

提言

化学・情報科学の融合による
新化学創成に向けて



令和2年（2020年）7月7日

日本学術会議

化学委員会

化学企画分科会

この提言は、日本学術会議 化学委員会 化学企画分科会 情報科学との融合による新化学創成小委員会での審議結果を踏まえ、化学委員会化学企画分科会において取りまとめ、公表するものである。

日本学術会議 化学委員会 化学企画分科会

委員長	加藤 昌子	(第三部会員)	北海道大学大学院理学研究院化学部門教授
副委員長	茶谷 直人	(第三部会員)	大阪大学大学院工学研究科教授、大阪大学環境安全研究管理センター長
幹事	君塚 信夫	(第三部会員)	九州大学大学院工学研究院応用化学部門主幹教授
幹事	関根 千津	(第三部会員)	株式会社住化技術情報センター取締役副社長
	相田 美砂子	(第三部会員)	広島大学理事・副学長
	阿尻 雅文	(第三部会員)	東北大学材料科学高等研究所教授
	菅原 洋子	(第三部会員)	北里大学理学部名誉教授
	所 千晴	(第三部会員)	早稲田大学理工学術院教授
	中村 栄一	(第三部会員)	東京大学総括プロジェクト機構特任教授、東京大学大学院理学系研究科特任教授、東京大学名誉教授
	橋本 和仁	(第三部会員)	国立研究開発法人物質・材料研究機構理事長、東京大学教授
	渡辺 芳人	(第三部会員)	総合研究大学院大学理事
	岡本 裕巳	(連携会員)	自然科学研究機構分子科学研究所教授
	川合 眞紀	(連携会員)	自然科学研究機構分子科学研究所所長
	酒井 健	(連携会員)	九州大学大学院理学研究院化学部門教授
	高原 淳	(連携会員)	九州大学先導物質化学研究所主幹

日本学術会議 化学委員会 化学企画分科会 情報科学との融合による新化学創成小委員会

委員長	阿尻 雅文	(第三部会員)	東北大学材料高等科学高等研究所教授
副委員長	松原 誠二郎		京都大学大学院工学研究科教授
副委員長	山下 善之		東京農工大学大学院工学研究院教授
幹事	阿久津 典子	(連携会員)	大阪電気通信大学工学部教授
幹事	佐藤 一彦		産業技術総合研究所触媒化学融合研究センターセンター長
	相田 美砂子	(第三部会員)	広島大学理事・副学長
	加藤 昌子	(第三部会員)	北海道大学大学院理学研究院化学部門教授
	君塚 信夫	(第三部会員)	九州大学大学院工学研究院応用化学部門主幹教授
	菅原 洋子	(第三部会員)	北里大学理学部名誉教授

関根 千津	(第三部会員)	株式会社住化技術情報センター取締役副社長
茶谷 直人	(第三部会員)	大阪大学大学院工学研究科教授
所 千晴	(第三部会員)	早稲田大学理工学院教授
中村 栄一	(第三部会員)	東京大学総括プロジェクト機構特任教授、東京大学大学院理学系研究科特任教授、東京大学名誉教授
橋本 和仁	(第三部会員)	国立研究開発法人物質・材料研究機構理事長、東京大学教授
渡辺 芳人	(第三部会員)	総合研究大学院大学理事
伊藤 耕三	(連携会員)	東京大学大学院新領域創成科学研究科物質系専攻教授
上村みどり	(連携会員)	帝人ファーマ株式会社生物医学総合研究所上席研究員
石原 司		産業技術総合研究所生命工学領域バイオメディカル研究部門主任研究員
江崎 宣雄		三井化学生産技術高度化推進室長
川尻 喜章		名古屋大学大学院工学研究科教授
杉本 邦久		(公財)高輝度光科学研究センター (JASRI) 主幹研究員
外輪 健一郎		京都大学工学研究科化学工学専攻教授
高橋 孝志		横浜薬科大学薬学部薬科学科特任教授
中川 敦史		大阪大学蛋白質研究所 附属蛋白質解析先端研究センター教授
永田 賢二		物質・材料研究機構統合型材料開発情報基盤部門主任研究員
長谷部 伸治		京都大学国際高等教育院教授
船津 公人		東京大学大学院工学研究科教授
山本 浩史		自然科学研究機構分子科学研究所教授

本提言の作成にあたり、以下の職員が事務を担当した。

事務局	犬塚 隆志	参事官 (審議第二担当)
	五十嵐 久留美	参事官 (審議第二担当) 付参事官補佐
	加藤 雅之	参事官 (審議第二担当) 付審議専門職付

要 旨

1 背景

今日、情報科学技術、特に人工知能 (AI) 技術の進展はめざましく、今後化学者と AI とがうまく協調できれば、これまでにない速度で新しい原理・原則の発見や新物質の開発が期待できる。化学の各分野での AI の利用は、化学の研究手法の変革のみならず、教育改革、さらに物質生産を通して化学産業構造にまで影響を与える可能性がある。

2 現状および問題点

AI は、構造決定、機能解析、材料設計、合成経路設計、反応予測などの化学の各面で、高効率化の鍵として、また研究者・技術者の思考を支援するツールとして、積極的に利用され始めている。このような分野で AI を活用するためには、信頼できるデータベースが不可欠であるが、我が国の科学データベースに対する意識は極めて低く、実用的なデータベース構築はほぼ海外に依存してきた。そして、近年データ駆動型科学の機運の盛り上がりにもとない、従来のエネルギー資源確保の戦略と同様に、AI 技術を活用するための自前のデータベースの確保が最重要と認識されるに至っている。今後、我が国においてもデータベース構築を推進していくべきではあるが、既に世界に大きく遅れをとっている現状や新たなデータベース構築にかけられる人的、経済的資源が限られていることを考えると、既存のデータベースと同じようなデータベースを新たに構築するのではなく、AI 利用を前提とした従来と異なる発想に基づくデータベースの構築が必要である。しかしながら、そのようなデータベース構築に関して、戦略的に検討する機関すら存在しない状況にある。

価値あるデータのデータベース化や、膨大なデータから新たな情報や新たな発想を引き出すためには、AI に関する基礎教育が必要であるが、化学を専門とする教員個人が独力で上述したような教育を行うシステムを構築して行くには限界がある。また、AI を利用した化学の展開においては、データの量とともにその信頼性の向上が重要である。しかしフラスコや試験管を用いる従来の化学実験では、多くのデータを入手することが困難であるだけでなく、得られた結果の信頼性が実験担当者の技量に大きく依存する状況にある。

我が国の化学プラントにおいては、プラントの経年劣化によるメンテナンス負荷の増大や製品の高品質化・高付加価値化への対応策として、AI 技術導入の要求が高まっている。現状においても、直接測れない変数を推定するシステムや、プラントの異常を検知するシステムなどが開発され、広く使われるようになってきているが、これらの支援システムは主に定常状態で機能するものであり、非定常状態での支援については十分ではない。また、異常などの非定常な状態はまれにしか生じないことから、非定常状態を対象として AI 技術を化学プロセスに適用するには、価値のあるデータが圧倒的に不足している状況にある。

3 提言

(1) 化学と情報科学の融合による新化学の創成 (日本化学連合および構成学会)

化学と情報科学の融合によって生まれる新化学が、産業・社会に与える影響の大きさを考えると、化学系学協会内での対応には限界があり産学官の連携による新化学創成を支援する施策が求められる。その中心施策である以下の項目 (2) から (4) の実施を支援する、産学官からなる「新化学創成協議会」を設置すべきである。

(2) AI 利用を意識した化学データの戦略的収集と戦略的創出 (経済産業省、文部科学省)

