NII Today

National Institute of Informatics News

Towards Creating Data-Centric Human and Social Sciences

The Transdisciplinary Research Integration Project kicked off in April 2010. Its objective is to create data-centric human and social sciences. The National Institute of Informatics (NII) and the Institute of Statistical Mathematics (ISM) joined hands in a bid to develop a technique for actively using a tremendous volume of data on the Web space and thereby to supporting the policymaking process for designing social system with resilience capabilities in the face of damage caused man-induced or other faults, attacks, human errors, and natural disasters. It is indeed an ambitious goal.

Takahashi I have heard that the NII and the ISM became part of the Research Organization of Information and Systems (ROIS) in 2004, together with the National Institute of Polar Research (NIPR) and the National Institute of Genetics (NIG). ROIS is an inter-university research institute corporation.

Sonehara Yes. Since then, ROIS has had three research pillars: global environment studies, life sciences and information infrastructure. Now, it has a fourth pillar—human and social sciences—in view of the recent expansion of the internet and the web. There is a tremendous volume of data in cyberspace, including the web space and SNS services. Originally, data were omnipresent, irrespective of whether it was in analog or digital form. Today, sensing of these data is possible. We think it is necessary to establish a spiral in which data are effectively analyzed to create models to be applied to different parts of physical society facing challenges as a means of improving the lives of people. We call this information flow infrastructure (Figure).

Tracking people's movements on a real time basis

Takahashi Specifically, what research is conducted in the project?

Sonehara We collect hotel and transport service booking data from the web. These data taken from flight, train, and hotel booking websites make clear on a real time basis how people are moving every day.

Takahashi Some hotels do not use the web, do they?

Sonehara Very few hotels are not using the web today. At first, we asked some personnel in the tourism department of the City of Kyoto Government to offer some data. They showed us that even the smallest hotel was using online booking sites. Of course, they accept phone booking as well. When a hotel is fully booked, it calls the booking website operator to report its no-vacancy status. In some seasons, including the Obon vacation season, accurate data cannot be obtained, but even so, the margin of error is 9% or less. There is little disparity between the fact-finding survey conducted by a public statistical body and our web analysis data. It takes three to four months until the statistical results are announced, while web analysis data is obtained instantly.

Takahashi I see. Data can be collected instantly without little effort. What are the data used for?

Sonehara Basically, they are useful in demand control. From the standpoint of hotels, they are helpful in pricing and customer attraction.

Takahashi But it seems already possible to grasp market demand without such data. Actually, accommodation charges are higher in the New-Year season and in the Obon vacation season.

Sonehara These scientific supporting data are beneficial not only for pricing but also for planning marathons, international conferences, and other events in slack seasons and to studies on where a new airport or a tourist attraction should be constructed.

**Takahashi** I see. I agree that policy decisions should be based on evidence.

In hope of institutional reforms

Sonehara The data are also helpful in the event of disasters. Suppose, for example, that a major earthquake occurs in the Kanto region. At present, we have no means of learning how many people there are in Shibuya or in Haneda. Actually, we could obtain these figures with considerable precision from the data on Pasmo and Suica smartcards and on mobile

Noboru Sonehara

Professor, Information and
Society Research Division, NII
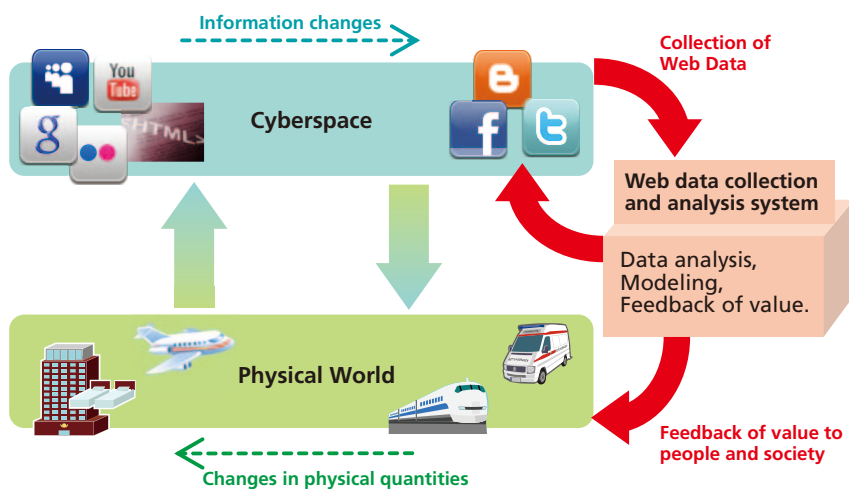
A Web Data-driven Information Flow System
Diverse data obtained and collected in large quantities from cyberspace are introduced to the real world as information with new value. The new information generated from the physical world is reflected in cyberspace. The information will thus circulate between cyberspace and the physical world. This circulation will help build a better and more efficient society.

phones, but our society is not designed to make this possible. Under the Telecommunications Business Act, telecom carriers must protect the secrecy of communications. That hinders positional information from being easily used for policymaking designed to confirm safety or provide immediate relief after a disaster, even if mobile phone carriers hold that information.

