

コンピュータ科学専攻各教員研究室紹介

研究室名：小林研究室

教授 小林 直樹

本研究室では、ソフトウェアに関する基礎理論とそのプログラム検証などへの応用についての研究を行っています。近年、交通システムや医療機器、電子商取引など世の中の多くのシステムがコンピュータソフトウェアによって制御されており、それらに欠陥があれば大惨事になりかねません。その一方でソフトウェアはますます大規模化・複雑化しており、テスト実行など従来のソフトウェア開発手法では品質の担保が難しくなっています。そこで、**数理科学的な手法を用いて機械的にプログラムの検証や変換を行うことによってソフトウェアの信頼性や性能を向上しよう**、というのが研究目的です。また、そのような目的を達成するためには、型理論、形式言語とオートマトン、定理証明など、理論計算機科学を深く学び、研究する必要があります。一見理論的興味の対象にしか見えない数学的概念が実は上記のような応用に結び付くことを実感できるのが本研究室で行っている研究の醍醐味でもあります。

以下、最近の研究テーマの例です。

1. 高階モデル検査：システム検証技術であるモデル検査の拡張です。世界初の高階モデル検査器の開発に成功して下の2,3のテーマに応用しています。
2. プログラムの自動検証：上記1の高階モデル検査や型理論などを利用して、OCamlやRust, C言語などのためのプログラム自動検証器を作っています。
3. データ圧縮：文字列や木構造データをそれを生成するプログラムの形式で圧縮し、1の高階モデル検査を利用して圧縮したままパターンマッチングなどのデータ操作を可能にする、というテーマです。
4. プログラミング言語および形式言語の理論：上記1～3の基礎理論として、型理論やラムダ計算を始めとするプログラミング言語理論や高階文法・不動点論理などについて研究しています。
5. セキュリティプロトコルの自動検証：インターネットショッピングなどで使われる、暗号を用いてパスワードなどの機密情報をやりとりする通信プロトコルの安全性を自動検証する、というテーマです。

詳細は <https://www.kb.is.s.u-tokyo.ac.jp> をご覧ください。

研究室名：五十嵐研究室

教員名：教授 五十嵐 健夫

(1) **ユーザインタフェース**：パソコンやスマホからロボットや自動運転車に至るまで、広く情報機器と人間のかかわる部分についての研究を行っている。これらの機器を使いやすくするための技術を開発する他、新しい使い方を提案したり、使われ方を分析して新しい知見を得たりすることも行っている。

- 機械学習や人工知能システムのためのユーザインタフェースの研究。特に、訓練データの効率的な作成手法や、対話的な学習手法などの開発。
- スマートフォンやスマートウォッチ、AIスピーカー、眼鏡型デバイスなど、新しい機器のためのインタラクション手法の開発
- ロボットや自動運転車といった実世界で活動する情報システムに効率的に指示を出したり、制御したりするためのインタラクション手法の開発。



(2) **グラフィクス**：主に対話的な形状処理を対象として研究を行っている。伝統的なグラフィクスとして、3次元形状を作成したりアニメーションさせたりするための技術開発に取り組む他、最近では3次元プリンタなどを利用して現実世界のものの形状を対象としたファブリケーションの研究も行っている。特に、専門家でない一般のユーザにとって使いやすいツールの開発を中心に行っている。

- 手書きスケッチや機械学習といった技術を活用して、簡単に3次元形状やアニメーションなどを作成するための技術の開発。医用画像を対象とした研究も行っている。
- 物理シミュレーションを活用して、楽器や紙飛行機、竹とんぼなど、実世界で機能を満たすモノの形状をデザインするためのシステムの開発
- 3次元形状をスキャンして計算機に取り込んだり、生成した形状を3次元プリンタで出力したりする際に必要となる技術の開発。



意欲のある学生には、海外の大学や研究機関、医療機関や映像プロダクション、企業などと共同で研究を行う機会を提供することも可能である。 <http://www-ui.is.s.u-tokyo.ac.jp/>

研究室名：宮尾研究室

教員名：教授 宮尾 祐介

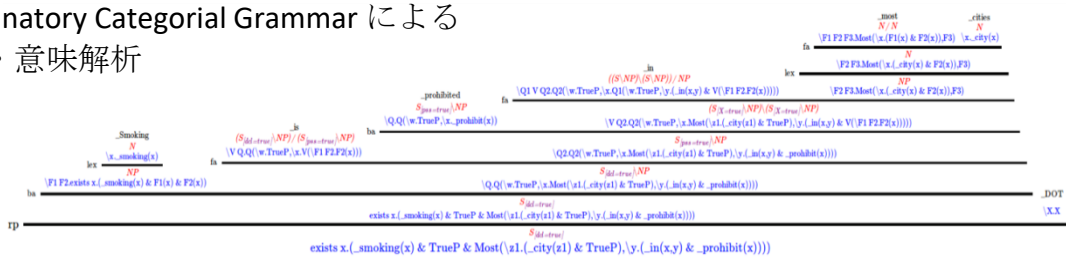
自然言語処理を中心に、人間が情報を理解するしくみについて研究しています。

本研究室では、自然言語処理・計算言語学とよばれる研究を行なっています。人間は、ことば（自然言語）を使ってコミュニケーションし、情報を理解し、行動しています。そのしくみを、数理モデルや言語理論を駆使して明らかにすることを目指しています。特に、コンピュータサイエンスや言語学の理論に基づいて自然言語を定式化・分析する基礎研究と、それに立脚した応用研究を中心としています。例えば、文の構造を計算する構文解析、意味を計算する意味解析、大規模言語モデルの言語理解能力の分析・評価などの研究を行っています。

