

# 研究室紹介(?)

たきがわ いちがく

瀧川 一学

<https://itakigawa.github.io/>

4/1～ 東京大学 新領域創成科学研究科 メディカル情報生命専攻

北海道大学 化学反応創成研究拠点 (クロスアポイント)

5/1～ 理化学研究所 革新知能統合研究センター 化学反応情報学チーム (兼務)

4/1～ 京都大学 iPS細胞研究所 (非常勤研究員)

# Takeaways #名前だけでも覚えてって下さい

- 瀧川は4/1着任予定
- 研究室自体はまだない！
- 人もまだいない😞 (卒研で来たら創業メンバー)  
→ 瀧川が直接丁寧に指導できる
- @柏キャンパス (チーバくんで言うと目と鼻のあいだ)
- 関心：「機械学習・機械発見」by 離散組合せ情報の統計科学  
→ “AI for Science” 科学の方法としての「AI」は今めっちゃ楽しい
- 「情報科学を使って生命科学を切り拓く」 のではなく  
「生命科学のための情報科学を切り拓く」



# 瀧川 一学 (たきがわ いちがく)

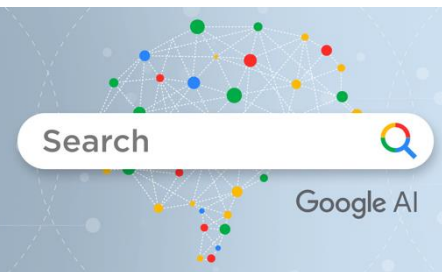
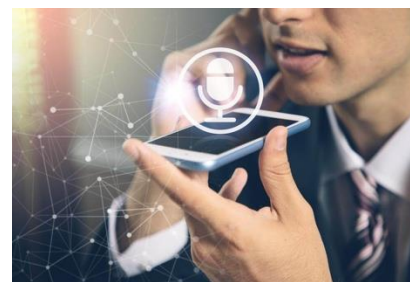
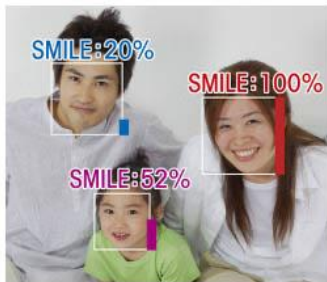


- 機械学習の研究者、かつ、機械学習のユーザ
- 四国うどん県生まれ
- なのに気づいたらなぜか北海道と京都を2往復

札幌 (10年)	北大 工学研究科	機械学習 × 信号処理(音源分離)
京都 (7年)	京大 化学研究所 / 薬学研究科	バイオインフォマティクス・創薬化学
札幌 (7年)	北大 情報科学研究科	機械学習 × 計算機科学(離散構造)
	JST さきがけ	材料インフォマティクス
京都 (6年)	理研 AIP 京大 iPS細胞研 連携G	機械学習 × 幹細胞生物学
	北大 化学 反応創成研究拠点	機械学習 × 化学
	京大 国際高等教育院	機械学習 × データサイエンス・統計学

# 機械学習は日常ツールとして身近に

スマホ、広告、検索、ネットショッピング、SNS、ゲーム、画像認識、音声認識、  
画像生成、音声生成、翻訳、要約、音声合成、自動運転、セキュリティ、…



# 画像・音声・言語の次なるフロンティア “AI for Science”

*Nature* **620**, 47–60 (2023)

*Science is changing,  
the tools of science are changing.  
And that requires **different** approaches.  
(Erich Bloch, 1925-2016)*

**Review**

---

**Scientific discovery in the age of artificial intelligence**

- データの変容により「**科学の方法**」の**再考**が迫られている
- 大手ビッグテックはどこもAI4Sの研究所・部署を作り大規模な投資も
- **今もっとも胸熱で楽しいタイミング！**  
チューリング賞、ノーベル賞、スターゲート計画、…
- 良い研究が出れば、皆さんも**いきなり世界に名が轟く有名人になれる**