化学情報・データは、これからの化学技術革新を生み出す貴重な資源である。国として、研究・教育機関に散在する化学データをデータベースとして集約し、広くより機能的に活用するための施策が重要である。そのためには、教育・研究機関における統一的な電子実験ノートの導入、得られるデータベースの構造化を早急に行い、その情報の有効活用を推進すべきである。また、誰が行っても再現性よく同じ結果が得られるデータの蓄積が質の高いデータベース構築の大前提であり、そのためには実験・分析の自動化が望ましい。新たな物質・材料の物性や機能と構造、構造と合成法、さらに合成法と生産プロセス間の関係について、ネガティブデータを含めた新規データベースとして戦略的に収集・創出し、それを重要な資源として管理、有効活用していく仕組みを構築することが重要である。そのためには、高分子、製薬、機能性材料等の分野で化学に特化したデータベースの構築法や活用法の開発や、高度な分析機器と自動合成機能を融合させ質の良いデータを提供できる高機能物質合成システムの開発などを通して、情報化時代の化学に携わる広い分野の研究者の研究力強化を支援する組織（新化学創成センター）を設置すべきである。

(3) 情報科学を活用した化学教育の変革（文部科学省）

情報科学の化学研究での活用を促進するためには、大学での教育の方法から改革する必要がある。AIの素養を有する研究者の育成には、学生の誰もがAIを使う方法を学べる環境づくりが重要であり、それを支援する施策が求められる。化学系学生・研究者に対するビッグデータに基づくAIを活用した研究手法、電子実験ノートの活用法、自動合成手法などの教育カリキュラムへの導入やその教育のための教員研修の早急な実施が求められる。このような化学とAIの融合の教育現場での推進により、我が国の研究者の研究力強化を図るべきである。AI技術は大規模なデータベースがあって初めて効力を発揮する。このような活用法を学習できる環境の構築を、個々の教育機関で行うことは非効率であり、その構築・運用の支援を、上記(2)で提案した新化学創成センター等を活用して実施すべきである。

(4) 情報科学活用による化学産業の高度化（経済産業省、文部科学省、日本化学連合）

新しい高付加価値物質の開発から生産までを迅速化し、安全かつ効率的に生産できるプロセス技術の構築は、我が国の化学産業の生き残りに不可欠である。分子設計から生産までをAI技術で繋ぐためには、機能-構造相関、構造-合成法相関に加え、合成法-プロセス相関、すなわち所望の構造を合成するプロセスの設計に直結するデータベースの構築が重要となる。また、未経験の事態に対応できる運転支援システム構築には、化学プロセス内の動的挙動を正確に表現する精密な物理モデルに基づく新規シミュレータが不可欠である。これらの機能設計からプラント設計までをシームレスにつなげるデータシステムの構築、そして新規シミュレータによるプラントビッグデータ創出を、新化学創成センターを中心に進める。このようなAI技術をベースとした支援システムの構築は、化学産業における生産性向上や設備保全に大きく貢献するだけでなく、熟練技術者の減少を補い、働き方改革にもつながる。営利企業が協力し合える我が国独自のAIと融合した高度生産システムの構築を産学官が協力し進められる体制を構築すべきである。

目 次

1	背景.....	1
2	現状と課題.....	2
(1)	化学と情報科学の融合の現状.....	2
(2)	化学データベースの現状と課題.....	3
(3)	情報化時代の化学人材教育と研究力強化.....	6
(4)	化学と情報科学の融合による化学産業の高度化.....	8
(5)	まとめ.....	11
3	提言.....	12
(1)	化学と情報科学の融合による新化学の創成.....	12
(2)	AI 利用を意識した化学データの戦略的収集と戦略的創出.....	12
(3)	情報科学を活用した化学教育の変革.....	12
(4)	情報科学活用による化学産業の高度化.....	13
	<用語の説明>.....	14
	<参考文献>.....	19
	<参考資料> 審議経過.....	22

1 背景

今日、情報科学技術の進展とその社会への浸透は目覚ましく、その影響により社会の文化も産業構造も変わろうとしている。計算機の回路の集積度は依然 18 ヶ月で倍になるといふムーアの法則に従って上がっており、また単位記憶量あたりのメモリーのコストも、大幅に下がり続けている。ハードウェアの進化に加え、深層学習などの人工知能 (AI) 手法も大きく進化してきている。このような背景から、ビッグデータを確保し、それを基盤として AI を活用するという新たな流れが、米国、中国、欧州を中心に世界中で進行中である。我が国でも、政府は日本経済団体連合会とともに、Society5.0 を提唱し、情報空間と実空間との融合による新たな社会の構築を目指した施策が強力に推し進められている。2018 年に公開された政府の AI 導入構想には、内閣府、経済産業省、文部科学省、総務省が連携しつつ、急ピッチでこの課題に取り組む姿勢と基本施策が示されている[1]。また、日本学術会議においても、情報科学の社会や産業への浸透にともなう科学・技術や社会・文化の変化について、市民と議論する場を作ってきた。また、情報科学の進展にともなう学術や教育のあり方についても議論が進められている。

AI は人間の仕事を代替または支援する技術としてさまざまな分野に進出しており、今後より高度な処理を行えるようになることが期待されている。一方で、既存の枠組みを超えた新規な事項の発見、創造を目指す学術における AI の活用は、その意義・意味を含めて未だ試行錯誤が続く状況である。学術における AI の活用は、人間の作業の単なる代替ではなく、これまで蓄積されてきた膨大な量の知識とデータの両者を用いて、人間が行う思考的生産活動を支援するためのものでなければならない。科学者と AI がうまく協調できれば、これまでにない速度で新しい原理・原則の発見や学術の進歩をもたらすことが期待できる。化学の各分野での AI の利用は、化学の研究手法の変革のみならず、教育改革、さらに物質生産を通して化学産業構造にまで影響を与える可能性がある。産学官で戦略を策定するとともに、それを実行推進する組織構築が必要である。

2 現状と課題

(1) 化学と情報科学の融合の現状

化学は、物質の構造や性質を明らかにし、目の前で起こる現象を解明して原理・原則を導き出すアナリシスと、目的とする機能を有する新物質を実際に創り出すシンセシスからなる。化学の学術・技術は、この両輪がうまく組み合わせることで発展してきた。世の中の有用な物質は、アナリシスの結果としての原理・原則に基づいて、望ましい構造や性質あるいは反応性などを予測しつつ、合成（シンセシス）される。そして、この両分野において、ビッグデータと AI の活用が新しい展開をもたらすことが期待されている。

アナリシスの分野では、機能と構造の相関に基づき、構造決定、物性予測、機能予測の各分野で、ビッグデータと AI の活用が進んでいる。物性および機能解析の基盤として欠かすことが出来ない構造決定は、大きく分光法と回折法に分けられるが、両手法で深層学習を中心に AI を活用する取り組みが既に開始されている。分光法においては分光データと構造の対応関係を AI に学ばせ、構造決定の効率化を図る試みが、また回折法では位相決定などに AI を利用する取り組みが既に開始されている。分析機器とコンピュータの発達により構造決定の測定段階の自動化、高速化はかなり進んできている。よって、同定段階の AI 利用法が確立すれば、構造決定の全段階の自動化が一気に加速することが予想される。この AI 利用に基づく構造決定の迅速化は、創薬や高機能材料創成を前進させる重要な技術の 1 つとしても期待されている。

材料開発（シンセシス）では、まず多くの候補物質から、求める物性を示す物質を選定する。この物質選定とその合成条件の多様性から材料開発には多大な時間と費用がかかっているのが実状である。このような状況を改善するため、望ましい物性を有する候補を AI を用いて導出する仕組みを構築する試みが、無機物、有機物、高分子そしてこれらの複合材料の各分野でスタートしている。そして、このような物性予測に基づく材料設計・試作のいくつかの成功例が既に報告されている[2-6]。また、特定の機能を有する材料の作成条件を確定する作業においても、AI 利用の模索が始まっており、結晶成長などにおいて既にその有用性が示されている[7, 8]。