自然言語処理の応用研究としては、自然言語を用いてコンピュータと情報をやりとりする対話システムや、画像や数値データなど世の中のさまざまな非言語情報を自然言語で説明するデータ・テキスト変換の研究を行なっています。

詳細はホームページを参照してください：<https://mynlp.is.s.u-tokyo.ac.jp/ja/>

Combinatory Categorical Grammar による 構文・意味解析



研究室名：佐藤一誠研究室

教員名：教授 佐藤 一誠

我々人間は、過去の経験や現在の状況を抽象化し学習した知識に基づいて行動を決定しています。機械学習は、過去の人間の経験や現在の人間とのやりとりから蓄積されたデータを利用して、機械の性能や行動を向上させる人工知能の一分野です。

現在では主に、以下の4つの研究テーマに取り組んでいます。

1. 汎化と記憶：汎化とは学習データに存在しない未知のデータの予測に関する性質です。記憶とは学習データを記憶領域に留めることを指しますが、単に丸暗記するだけでは汎化には結び付きませんし、学習データを抽象化しすぎても予測性能は低下します。両者の関係を分析することが機械学習研究で最も基本的な問題となります。
2. 摂動と不確実性:学習データやモデルパラメータに対する摂動や予測に対する不確実性が学習アルゴリズムにどのような影響があるのかを研究しています。
3. 表現学習：学習データをどのような抽象的表現で表すことが汎化につながるのかを解明することを目的とした表現学習の研究を行っています。
4. 頑健性：学習データや予測対象となるデータが想定されているものとは異なった場合でも学習が適切に行われるために必要な性質である頑健性について研究しています。

必要とされる知識：

我々の研究室では、確率・統計・線形代数・関数解析・最適化などを道具として機械学習の基礎理論構築を行っています。これらの数学の知識があることを前提に研究を進めます。また、理論を網羅的に実証的に分析することを行うため、Pythonを主としたプログラミング能力も必須となります。

詳細は <https://www.ml.is.s.u-tokyo.ac.jp/> をご覧ください。

研究室名：河原林研究室	教員名：教授 河原林 健一
-------------	---------------

本研究室では、「アルゴリズム」と「離散数学」、そしてその応用を中心に研究しています。とくに「グラフ」を対象として、以下を研究する予定です。

1. グラフを扱う問題に対して、理論的に高速でかつ正確なアルゴリズムを開発、あるいは、NP 困難性などを証明
2. 実世界に現れるグラフに対して、数学などの理論的な道具を利用して、高速に実装可能なアルゴリズムの開発
3. グラフの族を制限する（例えば、平面グラフ、あるいはソーシャルネットワーク）場合において、理論的（あるいは実用的に）高速に動作するアルゴリズムの開発
4. グラフアルゴリズムの知識・実装を機械学習、とくにオンライン学習、深層学習に応用
5. 自然言語処理、機械学習、データベース、データマイニング、プログラム言語分野などで現れるグラフを対象とした研究
6. （組合せ）最適化、および離散数学

本研究室では、意欲のある学生に対して、「理論的」な研究に対しては、海外の理論系トップ研究者との交流を予定し、その他の研究に対しては、企業などを含めた国内外共同研究者との研究も予定しています。

現在の情報検索技術（Google の PageRank）、セキュリティ技術（Apple の (Local) Differential Privacy）などのアルゴリズム革新は国家規模のビジネス創成につながっています。ここで重要なのは、PageRank も Differential Privacy もアルゴリズム、離散数学の基礎・理論研究であり、最初から応用を志向した仕事ではない点です。本研究室でも、最初から応用を念頭におくような研究を行わず、基礎・理論研究中心に行う予定です。 Email: k_keniti@nii.ac.jp

研究室名：品川研究室	教員名：教授 品川 高廣
------------	--------------

オペレーティングシステム (OS) や仮想化ソフトウェアを中心とした**システムソフトウェアの研究室**です。システムソフトウェアは、コンピュータのハードウェアとアプリケーションを繋ぐ重要な役割を持っておりその能力次第でシステムのパフォーマンスが大きく変わってきます。特に OS カーネルやハイパーバイザは、コンピュータ全体を制御できる特別な権限を持っており、他のソフトウェアとは違った能力があります。

