certain number of researchers claiming this; Compared to fully automatic driving, accidents are likely to occur if human operation intervenes halfway.

**INATANI:** Regarding level 3, I think that ultimately this kind of mindset may unconsciously be derived from the premise of modern philosophy that the control of a human being based on its free will over the object is better than the machine control over the human being. In modern law, the driver is a human being, so he/she can and should manage the danger of objects based on free will. Then the idea that there is responsibility as a human being for finding and controlling the danger of the vehicle is straightforward in a sense. Although the problems caused from it may be enormous.

**YOSHIDA:** Suppose that artificial intelligence that can solve any problem perfectly was driving a train or an aircraft, there is also a viewpoint on whether human beings want to ride or not. Some people may think that a human operator is necessary to cope with an emergency since they should be more responsible than a machine.

**INATANI:** There is a gap between the psychological sense of security and objective safety. It may be the biggest problem that is complicating the discussion.

**YOSHIDA:** Especially in the case of semi-automatic driving, it is important to consider (1) what kind of cognitive characteristics the driver has, (2) and how to maintain a sense of responsibility of the driver in a safe and comfortable manner. Based on the above (1) (2), the vehicle system may be designed to operate while keeping these human characteristics the best inside it.

**MATSUURA:** In terms of machines and humans working together, there are also approaches to the design of machinery systems to extend or support human actions and abilities. I expect technology to support and enhance our ability. The automatic driving technique may proceed in the direction of improving the ability to assist human driving ability rather than aiming for complete automation. It is the very design, I believe, which more than a few people really want, such as handicapped people and their supporters.





## Instead of pursuing problems, creating an ideal social vision

**INATANI:** Either way, I think that it is not very productive to pursue only the responsibilities of users and developers just because artificial intelligence causes danger. From now on, rather than trying to determine what the essence of machines and human beings is, everyone should think about what kind of society we want first, and then start to discuss the distribution of legal responsibilities appropriate for that purpose. In the case of autonomous driving, I think that it is better that everyone starts concretely thinking about how they want it to exist themselves and spread that idea into society individually. From that point of view, I feel that it is better to loosen the traditional way of establishing preliminary regulations by the government, based on its idealized and sometimes very fixed social images. I think that it is better not to discuss "how society should be" to fix every problem immediately or even ex-ante, but to adopt an approach that allows us to gradually develop a synthesized society between human beings and artificial intelligence centered on the debate over "what we want to be with them".

**YOSHIDA:** There are also ways to give quick prototypes to future users, to allow thinking while gathering feedback.

**INATANI:** I agree. Even after actually distributing the autonomous systems, if they cause problems, we should thoroughly discuss with users, companies, professional technicians and legal professions involved with them, like "Let's do this if we can control it", or "If you cannot yet control it, try this". I would like to prepare a legal system to embody the ideal mode of existence, little by little, while discussing the direction everyone is aiming for each time.

## How to incorporate people's voices

*Will artificial intelligence be acceptable to society? Will problems that cannot be divided by theory and institutions come up?*

**INATANI:** Who will be the subject to discuss it is important. For example, now you can also widely incorporate opinions from the public through SNS, and it is possible for artificial intelligence to analyze those voices as data. It is important not only to discuss experts' opinions but also to listen to the opinions of the general public. On the other hand, if you put too much emphasis on the user's opinion, the voice of uneasiness will rise too much, like "Let's eliminate all danger first", and there is a possibility that regulations may become too strict, so I think that it also needs attention.

**MATSUURA:** I strongly agree that it is necessary procedure to reflect the opinions in the social system, but I am concerned about two problems. One is "whose opinion is to be collected", and the other is "whether the consolidated opinions harmonize with current legal systems and culture". As AI's learning shows, the output of AI definitely depends on what kinds of input we give it in the learning process. If AI learns from violent people, the AI will produce violent output, and if you gather people who are interested in a particular religion as the teachers for the AI, it will produce an output that reflects that religion. It can be the same in the case of us. Likewise, I do not deny that aggregation of opinion from the public is one of the important processes, but the most important is the method of how to summarize their opinion. Otherwise, I am afraid that we would make different rules for each region or person's taste, or a law that conforms to an overly idealized human figure.

*Finally, please tell us about the possibility of your future collaborative research*

**YOSHIDA:** I think that making opportunities to listen to people in different fields and not ignoring their ideas is important. Since there are not many opportunities for the Tokyo Institute of Technology, especially Mechanical Engineering students and researchers to have contact with humanistic thinking, I will soon invite Dr. Matsuura and Inatani to the Tokyo Institute of Technology and apply law and philosophy, etc. I want to create the opportunity to talk with people in different fields. Also, I would like to involve more people in this type of discussion.

