#### TOKYO MEDICAL AND DENTAL UNIVERSITY







Tokyo Medical & Dental University M&D Data Science Center 5th Fl., Building 22, Surugadai Campus, Tokyo Medical and Dental University, 2-3-10 Kanda Surugadai, Chiyoda-ku, Tokyo 101-0062 Japan E-mail mddsc.dsc@tmd.ac.jp URL https://dsc.tmd.ac.jp/en/home/



October 2023





### **TOKYO MEDICAL AND DENTAL UNIVERSITY**



# A center of gravity for biomedical data science in the world







### Data Science will open up the innovative world of medicine and dentistry.

**Center Director** Specially Appointed Professor, Tokyo Medical and Dental University

### **Satoru Miyano**

In addition to the University Hospital, Tokyo Medical and Dental University has a Bioresource Research Center, an intractable diseases treatment cluster, and a large amount of data on clinical practice in medicine and dentistry. The M&D Data Science Center will serve as a national model for human resource development and research by utilizing the data held by the University and establishing three facets for basic, applied, and clinical development: the M&D Data Science Platform, M&D Data Science Implementation, and M&D Data Science Outcome, respectively.

The instructors at the Center are an array of researchers working in areas that include information science, statistical science, computational science, and ethical, legal, and social issues (ELSI).

Our goal is not only to promote cutting-edge research and clinical implementation, but also to develop human resources who will lead society as entrepreneurs.



Division of AI and Big Data Research Division of Multidimensional Biological Data Division of Health Intelligence Division of Data Science Algorithm Design and Analysis **Division of Medical Statistical Mathematics Research** Division of Data Science ELSI Research Division of Smart Hospital Division of Data Science Strategy

Medical and dental research and medical care are entering a new phase in terms of the as part of digitalization and globalization. This particularly applies to biomedical research, in which vast amounts of data can now be obtained in genome information analysis and one-cell analysis, and the possibility of creating new solutions through integrated analysis of this information is increasing. In the medical field, the advanced concept of a smart hospital is becoming a reality, with integrated analysis of vast amounts of medical information and its application in medical treatment, as well as the construction of networks within hospitals and with outside institutions. In addition to spatial expansions such as telemedicine using ICT technology and global public health, temporal expansions are also taking place in areas such as preventive medicine and preemptive medicine. In response to this situation, the M&D Data Science Center was established in April 2020 as a model case for spearheading new medical and dental research, medical care, and education at the University in the Society 5.0 era from the perspective of data science.

We have been carrying out research and collaboration based on and driven by data science through the promotion of collaborative research within the University and provisioning of data analysis infrastructure and services including SHIROKANE, a supercomputer installed at the Institute of Medical Science of the University of Tokyo, and SHIRAUME, a storage server running at the University.

Other than playing an important role in a university-wide effort to build a platform for the accumulation and provision of comprehensive, high-quality medical data owned by the university hospital and Bioresource Research Center, we are also promoting education in data science and AI for all undergraduate students and upgrading data science education in graduate schools.



The M&D Data Science Center is a university-wide organization under the direct control of the President, and consists of 8 divisions containing 5 core departments and 10 cooperative departments.

## **Divisions/Departments**

### **Division of AI and Big Data Research**

#### **Department of Integrated Analytics** (Core)





Associate Professor Takanori Hasegawa

Specially Appointed

Satoru Miyano

Professor



**Project Researcher** Yoko Tanaka

**Technical Assistant** Asako Suzuki

Visiting Professor Seishi Okamoto (Fujitsu Ltd.)

Visiting Professor Koii Maruhashi (Fujitsu Ltd.)

We develop new mathematical methods using bioinformatics and artificial intelligence technologies for integrated analysis of multidimensional and ultra-heterogeneous big data in the medical and biological field. Through the analysis using bioinformatics, statistical science, and machine learning on advanced information processing technologies such as supercomputers and GPUs, we aim to elucidate biological and life systems and immune profiles for cancer, aging, and chronic critical illness for the development of drug discovery and new therapeutic strategies.

#### **Research Topic**

- 1. Knowledge discovery from big data using supercomputers and artificial intelligence technology
- 2. Elucidation of biological/life systems and immune profiles of cancer, aging, and chronic critical illness by mathematical modeling
- 3. Large-scale data analysis of whole genome/exome, single-cell transcriptome, proteome data. etc.
- 4. Systems modeling and simulation of disease states

Development and application of mathematical methods for biological big data The pathophysiology of the disease is a situation in which control abnormalities affect each other in a complex manner due to the influence of multiple genes and the environment including the



intestine and skin, and in particular, cancer is a state that deviates from integrated control as a system. We integrate ultra-multidimensional and ultra-heterogeneous biological big data such as genome and other omics data, pathological conditions, and environmental data through the development of advanced computational science strategies and information processing technologies such as supercomputers. Our research will elucidate the mechanisms of biological/life system failures, called pathological conditions, and to develop drug discovery and therapeutics.

