

## Symposium | Co-Innovation Program (CIP) : Data-Driven Chemical Synthesis

📅 Mon. Mar 18, 2024 1:00 PM - 3:30 PM JST | Mon. Mar 18, 2024 4:00 AM - 6:30 AM UTC 🏢  
A1432(1432, Bldg. 14 [3F])

**[A1432-1pm] Data-Driven Chemical Synthesis**

Chair, Symposium organizer: Hikaru Takaya, hironobu iyama, Kazuhiko Sato, Kazuhiro Matsumoto

AI/デジタル技術を駆使した化学合成では、用いる/得られる様々な化学的・物理的情報を基盤としたデータ駆動型プロセス設計が必須となりますが、データ処理と理論構築は現場の実験者や情報科学を専門としない研究者が最も苦手とするところではないでしょうか。本セッションでは、実験台の前に立つ研究者が自らの手でデータ駆動型合成研究を推進するために必要となるインフォマティクスの基礎とHow Toについて知ってもらうために、研究の現場でデータサイエンスツールを使いこなし、活かしている研究者から「合成実験に活かすための化学データの取扱い」に関する様々なノウハウをご解説を頂きます。

聴講後の[アンケート](#)へのご協力をお願いいたします。

## 🎧 Japanese 🎧 Keynote Lecture

1:00 PM - 2:00 PM JST | 4:00 AM - 5:00 AM UTC

[A1432-1pm-01]

Handling of chemical data for use in experiments

○Hiromasa Kaneko<sup>1</sup> (1. Meiji University)

2:00 PM - 2:10 PM JST | 5:00 AM - 5:10 AM UTC

[1A143204-06-2add]

Incubation Time

## 🎧 Japanese 🎧 Invited Lecture

2:10 PM - 2:40 PM JST | 5:10 AM - 5:40 AM UTC

[A1432-1pm-02]

Development of a machine learning model that enables "extrapolative" search and its application to the search for new catalysts effective for CO<sub>2</sub> reduction reactions.

○Shinya Mine<sup>1</sup> (1. National Institute of Advanced Industrial Science and Technology (AIST))

2:40 PM - 2:50 PM JST | 5:40 AM - 5:50 AM UTC

[1A143204-06-4add]

Incubation Time

## 🎧 Japanese 🎧 Invited Lecture

2:50 PM - 3:20 PM JST | 5:50 AM - 6:20 AM UTC

[A1432-1pm-03]

Application of Chemoinformatics and Quantum Chemistry for Reaction Process Design

○Toru Yamaguchi<sup>1</sup>, Yuujirou Matsuo<sup>1</sup>, Kenji Hori<sup>1,2,3</sup>, Kimito Funatsu<sup>4</sup> (1. TS Tech. Co. Ltd., 2. AIST, 3. Yamaguchi Univ., 4. NAIST)

3:20 PM - 3:30 PM JST | 6:20 AM - 6:30 AM UTC

[1A143204-06-6add]

Incubation Time

## 実験に活かすための化学データの取扱い

(明大理工) ○金子 弘昌

Handling of Chemical Data for Use in Experiments (*School of Science and Technology, Meiji University*) ○Hiromasa Kaneko

By using chemical data, a mathematical model  $y=f(x)$  can be constructed between features  $x$  such as synthesis and process conditions including raw materials and reaction conditions, and objective variables  $y$  such as properties and activities of molecules and materials. Then, we can predict new experiments, propose efficient ways to proceed with experiments, and design  $x$  values for which  $y$  has target values. In addition to the research related to chemical data and models, networking among researchers and social awareness activities will be presented.