またシステムソフトウェア研究は、アプリケーションに対してコンピュータを抽象化・一般化した概念を提供するという科学的側面も持ちつつ、それを実際のハードウェア上で設計・実装・評価して有用性を実証するという工学的側面も併せ持っており、異なる領域の境界に位置する重要かつ魅力的な研究分野です。

品川研究室では、具体的には近年は以下のような研究テーマに取り組んでいます。

- **システムソフトウェアによるセキュリティ向上の研究**
アプリケーションの安全な隔離実行環境を研究しています。例えば AMD SEV-SNP や Intel TDX などの TEE、Cambridge 大学が研究開発した CHERI を活用した軽量 VM、FPGA ベースの RDMA を用いた OS など、高速で堅牢な保護を実現する手法を研究しています。また、Fuzzing によるネステッド仮想化の脆弱性発見や、CPU のアドレスマスク機能を活用した Use-after-free 対策など、様々な方法でセキュリティ向上を目指しています。
- **システムソフトウェアの性能・機能向上の研究**
OS のライブマイグレーションやコンテナの起動高速化、不揮発性メモリ向けファイルシステム、低遅延のクラスタスケジューラ、macOS や Windows 上で Linux バイナリを動作させる OS エミュレータなど、システムソフトウェア全般の性能や機能を向上させるための研究に取り組んでいます。

品川研究室では、興味を持って楽しく研究することを目指しています。低レイヤのシステムソフトウェア (Linux, KVM, BitVisor など) やシステムプログラミング (C, C++, Rust など) が好き and/or 得意な学生さんは特に歓迎しますが、興味があればどのレイヤでも楽しく研究ができると思います。

研究室ホームページ: <https://www.os.is.s.u-tokyo.ac.jp/> E-mail: shina@is.s.u-tokyo.ac.jp

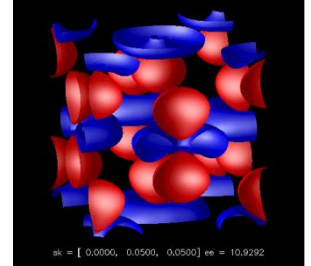
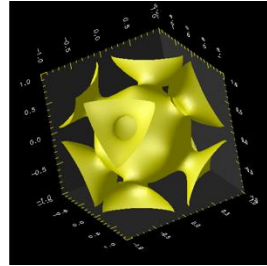
研究室名：吉本研究室

教員名：准教授 吉本 芳英

電子計算機の発明の動機の一つは、科学技術への応用でした。電子計算機が発明されて以降、半導体技術の飛躍的發展：ムーアの法則とともに計算機の能力は飛躍的に向上し、科学を計算によって推進する計算科学は大きな恩恵を受けてきたのですが、並列化技術など半導体技術の限界の顕在化に伴った計算機システムの複雑化のため、今後の発展には再び計算科学と計算機科学の間の密接な協力が必要になってきています。本研究室はこのような背景から計算機科学と計算科学の間をつなぐ教育研究を目指しています。

(<https://www.ep.is.s.u-tokyo.ac.jp/>)

吉本は、半導体、金属、誘電体、磁性体といった物の性質を微視的な観点から明らかにする物性分野を専門としていますが、その中でも物性の大部分を支配している電子の量子力学をコンピュータで写實的にシミュレートする計算手法、第一原理電子状態計算を専門としており、このためのプログラム群 xTAPP を開発、公開しています。右図は電子状態計算の可視化例です。



(<http://xtapp.cp.is.s.u-tokyo.ac.jp/>)

本研究室は、計算機科学と計算科学の境界を立ち位置として電子状態計算をホームタウンとしつつも、

1. 幅広い計算科学分野について、各分野で個々に発展している方法論を計算機科学の観点から捉え直してそれらの相互交流につなげること。
2. 計算科学のニーズを計算機科学の観点から理解してより本質的な解決法を提案すること。を目指したいと考えています。

研究室名：高前田研究室

教員名：准教授 高前田 伸也

コンピュータ・アーキテクチャ・システム研究室 (CASYS: Laboratory for Computer Architecture and Systems)

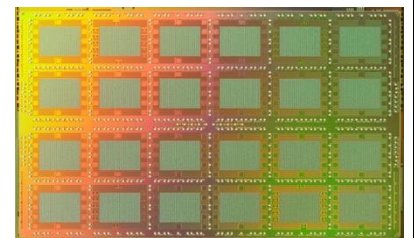
コンピュータ・アーキテクチャを中心に、FPGA や専用チップによるカスタムコンピューティング、機械学習処理などの計算対象アルゴリズムとハードウェアの協調設計、ハードウェア設計を支援する高位合成コンパイラなど、高効率で信頼できる次世代コンピュータに関する研究を進めます。ハードウェアに関する研究だけではなく、コンパイラやプログラミングモデル、コンピュータを扱いやすくするフレームワークなどのソフトウェア研究も取り扱っています。企業等との共同研究を活発に行っており、実際の課題を定量的解析に基づく新規アーキテクチャ技術で解決します。