# 技術屋から見てきた生命科学

分子レベル、細胞レベル、個体レベルの対象現象とデータを研究

- 代謝反応系の遺伝子発現制御

Bioinformatics 2007, 2008, 2009, 2010, Nucleic Acids Res 2011, PLoS One 2012, 2013, KDD'07, ISMB'07, Pattern Recognition 2011, In Silico Models for Drug Discovery 2013, Data Mining for Systems Biology 2013, Data Mining for Systems Biology 2018

- 分子ネットワーク・分子構造の機械学習

IEEE TPAMI 2017, Machine Learning 2011, BMC Bioinformatics 2020, IEEE BigData'19 WS, KDD '18 WS, AAAI'22 WS

- 薬剤-標的の相互作用ネットワーク(多重薬理)解析

PLoS One 2011, Drug Discov Today 2013, Brief Bioinform 2014

- ゲノム中の反復配列の解析

Discrete Appl Math 2013, 2016, AAAI 2020

- 糖鎖構造のパターン解析

ECCB 2008 / Bioinformatics 2008

- モジュレートプロテアーゼの基質解析・DB構築

Mol Cell Proteom 2016, Genome Informatics 2009

- メディエーター複合体による転写伸長制御

Nat Commun 2015, Nat Commun 2020

→ 生データからのChIP-Seq解析を担当

- がん細胞と正常細胞の細胞競合

Cell Reports 2018, Cell Reports 2022, Sci Rep 2015

- 翻訳停止時のリボソーム挙動の解析

Journal of Biological Chemistry 2014

- 多系統萎縮症・多発性硬化症のコピー数多型解析

Mol Brain 2017

→ 生データからのCNV解析を担当

- 乳がんの体細胞突然変異のパターン解析

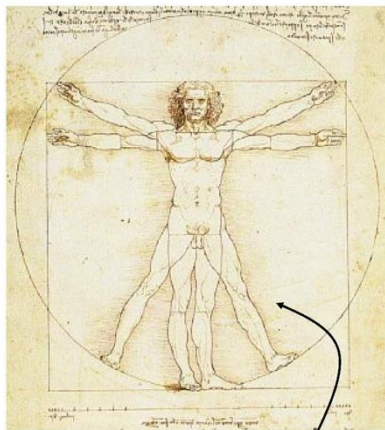
Brief Bioinform 2016, Big Data Analytics in Genomics 2016

- 歩行動画からの統一パーキンソン病評価尺度 (UPDRS)スコア予測

BMC Neurology 2023

# 関心：化学としての生命、情報としての生命

## The Human Molecule



C<sub>E27</sub>H<sub>E27</sub>O<sub>E27</sub>N<sub>E26</sub>P<sub>E25</sub>S<sub>E24</sub>Ca<sub>E25</sub>K<sub>E24</sub>Cl<sub>E24</sub>Na<sub>E24</sub>Mg<sub>E24</sub>Fe<sub>E23</sub>Fe<sub>E23</sub>  
Zn<sub>E22</sub>Si<sub>E22</sub>Cu<sub>E21</sub>Be<sub>E21</sub>I<sub>E20</sub>Sn<sub>E20</sub>Mn<sub>E20</sub>Se<sub>E20</sub>Cr<sub>E20</sub>Ni<sub>E20</sub>Mo<sub>E19</sub>Co<sub>E19</sub>Ve<sub>E18</sub>

Libb Thims

- 生物と無生物の違いは材料にはない
- 物質代謝・エネルギー代謝  
→ 絶え間なく変化し続ける(非平衡)
- 化学反応は細胞の命のあらわれだ (パスツール)
- 常温環境下なのに、何千もの異なる化学反応が全体として高度に連動し絶えず進行している
- 「ウェットウェア」→ 生きていないハードウェアシステムとは異なる柔軟性・適応性を伴う
- エントロピー増大則に抗って高度な秩序を保つ