機能性材料や創薬分野で対象となる分子は、複数の不斉炭素、官能基、置換基を含む複雑で巨大な分子であることが多く、そのような分子を効率良く、かつ実用性も考慮して合成する経路を見いだすことは容易ではない。1969 年、米国でコンピュータ支援による有機合成が発案されたが、当時はハードウェアの能力もデータベースも十分ではなく、実用には至らなかった[9]。近年、ソフトウェアとハードウェアの進展に加え、電子データも充実してきており、コンピュータ支援有機合成が再度注目を集めている。Merck 社では、Grzybowski が開発した Expert システム Synthia™ [10]を用い、有機合成ルートを高い精度で設計しうるシステムの運用を 2018 年より開始している。更に、このような有機合成経路設計にデータ駆動型の深層学習を組み合わせた手法の研究も拡がりつつある[11]。また、既存の実験結果に基づく合成経路探索だけではなく、未知の反応に対して量子化学計算を用いて機械学習のためのデータを作成するなどより高い信頼

度とより広い範囲での合成経路探索を可能にする手法も提案されている[12]。さらにこのような手法により、新規な触媒設計指針を与えうることも示されている[13]。

この現象の解明（アナリシス）と物質の合成（シンセシス）の双方を常に考慮する必要がある化学分野において、新規物質の開発から実際の工業生産までをシームレスに連携させるためには、物性・機能－構造相関、構造－合成法相関だけでなく、合成法－プロセス相関を明確にし、機能－構造－合成法－プロセスをつなげていく必要がある。

(2) 化学データベースの現状と課題

前節で述べてきたように、AI は構造決定、機能解析、材料設計、合成経路設計、反応予測などの各面で、高効率化とともに、研究者・技術者の思想的生産活動を支援するツールとなる可能性を有している。しかしながら、このような AI システム構築には、信頼できるデータベースとそれを活用する AI 手法を使いこなせる研究者や技術者の育成が不可欠である。

かねてより我が国の科学データベースに対する意識は極めて低く、実用的なデータベース構築はほぼ海外に依存してきた。近年データ駆動型科学の機運の盛り上がりにともない、ようやく、従来のエネルギー資源確保の戦略と同様に、AI 技術を活用するための自前のデータベースの確保が最重要と認識されるに至っている。化学分野では、情報科学がビッグデータを中核として取り入れられると、学術のみならず産業も含めて、既存データを利用して生まれる新たな知見・技術が、新たな良質のデータになるという正のスパイラルを生むことから、今後化学分野はより一層データに依存することになると想定される。このような点を踏まえ、経済産業省では、化学産業界からの要請を受けて、学術論文、特許、社内文書などから材料開発に必要なデータを自動抽出するプラットフォーム構築を支援するプロジェクトを 2019 年度からスタートさせている[14]。これまで遅れていた我が国のデータ基盤を一挙に充実させようとの狙いである。このほか、材料設計に有効に使える計測データの体系的收拾のための機器の開発にも取り組み始めている[15]。圧倒的に不足している化学分野のデータ取扱いに習熟した人材を育成するカリキュラムの作成とその実施についても、同じく経済産業省の支援により 2019 年度から行われる運びとなった[16]。データベースは、単にデータを集めれば良い訳ではなく、データクレンジングやデータキュレーションという専門的な作業が必要である。当然このような作業ができる人材が必須であるとともに、この先にあるデータのモデル化や解析を行える人材の育成が急務であることを考えると、上述したような政府の動きはまさに時期を得たものと言える。

しかしながら、化学データベースの構築においては、以下に示すように、データの量、質、構造、収集法等において、多くの課題が残されている。

① 既存のデータベース

化学関連のデータベースとしては、現時点で2億以上の物質情報を有する米国化学会の情報部門である Chemical Abstracts Service (CAS) の SciFinder、1億近くの

物質情報を有する米国国立衛生研究所傘下の国立生物工学情報センター（NCBI）が提供する PubChem、100 万余の有機物結晶構造に関する The Cambridge Structural Database、そして 15 万余りの蛋白質や核酸の立体構造に関する Protein Data Bank（PDB）等がある[17]。日本では、主に無機物質・材料・高分子について、国立研究開発法人物質・材料研究機構（NIMS）が、MAT Navi を提供している。また、上述した PDB は生命科学研究・創薬分野において欠くことのできない重要なデータベースとなっているが、我が国からは、大阪大学蛋白質研究所が運営する日本蛋白質構造データバンク（PDBj） [18]が PDB の世界 4 拠点の一つとして、アジア地区からの蛋白質原子構造情報の登録作業を担当し全データの約 1/4 をデータ処理している。それに加えて蓄積されたデータを基にした独自の検索サービスや 2 次データベースの提供なども行っている。

実験結果のみならず、計算化学も組み込んだデータベース構築も進められている。米国においては、2011 年から実験・データ科学・計算化学を連携させた手法を用いて、多数の物質・材料についての膨大なデータを生成し、新材料や新機能の開拓を加速することを目指した活動が、Materials Genome Initiative として始まっている。研究費総額は、日本円換算で 250 億円以上であり、マサチューセッツ工科大学（MIT）と Lawrence Berkeley National Laboratory による Materials Project を始め、複数の大型プロジェクトが実施されている。計算化学の導入により、実験的に確認されていない物質も大量にデータベースに組み込まれ、未知物質が知財化されてしまう危険も生まれつつある。

② 新しい考え方でのデータベース構築の必要性

上記①で述べたように、化学物質に関するデータベースは、我が国も寄与しているものの欧米を中心とした国々に依存しているのが実状である。エネルギー資源の確保が、これまでの産業や経済を支える上で重要であったことと同様に、情報・データは、将来の技術革新を生み出す貴重な資源である。巨大科学データベースが海外の限られた場所に集中している現状とその危険性を考えると、その対策は国家戦略として大きな課題といえる。データベース構築が既に海外にリードされている状況と、新たなデータベース構築に多くの人的、経済的資源が必要な点を踏まえると、我が国においては既存のデータベースと同じようなデータベースを新たに構築するのではなく、AI 利用を前提とした従来と異なる視点での化学データベースを設計していくことが望ましい。例えば、機能設計から分子設計、合成・プロセス設計までを見据えたデータベース構築が考えられる。機能設計結果を製造にまで結びつけるには、目的とする機能を発現させるための分子や材料構造を明らかにした後、その構造を有する物質の合成経路を推定し、さらに経済性を考慮して合成するためのプロセスの提示が必要となる。すでに、2（1）において、化学分野におけるシームレスな連動の重要性を述べたが、化学データベース構築においても同様の考え方が重要となる。すなわち、機能－構造相関、構造－合成法相関だけでなく、合成法－プロセス相関もとることにより、材料

設計から生産までを一貫して行うことが可能となるため、そのようなデータベースは、新たな機能性材料等の創造を迅速に行うための基盤技術となり得る。現時点では、機能-構造-合成法-プロセスの相関を考慮したデータベース構築例はほとんどない。その一因は、複合材料や相分離ポリマーのように、構造を必ずしも十分に規定できない無秩序系の材料も多く存在することである。結晶のように、構造を規定できる場合には、機能-構造-合成法-プロセスの相関が比較的とりやすいが、主な機能性材料であるアモルファスや複合材料、階層化された構造の場合、分子の化学構造や結晶構造とは異なり、そのような構造を記述する手段がこれまでなかった。最近、数学と材料の連携拠点 WPI-AIMR において、無秩序構造の表現法の開発が進められている[19]。このような、無秩序構造の構造表現法といった、日本オリジナルな方法を活用することにより、機能-構造-合成法-プロセス相関をイメージしたデータベースを構築できる可能性があるが、現状そのようなデータベース構築について、統一的に検討する組織が存在しない。