以下は研究テーマ例です。新規テーマも大歓迎です。

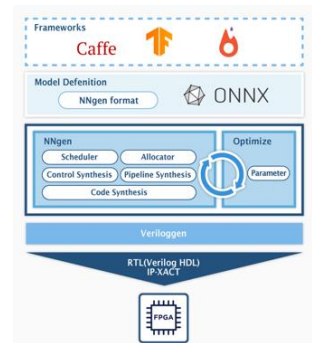
- セキュリティと速度を両立する CPU アーキテクチャ・メモリシステム
- オープンソースなハードウェア設計言語とコンパイラ
- ゆらぎの熱力学に基づく確率的コンピュータアーキテクチャ
- メモリの中で計算する Processing in Memory 回路とアーキテクチャ
- 信頼される AI のためのセキュアな Federated Learning
- 高性能・省エネルギーな AI チップ
- 軽量ハードウェアに適した機械学習アルゴリズム
- 低遅延・省エネルギーな FPGA アクセラレータ

研究室 WEB: <https://sites.google.com/view/casys-ja/>

GitHub: <https://github.com/casys-utokyo/> <https://github.com/PyHDI/> <https://github.com/NNgen/>



DNN プロセッサ QUEST



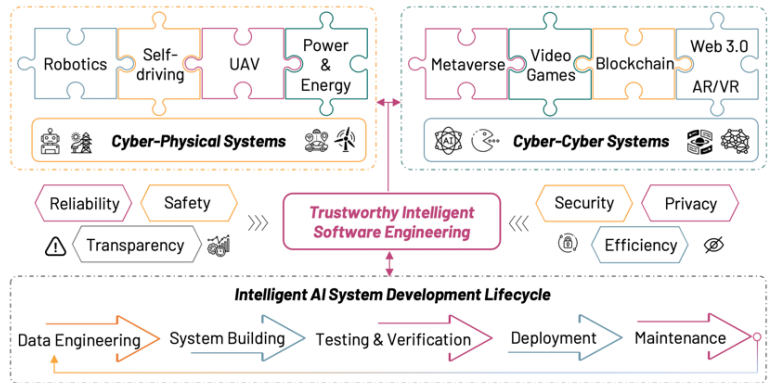
DNN-HW
高位合成コンパイラ
NNgen

研究室名：馬研究室

教員名：准教授 馬 雷

近年、データ駆動型 AI ソフトウェア（深層学習など）は、産業界の様々な分野で華々しい成功を収め、自動運転やロボットなど、社会基盤に関わる重要な分野での成果がますます期待されています。AI ソフトウェアは、産業界における技術イノベーションの継続的な原動力にもなってきており、社会や人々の日常生活のほぼすべての側面に浸透し、支えている一方で、AI ソフトウェアの品質、信頼性、及び安全性の懸念が依然として大きな課題となっています。そのため、ソフトウェアの品質保証は極めて重要であり、社会全体が安全・安心かつ効率的に運営される未来の知能社会である「Society 5.0」の実現に不可欠であると考えられています。本研究室は：

(1) ソフトウェア基礎理論と工学手法の双方に焦点を当て、ソフトウェア工学と AI の新しい学際的な分野（すなわち機械学習工学、AI システム工学、SE4AI とも称される）における、信頼できる AI ソフトウェアシステムを実現するための新しい方法論、工学技術、およびツールチェーンの設計・開発に取り組んでいます。



(2) 提案する信頼される AI システム工学手法と品質保証技術を、様々な実世界の Cyber-Physical Systems と Cyber-Cyber Systems に適用し、各産業分野における AI システムの品質、安全性、信頼性およびそれぞれの限界などを体系的に探求し、AI ソフトウェアシステムを継続的に検証・監視・強化する方法を提案し、安全・安心・信頼できる社会基盤の実現に取り組んでいます。

研究室ホームページ：<https://www.malei.org/lab.html>, Email: malei@is.s.u-tokyo.ac.jp

研究室名：谷中研究室

教員名：准教授 谷中 瞳

私たちは自然言語という記号を介してコミュニケーションを行い、世の中のデータの多くは自然言語で記述されています。情報処理技術の発展に伴い、情報検索や自動翻訳など人に代わって自然言語を処理してくれる人工知能技術は、日常生活の一部となりました。一方で、これらの技術が私たちのように言語の意味を理解して処理しているかという、自動翻訳だけではまだプロの翻訳家にはかなわないことから、その答えは明白でしょう。本研究室では、「より人間のように自然言語の意味を理解する言語処理技術」を目指して、以下の研究テーマを中心に国内外の研究者と協働して理論的・実証的な研究に取り組んでいます。

(1) 理論言語学を中心とした、統計的言語モデルの学際的・多面的分析

現在の自然言語処理技術は大量のテキストデータから統計的に言語モデルを学習する手法が中心ですが、言語モデルは言語の意味をどの程度理解できているのでしょうか？データから思わぬバイアスやヒューリスティクスを学習しているという信頼性の問題が指摘されています。そこで本研究室では、理論言語学、認知科学、哲学といった様々な分野の研究者と協働して統計的言語モデルのふるまいを多面的に分析しています。