**MATSUURA:** Like Dr. Yoshida, we scholars of literature should know more about the current technology, and have more opportunities to communicate with professional researchers of industry. Through this opportunity, I would like to have discussions beyond the boundaries of modern specialized fields. I would be happy if we could think of concrete methods of developing such discussions at educational institutions.

**INATANI:** In terms of education, I recently gave lectures for high school students. They, as children of the digital era, could understand the characteristics of information technology and artificial intelligence very quickly and they were very open to accept the social changes that they might bring about. Their ideas are very important even now, and there is a possibility that their flexible minds and frequent activities with artificial intelligence will innovate its future in our society. Also, since law is a field that can collaborate with all other fields, it has the ability to infiltrate everywhere. I hope to learn a lot from the specialists in other fields and to figure out a concrete legal proposal in the future.

**MATSUURA:** We are in the era in which we must go beyond the framework modern law has been built upon so far. Even if we present the best, we philosophers should present many options which we can choose from, and we can also consider how society would change. Today I thought that there was a lot to be in line with law specialists like Dr. Inatani. It seems that the time has come when we can create laws and societies that are more in line with our diverse values and lives.

**INATANI:** Based on this HITE initiative, I think that we may present a unique mode of law based on Japanese culture and philosophy. Applying the theory on the premise of an ambiguous division between "subject" and "object" to construction of the society, which we've been discussing today, is somewhat difficult in particular for the people in the West where they still try to maintain the division based on their cultural pedigree. Because Japanese culture and philosophy is not so firm on this problem, I think, there might be possibilities to innovate a new resolution to current and future problems caused by artificial intelligence, beyond the framework of Western modern philosophy. I would like to make this collaboration an opportunity to tackle the initiatives that stimulate and involve the whole world.

**Kazuya MATSUURA**
Principal Investigator of the HITE project "Consideration on the concept of "responsibility" between autonomous machines and citizenries". Associate Professor, Toyo University. Professional in Greek philosophy.
Formerly Full-Time Lecturer, Faculty of Teacher Education, Shumei University. Specializing in Philosophy.

**Takako YOSHIDA**
Principal Investigator of the HITE project "Which controls which? Sense of agency when humans and semi-automated systems co-operate". Associate Professor, Tokyo Institute of Technology. Specialty is applied brain science.

**Tatsuhiko INATANI**
Member of the HITE project "Legal being: electronic personhoods of artificial intelligence and robots in NAJIMI society, based on a reconsideration of the concept of autonomy". Associate Professor, Kyoto University. Specialized in criminal justice and criminology.

# A Thinking Method to Prepare for Unforeseen Futures

Yuichi Washida

Date of interview: November 14, 2017

**Professor Yuichi Washida has been conducting research on foresight in the HITE program, using "Future Scenario".**
**With information technology developing at a breakneck speed, we are now in need of a new foresight method. We asked him about the possibility and significance of Future Scenario, to prepare for unforeseen future events.**

*In your research, it is noted that using "Future Scenario" as a thinking method is effective for unpredictable futures with issues such as coexistence between AI and human beings.*

When we look back at cases of decision-making that have been made by the government and businesses in the past 30 years, we see a common factor lying behind their mistakes. To give you a few examples of unexpected future events that Japan has experienced in the past 30 years, there were the developments of the "Personal telecommunications device" and "Fifth generation computers". In 1991, the "2010 Technology Prediction Study Group" in the General Planning Bureau of the then Economy Planning Agency had estimated the market size of national personal telecommunications devices in the year 2010 to be around 500 billion yen. However, the actual market size had already reached over 2 trillion yen by the year 2005, more than five times the government prediction. This is one example of an upturned version of an unexpected case.

On the other hand, the fifth generation computer development was a downturned version of an unforeseen case. In 1982, the then Ministry of International Trade and Industry (present Ministry of Economy, Trade and Industry) launched a national project to develop the fifth generation computer, announcing that "Japan will aim to develop a pioneering computer equipped with AI". However, just a decade later, the project closed with scarce results. The capital invested for this reckless goal amounted to 57 billion yen.





Why were these schemes predicted erroneously? It may be because, since the time of the rapid growth period in Japan, the Japanese government and businesses have always been making future predictions based only on technological perspectives.
Future scenario making, in which interactive workshops are the key to work integration through its process, not only the technological perspectives, but also various other perspectives such as the cultural background of people, social conditions such as population composition, and changes in the natural environment is being adopted. By venturing to actively discuss more uncertain themes such as politics and cultures, I think that we can propose materials for alternative decision-making, which will not arise from simple linear predictions made only from a technological development basis.

