



国立大学法人
東京医科歯科大学
TOKYO MEDICAL AND DENTAL UNIVERSITY

TOKYO MEDICAL AND DENTAL UNIVERSITY

M&D

Data
Science
Center

M&Dデータ科学
センター

医歯学研究・医療・教育をデータサイエンス面で推進する

TOKYO MEDICAL AND DENTAL UNIVERSITY

M&D

Data
Science
Center



国立大学法人
東京医科歯科大学
TOKYO MEDICAL AND DENTAL UNIVERSITY

国立大学法人 東京医科歯科大学
M&D データ科学センター

〒101-0062
東京都千代田区神田駿河台 2-3-10
東京医科歯科大学駿河台キャンパス
22号館 5階
E-mail mddsc.dsc@tmd.ac.jp
URL <https://dsc.tmd.ac.jp/>



2023.10 発行

TOKYO MEDICAL AND DENTAL UNIVERSITY

M&D Data Science Center



ご挨拶

データサイエンスで 医歯学の世界を切り拓こう

M&Dデータ科学センター長
東京医科歯科大学特任教授

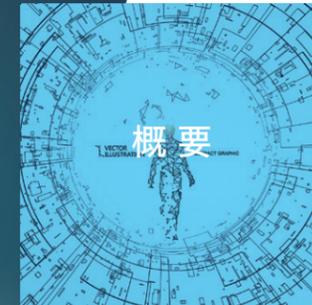
宮野 悟

東京医科歯科大学には大学病院の他、疾患バイオリソースセンター、難病治療クラスターなどがあり、医科と歯科の臨床に関わる大量のデータがあります。

M&Dデータ科学センターは基礎、応用、臨床での展開に向け、それぞれ「M&Dデータ科学基盤系」「M&Dデータ科学実践系」「M&Dデータ科学アウトカム系」の3つの系を置き、本学が保有するこれらのデータを活用し、全国モデルとなるよう人材育成や研究を担っていきます。

教員は、情報科学、統計科学、計算科学、倫理的・法的・社会的問題(ELSI)を含む領域の研究者らが集まっています。

最先端の研究や臨床への実装を進めるだけでなく、「アントレプレナー(起業家)として社会を牽引する」人材育成を目指します。



概要

- AI・ビッグデータ研究部門
- 生体多元情報研究部門
- ヘルスイテリジェンス部門
- データ科学アルゴリズム設計・解析部門
- メディカル統計数理研究部門
- データ科学 ELSI 研究部門
- スマートホスピタル部門
- データ科学戦略開発部門