(2) 機械学習と記号論理の融合による意味解析・推論技術

言語や画像などのデータの意味を計算処理できる形式で表し、データ間の関係を明らかにする意味解析・推論技術は、情報検索などの基盤を担う技術であり、様々なアプローチが検討されています。機械学習で意味をデータから統計的に学習するアプローチは否定や数量の意味の扱いや解釈性の課題があり、一方で、記号論理では否定や数量の意味を明示的に表現できますが、言い換え知識の扱いなどの課題があります。そこで、2つのアプローチを融合して各アプローチの課題を解決する意味解析・推論技術の研究を進めています。

(3) 人と推論システムとの相互作用による応用技術

人と推論システムとのインタラクションによって、医療テキストや社内文書といった、品質の担保が求められるデータチェックの効率化が期待できます。企業との共同研究を通して、人の暗黙知を形式知に変換して蓄積する手法や、データの種類や用途に応じて推論を行う推論システムの応用技術を研究しています。

研究室ホームページ：<https://ylab.mystrikingly.com/>

研究室名：杉山・石田研究室

教員名：教授 杉山 将
講師 石田 隆

情報通信技術の飛躍的な性能向上に伴い、これまで人間にしかできなかった知的な情報処理が、コンピュータによって実現できるようになりつつあります。本研究室では、「コンピュータはどこまで賢くなれるのか」をテーマに、人工知能分野の**機械学習**とよばれる知的データ処理技術に関する様々な研究課題に取り組んでいます。

(1) **学習理論の構築**

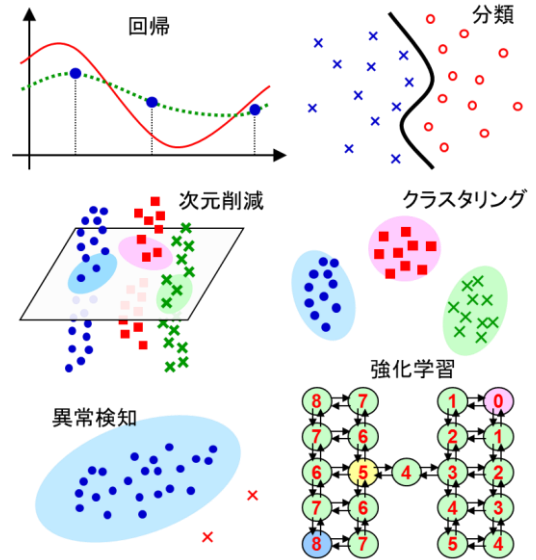
汎化とは、学習していない未知の状況に対応できる能力であり、コンピュータが知的に振る舞うために不可欠です。本研究室では、主に確率論と統計学に基づいて、汎化能力獲得のメカニズムを理論的に探求しています。

(2) **学習アルゴリズムの開発**

機械学習分野には、入出力が対になったデータから学習を行う教師付き学習、入力のみから学習を行う教師なし学習、環境との相互作用を通して最適な行動規則の獲得を目指す強化学習など、様々な課題があります。本研究室では、理論的な裏付けを持ちつつ、実用性の高い機械学習アルゴリズムを開発しています。

(3) **機械学習技術の実世界応用**

インターネットやセンサー技術の発達と普及に伴い、文書、音声、画像、動画、電子商取引、電力、医療、生命など、工学や基礎科学の様々な場面で膨大な量のデータが収集されるようになってきました。本研究室では国内外の企業や研究所と連携し、最先端の機械学習アルゴリズムを駆使して実世界の難問解決に挑戦しています。



<http://www.ms.k.u-tokyo.ac.jp/>

研究室名：横矢研究室

教員名：准教授 横矢 直人

本研究室では、画像処理と機械学習を活用して視覚的情報の取得と理解の高度化に関する研究を行っています。特に、リモートセンシングデータから都市デジタルツインを自動的に構築する基盤技術の開発に力を注いでいます。

(1) **画像逆問題**

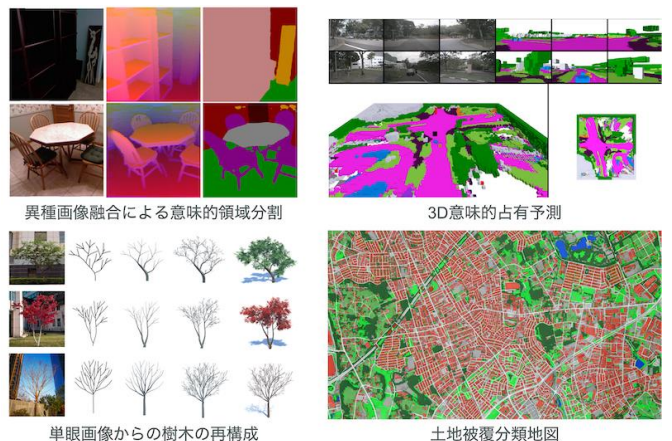
センシングと計算の融合により、ハードウェアだけでは得られない情報の取得が可能となるとともに、解像度やノイズなどのハードウェアの限界を克服することができます。機械学習・最適化・画像処理に基づき、不完全な観測データから元の信号を復元する数理モデルとアルゴリズムの開発に取り組んでいます。

