

# データ駆動材料研究の進展

～NIMSにおける実践を例として～

出村 雅彦\*

材料分野においてもデータ駆動研究への期待が大きい。本稿では、データ駆動材料開発の最近の取り組みを概説する。材料データは必ずしも豊富にあるわけではないものの、データ駆動による材料・プロセスの探索は、金属、高分子、無機化合物等々へ広がり、自動実験・計測との連携も始まっている。メカニズムの探究など高度な活用や独自のツール開発も進んできている。基盤となる材料データプラットフォームについては、今後、全国的な展開が期待される。

キーワード：データ駆動，材料，材料組織・構造，材料データプラットフォーム

本稿は、クリエイティブ・コモンズ表示 4.0 国際 (CC BY 4.0) ライセンスの下に提供する (<http://creativecommons.org/licenses/by/4.0/deed.ja>)。

## 1. はじめに

材料開発では欲しい性能から最適な元素の組み合わせや作り方を探索する。因果は元素の組み合わせと作り方から性能が決まるという方向なので、材料開発は逆問題といえる。元素の組み合わせの豊富さ、作り方の多様さを考えると、探索空間は広大である。通常は、専門家が、材料分野の知見や予備的な実験・計算に基づいて性能を制御する主な因子を数種類に特定し、最適化を図っていく。ときには「閃き」が無い降りてきて、これまで考えもしなかった答えが見つかることもある。材料専門家としては「経験と勘」に頼ってきた・・・とは言いたくないところだが、考える因子を網羅的に検討しているのかという反省は常にある。

前置きが長くなったが、人工知能やデータ科学に寄せる期待は大きい。これまで見逃していた思いもかけない材料を発見できるのではないかと、もっと速く最適なプロセス条件にたどり着けるのではないかと期待である。本稿では、材料分野の特徴を解説したのちに、データ駆動研究やその基盤となる材料データに関して、主に物質・材料研究機構 (NIMS) での取り組みを中心に紹介したい。

## 2. 材料分野の特徴

材料は物理学、化学、機械学に隣接している (図 1 (a))。物理学ではメカニズム探究が、化学では新しい物質の発見・合成、機械学では力学的な応答が主に対象となる。これら隣接学領域と比して材料学の特徴はどこにあるだろうか。

筆者は材料学の特徴は 2 つに集約されると考える。1 つ

目は不均一な構造への注目である。材料学ではこの不均一な構造のことを組織 (microstructure) と呼ぶ。組織を分類し、形成メカニズムを議論して、特性や性能の理解につなげるという具合に、組織への注目抜きで材料学は成り立たない。図 1 (b) に鉄鋼材料での例を示したように、材料の中では原子から部品までの様々なスケールで不均一が生じており、特性発現の主因や性能劣化の原因となっている。高分子では高次構造と呼ばれている。

もう一つ材料学を特徴付けるのが性能を取り扱う点である。材料は使われる前提に立つため、使用環境において変化する時間依存型の性能を考慮する必要がある。さらに、コストやリサイクル性、カーボンニュートラルなどへの対応も性能の要素に加わる。

材料学は、プロセス、構造 (組織)、特性、性能という 4 つの要素を重要なものとして取り上げ、それぞれの連関を議論する<sup>1)</sup>。データ駆動を適用する上でも、この視点は重要である。

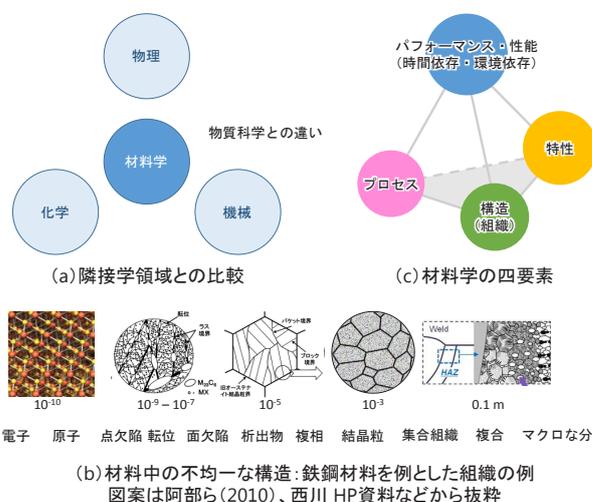


図 1 材料分野の特徴を示す模式図。組織と性能に注目する点が特徴。

\*でむら まさひこ 国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門 (MaDIS)