Takahashi The positional information of mobile phones involves so much privacy that anyone would feel concerned if it were known to the public. I would imagine there will be no problem if the number of people alone is disclosed in a way in which no personal information is revealed.

Sonehara It is a difficult problem. The recent Great East Japan Earthquake has revealed that it is impossible to readily make full use of any information system in the event of emergency until we are accustomed to operating it on a daily basis. Take the online safety confirmation system for example. It was designed to be unavailable unless you scroll down and press the agree button at the very bottom of the screen. I believe that a crisis-resistant infrastructure that enables instant use of personal information is essential for emergency policymaking.

Takahashi The citizen could possibly agree to conditional information disclosure confined to disaster situations.

Sonehara Yes. Taxis today are equipped with video cameras and sensors in order to record footage of the seconds before and after an accident. Something similar method should be done. At ordinary times, the positional information of mobile phones should be confidential. In the event of an earthquake, carriers should offer information on their whereabouts just before the quake. In this case, it has to be done by all mobile phone carriers. Train operators should also offer data. Then, a nationwide map will be instantly created. Presuming that this will be acceptable, we are now working to design a social system resistant to crises.

Takahashi We already have technologies that make this possible. If it is the Telecommunications Business Act and other institutions that hamper it, the problem lies rather in the institutions. So what technologies can do is limited, isn't it?

Sonehara Not really. Technologies help us discover new value in data. Consequently, they have an influence on institutions. The institutions won't change if the existence of that value is not demonstrated.

Takahashi So, the goal of the project is to uncover the value.

Sonehara That's right. If it is possible to instantly grasp who is where in the event of a disaster or other emergency, it is helpful to evacuation, rescue, and other relief efforts. There must be something that can be done in Japan, the world leader in mobile, SNS, and other ICT technologies. I think this is one of the tasks assigned to informatics. The privacy issue should not be dealt with at the level at which it would somewhat uncomfortable if personal information went public. We aspire to build a society where decisions can be made on the basis of scientific data.

Today, a very high compliance cost is incurred in the course of developing any new information service. It is fairly difficult to achieve consensus, but I think that the future policy support will be to address the challenge on the basis of data.

Comment from the Interviewer

There are many things he wants to do, but what he can do in reality is limited and that irritates him. I vividly feel his frustration from his words. Anyone would want a social system resistant to crises, but it would not come into being without collaboration among private businesses, national and local governments, politicians, and many others. I understand that this project is destined to go beyond the framework of conventional research projects.

Mariko Takahashi
Senior Staff Writer, Asahi Shimbun

Web Data-Driven Information Flow

Capturing human/social activities through massive web data

Psychological Reluctance to Personal Information Disclosure and Anti-Leak Technology

Today, so many people want their personal information undisclosed that it is necessary to ensure its anonymity for preventing personal identification. However, excessive anonymization leads to a deterioration of information quality and impedes effective use of information. How is it possible to bolster active use of information while protecting it?

A dilemma between anonymization and use of statistical data

What studies are you each undertaking in the domain of human and social sciences as part of the Transdisciplinary Research Integration Project?

Echizen My research seeks to make academic use of statistical data, including personal survey data that were collected with the assurance that they would be used only within the university or research institution, beyond organizational boundaries while maintaining anonymity to a certain degree.

For example, if medical data, which typically includes the patient's name, address, age, disease, and medication, were made publically available without change, the patient could be identified. To prevent this, it is necessary to blur the patient's attributes by deleting the name and address and by generalizing some of the information; for example, "Tokyo" could be generalized to "Japan," and "age 32" could be generalized to "thirties." Since more than one person would usually have the same general attribute, individuals could not be identified.

However, this generalization approach to ensuring anonymity impairs the value and accuracy of data. In other words, there is a trade-off between the level of data anonymity and the academic utility of the data. In this environment, the new Statistics Act came into force in Japan in 2009 to pave the way for academic use of questionnaire information. This means that higher priority was placed on academic utility than on data anonymity. Data anonymity has traditionally been emphasized on the assumption that anonymized data could be made freely available. Nowadays, the degree of anonymization has been lowered to increase data utility while more emphasis has been placed on measures to prevent data leaks.

Kobayashi In fact, many past statistical data published for secondary analysis were hard to use. For instance, they had address information merely at the prefectural level and hampered comparative analyses between urban and rural areas. Among the data governed by the Statistics Act, social survey data, such as the information that I use, are mostly personally unidentifiable. I do not think it is necessary to further anonymize individual data, unlike medical information. In fact, we recode any observations with extremely rare frequency as missing values. This is called top-coding. A considerable number of methodologies for boosting anonymity have already been established.

