

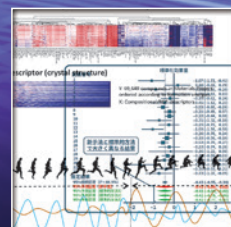
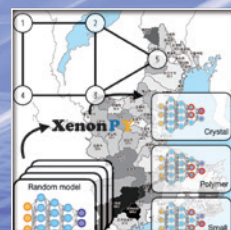
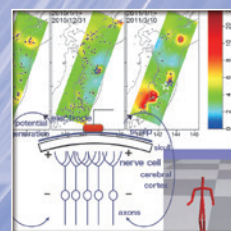
Research Organization of Information and Systems

# The Institute of Statistical Mathematics

2019-2020

# ISM

The Institute of Statistical Mathematics





# CONTENTS

■ Message from Director-General —	1
■ Institute Overview —————	2
■ Research Centers —————	4
■ ISM Projects —————	12
NOE (Network Of Excellence) Project	
Project for Fostering and Promoting Statistical Thinking	
■ Research Cooperation —————	16
■ Graduate School Program —————	18
■ Facilities and Equipment —————	21
■ Finance and Buildings —————	22
■ Organization —————	23
■ History —————	30



# Message from Director-General

---



I am Tsubaki Hiroe and I have taken office as the 12<sup>th</sup> Director-General of the Institute of Statistical Mathematics. While I carry on the tradition of the Institute that my predecessors have established and passed on, I intend to ensure that the Institute fulfill every role it is expected to play in society.

Today, the whole world is searching for the ideal state of data-driven social activities that utilize data as fundamental resources. Accordingly, people's lives are also rapidly changing. As a consequence, activities that generate intellectual and economic values using statistics as its core academic basis now have influence not only on the progress of various metric and mathematical sciences for academic researchers but also over social and economic activities for the public. It is not an exaggeration to say that a new era has come, where we should place statistical mathematics at the core of our society.

The Institute of Statistical Mathematics is the only research institute in Japan for statistical science and related mathematical sciences. Although I must admit that the Institute is not large enough to take full responsibility for such revolutionary social changes, we still have an important mission to perform,

precisely because the Institute is open to all universities for joint research.

Our task is to create a cooperative network to help promote the development and social expansion of so-called "data science," which focuses on the process of generating intellectual value. Furthermore, we have to establish a continuous management cycle of the network that the Institute and its collaborators have created, which should be evaluated by the degree of contribution to the academic world, a data-driven society, and the development of the next-generation of human resources. This is the mission that has been carried out by my predecessors, and I am humbly taking on the task at my own discretion.

This task can be formulated in my Director-General's policy as follows.

First, our goal is to create a research system where we provide joint support with researchers both in Japan and abroad to study statistical mathematic sciences that support each phase of the process for generating intellectual value as well as the common elements that underlie these phases. Through this system, we intend to deepen our understanding of basic mathematical science for data science. To this end, I will work with related communities to create a roadmap for statistical sciences, mathematic sciences, and basic mathematics for data science.

Second, I intend to determine how statistical mathematics, statistical mathematic sciences, and other academic areas should be integrated into the generative process of intellectual value in cooperation with prominent data scientists from the government, industry, and academia. Through this mechanism, I plan to help advance the academic and industrial world in Japan to make them ready for a data-driven society. To realize this, I will develop both an international and a domestic network within the NOE (network of excellence) project. While the international network aims to deepen cross-disciplinary research through core mathematics researchers, the domestic network aims to develop core human resources of statistical mathematics among intra-disciplinary researchers and businessmen and advance domestic intra-disciplinary activities. Given these clearly distinctive aims, the two networks should be able to exchange information whenever necessary and develop together into efficient networks.

Third, I plan to create a mechanism where industry, government, and academia collaboratively cultivate world class data scientists who can provide a foundation for the data science era, from the next generation of researchers, or the general public. Through this mechanism, we can develop a substantial stratum of professional data scientists and thereby contribute to the advancement of the data-driven era. As a first step, I intend to identify educators and their activities in developing statistical human resources in Japan.

In order to carry out these plans and meet the expectations of our future society, nothing is more important than creative and spontaneous activities performed by each staff member at the Institute. My main duty as Director-General is to encourage such initiative. The growth of our staff through self-realization is the most important resource for the development of the Institute and the communities that support us.

The Institute of Statistical Mathematics marked its 75th anniversary in June of this year. I truly hope that, by the 100th anniversary of the Institute, small ripples of our academic and social contributions will make big positive waves for the people in a data-driven society and that all our staff will do their best to realize that goal. I appreciate your continued guidance and encouragement.

**Hiroe Tsubaki**  
Director-General  
The Institute of Statistical Mathematics

# Institute Overview

## Basic Research

### Department of Statistical Modeling

The Department of Statistical Modeling works on structural modeling of physical phenomena related to numerous factors, and conducts research on model-based statistical inference methodologies. By means of model-based prediction and control, modeling of complex systems, and data assimilation, the department aims to contribute to the development of modeling intelligence in many fields.

#### ■ Prediction and Control Group

The Prediction and Control Group works on the development and evaluation of statistical models, which function effectively in terms of prediction and control of phenomena, decision making, and scientific discoveries. These efforts involve data analysis and modeling related to phenomena that vary across time and space.

#### ■ Complex System Modeling Group

The Complex System Modeling Group conducts studies aimed at discovering the structures of complex systems, such as nonlinear systems and hierarchical networks, through statistical modeling. For these purposes, the group also considers Monte Carlo simulations, discrete mathematics, and computer science.

#### ■ Data Assimilation Group

The Data Assimilation Group works on the development of data assimilation techniques, which are procedures aimed at combining information derived from large amounts of observations and a numerical simulation model. By developing computational algorithms and high-performance parallel computing systems, the group aims to build a next-generation simulation model that can predict the future in real time.

### Department of Statistical Data Science

The Department of Statistical Data Science conducts research on data design methods aimed at managing uncertainty and incompleteness of information, quantitative methods for evidence-based practice, and related data analysis methods. Moreover, the department investigates methods for inferring the latent structures in target phenomena from observation data.

#### ■ Survey Science Group

The Survey Science Group promotes research on the design of statistical surveys, development of statistical analysis methods on survey data, and applications. By exploring complex phenomena in various fields, the group also aims to contribute to practical applications in academia and policy-making through social surveys.

#### ■ Metric Science Group

The Metric Science Group conducts research aimed at identifying and evaluating statistical evidence through quantification of phenomena that have not been measured thus far, as well as efficient information extraction from large databases. The group investigates related methods and develops methods for analyzing the collected data. By working on applied research in various fields of real science, the group aims to advance practical, applied, statistical mathematical research based on evidence.

#### ■ Structure Exploration Group

The Structure Exploration Group conducts research on statistical science aimed at inferring the latent “structure” behind various target phenomena in biology, physics, and social science, based on observational data. The group focuses on machine learning, Bayesian reasoning, experimental design methods, and spatial-temporal analysis methods to investigate micro/meso/macrosopic and spatial-temporal dynamic structures in target phenomena.

### Department of Statistical Inference and Mathematics

The Department of Statistical Inference and Mathematics carries out research into general statistical theory, statistical learning theory, optimization, and algorithms for statistical inference.

#### ■ Mathematical Statistics Group

The Mathematical Statistics Group is concerned with aspects of statistical inference theory, modeling of uncertain phenomena, stochastic processes and their application to inference, probability and distribution theory, and the related mathematics.

#### ■ Learning and Inference Group

The Learning and Inference Group develops statistical methodologies to describe the stochastic structure of data mathematically and clarify the potential and the limitations of the data theoretically.

#### ■ Mathematical Optimization Group

The Mathematical Optimization Group focuses on mathematical theory and practical applications of optimization and computational algorithms together with underlying numerical or functional analysis and discrete mathematics.



---

## NOE-type Research

---

### Risk Analysis Research Center

Risk Analysis Research Center is pursuing a scientific approach to managing uncertainties and risks in society, which have increased with the growing globalization of society and economy. Our research projects are mainly seismology, finance, resources, environmentology, database development and risk mathematics. The Center also manages a network organization for risk analysis to facilitate research collaboration across different organizations with the common goal of creating a safe and resilient society.

### Research Center for Statistical Machine Learning

Machine learning is a research field associated with autonomous systems that can learn their behavior from data. This field is based on both the statistical science concerning inference from data and computer science concerning efficient algorithms. It can be applied to broad disciplines ranging engineering and information science to natural science, such as robotics and brain sciences. Our research center aims at supporting the academic community of this field, as well as producing influential research through various joint projects.

### Data Science Center for Creative Design and Manufacturing

We aim to foster new scientific methods for innovative design and manufacturing. Various fields in manufacturing are now facing a revolutionary period. Population reduction and globalization are bringing dramatic changes in the industrial structure in Japan. Countries around the world has actively developed their growth strategies utilizing data science as a driving force. We have accumulated state-of-the-art technologies in data science here. We are devoted to foster and practice advanced methods in data science for design and manufacturing.

### Research Center for Medical and Health Data Science

Research Center for Medical and Health Data Science aims to facilitate statistical data science research that covers medical studies, drug developments, health care, and public health. Our research projects involve fundamental mathematics and computational science for medical applications, applied methodology for basic, clinical and social medicine, and modern technology such as artificial intelligence, machine learning, and big data analyses. Furthermore, our research center aims at constructing a research network of the academic community of this field, as well as offering advanced statistical education programs.

---

## Professional Development

---

### School of Statistical Thinking

The mission of the School of Statistical Thinking is to plan and implement various programs for statistical thinking, from extension courses to a professional development program. The researchers affiliated with the school are often involved with specific data analysis projects, which help them to gain hands-on knowledge of data science. We expect such an experienced researcher will play an active role as a number-cruncher, as a modeler, or as a project coordinator.

---

## Research Support

---

### Center for Engineering and Technical Support

The Center for Engineering and Technical Support assists academics and their collaborators in many ways: managing computer systems and networks, editing and publishing journals, maintaining the library, and managing tutorial programs.

- **Computing Facilities Unit** The Computing Facilities Unit is in charge of managing computer facilities and scientific software.
- **Computer Networking Unit** The Computer Networking Unit is responsible for computer networking and its infrastructure, and network security.
- **Information Resources Unit** The Information Resources Unit is responsible for maintaining a library and an electronic repository, and is in charge of planning statistical tutorial programs open to the public.
- **Media Development Unit** The Media Development Unit is in charge of publishing and editing of research results and PR brochures.

## Computational Algebra and Sampling

### ■ Motivating problems in data analysis

The book entitled “Environmental Statistics” written by Vic Barnett consists of five parts: extremes and outliers, sampling and monitoring, stimulus and response, standards and regulations, and spatial and temporal processes. They may be chosen to present typical problems emerging in data analysis at the interface of statistical methods and environmental subjects. In this article, with keeping these problems in mind, a research activity of application of computational algebra to data analysis is introduced.

### ■ Toric model

A standard statistical model of count data is the toric model, where the logarithm of the probability that an observation falls into a cell is given by a linear model of marginal effects and interactions, where each term corresponds to a face of a simplicial simplex. A model whose facets correspond to the maximal cliques of a graph is called a graphical model. Figures 1 and 2 are the graphical model of the two-way contingency table with independence and a graphical model, respectively. In Figure 1, the facets are the two vertexes corresponding to the rows and columns.

### ■ Sampler

Let us call the probability that a data has the same as or smaller than the probability of an observed

data from a model the tail probability. In statistical test, a model is rejected if the tail probability is small. For example, we test independence of stimulus and response by setting that the rows of a contingency table as the stimulus and the columns as the response. If a closed expression of the tail probability is not available, it is estimated by simulated data from the model. The simulator is called sampler and evaluation of the tail probability is the typical use. Sampling is straightforward if a model is simple as Figure 1. However, graphical models as Figure 2 do not have a good property, and it has been considered that sampling from them is difficult.

### ■ Computational algebra and sampling

In these 20 years, application of computational algebra to statistical science has been a hot topic. For a toric model, it was shown that approximate samplers with Markov chain Monte Carlo can be constructed by using Gröbner basis of the polynomial ring and such samplers have been constructed for various models. However, the author showed that in principle sampling from any toric model is possible by using Gröbner basis of differential operators. The details are available in the book “Partitions, Hypergeometric Systems and Dirichlet Processes in Statistics” published in 2018 from Springer.

*Shuhei Mano*

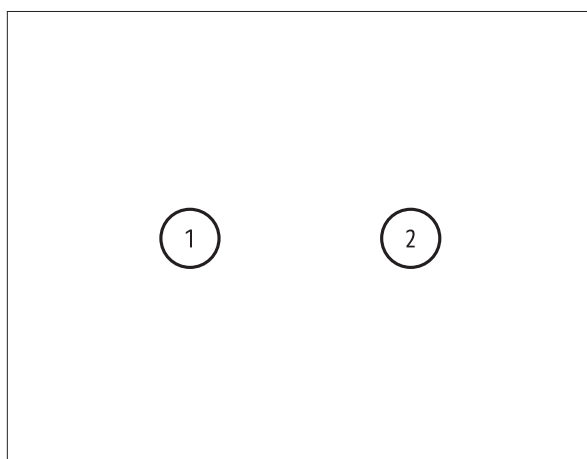


Figure1: The two-way contingency table with independence  $\{\emptyset, \{1\}, \{2\}\}$ .

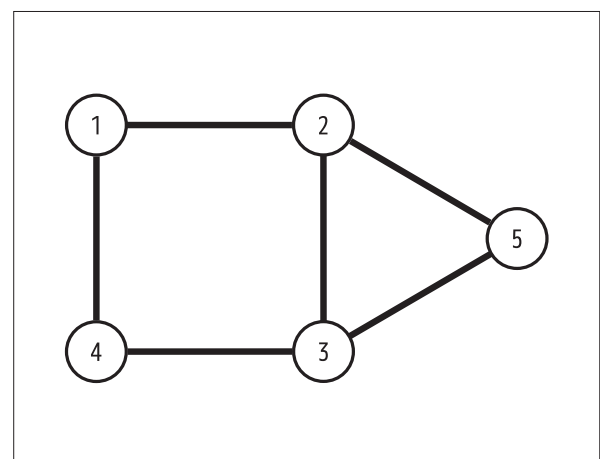


Figure2: A graphical model  $\{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}, \{2, 5\}, \{3, 5\}, \{2, 3, 5\}\}$ .

# Modeling of Recurrent Earthquakes and Long-Term Earthquake Forecast

## ■ Long-term forecast of earthquakes on active faults

There are so many active faults in Japan which have risks of earthquake disasters in the future. Although it is difficult to foretell earthquake occurrences so far, earthquakes in each active fault recur quasi-periodically and so their probability forecasts can be evaluated via statistical models. However, most active faults in Japan have only a few specified historical events, respectively, due to very long duration of earthquake recurrences. Such a small number of historical events may cause large errors in parameter inference. Especially, maximum likelihood estimates of aperiodicity parameters (coefficients of variation) in earthquake recurrence intervals have biases and may make significant biases in probability forecasts as well. Thus, we propose a Bayesian model to improve stability of parameter inference and predictive performance. Figure 1 shows estimates of aperiodicity parameters for respective active faults and we can find some spatial trends that, for example,

the estimates are relatively high in the center of Honshu island. Furthermore, when we compare the estimates with the active fault map in the right panel of Figure 1, we can find some correlation between the aperiodicity parameters and fault densities. Considering such spatial trends, we are challenging to improve long-term forecasts of active faults.