〒305-0044 茨城県つくば市並木 1-1

E-mail: DEMURA.Masahiko@nims.go.jp

<https://orcid.org/0000-0002-7308-3041>

(原稿受領 2021.3.24)

### 3. 材料分野におけるデータ駆動の広がり

データ時代の材料開発プロジェクトの系譜を図2にまとめた。データに基づく開発やデータベースの構築は以前から行われている。例えば、NIMSでは耐熱合金の開発において長年蓄積してきた実験データを元に特性を予測するプログラムを開発し、これをもとに優れた耐熱合金の開発に成功している。また、MatNavi<sup>2)</sup>という材料データベースを長年、構築・運用している。

データ駆動研究の日本での本格的な立ち上がりは2つの国家プロジェクトをもって嚆矢とする。2014年には内閣府SIP「革新的構造材料」において「マテリアルズインテグレーション」というコンセプトのもと構造材料の研究にデータ科学の適用が明確に謳われるようになり、第2期SIP「マテリアル革命」に引き継がれている。2015年には文部科学省のJSTイノベーションハブ構築事業において「情報統合型物質・材料開発イニシアティブ (MI2i)」が開始された。国家プロジェクトを基軸とした整理については拙稿を参考いただければ幸いである<sup>3)</sup> (なお、材料研究へのデータ科学適用の先駆例としては、15年以上前にベイズニューラルネットワークを用いた藤井らの仕事<sup>4)</sup>を挙げておきたい)。以下では、最近の取り組みを、材料・プロセスの探索、メカニズムの探究、材料分野独自のツールという視点でまとめる。

#### 3.1 材料・プロセスの探索

材料データは探索空間の広大さに比べて少ないことから、データ生成が焦点となる。データ生成が比較的容易な計算材料科学との組み合わせからデータ科学の適用例が多く出てきている。典型的なものは、第一原理計算とデータ科学を組み合わせたものであり、例えば、第一原理計算の結果をもとにデータ科学的な手法でスクリーニングを行っ

て新規の半導体を開発した例<sup>5)</sup>や、第一原理計算に基づいた特性予測計算を自動化してバイズ最適化とを組み合わせ、熱放射特性を最適化した多層膜構造を提案した例が報告されている<sup>6)</sup>。最近では、仮想的な物質の事例ではあるものの、探索に量子アニーリングマシンを活用した例も出ている<sup>7)</sup>。他に、機械学習の活用方法として、計算そのものを高速化する方向性もある。

これら計算材料科学との融合事例に加え、既存の(実験)データベースを活用した事例も報告され始めている。例えば、高分子データベースとして世界最大のPoLyInfo<sup>8)</sup>を活用して、熱伝導性の高い新規高分子を提案した例である。ここでは、データが少なく予測が難しい熱伝導性について、データが比較的豊富な他の物性の学習結果を活用した転移学習によって予測精度を高め、材料提案につなげた工夫がなされている<sup>9)</sup>。さらにデータが少ないケースとしては、熱抵抗の高い元素の組み合わせを探索した研究がある<sup>10)11)</sup>。ここでは、異種元素界面の熱抵抗を調べた論文から集めた約600程度のデータから、元素の記述子を専門家の知見で設計することで予測性の高いモデルの作成に成功している。特筆すべきは、予測モデルからスクリーニングすることで実際に熱抵抗の高い元素の組み合わせ(Bi/Si)を発見している点であろう。論文に掲載された公知データを再利用して材料開発につなげた好例といえる。

これまでは、元素の組み合わせや結晶構造から予測が可能な特性が主な対象であり、電子材料、半導体、熱制御材料、電池材料等の基礎物性に関する適用例が比較的多かった。最近では、これらの基礎物性に加えて、材料の構造が特性や性能を支配するより実用的な特性についても、データ科学を活用した探索が始まっている。内閣府SIP「マテリアル革命」<sup>12-16)</sup>では順問題を解くための計算手法としてマテリアルズインテグレーションシステムを開発し、これをもとに計算材料科学とデータ科学の組み合わせによっ