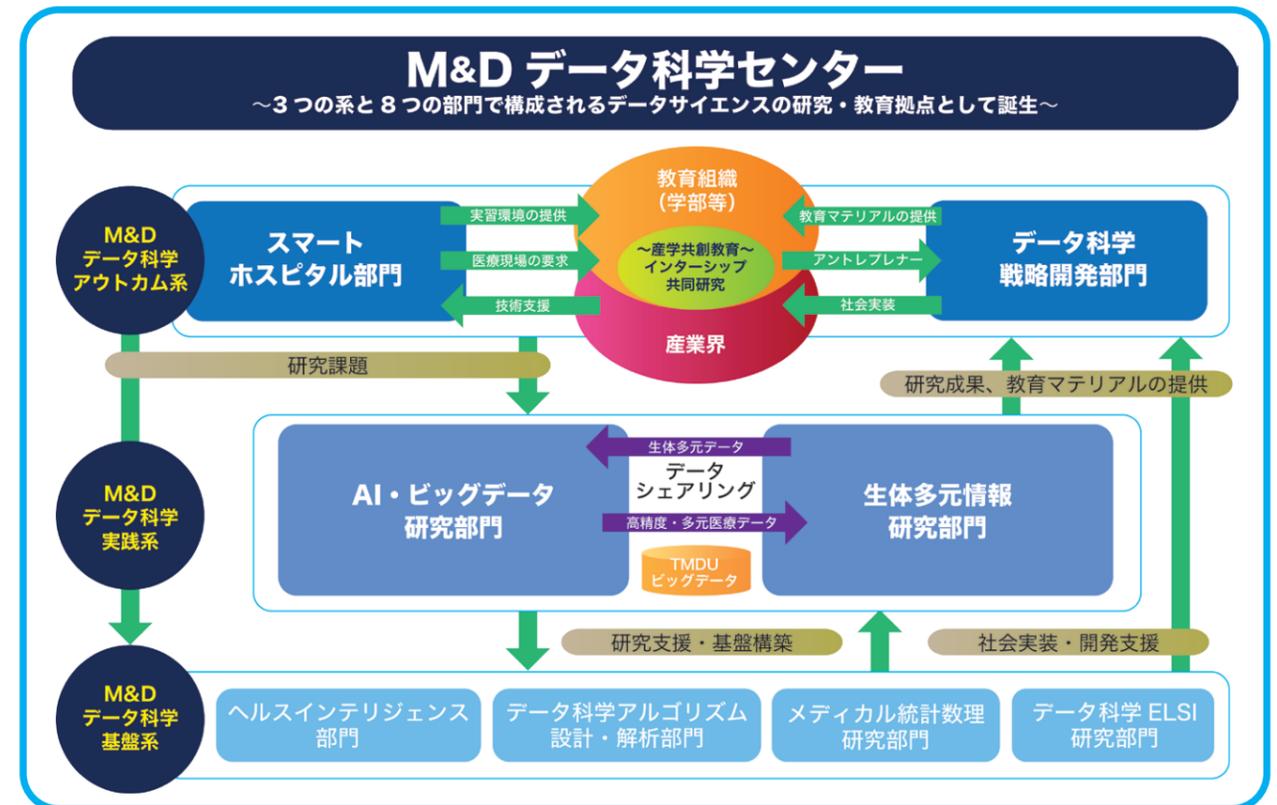
M&Dデータ科学センターは、
学長直轄の全学的組織であり、
5つの基幹分野と10の協力分野を含む
8部門から構成されている。

医歯学研究・医療は情報化・国際化の観点において、新たなフェーズに入っています。特に、バイオメディカル研究においては、ゲノム情報解析や1細胞解析などにおいて大量の情報が取得可能となっており、その情報を統合的に分析して新しい解決策を生み出す可能性が高まっています。また、医療においても、莫大な医療情報を統合的に解析し、診療に活用すると同時に、病院内及び他機関とのネットワーク構築など、先進的な「スマートホスピタル」構想が現実となりつつあります。これらにとどまらず、ICT技術を用いた遠隔医療や、地球規模での公衆衛生などの空間的な広がりに加えて、予防医学・先制医療などの時間的な広がりも起きています。

このような状況を踏まえ、Society5.0時代の新しい医歯学研究・医療・教育を牽引すべきモデルケースとして、本学における医歯学研究・医療・教育をデータサイエンス面で推進する中心的存在として、2020年4月「M&Dデータ科学センター」を設置しました。

現在までに、学内共同研究の推進や、東京大学医科学研究所に設置されているスーパーコンピュータSHIROKANE及び本学で稼働しているストレージサーバーSHIRAUMEを含むデータ解析インフラの構築・サービス提供等を通じ、データサイエンスを基盤かつ駆動力とした研究・連携を展開しています。

また、病院や疾患バイオリソースセンター等が保有する包括的で質の高い医療データを集積して提供するプラットフォームの構築という全学的な取組みにおいて本センターは重要な役割を担っているほか、教育面では全学部生を対象としたデータサイエンス・AI教育の推進や、大学院のデータサイエンス教育の高度化等にも取り組んでいます。



分野紹介

AI・ビッグデータ研究部門

基幹分野
統合解析分野



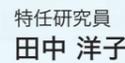
特任教授
宮野 悟



准教授
長谷川 高矩



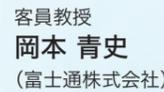
助教
伊東 聡



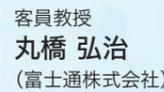
特任研究員
田中 洋子



技術補佐員
鈴木 麻子



客員教授
岡本 青史
(富士通株式会社)



客員教授
丸橋 弘治
(富士通株式会社)