So, my research focuses on how to encourage people to offer their personal data. Among other things, life logs on human behaviors and communications are now gathering the interest of those in academic and other circles. However, people are so worried about privacy that many are reluctant to offer them. I am exploring what constitutes their psychological reluctance and how to persuade them to offer socially valuable data.
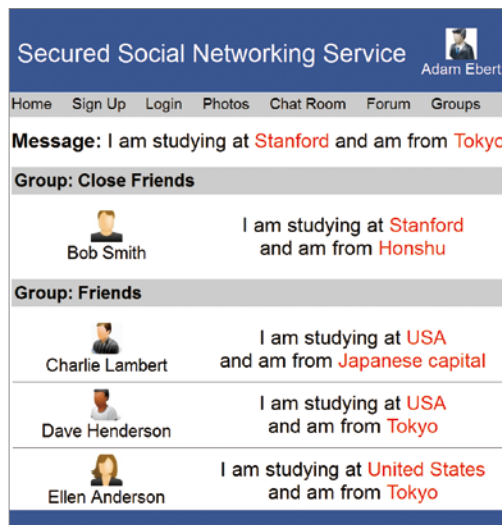
Isao Echizen

Associate Professor, Digital Content and Media Sciences Research Division, NII

Tetsuro Kobayashi

Associate Professor, Information and Society Research Division, NII

Introduction of fingerprinting to social networking services (SNS)
Messages from users are anonymized at different levels of anonymity, depending on their group. (In this example, their organizations and places of origin are anonymized.) In the group, different messages are generated at the same level of anonymity. After the individual anonymization processes are associated with user identification data, it is possible to identify the source of data leak from the message leaked.

Data leak prevention technologies and ideas for lowering psychological reluctance

Please give more details about your respective studies.

Echizen I have developed a method for identifying the source of leaked anonymized data that associates individual anonymization processes with user identification data. It is called "fingerprinting of anonymized data."(Figure) This approach capitalizes on the multiplicity of data anonymization processes, which means that there are many different processes for achieving the same degree of anonymization. Suppose that there are data consisting solely of birth date and gender. User A is provided with a data file containing "1971" and "male" while User B is provided with one containing "August 10, 1971" and "gender unknown." Our method prepares, for each user, a set of data generated by an anonymization process that varies with the user that has the same level of anonymization as every other prepared set. In the event of data leakage, the association between the anonymization process and the user ID is used to help identify the person responsible. Moreover, awareness of this identification method among users should make them more careful about data management. That is, the anonymization processes themselves deter data leakage.

Application of this method to social networking services (SNSs) and blogs would enable the source of a privacy leak to be identified from an analysis of the text containing the leaked information. In this application, not only would the anonymization process used vary with the user, but the degree of anonymization would vary with the group.

Kobayashi I conducted an experiment in which smartphone users were asked to disclose their life logs from their smartphones. The life logs were specifically classified into three categories: (1) locative information, (2) web browsing history; and (3) voice, SNS, and Gmail communications histories. In this event, three different levels of compensation were set for separate categories: 1,000 yen for category (1), 5,000 yen for (2) and 10,000 yen for (3). I introduced the conditions that combined the categories of life logs offered and compensation levels (i.e. 2*2*2*3=24 conditions) to see the effect of each factor on the life-log disclosure. As expected, the experiment confirmed that the lifelog disclosure was facilitated with the monetary compensation. On the other hand, it also found out that nearly 30% of the subjects disagreed to make any disclose irrespective of the level of compensation. Many showed strong reluctance to offer information about voice and other communications histories even if the content of communications was not to be recorded. On the other hand, their level of reluctance was lower with respect to GPS data. The subject may not have been very aware of the risk involved in offering locative data. Therefore, I think that it may be possible to gradually widen the scope of available data from which induces relatively minor reluctance, like GPS data, to remove the initial psychological barriers, and then move on to the types of information that people feel more reluctance to offer.

There is several major hurdles to the collection of personal information and it must never be leaked even after collection. Meanwhile, if everyone offers it its overall quality will improve and its utility as "public goods" will increase as well. So, it will be necessary to conduct further studies on the design of the incentive to facilitate information disclosure.

Echizen I agree. It is tough to deal with two mutually conflicting factors, namely data anonymity and data utility. We will continue our research to create a good technique for balancing them to contribute to effective use of information.

(Written by Madoka Tainaka)

Web Data-Driven Information Flow

Working Out Tourism Policies from Web Data

In recent years, the Japanese government has been emphasizing tourism as a promising industry. The NII embarked on research in this area, presuming that data-centered sciences would play a significant role in the revitalization of the tourism industry. With a view towards making effective use of Web Data, Dr. Yu Ichifuji, a project researcher for Transdiciplinary Research Integration Center, was interviewed on the current status and future prospects of the research.

Working to develop tourism into a leading industry of the country

In 2003, the Japanese government declared that Japan would be a tourism nation. Since then, powerful measures for revitalizing tourism, now defined as a key industry, have been adopted.

In 2008, the Japan Tourism Agency was created. It publishes a multi-day travel survey report every three months to help local governments develop their tourism policies. In response to this national government-level initiative, local governments are also making a huge effort. For example, when a major event is being held, they commission research firms to evaluate its economic effect in order to use the data for future policies.

