

Towards Full-scale Utilization of Big Data

Big data is considered to be the core of the digital revolution, but we are still working towards its full-scale practical utilization. This new paradigm of competitiveness is itself a site of fierce competition between nations. Is Japan sufficiently prepared? In this issue of *My Vision*, we consider the status of utilization of big data, based on discussions conducted at the joint DIJ-NIRA workshop “Big Data” (co-sponsored by the Japan Forum for Innovation and Technology at UC San Diego and the Max-Planck Institute for Innovation and Competition), held on November 1, 2018.

About this Issue

On the threshold of full-scale utilization – Seeking to obtain an accurate grasp of the current status of big data

Noriyuki Yanagawa

NIRA Executive Vice President, Professor, Graduate School of Economics, The University of Tokyo

Keywords...How should we use big data, and for what?, how far can specific technological innovations take us?

Expert Opinions

Towards Full-scale Utilization of Big Data

What are the possibilities for big data?

What should we do to promote its utilization?

Don't discard analogue technology in the era of IoT

Jun Nakamura

Professor, Graduate School of Engineering Management (MOT), Shibaura Institute of Technology

Keywords...IoT solutions that connect everything, brand-free IoT modules, analogue technology

A safe and efficient data trading market that promotes the utilization of big data

Hiroshi Mano

CEO, EverySense, Inc.

Keywords...Mediation of data provision to entities in need of specific data, neutrality, fairness and transparency

Persistent homology and big data in management research

Tristan W. Chong

Deputy Director, Research Institute of Big Data Analytics, Xi'an Jiaotong Liverpool University

Keywords...New business opportunities, data analysis method for business administration, incorporation of geometry

Which direction for corporate governance in a data-driven era?

Takuya Kitagawa

Chief Data Officer and Executive Officer, Rakuten, Inc.

Keywords...Discovering hidden value, influence on the future of management, governance of big data aggregation and utilization

Preparing for digital transformation and big data

Dietmar Harhoff

Director, Max-Planck Institute for Innovation and Competition

Keywords...Further enhancement of competitiveness, support for small and medium-sized enterprises, digital infrastructure, digital education, open government, establishment of future-oriented laws

On the Threshold of the Full-scale Utilization of Big Data

– Seeking to obtain an Accurate Grasp of the Current Status of Big Data



Noriyuki Yanagawa

NIRA Executive Vice President/
Professor, Graduate School of
Economics, The University of
Tokyo

The Utilization of Big Data – Idea Preceding Actuality

The term “big data” now appears frequently in the media, and its importance has come to be quite well recognized. Broadly speaking, there are three reasons for this.

The first is that advances in computer performance have made it possible to process large volumes of data very rapidly. The second is the fact that technological innovation, as exemplified by the Internet of Things (IoT), is making it possible to collect and transmit large volumes of data that would previously have been difficult to gather. Because these technologies have not come into full-scale use, we tend not to be fully aware of this fact, but a new era is coming in which, unlike today, the collection of “big data” will be a simple matter. The

third reason is the fact that the significance of analyzing large volumes of data has increased considerably with the development of data analysis technologies and artificial intelligence (AI). The evolution of machine learning in particular has made us strongly aware that large volumes of data are an important element in AI learning.

At the same time, in a certain respect the term “big data” has perhaps gotten ahead of itself. One sometimes encounters the argument that as long as a sufficiently large volume of data is collected, it will be possible to appropriately analyze it using AI and to derive an appropriate solution.

However, given that the basis of the functioning of AI is statistical processing, if it is not clear what type of data to collect and analyze in these large volumes, and for what purpose it is to be collected and analyzed, meaningful results will not be obtained no matter how “big” the data collected. It will therefore be essential for us to further discuss and examine precisely how and for what purpose we will use big data.

In order to conduct such an examination, we will be required to obtain a grasp of just what type of technological innovation is occurring, and how far specific innovations can take us. It will also be necessary to determine the extent to which systemic and legal infrastructure has been established to allow adequate exploitation of these technologies, or whether it will be necessary to put them in place. Unfortunately, however, in the case of big data the concept has conspicuously come in advance of the reality, and we are not adequately engaging in these types of detailed concrete discussions. Seeking to fill this gap, in this issue of My Vision, we discuss big data with Japanese and overseas experts from a variety of perspectives.