## ■ Monitoring of plate motion based on repeating earthquake activity

In recent years, development of seismic observation network enabled us to detect small repeating earthquakes on the same fault patch as well as large earthquakes on active faults. Especially, small earthquakes repeating on subduction zones of crustal plates attract much attentions as a key of earthquake prediction. Temporal accelerations of interplate slip rate called slow slips are often observed before megathrust earthquakes and can be monitored from changes in small repeating seismicity. Thus, we propose a new spatio-temporal stochastic model to estimate slip rate from repeating earthquakes on plate boundaries. Figure 2 shows spatio-temporal variation of slip rates estimated from small repeating earthquakes on the subduction zone of Pacific plate from 2008 to the 2011 Tohoku earthquake. Accelerated subduction can be seen in a wide region around the coseismic area of 2011 Tohoku earthquake. Using the proposed model, we are monitoring potential precursory changes in interplate slip rate for huge earthquakes.

**Shunichi Nomura**

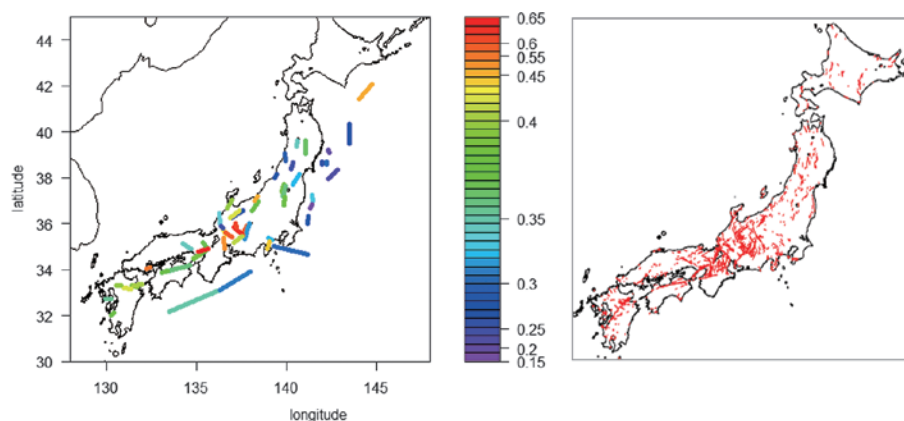


Figure 1: Estimates of aperiodicity parameters (left) and inland active fault map (right).

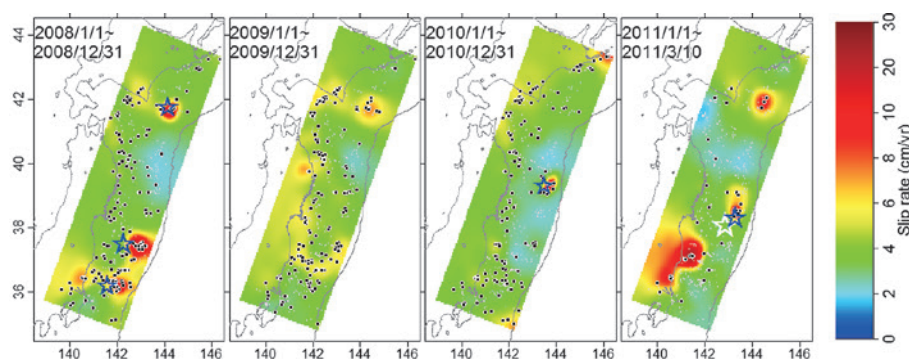


Figure 2: Estimated space-time distribution of interplate slip rate.

## Inducing “Motion” for Robots by a Statistical Learning

### ■ Relationship between robots and statistical model

As artificial intelligence based on statistical machine learning, i.e. AI, is changing the society, robotics is considered to be increasingly important in the future as one of the most important applications of machine learning. Robots are often thought to be simply operating mechanically, but in order to realize sophisticated intelligent operations on them, statistical models that can properly handle various uncertainties and inferences are necessary.

### ■ Learning “motion” for robots

As opposed to industrial robots that perform fixed movements, robots that cooperate with people to help them need to move in the same way as humans do and adjust their behavior appropriately. For robots, “motion” is nothing but the action of their motors, that is, the time series of the variation vectors of each joint (Figure 1). However, from the time series humans recognize motions such as “walking”, “pressing” and “lifting”. It is easy to understand if you think of a case of ballet. A ballet dance actually consists of a series of basic motions. In order to control movements at a high

level, it is necessary to automatically recognize these “motions” that correspond to “words” in natural language.

### ■ Gaussian process time series model and natural language

However, we do not hear from anyone what “walking” is. This is the same as knowing a word in agglutinative language, such as Chinese and Japanese, even if no one tells you that it is a word. In order to learn “motions” from data autonomously, we leverage a model for natural language (Mochihashi et al., 2009; Uchiumi et al., 2015) which statistically learns “words” purely from character strings of a language. Based on this, we have developed a statistical model to learn “motions” from a time series of various joint angle vectors, in collaboration with the University of Electro-Communications, Ochanomizu University, and the University of Tokyo.

From the viewpoint of statistics, this is a hidden semi-Markov model with Gaussian process output distributions which is estimated by a Bayesian forward-backward algorithm. By simultaneously estimating the

number of hidden states, i.e. how many “motions” there are, through a hierarchical Dirichlet process, it becomes possible to learn “motions” with high accuracy without any supervision (Figure 2). The result of this research was presented in the top international conference on robotics, IROS (Nagano et al., 2018).

### ■ Application to other areas

This basic technology that can learn “motion” from the time series of sensors at each joint has a wide range of possibilities besides robotics. For example, in joint research with the National Center of Neurology and Psychiatry, the application of this technology to the operation of a small ape, marmoset, will discover how “motions” will change in response to brain damage or diseases (to appear at JSAI 2019). Also, in dance science, I think that it will lead to automatic transcription and statistical modeling of dance scores.

**Daichi Mochihashi**

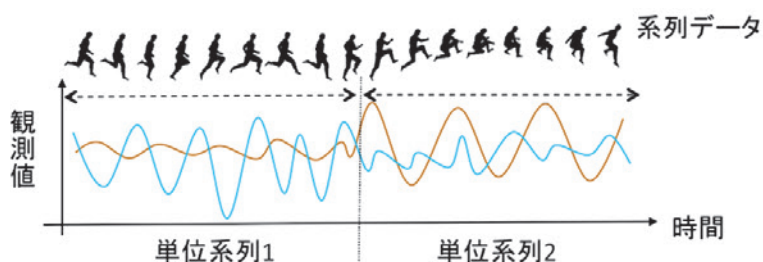


Figure 1: Sequence data corresponding to operations. We do not know in advance where and from where the unit of “motion” corresponds to.

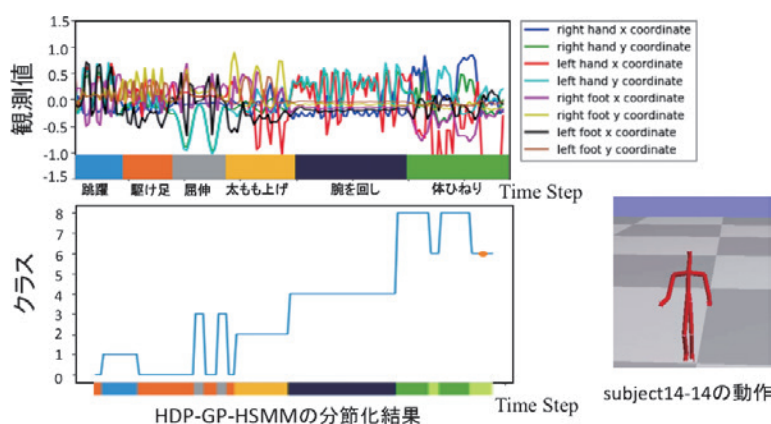


Figure 2: Recognition of “motion” by Gaussian process hidden semi-Markov model. Here, from the time series of the coordinates of the limbs, the “motion” and the number of their types are automatically inferred.



## Dipole Source Localization from EEG Signal By Particle Filter

### ■ Dipole estimation from EEG

In order to clarify the mechanism of the brain, diagnosis and treatment of diseases such as epilepsy, it is necessary to understand activity in the brain, and means to identify the activation position in the brain from biological signals is required. When nerve cells are activated, local currents are generated around them. Measurement of this local current on the scalp is EEG (electroencephalogram). It can be measured non-invasively and conveniently, and it has high temporal resolution, so it is being studied as a measurement method suitable for clinical application (Figure 1).

### ■ Problem formulation

As a method of identifying the active site in the brain from the EEG data, a method of replacing a bundle of local current in the brain with a current dipole, which is a 6-dimensional physical quantity having position and moment, is widely used. Using a model that takes into consideration the thickness and conductivity of the brain, skull, scalp, it is possible to calculate the potential (EEG) generated on the scalp by dipoles in the brain. However, determining the position, moment, and number of dipoles from the measured EEG is mathematically a difficult problem called inverse problem.

### ■ Approach

We formulated the dipole estimation problem from

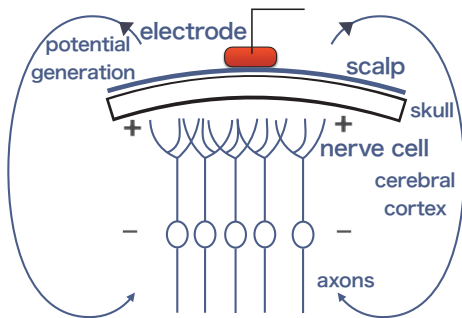


Figure 1: Mechanism of EEG.

transition model of dipoles	$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{z}_t$	$\mathbf{z}_t \sim \mathcal{N}(0, \mathbf{Q}_t)$	transition noise
	$K_t = K_{t-1} + \epsilon_t$	$\epsilon_t \in \{-1, 0, 1\}$	
observation model	$\mathbf{y}_t = f(\mathbf{x}_t) + \mathbf{w}_t$	$\mathbf{w}_t \sim \mathcal{N}(0, \Sigma_w)$	observation noise

Figure 2: State space model.

EEG as a state space model (Figure 2) and applied a Bayesian inference algorithm called particle filter. Specifically, at each point in time, approximating the probability distribution of the dipole position / moment, which is an unobservable internal state, with “ensemble of particles”. Although the dipole estimation method using such a particle filter has been proposed so far, the methodology for objectively determining the number of dipoles has not been sufficiently studied.

### ■ Information criterion and empirical Bayes method

We have developed a measure of validity of a model called information criterion on the premise that it is used in combination with a particle filter and to select the number of dipoles at each time. We also considered a process in which dipoles are generated or annihilated in the brain, and developed a method to estimate the number of dipoles by a method called empirical Bayes method (Figure 3). We applied the developed method to the actual data, results consistent with neurophysiological findings were obtained. Currently, we estimate the noise mixed in the time evolution process of the measurement system and particle by the framework of Bayesian inference to realize more accurate dipole estimation. We are aiming to applications such as clarify unknown brain function, brain computer interface, brain disease diagnosis.

*Hideitsu Hino*

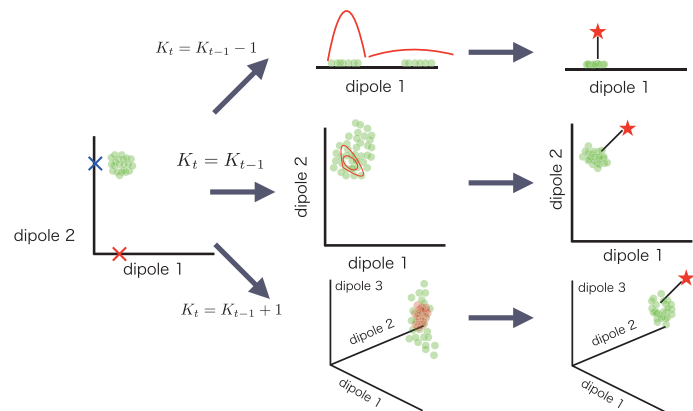


Figure 3: Process of estimating number/location/momentum of dipoles.

## Data Science Center for Creative Design and Manufacturing

### ■ Mission statement

We aim to foster new scientific methodologies for innovative design and manufacturing. Various fields in manufacturing are now facing a revolutionary period. Population reduction and globalization are bringing dramatic changes in the industrial structure in Japan, resulting in a rapid loss of global predominance in industry. Countries around the world have actively developed their growth strategies utilizing data science as a driving force, such as the Materials Genome Initiative in the US and the Industry 4.0. Following the global trend is no longer an effective way to survive in the intensive power game around the world. In 2017, the Institute of Statistical Mathematics has established a new research center - Data Science Center for Creative Design and Manufacturing. We have accumulated state-of-the-art technologies in data science here, for instance, machine learning, Bayesian inference, materials informatics, and so on. We are devoted to foster and practice innovative methods in data science for design and manufacturing through industry-academia collaboration.

### ■ Smart manufacturing

The developments of new materials depend largely on intuitions of highly experienced professionals, and time-consuming trial-and-error processes for laboratory synthesis and testing of designed materials based on computer simulation and experiments. On the other hand, recently, there are new attempts to substitute computational or real experiments in materials synthesis and testing by statistical models trained on given data. The enormous cost and time required in the characterization of material structures and physical properties has limited material studies within a small

set of candidates. For example, it has been proved that high throughput screening using techniques in data science significantly increases the chance of discovering innovative functional materials. This is a basic concept of smart manufacturing in the perspective of data science.

### ■ Creative design and manufacturing

We recognize the importance of being at the absolute leading edge position in the manufacturing industry. This cannot be done by data science alone. Most of the classical data science analysis tools are designed for interpolating predictions. Data science used to be a science of predictions based on pattern recognition from existing data. For example, we often assume that materials with similar chemical structure exhibit similar physical properties. However, by definition, a new material is not likely to be similar to any of the existing materials. Combination of experiment, theory and data science methods is an essential step to a new breakthrough in the current state. In other words, we adopt a stepwise approach to expand the region of accurate prediction of a statistical model. We achieve the goal by careful design of an optimal experiment or simulation schedule for new data points that efficiently improve an existing predictive model. We have accomplished preliminary success in materials science using an extrapolating prediction method based on a data science approach, and fostered new collaboration opportunities between academia and industry for the purpose of materials discovery. Our next step is to extend the application to various fields related to creative design and manufacturing.