## Non-linearly pictured, wide view futures

*What process do the Future Scenario workshops take?*

In the "First Foresight Workshop" held by the RISTEX HITE project, we made a future scenario themed for propagation of future AI and IoT. The participants were leaders and well informed persons mainly in marketing fields, and the workshop lasted two days. On the first day, "social change hypotheses" were built, which are scenario generations of social changes that may arise in 10 to 20 years, using the method called "Horizon Scanning". First, we prepared a database, the "scanning materials", that could be the signs of future changes. The participants discussed in depth using the database, picturing what kind of society we will see in the future. (The following page shows the "Chronological Table of the Future" of the designed social change hypotheses.) On the second day, we considered the problems that may arise after AI and IoT have propagated, naming them "Future Issues". In this workshop we came up with issues such as discords and bipolarization arising from the mixture of automatic driving vehicles and current cars in society. Then we further discussed the social change hypothesis and future issues given on the second day, and completed the "Future Scenario".

*Please tell us of any particular interesting cases of Future Scenario in your extensive experience with future scenario studies.*

In 2002, we held a Foresight Workshop to design the "Future vision of Akihabara in the year 2008" with KDDI. Looking back at that time, it is interesting to note that the idea of the "Eyeglass type mobile phone" came up in that workshop. Wearable devices using a location information system were already thought of, and privacy issues that would accompany it were also discussed. About a decade or so later, the "Google Glass" was launched, using the same kind of system, but the developers could not solve the personal data issues either. They soon gave up on releasing the product to the public and narrowed it to B to B uses only. This was an informative case of future prediction.

## 2025 Problem and the coming of a mosaic society

*With the rapid growth of information technology such as AI, we will be facing more and more unpredictable futures. What is your idea of the things we can do with the use of Foresight?*

By holding many workshops, we came to understand that there is a certain common pattern in the future change hypothesis. It is that after the year 2025, society will change directions, from optimism to pessimism. The same hypothesis arises in different groups and different themes. This shows that many people think that "In 2025, we will face an unforeseen event". We call it the "2025 Problem ".
We have held a Foresight Workshop in this "Human Information Ecosystem Project" and verified the hypothesis with the theme "Would the same kind of 2025 problem also arise in AI and IoT?" A brand new hypothesis came up. It is that, when society continues evolving, a collapse will occur at some point. We have named this phenomenon "Mosaicing society". What we mean by "Mosaicing" is that society will be like a mosaic with a mixture of AI implemented advanced areas and conventional areas. Society will not be visualized with a simple future prediction, as people today often question, "Will AI deprive humans of jobs? Will AI rule over us?"
Businesses and developers tend to picture an idealistic advancement of technology in simple linear growth, and they do not foresee this mosaicing propagation pattern. I think that Future Scenario can fill in the gaps.

*How will the Future Scenario contribute to the future of Japan?

Future Scenario can provide us with "Preparation" for unpredictable situations in society if we use it appropriately.
Especially for businesses involved in information technology, I think it will be an effective method to consider how information technology such as AI and IoT will spread into society.



In a conventional style, when people carry out a business plan, they tend to "conduct research, and then discover a trend from the current technological trend". However, much of information technology is consumer electronics based, in other words, it is a technology that is very close to consumers.
Information technology such as the typical smartphones has "Network Externality" (the effect that an increase of service users enhances the value of the service itself). It means that it can become a peerless product in a very short period of time. In other words, the first mover advantage is so large that any business that enters will become a "front runner". I am sure you know that many of those new services happen to come from the U.S.
With information technologies that have such a property, we have to be prepared for various futures that may unfold from all kinds of situations. I think that Future Scenario is an effective thinking method under such circumstances.

**Yuichi Washida**
Principal investigator of the HITE adopted project "Scenario generation of socio-technology problems in the information technology area by using the foresight method".
Professor of the Graduate School of Business Administration, Hitotsubashi University
Research fields: marketing, innovation research, future insight etc.

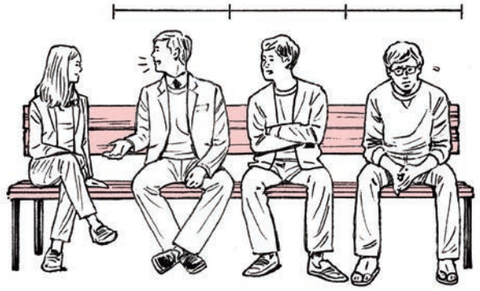
First Future Foresight Workshop

# How AI will spread in society

## Future road map until 2040

## 2018

### Young people can be divided into three types

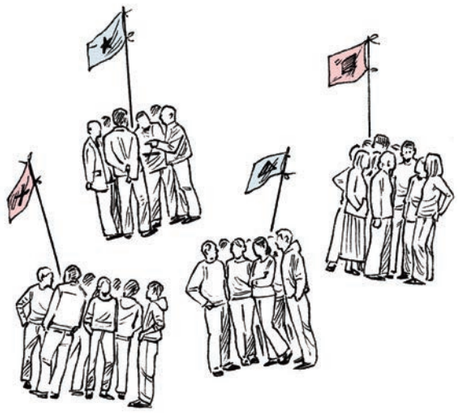


- Young people can be divided into three types, namely elite communicators, intermediate layer communicators, and poor communicators, and the separation between these types becomes prominent.
- Lateral communication weakens based on the age of young people, and the progression of individualization in communication.
- Sharing economy is enriched, and there is progress in the "Cloud computing system" or "Flow" that does not save or own anything, to manage entire lifestyles on the network.
- Schools will function only as community hubs.
- Business for people who feel disconnected will shift from "Knowledge" to "Comfort".