③ ネガティブデータの必要性

新たなデータベースを構築する際には、そこに入力するデータの種類や表現法、質についても検討が必要である。従来、化学分野の研究においては、実験データは、主として実験者や実験指導者（教員）個人の考え方に沿って採取され、実験ノートという形でグループ内のみで共有されてきた。失敗等のネガティブデータが外部に発表されることはない。よって、表にでるのは成功例だけであった。これらの全ての情報は、全体を管理するリーダー（教授や主任研究員）の頭の中で整理され、構造化され保持されてきたといえる。従って、リーダーが変われば、公表されていないデータやノウハウ等は散逸してしまう例が多かった。有能な合成研究者は、うまく進む反応経路の情報のみならず、うまく進まない反応経路（ネガティブデータ）も考慮して、新たな反応経路を探索する。AI を利用した合成経路探索においても、ポジティブな結果のみならず、合成研究者であれば意識せずに用いているネガティブなデータが重要な役割をはたす可能性が高い[20, 21]。よって、ネガティブデータなど AI 利用を前提とした時に重要となる可能性の高いデータを収集し構造化できれば、既存の超大規模データベースと視点の異なるデータベースとすることが可能である。しかしながら、現在電子実験ノートの活用法や上述したデータベース構築について、統一的に議論する場が存在しない状況にある。

④ データの質確保のための自動合成

AI 利用を前提としたデータベースでは、信頼できないデータの混入は、得られた結果の信頼性に大きく影響する。これまで、フラスコや試験管を用いてバッチ方式で様々な物質が合成されてきたが、このような方法では得られた結果が実験担当者の技量に少なからず依存する。誰が行っても再現性よく同じ結果が得られるデータを蓄積することが、質の高いデータベース構築の大前提となる。そのためには、実験・分析の自

動化は不可避である。すでに、Cronin らは、「有機化学のデジタル化」というコンセプトで、実験の条件の自動設定からインライン法によるデータ取得まで一貫して行えるシステムを構築している[22]。このような自動合成システム（自動実験ロボット）を活用することで、試薬の添加のタイミングや攪拌条件等、実験の再現性を向上させることができる。フラスコや試験管を用いた合成実験法は、基本的には数百年前のものと同じと言ってよい。それに対して、近年マイクロ化学プロセス、あるいはフローケミストリーと呼ばれる流通系の装置を用いて実験を行うことが提唱されている。バッチ装置での反応時間が、流通系装置では装置の位置に対応するため、実験の始点と終点のみならず、遷移過程のデータを容易に得ることができる。高速混合や急速加熱・冷却操作が可能であり、早い反応や高発熱の反応に対しても、フラスコでは取得が困難であったデータを容易に得ることができる。流通系自動合成設備を導入し、網羅的に合成実験を行い、紫外・可視・近赤外・赤外分光法やラマン分光法等のインラインでの測定法を介して遷移状態を含む多くの情報をデータとして取り込むことができれば、従来と異なるレベルの情報からなるデータベースを構築できる可能性があるが、このようシステムが一般的に普及している状況ではない。

今後、大学や研究機関の研究室に自動合成設備等が導入されることにより、ネガティブデータを含む膨大なデータが自動的に蓄積可能になる。それを電子実験ノート等を用いて適切に管理し、データを構造化（研究室内クラスターデータベース構築）していくことが望ましい。まずは、大学の研究室レベルからスタートし、協力が得られる複数の研究室から成るクラスターの形成、さらにその範囲を研究機関、企業研究所に拡張し、国レベルで集積することにより、従来と異なるビッグデータからなる我が国独自のデータベースが構築できる。ただし、このようなシステム構築には、大学や研究機関の意識改革や情報化時代にふさわしい化学人材教育が不可欠である。この点については、次節で説明する。

(3) 情報化時代の化学人材教育と研究力強化

AI を活用するためには、統計や情報科学に関する基礎的な知識が欠かせない。しかしながら、我が国の大学学部教育における統計学教育は諸外国と比較して著しく不足している[23]。文部科学省は、AI 人材育成における統計学の重要性を認識し、「数理・データサイエンス基礎」と位置づけて全学部・学科で実施することを検討している[24]。文部科学省から発表された基本計画[25]では、文系、理系、芸術に至るまで、あらゆる分野の大きな柱として情報科学教育を位置付けており、教育改革の重要項目の一つとなっている。情報科学と化学の融合を考えた時、情報化時代の化学人材に求められる情報処理能力は、膨大な量のデータを AI を用いて望ましい方法で処理し、必要な情報を抽出する能力（情報利用力）、実験データを AI で利用可能な価値ある情報として収集し発信する能力（情報生成力）、さらに、新しい知見や価値の創造のために新たなデータベース構築やデータ記述法を開発する能力（システム構築力）に分類できる。

情報化時代の新たな化学では、大学での研究開発と化学産業における技術開発、製品開発がより密接に連動する。それを支える大学と産業における人材の育成は、同時並行で進める必要があり、新たな教育の方法が求められる。そこには、情報利用力、情報発信力、システム構築力の各点において多くの課題が挙げられる。

① 情報利用力教育

化学者は、これまで限られた情報から飛躍的な思考を通して結論を仮定し、それを実証することで、新たな合成法や新規機能製品を生み出してきた。しかしながら、利用可能な情報量が増えるに従い、化学者個人の有する情報量の全情報量に対する比率は減少してきている。よって、大規模データベースとAI技術により、化学者が思考作業を行うのに必要な適切な情報が提供されれば、研究者の独創性・創造性は、これまでより活性化され、化学者の研究力強化が期待できる。言い換えれば、今後、化学最先端の研究には、多くの情報を適切に処理し、利用する技術が不可欠と言える。このような研究遂行スタイルの変化に応じられる対応力を教授することが、今後の大学、研究機関での教育に求められている。新しい研究遂行スタイルに対応した化学教育・研究の体制を早期に確立するためには、実規模の大規模データベースを用いてAI手法の活用法を学べ、かつ化学の特長を生かした情報科学教育システムを開発する必要がある。AI関連のソフトウェア技術は急速に発展しており、特別な情報科学の知識を持たない化学者であっても、機械学習のようなAI手法を利用できる状況にある[26]。しかしながら、現状において化学を専門とする多くの教員はAI利用に関する十分な訓練を受けていない。また、各個人の研究に加えて、非常に進歩の早い最先端のAI技術を常にフォローすることは困難であり、教員個人が独力でAI手法を取り入れた化学教育システムを構築するには限界がある。

② 情報発信力教育

前項①で提示したような化学の特長を生かした情報科学教育システムが構築できれば、それを活用することで、化学に特化した法則発見型・新着想創出型教育プログラムを構築でき、大規模データベースとAIを活用しつつ、そこから法則を見いだしていく力をもつ化学人材を養成できる。そのような教育は、化学に携わる研究者や技術者に資源としてのデータの価値を認識させることになる。その結果、これまで捨てられたり研究室内で死蔵されたりしていた実験で得られた多くのデータについて、研究者や学生がどのようなデータが情報化時代に価値のあるデータかを認識できるようになり、情報として活用できるようになる。前節で述べたように、ネガティブデータや遷移過程のデータなどの膨大なデータが電子実験ノート等を用いて研究室内で適切に管理され、あるいは研究室レベルでデータベース化されることは、将来我が国独自のデータベース構築の重要なインフラとなる。大学や研究機関において生成される情報のシステムティックな蓄積を図るためには、将来のAI利用を前提とした電子実験ノートのフォーマットや研究室内データベースシステムの構築、全体データベースの構

