Working on Industrial ICT Solutions

The Internet of Things (IoT) has triggered massive changes in industry. We at Toshiba are proceeding to change our business to one where we "Provide Products and Excellent User Experiences" (*"things and experiences"*) so that we not only provide physical products using our manufacturing technologies in various areas of businesses but also provide value as a service created by software.
To combine *"things and experiences"*, Toshiba will collaborate with its customers in creating a variety of new businesses using information and communications technology (ICT).
These feature articles introduce our efforts for *"things and experiences"* as well as our software and data technology that is supporting these efforts.

Shigeyoshi Shimotsuji
Corporate Vice President of Toshiba Corporation
Executive Vice President of Toshiba Corporation
Industrial ICT Solutions Company
General Manager of the IoT Technology Center

Co-creating New Businesses with Customers by Concentrating the Comprehensive Strength of the Toshiba Group

Accelerating the creation of new value using IoT technology and the co-creation of *"things and experiences"* businesses

The purpose of using ICT is to increase the value of products and businesses and to create new businesses. The Toshiba Group is currently in the process of promoting a reform of its business structure through ICT-based innovation and shifting a focus towards new service businesses. The key to success in all these initiatives lies in the IoT market, which is expected to expand in the future. Industrial ICT Solutions Company, an organization inside Toshiba Corporation, is committed to the area of IoT with the aim of creating *"things and experiences"* businesses in cooperation with Toshiba's in-house and group companies (hereinafter collectively referred to as "business divisions").

Noriyasu Okitani
Senior Manager
IoT Business Creation Division
Industrial ICT Solutions Company
Toshiba Corporation

Jointed Toshiba Corporation in 1985. Having worked on the development of computer systems for steelmakers, telecommunications carriers and social infrastructure, he is currently in charge of IoT- and IoE-based service businesses.

A shift from conventional business models of selling products to service-oriented models

The industry is redefining business models throughout the world, while utilizing cloud and big data to deliver various added-value solutions rather than simply selling products, services and their combinations. Catching the rising tide, the Toshiba Group is also striving to create different types of service business models.

At the same time, the Toshiba Group is looking beyond the ongoing shift. Our next step will be to create completely new value and business models, such as performance-based service models that are based on the value provided to customers. IoT, especially ICT technology, is the key to realizing them. In order to build new service models, it is crucial to monitor and control *"things"* using sensors and understand the actions and intentions of human beings.

Customers do not simply want to possess a "thing." Rather, they wish to solve a problem or realize new value by using it. To meet these customers' needs, we should not be restricted to using only our proprietary technologies and products. Instead, we should have an open mind about looking for external resources and offer the best mix of value to our customers. And, we also consider it is important to allow our customers to use "ONE Toshiba" service, as it offers them a way to get all the products and services they need in just one stop.

Supporting the creation of new value through building services and platforms

Our company reorganized in April 2015 with the aim of further improving all capabilities of ONE Toshiba. The purpose of this reorganization is to build services and an IoT platform in cooperation with the business divisions of the Toshiba Group. The main strategy of the former is co-creation of services through supporting consultation programs for the business divisions of the Toshiba Group as well as external customers. Each business division of the Toshiba Group has a deep understanding of its customers' business situations. The strength of the Toshiba Group is its assets of knowledge and experience in various industrial fields. By creating data models that reflect our broad expertise, we aim to provide new value through IoT technology.

Going one step further, we send our engineers to the customers' customers (end users in many cases) as part of our efforts to create value that will have a strong appeal to them.

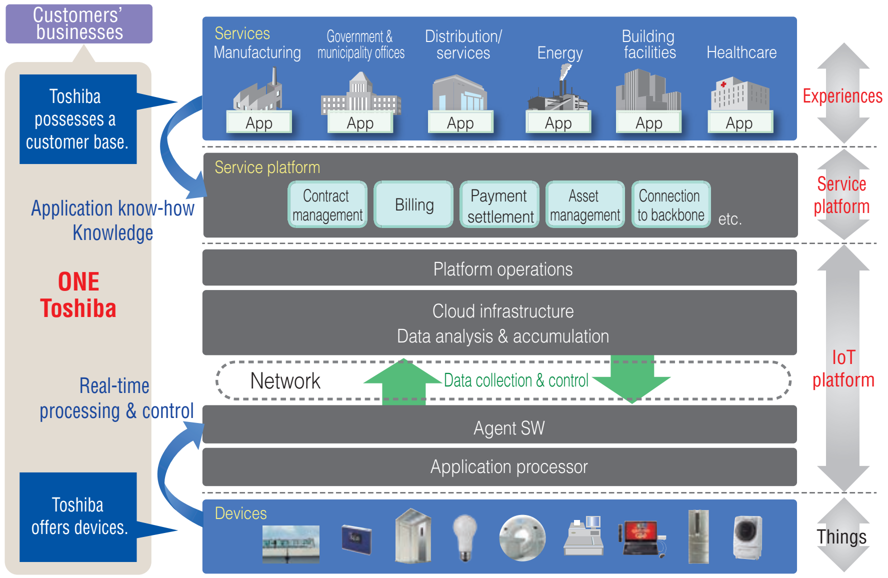
For example, our customers remotely monitor the installations and operations of process-related devices that they have provided to various plants around the world. We work in tandem with these customers based on this data to ensure their stable operation and timely maintenance[*1].

Our new vision is to create services for various industrial fields including energy and other social infrastructure, healthcare, distribution and manufacturing by utilizing our IoT resources.