*Ryo Yoshida*

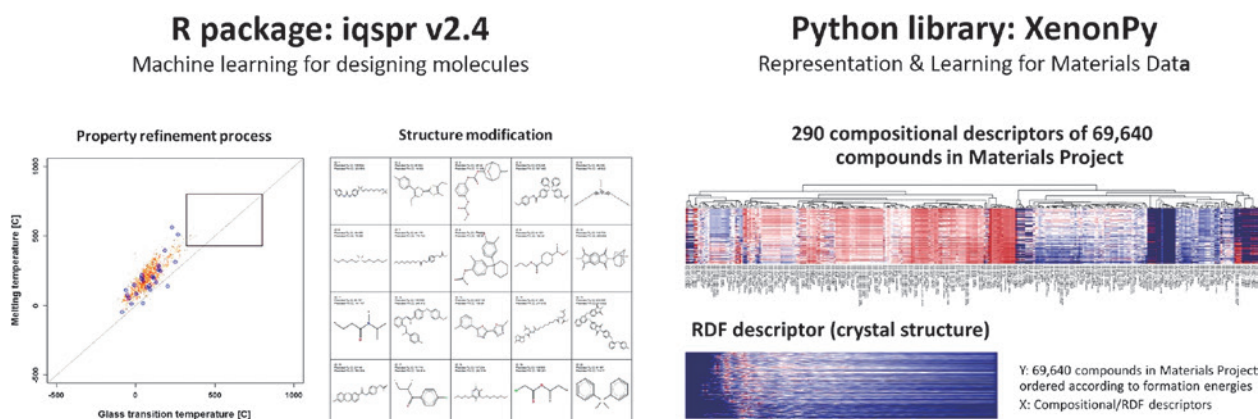


Figure: Software developments for materials informatics.

## XenonPy: The Key Foundation for New Generation Materials Science

### ■ Current status of materials informatics

Materials Genome Initiative, which began in 2011 under the presidency of Barack Obama, has triggered the Materials Informatics (MI) boom, i.e., the concept of using data science and machine learning methods to accelerate the discovery of high performance novel materials. By mapping the highly diverse materials space into a set of “descriptors” and developing a generator of material candidates, we are able to exploit the state-of-the-art machine learning technologies to solve the inverse material design problem. Despite the rapid development of many MI algorithms and toolboxes, successful discovery of truly novel materials is still very rare because of the difficulty to access a large data set for any MI applications in order to train a reliable predictive model. Such difficulty is mainly caused by (1) requirement of high cost and intensive labor, (2) conflict of interest between the industrial companies, and (3) highly diverse goals. The lack of data problem has become a bottleneck blocking the advancement of MI.

### ■ A solution for small data problem — transfer learning

Transfer learning is an indirect alternative to tack-

le the lack of data problem. Conceptually, we want to exploit or generalize knowledge learned from one task to another, for which human beings are naturally good at. For example, someone who knows how to play piano will be able to learn violin or flute faster than usual because the knowledge of playing different instruments are expected to be highly correlated. The same idea can be applied to MI in order to effectively train models with little data. However, we may not know the correlation between different material design tasks a priori. To benefit from the true power of transfer learning, it is necessary to gather a large library of pre-trained models that captures different knowledge in MI.

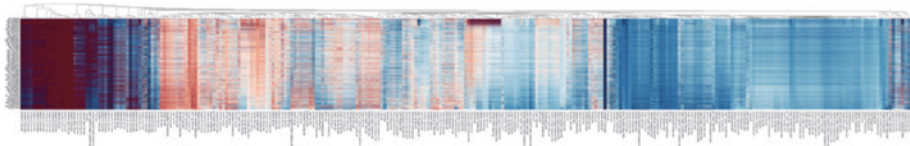
### ■ MI software — XenonPy

The Data Science Center for Creative Design and Manufacturing has developed an all-in-one package, called XenonPy, to provide easy-to-use tools for calculating various material descriptors, as well as machine learning model training. XenonPy unifies popular descriptors of both organic and inorganic compounds, including composition descriptor, structural descriptor, classical fingerprints, and neural fingerprints. Furthermore, the large collection of state-of-the-art

pre-trained models in XenonPy provides the best platform for transfer learning in various MI applications, such as, small molecules, polymers, crystals, etc.

*Hironao Yamada,  
Stephen Wu*

### Example of XenonPy descriptors



### List of models available in XenonPy

Model ID	Model Name	Model Type	Model Size	Model Version	Model Description	Model Status
001	Random Forest	Classification	1.5 MB	1.0.0	Random Forest model for classification	Available
002	Support Vector Machine	Classification	2.0 MB	1.0.0	Support Vector Machine model for classification	Available
003	Neural Network	Classification	3.0 MB	1.0.0	Neural Network model for classification	Available
004	Decision Tree	Classification	1.0 MB	1.0.0	Decision Tree model for classification	Available
005	Logistic Regression	Classification	1.0 MB	1.0.0	Logistic Regression model for classification	Available
006	Linear Regression	Regression	1.0 MB	1.0.0	Linear Regression model for regression	Available
007	Polynomial Regression	Regression	1.0 MB	1.0.0	Polynomial Regression model for regression	Available
008	Bayesian Linear Regression	Regression	1.0 MB	1.0.0	Bayesian Linear Regression model for regression	Available
009	Gaussian Process Regression	Regression	1.0 MB	1.0.0	Gaussian Process Regression model for regression	Available
010	Kernel Ridge Regression	Regression	1.0 MB	1.0.0	Kernel Ridge Regression model for regression	Available
011	Bayesian Ridge Regression	Regression	1.0 MB	1.0.0	Bayesian Ridge Regression model for regression	Available
012	Adaboost	Classification	1.0 MB	1.0.0	Adaboost model for classification	Available
013	Gradient Boosting	Classification	1.0 MB	1.0.0	Gradient Boosting model for classification	Available
014	XGBoost	Classification	1.0 MB	1.0.0	XGBoost model for classification	Available
015	LightGBM	Classification	1.0 MB	1.0.0	LightGBM model for classification	Available
016	CatBoost	Classification	1.0 MB	1.0.0	CatBoost model for classification	Available
017	AdaNet	Classification	1.0 MB	1.0.0	AdaNet model for classification	Available
018	Neural Network	Regression	3.0 MB	1.0.0	Neural Network model for regression	Available
019	Support Vector Machine	Regression	2.0 MB	1.0.0	Support Vector Machine model for regression	Available
020	Random Forest	Regression	1.5 MB	1.0.0	Random Forest model for regression	Available
021	Decision Tree	Regression	1.0 MB	1.0.0	Decision Tree model for regression	Available
022	Logistic Regression	Regression	1.0 MB	1.0.0	Logistic Regression model for regression	Available
023	Linear Regression	Regression	1.0 MB	1.0.0	Linear Regression model for regression	Available
024	Polynomial Regression	Regression	1.0 MB	1.0.0	Polynomial Regression model for regression	Available
025	Bayesian Linear Regression	Regression	1.0 MB	1.0.0	Bayesian Linear Regression model for regression	Available
026	Gaussian Process Regression	Regression	1.0 MB	1.0.0	Gaussian Process Regression model for regression	Available
027	Kernel Ridge Regression	Regression	1.0 MB	1.0.0	Kernel Ridge Regression model for regression	Available
028	Bayesian Ridge Regression	Regression	1.0 MB	1.0.0	Bayesian Ridge Regression model for regression	Available
029	Adaboost	Regression	1.0 MB	1.0.0	Adaboost model for regression	Available
030	Gradient Boosting	Regression	1.0 MB	1.0.0	Gradient Boosting model for regression	Available
031	XGBoost	Regression	1.0 MB	1.0.0	XGBoost model for regression	Available
032	LightGBM	Regression	1.0 MB	1.0.0	LightGBM model for regression	Available
033	CatBoost	Regression	1.0 MB	1.0.0	CatBoost model for regression	Available
034	AdaNet	Regression	1.0 MB	1.0.0	AdaNet model for regression	Available
035	Neural Network	Classification	3.0 MB	1.0.0	Neural Network model for classification	Available
036	Support Vector Machine	Classification	2.0 MB	1.0.0	Support Vector Machine model for classification	Available
037	Random Forest	Classification	1.5 MB	1.0.0	Random Forest model for classification	Available
038	Decision Tree	Classification	1.0 MB	1.0.0	Decision Tree model for classification	Available
039	Logistic Regression	Classification	1.0 MB	1.0.0	Logistic Regression model for classification	Available
040	Linear Regression	Classification	1.0 MB	1.0.0	Linear Regression model for classification	Available
041	Polynomial Regression	Classification	1.0 MB	1.0.0	Polynomial Regression model for classification	Available
042	Bayesian Linear Regression	Classification	1.0 MB	1.0.0	Bayesian Linear Regression model for classification	Available
043	Gaussian Process Regression	Classification	1.0 MB	1.0.0	Gaussian Process Regression model for classification	Available
044	Kernel Ridge Regression	Classification	1.0 MB	1.0.0	Kernel Ridge Regression model for classification	Available
045	Bayesian Ridge Regression	Classification	1.0 MB	1.0.0	Bayesian Ridge Regression model for classification	Available
046	Adaboost	Classification	1.0 MB	1.0.0	Adaboost model for classification	Available
047	Gradient Boosting	Classification	1.0 MB	1.0.0	Gradient Boosting model for classification	Available
048	XGBoost	Classification	1.0 MB	1.0.0	XGBoost model for classification	Available
049	LightGBM	Classification	1.0 MB	1.0.0	LightGBM model for classification	Available
050	CatBoost	Classification	1.0 MB	1.0.0	CatBoost model for classification	Available
051	AdaNet	Classification	1.0 MB	1.0.0	AdaNet model for classification	Available
052	Neural Network	Regression	3.0 MB	1.0.0	Neural Network model for regression	Available
053	Support Vector Machine	Regression	2.0 MB	1.0.0	Support Vector Machine model for regression	Available
054	Random Forest	Regression	1.5 MB	1.0.0	Random Forest model for regression	Available
055	Decision Tree	Regression	1.0 MB	1.0.0	Decision Tree model for regression	Available
056	Logistic Regression	Regression	1.0 MB	1.0.0	Logistic Regression model for regression	Available
057	Linear Regression	Regression	1.0 MB	1.0.0	Linear Regression model for regression	Available
058	Polynomial Regression	Regression	1.0 MB	1.0.0	Polynomial Regression model for regression	Available
059	Bayesian Linear Regression	Regression	1.0 MB	1.0.0	Bayesian Linear Regression model for regression	Available
060	Gaussian Process Regression	Regression	1.0 MB	1.0.0	Gaussian Process Regression model for regression	Available
061	Kernel Ridge Regression	Regression	1.0 MB	1.0.0	Kernel Ridge Regression model for regression	Available
062	Bayesian Ridge Regression	Regression	1.0 MB	1.0.0	Bayesian Ridge Regression model for regression	Available
063	Adaboost	Regression	1.0 MB	1.0.0	Adaboost model for regression	Available
064	Gradient Boosting	Regression	1.0 MB	1.0.0	Gradient Boosting model for regression	Available
065	XGBoost	Regression	1.0 MB	1.0.0	XGBoost model for regression	Available
066	LightGBM	Regression	1.0 MB	1.0.0	LightGBM model for regression	Available
067	CatBoost	Regression	1.0 MB	1.0.0	CatBoost model for regression	Available
068	AdaNet	Regression	1.0 MB	1.0.0	AdaNet model for regression	Available
069	Neural Network	Classification	3.0 MB	1.0.0	Neural Network model for classification	Available
070	Support Vector Machine	Classification	2.0 MB	1.0.0	Support Vector Machine model for classification	Available
071	Random Forest	Classification	1.5 MB	1.0.0	Random Forest model for classification	Available
072	Decision Tree	Classification	1.0 MB	1.0.0	Decision Tree model for classification	Available
073	Logistic Regression	Classification	1.0 MB	1.0.0	Logistic Regression model for classification	Available
074	Linear Regression	Classification	1.0 MB	1.0.0	Linear Regression model for classification	Available
075	Polynomial Regression	Classification	1.0 MB	1.0.0	Polynomial Regression model for classification	Available
076	Bayesian Linear Regression	Classification	1.0 MB	1.0.0	Bayesian Linear Regression model for classification	Available
077	Gaussian Process Regression	Classification	1.0 MB	1.0.0	Gaussian Process Regression model for classification	Available
078	Kernel Ridge Regression	Classification	1.0 MB	1.0.0	Kernel Ridge Regression model for classification	Available
079	Bayesian Ridge Regression	Classification	1.0 MB	1.0.0	Bayesian Ridge Regression model for classification	Available
080	Adaboost	Classification	1.0 MB	1.0.0	Adaboost model for classification	Available
081	Gradient Boosting	Classification	1.0 MB	1.0.0	Gradient Boosting model for classification	Available
082	XGBoost	Classification	1.0 MB	1.0.0	XGBoost model for classification	Available
083	LightGBM	Classification	1.0 MB	1.0.0	LightGBM model for classification	Available
084	CatBoost	Classification	1.0 MB	1.0.0	CatBoost model for classification	Available
085	AdaNet	Classification	1.0 MB	1.0.0	AdaNet model for classification	Available
086	Neural Network	Regression	3.0 MB	1.0.0	Neural Network model for regression	Available
087	Support Vector Machine	Regression	2.0 MB	1.0.0	Support Vector Machine model for regression	Available
088	Random Forest	Regression	1.5 MB	1.0.0	Random Forest model for regression	Available
089	Decision Tree	Regression	1.0 MB	1.0.0	Decision Tree model for regression	Available
090	Logistic Regression	Regression	1.0 MB	1.0.0	Logistic Regression model for regression	Available
091	Linear Regression	Regression	1.0 MB	1.0.0	Linear Regression model for regression	Available
092	Polynomial Regression	Regression	1.0 MB	1.0.0	Polynomial Regression model for regression	Available
093	Bayesian Linear Regression	Regression	1.0 MB	1.0.0	Bayesian Linear Regression model for regression	Available
094	Gaussian Process Regression	Regression	1.0 MB	1.0.0	Gaussian Process Regression model for regression	Available
095	Kernel Ridge Regression	Regression	1.0 MB	1.0.0	Kernel Ridge Regression model for regression	Available
096	Bayesian Ridge Regression	Regression	1.0 MB	1.0.0	Bayesian Ridge Regression model for regression	Available
097	Adaboost	Regression	1.0 MB	1.0.0	Adaboost model for regression	Available
098	Gradient Boosting	Regression	1.0 MB	1.0.0	Gradient Boosting model for regression	Available
099	XGBoost	Regression	1.0 MB	1.0.0	XGBoost model for regression	Available
100	LightGBM	Regression	1.0 MB	1.0.0	LightGBM model for regression	Available

### Concept of transfer learning

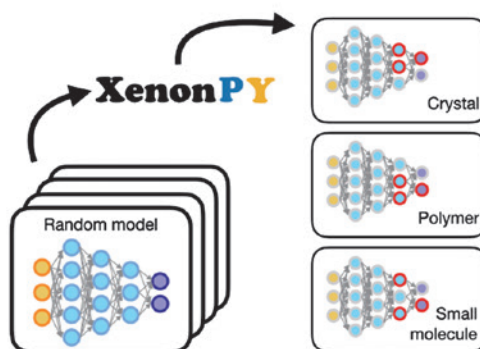


Figure: Functions of XenonPy.



## Prediction Intervals for Random-Effects Meta-Analysis

### ■ Meta-analysis

Meta-analysis is an important tool for synthesizing the results of a set of related studies. A common objective of meta-analysis is to estimate the treatment effect and its confidence interval, and random-effects models that assume the true treatment effects differ for each study have been widely applied. The average treatment effect across all studies and its confidence interval have been used together with heterogeneity measures that are important for generalizability. Recently, it is recommended that a prediction interval that can be interpreted as the range of the predicted true treatment effect in a new study would be reported alongside a confidence interval and heterogeneity measure.