Diversifying Methods of Collecting Data

The first question to be asked is who will be collecting data, and what type of data will they be collecting, using technological innovations such as the IoT. When we speak about the collection of big data, attention focuses on the collection of data by global platform companies such as Google and Amazon. In this issue, however, Professor Jun Nakamura of Shibaura Institute of Technology

indicates the potential for aggregation of data by IoT modules, without reliance on platform companies. And while discussing IoT modules, Professor Nakamura also admonishes Japan not to overlook the importance of the analogue technologies that will form part of them.

Nevertheless, the data that can be utilized are not restricted to data that are directly collected by a specific entity. Hiroshi Mano, CEO of EverySense, Inc., points to the importance of trading data that has been collected. Given that it is rather difficult for individual companies to collect big data, Mr. Mano projects a situation in which data is traded in the market. However, he also indicates that this is a field requiring the establishment of robust systems in order to ensure fair and proper trading, for example to protect privacy and ensure that proprietary knowledge is not leaked.

How will We utilize the Data?

How the collected data will be analyzed is also a significant issue. Taking into consideration multifaceted problems such as management issues, which cannot be solved by simple data analysis, Tristan W. Chong, the Deputy Director of Xi'an Jiaotong Liverpool University's Research Institute of Big Data Analytics, introduces us to the potential of a new analytic method that utilizes topology*.

Takuya Kitagawa, Chief Data Officer and an Executive Officer of Rakuten, Inc., discusses the forms in which it will be possible to make concrete use of big data, and investigates the effects on business management in an era of full-scale utilization of big data. Dr. Kitagawa predicts that maximization of the value of big data will affect business management, and data-driven management will continue to progress.

Establishing Regulations and Systems that accord with a New Era

How are these discussions regarding big data proceeding in other countries? Professor Dietmar Harhoff, Director of the Max-Planck Institute for Innovation and Cooperation, provides us with an overview of the situation in Germany. According to Professor Harhoff, Germany is lagging behind in the utilization of big data, and small and medium-sized enterprises in particular are not adequately utilizing this resource. In addition, he claims that it is necessary for Germany to establish regulations and systems that accord with this new era.

Without a detailed analysis of the situation, we cannot clearly say whether the utilization of big data is further advanced or is lagging behind in Japan as compared to Germany. However, it goes without saying that it is also essential to establish regulations and systems related to the utilization of big data in Japan. In particular, efforts to realize an appropriate balance between the use of big data and privacy and the protection of personal information, and to reduce poorly-delineated "grey zones," will be increasingly necessary in future. And in order to enable us to think about these issues, it will be increasingly necessary to understand the real situation, to obtain an accurate grasp of which ways different types of big data can actually be used.

*A method that incorporates geometry in data analysis, and conducts analyses based on the geometric structure of data sets and the characteristics of their topological space.

Professor Yanagawa is NIRA's Executive Vice President, and a Professor in the Graduate School of Economics of The University of Tokyo. He took his Ph.D. in Economics from The University of Tokyo. His research specializations are contract theory and financial contracts.

Don't discard Analogue Technology in the Era of IoT



Jun Nakamura

Professor, Graduate School
of Engineering Management
(MOT), Shibaura Institute
of Technology

The spread of smart devices, in particular smart phones, has allowed us to make use of an enormous number of applications and services in any way that we desire. The mutual linkage of these applications and services will allow the creation of a diverse range of IoT services. As an example, we might consider a shopping service for the elderly. The placement of robots fitted with cameras in shopping centers that users at home could control via their televisions would allow the elderly a highly realistic experience of shopping that also offers a high degree of freedom, all without leaving their own homes.

In order to realize services of this type, it will be necessary to connect devices and aggregate data in the cloud. The combination of IoT modules and service platforms will be the solution to realizing these necessary conditions. In the case of a shopping service like the one described above, an IoT module would encompass devices and functions including a Wi-Fi transceiver, an HDMI terminal, voice recognition, and a video conversion function, and would play the role of connecting a variety of devices, such as a camera, a television, and a robot. At the same time, data transmitted from these devices would be aggregated in the cloud, allowing the provision of necessary services through the combination of various applications. Users would be able to receive services such as remote window shopping, operating a controller or using voice control while watching images transmitted from the robot on the television in their home.