*1 For details, see page 8 of Volume 13 of T-SOUL.

IoT services are implemented in several layers (Figure 1). At the lowest layer lie the *"things"*. These are devices that reside at endpoints. The next higher layer is an IoT platform for collecting, accumulating and analyzing data. Then comes a service platform that provides functions such as contract management and billing. The top layer is experiences. It means specific services. To create top-layer services, we will forge close relationships with customers and the business divisions of the Toshiba Group.

Figure 1 Concentrating Comprehensive Expertise of the Toshiba Group (Toshiba IoT Architecture)



The Toshiba Group has a long experience with the *"things and experiences"* layers. Drawing on our knowledge and expertise, we will improve the IoT and service platforms to offer our customers comprehensive products and services in the framework of ONE Toshiba.

Sharing an IoT platform that supports various services

In order to offer IoT services, it is necessary to build IoT and service platforms. One approach is to create these fundamental functions for individual services. It is true that they will deliver optimal services; however, building custom solutions would be time-consuming and costly to accept for most customers.

Therefore, flexible IoT and service platforms that accommodate diverse services are crucial. At first, the IoT platform must be connected to a broad spectrum of products and equipment (hereinafter collectively referred to as "devices") and provide cloud-based functions for collecting and analyzing their operating data as well as availability, failure and other pieces of information and for controlling them accordingly. The application processor chips (ApP Lite*2) are embedded in the network gateways and devices on the customer's site (device side). The agent software runs on them (Edge Computing) and collects and processes data depending on the demand of cloud processing. Distributed cooperative processing by cloud and edge computing provides the combined benefits of real-time performance and robustness, making the platform suitable even for social infrastructure and other demanding applications. We will seek to form an alliance with global players with a view to establishing an industry-standard open platform.

Now, let's look at the service platform that provides commonly used service functions. That includes new billing-and-payment functions such as pay-per-use, pre-pay and pay-per-result; customer relationship management (CRM), a customer-oriented approach to acquiring and maintaining customers; management of service-level agreements (SLAs), or service contracts in which the scope and quality of a service is agreed upon between a service provider and a service user; and asset management for products and equipment that are set up at customer sites.

The Toshiba Group has many years of experience in providing various types of equipment to diverse facilities and infrastructure domains. The Toshiba Group will turn the first-hand knowledge built up through decades of its *monozukuri* experience into packages of expertise concerning on-site devices and their sensing capabilities as well as data collection, analysis and control. By sharing them across the Toshiba Group, we will promote the co-creation of new IoT services with customers.

Lean startup to accelerate business development

As described above, four layers—*"experiences,"* service platform, IoT platform, and *"things"*—comprise the entire platform called the Toshiba IoT Architecture. An important point to consider in building this platform is how to get the best mixture of our own and external resources.

For example, the service platform covers a vast range of service sectors, and it takes a long time to prepare all the necessary

*2 Toshiba applications processor IC

elements of it on our own. Therefore, it's important that we have an option to cooperate with an enterprise that has a strength in specific services such as proxy payment and leasing. It is necessary to evaluate external services properly and incorporate them into the overall architecture in a consistent manner.

We can say the same thing about the cloud platform that forms part of the IoT architecture. When you look at the world, you will find several services that have a strong presence in the field of cloud computing. Selecting appropriate ones from these services may be helpful. In the field of mission-critical applications, however, Toshiba's high-quality cloud platform will be considered optimal in many cases.

The Toshiba Group is involved in many mission-critical areas, including social infrastructure and healthcare. In order to utilize these mission-critical applications in this area, there are three requirements for their platforms: real-time performance, robustness and scalability. While building a platform that meets all these requirements, using external services might be beneficial; for example, it might be beneficial for areas that place higher priority on cost than robustness. The point is to have the capabilities to accommodate diverse services for a broad range of customers.

We are promoting the introduction of new development methodologies to co-create new services with customers using the Toshiba IoT Architecture.

Conventionally, it was common to take much time to listen to a customer's needs, desires and issues for each project and turn them into a workable concept for a system. In contrast, an approach called lean startup is attracting increasing attention since it emphasizes speed as a critical ingredient in product development.

Once we have shared a vision with our customers and identified the issues that must be solved to realize it, we will develop and release a minimum set of service features first. Then, based on the customer evaluation of and feedback on the measurement of its effect, we will modify and add service functions iteratively, gradually building overall services to perfection. This method, called lean startup, will offer customers a significant time-to-market advantage and help shorten the time to improve their services.

"Caring for people IoE" that recognizes speech and images

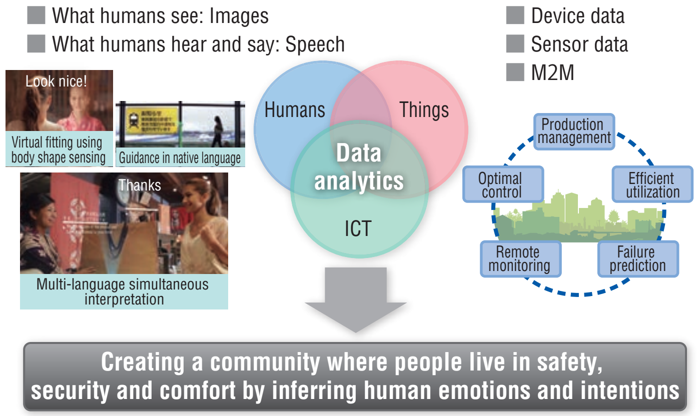
The widespread adoption of IoT will have a significant impact on a broad range of industrial sectors and social services.