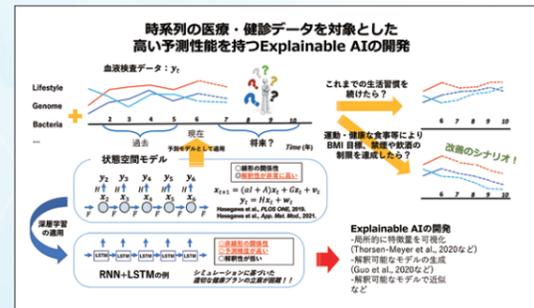
当分野では、医学・歯学などの医療・生命科学分野における多次元・超ヘテロなビッグデータを統合解析するために、生命情報学や人工知能技術を用いた新たな数理的解析手法を開発しています。スーパーコンピュータなどの先端の情報処理技術を用いることで、がんや老化、慢性疾患などを対象とした生体・生命システムや免疫プロファイルの解明と創薬、新たな治療戦略の策定を目指します。

研究

1. スーパーコンピュータと人工知能技術によるビッグデータからの知識発見
2. 数理モデリングによるがんや老化、慢性疾患などの生体・生命システムと免疫プロファイルの解明
3. 全ゲノムやシングルセルトランスクリプトーム・プロテオームデータなどの大規模データ解析
4. 病態のシステムモデリングとシミュレーションの研究

生体ビッグデータの数理的解析手法の開発と応用

疾患の病態は複数の遺伝子や腸内や皮膚を含む環境の影響により制御異常が複雑に相互に影響し合った状況であり、特にがんはシステムとしての統合的制御から逸脱した状態です。統合データ科学分野は、ゲノムをはじめとするオミクスデータ、病態、環境データなど超多次元・超ヘテロな生体ビッグデータをスーパーコンピュータなどの最先端計算科学戦略・情報処理技術を駆使して統合的に解析する数理的方法論の研究とその応用を研究します。それにより病態という生体・生命システムの破綻の仕組みを明らかにし、その成果を創薬や治療法開発へと発展させます。



コロナ制圧タスクフォースのメンバーとして研究を実施

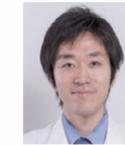
新型コロナウイルスから社会を守る緊急プロジェクトである、コロナ制圧タスクフォースのメンバーとして、全ゲノム・トランスクリプトーム解析の前段階解析の全てを実施しました。本タスクフォースから得られるデータは日々増加しており、本年度段階で1000検体を超える全ゲノムデータが解析されました。2023年度は更に4000検体程度の実施が想定され、コントロール群を加えた10,000単位の全ゲノムデータ解析を実施する予定です。

M&Dデータの統合的解析基盤の構築と協力体制の推進

医歯学・医療研究は、情報の質と量双方の観点から、新たな時代に入っています。ゲノム情報解析や一細胞解析を通して大量の生体情報を取得・利用するのは一般的になっており、それらの情報を統合的に解析することで、新たな知見や疾病の治療戦略を生み出すことが可能になっています。このような状況を踏まえ、統合解析分野では、既設の組織と有機的に連携し、本学の医療ビッグデータの統合的解析を実践すると共に、統合的解析に必要な情報教育と情報基盤の構築を推進します。

基幹分野
AI技術開発分野

特任教授(分野長代理)
宮野 悟



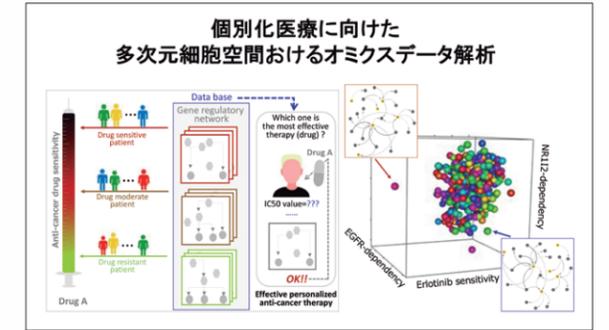
講師
鎌谷 高志

技術補佐員
杉山 桃子

現在、我々が直面している人口減少・超高齢化という大きな課題を解決するためには、メディカルビッグデータに基づく"Evidence-based"ヘルスケアや個別化医療は必修・不可欠です。AI技術開発分野では、統計科学、情報科学、人工知能などの理論・技術に基づき、健康・医療におけるメディカルビッグデータを解析し、がんなどの疾患の複雑なシステムを理解するためのAI・統計モデリング技術開発及び新手法を用いたバイオメディカル研究を行います。