In these circumstances, Dr. Ichifuji casts doubt on effectiveness of the tourism revitalization efforts. He sees both the multi-day travel survey report and the evaluation of economic effect as excessively time-and cost-consuming. He is a member of a group led by Professor Noboru Sonehara exploring the effective use of massive volumes of data in the cyber world. He has been studying the use of web data for tourism since 2009.

The hodgepodge nature of web data

The primary indicator of prosperity in the tourism industry is the number of visitors. Dr. Ichifuji first paid attention to hotel booking websites as a source of web data on visitor headcounts.

Nowadays, many hotels and other lodging facilities are registered on booking websites so that users can check their occupancy statuses at any time. They also provide a broad array of information, such as details about charges for individual lodging options. Dr. Ichifuji started by researching and studying accommodation facilities in Kyoto.

"To use web data, the data must be reliable. It is a prerequisite," he says. Booking websites have officially uploaded data with relatively higher reliability among the web data, which are said to be a mixture of valuable and unwanted information. However, hotels offer information on their availability to more than one such website in an attempt to attract more guests while taking care to avoid double-booking. This means that visitor headcounts may not be obtained by watching just one website and that it is necessary to collect data from multiple booking websites to understand the tourist situation. To accurately estimate the number of visitors from the data obtained, it is necessary that all lodging facilities in the city of Kyoto be registered with booking websites and that the uploaded booking statuses reflect the actual number of guests.

To assess this, Dr. Ichifuji compared the hotel booking data collected on the web with the information on lodging facilities officially collected by the Japan Tourism Agency.

He found that according to the web data, there were approximately 20,000 rooms at 196 lodging facilities with at least 10 employees in the city of Kyoto, whereas 191 lodging facilities and nearly 20,000 rooms were mentioned in the multi-day travel survey report in 2011.
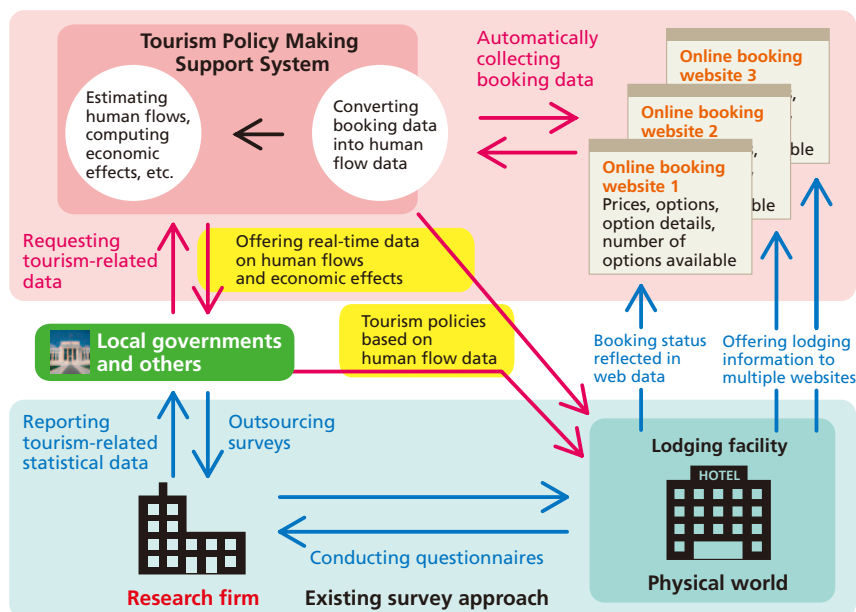
In addition, Dr. Ichifuji estimated the number of vacant rooms from the number of available options and computed the number of lodgers in each month to make a comparison with the survey report. In every month, there was a disparity of only around 9% at most between the online booking data and the survey report data.

Thus, the precondition that it is possible to estimate the number of visitors to the city of Kyoto from booking data for lodging facilities in the city has been estimated. Now, a study on specific uses of booking data is underway (Figure).

Yu Ichifuji

Project Researcher
Transdisciplinary Research Integration Center, NII

Tourism Policy Making Support System

Estimating human flows, computing economic effects, etc.

Converting booking data into human flow data

Automatically collecting booking data

Online booking website 3

Online booking website 2

Online booking website 1
Prices, options, option details, number of options available

Requesting tourism-related data

Offering real-time data on human flows and economic effects

Tourism policies based on human flow data

Local governments and others

Booking status reflected in web data

Offering lodging information to multiple websites

Reporting tourism-related statistical data

Outsourcing surveys

Lodging facility

HOTEL

Physical world

Conducting questionnaires

Research firm   **Existing survey approach**

Overview of the Tourism Policy Making Support System

The Tourism Policy Making Support System incorporates a computer to automatically collect tourism-related data from the web and to convert them into visitor headcounts and other data for tourism. It has paved the way for real-time collection and analysis of tourism data at low cost.

Offering data that are useful now

What is made possible based on the forecast of visitor headcounts from web data? Providing support for the research, the City of Kyoto Government says that it wishes to use web data as a supplement to the economic effect evaluation report and other information submitted from the research firm.

