

多様な知の創出を加速するデータサイエンス

講師紹介・講演概要

基調講演 ROISにおけるデータサイエンス促進の取り組みについて

機構長 藤井 良一



講演概要 情報・システム研究機構 (ROIS) のミッションは、極域科学、遺伝学、情報学、統計数理分野で、世界最高水準の先端研究を基盤に学術の発展と大学等の機能強化に貢献すること、そして現代社会の重要な複雑な課題を情報とシステムという視点から捉えなおし、データサイエンスを発展させ駆使することにより解決を目指すことにある。本講演では、本機構で展開する人文学、社会的課題への対応を含む極めて多様なデータサイエンスの取り組みを紹介する。

招待講演1 データサイエンスの可能性と不可能性

総合地球環境学研究所 所長 山極 壽一



講師紹介 1952年東京生まれ。京大大学院理学部卒、同大学院理学研究科博士課程退学、理学博士。人類進化論専攻。2020年まで第26代京都大学総長。現在、総合地球環境学研究所所長、環境省中央環境審議会委員を務める。

講演概要 人間を含む生物の行動や社会を扱う学問は、データサイエンスの進展によって可能になる領域と難しい領域がある。たとえば一見関係がなさそうな分野の情報でもデータを突き合わせてみると高い相関が出る場合がある。一方で、過去のデータのみでは予見が難しい行動や社会の動きがある。データサイエンスはゼロから何かを生み出すイノベーションには向いていないので、そこには人文学やアートの発想を組み合わせる必要がある。

一般講演1 生物多様性を生み出す遺伝基盤

国立遺伝学研究所 教授 北野 潤



講演概要 生物はいかにして多様な環境に適応してきたのか。いかにして様々な種が生まれてきたのか。これら生物学の素朴な問いに、野生生物に遺伝学・生態学・生理学などを導入した分野横断的アプローチで迫っている。本講演では研究室で主に対象としているトゲウオを用いた研究のいくつかを紹介する。

一般講演2 過去約100年間における北極域の気候変化と雪氷圏への影響 - ArCS II北極域研究加速プロジェクトの成果から -

国立極地研究所 特任教授 青木 輝夫



講演概要 北極域における現在の気温上昇は、地球全体の平均に比べて2~3倍の急激なペースで進行している。しかし、過去約100年間の気温変化は単純な右肩上がりではなく、20世紀前半の急激な温暖化とそれに続く一時的な寒冷化が起こったことが知られている。その有力な原因として気候モデルによる最新の研究成果を紹介する。また、そのような気候変化による雪氷圏への影響を、グリーンランド氷床で掘削されたアイスコアの解析と衛星リモートセンシングによる氷床表面状態の変化から考察する。数値モデル、現地観測、衛星データの解析から生み出される多圏相互作用の実態解明という視点でシンポジウムのテーマにアプローチする。

招待講演2 デジタル・ヒューマニティーズによる新たな知の創出 -人間文化研究機構の取組-

人間文化研究機構 機構長 木部 暢子



講師紹介 専門は言語学、日本語学。南九州をはじめとする日本各地のアフセント・音韻を研究するほか、消滅危機言語の記録・保存にも取り組む。2010年4月より国立国語研究所教授・副所長。2022年4月より人間文化研究機構長。

講演概要 人間文化研究機構の6機関は、歴史、民俗、民族、文学、言語、環境など多様な分野にわたる資料を有している。資料の種類も文字、絵画、写真、映像、音響、音声など多様である。これらの一部はデジタル化のうねりで公開しているが、そうでない資料も多い。これらをデジタル化し、相互に繋ぐことにより新しい発想を生み出そうというのが人間文化研究機構の第4期中期目標期間における取組である。本講演では人文データの特徴やデータ連結の際の問題等を紹介する。

一般講演3 相手とちょうどいい距離感でチャットできるポライトエージェント

国立情報学研究所 教授 山田 誠二



講演概要 談話におけるポライトネス理論の拡張である「ディスコースポライトネス理論」を応用して、チャット相手と自分の距離感を自動判定することで、円滑なチャットを行うチャットボットであるポライトエージェントの研究について報告する。まず、談話コーパスにラベル付けを行い、そのデータを訓練データとして、機械学習により距離感のクラシファイアを学習することでポライトエージェントのアルゴリズムを開発、システム実装の後、オンライン参加者実験により評価を行う。

一般講演4 生命医科学における多様なデータの統合と俯瞰からの探索的な知の創出に向けて

データサイエンス共同利用基盤施設 特任准教授 片山 俊明



講演概要 生命活動の基本であるタンパク質の立体構造が、膨大なデータを学習した人工知能によってある程度予測できる時代が来るなど、生命医科学においてもデータ科学への期待は大きい。毎年100万本以上の論文が発表される生命医科学分野において何が既知で何が未知なのかを把握し、新しい研究を促進するためにも、既存データの統合整備は研究のインフラとして必須の事業である。ライフサイエンス統合データベースセンターでは、日本人ゲノム変異の統合データベースTogoVarや、多様なデータを知識グラフで統合し俯瞰的に探索するTogoDXなど画期的な技術開発を行っており、ここから多分野にまたがるデータサイエンスを推進していきたい。

一般講演5 なぜ今、統計エキスパート人材育成か？

統計数理研究所 特任教授 千野 雅人



講演概要 「データサイエンス」から「統計学」が欠けると「危険な領域」に陥ると言われるように、統計学はデータサイエンスに不可欠な学問である。現に米国では「統計家」(Statisticians)が地位の高い職業として確立しており、統計家は、米国で成長が期待される職業の第4位(米国政府推計)、Best Business Jobs ランキングの第3位(U.S.News)である。一方、日本は、統計科学を専攻する人材が極めて少なく、危機的な状況にあると言われてきた。最近に至り、全国の大学にデータサイエンス系学部学科が次々と新設されるようになったが、ここで顕在化した大きな問題が統計学教員の不足である。これにいかに対応するかを紹介する。