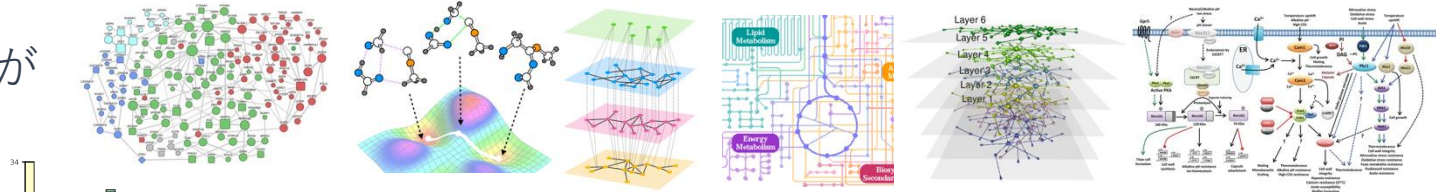
# 離散構造(情報の組合せ)の統計科学/機械学 習

離散構造 = 有限個の離散要素の組合せによって生じる情報  
→ 集合、論理、系列、ツリー、グラフ、プログラム、言語など

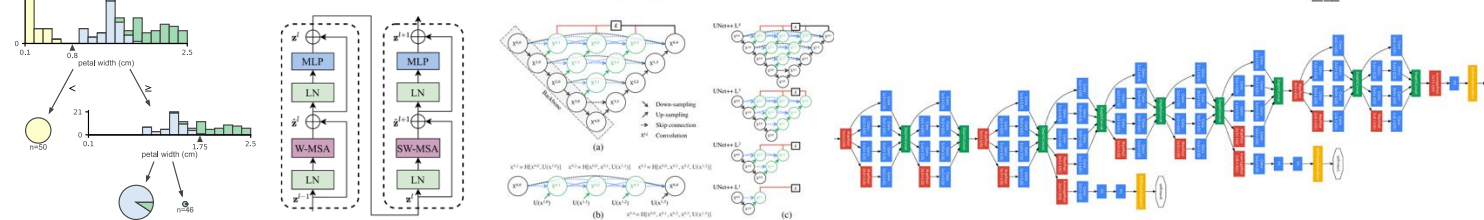
対象が  
離散構造



対象間の関係が  
離散構造



モデルが  
離散構造





# 事例：めちゃ強力な海中接着剤できた

1. 海洋生物やバクテリアなどの接着タンパク質の配列を網羅的に解析
2. 同定した接着力を発揮する配列雛形を模倣して180個のポリマー分子を合成
3. さらに機械学習モデルベースの最適化で接着力を最大化して超強力に！



# 卒研やラボローテのプロジェクトについて

- 瀧川は機械学習分野の運営に関わってきた（最新動向・実態に詳しい）
  - [NeurIPS](#) 2024 Area Chair, 2023 Top Reviewer (Top10%), 2022 Top Reviewer (Top10%), 2019 Best Reviewer, 2018 Top Reviewer (Top30%)
  - [ICML](#) 2022 Outstanding Reviewer (Top10%), Expert Reviewer 2021, Top Reviewer (Top33%);
  - [ICLR](#) 2022 Highlighted Reviewer
  - [IJCAI](#) 2022 Distinguished PC (Top 3%); 2019 Active Reviewer
- 会議では生命科学や分子科学に関わる論文も多数発表される
- 最近は[Benchmark and Dataset track](#)もでき、手がつけやすいベンチマークタスク・データセットがあり入り口としてとても良い  
→ 関連する論文を読み、SOTAに勝つ手法を考えれば研究になる

# 例：メタボロミクス×機械学習

- 代謝は生きているものすべての基礎
- 代謝の根幹は光合成  $6\text{CO}_2 + 12\text{H}_2\text{O} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{H}_2\text{O} + 6\text{O}_2$   
→ 人工光合成はDream Reactionの一つ
- 動物(従属栄養生物)は異化代謝、植物(独立栄養生物)では同化代謝
- 特に植物では代謝産物の総体(メタボローム)が細胞状態のスナップショットや植物の有効利用価値として重要
- 質量分析でメタボロームを同定するタスクは非常に難しいタスク  
→ 「これはAIの出番やろ」(ブレイクスルーが期待されている)
- 質量分析インフォマティクス  
→ 瀧川はコミュニティとゆるーい関係がある(プロテオ多め?)