For example, the concepts of total IoT utilization can be found in the Industrial Internet of General Electric (GE) and Industrie 4.0 in Germany.

IoT has two major roles: 1) enhancing not only the quality of *"things"* but also their use value (i.e., customer experience value), and 2) optimizing the efficiency of manufacturing *"things."* We believe that Toshiba, a key player in the manufacturing sector, can take advantage of its expertise to fulfill both these roles. We also believe that we will be able to offer even more sophisticated services by combining data related to *"things"* and human behavior (Figure 2). Many of these services are getting closer to practical use, for example borderless and barrier-free communication services that recognize speech and images; work assistance at medical and nursing care facilities; and vehicle and machine operation assistance. We have worked on the R&D of multi-language recognition, machine translation, image recognition and other related technologies as well as the development of their application systems from an early time.

It is humans who touch and use *"things."* Unlike *"things,"* humans have emotions and intentions. Here comes our media intelligence technology that can utilize knowledge processing to infer human emotion and intention. We will make full use of our "Caring for people Internet of Everything (IoE)" technology to help create an ideal Human Smart Community, where people live in a safe, secure and comfortable society.

Figure 2 **Shift from IoT to "Caring People IoE"**



Toshiba aims to realize a caring people IoE that connects all people and things to infer human emotions and intentions from data related to *"things"* and human behavior.

Creating New Value of Social Infrastructure by Using IoT Technology and Expertise

Hideaki ishii
Chief Technology Executive
Social Infrastructure Systems Company
Toshiba Corporation

Toshiba's Social Infrastructure Systems Company consists of four business divisions: Transmission & Distribution Systems Division, Railway & Automotive Systems Division, Security & Automation Systems Division, and Defense & Electronic Systems Division. In recent years, we have been engaged in the *"things and experiences"* business to increase its presence in the service sector.

As our growth strategy, we try to enter new markets and launch new products while, at the same time, we try to improve the strength of our existing products in the marketplace. When new products are developed for new markets, *"things and experiences"* businesses tend to be built. Examples of such *"things and experiences"* businesses include energy, urban transportation, physical distribution, and weather and disaster-prevention solutions.

Knowledge of and expertise in ICT and IoT are essential ingredients for promoting *"things and experiences"* businesses. Support of skilled persons is highly beneficial to collaborate closely with Industrial ICT Solutions Company, which is specialized in ICT and IoT. New businesses begin to be created as professionals in diverse technical and business areas interact with different divisions of the Toshiba Group and external customers.

There are two key points for forming business alliances: 1) enhancing the competitiveness of an existing business using IoT technology, and 2) creating a *"things and experiences"* business.

A major challenge in realizing the first point is incorporating differentiating IoT technologies into power inverters, storage batteries and other power electronics products. As for the second point, the Social Infrastructure Systems Company is endeavoring to develop services using IoT technology. For example, IoT technology will be able to make new services such as a demand response control that manages customer consumption of electricity in response to supply conditions using smart meters.

Creating a smart city with enhanced energy and transportation efficiencies

Value creation using IoT technology has already begun. One example of such value creation can be seen in the Yokohama Smart City project in which the Toshiba Group has played a major role. In this project, we have successfully reduced overall power consumption in the project area by using home, building and community energy management systems together with storage batteries and other devices.

Various issues caused by urbanization have become increasingly serious concerns throughout the world, especially in emerging countries. In addition to energy management, smart urban transportation is an important point to focus on. In this respect, the Social Infrastructure Systems Company can also offer new value that will help alleviate traffic congestion by combining IoT technology with its expertise built up through the social infrastructure business.

We will contribute to the local community and people's lives by making our social infrastructure more sophisticated using IoT technology. We will address the challenges we face one by one and pursue the potential of the *"things and experiences"* business that will be beneficial to our customers.

Applying Software Technology to Create Experience Value

Enhancing software technology that supports platforms and processes

Software technology is crucial in the *"things and experiences"* business. To establish a common platform and a common process by enhancing software technology leads to the launch of new *"things and experiences"* services for the Toshiba Group. Taking advantage of the software technology cultivated through many years of experience in devices and services, the Industrial ICT Solutions Company is developing new software technology required for the era of the Internet of Things (IoT). As IoT services are expected to expand overseas, we are reinforcing software development centers in India, Vietnam and elsewhere.

Takashi Amano

Technology Executive
Industrial ICT Solutions Company
Toshiba Corporation

Jointed Toshiba Corporation in 1991. Having worked on the development of hardware and software for mobile devices and PCs and then the development of consumer cloud services, he is now involved in the development of software technology for IoT and IoE applications. He serves as Toshiba's representative in the Industrial Internet Consortium (IIC).

Promoting a common platforms and a common process

Software technology is essential in the *"things and experiences"* business being promoted by the Toshiba Group. The key to success in creating Toshiba-original IoT services or co-creating new IoT services with customers lies in platforms and processes supported by advanced software technology.

A platform means a foundation that supports a *"things and experiences"* business and it consists of four layers—*"experiences,"* service platform, IoT platform and *"things"* (see page 3).

The word "process" is used here to mean methodologies and facilitation techniques for creating a *"things and experiences"* business. We are working to further enhance software technology and competitiveness by promoting a common platform and a common process.