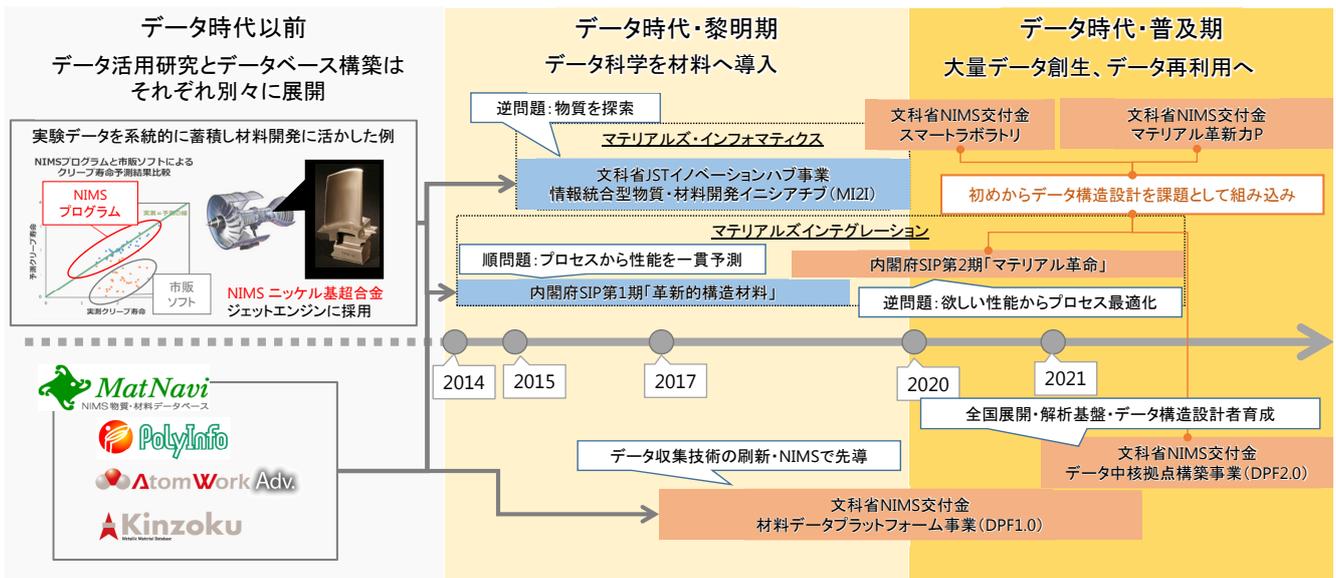


図2 材料分野におけるデータ駆動研究・データ基盤構築の系譜

て、組織に着目した材料・プロセスの最適化、あるいは、最適な材料組織の提案に取り組んでいる。特に、金属系では順問題を解くために複数の計算モジュールを組み合わせるワークフローを自動計算するシステムを構築し、これをMIInt (Materials Integration by Network Technology) システムと呼んでいる<sup>17)</sup>。例えば、耐熱鉄鋼材料の溶接の事例では、溶接熱の影響で材料が変質して高温でのクリープ寿命が低下する現象を取り扱っている。溶接プロセスを模擬し、変質を受けた材料領域を計算機上で特定した上で、溶接継手全体に対してクリープ損傷解析を行うという多段階の計算シミュレーションを組み合わせ、溶接条件からクリープ寿命を予測するワークフローを構築する。これを実験の代替としてベイズ最適化等のデータ科学手法を適用することで、クリープ寿命を最大化する溶接条件を求めることができる。他にも、次世代鉄鋼材料を対象に組織の形態から強度や延性を予測する計算シミュレーションを開発し、ベイズ最適化手法と組み合わせることで強度と延性を最大化する組織を提案するなどの試みが行われている。プロジェクトは途上であり成果の公表はこれからであるが、組織 (= 構造) に着目したデータ駆動研究の今後の展開が期待される。