#### **Joint Research Coronavirus Task Force**

As a member of the Collaborative Research Coronavirus Task Force, an emergency project to protect society from novel coronaviruses, we have conducted all preliminary whole genome and transcriptome analyses. Data from this task force is increasing daily, and by FY2022, over 1,000 whole genome samples have been completed; another 4,000 samples will be analyzed in FY2023, resulting in 10,000 whole genome data including controls.

#### **Construction of integrated analysis platform for** M&D data and promotion of cooperation system

Medical and dental sciences and medical research are entering a new era in terms of both quality and quantity of information. It is common to acquire and use large amounts of biological information, such as genomic information analysis and single cell analysis, which can be integrated to obtain new knowledge and new findings. We will then promote the construction of the necessary information education and infrastructure.

#### **Department of** AI Technology **Development (Core)**

Specially Appointed Professor (Acting Department Head) Satoru Miyano



Technical Assistant

Momoko Suguyama

biomedical data analysis.

### medicine

To improve efficiency of medical treatment and preventive care, genomic personalized medicine has been drawn a large amount of attention. One of our research topics is developing computational strategies for patient specific analysis and provide data-driven evidence to achieve effective precision medicine.

### **Network Biology**

Gene regulatory network is crucial for understanding complex disease mechanisms because the molecular mechanisms of diseases involve many genes intricately connected in a molecular network rather than the abnormality of a single gene. We conduct network biology studies, such as development of computational approaches for gene regulatory network estimation and network-based marker identification, especially we focus on interpretation of multilaver massive networks.

#### **Department of Medical Data Analytics (Cooperative)**

Ukihide Tateishi Professor (Department of Diagnostic Radiology and Nuclear Medicine) Yoshikazu Nakajima Professor (Department of Biomedical Informatics)

### **Division of Multidimensional Biological Data**

#### **Department of Multiomics (Cooperative)**

Tetsushi Furukawa Executive Director (Director, Institute of Research) Jun Nakabayashi Professor (Department of Mathematics)

#### Department of Genomics (Cooperative)

Toshihiro Tanaka Professor (Department of Human Genetics and Disease Diversity)

#### **Department of Transcriptomics (Cooperative)**

Itoshi Nikaido Professor (Department of Functional Genome Informatics)



Medical big-data analysis and evidence-based personalized medicine are crucial to settle on social issues that we face, such as the population decline and super-aging challenge. Our mission of Department of AI Technology Development is to develop the statistical and artificial intelligence methodologies that can be used in biomedical research and to identify evidence if precision medicine based on biomedical data analysis.

The main research areas of the Department of AI Technology Development include: 1. Development of the statistical and artificial intelligence methodologies that can be used in biomedical research; 2. Uncovering complex mechanism of diseases (e.g., cancer) based on



#### **Develop statistical and artificial methodologies for precision**

## **Divisions/Departments**

# **Division of Health Intelligence**

#### **Department of Al Systems Medicine** (Core)



Junior Associate Professo Satoshi Ohno

Professor

Hidevuki Shimizu



**Project Researcher Daisuke Koga** 

Technical Assistant Yasunari Wakizaka We approach questions from the viewpoint of medical professionals by integrating largescale data acquisition, mathematical modeling, and validation experiments. We are also investing heavily in science and medicine's future by emphasizing education inside and outside our laboratory. Please visit our website (https://shimizuhideyuki-lab.org/) for more information.

#### **Research Projects**

#### AI-based stratification for personalized medicine (clinical informatics)

Based on our expertise as physicians, we are developing analytical methods and tools to address issues unique to healthcare data and conducting research for personalized medicine. For example, by using AI to learn breast cancer gene expression data, we developed a method to stratify stages based on only 23 genes, which was broadcasted in NHK news and Nikkei newspaper.

#### **AI Drug Discovery**

We developed LIGHTHOUSE, an AI drug discovery system that "illuminates" target compounds from a vast chemical space. This study was widely broadcasted in many media. We are extending LIGHTHOUSE to develop a drug discovery platform that incorporates the laws of physiochemistry. We have begun to take on the challenge of designing not only small-molecule drugs but also more complex molecules such as nanobodies.



**Understanding Multifactorial Diseases by Integrating Multi-omic Data** 

We aim for a comprehensive understanding of diseases as a failure of biological systems using large-scale data.

We take a systems biology approach with differential equations and other applied mathematical methods to elucidate disease mechanisms and to predict drug targets.

### **Division of Data Science Algorithm Design and Analysis**

#### **Department of Data Science Algorithm Design and Analysis** (Core)



Assistant Professo Diptarama Hendrian

Our research focuses on developing algorithms and data structures for efficiently processing and analyzing large scale data.

#### String Algorithmics

Algorithms are an essential component of data science, especially for handling large data sets. Our aim is to design algorithms and data structures that are both effective and efficient, to help manage and analyze various types of large-scale medical data such as sequence and sensor data. Our main research topic is on time and space efficient algorithms for pattern matching/search/discovery, data compression, and compressed data processing. We also study combinatorics on words as a theoretical background for algorithmics on strings.