(2) **シーン理解**

光学カメラやLiDARなど様々な種類のセンサデータを統合的に処理し、シーンの意味的情報と3次元情報をより詳細かつ正確に理解する手法を研究しています。また、限られた学習データから機械学習モデルを構築し、計算効率を向上させる方法を探求しています。

(3) **リモートセンシング**

都市デジタルツインの自動構築を目指して、衛星や航空機に搭載されるセンサから得られるリモートセンシングデータを活用し、土地被覆や標高モデルなどの地図情報を自動抽出する知的情報処理の研究に取り組んでいます。

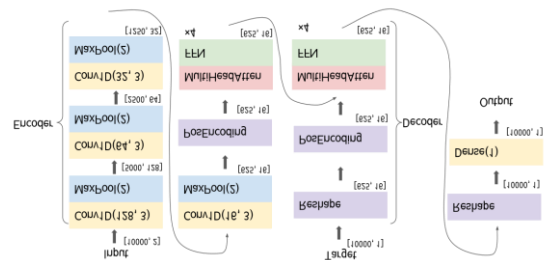


研究室名：中井・朴研究室	教員名：教授 中井 謙太 准教授 朴 聖俊
--------------	--------------------------

白金台キャンパスに位置するバイオインフォマティクスを専門とする研究室の一つです (<https://fais.hgc.jp/>)。私達の研究の原点は、生命の情報がどのような形で一次元文字列のDNA配列として記述されているのかを解き明かしたいという素朴な願いです。その中でも主に、ゲノムにコードされた遺伝子の制御様式を様々な情報科学的アプローチで理解しようとしています。つまり、私たちの体は多種多様な細胞からできていますが、それらはすべて共通のゲノム情報によって指定されており、どのようにしてそんなことが実現されているのかを知ろうとしています。

例えば、自然言語処理と深層学習の組み合わせをスパコン上で設計・実行することで、遺伝子発現を調節する高次ゲノム構造の特徴解析、単一細胞レベルでの時系列・空間的遺伝子発現パターンの解析、遺伝子転写後のRNAスプライシングのメカニズム解明などに挑戦しています。また、実験系専門家との共同研究を通じて、私達の情報計算結果の実験的検証、さらには、様々な難治性疾患治療・再生医療の実用化に役立つ情報解析とアルゴリズム開発を行なっています。これらの研究に欠かせないデータベースの設計と公開・運用にも従事しています。

私達の研究室では、多様な専門分野出身の国際色豊かな構成員が、先端ゲノム情報解析のキーワードを合言葉に、どちらかという基礎生物学志向の強い個別研究を遂行しながら、お互いの研究にも興味をもつような方向を目指しています。すなわち、学生さんの研究テーマ選定にあたっては、なるべく本人の興味や希望を尊重したいと考えています。



研究室所属学生による、mRNA前駆体配列から全てのスプライス部位を予測するモデルの構成案 (宮地、未発表データ)

研究室名：井元・片山・張研究室	教員名：教授 井元 清哉 准教授 片山 琴絵 准教授 張 耀中
-----------------	---------------------------------------

健康医療分野のデータサイエンスの研究室です。全ゲノム、トランスクリプトーム、エピゲノム、腸内細菌叢のメタゲノムなど多種多様なゲノム関連のビッグデータと医療画像・レセプト情報・特定健診・ウェアラブルデバイス等の空間・時間軸を有する健康・医療に関する大規模データに対する統計科学的なデータ解析技術の開発を行います。また、スーパーコンピュータやクラウドを用いた大規模解析を行います。国内外のさまざまな企業、多様な分野の研究者と協力し、健康医療のビッグデータ解析の社会実装を進めています。主要な研究テーマを下に記載します：

研究テーマ1：ゲノムビッグデータ解析技術の開発 ゲノム関連データに特化した機械学習モデルの開発に取り組んでいます。DNA解析技術の急速な進歩により、全ゲノム解読のコストは200ドル未満に大幅に低下し、膨大なデータ量が生み出され、高効率・高精度の解析技術が必要となっています。私たちは、ゲノム関連データに特有の特性をモデルに取り込むことで、ゲノム分野で有用な深層学習モデルを開発しています。最近では、DNA配列やタンパク質配列のための基礎モデルにおいて成果が上がっています。さらに、ロングリードデータにおける構造変異の検出、メチル化異常の検出、シングルセル解析における細胞分類など最新技術によって取得されたデータに対する機械学習モデルの開発にも取り組んでいます。