*Keywords : Chemical Data, Machine Learning Model, Experimental Design, Collaborative Research, Online Salon*

高機能性材料を研究・開発・製造する際、化学データや化学工学データを活用して機械学習を行い、分子設計・材料設計・プロセス設計・プロセス管理を効率化することが一般的となっている。分子設計では、化合物の分子構造を数値化した分子記述子  $x$  と活性・物性・特性  $y$  との間で数理モデル  $y = f(x)$  を構築する。構築されたモデルに基づいて、新たな分子構造に対する活性・物性の値を予測したり、活性・物性が所望の値となる分子構造を設計したりする。材料設計では、材料の物性・活性・特性  $y$  と材料の合成条件や製造条件  $x$  との間で数理モデル  $y = f(x)$  を構築する。モデルに基づいて、実験前もしくは製造前に材料の合成条件や製造条件から材料の物性・活性・特性を予測したり、目標の材料を達成するための合成条件や製造条件を設計したりする。プロセス設計では、目標の性能をもつ装置やプラントを設計するため、実験もしくは各種シミュレーションを行う際のプロセス条件  $x$  と装置・プラントの性能  $y$  との間で数理モデル  $y = f(x)$  を構築することで、装置やプラントで合成・製造する際の目標を達成するためのプロセス条件を効率的に設計できる。

実験結果である物性・活性・特性などの  $y$  と実験条件  $x$  との間で数理モデル  $y = f(x)$  を構築する際、適切に  $x$  を準備することが重要である。例えば化合物においては分子構造を扱うため、分子構造から  $x$  を計算する必要がある。また組成や温度のように数値で表される実験条件であっても、それらを適切に変換して  $x$  とすることで、モデルの予測精度が向上したり、既存のデータにない新たな材料 ( $x$  の外挿領域) を予測したりすることが可能となる。 $x$  を準備した後も、予測精度の高いモデルを構築するためには、少数の材料サンプルを駆使して、多数のモデル構築手法の中から適した手法を選択する必要がある。

以上を踏まえ今回は、実験とデータ解析・機械学習を組み合わせる際の化学データや数理モデルの扱いについて講演する。さらに、研究者のネットワーク形成や社会的な啓蒙活動についても紹介する。

## 「外挿的」探索を可能にする機械学習モデルの開発と CO<sub>2</sub> 還元反応に有効な新規触媒探索への展開

(産総研) ○峯 真也

Development of a machine learning model that enables "extrapolative" search and its application to the search for new catalysts effective for CO<sub>2</sub> reduction reactions (*Research Institute for Chemical Process Technology, National Institute of Advanced Industrial Science and Technology (AIST)*) ○Shinya Mine

Data-driven materials development, which uses machine learning (ML) and other artificial intelligence (AI) techniques to accelerate the development of materials and catalysts, is attracting attention. Successful examples of "optimization research" with defined rules and goals are already emerging, including significant reductions in development time and cost. On the other hand, the development of truly innovative materials and catalysts has yet to be achieved. This is because conventional AI technology cannot suggest information that is not contained in the data. In addition, the significant time and cost involved in collecting high-quality experimental data for learning also hinders the spread of data-driven research. In other words, there is a need to develop new ML approaches that can explore "extrapolated regions" of the training dataset and that can be used even with relatively small datasets.

In this talk, we will present the ML model we developed that enables "extrapolative" search. Furthermore, we will present a case study in which we found a multi-element catalyst effective for the reverse aqueous gas shift reaction to synthesize CO from CO<sub>2</sub>/H<sub>2</sub> by using the developed ML model and running a closed loop of experiment (evaluation) ↔ machine learning starting from a small set of training data.

*Keywords : Machine Learning; Heterogeneous Catalysts; CO<sub>2</sub> Reduction*

機械学習をはじめとする人工知能(AI)の技術を用いて材料・触媒開発を加速化する、「データ駆動型」材料開発が注目を集めており、ルールとゴールが明確な「最適化研究」においては、開発にかかる時間とコストを大幅に削減する等の成功例がすでに出始めている。一方、真に革新的な材料・触媒の開発は未だ達成されていない。革新的な材料・触媒の発見は、今までのデータ中に無い、もしくは極めて少ない元素の活用等から生まれるものであるが、従来型の AI 技術ではデータの中に含まれていない情報は提案できないためである。また、学習に用いる良質な実験データを収集するために多大な時間的・金銭的コストがかかることも、データ駆動型研究の普及を妨げている。すなわち、学習データの「外挿領域」まで探索が可能で、かつ、比較的小さなデータセットでも利用できる新しい機械学習手法の開発が求められている。