個別化医療向けの統計モデリング・AI技術開発

近年、医療分野においては、患者個々のDNAやRNAを読み取り、得られたデータの解析から抽出された情報の活用に基づいて治療の成功率の向上を目指すゲノム個別化医療 (genomic personalized medicine) の研究が急速に進んでいます。本研究室では、個別化医療に役立てることを目指し、ゲノム情報などの膨大なパーソナルオミクスデータの解析を通じて疾患の複雑なシステムを理解し、個別化医療へのエビデンスを得るための統計モデルやAI技術開発の研究を行っています。



ネットワークバイオロジ

遺伝子の発現制御関係を表す遺伝子ネットワークは、がんなどの複雑な疾病におけるドライバー遺伝子変異の探索などの複雑な疾患のシステムの理解に向けて活用されています。本研究室では、メディカルビッグデータに基づき遺伝子ネットワークを構築し、遺伝子制御パターンによるがんの進化などの疾患のメカニズムの理解、分子標的薬の感受性予測などに関するネットワークバイオロジ研究を行っています。

協力分野 メディカルデータ解析分野

立石 宇貴秀 教授 (画像診断・核医学分野) 中島 義和 教授 (情報医工学)

分野紹介

生体多元情報研究部門

協力分野 多元情報解析分野

古川 哲史 理事 (統合研究機構長) 中林 潤 教授 (数学分野)

協力分野 ゲノム情報解析分野

田中 敏博 教授 (疾患多様性遺伝学)

協力分野 遺伝子発現情報解析分野

二階堂 愛 教授 (ゲノム機能情報学)

ヘルスインテリジェンス部門

基幹分野
AIシステム医科学
分野

教授
清水 秀幸

講師
大野 聡

特任助教
麻生 啓文

特任研究員
古賀 大介

技術補佐員
脇阪 康成

医療者の視点から見た課題に対し大規模データ取得・数理データ科学・生命科学実験による検証という三位一体研究によってアプローチします。また、ラボ内外の教育を重視することで将来の科学・医療界に大きな投資をしています。AIシステム医科学分野では独自の Web サイト (<https://shimizuhideyuki-lab.org/>) も運営しておりますので合わせてご覧ください。

研究プロジェクト

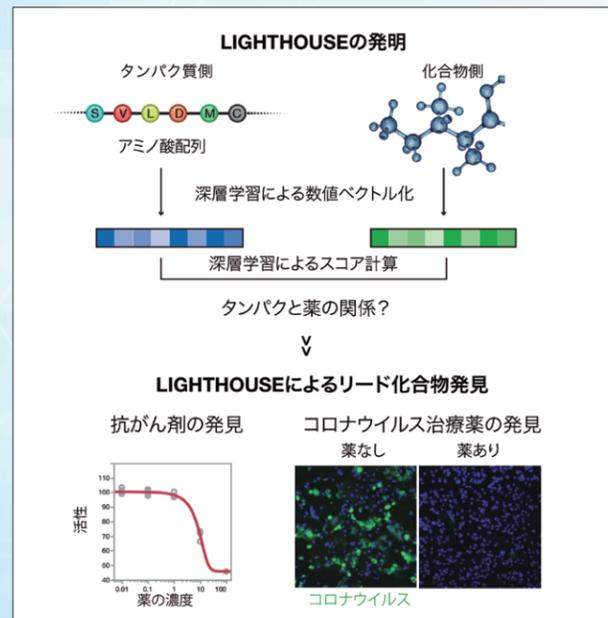
AIを使った臨床患者データ解析 (クリニカルインフォマティクス)