外挿領域に向かって材料・プロセスの探索を行うためには、実験との組み合わせが欠かせない。いわゆるアクティブラーニングの手法を用いて、実験研究者に対して候補を提示し、実験結果をもとにさらに次の候補を提案するという方法を試みている。うまくいった事例として、熱電素子のプロセス最適化<sup>18)</sup>、新規接着剤の開発<sup>19)</sup>、新規アルコール酸化触媒の探索<sup>20)</sup>、金属微粉末作製条件の最適化<sup>21)</sup>などが出てきている。これらの例はいずれも、実験が定型化され、提案された条件を実際に試す方法が確立している点が共通している。アルコール酸化触媒の例では、従来触媒特性が高いと考えられていた組成領域から大きく外れたところにより高い性能を示す組成が見つかり、実験との組み合わせによる外挿領域探索の効果が出ている。

さらに最近では、自動実験との組み合わせによる材料・プロセスの探索が行われるようになってきた (例えば、参考文献<sup>22)</sup> など)。NIMS ではスマートラボラトリーと称して、合成、計測、データ収集、解析の自動化を進めている。

以上のように、材料・プロセスを探索する観点での応用は、機能性材料から構造材料、金属、酸化物等の無機化合物、有機材料等、幅広く拡大している。そこでは、様々な機械学習による回帰や探索が適用され、必ずしもビッグデータといえない量のデータしかない中で、材料領域の専門知識も活用しながら記述子のデザインやモデルの設計などが工夫されている。

### 3.2 メカニズムの探究

メカニズムの探究にデータ駆動が役立っている事例は、まだ、それほど多くないものの、いくつかの取り組みが進んでいる。手法の観点から、3つに分けて説明する。

第一はモデル選択の枠組みの活用である。特に事後確率

を求めてモデルの妥当性を推定するベイズ推定の枠組みは、材料分野においても威力を発揮すると考える。例として、金属のクリープ現象に適用した事例を紹介する<sup>23)24)</sup>。クリープは10年以上にもわたる非常に長期の時間依存型の事象であり、その理解は火力発電所等の高温設備の安全運用には欠かせない。クリープは初期の変形速度が減速する第一領域を経て、見かけ上一定の変形速度が達成したように見える第二領域を経て、加速的に変形速度が大きくなって破断を迎える第三領域に大別される。寿命の大半を占める第二領域において、実際に材料中で何らかの定常状態が達成されているのか、あるいは、第一領域の減速と第三領域の加速が釣り合って見かけ上定常的な状態が達成されているかで議論があり、これまで結論が出ていない。実際には、材料の種類や試験条件 (温度、負荷応力) によって定常状態の有無も変わる可能性があり、データに基づいて検討したいところである。我々のグループでは、定常状態を仮定する構成方程式と、定常状態を仮定しない構成方程式の2つについて、実際のクリープデータに基づいてベイズ推定の枠組みでモデルの妥当性を検証した。その結果、対象とした鋼種 (高クロム鋼) においては、定常状態を仮定する構成方程式の事後確率がほぼ1であり、定常状態の存在を強く支持する結果が得られている。従来は、本鋼種では組織の時間変化が報告されているが、定常状態がある前提で見直すと、例えば、負荷応力が小さいところでは組織の時間変化が極めて小さいこと、負荷応力が大きいところでは組織の変化が定常的であることなどが見えてきており、メカニズムの理解につながる知見が得られている。この他、同じクリープ現象を対象とするものでは、寿命を支配する元素をモデル選択の枠組みで特定することも行われている<sup>25)</sup>。モデルとデータという前提条件のもとであるが、ベイズ推定によるモデル選択はデータを生成しているメカニズムに近づく有力な方法であると考えている。

第2は機械学習の中で記述子の重要度によって考察を進めるという方法である。ランダムフォレスト、勾配ブースティングなどの決定木を活用したアンサンブル手法では、分岐に使用される記述子の頻度から重要度を判定する方法が提案されており、材料分野でも重要度による支配因子の考察が行われている。例えば、固体リチウムイオン蓄電池の固体電解質について、論文データをもとにリチウムイオンの伝導度を解析した例が報告されている<sup>26)</sup>。ここでは、焼成条件など従来から重要と考えられてきた因子が抽出されるとともに、界面伝導度の予測に理想結晶密度が重要という意外な因子も抽出されている。