## 2020

### Wealthy elderly will become 'Daddy-Long-Legs'

- As a result of an extreme trend towards nuclear families, the wealthy elderly people with excess money will economically support the young people regardless of blood relation = They will impact society as 'Daddy-Long-Legs'.
- With the growing lifespan, the living time in old age will be longer than the young age, and many people will lead a life of "Live for old age"
- Smart device interfaces will get simpler, and there will be progress in the trend towards universal service of AI.
- "Service matching society" that can automatically match using AI, rather than searching the products or services based on the individual, will arrive.



## 2020

### Course of the nation will change

- Triggered by the collapse of the pension system and the increase in self-pay burden of medical expenses, the welfare provided by the government to the citizens until now will become limited. As a result, the influence that the nation can have on its citizens will get weakened.
- A diverse community will form like an alternative nation. Momentum to go independent from the present nation will increase.
- Friction will occur with the existing national system. Businesses that use the friction and information arbitrage will be created.





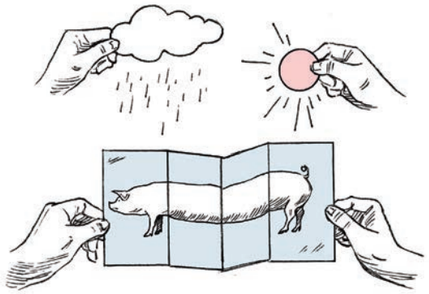
## 2030

### AI will form a 'Mosaic society'

- AI will be introduced in industry, society, life and capital. People who "Entrust" decision-making to AI, and people who "Stick" to their own decisions, will mix in society, which will continue creating a "Mosaic".
- Technology that can analyze big data of the real world, and then extract and use the meaningful data, will dramatically improve.
- The Cyber world will collaborate closely with the physical, real world, and the trend of "Cyber physical" of society will progress.

## 2035

### Human beings will be able to freely handle nature



- Genetic engineering of agricultural and livestock products will progress, and nature itself will be artificially created, such as the artificial construction of its cultivation environment.
- New sustainable energy resources will be available by controlling nature, such as with the possibility of certain types of weather manipulation.
- Technology to artificially create food will be established and spread. A food revolution will occur.

## 2040

### Great ethnic migration will occur in the virtual and real world


- A New Northern era (prosperity in high-latitude regions) will arrive due to the global climate change. Ethnic units in the real world, or migration between individuals, will be generated.
- At the same time, a virtual nation on the cloud will also prosper. This will lead to a sustainable social change.
- Terrorism and war due to the development of information technology, or unknown accidents and diseases due to excessive progress in biotechnology will lead to unexpected negative impacts. Global scale risk will increase, and a society that utilizes these technologies will be born.

## AI technology will spread to mosaic type

Professor Yuichi Washida, who researches and practices the methods of future foresight, invited experts and assembled the "First Future Foresight Workshop" to predict the future spread of information technologies such as AI and IoT, for the HITE project. As a result, a future scenario in which a "Mosaic diffusion" of AI and IoT that will occur during the years 2025~30 was created.

It was suggested that the social divide will progress in the background, and welfare services by the nation will weaken. As a result, a scenario has emerged in which there will not be uniform change, such as the one envisioned in the current topic of interest where "AI steals humans' jobs", and "Society converts to IoT all together", but technology penetration will progress with considerable "variation" within society. Professor Washida and others called this variation phenomenon "Mosaicing of society", but it will be a big challenge to face future technology development, after anticipating the problem. The conception training of imagining every future scenario, and flexibly preparing for each of these futures, is called "foresight".

* The results of this workshop are also published in the "Marketing Journal Vol. 37 No. 1 (2017)".

Date of interview: November 14, 2017

# Using "Power Laws" to assess the degree to which AI adapts to human society

**Kumiko Tanaka-Ishii**

**Professor Kumiko Tanaka-Ishii has been trying to comprehend the mathematical universal properties underlying human communication and natural language. She uses the perspective of a statistical model called the "power law" as the key to assessing how well AI is adapting to human society. We asked her about her research and her vision for society.**

**Kumiko Tanaka-Ishii**
Professor of the Research Center for Advanced Science and Technology, the University of Tokyo
Area of Specialization: computational linguistics, complex systems science for natural language, communication and social systems. Principal investigator of the HITE adopted project "A coevolutionary study on society with respect to power laws: - Can AI replicate the behavior of a non-equilibrium complex system?"