医師としての専門的な知見をベースにして、健康・疾患データならではの課題に対処するための解析手法・ツールの開発や、患者さんの層別化に関する研究を行っています。例えば乳がんの遺伝子発現データを AI で学習することで、23 の遺伝子のみから TNM 分類以上にステージを層別化する手法を開発し、NHK ニュースや日経新聞で報道されました。

AI創薬

広大な化合物の「海」からほとんど手がかりなしに目的の化合物を「照らし出す」AI 創薬システム LIGHTHOUSE (灯台) を開発し、いくつもの新薬を実験的に見つけたことで、NHK などテレビ各社や朝日新聞、読売新聞、ネットニュースに大きく取り上げられました。

現在は LIGHTHOUSE をさらに拡張させ、物理化学法則を取り込んだ創薬基盤プラットフォームの開発をしています。低分子医薬品だけでなく、中分子・抗体医薬といったより複雑な分子デザインにも挑戦を始めました。



大規模データ時代のシステムワイドな疾患研究

生体システムの破綻としての疾患に対し、大規模計測データの数理・情報学的解析によってその病態の理解に深く切り込んでいます。システム生物学的アプローチを指向し、微分方程式やその他の応用数学的手法を駆使して、疾患のメカニズムの解明や創薬ターゲットの予測を行っています。

データ科学アルゴリズム設計・解析部門

基幹分野
データ科学
アルゴリズム設計・
解析分野

教授
坂内 英夫

助教
ディプタラマ
ヘンリアン

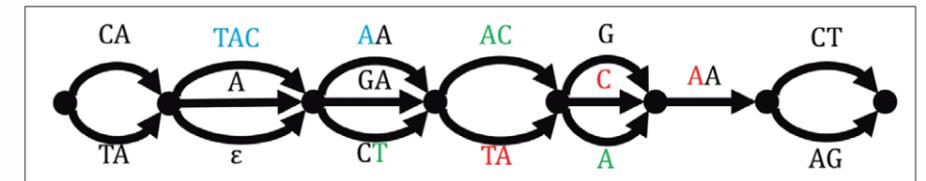
大規模データを迅速に処理・分析・活用するためのアルゴリズムとデータ構造の研究を行っています。

M&D 分野データの処理のためのアルゴリズム設計・解析

M&D 分野において生成される各種シーケンスデータ・センサデータなどの大規模データを迅速に処理・分析・活用するための手法を研究しています。主に文字列・系列データに対して、パターンの照合・検索・発見、データ圧縮・圧縮データ処理など、様々な処理を高速かつ省領域に行うアルゴリズムとデータ構造の設計やその性能解析を行っています。また、それらの手法の理論的基盤となる文字列組合せ論の研究も行っています。

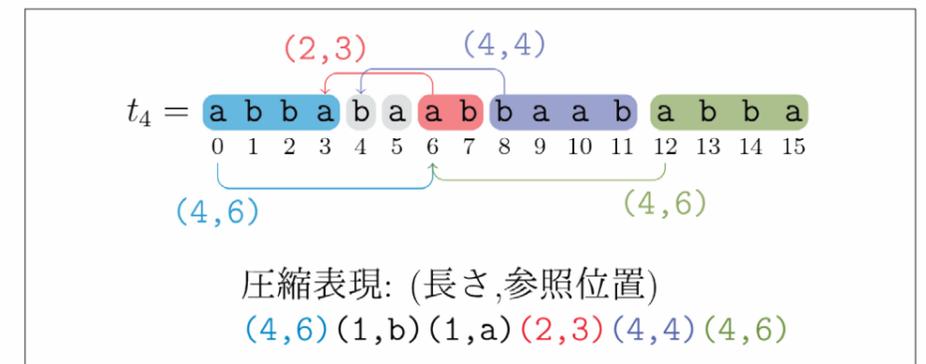
パターンの照合・検索・発見

キーワード検索に代表されるパターン照合・検索の問題やデータの特徴づけるパターンを発見する問題に対し、データの種類や解析の目的に応じてパターンがデータに「現れる」ことを適切に定義した上で効率的なアルゴリズムや索引構造の研究を行っています。