# 事例：メタボロミクス × 機械学習

- MassSpecGym (NeurIPS 2024)  
<https://neurips.cc/virtual/2024/poster/97823>
- Tomáš Pluskal, Wout Bittremieux と湘南会議を2026年に計画中

---

## MassSpecGym: A benchmark for the discovery and identification of molecules

---

Roman Bushuiev<sup>1,2</sup>, Anton Bushuiev<sup>2</sup>, Niek F. de Jonge<sup>3</sup>, Adamo Young<sup>4</sup>,  
Fleming Kretschmer<sup>5</sup>, Raman Samusevich<sup>1,2</sup>, Janne Heirman<sup>6</sup>, Fei Wang<sup>7,8</sup>,  
Luke Zhang<sup>9</sup>, Kai Dührkop<sup>5</sup>, Marcus Ludwig<sup>10</sup>, Nils A. Haupt<sup>5</sup>, Apurva Kalia<sup>11</sup>,  
Corinna Brungs<sup>1</sup>, Robin Schmid<sup>1</sup>, Russell Greiner<sup>7,8</sup>, Bo Wang<sup>4</sup>, David S. Wishart<sup>7,12</sup>,  
Li-Ping Liu<sup>11</sup>, Juho Rousu<sup>13</sup>, Wout Bittremieux<sup>6</sup>, Hannes Rost<sup>9</sup>, Tytus D. Mak<sup>14</sup>,  
Soha Hassoun<sup>11,15</sup>, Florian Huber<sup>16</sup>, Justin J.J. van der Hooft<sup>3,17</sup>, Michael A. Stravs<sup>18</sup>,  
Sebastian Böcker<sup>5</sup>, Josef Sivic<sup>2</sup>, Tomáš Pluskal<sup>1</sup>

# テーマ：生命科学を試験台に情報科学を研究

- 化学理論や幾何制約などの厳密制約を満たす離散機械学習手法の研究
- 複数分子の相互作用や動的状態遷移のための表現学習手法の研究  
(特に遷移状態やキラリティに関わる問題)
- 機械発見：機械学習に基づく訓練分布外への探索法、UQ法の研究
- 量子化学計算など計算化学データからの機械学習の研究
- 幾何グラフに対する群同変Graph Transformerの設計と改良
- 分子表現に対する基盤モデル/世界モデル開発
- StateやMemoryを伴うポストTransformerの設計と高効率化  
(Linear Transformer, SSM/Mamba, xLSTM, RWKV, Titans, etc..)

# 科学の方法の科学：科学的理解とは何か？

- AI for Scienceは「科学の方法」の科学  
→ 必然的にメタサイエンスの側面を伴う
- 「生命の仕組みを理解したい」と言うが「理解」とは何を指す？
- 現行のLLMはかなり自然に言語を扱え、部品の詳細も組み上げ方も100%明らかなだが(人工物なので)、それは「私たちがなぜ言語を扱えるのか」の理解にはほとんど寄与していない…
- 現象を再現するシステムを「もし作れたら理解できる」という構成論的アプローチに懐疑を突きつける結果
- これは「機械学習」と言う技術の特殊性に深く関係

1. 瀧川一学, なぜ経験則は説明の論理として受け入れがたいか. 岩波「科学」2023年12月.
2. 瀧川一学・折田奈甫, 大規模言語モデルは人間の言語能力の解明に役立つのか？[後編] 岩波「科学」2024年8月.



# Takeaways

たきがわ いちがく

瀧川 一学

<https://itakigawa.github.io/>

このスライドのPDF

<https://itakigawa.github.io/data/lab2025.pdf>



- 4/1着任予定
- 人もまだいない🙄  
→ 瀧川が直接丁寧に指導
- @柏キャンパス
- 関心：「機械学習・機械発見」by 離散組合せ情報の統計科学  
→ “AI for Science” 科学の方法としての「AI」は今めっちゃ楽しい
- 「情報科学を使って生命科学を切り拓く」のではなく  
「生命科学のための情報科学を切り拓く」
- われわれはデータに溺れながら、知識を渴望している（ブレナー）