### ■ Problem in conventional prediction intervals

Conventional prediction intervals have a problem of

severe under-estimation of the range when the number of studies is small, as is often the case in clinical practice. The reason of the under-estimation has not been studied. No explicit solution to this problem has been found thus far.

### ■ The proposed prediction interval

We discussed the reason of the under-estimation and proposed a new prediction interval that is valid under more general and realistic settings of meta-analyses in medical research. Conventional prediction intervals apply a large sample approximation, but this approximation has a detrimental impact on the estimation performance. Therefore, we develop an alternative that is not based on such approximation and show by simulations that its performance is better than conventional prediction intervals. An application to a published

random-effects meta-analyses concluded that substantially different result and interpretation might be obtained from the proposed and conventional prediction intervals, as shown in Figure 2. Conventional prediction intervals were narrower than the proposed prediction interval, which means that an under-estimation occurred in conventional prediction intervals. We should be cautious in using and interpreting conventional approaches. An R package implementing the new method is available (<https://cran.r-project.org/package=pimeta>, CRAN Task View: MetaAnalysis).

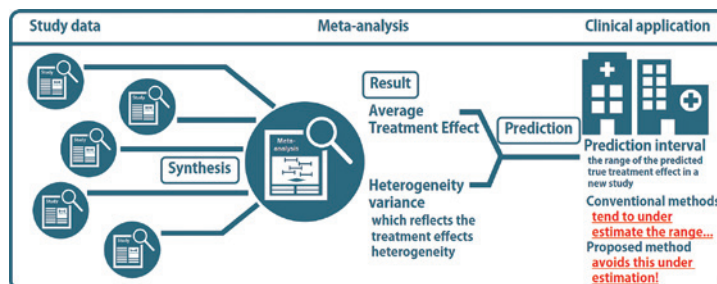


Figure 1: The random-effects meta-analysis and prediction interval.

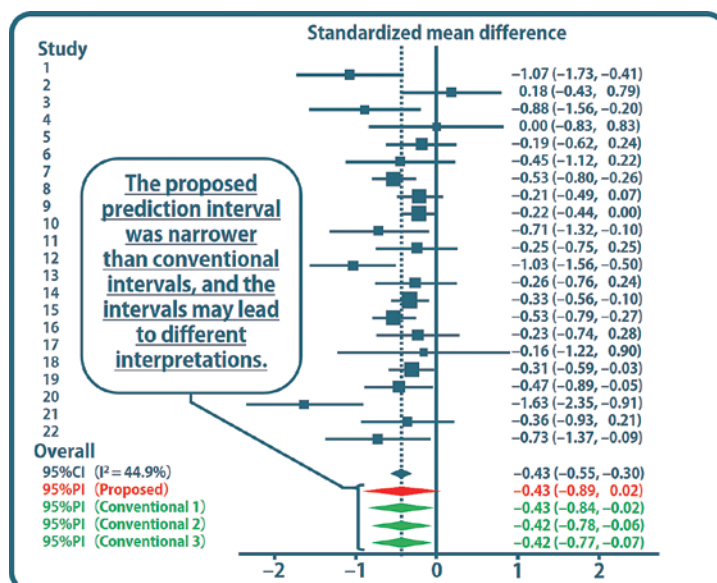


Figure 2: An application to a published random-effects meta-analysis that included 22 studies comparing the treatment effect of antidepressants on reducing pain in patients with fibromyalgia syndrome (Häuser et al. JAMA 2009; 301: 198–209; Riley et al. BMJ 2011; 342: d549).

### ■ Statistical methods for medical and health data science

Since results of meta-analyses have been utilized for formulation of health care policies and medical guidelines, the proposed method can provide more precise scientific evidence and useful information for clinical practice. This is a collaborative study with Professor Toshi A. Furukawa, Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health. We aim to provide more precise scientific evidence for clinical practice by developing statistical methods for medical and health data science.

**Kengo Nagashima**

## Quantifying Environmental Characteristics in Social Epidemiology — Devising a Method for Estimating “Alleyss”

### ■ Quantifying environmental characteristics

In the field of social epidemiology, physical and mental health problems that often occur in certain groups are examined in relation to their living environment. In order to test the hypothesis that particular environmental characteristics affect health, it is first necessary to accurately quantify these characteristics by devising appropriate indices. In this research, we will devise and evaluate an index that is useful for probing the relationship between space-structure characteristics (such as the topography, shape of roads, and density of houses in each municipality) and suicide incidence.

Previous research has shown that the ability of residents to seek help is significantly higher in areas with rare suicide incidence. One notable space-structure characteristic common to areas with rare suicide incidence

is the high number of alleys in these areas. As such, it is hypothesized that alleys promote casual conversation and exchange of information between residents that result in their individual and

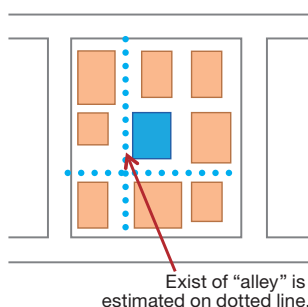


Figure 1: Description of estimating the “alley”.

social problems trickling out subconsciously and becoming manifest, which in turn allows for timely external intervention.

### ■ Proposing a method for estimating alleys

We obtained the assistance of civil and urban engineering researchers and map companies to devise a method and the parameters required to estimate the number of alleys in areas where geospatial data and GIS (Geographic Information System) suggest the existence of alleys.

According to the Building Standards Act, any road wider than 4m around a building site must have at least 2m of it connected to the site itself. This implies that buildings that do not appear to be connected to a road in the geospatial data are highly likely to have alley access. In other words, by cross-checking this information obtained from geospatial data with GIS, we are able to conclude if alleys exist in the immediate vicinity of these buildings of interest (Figure 1).

### ■ Relationship between alley density and suicide-SMR

An analysis of the relationship between the alley density and suicide-SMR (standardized mortality ratio, averaged over 30 years) in each of the 69 municipalities of Mie prefecture reveals a significant negative correlation. When we conducted multiple regression analysis using an interaction term to represent the strong correlation between alley density and coastal regions, the results showed that even for municipalities in the coastal regions, a higher number of alleys significantly reduce suicide-SMR (Figure 2).

*Mayumi Oka*

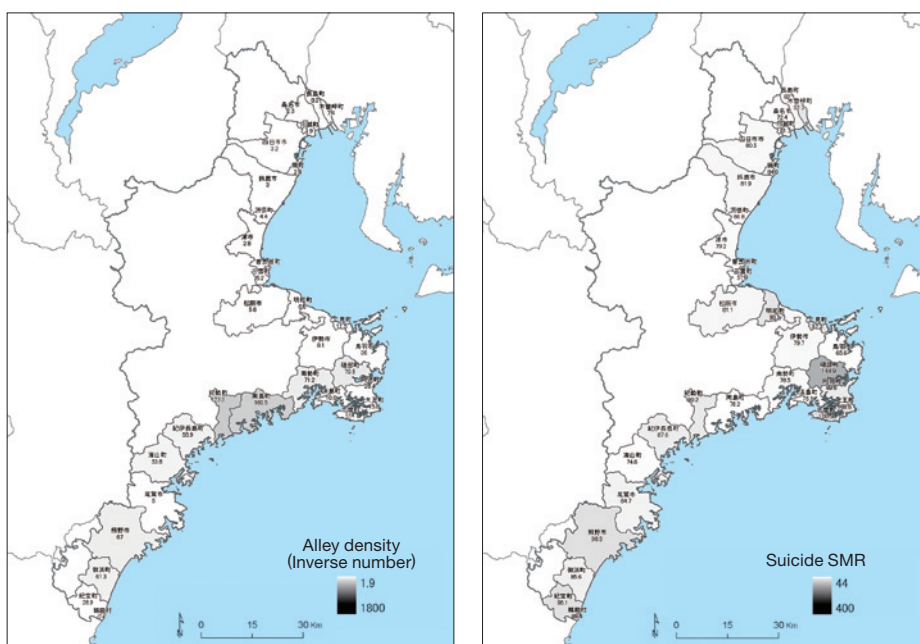


Figure 2: Alley density and Suicide-SMR.

# NOE (Network Of Excellence) Project

## Building a Framework for Advancing Strategic Research and Pursuing a New Approach to Collaborative Research

### ■ Research and Educational Activities as a Biaxial Structure

The Institute of Statistical Mathematics (ISM) pursues research and education along two lines of basic research, as well as NOE (Network Of Excellence)-type research and professional development. Research and education efforts are conducted by basic research departments along a horizontal axis, and the NOE-type research centers and the school for professional development are organized along a vertical axis. By its nature, the basic research departments (along the horizontal axis) cuts across and links various disciplines, with the goal of developing tools for interdisciplinary research. The field of statistical mathematics must itself evolve to meet the changing needs of society and the data environment, and is therefore constantly evolving as a field of study. At the same time, there are approaches and directions that have remained unchanged as the field evolves. For that reason, we have chosen not to refer to these efforts as “fundamental research” or “foundational research,” but instead as “basic research”, in order reflect both the fixed and evolving qualities of statistical mathematics. There are three basic research departments: Statistical Modeling, Statistical Data Science, and Statistical Inference and Mathematics. All tenured research staff in ISM are assigned in principle to one of these basic research departments. These departments engage in cutting-edge research aimed at developing methodologies for rational prediction and decision making, based on data and existing knowledge. We regularly assess whether our research system is functioning effectively from the

viewpoints of research trends and prospects in statistical mathematics and we reorganized several groups and transferred research staff within basic research departments on April 1, 2018.

On the other hand, the NOE-type research centers and the school for professional development (along the vertical axis) are staffed by permanent researchers within ISM, project professors/researchers (post-doctoral staff), and visiting professors and researchers. There are four NOE-type research centers: the Risk Analysis Research Center, the Research Center for Statistical Machine Learning, the Data Science Center for Creative Design and Manufacturing, and the Research Center for Medical Health Data Science. These centers conduct research activities that interface statistical mathematics with individual scientific disciplines in order to find solutions to urgent social problems. And the School of Statistical Thinking which is for professional development also provides multiple programs aimed at fostering and promoting statistical thinking. Please refer to the page “Project of Fostering and Promoting Statistical Thinking” for more details.

### ■ NOE (Network Of Excellence) Project

In accordance with the second medium-term plan for Research Organization of Information and Systems (ROIS, ISM's parent organization), ISM had set as a goal the establishment of NOEs (Networks Of Excellence) in statistical mathematics. This FY: 2019-2020 is the 10th year since the start of the project. Initially, we established NOEs in Risk Research, Next-Generation Simulation, Survey Science, Statistical Machine Learning, and Service Science.

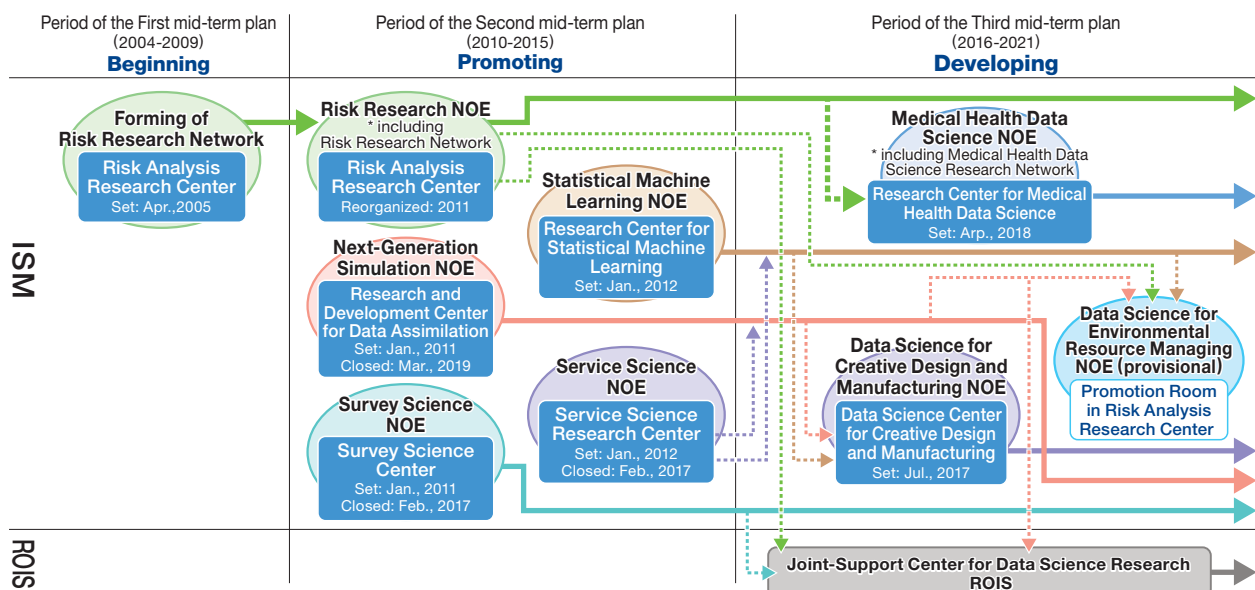


Figure 1: Brief history of the NOE (Network Of Excellence) Project.



We reorganized domains of NOEs or NOE-type Research Centers, considering the needs of each community as well as modern society as a whole. We closed the Survey Science Center, whose projects were transferred to the research activities of the Center for Social Data Structuring in the Joint Support Center for Data Science of ROIS. At the same time, the Service Science Research Center was also closed, and its projects were distributed according to their methodologies to other NOE-type Research Centers at the end of February 2017. In July 2017, we launched the Data Science Center for Creative Design and Manufacturing, and initiated the establishment of “Data Science for Creative Design and Manufacturing NOE”. The Research Center for Medical Health Data Science was set up in April, 2018. The Center, which was launched from one of the projects of the Risk Analysis Research Center, is promoting research and educational activities in Medical Health Data Science domain based on the knowledge and research networks that ISM has obtained in statistical mathematics both inside and outside Japan (Figure 1). As of this year, ISM is promoting the NOE Project in six fields: Risk Research, Next-Generation Simulation, Survey Science, Statistical Machine Learning, Data Science for Creative Design and Manufacturing, and Medical Health Data Science (Figure 2).

### ■ Future Perspective of NOE Activities

This NOE Project is a core theme of ISM. To fulfill the goal of establishing new scientific methodologies in a knowledge-based society, in which the importance of knowledge goes beyond merely solving individual problems, NOE activities are being systematically pursued under the unified guidelines formulated by the Managing Committee of the NOE Project. We also commission experts from the industrial, academic, and government sectors to be members of the Advisory Board of the NOE Project, and their advice helps us promote the project much more effectively. The NOE Project, which is made possible by ISM’s special focus in the cross-disciplinary field of “statistical mathematics,” is attracting strong support from each of these communities. On the basis of this project, ISM, as an Inter-University Research Institute, will be providing the industrial, academic, and gov-

ernment communities with further opportunities for joint usage (of facilities) and joint research. ISM continues to promote this NOE Project. As described above, each NOE research center serves as core hubs in their respective fields. ISM is promoting the signing of MOUs (Memorandum of Understanding) with research organizations within Japan and overseas, and the number is increasing each year, including ones that span multiple research fields. The goal of ISM is general research in statistical mathematics, which is in demand by various research fields in both the humanities and science. ISM must respond flexibly to requests from each community and contribute to them. Reorganization over recent years, based on such needs, has aimed at much deeper promotion of the Project.