第3は、機械学習で得られたモデルからメカニズムにつながる知見を汲み出すという取り組みである。定型的なやり方がなく、ここでは2つの例を紹介する。1つは再び金属のクリープの例である。NIMSには構造材料データシートという40年以上にわたる試験研究のデータが収録された刊行物があり<sup>27)28)</sup>、貴重なデータ資源となっている。このデータを活用してフェライト耐熱鋼のクリープ寿

命を回帰するモデルを作成し、これをもとにある特定の材料規格における微量元素の寿命への感受性を調べた例がある。材料規格で制約されていた範囲を超えた場合に寿命が低下するという興味深い応答が観察され、従来の規格の基準がデータによって追認された結果となった。加えて、規格上は制約が課されていないにもかかわらず感受性のある元素が見出されており、新たな支配因子が発掘された可能性が示唆されている<sup>29)</sup>。

さらに、機械学習の結果を深く用いる例も報告されている。新しい準結晶を探索するために、化学組成から準結晶(や近似結晶)となるかを予測する分類器を作った研究では、記述子の貢献度について解析を加えて、専門家にも理解しやすい準結晶となる必要条件の抽出に成功している<sup>30)</sup>。

### 3.3 材料分野独自のツール

独自のデータ駆動ツールの開発も進んでいる。例えば、化合物の特性予測に使えるツールとしては SMILES-X シリーズがあり、これは化合物構造の表記の一つである smiles 表記を自然言語処理技術によって記述子化し、特性を予測するものである<sup>31)</sup>。当該ツールの予測精度は物性によっては分子動力学法などの計算コストの高い方法を凌駕している。リチウムイオン蓄電池の電解質の探索などへの活用が期待されている。このような機械学習を活用した材料特性予測のプラットフォームとして XenonPy は先駆的な役割を果たしており、材料分野のデータの少なさに対応すべく、転移学習を容易にしている点に特徴がある<sup>32)</sup>。

材料研究に欠かせない計測情報へのデータ科学適用も進んでいる。例えば、材料の化学状態を調べることができる X 線光電子分光法に対しては、スペクトルの分解や試料に含まれる化合物の特定などを情報量基準に基づいて実施できるツールの開発が進んでいる<sup>33)34)</sup>。スペクトルデータの解析は属人性が高く、データ駆動のアプローチが期待さ

れる領域といえる。

## 4. データプラットフォームの構築

データ時代のプラットフォームは他分野を見る限り寡占化が著しく、材料分野におけるデータ戦略にとって今後の10年が重要である。NIMS はこれまでに MatNavi<sup>2)</sup> を世界最大級の材料データベースへと育ててきたが、学術出版や学会母体の材料情報サービスを業とする民間企業の活動が活発化しており、このままでは優位性を保つことは難しくなると認識している。加えて、IT 企業がものづくり企業からデータを受け取って解析結果を提供するといった動きも始めている。これら情報産業の有力プレイヤーにはない我々の特徴は材料の研究を日々行っているということであると考え、日々の研究活動で生み出されている研究データを集約し、これを利活用していくためのプラットフォームを開発することとした。2017年に開始した材料データプラットフォーム事業<sup>35)</sup>は3年が経過し、計測データを中心としたIoTによるデータ自動収集システムの開発<sup>36)</sup>、自然言語処理技術の活用による文献データのキュレーション効率化技術の開発<sup>37)</sup>等の要素開発、データを集約するためのデータ基盤システムの開発を行い、DICE という名称で試験公開を始めている<sup>38)</sup>。

実験・計測の自動化としてのスマートラボラトリーについて3.1で述べたが、これはデータ収集のシステム化と融合させることが肝要である。さらに、データ利活用の促進は、実際の材料プロジェクトの中で実践されていく必要がある。つまり、データ利活用プロジェクト、データ創製スマートラボラトリー、データ基盤の3つが一体となって、材料研究のデータ駆動化を促進していくものとする。

これら NIMS での実践を雛形としてマテリアル DX プラットフォーム構想が提案され、文科省事業として構築が進められようとしている。その中で、NIMS は全国の材料研究のデータを集約していくデータ中核拠点(図3<sup>39)</sup>)