データ圧縮と圧縮情報処理

データを圧縮することでデータの記録や受け渡しに必要な記憶領域・通信帯域を節約するだけでなく、圧縮したまま処理を行うことで処理時間までも「圧縮」することを目指した圧縮方法や処理アルゴリズムの研究を行っています。



分野紹介

メディカル統計数理研究部門

基幹分野
生物統計学分野



教授 高橋 邦彦



講師 安齋 達彦



助教 西田 優紀

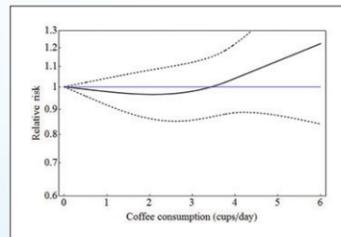
生物統計学 (Biostatistics) の視点から、医学・歯学・医療にフォーカスしたデータ分析のための方法論を開発し、さらに統計学の応用を通して実践的な M&D 領域研究を推進していきます。

M&D 分野データ分析のための方法論開発

M&D 分野のデータ分析のための生物統計学的方法論の開発を行っています。特に地理・時間情報を利用した空間疫学研究、感染症サーベイランス・モニタリング、複雑かつ精緻でない医療リアルワールドデータ (RWD) 分析、多面的なエビデンスやモデルの統合に向けたメタアナリシス、ダイナミックモデリングなどにおける問題解決のための新たな統計的方法の開発とともに、実践的な応用研究、普及に向けたソフトウェアの開発も進めています。



新型インフルエンザ発生時における NYC 保険局サーベイランスデータのクラスタリング解析



用量反応関係のメタアナリシス (コーヒー摂取量と大腸がんの関係)

実践的なM&D 領域研究の推進

医学・歯学・医療分野の専門家と協力し、M&D 領域の研究の発展・さらなる向上を目指した共同研究を推進しています。臨床研究・試験、観察研究、基礎実験研究など様々な研究において、研究設計から統計解析まで協業し、実践していきます。さらにデータサイエンスとしての視点から、新たな横断的研究の提案と実践、健診・検査・診療データなど、M&D データの有効的な活用研究の推進も目指しています。

- 協力分野 生物統計学分野 平川 晃弘 教授 (臨床統計学分野)
- 協力分野 遺伝統計学分野 高地 雄太 教授 (ゲノム機能多様性分野)
- 協力分野 疫学分野 藤原 武男 教授 (国際健康推進医学分野) 相田 潤 教授 (健康推進歯学分野)

分野紹介

データ科学ELSI研究部門

協力分野 データ科学ELSI研究分野
吉田 雅幸 教授 (先進倫理医学分野)

分野紹介

スマートホスピタル部門

協力分野 スマートホスピタル開発分野
小池 竜司 教授 (ヘルスサイエンスR&Dセンター) 藍 真澄 教授 (保険医療管理学分野)

分野紹介

データ科学戦略開発部門

協力分野 データ科学戦略開発分野
飯田 香緒里 教授 (オープンイノベーションセンター) 木下 淳博 教授 (教育メディア開発学分野)

教員紹介

センター客員教員

Heewon Park 客員教授 (韓国・誠信女子大学 准教授)

