

戦略的創造研究推進事業 CREST
研究領域「多様な天然炭素資源の活用に資する
革新的触媒と創出技術」
研究課題「実験・計算・データ科学の統合によるメタン
変換触媒の探索・発見と反応機構の解明・制御」

研究終了報告書

研究期間 2017年10月～2023年3月

研究代表者：高橋啓介
（北海道大学
理学研究院化学部門、教授）

§1 研究実施の概要

(1) 実施概要

実験・計算・データ科学を統合した**触媒インフォマティクスを推進し、「メタン酸化カップリング反応」と「メタノール合成」触媒の設計・開発を進める。**研究は「独自触媒データベースの構築」「インフォマティクスによるデータから触媒設計」「触媒インフォマティクス統合プラットフォーム」の3本柱を中心に進めている。研究グループは**触媒インフォマティクスとプラットフォーム(高橋 G)、ハイスループット実験(谷池 G)、実験実証(大山 G・西村 G)、データマイニング(宇野 G)の5つのチーム**からなり、データ構築、データ科学による知識の抽出、実験実証まで一貫通貫する体制となっている。

・データベース

OCMハイスループット実験装置(谷池 G)の開発に成功し、約11万件超のOCM触媒データの取得が完了した。メタノール・ハイスループット実験装置の開発(谷池 G)にも成功した。また独自開発したハイスループット計算(高橋 G)においても、約1万件の計算OCMデータ取得が完了している。大山 Gと西村 GではXPS,XAFS,XRD等によるOCM触媒の計測データも習得した。また文献においても、約2000のOCMデータを扱っている。

・データからの触媒設計

インフォマティクスからの触媒設計では、**ディープラーニング(高橋 G)とハイスループット実験データ(谷池 G)の連携により20%近くの収率を出す文献未報告のOCM触媒「TiKW-SiO₂」の発見や、新規触媒としてMo-Cs-W/BaO(C₂選択率83.1%・収率16.2%)が発見された。**また独自開発した**アイテムセットマイニング(宇野 G)を使い、未報告低温OCM触媒の開発(大山 G)に成功し、500°CでOCM活性を示す未報告の特異な触媒「Al-Y₂O₃」と「Ag-Y₂O₃」と500°Cで収率18%のSr/La₂O₃が発見された。**また機械学習とハイスループット実験データを組み合わせることにより、**450°Cで活性なOCM触媒Ga-Sr-Yb-La₂O₃を発見(西村 G 高橋 G)し、特許を出願した。**またハイスループット実験データを活用しC₂選択率を決定する触媒記述子の設計に成功(高橋 G)し、未発見OCM触媒La-Pr-Hf-BaO(C₂収率16%、C₂選択率84%)の実験実証(谷池 G)にも成功した。

触媒インフォマティクスの方法論の開発も進め、**触媒ネットワーク、触媒遺伝子、計算データからの触媒設計、ベイズによる触媒最適化**も提案した。ネットワークでは、オントロジーを導入しOCM触媒ビッグデータにおけるデータ内の関係性をネットワークとして記述することに成功(高橋 G)し、機械学習では達成不可能であった理解を伴う触媒設計を実現し、K-V-Eu-BaO等のC₂収率20%を超えるOCM触媒の実験実証(谷池 G)に成功した。またOCM触媒ビッグデータにおいて、組成が異なる触媒が類似した触媒活性・反応条件応答性を示すことに着目し、信号解析・パターン認識記述を使い、触媒に特有の記号を与えた触媒遺伝子を提案し、触媒の新たな記述方法を提案した。ハイスループット計算データとインフォマティクスを組み合わせた触媒設計を提案し、発見された高活性触媒Co-Ag-TiO₂等は特許出願した。XPSデータと触媒活性を機械学習によってつなげる計測インフォマティクスも提案(高橋 G 大山 G)し、Ba-Hf-Sm-Yb₂O₃(C₂収率17%)も実験実証に成功(大山 G)し、計測データの活用方法を見出した。ベイズ最適化(高橋 G)と銅ゼオライトの調製方法(大山 G)を組み合わせ、少ないデータからの触媒最適化も行い、高いメタノール選択率とターンオーバー数を持つ銅ゼオライト触媒の実験実証にも成功した。

・プラットフォーム

触媒プラットフォーム「CADS」の開発(高橋 G)にも成功し、簡単な操作で触媒インフォマティクスを実践できるプラットフォームを運用している。機械学習、データ可視化、画像解析だけでなく、**世界初の実験触媒データセンター**としての機能も兼ねている。利用者は国内外合わせて400ユーザーを超えている。CADSは企業との連携を進め、2つの特許を取得し、社会実装を進めている。