In order to bring IoT service to fruition, in addition to facilitation techniques and tools that help clarify customers' business issues, it is important to have an efficient software development processes such as agile software development. This article introduces our technology to properly manage and support the entire agile development process ranging from requirements analysis and design to implementation and testing (Figure 1).

Figure 1 Process and Quality Technologies

Requirements Analysis Phase

Customer requirements-to-specifications translation
Technologies that allow us to effectively create requirements of specifications without ambiguous, missing or inconsistent requirements, and that manage changes of the specifications and implementation status throughout the development cycle

Formal methods
Technologies that allow us to eliminate ambiguous, missing and inconsistent specifications, and perform a mathematical analysis and an exhaustive search over a state space to uncover errors at an early stage in development

Design Phase

Product line engineering
Technologies that allow us to efficiently build a common software platform to develop a wide range of product series.

Model-based design
Technologies that allow us to design software at a higher level of abstraction than general-purpose programming languages such as C to check its quality in areas upstream of the development cycle and ensure reliable quality of source codes

Implementation Phase

Software structure diagnosis
Technologies that allow us to quickly identify all structural problems in software that could lead to extra cost due to changes or extensions to specifications and minimize the amount of alterations to its design and implementation

Development infrastructure building
Technologies that allow us to improve the efficiency of the build and intra-group release processes as well as communicating information, and to build a development infrastructure that will reduce loss of time

Agile methodologies that support efficient software development are backed up by a suite of process technologies cultivated by the Toshiba Group.

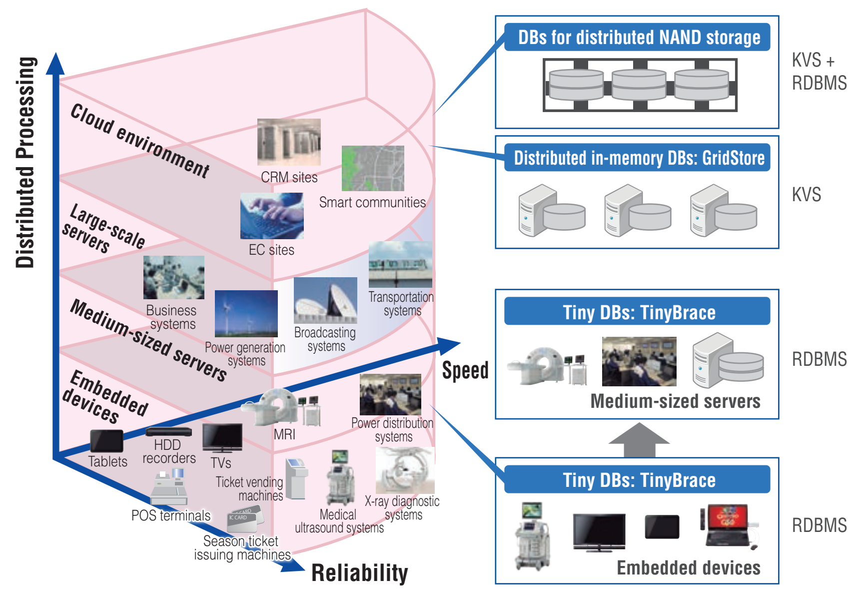
Software process technology to improve the quality and efficiency of agile development

Software development begins with requirements analysis. Customer requirements are not always clear and consistent, and it is often necessary to resolve ambiguous or missing requirements. Customer requirements may also include conflicts that are not plainly apparent.
Requirements specification techniques help uncover these problems at an early stage and improve the quality and efficiency of software development. These techniques are also used for the management of specification changes.
Formal methods, which are similar to these requirements specification techniques, are also useful. Formal methods perform a mathematical analysis and an exhaustive search over a state space to eliminate inconsistent and missing requirements in the early stages.
After the requirements analysis phase comes the design phase. In this phase, product line engineering techniques play more significant roles for agile development. Product line engineering makes it possible to develop a series of software products efficiently by using a common underlying platform rather than applying separate processes to individual products. A model-based design technique handles software design at a higher level of abstraction than general-purpose programming languages such as C. Software is written in model description languages such as the Unified Modeling Language (UML) to ensure that its design matches customer requirements. Increasing the level of abstraction sometimes helps us find the true essence of a design.
Finally, several techniques are indispensable in the implementation phase, such as software structure diagnosis for quantitatively evaluating the quality of a software design and a methodology for building a development platform comprised of a toolset or a tool chain that supports the entire development process. All these software processes and quality techniques form the essential foundation for agile development.

Leading-edge software technology and media intelligence technology

Apart from process technologies, the Toshiba Group is also committed to the development of leading-edge software technologies. For example, in the field of embedded technology, the Toshiba Group is working on high-speed processing, power-saving, code size optimization and so on while supporting Linux and other general-purpose operating systems.
For the storage and management of the collected big data, the Toshiba Group is also involved in the development of database technology and products (Figure 2).
IoT data is collected from many sources including social infrastructure domains, business activities and daily lives, and it comes in various types and formats. To provide IoT services, storing data is not enough. Rather, it is more important to quickly retrieve data, analyze it and utilize the final result. In order to address the most demanding needs to build, manage and analyze huge databases for smart grids and smart cities, the Toshiba Group supports IoT services with wide-ranging technologies.
Furthermore, the Toshiba Group is developing application frameworks for graphical user interfaces (GUI) on the Web as well as user interface technologies. User-friendly human-

Figure 2 Database Technology to Reliably Store Data



Data is accumulated and managed according to the specific performance required for different scenarios ranging from individual devices to smart communities.Toshiba develops and offers its own products such as GridStore, a distributed in-memory database.

machine interfaces that provide intuitive operation and visual feedback are becoming increasingly important to offer the premium user experience (UX) to customers.