サービス

大規模計算資源の有償利用

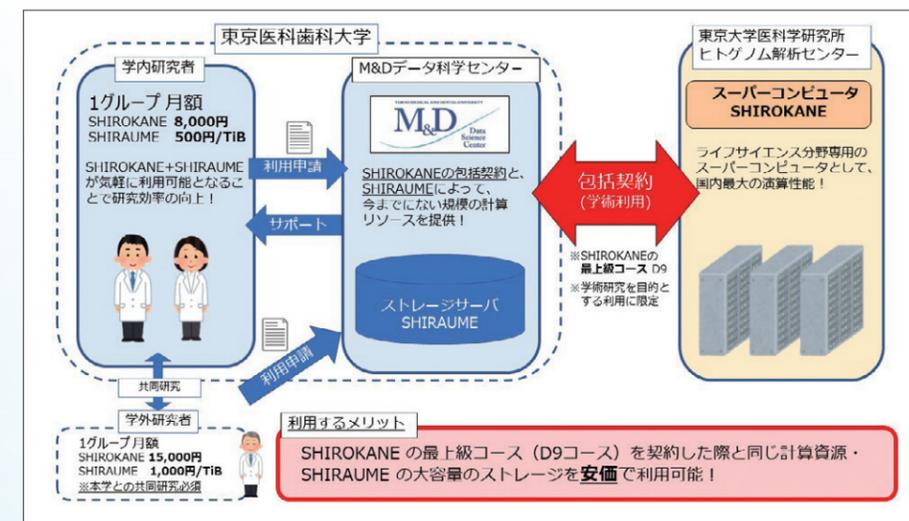
SHIROKANE

SHIROKANE は東京大学医科学研究所に設置されているライフサイエンス分野専用のスーパーコンピュータ (<https://gc.hgc.jp/>) であり、専用コンピュータとしては国内最大の演算性能を持ちます。演算性能は合計で 1,250TFLOPS 以上、ストレージ容量は約 30PB、そのほかにもデータ長期保存用のアーカイブディスクや GPU サーバーなど昨今の生命科学研究に必要不可欠である計算資源を備えています。学術・民間機関向けにサービスを提供しており、多くの生命科学者が共同で利用しています。

SHIRAUME

SHIRAUME は 2021 年 3 月より稼働している、大規模生命・医療データ格納向けに設置されたオブジェクトストレージサーバーです。総容量は 2.5PB であり、全ゲノムデータなどのアラインメントなどを実行させることのできる計算機サーバーも備えています。大学内ネットワークにおいてセキュアに運営されており、SHIROKANE と SHIRAUME はデータの直接的な送受信が可能のため、全ゲノムデータなどの大規模データを SHIRAUME 内で管理しつつ、必要に応じて SHIROKANE に転送し、解析を行うことが可能です。

M&D データ科学センターでは上述の SHIROKANE の利用に関する包括契約を締結しました。本契約の締結に伴い、大規模計算リソースを安価に利用することが可能です。これに合わせて学内ストレージサーバーである SHIRAUME を連携させることにより、大規模生命・医療データの計算・格納資源を十分に提供することが可能です。



お問い合わせ先 SHIROKANE & SHIRAUME サポート係
メールアドレス: dsc_support@ml.tmd.ac.jp

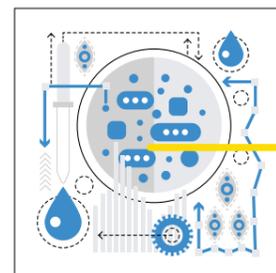


<https://dsc.tmd.ac.jp/service/>

大学院入進学情報

M&D データ科学センターの各分野は本学大学院医歯学総合研究科に教育研究分野を設置しており、大学院生・研究生を受け入れています。

本学大学院医歯学総合研究科では医歯学に限らず様々なバックグラウンドの学生を受け入れるために修士課程医歯理工保健学専攻（2年間）から博士課程医歯学専攻（4年間）もしくは生命理工医療科学専攻（3年間）を継続的に履修する体制が整っています。

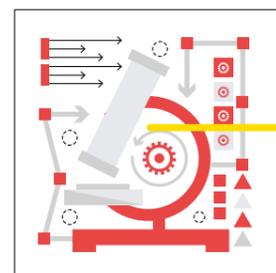


修士課程 医歯理工保健学専攻

統合解析分野 / AI技術開発分野 / AIシステム医科学分野 / データ科学アルゴリズム設計・解析分野 / 生物統計学分野

4年制の学部を卒業あるいは卒業見込みの者（または同等以上の学力があると認められた者）が対象となります。

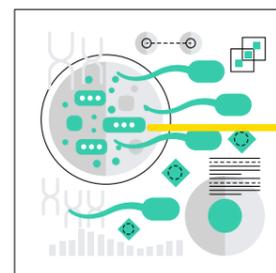
修士課程医歯理工保健学専攻では、出願に先立ち必ず専攻分野の指導教員と面談し、内諾を得たうえで入学試験を受験できます。合格すれば希望分野で研究を行います。（第1志望入学に添えない場合、第2志望入学を認めることがある）年限内に修士論文を提出し、学位は修士（医科学・歯科学・口腔保健学・理学・工学・保健学）を取得することが可能です。