On the other hand, "It is a computer program that collects data on the web. It is a faster and less expensive means of data collection than a research firm deploying survey personnel," says Dr. Ichifuji, explaining the advantage of online data. Among other things, the point that data can be collected quickly is significant. According to Dr. Ichifuji, it may give impetus to the revitalization of the tourism industry. For instance, the visitor attraction status can be learned just before an event takes place. If it is found to be poor, it is possible to take action such as placing ads.

While conventional survey reports provide historical information, online data reflect the present state and pave the way for real-time control. To stimulate the tourism industry effectively, these data will facilitate the organiza-

tion of events in off seasons.

Dr. Ichifuji started studying the data on the number of passengers in Shinkansen bullet trains. By analyzing them in combination with the hotel booking data, he can forecast the movements of tourists in more detail.

He continues, "If we could use SNS posts from tourist destinations, they could help planning a highly appealing tour. At the moment, they cannot be used as it is difficult to evaluate their credibility." He considers combining diverse web data to offer various services to travel agencies, hotels and lodging facilities, and individuals in the future.

Information that is effective at the time of emergency

"We are seeking to use the data collected this way at the time of disasters," he says. In the event of the Great East Japan Earthquake in March 2011, many emergency systems did not function. He explains the reasons behind the failure by saying, "A university sent a mail to students to find out about their safety, but they saw it as spam. In a chaotic situation, anyone will misunderstand a safety confirmation message from an

unknown email address."

Dr. Ichifuji concludes that it is important to ensure that the system in everyday use is operational in the event of emergency as well, to prevent a recurrence of the failure.

With this thinking, he studied what could be done at the time of emergency with the use of online hotel booking data in the city of Sendai. Graphical presentation of the data on the hotel reopening dates demonstrated that it was possible to visualize the progress of reconstruction in the affected area. Resumption of hotel operation implies that electricity, gas, and water utility services have been restored and that the supply of bed linens and other equipment has been restarted. More facts than expected can be seen from these data.

Despite the hodgepodge mixture, it is now evident that web data can be used to invigorate the tourism industry and for understanding the circumstances at the time of emergency, provided that they are carefully sorted and assessed. Future suggestions for tourism policies may originate from online information. Web data will be of greater importance than ever before.

(Written by Akiko Ikeda)

Product of Integration between Informatics and Statistics

The Research Organization of Information and Systems established the Transdisciplinary Research Integration Center in 2005 to push ahead with transdisciplinary studies based on diverse and massive data acquired from the NII, the Institute of Statistical Mathematics (ISM), the National Institute of Genetics (NIG) and the National Institute of Polar Research (NIPR). For the second phase that commenced in 2010, it has added a new discipline of human and social sciences to the conventional ones, namely life, the global environment, and information infrastructure. Transdisciplinary research conducted by the NII and the ISM is in progress.

What is behind the necessity to integrate informatics and statistics?

Why is transdisciplinary research between informatics and statistics necessary now?

Sonehara In the past, most of the data relating to humans and society were static data collected offline, such as those obtained by the Population Census every five years, manned surveys and questionnaire surveys. We are now working to encourage the active use of dynamic data obtained from the internet and mobile terminals, including the web and social networking services (SNS). Web data lack reliability but we believe we can collect and analyze information closer to reality by calibrating it and by integrating ever-changing dynamic data and offline data (Figure).

Tsubaki As you mentioned, principal statistics taken by the state government are stationary measurement and do not reflect dynamic shifts. On the other hand, since the Act on the Protection of Personal Information came into force in 2005, it has been difficult to secure cooperation in surveys. To supplement the traditional stationary measurement under these circumstances, we are stepping up the use of real-time and dynamic web data.

Moves towards actively using web data are already emerging in the business sector. For example, Komatsu, a construction machinery manufacturer, monitors its machines at remote locations originally for maintenance purposes to view how they work, but these data suggest where economic activities are brisk. In other words, it is possible to acquire information close to principal survey data sought by the state government from information in the cyberspace.

Sonehara However, the data from cyberspace and the principal data differ in quality. Our research aims to contribute to the areas of safety and reassurance, disaster control, and the environment by comparing and transforming them into more valuable information.

## Need to utilize scientific supportive data in policymaking

How are the principal data used?

Sonehara Essentially, the purpose of the principal data is for use in policymaking. These days, data are also being used in education and research. Located in Tachikawa, the Onsite Data Analysis

Hiroe Tsubaki

Vice Director-General,
The Institute of Statistical Mathematics (ISM)