造化等が必要であるが、現状そのような課題解決を推進する組織が存在しない。

③ AI に適した物質表現とシステム構築力

化学では、実験で得られる結果を、分子の構造変化ととらえ、それを化学式としてあらわしてきた。このような分子の表現方法が、「人」の現象理解を深め、発想を刺激し、新規な発見、発明を生み出してきた。しかし、情報科学が導入されたことで、化学構造の表記法についても、異なる視点から再考する動きが見えつつある。現在、AI を用いた反応経路探索や反応予測においても、分子構造のインプットとアウトプット、およびその変化をどのように表現するかという点で、議論が進んでいる。例えば、化学構造式を、CHONSP といった元素の結合の有無を示す行列で記述し、反応を変換行列として記述する方法は古くから広く使われてきた。最近では、グラフ理論も導入され(分子グラフ)、さらにグラフの構成成分に特徴量(結合角度や原子の属性など)を与えるなど、その表現にも新たな工夫が行われている。AI の活用という意味では、反応の記述は化学式である必要はなく、種々分析で得られたシグナル(スペクトル)であっても良い。扱う情報は膨大とはなるが、分析シグナルは分子の特徴量そのものである。最近の種々の機械学習を用いた研究によれば、分子を何らかのベクトル量で表現した方が良い推定結果が得られるという報告もある[27]。今後、化学の各分野において、情報科学に精通し、このような新たな物質表現を含めた新たな研究開発を行える研究者を育成していく必要がある。

(4) 化学と情報科学の融合による化学産業の高度化

上記 2 (2)、(3) 項では大学や研究機関での状況を中心に述べてきたが、情報科学と化学の融合は、最終的にものを生産する化学産業にも大きな影響を与える。我が国の化学産業においては、団塊世代の経験豊富な技術者の大量退職や少子化による人手不足、新規プラントの建設の減少や連続運転の長期化による技術者の経験不足が大きな問題となっている。また、プラントの経年劣化によるメンテナンス負荷の増大や製品の高品質化・高付加価値化への効率的な対応の要求が高まっている。さらに、エネルギー多消費産業であることから今まで以上の省エネルギー化が望まれている。これらの問題に対応するために、AI 技術を活用し必要なときに適切な情報を提示することで、経験の少ない少人数でもプラントを安全かつ効率的に運転・保守できる高度な運転支援システムや自動制御システムの構築が期待されている。現状においても、ある規模以上の化学プラントは、連続自動運転されており、プラント内の多くの情報が、ネットワークを介して分散制御システム(DCS)やより上位のコンピュータに送られ、最適運転や監視、制御に使われている。また、装置の温度や圧力の情報を用いて直接測れない変数を推定するシステム(ソフトセンサー)の開発や、プラントが通常と異なる状態になったことを知らせる異常検知システム、さらに異常箇所を推定する異常診断システムなどが開発され、広く使われるようになってきている。しかしながら、上述した支援システムは主に定常状態で機能するものであり、プラントのスタートアップやシャットダウン、異常発生後

の対応等、非定常状態に対する支援については十分ではなく[28]、AIを用いた支援システム開発の必要性が叫ばれている状況にある。このようなシステム開発には、化学プラントと情報処理技術の双方に精通した技術者が必要であるが、化学プロセスの特徴を踏まえた AI 利用に関して、十分な知識を有する人材は不足している[29]。その解決のためには、AI のわかる人材の早期育成が不可欠であり、化学やプロセス技術の専門家が情報通信技術を学び、データを有効に活用できるスキルを身につけた人材となるように教育することが望まれている。それに対し、経済産業省の第四次産業革命スキル習得講座認定制度において、「プラント運転・保安 IoT/AI 人材育成講座」[30]が実施されている。また、産業データ活用推進事業が 2017 年度の補正予算で盛り込まれ[31]、Connected Industries のもとでデータ活用の推進事業[32]も始まっている。

このように AI 人材育成に関する事業は展開されているが、将来の AI 活用として必須の非定常状態への対応という観点では、以下に示すように未解決の課題が多い。

① 非定常データの不足

AI 技術の進歩は、これまで経験豊かな技術者に依存していた異常時対応等に対して、それを支援するシステムを構築できる可能性を有している。異常時の最適操作手順の導出は、ある状態から正常な状態へ、ある評価（時間や消費エネルギー）を最適にするように遷移させる問題であり、連続変数を扱う大変さはあるが最終形が定まった詰将棋のようなものと考えられ、AI 技術の活用により大きな成果が期待できる。またそのようなシステムが構築できれば、異常時対応のみならず、スタートアップやシャットダウン操作の最適化、目的製品変更時の対応最適化等に広く展開できる。現状でも AI ツールとして販売されている商品は数多くあるが、これらのツールは価値のあるデータが大量に存在していることを前提としており、データが大量にあっても価値のあるデータが不足している場合にはほとんど役に立たない。プラントの運転のために必要な温度や圧力、流量といったデータについては、情報通信技術の進歩によって着実に蓄積されるようになってきたが、データの大部分は定常状態での値である。AI 技術を用いて非定常状態を対象とした支援システムを開発するには、大量の非定常データが必要であるが、異常などによる非定常な状態はまれにしか生じないため、価値のあるデータが圧倒的に不足している状況にある。

② 動的シミュレータの活用限界

非定常データ不足を補うための有効な手法の一つは、シミュレーション技術の活用である。現時点においても、デジタルツイン、ミラープラントといったシミュレータをプラントの運転と同期させ、シミュレーションに必要なパラメータを合わせこみつつ将来を推定する方法が導入されつつあり、これにより、測定されていないプラントの状態や近未来のプラントの状態を推定することが可能となってきている。また、実プラントとシミュレータのずれを利用したプラント制御の支援ツールも導入されつつある。しかしながら、現状の動的プラントシミュレータは、物理現象を忠実に再現す

るものではなく、予め想定した操作に対して、その挙動が現実らしく見えることに力点が置かれて構築されている。プロセスシミュレータを、異常時等の予め設定されていない状況に対しても、実プラントと同様の挙動を示すものとするためには、装置内の各箇所で見られる物理現象、さらにはその非定常過程をも考慮するなど、現在のシミュレータとは次元の異なる精度と機能が要求される。このような物理法則に基づく新たなシミュレータが開発できれば、それを様々な設定で実行させることにより大量の仮想データを得ることができる。そして、そのデータをAIに学習させることによって異常診断や異常からの回復シーケンスを自動的に導出できる可能性がある。このようなシミュレータ開発は、単独の企業や大学では困難であり、開発に着手されていない状況にある。

③ データベース

機能性材料や創薬の分野では、開発から生産までのリードタイムの短縮が競争力強化の一つの要点である。現状多くのケースでは、製品開発はビーカーやフラスコを用いてバッチ式に行われており、そこで得られたデータのみでは実生産プロセスを設計できず、実生産までにスケールアップや連続化、最適運転条件設定等に多くの時間を要している。製品開発時には最大収率となる操作条件や使用溶媒のみが報告される場合も多く、生産設備を設計するために再度同じような実験が繰り返されている。化学者が行った多くの試行錯誤情報は、データベースとして整理されておらず、生産設備設計を任された技術者はその情報にアクセスできる状況にない。2(2)項で述べたように、化学者の行った最適でない条件での実験結果やネガティブデータ、合成の遷移段階での情報などがデータベースとして整備されれば、プロセス開発期間を大きく削減できる可能性がある。そのためには、2(2)項で述べたような、機能—構造—合成法—プロセス相関をイメージしたデータベースの構築とその活用法の検討を、産学官共同で、また、純粋化学、応用化学、化学工学技術者が共同して進めていく必要があるが、そのような組織が構築できていないのが実状である。