Because it would be possible to fit IoT modules in devices produced by any manufacturer, there would be no concern that they might be tailored to a specific brand's edge devices. It would become possible to aggregate data from all devices, and the analysis of this data would enable the provision of interrelated services in an integrated fashion through the cloud. It would be important, while making use of platforms such as Google and Microsoft, to also be able to extend these platforms in order to create unique service platforms. The use of open source software would enable a wide variety of functions (increased speed, bandwidth, interfaces, etc.) to be developed and added. But while software will be important to the realization of IoT solutions that are able to connect everything, Japan should also not discard the analogue technologies that will be fitted in IoT modules.

In addition, looking back at our previous example, while the importance of image and voice conversion and recognition processing will no doubt increase in future, because the information that humans are able to see, touch and feel is exclusively analogue, technologies that connect analogue data and service platforms will certainly also be essential. I believe that Japan will be able to make use of its strengths by creating devices for the IoT era that utilize analogue electrical design technologies.

In addition to his academic position, Professor Nakamura also serves as a board member of Persol-AVC-Technology Co., Ltd., a company that is mainly involved in the design of audio, video and communication products and those technological developments, chiefly for Panasonic. Persol-AVC-Technology is also working together with Taiwan's Dynalab Inc. in the IoT field, creating service platforms that utilize big data. Professor Nakamura's field of specialization is cognitive science, and his particular interest is the visualization of human thought process. He holds a Ph.D in Engineering from The University of Tokyo's Graduate School of Engineering, and completed coursework in the same institution's Department of Technology Management for Innovation, graduating with the highest honors. He worked for a general trading company and was involved in management consulting before taking up his present position. Professor Nakamura also serves as a Managing Director of the International Academy of Strategic Management.

A Safe and Efficient Data Trading Market that promotes the Utilization of Big Data



Hiroshi Mano

CEO, EverySense, Inc.

While companies have been urged to make use of big data for quite some time, it is difficult for individual companies to collect big data. The key to solving this problem is *the market*.

The data trading market is made up of data providers, data users, and data trading market-forming companies that mediate between the two and provide data trading platforms. My company is one of these market-forming companies.

Data providers collect and organize data from their own business or other businesses, and provide it to companies, research institutions and other entities that need this data. The data provider is able to freely set the extent to which the data can be made public, and the exchange price of the data is determined between the data provider and the data user that receives the data. When a data user makes a request regarding the type of data that it requires, my company conducts a search for data providers that match the specific conditions, and automatically matches the data user with data providers. While we are a private enterprise, we do not buy, sell or retain data, and we do not have the right to set prices. It is essential for a service that acts as an intermediary to ensure neutrality, fairness, and transparency in transactions.

Data providers in this market may experience anxiety over the possibility of specific aspects of their professional expertise being made public via the market. For example, farmers may not wish aspects of their professional expertise such as the timing for watering of their crops or their methods of cultivation to be made public. But what the data trading market deals in is *data* and nothing more. For example, if the data user's focus is on temperatures, then only temperatures are included in the data – as this indicates, only primitive data is made public.

Naturally, the market is not entirely perfect. While market-forming companies mediate data transactions, they do not survey individual data from the standpoint of the protection of privacy, and it can therefore be the case that low-quality data is sold, or that undesirable companies collect data. However, in the case of the service provided by EverySense, all transactions are tracked, and data is only initially released to the data user when both the data provider and the data user have approved the transaction. If we are so instructed by either party, data disclosure and data transmission and reception can be stopped instantly.

The Data Trading Alliance, which was launched in November 2017, has formulated and published standards for the certification of data trading market-forming companies. I believe that this mechanism will make it possible to ensure the reliability of these companies.