## "Power laws" everywhere

*What is the "power law" model?*

A power law is a physical law that is observed empirically in various natural and social systems. In essence, it is a model that forms a proportional relationship between two statistical variables (excluding constant terms), when plotted on logarithmic axes.
Curiously enough, this power law distribution can be observed in various phenomena, such as earthquakes, market prices, natural language, and the population distribution of cities. For example, in earthquake statistics, plotting the energy on the horizontal axis and the frequency on the vertical axis, we see a graph that shows a clear power law distribution. Power laws can also be found between the size of cities and their population, or the number of references contained in technical papers and their frequency. Plotting the annual income of millionaires against their income ranking also produces a linear relationship.

*When a graph shows a "power law", will it hold without exception?*

Surprisingly enough, in most cases a power law seems to hold without exception, but this depends on the measuring technique. If the measuring technique is not good enough, a method might not show the underlying properties. I started my career studying natural language. Statistics of the words used in English texts show that the most frequent word is "the", followed by words such as "of" and "and". When the frequencies of words are plotted against their frequency rankings, a power law graph is produced. This power law has been studied to hold in various types of texts, across languages, genres, cultures, production times, authors' age and gender. It even holds with the child-directed speech of three year-olds. No one is consciously aware of talking in a way that produces a power law, but curiously enough it happens anyway. What's more, the reason for this distribution remains unknown.

*How can "power laws" be utilized in the study of AI?*

Machines have simple designs, so it is not clear whether a power law can be produced with machine-generated texts. For example, until a decade ago, machine-generated texts hardly show power law distribution in ways that were similar to human language. Twitter bots are designed to quote other people's words automatically, but it is still questionable whether they display a power law. Recently, however, when using deep learning, these machine-generated texts have started to display "power law"-like behavior, as shown in Fig 1. However, on the whole, AI performance has a long way to go.

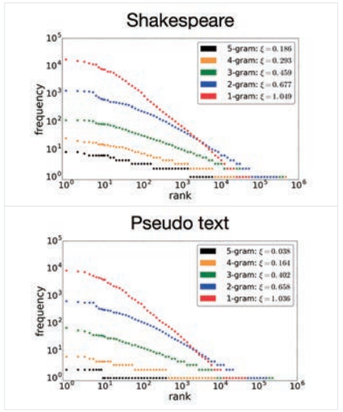



fig1: [Above] Power Law of the Shakespeare whole sentence. [Below] Power Law of the document which the deep learning that learned it generated.

This was the starting point for our project. Accordingly, we set out to consider the question "How natural are the results produced by AI?", or in other words "How well has AI adapted to human society?" using the following question as a gauge - "Do machine-generated results produce a power law?".

*What is implied by the formation of a power law?*

A power law presents a statistical self-similarity, similar to "fractal structures" like the Koch curves. A self-similar structure contains many similar structures but at smaller scales within the whole structure. Likewise, a power law presents the same exponents when the scale is magnified or minimized. As mentioned, this is a property that can be seen widely in nature and social phenomena, but because machines are man-made, they do not have such structures in general. Therefore, we look for power law formation as an indicator to measure whether self-similarity exists in the systems produced by machines.

## Power laws as a new indicator for assessing AI

*What are the merits of applying this measure to AI?*

We are considering how to create a system that detects any AI systems that do not produce power laws, so that for example we could exclude AI with the potential to behave recklessly in the field of investment. In stock investments, if an AI keeps on pursuing short-sighted maximum gains, it is likely that its behavior will destroy the power law, and this could risk the whole market. We believe it might be possible to detect these signs in advance by using methods that use power laws as the criteria for making judgments.
Moreover, the power law method serves to maintain the diversity that underlies human society. In human society, we have what is called a "long tail" property, which corresponds to the notion of the power law. Until recently, a business that deals with artworks such as movies and music, was considered to be successful if it produced a major hit. However, charts composed only of big hits do not form a power law. It is known that their sales are certain to fall at some point. As an example, in the case of a digital juke box containing 10,000 pieces of music offered via broadband for three months, it was reported that 98% of the pieces were accessed by users. This shows how vastly varied people's tastes are. Evidence of such phenomena can be observed even more clearly in situations such as online shopping. Online shopping does not just provide popular items: it is known that the more varied products an online shopping website offers, the more successful its sales will be. This shows how a small number of users always want something specific, but that there are many such small users, and the accumulation of these users is an important factor.

*What exactly are you going to be doing in this project?*

In brief, we are dealing with two themes - a technical study, and social implementation. In the technical study we will first analyze the performance of AI to see if a power law is formed. Generally speaking, AI does not form a power law. Even when there seems to be a power law, there are often crucial differences from human behavior. Given these results, we are interested in improving the current mathematical models underlying AI, so that their outputs satisfy power laws. As far as social implementation is concerned, we will try to consider the problems in AI investment and propose new social norms that can serve to maintain sound markets.