Additionally, media intelligence technology is one of the important software ingredients to realize IoT services that include knowledge processing of text, speech, image and other types of data. The Toshiba Group has been working on this technology for a long time, and its combination with IoT technology is expected to bring about new possibilities. "Human behavior" data is a subject of knowledge processing. Its results can be combined with operating data of *"things"* to open the door to human "intention understanding" that is impossible to achieve with sensing alone.

For example, assistance can be given for field work, such as a situation in which a serviceperson has come to repair a piece of equipment. It is possible to combine his or her utterances with IoT data to produce advanced maintenance techniques. Accumulating data on both "human behaviors" and *"things"* will enable to create novel services.

Increasing importance of data security in the IoT era

As IoT services gain momentum, security is attracting much attention. As all *"things"* become interconnected via the Internet and more services are being provided over the Internet, total security is expected to cover not only the area of information technology (IT) but also operational technology (OT). Especially data security is crucial in the social infrastructure and healthcare fields.

The Toshiba Group has its own unique advantage that specialized information security providers do not. For example, we can implement security functions in our endpoint embedded devices. We consider that we can provide security solutions at all levels from the cloud to the endpoints.

Last but not least, let me mention global sourcing of software development. In order to expand our IoT service business globally, we have established development centers in India and Vietnam where over 1,000 engineers are working on software development and data analysis.

In the future, Indian and Vietnamese engineers will be more actively involved in overseas social infrastructure projects, especially in Asia. We can support customers' global businesses strongly by these overseas engineers collaborating with domestic engineers, and that will stimulate both engineers.

Forefront of Data Analytics

Optimization of O&M and electricity demand forecasting
Evolving data analytics landscape in the IoT era

Analysis and utilization of data for products and services is expected to produce many positive effects, including cost optimization, enhanced customer satisfaction, and even an improvement in the overall efficiency of society.
The Toshiba Group is working on data analytics using data captured from manufacturing processes and customer sites in various business fields. This article describes the landscape at the forefront of IoT and data analytics, focusing on optimization for operation and maintenance (O&M) of equipment and electricity demand forecasting.

Koji Okada
Ph.D. (in Engineering)
Senior Manager
Data Science & Analytics Department
IoT Technology Center
Industrial ICT Solutions Company
Toshiba Corporation

Jointed Toshiba Corporation in 1999. Having worked on R&D of information security technology, he is currently leading a group of data scientists specialized in R&D of data science and analytics. He is also promoting the solution business for the utilization of big data in IoT.

Using IoT for manufacturing and O&M to reduce costs and increase sales

The Toshiba Group is engaged in a wide range of R&D projects in the field of the Internet of Things (IoT). Among them, Industrial ICT Solutions Company is working on the analysis and utilization of data captured from sensors and other *"things."*
The Toshiba Group has many factories in Japan and overseas, and it is relatively easy to attach sensors to its own factories to collect data. Additionally, some of our products that have been delivered to our customers such as elevators and multi-function peripherals (MFPs)*[1] are under remote monitoring by the constant data acquisition of operational status. Thus, Industrial ICT Solutions Company has been at the forefront of the R&D of data analytics in the fields of manufacturing and O&M.
At the site of manufacturing and O&M, there are many issues that can be solved with data science. For example, we need to predict the potential failure of manufacturing equipment in advance in order to prevent an abrupt shutdown of a production line. Manufacturers also wish to shorten the repair time and optimize parts inventory. Solving these issues will lead to cost reduction.
Additionally, IoT may contribute to an increase in sales. If you can identify the cause of a product failure and provide feedback to people who are in charge of the production and development it will be possible to improve quality or appropriately update the next model. For performance-based services that are charged at a per-volume or per-hour basis, a reduction of downtime will affect sales positively.

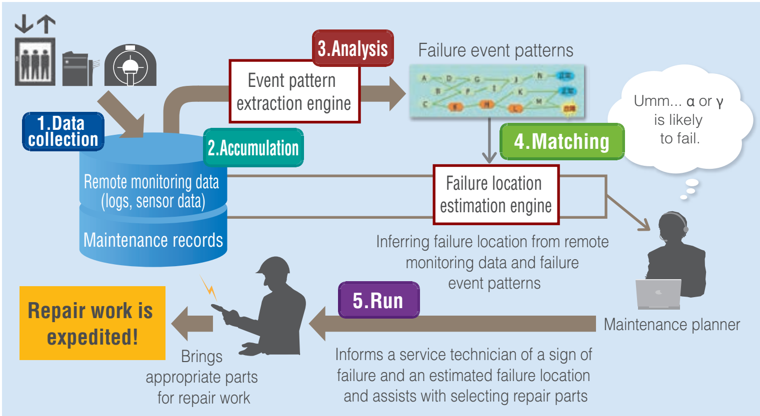
Preventive maintenance using event pattern extraction technology brings a win-win relationship (between a customer and a vender).