博士課程 医歯学専攻

統合解析分野 / AIシステム医科学分野 / 生物統計学分野

医学部・歯学部・獣医学部（6年制）・薬学部（6年制）の学部卒業生、卒業見込みの者、修士課程修了者、修了見込みの者（または同等以上の学力があると認められた者）が対象となります。年限内に博士の学位論文を提出し、博士（医学・歯学・数理医科学・学術）の学位を取得することが可能です。



博士課程 生命理工医療科学専攻

AI技術開発分野 / データ科学アルゴリズム設計・解析分野

修士の学位を有する者又は取得見込みの者（または同等以上の学力があると認められた者）が対象となります。年限内に博士の学位論文を提出し、博士（理学・工学・保健学・口腔保健学）を取得することが可能です。

本学大学院でデータサイエンスを学ぶ意義



ソフトバンク株式会社 AI戦略室AIエンジニアリング部AIエンジニアリング2課課長

金 大柱氏

2013年3月 九州大学大学院にて博士号（機能数理学）取得後、ソフトバンク株式会社に入社
2013年4月～2018年3月 セキュリティ部門で社内システム開発業務に従事
2018年4月～2019年3月 AI部門で機械学習システム開発業務に従事
2019年4月～2020年9月現在 同部門課長としてOCR製品開発および管理業務に従事

一分野の研究に没頭した経験が広い視野を培い、次の挑戦を成功に導く

Q. データサイエンティストに求められること

メディカルサイエンスでいうと、まずはAIによる画像データ解析によって腫瘍の発現箇所を見極めるといった技術が進んでいます。AIモデルの腫瘍検出精度で本物の医師に匹敵するほどの精度が出ている領域もあり、このような技術は医療の課題を解決する一助になると考えます。また、ゲノムデータの解析では、高次元の統計理論を使い効率的に情報を抽出することで、特定の患者群の抽出を行うことが可能です。さらに、カルテについても完全に電子化して、関係する医療従事者であればセキュアな環境で誰でも見られるようにするといった取り組みも期待されます。

画像診断などで病気の発現箇所を見極めたいなどの予測性能を重視したい場合には、機械学習あるいは深層学習が有効です。事象の解釈が求められる場面では統計的手法を使えばいいでしょう。場面に応じてそれらの手法を使い分けられるスキルが求められていて、かつ適切な手法でデータを分析し結論を導く能力が求められていると思います。

Q. M&D データ科学センターで学ぶメリット

研究志向が強い学生、一つ分野を究めたいという学生には大学院進学を勧めます。また、その大学院で取り扱う内容が独創的であればあるほど、希少価値の高いスキルが身に付くことが期待されます。その点、東京医科歯科大学はメディカル分野・デンタル分野に特化した知見やデータが豊富に蓄積されており、それらとデータサイエンスを融合した研究ができる日本で唯一ともいえる環境です。東京医科歯科大学大学院に行かなければできない貴重な研究に挑戦してほしいですね。

Q. 入学を検討中の方へアドバイス

私が大学院にいたころとは違い、新しい論文が発表されれば論文はWEBで簡単に閲覧でき、その実装が同時に公開されることも多く、運が良ければ動画で内容を詳細に説明してくれている場合もあります。このように勉強や研究の環境が整ってきている反面、自分の独創性を出すのが難しくなっているとも言えますよね。だからこそ他者との差別化を意識し、自分が戦っていく分野を意識しながらスキルを伸ばして行ってほしいと思います。独創性は、自分がとことん興味を持った分野で研究を主体的に進める過程で養われ、発揮できると考えます。

自分が興味を持った一つ分野をとことん研究しつくした経験があると、ほかの分野に挑戦するときにも広い視野で見渡せるようになります。私自身も博士後期課程まで行って研究に没頭したことが、ほかの場面でも生きていくという実感がありますし、弊社内でも、そういう経験がある人は伸びる傾向にあると感じています。

取材日：2020年9月4日