We held the International External Evaluation last year and received valuable opinions regarding ISM’s position, research activities, and presence in an international context, enabling us to examine the appropriateness of our research structure, as well as other issues. ISM is expanding and developing this NOE projects with the aim of establishing new scientific methodologies (“Fourth Paradigm”), creating new research disciplines, and developing new styles of joint research. For up-to-date information on the research structures or activities of the Project, please visit the website. We very much appreciate your continued support for this Project.

<https://www.ism.ac.jp/noe/project/en/>

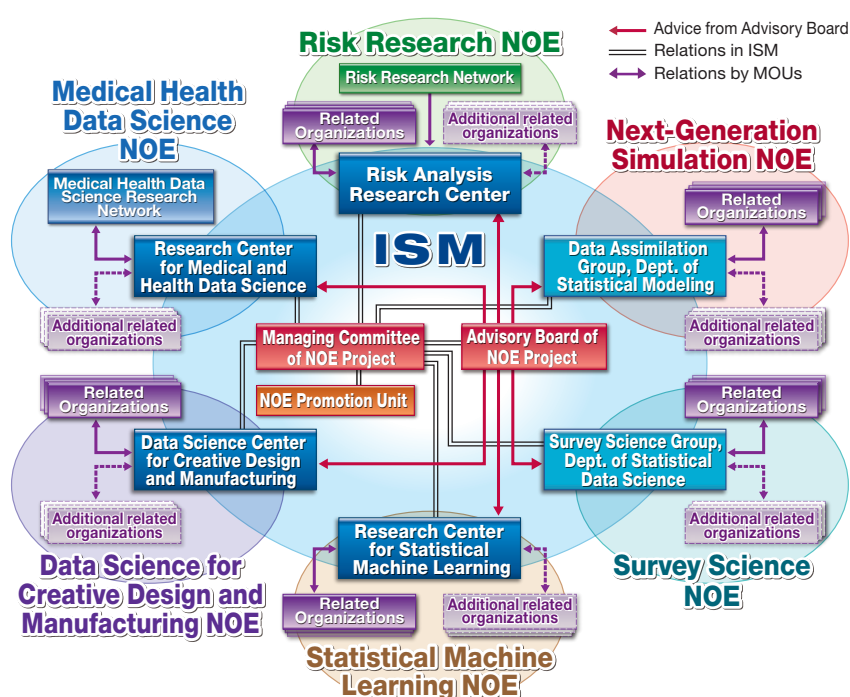


Figure 2: Relationship diagram of the NOE (Network Of Excellence) Project.

# Project for Fostering and Promoting Statistical Thinking

---

Rapid development of information and communication technology has led to the explosion of data. Now surrounded by “Big Data”, everybody is expected to “think statistically”. More than ever, there is a need for data scientists who can handle such big data and are able to extract useful knowledge from it. Meanwhile, Japanese higher education is exhibiting a deplorable lack of production capacity in terms of data scientists. This can be accounted for by the fact that no academic institution other than ISM has a Ph. D. course in statistics and the small number of statisticians in academia are isolated from each other, being scattered over various disciplines. Hoping to gain a little traction on this problem, ISM established the School of Statistical Thinking, into which we integrated all of our educational resources. In FY 2016, ISM established the Managing Committee of School of Statistical Thinking, inviting contributions from outside experts, and in FY 2017 we launched the Leading DAT program by adopting the suggestions by the committee. The following are the principal projects.

## Research Collaboration Startup

---

The Institute had already been providing a consultation service for statistical science, but along with the launch of the School of Statistical Thinking in November 2011, this service was reorganized as a research collaboration startup. This program, being one of the projects to foster and promote statistical thinking, is mainly aimed at supporting applied scientists and other non-experts. Expert statisticians affiliated with the Institute give them advice on statistical modeling, data analysis, and research. Some cases have developed into official research collaborations, which are our primary duty as an inter-university research institute. The

Institute accepts more than 30 cases annually, some of which benefit society in diverse ways.



## Open-type Professional Development Program

---

This is a spin-out program from ISM cooperative research projects. Establishing a goal is an indispensable element of the proposal of a cooperative research project. On the other hand, such goal setting is irrelevant for a summer school program, study session, or retreat. Since the launch of the School of Statistical Thinking, organizers of such

group-oriented study programs can apply to the Open-type Professional Development Program. There are two categories under this program: one is ‘workshop’ and the other is ‘intensive training for young researchers’. For FY 2018, five workshops and one intensive training for young researchers have been accepted after review.

## Open Lecture

---

We hold an open lecture to introduce the Institute’s activities and to promote statistical science. We invite lecturers to speak on a timely topic relating to statistical science. The lecture is open to the

general public. For further information, please visit the website of the Institute of Statistical Mathematics.

<https://www.ism.ac.jp/kouenkai/>

## Statistical Mathematics Seminar Series

The Institute holds weekly seminar series on statistical mathematics every Wednesday. The seminars are led by in-house and external lecturers to showcase their latest studies. These lectures

are free to attend. To view the seminar schedule and learn more about the program, please visit the Institute of Statistical Mathematics website.

[https://www.ism.ac.jp/index\\_e.html](https://www.ism.ac.jp/index_e.html)

## Data Science Research Plaza

Researchers funded by private-sector firms can maintain a desk and phone in the School of Statistical Thinking. This program is subject to fees, and the contract can be renewed annually. A faculty mentor gives advice to the accepted funded

researcher so that he or she can freely attend various events, such as seminars, workshops, conferences, and extension courses. After learning the expertise of the ISM research staff, participants in this program are invited to take advantage of paid consultations and funded research collaboration.

## Leading DAT

In FY 2017, the School of Statistical Thinking launched a program called “Leading DAT” aimed at training data scientists with the knowledge and skills in statistical mathematics required by modern society. In FY 2018, we organized four Leading DAT lectures entitled “L-A Introductory Data Science”, “L-B1 An Introduction to Statistical Modeling”, “L-B2 Machine Learning and Modern Methodologies in Data Science” and “L-S Geographic Information and Spatial Modeling”. Around one third of the attendants of L-B1 and L-B2 enroll the Leading DAT Training Course, in which we grant certificates to participants who have fulfilled the course requirements, including attendance in L-B1

and L-B2 lectures and submission of reports. A total of 27 people have been granted the certificate of completion.



Certificate ceremony

## Tutorial Courses

The education program at ISM dates back to 1944, the year of founding. The Ministry of Education installed a training center within ISM to foster technicians in numerical computation. After the World War II, this training center was relaunched in 1947 to develop pollsters and census takers. It helped to cultivate professionals in the field of statistical surveys, while a growing number of entries from business and industry coming for various types of training were also observed.

Now the tutorial courses are operated by the

School of Statistical Thinking, which was established in 2011.

In FY 2018, 10 lectures and one course (including Leading DAT) were held and the number of participants was 993. The total numbers of lectures and courses held from 1969 to March, 2019 were 373 and 2 respectively, with a total of 27,021 participants. These lectures covered a wide range of fields from basic to applied statistics. The schedule of tutorial courses can be found on the website of the Institute of Statistical Mathematics.

<https://www.ism.ac.jp/lectures/kouza.html>



# Research Cooperation

## International Cooperation

### ■ Associated Foreign Research Institutes

Organization name	Address	Conclusion day
The Statistical Research Division of the U.S. Bureau of the Census	USA (Washington)	July 27, 1988
Stichting Mathematisch Centrum	The Kingdom of the Netherlands (Amsterdam)	May 10, 1989
Institute for Statistics and Econometrics, Humboldt University of Berlin	Germany (Berlin)	December 8, 2004
The Steklov Mathematical Institute	Russia (Moscow)	August 9, 2005
Central South University	China (Changsha)	November 18, 2005
Soongsil University	The Republic of Korea (Seoul)	April 27, 2006
University of Warwick	The United Kingdom (Coventry)	January 16, 2007
Indian Statistical Institute	India (Kolkata)	October 11, 2007
Institute of Statistical Science, Academia Sinica	Taiwan (Taipei)	June 19, 2008
Department of Empirical Inference, Max Planck Institute for Biological Cybernetics	Germany (Tubingen)	August 11, 2010
Department of Communication Systems, SINTEF Information and Communication Technology	Norway (Trondheim)	January 30, 2012
University College London	The United Kingdom (London)	February 16, 2012
Department of Electronics and Telecommunications, Norwegian University of Science and Technology	Norway (Trondheim)	May 22, 2012
Department of Probability and Mathematical Statistics, Charles University in Prague	Czech Republic (Prague)	October 10, 2012
Department of Ecoinformatics, Biometrics and Forest Growth of the Georg-August University of Goettingen	Germany (Goettingen)	October 18, 2012
Korean Statistical Society (KSS)	The Republic of Korea (Seoul)	July 9, 2013
Toyota Technological Institute at Chicago	USA (Chicago)	February 10, 2014
Australian National University	Australia (Canberra)	May 15, 2014
RiskLab ETH Zurich	Switzerland (Zurich)	February 7, 2015
Institut de Recherche en Composants logiciel et matériel pour l'Information et la Communication Avancée	France (Paris)	February 9, 2015
Le laboratoire de mathématiques de l'Université Blaise Pascal	France (Clermont-Ferrand)	February 11, 2015
Centre de Recherche en Informatique, Signal et Automatique de Lille	France (Paris)	February 12, 2015
University College London Big Data Institute	The United Kingdom (London)	February 26, 2015
The Institute of Forestry, Pokhara of Tribhuvan University	Nepal (Pokhara)	March 6, 2015
The Institute of Forest and Wildlife Research and Development of the Forestry Administration of Cambodia	Cambodia (Phnom Penh)	March 6, 2015
The Chancellor masters and Scholars of the University of Oxford	The United Kingdom (Oxford)	March 10, 2015
Forest Inventory and Planning Institute of Vietnam	Vietnam (Hanoi)	June 2, 2015
Zuse Institute Berlin	Germany (Berlin)	June 20, 2016
The University of Porto	Portugal (Porto)	June 22, 2016
National University of Laos	Laos (Vientiane)	March 15, 2017
Institute of Geophysics China Earthquake Administration	China (Beijing)	April 28, 2017
Hong Kong Baptist University	Hong Kong (Kowloon Tong)	August 28, 2017
University of Malaya	Malaysia (Kuala Lumpur)	September 18, 2017
Universidade de Évora	Portugal (Evora)	November 30, 2017
Universität Ulm	Germany (Ulm)	December 8, 2017
The Korean Association for Survey Research	The Republic of Korea (Seoul)	February 14, 2018
The Jean Golding Institute for data-intensive research, University of Bristol	The United Kingdom (Bristol)	January 15, 2019
Survey Research Center, Sungkyunkwan University	The Republic of Korea (Seoul)	February 25, 2019
University of Lampung	Indonesia (Lampung)	March 6, 2019
Department of Earth and Space Sciences, Southern University of Science and Technology	China (Shenzhen Shi)	March 25, 2019
Université Bretagne Sud	France (Lorient)	March 29, 2019

\* There are two more agreements concluded.

## Research Collaboration

ISM performs many activities for collaborating with researchers in the various fields of statistical science, from the individual level to the national level. The ISM cooperative research program regularly performs research activities to provide the research resources of ISM to researchers at universities or research institutes in order to advance their academic research. Available research resources include books, journals, supercomputers, some commercial statistical software packages, as well as statistical packages developed by ISM, and also the researchers in ISM themselves, who have abundant professional knowledge and experience in statistical science and data analysis. The ISM cooperative research program provides not only research support funds but also opportunities for the various researchers in many fields who require statistical knowledge to make use of the resources available at ISM. ISM's aim is to be a place for interaction and fusion among researchers inside and outside of ISM, and to contribute to multidisciplinary development of both the theory and the application of statistical science.

### ■ Number of Activities

Year	2013	2014	2015	2016	2017	2018
Number of Activities	181	177	183	187	161	166

### ■ Fields of Research Collaboration

Research collaboration is classified by research field as follows. Applicants can use the table below to find the most appropriate type of project.

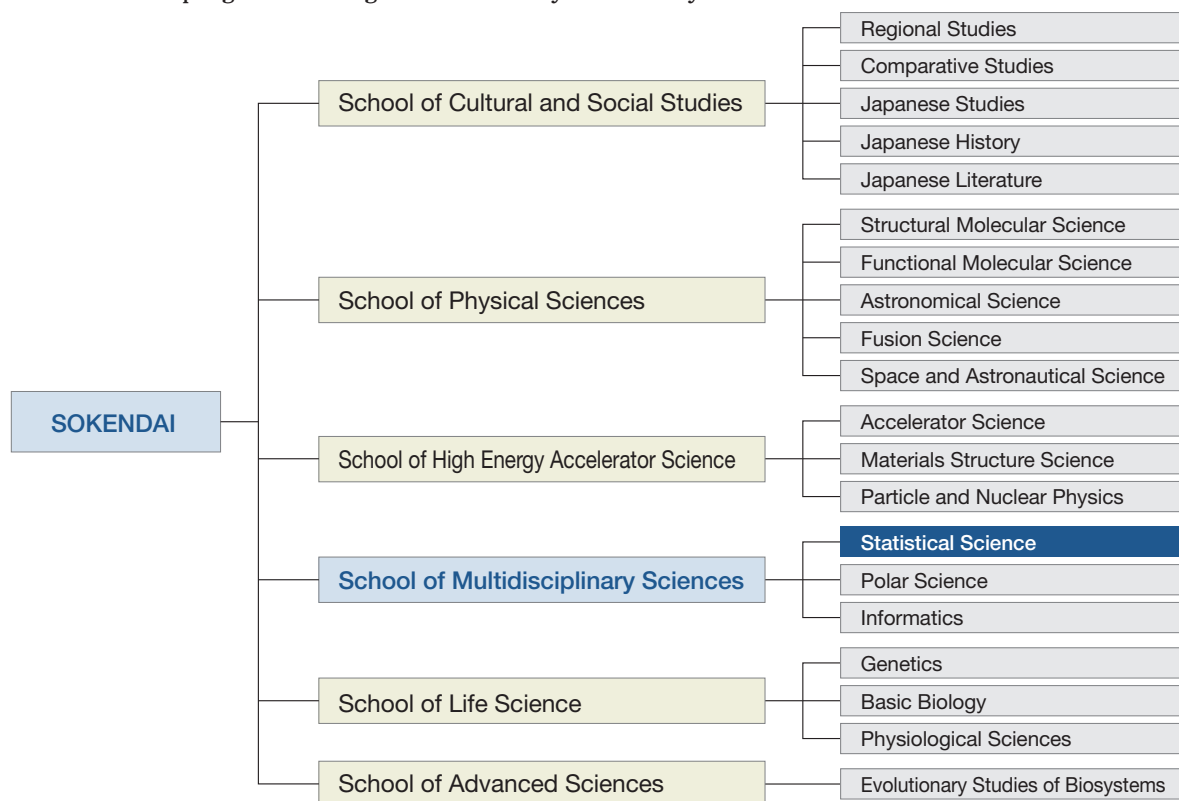
ISM Fields			
Number	Fields	Number	Fields
a	Prediction and Control Group	f	Structure Exploration Group
b	Complex System Modeling Group	g	Mathematical Statistics Group
c	Data Assimilation Group	h	Learning and Inference Group
d	Survey Science Group	i	Mathematical Optimization Group
e	Metric Science Group	j	Others

Major Research Fields		
Number	Fields	Major Research Domains
1	Statistical mathematics	Mathematical theory of statistics, optimization, etc.
2	Information science	Algorithms, use of computer in statistics, etc.
3	Biological science	Medicine, pharmacy, epidemiology, genetics, etc.
4	Physical science	Space, planet, earth, polar region, materials, etc.
5	Engineering	Mechanics, electronics, control, chemistry, architecture, etc.
6	Human science	Philosophy, art, psychology, education, history, geography, culture, language, etc.
7	Social science	Economics, law, politics, society, management, official statistics, population, etc.
8	Environmental science	Environmental statistics, environmentrics, agricultural statistics, statistical meteorology, land economics, landscape management, forest management, etc.
9	Others	Other research fields

# Graduate School Program

## Organization

The Institute of Statistical Mathematics is one of the platforms of SOKENDAI (The Graduate University for Advanced Studies; the headquarters in Hayama, Kanagawa), which was opened in October 1988 to offer graduate education. Since its opening, the Institute has included the Department of Statistical Science and, since April 1989, has accepted students for education and research in doctoral programs. In 2006, the Institute adopted a five-year system, offering either a five-year education and research program, or a three-year education and research program starting from the third year of study.