研究テーマ2：生命システムのモデリング技術開発 ヒトゲノムや共生微生物由来の数万種類の分子が形成する生命システムのモデリング技術、シミュレーション技術を開発しています。薬剤の効果予測や病気の今後の進展の予測等を行う数理解析の技術を開発しています。

研究テーマ3：ゲノムビッグデータを医療に応用するための人工知能の開発 がんゲノムの解析では数千から数万箇所のゲノム変異が同定されます。この中からがんの原因のゲノム変異を同定するには、論文として公開されている情報が用いられます。しかし、生命科学分野における論文は、3,600万報を超え、その全てを網羅することは既に人知人力を越えています。この問題に対して、人工知能を用いた解決について、企業との共同研究や医科研病院での医療実践を伴う研究を行っています。

研究室のホームページ <https://www.hm-intelligence.com/> には、最近の研究室の活動状況や研究内容(上記のテーマ以外にも多数あります)を紹介しています。興味のある方は是非ご覧下さい。

研究室名：渋谷研究室	教員名：教授 渋谷 哲朗
------------	--------------

近年、医科学が急激に変貌しています。その最大の理由は、次世代シーケンサーとよばれる DNA の計測技術の飛躍的進展によってきわめて低コスト・高速に大規模なゲノムデータが得られるようになったことですが、このような超大規模ビッグデータを扱うにあたっては、最先端の情報科学技術の活用が不可欠です。本研究室では、最先端医科学研究をさらにドライブするための新しいビッグデータ基盤技術の開発と、そのような情報技術の実際の臨床研究などへの活用を目標に研究を行っています。

主な研究テーマとしては以下のようなものがあります。

(1) 超高速検索・索引

今後、国民全員規模の超巨大ゲノムデータベースなども出現する可能性もあり、そのような超大規模データを最大限活用することを可能とする超高速検索・索引アルゴリズムの実現が求められています。

(2) プライバシー保護

ゲノムデータは究極の個人情報ともいわれ、その検索、解析、研究においてはゲノムデータの要求する倫理基準を踏まえた最大限のプライバシー保護が求められます。超大規模ゲノムデータにスケラブルな差分プライバシー技術、秘匿計算技術など、新時代のプライバシー保護アルゴリズムの開発が必要です。

(3) 医科学データ解析・学習

現代の医科学研究では、医療データや分子生物学データなどを正確に解析するために様々な情報科学技術の活用することが必須です。実際のデータの性質を最大限活用した医科学データ解析アルゴリズムや機械学習など人工知能技術の開発が求められています。

(4) これらの臨床シーケンス研究への応用

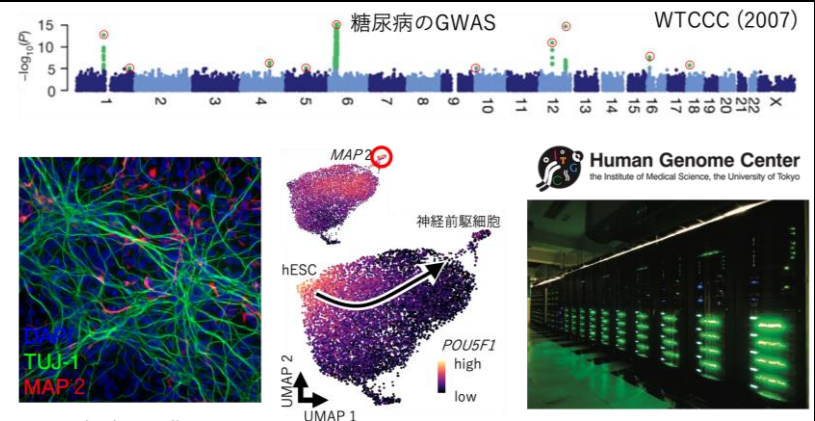
そして、これらの研究を今後の実際の医療の発展、特に実際のゲノムを用いた個人医療などにつなげていくことが期待されています。

基礎医学、臨床医学、分子生物学、創薬といった様々な分野に対し、これらのテーマに限らず、情報科学の観点から新しい世界を切り拓いていきたいと考えています。

ホームページ → <http://shibuyalab.hgc.jp/index-j.html>

研究室名：熊坂研究室	教員名：教授 熊坂 夏彦
------------	--------------

ゲノムワイド関連解析 (GWAS) という手法によって、私達のゲノム上に存在する疾患に関連した責任遺伝子変異が次々と同定されています (図上部)。熊坂研究室では大きくわけて2つのプロジェクトを遂行しています：(1) 小児の発達と疾患発症に関連するような遺伝子変異を網羅的に探索すること、(2) 同定された変異がなぜ疾患の発症や進行に関係しているのかを *in-vitro* (試験管内) のシステムを使って検証すること (東海大学・福田篤研究室との共同研究)、を行っています。どちらの研究プロジェクトも数万にのぼるサンプル数をもった非情に大規模なオミックス (ゲノム、エピゲノム、遺伝子発現) ・データを最新の統計・機械学習技術を用いて解析することが求められます。