Noboru Sonehara

Professor, Information and Society Research Division, NII
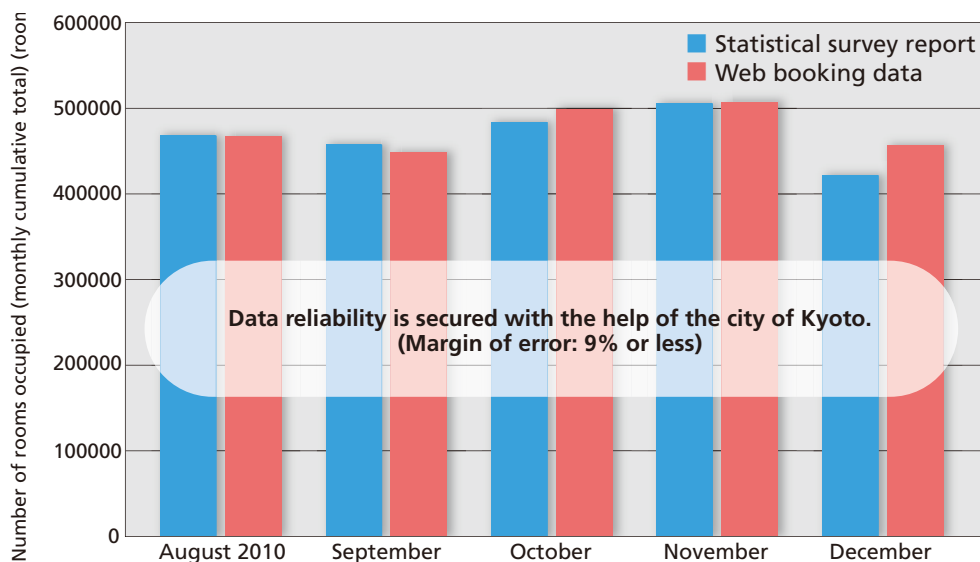
Comparison between estimates from online booking data and statistical survey data

This comparison has clarified that the number of hotel rooms occupied based on web data agrees with that stated in the multi-day travel survey report. This suggests that web data provides cost-effective support for the tourism policymaking process.

Base of the Transdisciplinary Research Integration Center now grants access to principal data for academic use. However, networking is indispensable and I suspect that these data would be better used in the business sector.

Tsubaki I agree. Following the introduction of the Act on the Protection of Personal Information, access to the Residential Basic Book has been denied and it is very difficult to conduct marketing surveys in the private sector. Certainly, some restrictions are needed, but if the principal data are shared in a form with higher utility and if the private sector has modeling and analyzing technologies, which will add a major boost to Japan's competitiveness.

Sonehara There is a fundamental question as to who owns the principal data. Essentially, members of the public have an obligation to cooperate in the creation of key statistics. The collected information should be returned to the public at large. It is very wasteful to leave unused the data collected for academic purposes.

Tsubaki Today, information is among the most important aspects of national politics. Given this, some information is nearly as valuable as tax and it should be regarded as an asset shared by the public. If it were to be concealed by an individual or business, the government could not do anything. Meanwhile, the government has the obligation to properly use the data collected. The information should be appropriately incorporated into national politics and the analyzed information should be returned to the private sector.

Sonehara In today's society, a trivial matter can result in a loss of opportunity. I strongly believe that policymaking and institutional design based on scientifically supportive data will be essential.

Data-centered sciences as a hub that links different disciplines

Specifically what will data integration enable?

Sonehara We are now engaged in the development of systems for real-time tourism policy support based on web data and for comprehensive access to industrial environment data on the web. We are also working to create a system that communicates patients' images and positional information from ambulances to medical facilities to speed up emergency transportation. This effort will soon produce results.

Tsubaki However, there are some problems. For example, suppose that statistical data show that severe adverse effects are seen from patients taking two different drugs. However, data anonymity is an impediment to directly contacting these patients. Some techniques will be needed to prevent the deterioration of data quality attributable to anonymization.

How to actively use the results of analysis is another issue. For instance, Keio University's research on suicide shows that the suicide rate is high in areas with slopes while it is low in coastal regions. However, reconstruction after the Great East Japan Earthquake is placing the focus on the relocation of communities to higher ground. It will be necessary to build a structure in which a certain organization undertakes overall monitoring of information and puts it to active use.

Sonehara As we are experts in informatics and statistics, it is difficult for us to deal with institutional design. In this sense, transdisciplinary integration involving a broader array of disciplines, not confined to informatics and statistics but including law, politics, education, and many other elements, will be required. And I think data-centered sciences will be at the heart of the integration. Data are indispensable to any domain and they exist irrespective of domains.

Tsubaki As specialists in statistics and informatics, we more than anyone else may have to play the role of a hub that links different disciplines.

Sonehara I think it is difficult to expect this role from universities. Inter-University Research Institute Corporation should play this rule.

To fulfill our duties, we will continue to forge ahead with the transdisciplinary research.

(Written by Madoka Tainaka)

Problems with ICT and Legal System

It is still fresh in our memory that social networking services (SNS) played significant roles at the time of the Great East Japan Earthquake. Progress in information and communication technology (ICT) has enabled anyone to freely send out information all over the world. On the other hand, it has given rise to serious problems such as privacy invasion and personal information leakage. At Graduate Schools of Law and Politics, the University of Tokyo, Associate Professor George Shishido explains the current issues of ICT and measures from the perspective of a legal expert.