(2) 顕著な成果

<優れた基礎研究としての成果>

1. 機械学習による触媒設計方法論の確立

概要:機械学習を用いた触媒開発方法論を確立した。ハイスループット実験データと機械学習を組み合わせることにより、450°Cで活性な OCM 触媒 Ga-Sr-Yb-La₂O₃を発見し特許出願をした。また新規触媒として Mo-Cs-W/BaOも発見され、C₂選択率 83.1%・収率 16.2%を実現した。触媒科学において世界で初めてディープラーニングを適用し 20%近くの収率を出す文献未報告の OCM 触媒「TiKW-SiO₂」の発見も達成した。少ないデータに対して機械学習を適用し、高いメタノール選択率とターンオーバー数を持つ銅ゼオライト触媒の実験実証にも成功した。このように機械学習による触媒設計方法論を世界に先駆けて実践してきた。

2. 触媒の新たな記述方法の確立

概要:OCM 触媒ビッグデータにおいて、組成の異なる触媒が類似した触媒活性・反応条件応答性を示すことが発見された。そこで信号解析・パターン認識を使い、触媒の活性や反応条件応答性を示す特有の記号を与えた触媒遺伝子を提案し、従来の元素記号によるものとは異なる触媒の新たな記述方法を確立した。この触媒遺伝子と自然言語処理を用いることにより、同等の触媒性能を持つ触媒を容易に特定できるだけでなく、同じ遺伝子内の元素を組み替えることにより同性能の触媒を設計することも可能になる。またバイオインフォマティクスで使用される解析手法の適用も可能になり、触媒開発の新しい方法論を提案した。

3. ハイスループット計算と触媒ネットワークの融合

概要:計算化学からの触媒設計は困難を極める。理由として計算化学では局所的な触媒活性を見ているが実験では複雑な触媒活性の平均情報が結果として得られているという仮説が考えられる。そこで独自開発したハイスループット計算を実行し、約 1 万の触媒表面でメタン酸化反応を計算しデータ化した。この計算データに対して開発した触媒ネットワーク方法とデータ可視化を組み合わせることにより平均的な触媒情報に変換することに成功した。発見された高活性触媒 Co-Ag-TiO₂等は特許出願し、成果は Journal of the American Chemical Society で発表した。

<科学技術イノベーションに大きく寄与する成果>

1. 統合データプラットフォーム「CADS: Catalysts Acquisition by Data Science」の開発

概要:触媒インフォマティクスのための統合データプラットフォーム「CADS: Catalysts Acquisition by Data Science」の開発に成功し、2019 年度より本格運用を開始している。URL は(<https://cads.eng.hokudai.ac.jp/>)である。現在まで登録アカウント数 400、平均月 8000 程度のアクセス数となっており、国内外の大学・企業など幅広い方々に利用されている。CADS の機能は「触媒データの公開・管理」「データの可視化」「機械学習による学習・予測」「画像解析」が実装されており、ウェブ上でマウスのクリック 1 つでインフォマティクスの一連の解析を行うことが可能となっている。このような触媒データセンターはこれまでになく、世界中からアクセスがあり、企業から実装したいという問い合わせが多数ある。CADS に関する論文は RSC の Reaction Chemistry & Engineering で発表し、プレスリリースも行っている。

2. OCM ハイスループット実験装置の開発と OCM 触媒データの活用

概要:1 日に 4000 点もの触媒データを自動取得可能な OCM ハイスループット実験装置の開発に成功し、過去 30 年で蓄積されたデータ数を一桁上回る 12000 点ものデータをわずか 3 日で取得することに成功した。この規模の触媒データはこれまでに存在せず、世界初である。得られた触媒ビッグデータを機械学習などによって分析し、実験条件による C₂ 収率の変化や鍵となる触媒組成を突き止めることに成功した。CREST では 11 万を超える OCM 触媒データを取得し、世界最大の触媒データとなっている。成果は ACS Catalysis などに掲載され、プレス発表も行った。

3. 触媒ビッグデータのネットワーク化「触媒世界地図」

概要: 機械学習では実現し難い「理解を伴う触媒インフォマティクス」を推進するため、オントロジーを用いた触媒ビッグデータのネットワーク化を提唱した。ハイスループット実験データにおいて、触媒組成、実験条件、 C_2 収率の関係性を知識のネットワークとして可視化することにより、元素組み合わせや反応条件と C_2 収率の関係性が提示され、理解を伴う触媒設計を可能にした。このネットワークを「触媒世界地図」と名付け、K-V-Eu-BaO 等の C_2 収率 20% を超える触媒の実験実証にも成功した。成果は RSC の *Chemical Science* に採択され、“pick of the week”にも選出された。本方法論は触媒開発の新たな方法論を提案するものであり、多くの触媒データに適用可能な技術である。

<代表的な論文>

1. Thanh Nhat Nguyen, Thuy Tran Phuong Nhat, Ken Takimoto, Ashutosh Thakur, Shun Nishimura, Junya Ohshima, Itsuki Miyazato, Lauren Takahashi, Jun Fujima, Keisuke Takahashi, Toshiaki Taniike, “High-Throughput Experimentation and Catalyst Informatics for Oxidative Coupling of Methane” *ACS Catalysis* (2020) 10(2), pp 921–932.