## Outline of Education and Research

The Department of Statistical Science, which is based on the Institute of Statistical Mathematics (ISM) serving as its underlying platform, aims to cultivate individuals who possess creative research skills to contribute to solving various important intricately-intertwined problems. To this end, the Department conducts education and research related to the basis, mathematics and applications of data collection designs, modeling, inference and prediction, and equip students with the ability to extract information and knowledge from the real world based on the effective use of data.

Field of Education and Research	Contents
Statistical Modeling	Education and research focuses on dynamic modeling such as spatial and space-time modeling, graphical modeling of temporally and/or spatially inter-related complex phenomena, and intelligent information processing. We also provide education and research on statistical inference based on various models, methods of calculation for inference, and evaluation of models based on data.
Statistical Data Science	We provide education and research on data design, investigation and analytical methods to cope with the uncertainty and incompleteness of information, as well as on computational statistics.
Statistical Inference and Mathematics	We provide education and research on the theory of statistics and related basic mathematics, statistical learning theory to extract information from data by automated learning and inference techniques; as well as theory and applications of optimization and computation algorithms which serve as the basis for computational inference.



## Features of Education and Research

- The course is the only integrated doctoral program on statistical science in Japan. It has received students from a wide variety of disciplines and has offered education and research on statistical science by professors specialized in many different fields, from theory through to practical applications.
- The Institute of Statistical Mathematics, the platform for the course, is equipped with a world-class super computer, high-speed 3D graphic computers and simulators to generate physical random numbers, as well as a variety of software, including original statistical software developed by the Institute.
- The academic publications and books on statistical and mathematical sciences produced are some of the best in the world.
- In its role as an inter-university research institute, the Institute holds frequent workshops and seminars by visiting professors and researchers from both Japan and abroad. Students are free to attend and participate.
- It is possible to collaborate with researchers from other universities and institutions. It is also possible for students to develop their own projects by participating in research projects with other institutions.

## Course Requirements and Type of Degree Granted

- Requirements to complete the doctoral course are as follows:  
Completion of at least 40 credits, including the required ones, by a student in the five-year program, or completion of at least 10 credits by a three-year doctorate student who previously completed a Master's course; meeting all the criteria set by the thesis committee of the Institute; and successfully completing the examination.
- On completion of the course, either a Doctorate in Statistical Science or, if the thesis deals mainly with an inter-disciplinary field related to statistical science, a Doctorate of Philosophy is awarded.
- The required number of years of study will be flexible if a student demonstrates outstanding research results.

## Number of Students (As of April 1, 2019)

■ 5-year doctoral course:Quota,2	Year of enrollment	2010	2013	2016	2017	2018	2019
	Number of students	1	1	1	2	1	3
■ 3-year doctoral course:Quota,3	Year of enrollment	2014	2015	2016	2017	2018	2019
	Number of students	1 ①	1 ①	4 ③	4 ③	7 ⑤	9 ⑦

\* The figures in circles indicate those who are employed by other organizations.

## University Background of Students

### National and public universities

● Hokkaido University (5) ● Tohoku University (3) ● Fukushima University (1) ● University of Tsukuba (6) ● Saitama University (1)  
 ● Chiba University (1) ● Ochanomizu University (1) ● National Graduate Institute for Policy Studies (1) ● Tokyo Medical and Dental University (1) ● Tokyo University of Marine Science and Technology (1) ● Tokyo Gakugei University (2) ● Tokyo Institute of Technology (6) ● The University of Tokyo (20) ● Tokyo Metropolitan University (1) ● Tokyo University of Agriculture and Technology (1)  
 ● Hitotsubashi University (6) ● Shizuoka University (1) ● Japan Advanced Institute of Science and Technology (1) ● Nagoya University (3)  
 ● Toyohashi University of Technology (2) ● Kyoto University (7) ● Osaka City University (1) ● Osaka University (3) ● Nara Institute of Science and Technology (1) ● Okayama University (2) ● Shimane University (3) ● Kyushu University (3) ● Oita University (1)

## University Background of Students

### Private universities

• Kitasato University (1) • Keio University (8) • International Christian University (1) • Shibaura Institute of Technology (1) • Chuo University (9) • Tokyo University of Science (7) • Toyo University (1) • Japan Women's University (1) • Nihon University (2) • Hosei University (7) • Waseda University (9) • Nanzan University (1) • Osaka Electro-Communication University (1) • Kansai University (1) • Kyoto Sangyo University (1) • Ritsumeikan University (1) • Okayama University of Science (1) • Kurume University (1)

### Foreign universities

• Aston University (1) • University of California, Irvine (1) • California State University, Long Beach (1) • University of Campinas (1) • University of Colorado Boulder (2) • University of Dhaka (2) • University of Hawaii (1) • Jahangirnagar University (2) • University of Malaya (1) • Northeast Normal University (1) • Ohio University (2) • University of Rajshahi (2) • Stanford University (1) • The University of Nottingham (1) • Institute of Applied Mathematics, AMSS, CAS (1) • University of Science and Technology of China (1) • Center for Analysis and Prediction, China Seismological Bureau (1) • Northeastern University (1) • The Hong Kong University of Science and Technology (1)

## Degrees Awarded

Year	2013	2014	2015	2016	2017	2018
Doctor of Philosophy	6	5	5	7	5	5

## Alumni

### National and public universities, and public organizations

• Obihiro University of Agriculture and Veterinary Medicine • University of Tsukuba • University of Hyogo • The University of Tokyo • The University of Electro-Communications • Saitama University • Nagoya University • Kyushu University • Kyushu Institute of Technology • University of the Ryukyus • The Institute of Statistical Mathematics • Tohoku University • Yokohama National University • Hokkaido University • Tokyo Institute of Technology • Hiroshima University • Oita University of Nursing and Health Sciences • JAXA's Engineering Digital Innovation Center • Kyoto University • Nara Institute of Science and Technology • Bank of Japan • Japan Broadcasting Corporation • Railway Technical Research Institute • Statistical Information Institute for Consulting and Analysis • Government Pension Investment Fund • Public School • RIKEN • Statistics Bureau of Japan

### Private universities

• Sapporo Gakuin University • Tokyo Health Care University • Meiji University • Doshisha University • Josai University • Nihon University • Komazawa University • Aichi University of Technology • Tokyo University of Information Sciences • Shibaura Institute of Technology • Rikkyo University • Waseda University • Keio University

### Foreign universities

• Jahangirnagar University • Victoria University • Massey University • University of Otago • Statistics New Zealand • University of Rajshahi • University of California, Los Angeles • Asia-Pacific Center for Security Studies Department • Central South University • Hong Kong Baptist University • University of South Carolina • The University of Warwick

### Private companies, etc.

• Hitachi, Ltd. Central Research Laboratory • NTT Communication Science Laboratories • Seiwa Kikaku • NLI Research Institute • Mizuho Trust and Banking • Nomura Securities Co., Ltd. • ATR Computational Neuroscience Laboratories • Toyota Motor Corporation, Higashi-Fuji Technical Center • Schlumberger Limited • Macquarie Securities, Japan • Non-Life Insurance Rating Organization of Japan • Barclays Global Investors • Open Technologies Corporation • Yamaha Corporation • Goldman Sachs Asset Management L.P. • CLC bio Japan, Inc. • Bank of Tokyo-Mitsubishi UFJ • Pfizer Japan Inc. • Doctoral Institute for Evidence Based Policy • Sony Corporation • NTTIT Corporation • Sompo Japan Insurance Inc. • Qualicaps Co., Ltd. • Bridgestone Corporation • Brain Pad Inc. • Sumitomo Chemical Co., Ltd. • PricewaterhouseCoopers Aarata • Mitsubishi Tanabe Pharma Corporation • Daiichi Sankyo Co., Ltd. • Shizuoka Cancer Center • CPC Clinical Trial Hospital, Medipolis Medical Research Institute • CRD Association • Japan Society for the Promotion of Science • Tokyo Electric Power Company Holdings, Inc. • Asahi Kasei Corporation • Honda R&D Co., Ltd. • Yokogawa Electric Corporation • Kao Corporation • Advanced Smart Mobility Co., Ltd.

# Facilities and Equipment

## Computational Resources (As of April 1, 2019)

ISM is now operating two different supercomputer systems. One is the Supercomputer System for Statistical Science, which was introduced in FY2018. The other system is Communal Cloud Computing System (called System “C”) introduced at the end of FY2013.

The Supercomputer System for Statistical Science, an HPE SGI 8600 system, has been operated since October 2018. The system is a distributed-memory parallel computer that has total theoretical peak performance of 1.49 petaflops. The system is liquid cooled and consists of 384 compute nodes. Each node has two CPU chips (Intel Xeon GOLD 6154) with 18 cores and has 384 GB memory. The system includes hardware random number generator. For visual representation of research results, ISM has a 200-inch wide screen and a projector capable of showing 3D movies in 4K resolution in the historical computers exhibit room.

System “C” consists of 64 Dell PowerEdge R620 (two 10-core Xeon E5-2680v2, 256GB memory). This system

provides easy-to-use computing environments such as distributed-memory statistical computing environments and Web servers, running on Apache CloudStack software. This private cloud system is also used in hosting data analysis competition events.

In the main office building, the primary local area network (LAN) consists of an Ethernet network using 10GBase-SR for the main trunk and 1000Base-T for branches. The personal computers in researchers’ offices, and the two supercomputer systems are all connected to this network. A wireless LAN system is also available in the immediate area of the building occupied by ISM. These LAN systems enable distributed processing and allow computing resources and statistical data to be used effectively. Comprehensive network security methods have been implemented, such as a firewall system, anti-virus software, and an intrusion prevention system. To encourage joint research with researchers both in Japan and abroad, as well as the exchange of e-mails, the network is connected to the Internet through SINET5 (40 Gbps).



Supercomputer System for Statistical Science (HPE SGI 8600)



200-inch wide screen showing 3D movies in 4K resolution

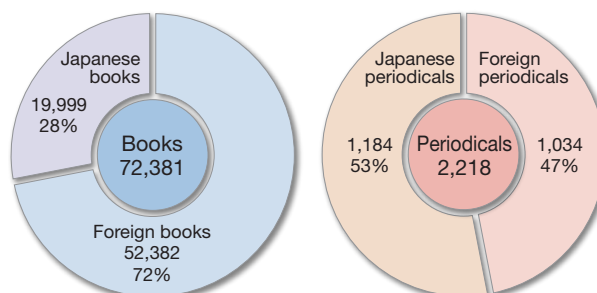
## Library and Materials (As of April 1, 2019)

We have a large number of major Japanese/foreign journals covering a wide variety of fields including statistics, mathematics, computer science and informatics. In addition, we also have a large library consisting of books on humanities, social science, biology, medical science, science and engineering.

Besides contributed to Japanese and foreign publications, we also have a collection of journals that we publish ourselves: Annals of the Institute of Statistical Mathematics (English; Springer), Proceedings of the Institute of Statistical Mathematics (Japanese), ISM Survey Research Report (Statistical Researches mainly related to the Japanese National Character), Computer Science Monographs, Cooperative Research Reports (for collaborative research projects), Research Memorandum, ISM

Reports on Statistical Computing, and ISM Report on Research and Education.

All materials are properly catalogued and can be searched from the web in order to meet the needs of researchers working in a wide of fields. We also accept photocopy requests.





# Finance and Buildings

## Administration Subsidy and Others (2018)

Type	Personnel expenses	Non-personnel expenses	Total
Expenditure	665,753	953,057	1,618,810

Unit: 1,000JPY

## Accepted External Funds (2018)

Type	Joint research	Subcontracted research, Trustee business	Academic Consulting	Contract researchers	Contribution for scholarship	Total
Items	23	24	10	3	7	67
Income	54,913	137,514	10,680	1,672	8,777	213,556

Unit: 1,000JPY

## Grant-in-Aid for Scientific Research “KAKENHI” (2018)

Research Category	Items	Amount Granted
Grant-in-Aid for Scientific Research on Innovation Areas	—	—
Grant-in-Aid for Scientific Research (S)	—	—
Grant-in-Aid for Scientific Research (A)	5	63,180
Grant-in-Aid for Scientific Research (B)	9	33,670
Grant-in-Aid for Scientific Research (C)	20	19,711
Grant-in-Aid for Challenging Research (Exploratory)	3	6,370
Grant-in-Aid for Young Scientists (B)	11	12,306
Grant-in-Aid for Early-Career Scientists	5	4,179
Grant-in-Aid for Research Activity Start-up	1	1,430
Grant-in-Aid for JSPS Fellows	1	1,690
Total	55	142,536