*What problems lie in AI investment behavior?*

In May 2017, the Financial Instruments and Exchange Act was amended in Japan in order to better accommodate AI in markets. However, the amendment only focused on keeping logs and submitting these records to the Financial Services Agency in the event that trouble arises. This would be nothing more than a response to something that had already happened. If a market crash similar in scale to Black Monday should occur as a result of AI, even if we analyzed the cause of such an event afterwards, it would be too late to avoid the huge impact. This is why we need measures to prevent such an event from occurring, or at the very least to stop it while it is happening. For this reason, we are discussing how to assess AI behavior with the help of power laws by detecting the moment when the power law stops holding. I hope this project will be a good start to considering questions such as "How can we stop AI that is motivated purely by profit?", and "How can we design evaluation methods to prevent such behavior?"
I would like to offer a new dimension to maintain the soundness of our society by assessing AI through power laws in order to help people working in both economic and other fields.

*Basic studies have been developed since 2014 by the PRESTO team of JST Strategic Basic Research Programs led to this project.

# **PROJECTS** OF HUMAN-INFORMATION TECHNOLOGY ECOSYSTEM

## 01

### Which controls which? Sense of agency when humans and semi-automated systems co-operate

When a semi-automated system, including artificial intelligence (AI) or a mechanical system, co-operates with a human and commits a socially undesirable action such as an incident or accident, which idea should be adopted: should the human user be responsible as the subject for the action, or should the machine, system, or manufacturer be responsible as the subject? What are the scientific definitions of action and subject? What does it mean to be a social agent or subject who has free will to act and hold responsibility for his/her actions? We propose a scientific definition and answer for these questions from the perspective of psychology and neuroscience.

In particular, we identify that there is an illusion in which human users misattribute their actions as their own instead of belonging to other agents, which can lead to unnecessary feelings of responsibility for an action. Through the brain science based on the understanding of this illusion and the development of its application in human-machine interactions, we propose the idea that most of the human sensation of "agency" (i.e., that oneself caused an action) is based on postdicted subjective sensation or illusion. In this sense, humans cannot be fully autonomous.

**Takako Yoshida**
Associate Professor, School of Engineering,
Tokyo Institute of Technology
Psychology

## 02

### Consideration on the concept of "responsibility" between autonomous machines and citizenries

While we are expecting that autonomous machines mounting artificial intelligence would make our life convenient, we may be anxious because we hardly identify the one(s) who should take responsibility when autonomous machines cause serious accidents and harm us.

In order to prepare to the coming age of autonomous machines it is necessary for us to design, our project believes, convincing concept of "responsibility", which not only explains social position of the machines, but also has ability to persuade citizens as non-specialists. Therefore, the concept which we seek must be the one which reflects the concept of responsibility which human beings have developed in the history and various culture. Our project, being based on current condition of society and technology, will propose the concept of "responsibility" by considering "what is necessary condition for autonomous machines to be recognized as human being in a society"; this question inheres more fundamental and philosophical question, namely, "what is, or who has, human nature". The answer to the question will provide us appropriate relationship between autonomous machines and citizenries in a society.

**Kazuya Matsuura**
Associate Professor, Faculty of Letters,
Toyo University
Philosophy

## 03

### Legal being: electronic personhoods of artificial intelligence and robots in NAJIMI society, based on a reconsideration of the concept of autonomy

The AI technology and its development in recent years have been enabling a certain kind of autonomy of artificial systems and robots. Just like a child who tries to be independent from its parents, or, because of the complexity of their calculations they perform, they may behave beyond the expectations of their designer. As a result, they appear to be given a certain autonomy. However, the designer or the user may be subjected to an undue legal responsibility under current legal systems, that may interfere with desirable progress of science and technology.

In our project, based on the extent of their ability to set their object, to rewrite their programs by themselves, or the level of complexity of their calculations, we assume the three stages of the concept of autonomy of artificial systems. Then we figure out the legal devise applicable to the three stages, through the analysis of the traditional legal personality doctrines and of the history of the concept, persona. Moreover, we demonstrate the major legal and social problems provoked under current legal responsibility theories, especially in criminal law field, and then we propose a new legal responsibility theory and an institution for artificial systems. Also, we propose the legal provisions that realize the "NAJIMI society" through mock trials where using android in order to deepen the concept of autonomy and propose the preferable artificial systems in the future society and environment.

**Minoru Asada**
Professor, Graduate School of Engineering,
Osaka University
Robotics

## 04

### Co-Creation and Communication for Real-Time Technology Assessment (CoRTTA) on Information Technology and Molecular Robotics

In our project, by extracting agendas of discussion by media analysis and by the predictive estimation method (horizon scanning) regarding the Ethical, Legal and Social Issues (ELSI) of high-tech information technology and molecular robotics. Furthermore, through development of a "subject of discussion co-creation platform (NutShell)," in which high-tech information technology specialists as well as various stake holders participate, we formulate a "Real-Time Technology Assessment (RTTA)" system to swiftly focus social discussions in the relevant field.