概要: 1日に4000点もの触媒データを自動取得可能な OCM ハイスループット実験装置の開発に成功し、この装置を用いて過去30年で蓄積されたデータ数を一桁上回る12000点ものデータをわずか3日で取得することに成功した。得られた触媒ビッグデータを機械学習などによって分析し、実験条件による C_2 収率の変化や鍵となる触媒組成を突き止めることに成功した。

2. Jun Fujima, Yuzuru Tanaka, Itsuki Miyazato, Lauren Takahashi, Jun Fujima, Keisuke Takahashi, “Catalyst Acquisition by Data Science (CADS): a web-based catalyst informatics platform for discovering catalysts” *Reaction Chemistry and Engineering* (2020) 5(5), 903-911.

概要: 触媒インフォマティクスのための統合データプラットフォーム「CADS: Catalysts Acquisition by Data Science」の開発に成功し、2019年度より本格運用を開始している。URLは(<https://cads.eng.hokudai.ac.jp/>)である。現在まで登録アカウント数76、平均月8000程度のアクセス数となっており、国内外の大学・企業など幅広い方々に利用されている。現時点でのCADSの機能は「触媒データの公開・管理」「データの可視化」「機械学習による学習・予測」が実装されており、ウェブ上でマウスのクリック1つでインフォマティクスの一連の解析を行うことが可能となっている。このような触媒データセンターはこれまでになく、世界中からアクセスがあり、企業から実装したいという問い合わせが多数ある。CADSに関する論文は RSC の *Reaction Chemistry & Engineering* で発表し、プレスリリースも行っている。

3. Lauren Takahashi*, Thanh Nhat Nguyen, Sunao Nakanowatari, Aya Fujiwara, Toshiaki Taniike, Keisuke Takahashi “Constructing catalyst knowledge networks from catalyst big data in oxidative coupling of methane for designing catalysts”, *Chem. Sci* (2021), 12, 12546-12555

概要: 機械学習では実現し難い「理解を伴う触媒インフォマティクス」を推進するため、オントロジーを用いた触媒ビッグデータのネットワーク化を提唱した。ハイスループット実験データにおいて、触媒組成、実験条件、 C_2 収率の関係性を知識のネットワークとして可視化することにより、元素組み合わせや反応条件と C_2 収率の関係性が提示され、理解を伴う触媒設計を可能にした。このネットワークを「触媒世界地図」と名付け、K-V-Eu-BaO 等の C_2 収率 20% を超える触媒の実験実証にも成功した。成果は RSC の *Chemical Science* に採択され、“pick of the week”にも選出され、プレスリリースも行った。本方法論は触媒開発の新たな方法論を提案するものであり、多くの触媒データに適用可能な技術である。

§2 研究実施体制

(1) 研究チームの体制について

① 「ハイスループット計算・データ科学」グループ

研究代表者: 高橋 啓介 (北海道大学理学研究院化学部門、教授)

研究項目

・ データベース構築・整備

- ハイスループット計算プログラム開発・運用
- 文献データの集積

・ データベースからの知識

- インフォマティクス(機械学習・データ可視化など)による触媒データからの材料設計
- 活性化エネルギーの高速予測技術開発
- 触媒記述子探索と触媒提案技術の確立
- インフォマティクスによる反応経路理解

・ データプラットフォーム

- データベース環境整備
- プラットフォーム「CADS」の開発・運用

② 「データ科学・反応経路」グループ

主たる共同研究者: 宇野毅明 (情報・システム研究機構国立情報学研究所、教授)

研究項目

・ データベース構築

- 反応経路マップ(メタンの化合プロセスにおける中間物質ネットワーク)作成

・ データベースからの知識

- 機械学習手法開発(触媒の作用の予測に対するデータマイニング的な手法の開発)
- クラスタリングによる触媒データからの知識の抽出

③ 「触媒開発・解析」グループ

主たる共同研究者: 大山 順也 (熊本大学大学院先端科学研究部、准教授)

研究項目

・ データベース構築

- 実験によるデータ集積

・ データベースからの知識

- 実験実証(触媒開発・中間体・最終体の特定・触媒構造解析)
- 計測データの集積
- 実験プロセスの最適化

④ 「ハイスループット実験」グループ

主たる共同研究者: 谷池 俊明 (北陸先端科学技術大学院大学先端科学技術研究科(物質化学領域)、教授)

研究項目

・ データベース構築

- ハイスループット OCM 及びメタノール実験システム開発
- ハイスループット実験システムによる OCM・メタノールデータ集積

・ データベースからの知識

- データ科学による触媒設計方法の確立

⑤ 「実験・学習検証」グループ

主たる共同研究者: 西村俊 (北陸先端科学技術大学院大学先端科学技術研究科(物質化

学領域)、准教授)

研究項目

・ **データベースからの知識**

- ハイスループット実験データに基づく OCM 触媒提案の実験実証および学習モデル鍛錬
- 文献データに基づく OCM 触媒提案の実験実証および学習モデル鍛錬

(2)国内外の研究者や産業界等との連携によるネットワーク形成の状況について

谷池・西村は、北陸先端科学技術大学内にマテリアルズインフォマティクス国際研究拠点を形成し、触媒インフォマティクスを含む国際共同研究・教育を推進している。