Our event pattern extraction technology is one of the technologies that will play a pivotal role in solving the challenges in the fields of manufacturing and O&M. This technology can be used to extract a characteristic trend in time-series data that will lead to a specific result. For example, it estimates that there is an 80% probability of trouble arising when four events, A, B, C and D, have occurred in this order within 10 minutes. Facts like that can be translated into a set of rules to incorporate into remote monitoring systems. This makes it possible to detect a sign of equipment trouble even before service technicians become aware of it and thus to implement preventive maintenance. Reducing downtime not only helps us improve customer satisfaction but also provides benefits for manufacturers running a support team such as fewer emergency calls and thus a reduction in operating cost.
Furthermore, the event pattern extraction technology comes in handy even in the event of equipment trouble. Since this technology can use an operating log or sensor data to infer a possible cause of the trouble, service technicians can prepare necessary parts and tools before heading for a customer site and thus they can increase maintenance efficiency (Figure 1).

*1 A machine that incorporates the functionality of multiple devices such as an all-in-one printer

It is known that experienced O&M engineers possess a keen sense of foreseeing certain equipment troubles and tend to notice subtle changes in the rhythm or sound of equipment. They can predict an impending failure of a certain part of equipment or conjecture its cause, and take an appropriate measure accordingly. Nevertheless, manufacturing systems are increasingly becoming complicated. This means myriad permutations of events are conceivable as the causes of trouble before it becomes apparent; so there are limits to what even experienced engineers can do. Using event pattern extraction technology is highly effective in this situation.

Figure 1 Estimating a failure location by analyzing an operating log



To assist service technicians in choosing repair parts, we infer a failure location from a pattern of events that has led to a failure.

Forecasting electricity demand of each household by categorizing usage patterns

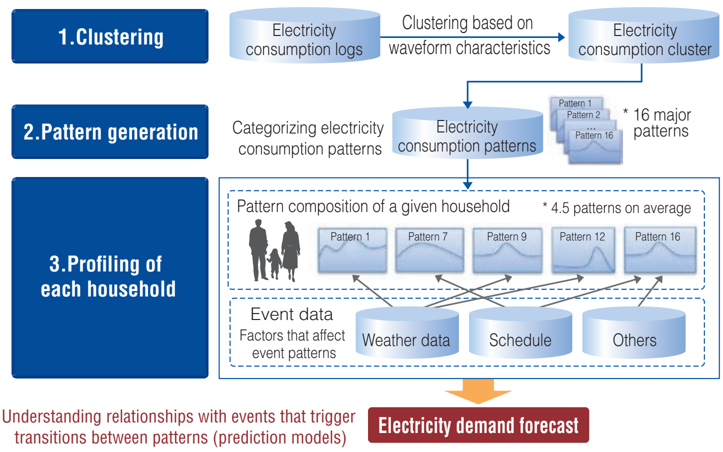
Now, let's look at electricity demand forecasting. Here too, data analytics is the key.

It is possible to aggregate and visualize total electricity consumption if houses, office buildings and other electricity consumers are equipped with smart meters and/or energy management systems (EMS). Daily electricity consumption curves are closely related to people's life patterns, and they provide various types of information. Let me describe an aspect of our research, taking general households as an example.

Household electricity consumption curves can be classified into several types. For men in their thirties living alone, the most common pattern is "up early morning and out during the day." In other words, they consume more electricity in the morning and are not at home in the daytime on weekdays. On weekends, their electricity consumption increases slightly. Households of four persons or so may be categorized into the "high-income and high-electricity-consumption" group. This group consumes a large amount of electricity in the daytime, and the graph made based on the data of its electricity consumption looks like a mountain. A typical electricity consumption curve of households of couples over 60 years of age is characterized by "three peaks at mealtimes."

As a result of analyzing various types of households, we have found 16 major patterns of electricity consumption. However, a given household does not follow the same pattern every day and often transitions from one pattern to another. On average, the variation of electricity consumption patterns are 4.5 in one season and in one household. Conversely, household attributes can be inferred from the composition of these patterns. Forecasting electricity demand is possible if we understand the composition of consumption patterns and the relation with the events that trigger the transitions between these patterns (Figure 2). One of the contributing events is weather. For example, given the information "a sunny weekend with a temperature of 20°C at 10 a.m.," the behavior of a given family is predictable to a certain extent. If a family is likely to go out, its electricity consumption will remain low during the daytime. The electricity demand for an area can be estimated by collecting a significant amount of such information. If the accuracy of

Figure 2 Electricity demand forecasting based on consumption patterns



Electricity demand is forecast based on the composition of household electricity consumption patterns and the events that trigger transitions between patterns.

estimation is improved, this information will be useful for electric utilities.
Furthermore, household profiles inferred from electricity consumption patterns can be utilized for advertising purposes because they include not only the household size but also family lifestyles. The use of household profiles makes it possible to predict the overall behavior of a family to a certain degree, for example, how many families will not be home until evening. Bargain sales and other advertisements can be emailed to the selected people who are likely to go out.

Prior data analytics examples in the medical, physical distribution and advertisement fields

The Toshiba Group is working on services using data analytics in various fields.
For example, in the healthcare field, the Toshiba Group has started to collaborate with its health insurance union for a service designed to prevent illnesses from getting worse. At present, this service checks the results of a physical checkup and the statements of medical expenses to identify people with a high risk of developing diabetes and urge them to consult a doctor. In the field of distribution, the Toshiba Group is conducting an experiment at a warehouse in which workers wear a wristband sensor for tracking their movement to get feedback on more efficient and shorter routes. An experiment in the retail field is designed to use facial images to estimate the gender and age group of individuals who are looking at digital signage screens and show them advertisements chosen for specific groups.
The key to expanding the application areas of data analytics is automating the analysis process using a machine learning technology, which allows a system to learn from input data and produce a model automatically. Industrial ICT Solutions Company has already achieved significant results with machine learning in the fields of image and speech recognition. We will also use machine learning to retrieve data that serves a particular purpose from a huge amount of diverse data. This is a prerequisite for taking advantage of big data. Moreover, we will tackle deep-learning and other leading-edge machine-learning methodologies in order to tune in to rapidly changing features of the environment and optimize an analysis accordingly.
In order to develop IoT solutions that will contribute to society, Industrial ICT Solutions Company will continually conduct research on advanced technology and deepen its knowledge of each industrial sector.