In addition to conducting research and development in the field of IoT, EverySense Japan Co., Ltd. offers the IoT data distribution platform "EverySense." The company provides a service that matches IoT data and entities including companies and research institutions that will make use of that data for purposes such as business development and academic research, acting as an intermediary in the sale and purchase of data. In 2014, EverySense established EverySense Inc. in Silicon Valley. Mr. Mano also serves as the President and Secretary General of the Data Trading Alliance, a general incorporated association. He works tirelessly for the realization of sound growth and the establishment of an appropriate technological and institutional environment in the data distribution market.

Persistent Homology and Big Data in Management Research



Tristan W. Chong

Deputy Director, Research
Institute of Big Data
Analytics, Xi'an Jiaotong
Liverpool University

An increasing number of Asian countries have adopted Internet economic conducive to rapid growth and a presence in different business industries. In the age of the Internet of Things, the development of big data technology has brought significant changes in recent years. For instance, in China, consumers are exposed to an expanding, fragmented array of management touch points across digital media and sales channels, little is known the challenges faced by businesses in this regard. Powered by the integration of social media, e-commerce, cloud computing, and mobile computing, big data technologies provides fundamentally new insights into business's effect on revenue. Therefore, big data analytics is increasingly recognised as an emerging approach for integration between public/private sector organisations and as a way to identify new

business opportunities in the region.

However, it is unclear in current research what approaches are available and optimal for data analytics in management, which needs to be clarified and comparatively evaluated. To seek a better analytical method in order to have a better understanding in management studies, theoretical development is in a need to provide general guidance for researchers to implement big data methods. Dealing with vast amount of data is difficult, and it is even more challenge if the data structure with noise, high-dimensional, and incomplete. Therefore, to include other elements such as computational geometry and topology in "traditional" technique will enhance the visibility of the complicated structure of data.

In the past decades, scholars have found that persistent homology (PH) is an effective computational tool for characterisation with topological features for large data sets. In comparison with other traditional techniques, PH has unique ability to compute data at different topological space and spatial resolutions. Over a wide range of geometric data analysis, more topological features are detected and are deemed more likely to reflect true features of the underlying space, rather than artifacts of noise, sampling or particular selection of parameters. Academics can uncover a new and more insight characteristics that can deepen the understanding of big data. It would be easy for scholars to experiment and study on various business issues because it offers topological insight for geometric structures and/or intrinsic features against noise. In other words, PH can provide excellent results for the identification, characterisation and classification of various business behaviours. Consequently, PH is a better tool for topological characterisation from vast amount of complex data which is a realistic alternative and can contribute to innovative theory development.

(Submitted manuscript)

Dr. Tristan W. Chong is Deputy Director of the Research Institute of Big Data Analytics and Associate Professor of Marketing Analytics at the Xi'an Jiaotong – Liverpool University, Suzhou, China. His research interests are in the areas of B2B e-marketing, topological data analysis and applications, and marketing analytics. His research has been continuously funded by both internal and external funding bodies including the Natural Science Foundation of China. He is also a Visiting Scholar at the National University of Singapore, Singapore; Deakin University, Australia; Heriot Watt University, UK; and South University of Science and Technology of China.

Which Direction for Corporate Governance in a Data-driven Era?



Takuya Kitagawa

Chief Data Officer and
Executive Officer, Rakuten,
Inc.

What does “the utilization of big data” refer to in concrete terms? For me, it is the discovery of hidden value. One example is the estimation of corporate value. Investors today want to obtain corporate data quickly, without waiting for quarterly results. A US company that provides a service analyzing companies’ production status and their level of appeal to customers based on satellite images became a major talking point in 2018 when it identified the status of production of electric vehicles by Tesla. Credit assessment provides another suggestive example. No-one trusts a friend based on the amount of money that they have. Trust is earned by keeping promises – arriving at appointments on time, returning loans of money or property, etc. In other words, we could obtain a more accurate credit assessment by analyzing data including the history of repayment. And while this is not given much attention, in the future, analyses of where user interest is tending will generate considerable value. Web businesses optimize their sites by analyzing the direction of user interest from user behavior, and it can be predicted that the use of such methods will expand in future.