本講演では、我々が開発した「外挿的」探索を可能にする機械学習モデルについて紹介する。さらに、開発した機械学習モデルを用い、少数の学習データから出発して実験(評価)↔機械学習のクローズドループを回すことで、CO<sub>2</sub>/H<sub>2</sub> から CO を合成する逆水性ガスシフト反応に有効な多元素触媒を見出した研究事例についても紹介する。

## 反応プロセス設計に向けたケモインフォマティクスと量子化学の応用

(TS テクノロジー<sup>1</sup>・産総研<sup>2</sup>・山口大<sup>3</sup>・奈良先端大<sup>4</sup>) ○山口 徹<sup>1</sup>・松尾 勇二郎<sup>1</sup>・堀 憲次<sup>1,2,3</sup>・船津 公人<sup>4</sup>

Application of Chemoinformatics and Quantum Chemistry for Reaction Process Design (<sup>1</sup>*Transition State Technology Co. Ltd.*, <sup>2</sup>*National Institute of Advanced Industrial Science and Technology*, <sup>3</sup>*Graduate School of Sciences and Technology for Innovation, Yamaguchi University*, <sup>4</sup>*Data Science Center, Nara Institute of Science and Technology*) ○Toru Yamaguchi,<sup>1</sup> Yuujiro Matsuo,<sup>1</sup> Kenji Hori,<sup>1,2,3</sup> Kimito Funatsu<sup>4</sup>

In recent years, with the development of AI technology, computer chemistry, represented by chemoinformatics and quantum chemical calculations, has been attempted to develop new compounds and their synthetic routes. Although there is still a gap between computed virtual results and practical design and operation of production equipment, we aim to resolve this gap by combining computer chemistry focusing on the TS motif method<sup>1)</sup> with minimal experiments utilizing flow and droplet methods. We call this approach “Digital Driven Chemistry”. In “Digital Driven Chemistry”, computer chemistry plays (1) synthesis routes designing, (2) screening by quantum chemical calculations, and (3) accurate rate simulations based on reaction kinetics<sup>2)</sup>. In this lecture, we will introduce examples of utilities of the TS motif method with Transition State Database and self-organizing maps, and of reaction kinetics simulations with quantum chemical calculations in relation to (2) and (3) respectively.

The part of this research is based on results obtained from a project, JPNP19004, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

**Keywords :** *Digital Driven Chemistry; Transition State Database; TS Motif Method; Self-organization Map; Kinetics Simulation*

近年、AI 技術の発達に伴い、ケモインフォマティクスや量子化学計算に代表される計算機化学を用いて、新規化合物やその合成経路の開発が試みられるようになってきた。コンピュータが推測する仮想的な結果と、現実で用いられる生産装置の設計や運用との間には依然隔たりがあるが、我々は、TS モチーフ法<sup>1)</sup>を中心とした計算機化学と、フロー法やドロップレット法などを活用した最小限度の実験とを組み合わせることにより、これを解消することを目指しており、この手法を「デジタル駆動化学」と呼んでいる。「デジタル駆動化学」における計算機化学の役割には、(1)合成経路の創出、(2)量子化学計算によるスクリーニング、(3)反応速度論による精密な速度シミュレーション<sup>2)</sup>、などが挙げられる。今回は、特に (2)、(3)に関連し、それぞれ、遷移状態データベースと自己組織化マップによる TS モチーフ法の活用例と、量子化学計算を用いた反応速度論シミュレーション例について紹介する。

1) K. Hori, T. Yamaguchi, *Chemistry and chemical industry*, **2019**, 70, 12, 858.

2) T. Yamaguchi, H. Yamada, K. Hori, *Kagaku Kougaku Ronbunshu*, **2017**, 43, 2, 111