The amendment to the Personal Information Protection Act increases the importance of data anonymization technology

Recently, a proposed amendment to the Personal Information Protection Act was passed by the National Diet of Japan. One of the noteworthy points in the amendment is a provision for anonymized data, which is the first data type to be legislated in the world. Data anonymization makes it possible to utilize big data including personal information. However, too much anonymization reduces the usefulness of data. Although balancing the anonymity and usefulness of data is difficult, it is possible to make the best trade-off between these two conflicting goals. To address this issue, Industrial ICT Solutions Company has developed an original algorithm for k-anonymization.

Masanobu Koike
Ph.D. (in Engineering)
Senior Specialist
System & Security Technology Department
IoT Technology Center
Industrial ICT Solutions Company
Toshiba Corporation

Jointed Toshiba Corporation in 1998. Since then, he has been working on the R&D of information security technology, especially cryptography. In recent years, he has been committed to the development of privacy protection technology for enterprises that utilize personal data.

Anonymized data opens up new paths for big-data utilization

In 2015, a bill for amending the Personal Information Protection Act was passed by the National Diet of Japan. Several major revisions will shortly be made to the Act.

First, the amendment now provides a more clear definition of "personal information." Previously, the term "personal information" was defined as "information that can identify a specific individual." Now, the amendment provides that "personal information" includes "any codes that distinguish personal identity."

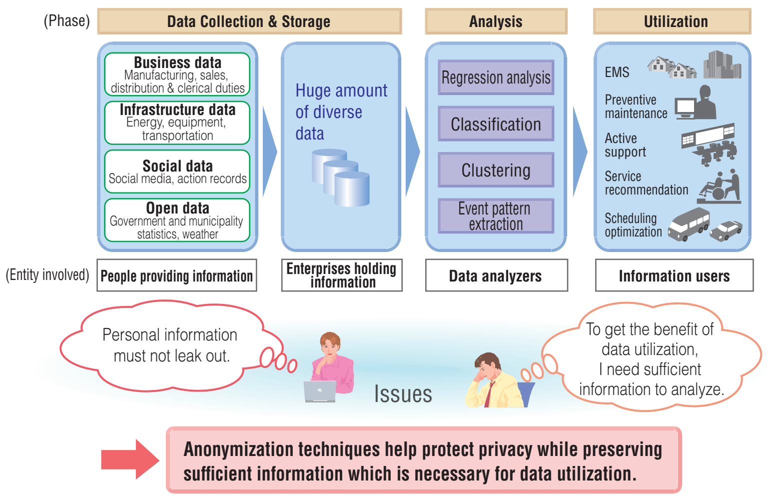
Second, the amendment defines a new concept called anonymized data, which refers to data that has been converted to remove personally identifiable information in an irreversible form. Data anonymization will open up new paths for the utilization of big data that contains personal information.

Additionally, the amendment provides for tighter regulations on name-list brokers and calls for the establishment of a government personal information protection committee.

The government's growth strategy is the driving force that lies behind these revisions. In the *Japan Revitalization Strategy: Revised in 2014* issued in June 2014, there is a clause that reads "To promote the utilization of personal data while protecting personal information and privacy in the era of big data." Data anonymization is crucially important to achieve this purpose.

In order to utilize a huge amount of personal information, it is necessary to meet the requirements for several phases. Four entities are involved in these phases (Figure 1):
1) consumers and other people who provide information, 2) business operators that hold personal information collected through their business activities, 3) specialist firms and other businesses that conduct data analysis, and 4) information users who wish to create value based on the analysis results.

Figure 1 **Data utilization and privacy protection**



In order to utilize big data, anonymization techniques are important to protect privacy while preserving sufficient information.

What is k-anonymization necessary for the IoT era?

If it were not for legal provisions for anonymized data, it would be impossible for one party to share both front-end and back-end phases with an external party. For example collecting and accumulating information as a front-end phase and analyzing and utilizing that information as a back-end phase. It is permitted for an information holder to analyze and utilize the information only within their own company. However, many business operators tend to give up using information due to a lack of analysis expertise. Suppose that you want to contract an outside party to analyze data collected by IoT sensors. According to the law now in force, you have to obtain approval from each individual included in the dataset before providing it to a third party. In some cases, it is difficult or impossible to ask for approval because IoT collects data without your knowledge and consent. In such cases, you cannot farm out the data analysis to an outside party.

In this situation, anonymized data works as a bridge between the front and back ends. In the case of the sensor data described above, you can provide anonymized data to a third party without approval from all parties involved as long as you have announced which pieces of information will be anonymized and how they will be provided to a third party. Therefore, data anonymization will make it possible for expert data analysts to work in a broader field, and help promote new IoT services using big-data analytics.

It should be noted, however, that data anonymization does not provide absolute protection for personal information. Data anonymization is a technique that makes it difficult, but not impossible, to distinguish personal identity.