Problems uncovered by the massive earthquake

In the Great East Japan Earthquake, safety confirmation and instructions for emergency evacuation were delayed because voice and text communications using mobile phones were controlled. In contrast, Twitter, Facebook, and other SNSs were found to be very useful as means of sending out information.

"It is a good example of a case in which freedom of information is protected by the Telecommunications Business

George Shishido

Associate Professor
Graduate Schools for Law and Politics,
The University of Tokyo

Act even under emergency circumstances such as the disastrous earthquake," says Associate Professor Shishido. This Act obliges telecom carriers and internet service providers (ISPs) to maintain communication confidentiality and also prohibits them to infringe on users' freedom of transmitting and exchanging information (Figure).

These operators are technically capable of storing and managing information exchange on the internet, but the legislation prohibits them from acquiring such information, using it at their convenience, or controlling it at their discretion by abusing their special positions.

In addition, personal information on networks is also strictly protected under the Act on the Protection of Personal Information. All personal information is generally obtained with the consent of the individual concerned. Here "personal information" refers to all information from which any individual is identifiable, including names, dates of birth, and addresses.

While the Great East Japan Earthquake reminded us how important roles ICT plays, some began to insist that the utilization of personal information via ICT promotes the public interest especially in the case of emergencies.

For example, in case we want to know accurately who live in the affected area and where they are at the moment, it would be technically possible to track such information on mobile terminals

equipped with Global Positioning System (GPS) functions which disaster victims possess.

And when offering medical treatment to victims at evacuation centers, the rescue team could identify drugs that they usually take and those that must not be administered by accessing their medical records and chronic diseases via ICT.

However, under the current legislation, such information may not be acquired or used by third parties, including telecom carriers and ISPs, in principle, without the consent of the individual concerned because it is deemed as personal information. This is why some start to argue we need to reconsider the rules on handling personal information in emergencies.

The need for new rule-making

In fact, the Act on the Protection of Personal Information stipulates in item (ii) of paragraph (3) of Article 16 that personal information may be obtained without the consent of the individual concerned provided that his or her consent is difficult to obtain and besides that the life, body, or property of an individual is under threat. In reality, however, it is very difficult to say who should decide whether it is permissible to acquire personal information and on the basis of what criteria in the event of an emergency, given that it is impossible to retrieve personal information once it
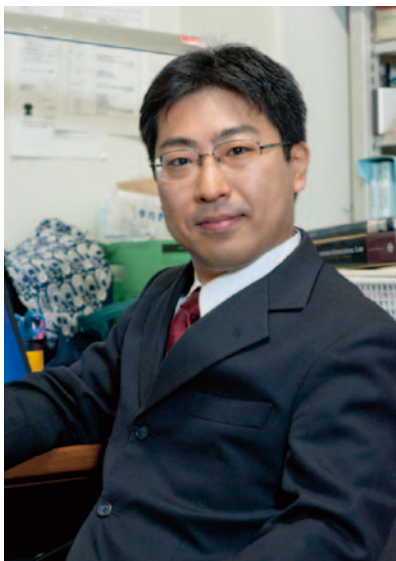
has been made public.

"So, national and local government officials, telecom carriers, legal experts, and other relevant parties should immediately undertake exhaustive deliberation and study new rules on who may acquire and use personal information, and in what manner, in emergencies ," says Associate Professor Shishido.

He mentions two significant points in such rule-making.

The first is to define the criteria for authorized exceptions to the prohibition by the Act, such as the scale of disaster. The second is the necessity to introduce a new system that, for example, makes illegible or automatically deletes the personal information acquired under exceptional authority without the consent of the person concerned after the situation returns to normal.

"Easy authorization of acquisition and use of personal information in the event of an emergency simply in view of the technical possibilities could lead to irrevocable outcomes. However, if this goes on, it will make totally useless information that can be very useful when properly handled. To break the current deadlock, it is necessary to make new rules as soon as possible," he stresses.

Emerging problems about lifelogs and the Act on the Protection of Personal Information

Meanwhile, problems with lifelogs have emerged in relation to the Act on the Protection of Personal Information. "Lifelog" refers to a record of the life and behavior of an internet user.

Today, lifelogs containing internet users' histories of web browsing and online shopping on computers and smartphones are accumulated and kept stored on servers under the control of telecom operators. From GPS-equipped terminals, it is possible to obtain data on users' positions and movements. Consequently, tourism- and transport-related services and advertising matching personal preferences based on lifelogs are constantly increasing.

Future advancement of the ubiquitous network society will expand such services, and it will also make them more sophisticated.

That would occur no problem if it were based on users' consent and were beneficial to users. In reality, many lifelogs are acquired and used or even abused without permission. That fuels anxiety and discomfort among users and even provides a hotbed for criminal activity.

However, it is terminals used, and not personal information, that are identifiable from browsing and shopping histories. For this reason, these data are not protected under the Act on the Protection of Personal Information. As a result, misrule has occurred.

To address this problem, "First of all, businesses using lifelogs should announce their purposes in collecting information or enhance their own transparency," he suggests.