ヒトES細胞から作られた神経前駆細胞 GPLVMによるデータ解析 ヒトゲノム解析センター スパコンSHIROKANE

一例として、統合失調症のモデル細胞となる神経前駆細胞をヒト iPS/ES 細胞から作り出し、疾患に関連する遺伝的領域を探索しています (図下段左)。この解析では**ガウス過程潜在変数モデル (GPLVM)** と呼ばれる機械学習手法を用いており (図下段中央)、医科研ヒトゲノム解析センターにあるスパコン SHIROKANE 上でソフトウェアの実装とデータ解析が行われています (図下段右)。

当研究室は 2023 年 10 月に創設された新しい研究室です。情報学の分野からゲノム医科学研究の発展に貢献できるような人材を育成していくことを目指しています。詳しくは研究室 HP をご覧ください：

https://www.ims.u-tokyo.ac.jp/imsut/jp/lab/hgclink/page_00406.html

研究室名：相澤研究室

教員名：教授 相澤 彰子

※2025年10月に博士後期課程入学希望の学生は、相澤研究室は選択できません

本研究室では、「人間の知的活動を支援する言語処理技術」に焦点をあてて、言語テキストを中心とするコンテンツとメディアに関する研究に取り組んでいます。深層学習を含む機械学習、統計モデルと分析、アノテーションとコーパス分析などの手法を基盤として、以下の分野を中心にテーマを設定しています。

- (1) テキストマイニング：文書構造解析や情報抽出、言語資源の自動構築
- (2) 人工知能による言語理解：言語解析による意味構造の抽出や知識の獲得
- (3) 人の言語活動のモデル化：テキストを介した人間やコミュニティ活動の計測とモデル化

具体的な研究課題として、コンピュータによる文章読解、言語理解タスクの設計と分析、対話システムの基盤化、実文書の解析、数式および数学概念の検索と理解支援、情報統合のための情報同定基盤技術、研究者のための情報推薦および英文執筆支援などがあります。また、言語処理や情報検索に関連した新しい研究テーマへの挑戦も歓迎します。

本研究室では、国立情報学研究所が有する大規模な情報資源や分散計算機環境などを利用した研究が行えます。研究室は博士学生から修士学生、海外からの交換留学生やインターンシップ生、訪問研究員など多様なメンバーで構成されていて、さらに研究所内でのセミナーや輪読会への参加などを通して、学際色・国際色豊かな研究者との交流が可能です。また、所内外の共同研究プロジェクトへの参加を通して、研究室という単位にとどまらず、一人の研究者・技術者として活動の幅を広げることを目標としています。

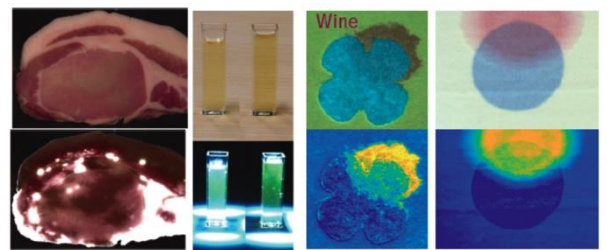
研究室ホームページ：<http://www-al.nii.ac.jp/>

研究室名：佐藤いまり研究室

教員名：教授 佐藤 いまり

光データの獲得／解析による実世界理解

私たちは、ある物体を見ただけで、その物体の材質（金属、ガラス、布等）を推測することができ、その物体の手触りや柔らかさといった付加的な情報も想像することができます。観察される明るさには、対象物体に関する情報（材質や形状など）と周りの環境が複雑に絡み合っているにも関わらず、光情報から対象となる物体の材質や形状を容易に読み解くことができるのはなぜでしょう？人間にとって容易な作業であつても、同じような解析をコンピュータにやらせることは困難です。本研究室



蛍光指紋に基づく
バクテリア可視化 液体識別

散乱解析に基づく状態推定
(上)入力画像、(下)濡れ度合い

では、照明条件（照明色、空間方向、パターン照明等）や観察条件（カメラの方向、紫外から近赤外までの広い波長を考慮）を変えて撮影した画像を通し、被写体となった対象世界がどうなっているのかを最先端の機械学習や人工知能技術を駆使して、人間の視覚情報処理のメカニズムの解明や実世界の情報抽出していきます。医療診断、品質鑑定、wet/dryの解析などの内部状態など基礎理論から実応用も幅広く研究しています。



アクティブラーニングによる血管抽出

研究室からのメッセージ

実世界は複雑だからこそ面白い！画像を通して実世界の事象を解明することは、驚きと発見があります。本研究室は、医学、視覚心理、芸術の研究者との学際研究も積極的に進め、海外からの博士研究員やインターンシップの学生など、国際色豊かなメンバーにより構成されています。充実した光学実験設備も整っています。光を通して実世界を知る楽しみを一緒に味わいましょう。

研究室：<http://research.nii.ac.jp/pbv/index.html>