This maximization of the value of big data will affect the future of management, and a data-driven management orientation is continuing to advance in companies. It is already the case that the role of top management is no longer to issue the company’s quarterly results. The role of top management is to predict what the situation will be three years in the future, and to make decisions as to what type of data to utilize and retain. Nevertheless, aggregating data from multiple services and group companies is not an easy matter. Rakuten itself offers more than 70 services, and there are more than 100 million Rakuten members using these services. From 2017, therefore, Rakuten has been testing a hub and spoke model in relation to its Chief Data Officer (CDO) positions. This is a model in which the CDO for the entire Group (a position that I fill at present) cooperates with the CDOs of each Group company. Our approaches include data collection and data governance. In addition, every quarter we hold conferences attended by the CEOs and executives of each Group company, at which we spend an entire day discussing data strategy. All data utilization must be something that provides value to the customer. Leadership by top management to create a customer-oriented corporate culture will be essential to business into the future.

Dr. Kitagawa is involved in the establishment of a data science organization within Rakuten. He serves as both an Executive Officer and the Chief Data Officer for the Rakuten Group as a whole, and is responsible for the Group’s data strategy. Dr. Kitagawa also serves as a Director of Rakuten Data Marketing Co., Ltd. He is currently engaged in efforts including the creation of data infrastructure, the expansion of the customer experience based on scientific analysis, the creation of an advertising business using data, and the stimulation of business innovation. Dr. Kitagawa was previously a theoretical physicist, and took his doctorate through coursework in Harvard University’s Graduate School of Physics. He has had more than 20 papers published in scientific journals including Science, Nature Physics, and Physical Review Letters.

Preparing for Digital Transformation and Big Data



Dietmar Harhoff

Director, Max-Planck
Institute for Innovation
and Competition

Digital change is an extremely rapid process. Artificial intelligence, big data, cloud computing and the disruptive business models associated with them challenge Germany's specialization advantages which were attained over decades. In the next few years, many countries, and Germany in particular, must develop new technical and economic strengths. This will require consistent and prompt policy measures.

Big data and related innovative services represent the kind of innovations that supersede existing technologies, products or services. Frequently, they lead to the emergence of entirely new markets. At this point, Germany lags behind in the use of big data. Furthermore, considerable differences between large companies and SMEs in the use of big data approaches pose the risk of a 'digital divide' in the corporate sector. Many SMEs do not seem to be fully aware of the importance of the imminent changes or a lack of financing makes it difficult to tackle them. SMEs may need special support with digital change.

Other important areas for policy-making will be the development of a future-proof digital infrastructure suitable for high performance and upgradability. Moreover, digital education should be strengthened for all age groups and levels of training, but most notably in elementary and secondary schools. Computer science should be understood as a new key discipline and incorporated more into the curricula of other training courses.

Administrations can also foster the use of big data by providing public sector data as open government data. In Germany these are not yet automatically made available via well-structured access systems. When it comes to digital governmental and administrative processes, Germany still has a lot of catching up to do. Recently, promising legislative framework conditions have been created for the establishment and operation of efficient public data and e-government portals.

It is tantamount to create a future-oriented legal framework for the digital economy - e.g., in the fields of copyright, data protection and consumer protection. In the long term, grandfathering and perks for established business models will jeopardize a country's competitiveness in the digital age. Thus, legislation must not be geared towards building protective fences around established sectors. Rather, the framework should enable the fast introduction of models of the digital economy.

(Submitted manuscript)

Professor Harhoff is Professor for Entrepreneurship and Innovation at the Ludwig-Maximilians-Universität (LMU) Munich. His research focuses on innovation, entrepreneurship, intellectual property, and industrial economics. He holds a Diploma degree in mechanical engineering from the Technical University of Dortmund, a Master's degree in Public Administration from Harvard University, and a doctoral degree from the MIT Sloan School of Management. Since 2004, he has been a member of the Scientific Advisory Boards of the Federal Ministry of Economics. Professor Harhoff has also been a member of the Economic Advisory Group on Competition Policy (EAGCP), the Chair of the Economic and Scientific Advisory Board (ESAB) at the European Patent Office, and the Chairman of the Commission of Experts for Research and Innovation (EFI) of the German Federal Government. He has also published numerous books and papers.