Moreover, SNSs have become so popular that anyone can now freely, globally, and easily send out information. This aggravates such problems as defamation and privacy invasion. The legal system development fails to keep up with the speed of ICT progress. Associate Professor Shishido concludes: "ICT is a double-edged sword. It brings greater convenience to our life but it has a risk of causing serious problems. From now on, we have to swiftly develop laws and guidelines and to take technical actions so that we can build an ICT-based, safe and secure future."

(Written by Kumi Yamada)

1947	The Japanese Constitution comes into effect.
	·It provides for freedom of speech and secrecy of communication in Article 21.
1950	The Radio Act and the Broadcast Act are enacted.
1964	The court ruling the After the Banquet case is handed down.
	·It is the first court ruling in Japan that acknowledged the protection of privacy.
1984	The Telecommunications Business Act is enacted
1985	Nippon Telegraph and Telephone Public Corporation (NTT) is privatized.
1988	The Act on the Protection of Personal Information Pertaining to Electronic Data Processing Held by Administrative Organs is enacted.
1999	The Residential Basic Book Act is amended.
	·Introduction of the Residential Basic Book Network is determined.
2000	The Basic Act on the Formation of an Advanced Information and Telecommunications Network Society is enacted.
	·The IT Strategic Headquarters is established.
2001	The Provider Liability Limitation Act is enacted.
2002	Three laws for introducing online administrative procedures for promoting electronic government and local administration are enacted.
2003	The Act on the Protection of Personal Information and the Act on the Protection of Personal Information Held by Administrative Organs are enacted.
	The Residential Basic Book Network begins full operation.
2007	A fully amended version of the Statistics Act is enacted.
	·It authorizes the administrative agencies to furnish the anonymized statistical data.
2008	The Act on Development of an Environment that Provides Safe and Secure Internet Use for Young People is enacted.
2010	The Broadcast Act is amended.
2011	The Great East Japan Earthquake occurs.

An abridged chronology of legislative development relating to information

Information-related laws have been developed by making new legislation and by revising existing laws in line with the evolution and expansion of information and communication technologies (ICT). Among others, the Great East Japan Earthquake has reminded us how important roles that the ICT plays. On the other hand, revisions to the handling of personal information in emergencies are now highlighted. With respect to the protection of personal information, problems concerning lifelogs have also emerged. Today, ICT is making such rapid progress that there is an urgent need for new laws and rules to keep pace.

Individuals Need to Decide on Their Own in Various Situations

Norio Katayama

Associate Professor, Digital Content and Media
Sciences Research Division, NII

Should I use or avoid the expression of "omedeto" (or "happy") in New Year's greeting cards? Of course, this puzzled me in the wake of the Great East Japan Earthquake. The impact of the disaster still remained so serious that I found myself unable to feel the joy of having a New Year. When I saw Q&A websites, some argued that self-restraint was unnecessary while others suggested using an alternative expression like "Season's Greetings." It is not a question about which is right. I think that this decision should be made according to the position and circumstances of each individual.

When I think back, I realize that the earthquake has forced individuals to make difficult decisions. I suppose that those in the affected regions had to make tough decisions in different forms. Let me hereby express my sympathy for them. Meanwhile, people living in Tokyo and its suburban areas are also somewhat affected by the radioactive contamination following the nuclear plant accident. What action should be taken is left to individuals.

The recent earthquake is characteristic in that there are numerous victims and that the affected areas are vast. Under these diverse circumstances, individuals are challenged to make decisions on their own. I am a member of a research group that stores and analyzes TV broadcasts. We have closely watched how the disaster is reported on different TV networks. Before the quake, multiple stations often used similar footage. Just after the disaster, the images commonly used were limited to press conferences at the Prime Minister's Official Residence, video footage taken by the Japan Ground Self-Defense Force, and suchlike. Broadcasters aired the images of the tremors and tsunami that they shot independently. Several days after the quake, the nuclear power plant was in a serious situation. At this point, the broadcasters scaled back on-site footage and, instead, mainly showed government briefings and illustrated explanations.

Since just after the earthquake, broadcasters have provided different kinds of information in special programs. However, they failed to fully respond to personal diversity, given that too much of the information was uncertain and that the circumstances surrounding individuals were too different. As a result, it is essential for people to independently gather information from the internet and other sources while turning eyes to information from television as well. Some information available online may be inaccurate or lack integrity, while there is also some detailed information that is unavailable from television. More than ever, we are now required to collect information and to make decisions on our own. I feel that informatics now faces a new challenge. In other words, academic professionals must review and further improve the transmission and use of information in order to contribute to the recovery and reconstruction.

I ultimately decided to avoid "omedeto" on my New Year greeting cards. Beside the commonly used "Season's Greetings", I added a message that reads "May the New Year be brighter and better." Here I made a personal decision too but I believe that the wishes for a better year should be common to everyone. I am hoping to contribute to achieving the earliest possible recovery and reconstruction.

Cover illustration

In agriculture, management based on temperature, water, wind, fertilizer, and other data collected with various sensors results in good crops. Similarly, the use of massive data collected with the sensor network will make human and social activities clearer and help build a better society.