For example, there is a possibility that individuals might be identified by combining one anonymized dataset with another. In the United States, a DVD rental firm announced a prize for a movie recommendation algorithm. To aid contestants, the DVD rental firm publicly disclosed an anonymized dataset containing movie ratings created by 500,000 subscribers (which consisted of movie titles, user IDs, dates and times, and reviews). However, subscribers were de-anonymized by matching this dataset to another on a movie site. Their privacy was compromised because many subscribers submitted similar posts to the movie site at around the same time as the ratings for the DVD rental firm.

Various R&D projects on anonymization technology have been conducted to develop solutions to this problem. One of the most commonly used anonymization methods is called k-anonymization. A dataset is said to have the k-anonymity property if the identity for a certain record cannot be narrowed down to fewer than k persons. For example, if k = 10, it is known that a record of a particular individual is one of these records of 10 people, but his/her identity cannot be further narrowed down. Thus, k-anonymization is widely used as a means of privacy protection.

A higher level of anonymization reduces the usefulness of data. Having a good balance between them is the key.

Figure 2 shows a table of non-anonymized medical records and a table of anonymized records with 2-anonymity. The non-anonymized records include specific data on age, gender and other attributes. In the anonymized records, the values of the attribute “Age” are replaced by broader categories, and the attribute “Place of Domicile” is generalized using prefecture names instead of city names. The values of the attribute “Medical Expenses” are also replaced by wider values.

For example, look for women aged 25–29 in Kanagawa who spent 5,000 yen or less on medical expenses. The anonymized database includes two persons who exactly match these attributes. It is therefore impossible to single out one person. The greater the k value, the higher the anonymity level. At present, it is difficult to say which k value is objectively appropriate to provide secure privacy

Figure 2 Example of applying k-anonymization

Before data anonymization

Age	Gender	Place of Domicile	Medical Expenses
25	F	Yokohama	5,000
28	F	Yokohama	1,000
26	M	Fuchu	800
28	M	Fuchu	2,000
32	M	Kawasaki	10,000
32	M	Kawasaki	6,000
33	M	Kawasaki	8,000
37	F	Ome	2,000
38	F	Fuchu	15,000
43	F	Kawasaki	30,000
44	F	Yokohama	20,000
46	M	Minato	3,000
49	M	Minato	100,000

k-anonymization (k = 2) →

After data anonymization

Age	Gender	Place of Domicile	Medical Expenses
25-29	F	Kanagawa	≤ 5,000
25-29	F	Kanagawa	≤ 5,000
25-29	M	Tokyo	≤ 5,000
25-29	M	Tokyo	≤ 5,000
30-34	M	Kanagawa	5,000-10,000
30-34	M	Kanagawa	5,000-10,000
30-34	M	Kanagawa	5,000-10,000
35-39	F	Tokyo	≤ 15,000
35-39	F	Tokyo	≤ 15,000
40-44	F	Kanagawa	20,000-30,000
40-44	F	Kanagawa	20,000-30,000
45-49	M	Tokyo	≤ 100,000
45-49	M	Tokyo	≤ 100,000

There are at least k (two) persons with the same data.

Anonymizing records of medical expenses with k = 2. When the age, place of domicile and medical expenses are generalized, the anonymized data includes at least two persons with the same data. (It is impossible to identify a single person.)

protection. For example, suppose that a retail chain wants to sell a set of anonymized purchase records of its members to a third party. The value of k has to be determined based on qualitative and ambiguous attributes such as the local characteristics of the retail chain's service areas. This is the puzzling part of k-anonymization.

It should be noted that the greater the k value, the higher the anonymity level and the less information becomes available. It reduces the usefulness of data.Thus, the question is how to strike a balance between usefulness of data and anonymity. Optimal usefulness and anonymity can never be achieved simultaneously. Nonetheless, it is still possible to make a good trade-off between the two opposing goals. That is our challenge of improving the k-anonymization technique.

Original algorithm that reduces the loss of information by 30%

The k-anonymization algorithm originally developed by Industrial ICT Solutions Company has succeeded in reducing the loss of information due to anonymization while increasing anonymity.

Our new k-anonymization algorithm reduces the loss of information by approximately 30% compared to the conventional technique at the same k value. This means, at the same anonymity level, the new algorithm provides more useful anonymized data than the conventional technique. Conversely, at the same usefulness level (i.e., the same level of information loss), the new algorithm provides a higher level of anonymity.

The innovative aspects of our new algorithm lie in the arrangement of data and the selection of attributes. A general principle is to categorize persons having as similar attributes as possible into the same group. Take the Age and Place of Domicile attributes for example. Suppose there are (k-1) persons aged 23 in Tokyo. In this case, it does not make sense to create a group of k people by adding a 30-year-old person in Hokkaido to this group. Instead of doing so, adding a 24-year-old person in Tokyo to the group helps reduce the overall loss of information.

Our k-anonymization algorithm and other anonymization techniques are utilized in various fields. For example, in the healthcare field, our techniques are used to anonymize data on physical checkups as well as the statements of medical expenses. At present, their application areas are being expanded to insurance and other industrial sectors.

Anonymization technology plays an important role in creating new value from big data. It is one of the essential technologies to enhance the *"things and experiences"* services of the Toshiba Group.

Published by: **Toshiba Corporation**
URL: http://www.toshiba.co.jp /iot/en
E-mail: INS-info@ml.toshiba.co.jp

The feature articles herein are translations from the Japanese edition of ***T-SOUL***, Volume 15, a Toshiba journal.