Unit: 1,000JPY

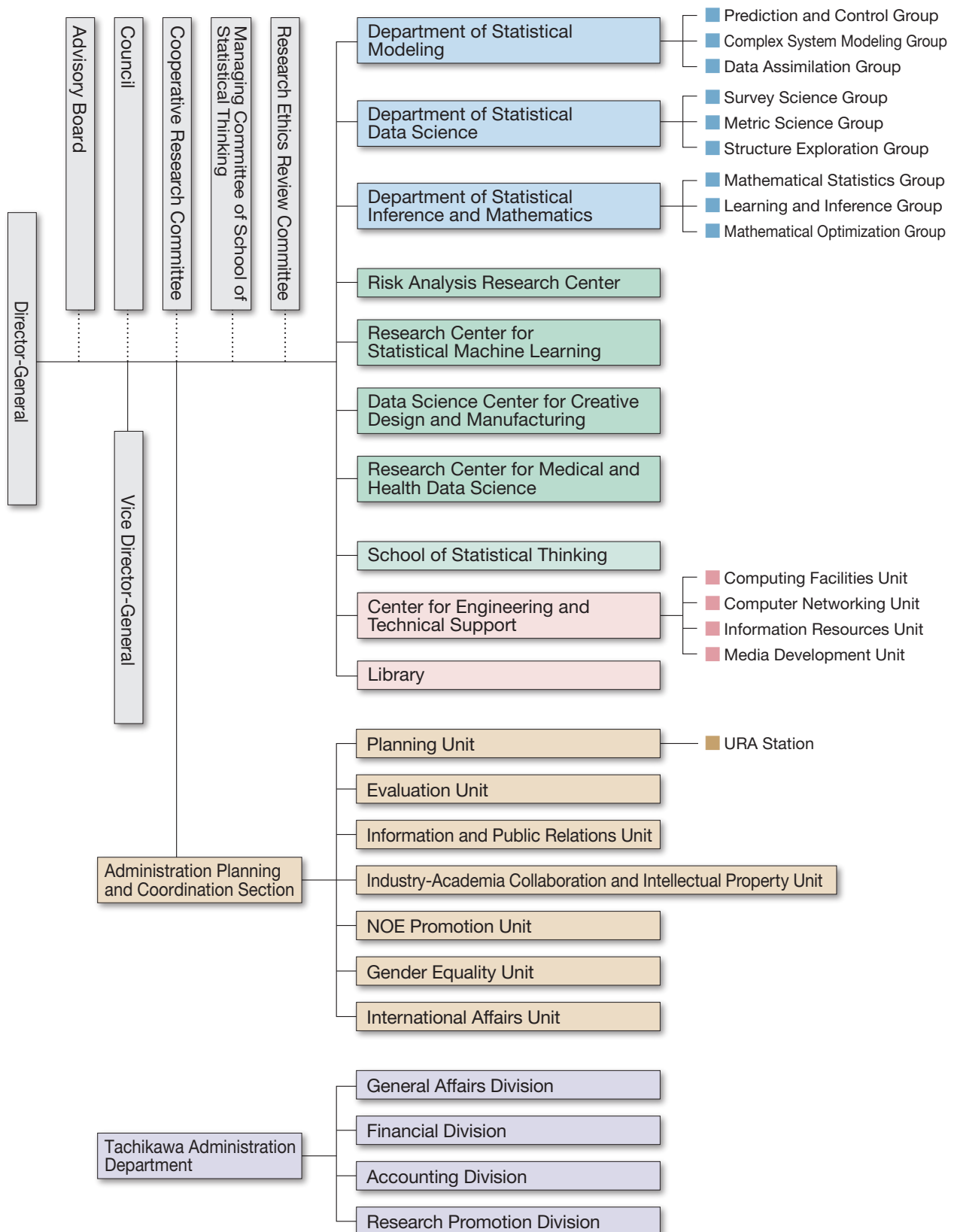
## Site and Buildings (As of April 1, 2019)

Site Area	62,450m <sup>2</sup>
Area for Buildings (total)	16,209m <sup>2</sup>



# Organization

## Organization Diagram (As of April 1, 2019)



## Number of Staff (As of April 1, 2019)

Type	Director-General	Professor	Associate Professor	Assistant Professor	Administrative Staff	Technical Staff	Total
Director-General	1						1
Department of Statistical Modeling		6	7	2			15
Department of Statistical Data Science		5	6	3			14
Department of Statistical Inference and Mathematics		7	5	3			15
School of Statistical Thinking							0
Center for Engineering and Technical Support						11	11
Administration Planning and Coordination Section					1		1
Tachikawa Administration Department					(33)		(33)
Total	1	18	18	8	1(33)	11	57(33)

( ) Total number of staff of Tachikawa Administration Department.

The number under Technical Staff at the Center for Engineering and Technical Support and Administrative Staff at the Tachikawa Administration Department include two each staff member who retired because of age but was reemployed in a different position.

## Staff (As of August 1, 2019)

Director-General	Hiroe TSUBAKI		
Vice Director-General	Satoshi ITO	Vice Director-General	Satoshi YAMASHITA
		Vice Director-General	Yoshihiko MIYASATO

### Department of Statistical Modeling

Director Tomoko MATSUI

#### Prediction and Control Group

Prof.	Yoshinori KAWASAKI	Prof.	Yoshihiko MIYASATO	Prof.	Atsushi YOSHIMOTO
Assoc. Prof.	Jiancang ZHUANG	Assoc. Prof.	Yumi TAKIZAWA	Assoc. Prof.	Fumikazu MIWAKEICHI

#### Complex System Modeling Group

Prof.	Tomoko MATSUI	Prof.	Yukito IBA	Assoc. Prof.	Kazuhiro MINAMI
Assoc. Prof.	Shinsuke KOYAMA	Assoc. Prof.	Hideitsu HINO	Assist. Prof.	Momoko HAYAMIZU

#### Data Assimilation Group

Prof.	Genta UENO	Assoc. Prof.	Shinya NAKANO	Project Assoc. Prof.	Masaya SAITO
Assist. Prof.	Shunichi NOMURA	Project Assist. Prof.	Takashi YAMAMOTO	Visiting Prof.	Tadahiko SATO
Visiting Prof.	Shinichi OTANI	Visiting Prof.	Kazuyuki NAKAMURA	Visiting Prof.	Tomoyuki HIGUCHI
Visiting Prof.	Masako KAMIYAMA	Visiting Assoc. Prof.	Yosuke FUJII	Visiting Assoc. Prof.	Hiromichi NAGAO
Visiting Assoc. Prof.	Hiroshi KATO				

### Department of Statistical Data Science

Director Koji KANEFUJI

#### Survey Science Group

Prof.	Ryozo YOSHINO	Assoc. Prof.	Tadahiko MAEDA	Assoc. Prof.	Yoo Sung PARK
Project Assist. Prof.	Yusuke INAGAKI	Project Assist. Prof.	Kiyohisa SHIBAI	Project Assist. Prof.	Le Duc ANH

## Staff

### Department of Statistical Data Science

Project Researcher	Naoko KATO	Visiting Prof.	Toru KIKKAWA	Visiting Prof.	Yoshimichi SATO
Visiting Prof.	Wataru MATSUMOTO	Visiting Prof.	Kazufumi MANABE	Visiting Prof.	Takatoshi IMADA
Visiting Prof.	Masahiro MIZUTA	Visiting Assoc. Prof.	Taisuke FUJITA	Visiting Assoc. Prof.	Koken OZAKI

### Metric Science Group

Prof.	Satoshi YAMASHITA	Prof.	Koji KANEFUJI	Prof.	Yoichi ITO
Assoc. Prof.	Ikuko FUNATOGAWA	Assoc. Prof.	Hisashi NOMA	Assist. Prof.	Nobuo SHIMIZU
Project Researcher	Hiroka HAMADA				

### Structure Exploration Group

Prof.	Ryo YOSHIDA	Assoc. Prof.	Jun ADACHI	Assoc. Prof.	Kenichiro SHIMATANI
Assist. Prof.	Stephen WU	Assist. Prof.	Daisuke MURAKAMI		

### Department of Statistical Inference and Mathematics

Director Satoshi KURIKI

### Mathematical Statistics Group

Prof.	Satoshi KURIKI	Prof.	Yoshiyuki NINOMIYA	Assoc. Prof.	Shuhei MANO
Assoc. Prof.	Shogo KATO	Assoc. Prof.	Takaaki SHIMURA	Visiting Prof.	Akimichi TAKEMURA

### Learning and Inference Group

Prof.	Shinto EGUCHI	Prof.	Kenji FUKUMIZU	Prof.	Hironori FUJISAWA
Assoc. Prof.	Daichi MOCHIIHASHI	Assoc. Prof.	Masayuki HENMI	Assist. Prof.	Ayaka SAKATA
Assist. Prof.	Masaaki IMAIZUMI				

### Mathematical Optimization Group

Prof.	Satoshi ITO	Prof.	Shiro IKEDA	Assist. Prof.	Mirai TANAKA
Visiting Prof.	Eitarou AIYOSHI				

### Risk Analysis Research Center

Director Satoshi YAMASHITA

Vice Director Shogo KATO

Prof.	Satoshi YAMASHITA	Prof.	Satoshi KURIKI	Prof.	Koji KANEFUJI
Prof.	Tomoko MATSUI	Prof.	Atsushi YOSHIMOTO	Prof.	Yoshinori KAWASAKI
Prof.	Yoshiyuki NINOMIYA	Assoc. Prof.	Masayuki HENMI	Assoc. Prof.	Jianchang ZHUANG
Assoc. Prof.	Kenichiro SHIMATANI	Assoc. Prof.	Shogo KATO	Assoc. Prof.	Yumi TAKIZAWA
Assoc. Prof.	Shuhei MANO	Assoc. Prof.	Kazuhiro MINAMI	Assoc. Prof.	Takaaki SHIMURA
Assist. Prof.	Shunichi NOMURA	Assist. Prof.	Stephen WU	Assist. Prof.	Daisuke MURAKAMI
Project Assist. Prof.	Yicun GUO	Project Assist. Prof.	Junchao ZHANG	Project Assist. Prof.	Yuma UEHARA
Project Assist. Prof.	Hideaki NAGAHATA	Visiting Prof.	Yo SHIINA	Visiting Prof.	Hisayuki HARA
Visiting Prof.	Toshinao YOSHIBA	Visiting Prof.	Nakahiro YOSHIDA	Visiting Prof.	Rinya TAKAHASHI
Visiting Prof.	Toshikazu KITANO	Visiting Prof.	Sadaaki MIYAMOTO	Visiting Prof.	Shigeyuki MATSUI
Visiting Prof.	Takahiro HOSHINO	Visiting Prof.	Shinsuke ITO	Visiting Prof.	Hitoshi MOTOYAMA
Visiting Prof.	Takashi KAMEYA	Visiting Prof.	Shunji HASHIMOTO	Visiting Prof.	Naoki SAKAI
Visiting Prof.	Satoshi TAKIZAWA	Visiting Prof.	Mihoko MINAMI	Visiting Prof.	Toshihiro HORIGUTI



#### Risk Analysis Research Center

Visiting Prof.	Naoto KUNITOMO	Visiting Prof.	Hideatsu TSUKAHARA	Visiting Prof.	Toshio HONDA
Visiting Prof.	Masaaki FUKASAWA	Visiting Prof.	Yasutaka SHIMIZU	Visiting Prof.	Takaaki YOSHINO
Visiting Prof.	Michiko MIYAMOTO	Visiting Prof.	Tadashi ONO	Visiting Prof.	Hiroshi TSUDA
Visiting Prof.	Satoshi FUJII	Visiting Prof.	Masakazu ANDO	Visiting Assoc. Prof.	Masao UEKI
Visiting Assoc. Prof.	Takafumi KUBOTA	Visiting Assoc. Prof.	Yukihiko OKADA	Visiting Assoc. Prof.	Junichi TAKAHASHI
Visiting Assoc. Prof.	Seisho SATO	Visiting Assoc. Prof.	Yuta KOIKE	Visiting Assoc. Prof.	Kenichi KAMO
Visiting Assoc. Prof.	Masashi KONOSHIMA	Visiting Assoc. Prof.	Tetsuji TONDA	Visiting Assoc. Prof.	Bogdan Dumitru ENESCU
Visiting Assoc. Prof.	Kazuyoshi NANJO	Visiting Assoc. Prof.	Takaki IWATA	Visiting Assoc. Prof.	Teppey OGIHARA
Visiting Assoc. Prof.	Dou XIAOLING				

#### Research Center for Statistical Machine Learning

Director Kenji FUKUMIZU

Vice Director Tomoko MATSUI

Prof.	Kenji FUKUMIZU	Prof.	Tomoko MATSUI	Prof.	Shinto EGUCHI
Prof.	Yoshihiko MIYASATO	Prof.	Satoshi ITO	Prof.	Shiro IKEDA
Prof.	Satoshi KURIKI	Prof.	Hironori FUJISAWA	Assoc. Prof.	Daichi MOCHIHASHI
Assoc. Prof.	Shinsuke KOYAMA	Assoc. Prof.	Kazuhiro MINAMI	Assoc. Prof.	Hideitsu HINO
Project Assoc. Prof.	Shinichiro GOTO	Assist. Prof.	Mirai TANAKA	Assist. Prof.	Daisuke MURAKAMI
Assist. Prof.	Ayaka SAKATA	Assist. Prof.	Masaaki IMAIZUMI	Project Assist. Prof.	Mikio MORII
Project Assist. Prof.	Sho SAITO	Project Assist. Prof.	Yoichi MOTOTAKE	Project Assist. Prof.	Ye CHEN
Project Researcher	Takuo HAMAGUCHI	Visiting Prof.	Yoichi MOTOMURA	Visiting Prof.	Nobuhiko TERUI
Visiting Prof.	Yoshiki YAMAGATA	Visiting Prof.	Masataka GOTO	Visiting Prof.	Arthur GRETTON
Visiting Prof.	Takashi TSUCHIYA	Visiting Prof.	Akiko TAKEDA	Visiting Prof.	Katsuki FUJISAWA
Visiting Prof.	Yuji SHINANO	Visiting Assoc. Prof.	João Pedro PEDROSO	Visiting Assoc. Prof.	Tsutomu TAKEUCHI
Visiting Assoc. Prof.	Eiji MOTOHASHI	Visiting Assoc. Prof.	Tsukasa ISHIGAKI	Visiting Assoc. Prof.	Makoto YAMADA

#### Data Science Center for Creative Design and Manufacturing

Director Ryo YOSHIDA

Vice Director Hironori FUJISAWA

Prof.	Ryo YOSHIDA	Prof.	Hironori FUJISAWA	Prof.	Kenji FUKUMIZU
Assoc. Prof.	Shinya NAKANO	Assoc. Prof.	Daichi MOCHIHASHI	Assist. Prof.	Stephen WU
Project Assist. Prof.	Liu CHANG	Project Researcher	Yo NOGUCHI	Visiting Assoc. Prof.	Terumasa TOKUNAGA

#### Research Center for Medical and Health Data Science

Director Yoichi ITO

Vice Director Hisashi NOMA

Prof.	Yoichi ITO	Prof.	Satoshi YAMASHITA	Prof.	Shinto EGUCHI
Assoc. Prof.	Hisashi NOMA	Assoc. Prof.	Masayuki HENMI	Assoc. Prof.	Ikuko FUNATOGAWA
Assoc. Prof.	Fumikazu MIWAKEICHI	Project Assoc. Prof.	Kengo NAGASHIMA	Project Assist. Prof.	Mayumi OKA
Project Assist. Prof.	Naomi TAMURA	Project Researcher	Yasuyuki HAMURA	Visiting Prof.	Yasuo OHASHI
Visiting Prof.	Senichiro KIKUCHI	Visiting Prof.	Ken KIYONO	Visiting Prof.	Tatsuhiko TSUNODA
Visiting Prof.	Hisateru TACHIMORI	Visiting Prof.	Satoshi TERAMUKAI	Visiting Prof.	Toshiya SATO
Visiting Prof.	Satoshi HATTORI	Visiting Prof.	Manabu IWASAKI	Visiting Prof.	Shigeyuki MATSUI
Visiting Prof.	Michiko WATANABE	Visiting Assoc. Prof.	Ryoichi KIMURA	Visiting Assoc. Prof.	Masataka TAGURI

## Staff

### Research Center for Medical and Health Data Science

Visiting Assoc. Prof. Yasunori SATO	Visiting Assoc. Prof. Ryota NAKAMURA	Visiting Assoc. Prof. Kazushi MARUO
Visiting Assoc. Prof. Atsushi GOTO	Visiting Assoc. Prof. Kunihiko TAKAHASHI	