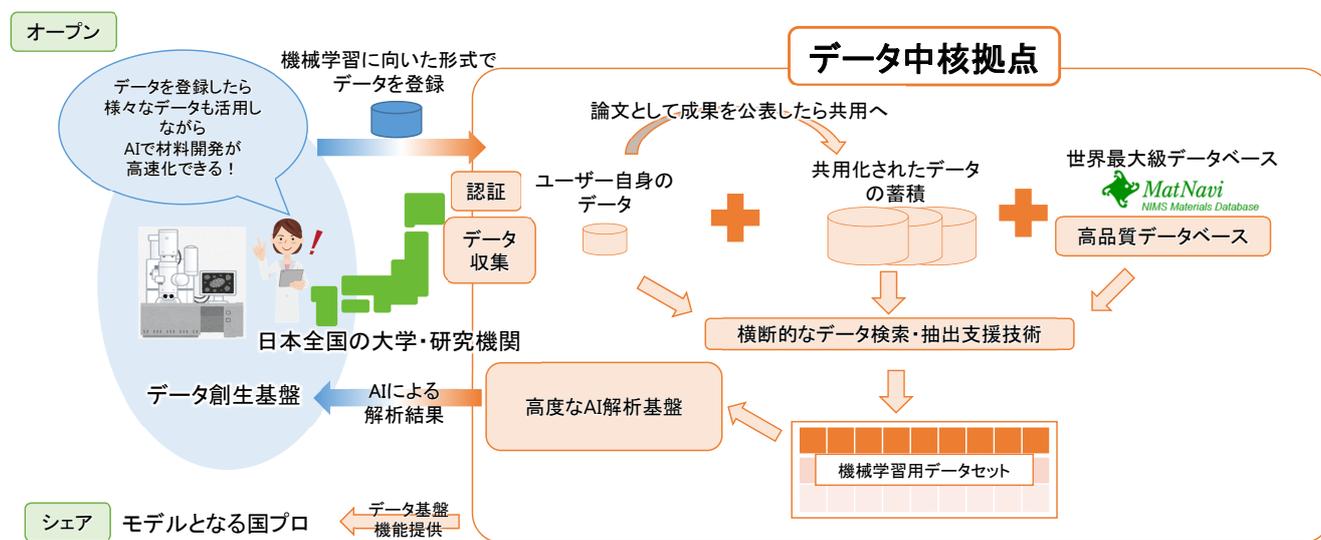


図3 データ中核拠点の構想(初出は参考文献<sup>39)</sup>)

を担うことになる。我が国は材料の研究開発に依然として強みを有している。今ならば、研究から生み出される実データを組織的に収集し再利用を進めていくことで、研究開発力を高めていくことができると期待される。

## 5. おわりに

材料分野におけるデータ駆動研究の進展を中心に述べた。また、材料データプラットフォームの取り組みが全国展開する将来像についても簡単に触れた。材料分野におけるデータ駆動研究は日進月歩で、どんどん新しい試みが生まれている。一方で、データ駆動手法をより広い材料研究者が使っていく普及も求められている。つまり、山の高さを引き上げつつ、裾野を広げるという段階にあると認識している。そのためには、情報の専門家の方とのより一層の連携が欠かせない。本稿が、情報専門家の方々に、データ駆動の適用先として材料分野が面白いと思っていただく一助となることを祈りたい。

## 謝辞

本研究で紹介した一部の研究は、内閣府総合科学技術・イノベーション会議の戦略的イノベーション創造プログラム (SIP)「革新的構造材料」及び「統合型材料開発システムによるマテリアル革命」(管理法人: JST) によって実施された。ここに記して感謝する。