④ 働き方改革

化学産業においては、情報の集約化は既に達成されているとみなして良く、どこでもプラントの状態を把握し、操作可能である。特にAIの支援により、誰もが場所と時間にとらわれずに働けるようになれば、今までとは全く違った「働き方」となり、様々な制約により未就労状態におかれている有用な人材の登用が可能となる。育児、家族との時間、生涯教育、社会コミュニティといった生活や文化も大きく変わることが予想される。また、高度な運転支援システムの導入は、経験値の少ない若手や外国人労働者にも働く場を提供することにもなる。このような将来構想については、セキュリティ問題など運転支援システム構築を越えた様々な課題を含んでいるが、統一的に議論されていない状況にある。

(5) まとめ

以上、研究開発の場、データベース、教育、化学産業の視点から情報科学の導入の現状と課題を述べてきた。化学と情報科学の融合は、化学に関する教育や実験手法、研究開発、さらには産業技術の展開、働く場の環境を大きく変革し、また新たな学術分野を創成する可能性も秘めている。

3 提言

以上の議論を提言として以下にまとめる。また、それぞれの提言に関連する主な省庁・機関を付記した。

(1) 化学と情報科学の融合による新化学の創成（日本化学連合および構成学会）

化学と情報科学の融合によって生まれる新化学が、産業・社会に与える影響の大きさを考えると、化学系学協会内での対応には限界があり産学官の連携による新化学創成を支援する施策が求められる。その中心施策である以下の項目（2）から（4）の実施を支援する、産学官からなる「新化学創成協議会」を設置すべきである。

(2) AI 利用を意識した化学データの戦略的収集と戦略的創出（経済産業省、文部科学省）

化学情報・データは、これからの化学技術革新を生み出す貴重な資源である。国として、研究・教育期間に散在する化学データをデータベースとして集約し、広くより機能的に活用するための施策が重要である。そのためには、教育・研究機関における統一的な電子実験ノートの導入、得られるデータベースの構造化を早急に行い、その情報の有効利用法を推進すべきである。また、誰が行っても再現性よく同じ結果が得られるデータの蓄積が質の高いデータベース構築の大前提であり、そのためには実験・分析の自動化が望ましい。新たな物質・材料の物性や機能と構造、構造と合成法、さらに合成法と生産プロセス間の関係について、ネガティブデータを含めた新規データベースとして戦略的に収集・創出し、それを重要な資源として管理、有効活用していく仕組みを構築することが重要である。そのためには、高分子、製薬、機能性材料等の分野で化学に特化したデータベースの構築法や活用法の開発や、高度な分析機器と自動合成機能を融合させ質の良いデータを提供できる高機能物質合成システムの開発などを通して、情報化時代の化学に携わる広い分野の研究者の研究力強化を支援する組織（新化学創成センター）を設置すべきである。

(3) 情報科学を活用した化学教育の変革（文部科学省）

情報科学の化学研究での活用を促進するためには、大学での教育の方法から改革する必要がある。AI の素養を有する研究者の育成には、学生の誰もが AI を使う方法を学べる環境づくりが重要であり、それを支援する施策が求められる。化学系学生・研究者に対するビッグデータに基づく AI を活用した研究手法、電子実験ノートの活用法、自動合成手法などの教育カリキュラムへの導入やその教育のための教員研修の早急な実施が求められる。このような化学と AI との融合を教育現場で推進により、我が国の研究者の研究力強化を図るべきである。AI 技術は大規模なデータベースがあって初めて効力を発揮する。このような活用法を学習できる環境の構築を、個々の教育機関で行うことは非効率であり、その構築・運用の支援を、上記（2）で提案した新化学創成センター等を活用して実施すべきである。

(4) 情報科学活用による化学産業の高度化（経済産業省、文部科学省、日本化学連合）

新しい高付加価値物質の開発から生産までを迅速化し、安全かつ効率的に生産できるプロセス技術の構築は、我が国の化学産業の生き残りに不可欠である。分子設計から生産までを AI 技術で繋ぐためには、機能－構造相関、構造－合成法相関に加え、合成法－プロセス相関すなわち所望の構造を合成するプロセスの設計に直結するデータベースの構築が重要となる。また、未経験の事態に対応できる運転支援システムの構築には、化学プロセス内の動的挙動を正確に表現する精密な物理モデルに基づく新規シミュレータが不可欠である。これらの機能設計からプラント設計までをシームレスにつなげるデータシステムの構築、そして新規シミュレータによるプラントビックデータ創出を、新化学創成センターを中心に進める。このような AI 技術をベースとした支援システムの構築は、化学産業における生産性向上や設備保全に大きく貢献するだけでなく、熟練技術者の減少を補い働き方改革にもつながる。営利企業が協力し合える分野で、我が国独自の AI と融合した高度生産システムの構築を産学官が協力し進められる体制を構築するべきである。

<用語の説明>

回帰式

Y が連続値の時にデータに $Y = f(X)$ というモデル（「定量的な関係の構造」）を当てはめる事を回帰分析という。回帰分析は統計学から生まれたが、現在では機械学習の一分野である。回帰分析の結果得られた関係式を回帰式という。

回折法

波はおもに結晶の表面の原子と衝突して多方向に散乱されるが、原子が規則的に配列されていると、各原子から特定の方向に散乱された成分が干渉して強め合い、反射光に模様が生じる。その模様から物質内部の原子配列を推定する。

回復シーケンス

プラントの運転において、その状態を望ましい状態に回復するために行う一連の操作のことを回復シーケンスと呼ぶ。通常は、オペレータが知識と経験に基づいて実行している。

官能基

構造が確認されている有機化合物の数は 300 万以上といわれるが、その性質（反応性）によっていくつかのグループに分類される。同じグループに属する化合物が示す共通の反応性の原因となる原子団、または結合様式を官能基という。

機械学習

コンピュータやロボットなどの機械に自動的に概念や行動プログラムを学習させる研究分野。さまざまな分野で多岐にわたる手法が開発されており、特に、深層学習（後述）の発展により、人間による規則の入力無く学習することが可能となってきた。

深層学習

ディープラーニング。コンピュータによる機械学習の一手法で、人間の脳神経回路を模したニューラルネットワークを多層的にすることで、コンピュータ自らがデータに含まれる潜在的な特徴をとらえ、より正確で効率的な判断を実現させる技術や手法。音声認識や自然言語処理、画像認識・判別などの分野を中心に広く用いられている。

データキュレーション

デジタル資産を選択、保存、維持管理、組織化そしてアーカイブする一連の行為である。

データ駆動型

得られたデータを総合的に分析しモデル化することによって、予測や意思決定などに役立てること。特に、ビッグデータを対象とし、各種データを解析して課題解決に結びつけることを指すことも多い。

データクレンジング

データベースの中から誤りや重複、欠損を洗い出し、異質なデータを取り除いたり補填して整理すること。データベースの精度を高めることにより、データ解析の誤りを減らし精度を大幅に高めることができる。

電子実験ノート (Electronic Laboratory Notebook, ELN)

実験ノートを電子化することで、紙では困難であった過去のデータの検索性が向上することにより、個人の中でしか蓄積し難かった知識を、組織全体の知識として活用することも可能となり、紙のノートを保管するための場所と費用の削減を実現することができる。更に不正行為やデータの改ざん防止のために、実験ノートに対して、誰がいつ何をどのように記録したかのログを自動的に記録するような機能がついた電子実験ノートも出てきている。

