

TS-02 日時：2021 年 10 月 25 日 13:00-17:00

半日で知る、化学分野のデータサイエンスおよび人工知能概要：

「FS-08:化学データサイエンスおよび人工知能討論、勉強会」立ち上げ会協賛

2021 年 10 月 28 日（木） 13:00-14:30

A Half-day Overview of Data Science and Artificial Intelligence in the Field of Chemistry:

"FS-08: Chemical Data Science and Artificial Intelligence Discussion, Work Shop"

Kick-off Meeting Sponsorship

モデレーター：湯田 浩太郎 Kohtaro Yuta

株式会社インシリコデータ（In Silico Data, Ltd.）

<https://cbi-society.org/taikai/taikai21/TS/TS-02.pdf>

<フォーカストセッション> 2021 年 10 月 28 日 (木) 13:00-14:30

FS-08 「化学データサイエンスおよび人工知能討論、勉強会」立ち上げ会：
計算毒性学研究会主催

モデレーター：

湯田 浩太郎 (株式会社インシリコデータ)

植沢 芳広 (明治薬科大学)

金子 弘昌 (明治大学 理工学部)

講師：

金子 弘昌 (明治大学 理工学部)

「人工知能を活用した分子設計・材料設計」

結城 伸哉 (株式会社 Elix)

「はじめてのAI創薬とElixにおける事例紹介」

古田 一匡 (富士通株式会社)

「富士通が考えるニューノーマル時代の研究スタイル」

大内山 浩 (インテル株式会社)

「インテルのAIに向けての取り組み」

化学研究実施のための三大基本技術

データサイエンス

コンピュータが生まれる前からデータ整理や解析技術として展開されていた。

シミュレーション

MDの最初の文献は1957年である。現在のドッキングはコンピュータグラフィック技術の展開およびタンパク質の座標データの存在が必須である。これらの条件がクリアされたことで、急速に普及した。

AI（人工知能）

世界最初の人工知能システム
1965年にDENDRALが発表
エドワード・ファイゲンバウム
(Edward Albert Feigenbaum)
人工知能の父と呼ばれる。
最初の人工知能はルールベース型であり、現在のニューラルネットワーク型とは異なる。

今日のチュートリアルテーマ

◇本日の講義スケジュール

序章 13:00～13:20

- ・自己紹介
- ・時代の流れと、研究環境および研究内容や研究スタイルの変化

第一章 13:20～13:40 (10分休憩)

化学データサイエンスおよび人工知能の適用イメージ

第二章 13:50～15:00 (10分休憩)

化学データサイエンスおよび人工知能の適用事例：様々なアプローチが可能

第三章 15:10～15:50 (10分休憩)

化学データサイエンスおよび人工知能の適用手順やパターン

第四章 16:00～16:50

化学データサイエンスおよび人工知能の適用上での留意事項

- ①化学分野の留意点
- ②データサイエンス実施上での留意点

まとめと提案 16:50～17:00

- ①化学データサイエンスおよび人工知能のまとめ
- ②今後の展開についての「**オートノマス創薬**」の提案
 - ・KY法の展開（クラス分類と重回帰型）
- ③自由討論

◆なんとなく取り付きにくい

- ・「データサイエンス基本」と言っても、本の中には**馴染みのない数式が多く理解しにくい**

◆どのように使うのかイメージがわきにくい

- ・汎用的な学習書で取り上げられる適用事例は化学とはかけ離れており、**自分の研究に適用する時のイメージがわからない**

◆討論内容がわからない

- ・本分野の研究発表では専門用語や、前提知識等がないと理解できないことが多い
データサイエンスおよびAI関連の本では、基本や基礎編といっても**ある程度の基本用語が理解できる前提で書かれている**

◆化学とデータサイエンスやAI

- ・化学分野のアナログデータをコンピューターのデジタルデータとの関係が良く分からない
日常の研究では化合物構造式や安全性に関する内容でコミュニケーションするが、**データサイエンスやAIの研究者との討論が出来ないし、討論がかみ合わない**

◆参考及び理解できる資料が少ない

- ・データサイエンスやAIの基本を説明する資料は多いが、**化学分野での適用に関するものは少ない**。学会での研究発表事例は特定の分野で、発表は基本を理解できないと把握困難。

●データサイエンスやAIには興味がある

- ・実験データ等の整理や精度向上、更なる方向性の探索、戦略策定等に使える手法と考える

●研究分野や内容の変化

- ・自分の研究分野自体にコンピュータの導入が必要となっている。

- ①実験データの増大とその処理
- ②研究自体に他の研究分野との融合や連携が必要となりつつある
- ③実験結果の整理のみならず、実験の精度向上や改良、新たな研究の方向性探索に使いたい

●自分の研究内容と時代の変化

- ・会社や研究所での自分の研究がこのまま5年10年20年と続けられるとは考えにくい

●化学データサイエンスとAIの、時を超えた汎用性

- ・化学データサイエンスとAIは殆どの研究分野の基本アプローチである。従って、時代の変化に従った研究内容や研究分野の変化に耐えるロバスト性が高い

化学データサイエンス及びAIはなぜ取り付きにくいのか

環境は大きく変化して、実施環境が改善している

×化学データサイエンスやAIの必要性が少ない

- ・現場の研究者にかなわないレベル（当初の状態）：高いシステム購入して、効果はこんなものか
- ・無理に研究に変化を加えるほどの必要性は無い
- ・新しい研究スタイルに変えて、失敗すると後が怖い
- ・新しいことをやる時に、周りにアドバイザーがいなく、上司からの理解が得られにくい
- ・分野の異なる研究者とチームを組んでも、相手との相互コミュニケーションが困難である

◎時代の変化による化学データサイエンスやAIの実施必要性の増大と実施の壁が低くなった

・コンピュータ関連環境の急激な進歩

- ①PCが普及し、値段の高いコンピュータがなくなかった。気軽に実施可能。
- ②CPUが早くなり、従来は時間がかかった計算が急速に早くなった。
- ③データストア機能が向上し大量のデータを扱えるようになった。
- ④ネットワークの進歩により、ビッグデータレベルの情報を集めやすくなった。

・関連環境の変化

- ①様々な機器が進歩して、情報量が急速に拡大し、紙ベースの媒体から電子媒体に変化した。
- ②研究内容が、単一分野から境界領域や複合領域研究の必要性が高まっている。
- ③化学データサイエンスやAI研究の実施支援システムが多数開発されてきた。

TS-02 序章に続く