Regarding such formulation of subjects of discussion using the RTTA system, by extracting the subjects of discussion from actual practice via case examples in the molecular robotics and artificial intelligence fields, we propose realization of a better agenda building process concerning ELSI and how feedback of the knowledge to the researchers at field sites should be channeled.

**Ryuuma Shineha**
Associate Professor,
Seijo University
Science, Technology and Society

## 05

### Co-creation of Molecular Robot ELSI and Real-time Technology Assessment Research

We aim at realizing a society in which molecular robot technology and human beings are familiar with one another by spirally promoting the study (Konagaya G) of molecular robot's Ethical, Legal and Social Issues (ELSI) and the real-time technical assessment study (Shineha G) to gather a wide range of knowledge/opinions about both technology and society utilizing the Internet.

In our project, co-creating with the Shineha G, we proceed with the molecular robot ELSI study by repeating the process of: (1) aggregation of a wide range of opinions from society utilizing real-time technical assessment technology; (2) formulation of a draft molecular robot guideline reflecting such opinions and (3) discussion at study meetings/symposia concerning the formulated draft. Also, we support in the molecular robot international student contest from the viewpoint of ELSI in order to further young molecular robot researchers' and students' understanding of molecular robot ELSI.

**Akihiko Konagaya**
Professor, School of Computing,
Tokyo Institute of Technology
Intelligent Informatics

## 06

### A coevolutionary study on society with respect to power laws: - Can AI replicate the behavior of a non-equilibrium complex system?

We study a new methodology for applying statistical analysis to assess an AI's performance and adaptability to society. Human activities such as natural language and financial markets have statistical properties characterized by power laws. Using these properties, we focus on the differences in behaviors and outcomes between AIs and humans to examine the societal impacts, capacities, and limitations of AIs.

Achieving these goals involves two themes. The first theme investigates quantitative analysis with respect to power laws on models of human activities such as natural language and financial markets. We seek to clarify the divergences between pseudo-data generated by AIs and real data on language and investment. The second theme applies insights from the first theme to the investment field. We suggest improvements in social design and legal systems in the AI era. Because the first theme involves a technical study and the second applies the resulting understanding to society, the two themes constitute an example of a coevolutional research platform.

**Kumiko TANAKA-ISHII**
Professor, Research Center for Advanced Science and Technology,
The University of Tokyo
Communication Science

**PROJECTS** OF
HUMAN-INFORMATION
TECHNOLOGY
ECOSYSTEM

## 07

### Preliminary research on methodology for improving the information access literacy from the viewpoint of “benefits of inconvenience”

This project aims to enhance “information access literacy”, which is defined as the ability to critically make decisions by searching for appropriate information through information access systems. Information access technologies designed only for high efficiency and convenience can bring potential danger of causing low cognitive capacity of humans. Thus, we pay attention to “benefits of inconvenience,” which is obtainable only by users’ active contributions to tasks that require mental exercises.

We conduct a survey on medical professionals, who are expected to use decision-making support systems more frequently in the future, for maintaining and enhancing the information access literacy based on the benefits of inconvenience. More specifically, our project conducts (1) in-depth analysis on the problems and needs concerning the information access literacy in a realistic medical setting through hearing-based and questionnaire-based surveys; (2) development of a literacy diagnosis tool; and (3) a study on the inconvenience acceptability through workshops with medical professionals.

**Hiroshi Kawakami**
Professor, Unit of Design,
Kyoto University
System Engineering

## 08

### Research on the task models to cooperate with the human and AI systems

In our project survey, by organically connecting the knowledge in 3 fields of labor economics, service management and intelligence information science, we conduct an analysis to qualitatively/quantitatively access the kinds and quantities of the tasks that workers and new information technologies engage. We conduct interviews about case examples of preceding introduction/experiment of new information technology and questionnaire surveys through the Internet. We then elaborate methods to drive responsive measures or desirable directionality in business and policy by exploring the key points to be heeded in the cooperative/competitive relationship between information technology and human/society.

Through such analysis, we proceed with preparation for implementing a study based on two analyses: the case studies on the fields introducing/experimenting new information technology and the quantitative survey/analysis on the consumers and workers.

**Isamu Yamamoto**
Professor, Faculty of Business and Commerce,
Keio University
Labour Economics

## 09

### Aritificial intellegence and human labor: Substitutability and complimetarity

It has been a long time since concern was first expressed that development of artificial intelligence would deprive us of our work. For example, a report published by a consulting company at the end of 2015 states that 49% of Japanese employment can be replaced by machines. These studies assumes the substitutable/complementary relationship between labor and artificial intelligence technology along the axis of the traditional occupation database's labor characteristics; thus the current framework is not such that the substitutable/complementary relationship with labor is measured upon grasping the substance of artificial intelligence technology.