ネガティブデータ

実験者が、あらかじめ設定した目標に達せず採用されなかったデータ。

ソフトセンサー

温度や湿度、圧力を測る装置は、それぞれ「温度計」「湿度計」「圧力計」である。これら実測値を直接測る装置をハードセンサーと呼ぶ一方で、化学プラントで化学反応が進んでいく最中の「密度」や「重合度」といったパラメータは直接測定することができない。しかし、サンプリングして間欠的に分析した測定値とハードセンサーとの間の関係をモデル化しておくことによって、その値をリアルタイムで推定できるようにしたソフトウェアのことを「ソフトセンサー」と呼ぶ。

ビッグデータ

情報通信技術 (ICT) の進歩によって収集、蓄積、分析できるようになった膨大かつ多様なデータ。

フローケミストリー

連続フローまたはプラグフローケミストリーとも呼ばれる化学反応で、連続的なフローの中で稼働する。反応体が混合装置に注入されて、温度制御されたパイプやチューブ、マイクロリアクターの中を流れていき、このフローは反応が終了するまで続く。

不斉炭素

一分子中の炭素原子に四個の互いに異なる原子または原子団が結合しているとき、この炭素原子を不斉炭素原子といい、*C で表わす。分子中に不斉炭素原子が1個存在すると、一对の光学異性体が存在する (たとえば乳酸)。

分光法

物質に種々の波長の光を当てて、その吸収などから物質の特性を調べたり、定量したりする方法。

分散制御システム (DCS)

制御システムの一つで、制御装置が脳のように中心に一つあるのではなく、システムを構成する機器ごとに制御装置があるもの。制御装置はネットワークで接続され、相互に通信し監視し合う。化学工場や製鉄プラントなどによく使われている。

分子グラフ

原子をノード、結合を辺に対応させて表記した化合物の表現方法。AI 処理に適合した「化学式」の表現。

マイクロ化学プロセス

微小空間の特性を利用して化学反応、細胞培養、物質分析などを有利に進め、ある資源からより付加価値の高い製品を生産する工程。

マテリアルズインフォマティクス

コンピュータによる情報科学の手法を材料科学に取り入れた学問分野。データマイニングや人工知能を用いることで、さまざまな材料を組み合わせることで繰り返し実験を行う従来の手法に比べ、新規材料や代替材料の探索などを効率よく行うことが可能となる。

ムーアの法則

大規模集積回路 (LSI IC) の製造・生産における長期傾向について論じた一つの指標であり、経験則に類する将来予測である。米インテル社の創業者のひとりであるゴードン・ムーアが 1965 年に自らの論文上に示したのが最初であり、その後、関連産業界を中心に広まった。

溶媒効果

反応性もしくは分子の会合に対して溶媒が及ぼす影響。

量子化学計算

量子化学では、(通常の化学の範囲内で) 物質の構成最小要素である原子核や電子を取り扱う Schrödinger 方程式を、可能な限り精密に解くことに基づいている。この方程式を精密に解くことにより (原理的には) 分子構造と電子数の情報を与えるだけで、物質の物性や反応性を評価できる。

AI

人工知能。artificial intelligence の略。1950 年代半ばから研究がはじまった。コンピュータの目覚ましい発達を背景にし、学習、推論、認識、判断など、人間の脳の役割の一部を機械に代替させようという研究分野、あるいはそのコンピュータシステムをいう。

Expert システム

人工知能研究から生まれたコンピュータシステムで、人間の専門家（エキスパート）の意思決定能力をエミュレートするものである。専門家のように知識に基づく推論によって複雑な問題を解くよう設計されており、通常のプログラミングのようにソフトウェア開発者が設定した手続きに従うわけではない。1970 年代に人工知能の研究者によって開発され、1980 年代にわたって商業的に適用され、AI ソフトウェアとして最初に成功を収めた形態である。日本語訳では専門家システムという場合もある。

Expert システム Synthia™

有機化合物の合成に関する過去の膨大な報告例から、目的の物質の実用的な有機合成ルートを見出すメルクグループの expert システム。

Mat Navi

物質・材料に関する世界最大級のデータベース。新材料の開発、材料の最適な使用、最適な材料の選択に役立つソリューション獲得を支援。

Materials Genome Initiative

最新の情報技術や計算科学を駆使し、優れた新材料の発見から実用化までのスピードを 2 倍に早める計画。アメリカ政府がホワイトハウス主導で 2011 年から始めた国家プロジェクトである。英語名の頭文字をとって MGI と略してよばれるほか、日本では材料ゲノム計画といわれることもある。2000 年代に、膨大な DNA データをコンピュータ解析技術やデータ管理技術を駆使して解析し、生命科学に革新をもたらしたヒトゲノム（遺伝情報）解析計画になぞらえ、情報技術や計算科学を応用して材料分野に技術革新をもたらすため、マテリアルゲノム計画とよばれる。計画にゲノムの名がついているが、必ずしも遺伝情報を利用するわけではない。

Protein Data Bank (PDB)

「JST-NBDC」と「大阪大学蛋白質研究所に措置された共同利用・共同研究拠点経費（文部科学省）」の支援を受け、米国 RCSB、BMRB、および欧州 PDBe と協力して、生体高分子の立体構造データベースを国際的に統一化された PDB アーカイブとして運営するとともに、様々な解析ツールを提供。

PubChem

化学分子データベースの一つ。このシステムは、アメリカ国立衛生研究所 (NIH) の下の国立医学図書館 (NLM) の一部門である国立生物工学情報センター (NCBI) によって維持管理されている。ウェブユーザインタフェースを通して自由に接続する事ができ、数百万の化合物構造および記述のデータセットを FTP 経由でダウンロードすることが可能である。PubChem に集積されているのは 1000 原子および 1000 結合より少ない小さな分子である。

SciFinder

物質科学関連分野に強い情報検索ツール。

Society5.0

日本政府により閣議決定された科学技術政策の基本指針の一つ。人工知能・ビッグデータ・ユビキタス関連の情報技術を従来の技術と組み合わせ、社会のあらゆる分野で新しい製品やサービスを提供できるよう、研究や開発、投資を進めようとするもの。

<参考文献>

- [1]人工知能技術戦略会議、「人工知能技術戦略実行計画」(平成30年8月17日)、
<https://www8.cao.go.jp/cstp/tyousakai/jinkochino/keikaku.pdf>
- [2] “Computer-Aided Screening of Conjugated Polymers for Organic Solar Cell: Classification by Random Forest” S. Nagasawa, E. Al-Naamani, A. Saeki, *J. Phys. Chem. Lett.* 2018, 9, 2639.
- [3]https://jpn.nec.com/press/201802/20180209_04.html
- [4]<https://pr.fujitsu.com/jp/news/2018/03/16.html>
- [5] “Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments” F. Ren, L. Ward, T. Williams, K. J. Laws, C. Wolverton, J. Hattrick-Simpers, A. Mehta, *Science Advances*, 2018, 4, 1566.
- [6] “Hunting for Organic Molecules with Artificial Intelligence: Molecules Optimized for Desired Excitation Energies” M. Sumita, X. Yang, S. Ishihara, R. Tamura, K. Tsuda, *ACS Cent. Sci.* 2018, 4, 1126.
- [7] “High-speed prediction of computational fluid dynamics simulation in crystal growth” Y. Tsunooka, N. Kokubo, G. Hatasa, S. Harada, M. Tagawa, T. Ujihara, 2018, 20, 6546.
- [8] “Classification of crystallization outcomes using deep convolutional neural networks” A. E. Bruno, P. Charbonneau, J. Newman, E. H. Snell, D. R. So, V. Vanhoucke, C. J. Watkins, S. Williams, J. Wilson, *PLoS ONE*, 2018, 13, e0198883.
- [9] “Computer-Assisted Design of Complex Organic Syntheses” E. J. Corey, W. T. Wipke, *Science*, 1969, 166, 178.
- [10] “Computer-Assisted Synthetic Planning: The End of the Beginning” S. Szymkuć, E. P. Gajewska, T. Klucznik, K. Molga, P. Dittwald, M. Startek, M. Bajczyk, B. A. Grzybowski, *Angew. Chem. Int. Ed.* 2016, 55, 5904.
<https://www.sigmaaldrich.com/japan/chemistry/chemical-synthesis/synthesis-software.html>
- [11] “Planning chemical syntheses with deep neural networks and symbolic AI” M. H. S. Segler, M. Preuss, M. P. Waller, *Nature*, 2018, 555, 604.
- [12] “Interrogating selectivity in catalysis using molecular vibrations” A. Milo, E. N. Bess, M. S. Sigman, *Nature* 2014, 507, 210.
- [13] “Machine Learning Approach for Prediction of Reaction Yield with Simulated Catalyst Parameters” A. Yada, K. Nagata, Y. Ando, T. Matsumura, S. Ichinoseki, K. Sato, *Chem. Lett.* 2018, 47, 284.
- [14]新エネルギー・産業技術総合開発機構、「Connected Industries 推進のための協調領域データ共有・AI システム開発促進事業」、
https://www.nedo.go.jp/activities/ZZJP_100157.html
- [15]新エネルギー・産業技術総合開発機構、「省エネ製品開発の加速化に向けた複合計測分