### School of Statistical Thinking

Director Yoshinori KAWASAKI Vice Director Yukito IBA

Prof. Yoshinori KAWASAKI	Prof. Satoshi ITO	Prof. Yukito IBA
Prof. Kenji FUKUMIZU	Prof. Yoshiyuki NINOMIYA	Project Prof. Yoshiyasu TAMURA
Assoc. Prof. Kenichiro SHIMATANI	Assoc. Prof. Masayuki HENMI	Assoc. Prof. Hideitsu HINO
Project Assoc. Prof. Naoki KAMIYA	Assist. Prof. Shunichi NOMURA	Assist. Prof. Masaaki IMAIZUMI
Project Researcher Tsubasa ITO	Project Researcher Kazuhei KIKUCHI	Visiting Assoc. Prof. Osamu KOMORI
Visiting Assoc. Prof. Kei TAKAHASHI		

### Center for Engineering and Technical Support

Director Genta UENO Vice Director Kazuhiro MINAMI  
Deputy Manager Yuriko WATANABE

Unit Leader of Computing Facilities Unit Mitsuru HAYASAKA	Unit Leader of Computer Networking Unit Kazuhiro NAKAMURA
Unit Leader of Information Resources Unit Yuriko WATANABE	Unit Leader of Media Development Unit Akiko NAGASHIMA

### Library

Head Genta UENO

### Administration Planning and Coordination Section

Chief Director Hiroe TSUBAKI

Director of Planning Unit Satoshi ITO	Director of Evaluation Unit Yoshihiko MIYASATO
Director of Information and Public Relations Unit Satoshi ITO	Director of Industry-Academia Collaboration and Intellectual Property Unit Satoshi YAMASHITA
Director of NOE Promotion Unit Satoshi YAMASHITA	Director of Gender Equality Unit Satoshi ITO
Director of International Affairs Unit Tomoko MATSUI	

#### URA Station

Senior URA Kozo KITAMURA	Chief URA Motoi OKAMOTO
Chief URA Keisuke HONDA	

### Tachikawa Administration Department

Director Isao OSHIRO

#### General Affairs Division

Manager Tsuyoshi ABE Deputy Manager Masatoshi NAKAMURA  
Deputy Manager Yasunori SAITO

Head, General Affairs Team Masatoshi NAKAMURA	Head, Personnel Team Yasunori SAITO
Head, Labor Management Team Yasunori SAITO	

## ■ Financial Division

	Manager Katsuhiro OZAKI	Deputy Manager Koji SAKAMOTO
Head, General Affairs and Audit Team	Koji SAKAMOTO	Head, Budget and Account Settlement Team (NIPR) Michihito SAKURAI
Head, Budget and Account Settlement Team (ISM)	Ichiro KAWAJI	Head, Assets Management and Acceptance Team Akiko MAEKAWA

## ■ Accounting Division

	Manager Yuji TAHARA	Deputy Manager Hiroki TERAUCHI
		Deputy Manager Masayuki KOBAYASHI
Head, Accounting Team	Hiroaki ARAI	Head, Contract Team1 Toshiaki TAKASAKI
Head, Contract Team2	Takeshi KUWAHARA	Head, Facilities Team Takuya SAITO

## ■ Research Promotion Division

	Manager Motoki NISHIJIMA	Deputy Manager Motokazu TOYODA
Head, Research Promotion Team	Hideaki ASANO	Head, Cooperative Research Team Hitoshi HIRAYAMA
Head, Graduate School Team	Ayako NARITA	

## Council of The Institute of Statistical Mathematics (As of April 1, 2019)

Keiko TAKAHASHI	Director-General, Japan Agency for Marine-Earth Science and Technology (JAMSTEC) Center for Earth Information Science and Technology (CEIST)
Masahiro MIZUTA	Professor of Information Initiative Center/ Graduate School, Faculty of Information Science and Technology, Hokkaido University
Shigeru OBAYASHI	Director, Institute of Fluid Science, Tohoku University
Nakahiro YOSHIDA	Professor, Graduate School of Mathematical Sciences, University of Tokyo
Masayuki UCHIDA	Professor, Graduate School of Engineering Science, Osaka University
Hiroshi SAIGO	Professor, Faculty of Political Science and Economics, Waseda University
Yasuhiro OMORI	Professor, Faculty of Economics, University of Tokyo
Kikuo MAEKAWA	Professor, Spoken Language Division, Director, Center for Corpus Development National Institute for Japanese Language and Linguistics
Hideki ASOH	Director, National Institute of Advanced Industrial Science and Technology Department of Information Technology and Human Factors Artificial Intelligence Research Center Deputy
Mihoko MINAMI	Professor, Department of Mathematics, Faculty of Science and Technology, Keio University
Satoshi ITO	Professor (Vice Director-General, ISM)
Satoshi YAMASHITA	Professor (Vice Director-General, ISM)
Yoshihiko MIYASATO	Professor (Vice Director-General, ISM)
Tomoko MATSUI	Professor (Director of Department of Statistical Modeling, ISM)
Koji KANEFUJI	Professor (Director of Department of Statistical Data Science, ISM)
Satoshi KURIKI	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)
Genta UENO	Professor (Director of Center for Engineering and Technical Support, ISM)
Yoshinori KAWASAKI	Professor (Director of School of Statistical Thinking, ISM)
Ryozo YOSHINO	Professor (Department of Statistical Data Science, ISM)
Kenji FUKUMIZU	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)
Hironori FUJISAWA	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)

## Cooperative Research Committee (As of June 1, 2019)

Hiroyuki MINAMI	Professor, Information Initiative Center, Hokkaido University
Kunihiko TAKAHASHI	Associate Professor, Department of Biostatistics, Nagoya University Graduate School of Medicine
Takahiro TSUCHIYA	Professor, School of Data Science, Yokohama City University
Toshikazu KITANO	Professor, Department of Architecture, Nagoya Institute of Technology
Nagatomo NAKAMURA	Professor, Department of Economics, Sapporo Gakuin University
Atsushi YOSHIMOTO	Professor (Director of Department of Statistical Modeling, ISM)
Koji KANEFUJI	Professor (Department of Data Science, ISM)
Shiro IKEDA	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)
Junji NAKANO	Project Professor (School of Statistical Thinking, ISM)

## Managing Committee of School of Statistical Thinking (As of April 1, 2019)

Manabu IWASAKI	Dean/Professor, School of Data Science, Yokohama City University
Toshiyasu MATUSHIMA	Director, Center for Data Science, Waseda University
Hiroshi YAMADA	Professor, Graduate School of Social Sciences, Hiroshima University
Seiji YAMADA	Professor, Digital Content and Media Sciences Research Division, National Institute of Informatics
Junichi SHIOZAKI	Senior Manager, Insight Signal Business Department, Nomura Research Institute, Ltd.
Yoshinori KAWASAKI	Director (School of Statistical Thinking, ISM)
Yukito IBA	Vice Director (School of Statistical Thinking, ISM)
Satoshi ITO	Professor (Vice Director-General, ISM)
Kenichiro SHIMATANI	Associate Professor (Department of Data Science, ISM)

## Research Ethics Review Committee (As of April 1, 2019)

Specialist on epidemiology and social research	Masayuki KANAI	Professor, School of Human Sciences, Senshu University
Specialist on epidemiology and social research	Keiko SATO	Associate Professor, Kyoto University Hospital, Institute for Advancement of Clinical and Translational Science Department of EBM Research
Specialist in the field of ethics and law	Hitomi NAKAYAMA	Lawyer, Kasumigaseki-Sogo Law Offices
Person in citizen's position	Yutaka KURIKI	Kindergarten Director, Nishikokubunji Nursery School
Research education staff of ISM	Tadahiko MAEDA	Associate Professor (Department of Data Science, ISM)
Research education staff of ISM	Yoo Sung PARK	Associate Professor (Department of Data Science, ISM)
Research education staff of ISM	Koji KANEFUJI	Professor (Department of Data Science, ISM)
Research education staff of ISM	Ikuko FUNATOGAWA	Associate Professor (Department of Data Science, ISM)
Research education staff of ISM	Shuhei MANO	Associate Professor (Department of Mathematical Analysis and Statistical Inference, ISM)

## Professor Emeritus (As of April 1, 2019)

Sigeki NISHIHARA	Tatsuzo SUZUKI	Giitiro SUZUKI	Ryoichi SHIMIZU
Noboru OHSUMI	Masakatsu MURAKAMI	Kunio TANABE	Tadashi MATSUNAWA
Masami HASEGAWA	Yoshiyuki SAKAMOTO	Takemi YANAGIMOTO	Yoshiaki ITOH
Yasumasa BABA	Katsuomi HIRANO	Masaharu TANEMURA	Makio ISHIGURO
Yosihiko OGATA	Hiroe TSUBAKI	Genshiro KITAGAWA	Nobuhisa KASHIWAGI
Takashi NAKAMURA	Yoshiyasu TAMURA	Tomoyuki HIGUCHI	Junji NAKANO

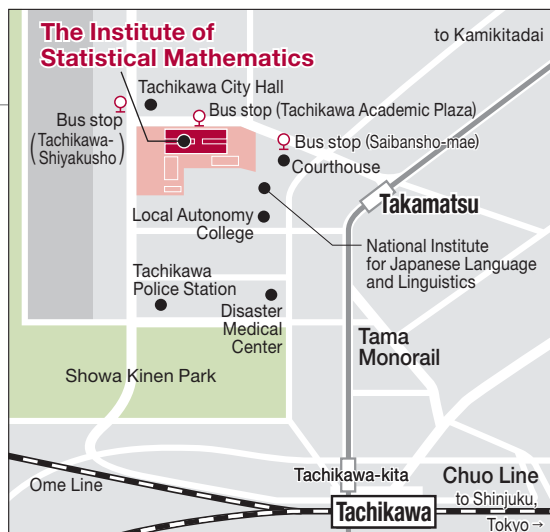


# History

1944	June	●	Based on a proposal submitted at an academic study conference in December 1943, the organization was founded as an institute under the direct control of the Ministry of Education. This proposal aimed to provide supervision for studies looking into the mathematical principles of probability and their application, and was also intended to facilitate, unify and promote the publication of research results.
1947	April	●	The affiliated statistical specialists' school was opened.
	May	●	The Institute was divided into the 1st Research Dept. (fundamental theories), the 2nd Research Dept. (statistical theories for the natural sciences), and the 3rd Research Dept. (statistical theories for the social sciences).
1949	June	●	The Institute was placed under the control of the Ministry of Education because of the enforcement of the Ministry of Education Establishment Law.
1955	September	●	Reorganized into the 1st Research Dept. (fundamental theories), the 2nd Research Dept. (natural and social science theories), and the 3rd Research Dept. (operations research, statistical analysis theories). The laboratory system, comprising 9 laboratories and the research guidance promotion room, was adopted.
1969	October	●	A new office building was constructed in Minato Ward.
1971	April	●	The 4th Research Dept. (informatics theories) was instituted.
1973	April	●	The 5th Research Dept. (prediction and control theories) was instituted.
1975	October	●	The 6th Research Dept. (statistical theories of human behavior) was instituted.
1979	November	●	The Information Research Building was constructed.
1985	April	●	Repositioned as a National Inter-University Research Institute due to the regulation change. The new mission includes providing facilities and skills to other universities, in addition to conducting cutting-edge research on statistical mathematics. Accordingly, the institute was reorganized into four basic research departments (Fundamental Statistical Theory, Statistical Methodology, Prediction & Control, and Interdisciplinary Statistics) and two strategic centers (Statistical Data Analysis Center and Statistical Education & Information Center). The Statistical Technical Training Center was terminated.
1988	October	●	The Dept. of Statistical Science was instituted in the School of Mathematical and Physical Science, part of the Graduate University for Advanced Studies (SOKENDAI).
1989	June	●	The Institute was reorganized as an Inter-University Research Institute based on the National School Establishment Law.
1993	April	●	The Planning Coordination Chief System was instituted.
1997	April	●	The affiliated Statistical Data Analysis Center was reorganized into the Center for Development of Statistical Computing, and the Statistical Education and Information Center was reorganized into the Center for Information on Statistical Sciences.
2003	September	●	The Prediction and Knowledge Discovery Research Center was instituted.
2004	April	●	The Institute was reorganized into the Institute of Statistical Mathematics, part of the Research Organization of Information and Systems of the Inter-University Research Institute based on the National University Corporation Law. The Planning Coordination Chief System was abolished and the position of Vice Director-General was instituted instead. The Dept. of Statistical Science in the School of Mathematical and Physical Science, SOKENDAI, was reorganized. In addition, the Dept. of Statistical Science and the School of Multidisciplinary Sciences were instituted.
2005	April	●	The research organization was reorganized into three research departments (the Department of Statistical Modeling, the Department of Data Science, and the Department of Mathematical Analysis and Statistical Inference). The affiliated Center for Development of Statistical Computing, the Center for Information on Statistical Sciences, and the Engineering and Technical Services Section were integrated into the Center for Engineering and Technical Support. The Risk Analysis Research Center was instituted.
2006	April	●	The Administration Planning Coordination Unit was instituted.
2008	April	●	The Research Innovation Center was instituted.  The Administration Planning and Coordination Unit was reorganized into the Administration Planning and Coordination Section (hereafter APCS), within which the Intellectual Property Unit, the Evaluation Unit and the Information and Public Relations Unit were instituted.
2009	January	●	The Planning Unit was instituted within APCS.
	October	●	The Institute was moved to 10-3 Midori-cho, Tachikawa, Tokyo.

2010	June	●	Officially opened the Akaike Guest House.
	July	●	Reorganized the Administration Office to create the NIPR/ISM Joint Administration Office and launch the General Service Center. The NOE Forwarding Unit (now we call “NOE Promotion Unit”) was instituted within APCS.
2011	January	●	Research and Development Center for Data Assimilation was instituted. Survey Science Center was instituted.
2012	January	●	Research Center for Statistical Machine Learning, Service Science Research Center and School of Statistical Thinking were instituted.
2014	July	●	The URA Station was instituted within the Planning Unit.
	December	●	The Office of Female Researcher Development was instituted within the Planning Unit.
2017	January	●	Survey Science Center and Service Science Research Center were closed.
	July	●	Data Science Center for Creative Design and Manufacturing was instituted.
	December	●	The International Affairs Unit were instituted, and the Gender Equality Unit, which had been within the Planning Unit, reorganized within APCS. The Intellectual Property Unit was reorganized as the Industry-Academia Collaboration and Intellectual Property Unit within APCS.
2018	April	●	Research Center for Medical and Health Data Science was instituted. The NIPR/ISM Joint Administration Office was reorganized as the Tachikawa Administration Department of the Research Organization of Information and Systems (ROIS).
2019	March	●	Research and Development Center for Data Assimilation was closed.

# The Institute of Statistical Mathematics



#### Access to the ISM

- ◎ Tachikawa Bus
  - Tachikawa Academic Plaza bus stop
  - 5 min walk from Saibansho-mae or Tachikawa-Shiyakusho bus stop
- ◎ Tama Monorail
  - 10 min walk from Takamatsu Sta.
- ◎ JR Chuo Line
  - 25 min walk from Tachikawa Sta.

 Research Organization of Information and Systems  
**The Institute of Statistical Mathematics**  
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan  
Tel: +81-(0)50-5533-8500 Fax: +81-(0)42-527-9302  
<https://www.ism.ac.jp/>