In our research and development project, conceptualizing the fundamental causes of determining the alternative/complementary relationship between machines, as represented by artificial intelligence, and labor; through the conceptualization, we further develop a method to grasp them by survey question items. We speculate that the key concept in that case lies in whether large-scale electronic data exists or not and in the difficulty of grasping the cause-effect relationship. Implementing a large-scale survey upon sophisticating this concept through interviews with scientists/engineers, we develop a new occupation database beyond the existing occupation database.

**Daiji Kawaguchi**
Professor, Graduate School of Economics,
University of Tokyo
Labour Economics

## 10

### Development of behavior changing library for implementation of AI-based precision monitoring technology

Monitoring technology being upgraded thanks to advancement of IoT technology as well as AI technology, it is expected to solve problems in communities where supervision is deemed necessary, e.g. day-care centers for children and nursing homes for the elderly. In communities, however, there being, along with monitoring technology's reliability, the implementation of its technologies is actually not progressing. For the monitoring technology developers, the community's needs are hard to understand and what is acceptable for a community is unknown, so actually they don't venture on the development.

In our project, we take the initiative to clarify the range of acceptance or the acceptable state for monitoring technology at communities by understanding for the problems and the needs for monitoring. We aim at developing a new methodology for implementation of new technologies and accelerating the social implementation of monitoring technology.

**Koji Kitamura**
Senior Researcher, Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology
Computer Science





# ADVISOR

Program Supervisor and Advisors
(as of April 1, 2018)

| | | |
|---|---|---|
| Program Supervisor : | **Jiro Kokuryo** | Professor, the Faculty of Policy Management, Keio University |
| Deputy Program Supervisor : | **Hideaki Shiroyama** | Professor, Graduate Schools for Law and Politics, The University of Tokyo |
| Advisor : | **Kazuhiko Kato** | Provost and Executive Office, Professor, Faculty of Engineering and Systems, University of Tsukuba |
| Advisor : | **Koichi Kume** | Associate professor, Department of Policy Studies, Faculty of Economics, Toyo University |
| Advisor : | **Yasuko Kohno** | Director, Japan Consumers' Association |
| Advisor : | **Kaoru Sunada** | Executive Research Fellow, Center for Global Communications, International University of Japan |
| Advisor : | **Norihisa Doi** | Professor Emeritus, Keio University |
| Advisor : | **Toru Nishigaki** | Professor, Faculty of Communication Studies, Tokyo Keizai University |
| Advisor : | **Yukihiro Nobuhara** | Professor, Graduate Schools of Arts and Sciences, The University of Tokyo |
| Advisor : | **Hitoshi Matsubara** | Vice-President, Future University Hakodate |
| Advisor : | **Tsuyoshi Maruyama** | Specially Appointed Professor, Faculty of Science and Engineering, Chuo University |
| Advisor : | **Fumihiro Murakami** | Research Director, ICT and Media Strategy Group, Social ICT Innovation Division, Consulting Unit, Mitsubishi Research Institute, Inc. |
| Advisor : | **Yuko Murakami** | Specially Appointed Professor, College of Science, Rikkyo University |

**Contact details for inquiries**

Human-Information Technology Ecoeyetem
Japan Science and Technology Agency (JST)
Research Institute of Science and Technology for Society (RISTEX)
Science Plaza 4F 5-3, Yonbancho, Chiyoda-ku, Tokyo 102-8666, Japan
TEL: +81 (0)3-5214-0133
Email: info-ecosystem@jst.go.jp



Human-Information Technology Ecosystem Vol.02

| | |
|---|---|
| Management: | **Akiko Kaya, Kenichi Haga, Hitomi Uehara** (JST) |
| Creative Director, Editor: | **Arina Tsukada** |
| Art Directer, Designer: | **Hiroka Hasegawa** (H Inc.) |
| Writer ( P.3-6,11-12 ): | **Mirei Takahashi** |
| Writer ( P.7-10 ): | **Akihico Mori** |
| Photographer: | **Rakutaro Ogiwara** |
| Illustrator: | **Yoshifumi Takeda** |
| Texts of cover visual*: | **Toru Urakawa** (Qosmo) |
| Published by | **Japan Science and Technology Agency** (JST) |
| | **Research Institute of Science and Technology for Society** (RISTEX) |

Date of publication: September 1st, 2018

The cover of this magazine was designed using the analytical text data related to HITE program which the language analysis program "a dictionary of the consciousness" that programmer Toru Urakawa developed analyzed.

Japan Science and Technology Agency

RISTEX
Research Institute of Science and Technology for Society