析システム研究開発事業」 https://www.nedo.go.jp/activities/ZZJP_100136.html

[16]新化学技術推進協会、「化学×デジタル人材育成講座」

<http://www.jaci.or.jp/jo/pub/chemdigi.html>

[17]Protein Data Bank: the single global archive for 3D macromolecular structure data
wwPDB consortium Nucleic Acids Research, 2019, 47(D1): D520

[18]New tools and functions in Data-out activities at Protein Data Bank Japan (PDBj)
Kinjo, A. R., Bekker, G.-J., Wako, H., Endo, S., Tsuchiya, Y., Sato, H., Nishi, H.,
Kinoshita, K., Suzuki, H., Kawabata, T., Yokochi, M., Iwata, T., Kobayashi, N.,
Fujiwara, T., Kurisu, G. and Nakamura, H. Protein Science, 2018, 27, 95

[19]“Hierarchical structures of amorphous solids characterized by persistent
homology” Y. Hiraoka, T. Nakamura, A. Hirata, E. G. Escobar, K. Matsue, Y. Nishiura,
Proceedings of the National Academy of Sciences USA, 2016, 113, 7035.

[20]“Unpublished results hide the decline effect”, J. Schooler, Nature, 2011, 470,
437.

[21]“Chemistry’s web of data expands” R. Van Noorden, Nature, 2012, 483, 524.

[22]“Controlling an organic synthesis robot with machine learning to search for new
reactivity” J. M. Granda, L. Donina, V. Dragone, D.-L. Long, L. Cronin, Nature, 2018,
559, 377.

[23]日本学術会議数理科学委員会数理統計学分科会、提言「ビッグデータ時代における統
計科学教育・研究の推進について」 2014年8月20日

[24]未来投資会議構造改革徹底推進会合「企業関連制度・産業構造改革・イノベーション」
会合（雇用・人材）第5回

[https://www.kantei.go.jp/jp/singi/keizaisaisei/miraitoshikaigi/
suishinkaigo2018/koyou/dai5/siryoku4.pdf](https://www.kantei.go.jp/jp/singi/keizaisaisei/miraitoshikaigi/suishinkaigo2018/koyou/dai5/siryoku4.pdf)

[25]文部科学省林文部科学大臣9月28日提出資料

www.kantei.go.jp/jp/singi/tougou-innovation/dai2/siryoku2.pdf

[26]「進むAI技術の民主化—AI活用の「アイデア」が問われる時代に」、「未来のテク
ノロジーが実現する新しいビジネスモデルを徹底議論」9回シリーズ(その3):テーマは、
ビジネスにAIをどう使用しているか、2018年7月17日 Web掲載

<https://industry-co-creation.com/industry-trend/33823>

「未来のテクノロジーが実現する新しいビジネスモデルを徹底議論」, ICC サミット
FUKUOKA 2018, (2018年2月20-22日)

[27]“Prediction of Major Regio-, Site-, and Diastereoisomers in Diels-Alder
Reactions by Using Machine-Learning: The Importance of Physically Meaningful
Descriptors” W. Beker, E. P. Gajewska, T. Badowski, B. A. Grzybowski, Angew. Chem.
Int. Ed. 2019, 58, 4515.

[28]経済産業省 産業保安グループ 高圧ガス保安室、「Connected Industries プラント・
インフラ保安分科会の取り組み状況」(平成31年3月15日)

https://www.meti.go.jp/shingikai/sankoshin/hoan_shohi/koatsu_gas/pdf/014_05_00.pdf

[29] マイクロソフト 「Future-Ready Bussiness: AI によるビジネスの可能性について」
(2019年3月1日)

https://news.microsoft.com/wp-content/uploads/prod/sites/47/2019/02/190301_AI_report_Japan.pdf

[30] 経済産業省 「プラント業界における「IoT人材」を育成する講座の開発」

<https://www.learning-innovation.go.jp/verify/detail/c0034>

[31] 経済産業省平成29年度補正予算 「産業データ共有促進事業費補助金」

<https://www.meti.go.jp/press/2018/06/20180601008/20180601008.html>

[32] 経済産業省 商務情報政策局 製造産業局、「Connected Industries 経済対策について」(平成30年3月)

http://www.meti.go.jp/policy/mono_info_service/connected_industries/pdf/economic_measures.pdf

<参考資料>審議経過

平成 29 年

- 12 月 27 日 化学委員会 (第 24 期・第 2 回)
化学委員会化学企画分科会 (第 24 期・第 1 回) 合同開催
役員の確認、新委員の紹介
第 23 期の活動方針
第 24 期の活動方針
情報科学との融合による新化学創成小委員会について
分子研所長招聘会議について

平成 30 年

- 5 月 29 日 「化学と情報科学との融合による新化学創成」
キックオフミーティング、講演を通して意見交換
- 5 月 30 日 分子研所長招聘会議「化学の近未来：化学と A I ・大学の質保証」
- 9 月 30 日 情報科学との融合による新化学創成小委員会 (第 24 期・第 1 回)
役員選出・提言文案資料を持寄り議論
- 12 月 23 日 情報科学との融合による新化学創成小委員会 (第 24 期・第 2 回)
提言・講演に向けての議論 提言 Ver. 1
- 12 月 27 日 化学委員会 (第 24 期・第 6 回)
化学委員会化学企画分科会 (第 24 期・第 2 回) 合同開催
講演・提言案を小委員会から報告し外部から意見聴取
化学委員会・総合工学委員会・材料工学委員会合同
触媒化学・化学工学分科会 (第 24 期・第 5 回)

平成 31 年

- 2 月 11 日 情報科学との融合による新化学創成小委員会 (第 24 期・第 3 回)
提言に向けての意見交換
- 3 月 10 日 情報科学との融合による新化学創成小委員会 (第 24 期・第 4 回)
提言に向けての意見交換

令和元年

- 5 月 29 日 分子科学研究所所長招聘会議「化学の近未来：化学と情報科学との融合」
提言 Ver. 3 に向けて広く意見聴取

5月30日 分子研研究会「化学・情報科学の融合による新化学創成に向けて」
提言 Ver. 3 に向けて外部意見反映

10月16日 化学委員会（第24回・第9回）
提言最終案に向けて意見聴収

令和2年

6月11日 日本学術会議幹事会（第292回）
提言「化学・情報科学の融合による新化学創成に向けて」について承認