#### Pattern matching/search/discovery

We study efficient algorithms and index structures for pattern matching and string information retrieval, used for example in keyword search, as well as finding patterns that characterize the data, based on various appropriate definitions of "occurrence".



#### Data compression and compressed data processing

Through data compression, we aim not only to save storage space and communication bandwidth required for storing and transferring the data, but also to "compress" even the time required for computation, by processing the data "as is" in compressed form.





### **Divisions/Departments**

### **Division of Medical Statistical Mathematics** Research

We aim to contribute to M&D sciences,

through developing and practicing effective statistical methods for data

#### **Department of Biostatistics (Core)**



Junior Associate Professo Tatsuhiko Anzai

Professor



analysis to solve statistical problems from a wide spectrum of medical, Kunihiko Takahashi dental, and healthcare research.

#### **Development of biosta**tistical methods for M&D data science

Our department is carrying out methodological research in biostatistics, including geographical data analysis in spatial epidemiology, disease surveillance and monitoring of infectious, real-world data analysis of complex and unistructural data in medicine, meta-analysis for integration of evidence and models, dynamic modeling in health science. We also develop software and are exploring application thereof in various fields in healthcare.

#### **Practice of Biostatistics** in M&D science

Our department are engaged in many projects in collaboration with medical,

dental, and healthcare professionals and scientists. We are bridging biostatistical design and analysis to a wide variety of statistical problems in clinical research, clinical trials, observational studies, and laboratory research in medicine, and aims to contribute to health sciences through biostatistics in M&D data science by effective use of data.

**Department of Biostatistics (Cooperative)** Akihiro Hirakawa Professor (Department of Clinical Biostatistics)

**Department of Statistical Genetics (Cooperative)** Yuta Kochi Professor (Department of Genomic Function and Diversity)

**Department of Epidemiology (Cooperative)** Takeo Fujiwara Professor (Department of Global Health Promotion) Jun Aida Professor (Department of Oral Health Promotion)







### **Division of Data Science ELSI Research**

**Department of Data Science ELSI Research (Cooperative)** Masayuki Yoshida Professor (Department of Life Sciences and Bioethics)

# **Division of Smart Hospital**

**Department of Smart Hospital Development (Cooperative)** Rvuii Koike Professor (Health Science R&D Center) Masumi Ai Professor (Department of Insured Medical Care Management)

### **Division of Data Science Strategy**

**Department of Data Science Strategy (Cooperative)** Kaori lida Professor (Open Innovation Center) Atsuhiro Kinoshita Professor (Department of Educational Media Development)

### Visiting Academic Staff

Heewon Park Visiting Professor (Sungshin University)

M&D Data Science Center 9



### **Graduate School Admissions**

Each of the M&D Data Science Center's core departments has established education and research departments in the University's Graduate School of Medical and Dental Sciences, and accepts graduate and research students. The Graduate School of Medical and Dental Sciences welcomes students from diverse educational backgrounds, not limited to medicine and dentistry and students enrolled in the Master's Program: Health Sciences and Biomedical Engineering (2 years) can continue their studies to the Doctoral Program: Medical and Dental Sciences Track (4 years) or Doctoral Program: Biomedical Sciences and Engineering Track (3 years).



#### Master's Program: Health Sciences and Biomedical Engineering

Integrated Analysis/Biostatistics/AI Systems Medicine/AI Technology/Data Science Algorithm **Design and Analysis** 

Applicants must have graduated or be expected to graduate from a four-year undergraduate program (or be recognized as having equivalent or higher academic ability).

For the Master's Program: Health Sciences and Biomedical Engineering, students must contact and receive permission from prospective supervisors in their desired department before applying to take the entrance examination. Students can obtain a Master's degree (Medical Science, Dental Science, Oral Health Care Science, Science, Engineering, Medical Laboratory Science) by submitting a master's thesis within the term.

#### **Doctoral Program: Medical and Dental Sciences Track**



Integrated Analysis/Biostatistics/AI Systems Medicine

Applicants must have graduated or be expected to graduate from the School of Medicine, School of Dentistry, School of Veterinary Medicine (6-year program), or School of Pharmacy (6-year program), or have completed or be expected to complete a master's degree (or be recognized as having equivalent or higher academic ability). Students can obtain a Doctoral Degree (Medical Science, Dental Science, Mathematical Medical Science, Philosophy) by submitting a doctoral thesis within the term.



#### **Doctoral Program: Biomedical Sciences and Engineering Track**

AI Technology/Data Science Algorithm Design and Analysis

Applicants must have completed or be expected to complete a master's degree (or be recognized as having equivalent or higher academic ability). Students can obtain a Doctoral degree (Science, Engineering, Medical Laboratory Science, Oral Health Care Science) by submitting a doctoral thesis within the term.



Haneda Airpo