## 参 考 文 献

- 1) Flemings, M. C.; Cahn, R. W. Organization and trends in materials science and engineering education in the US and Europe. *Acta Materialia*. 2000, vol.48, p.371-383.
- 2) MatNavi. <https://mits.nims.go.jp>, (参照 2021-04-20)
- 3) 出村雅彦. AIを活用した材料開発における我が国の取り組み. 第70回白石記念講座 (日本鉄鋼協会編, SMS-ISIJ70, ISSN: 1344-0942). 2018.
- 4) Fujii, H. et al. Bayesian Neural Network Analysis of Fatigue Crack Growth Rate in Nickel Base Superalloys. *ISIJ International*. 1996, vol.36, p.1373-1382.
- 5) Matsuzaki, K. et al. High-Mobility p-Type and n-Type Copper Nitride Semiconductors by Direct Nitriding Synthesis and In Silico Doping Design. *Adv. Mater.* 2018, vol.30, 1801968.
- 6) Sakurai, A. et al. Ultranarrow-Band Wavelength-Selective Thermal Emission with Aperiodic Multilayered Metamaterials Designed by Bayesian Optimization. *ACS Central Science*. 2019, vol.5, p.319-326.
- 7) Kitai, K. et al. Designing metamaterials with quantum annealing and factorization machines. *Physical Review Research*. 2020, vol.2, 013319.
- 8) PoLyInfo. <https://polymer.nims.go.jp>, (accessed 2021-04-07)
- 9) Wu, S. et al. Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. *npj Comput.Mater.* 2019, vol.5, 66.
- 10) Wu, Y.-J. et al. Electrically Conductive Thermally Insulating Bi-Si Nanocomposites by Interface Design for Thermal Management. *ACS Applied Nano Materials* 2018, vol.1, p.3355-3363.
- 11) Wu, Y.-J. et al. Predicting interfacial thermal resistance by machine learning. *npj Comput. Mater.* 2019, vol.5, 56.
- 12) Demura, M.; Koseki, T. SIP-Materials Integration Projects. *Materials Transactions*. 2020, vol.61, p.2041-2046.
- 13) Koyama, T. et al. Development of Microstructure Simulation System in SIP-Materials Integration Projects. *Materials Transactions*. 2020, vol.61, p.2047-2051.
- 14) Enoki, M. Development of Performance Prediction System on SIP-MI Project. *Materials Transactions*. 2020, vol.61, p.2052-2057.
- 15) Inoue, J. et al. Development of Data-Driven System in Materials Integration. *Materials Transactions*. 2020, vol.61, p.2058-2066.
- 16) Minamoto, S. et al. Development of the Materials Integration System for Materials Design and Manufacturing. *Materials Transactions*. 2020, vol.61, p.2067-2071.
- 17) MInt. <https://www.nims.go.jp/MaDIS/MIconso/MInt.html>, (accessed 2021-04-20)
- 18) Hou, Z. et al. Machine-Learning-Assisted Development and Theoretical Consideration for the Al<sub>2</sub>Fe<sub>3</sub>Si<sub>3</sub> Thermoelectric Material. *ACS. Appl. Mater. Inter.* 2019, vol.11, p.11545-11554.
- 19) Pruksawan, S. et al. Prediction and optimization of epoxy adhesive strength from a small dataset through active learning. *Science and Technology of Advanced Materials*. 2019, vol.20, p.1010-1021
- 20) Nugraha, A. S. et al. Mesoporous trimetallic PtPdAu alloy films toward enhanced electrocatalytic activity in methanol oxidation: unexpected chemical compositions discovered by Bayesian optimization. *Journal of Materials Chemistry*. 2020, vol.A 8, p.13532-13540.
- 21) Tamura, R. et al. Machine learning-driven optimization in powder manufacturing of Ni-Co based superalloy. *Materials & Design*. 2021, vol.198, 109290.
- 22) Matsuda, S. et al. High-throughput combinatorial screening of multi-component electrolyte additives to improve the performance of Li metal secondary batteries. *Scientific Reports*. 2019, vol.9, 6211.
- 23) Izuno, H. et al. Data-based selection of creep constitutive models for high-Cr heat-resistant steel. *Science and Technology of Advanced Materials*. 2020, vol.21, p.219-228.
- 24) Mototake, Y. et al. A universal Bayesian inference framework for complicated creep constitutive equations. *Scientific Reports*. 2020, vol.10, 10437.
- 25) Sakurai, J. et al. Descriptor extraction on inherent creep strength of carbon steel by exhaustive search. submitted. 2021.
- 26) Wu, Y.-J. et al. Essential structural and experimental descriptors for bulk and grain boundary conductivities of Li solid electrolytes. *Science and Technology of Advanced Materials*. 2020, vol.21, p.712-725.
- 27) Sawada, K. et al. Catalog of NIMS creep data sheets. *Science and Technology of Advanced Materials*. 2020, vol.20, p.1131-1149.
- 28) Furuya, Y. et al. Catalogue of NIMS fatigue data sheets. *Science and Technology of Advanced Materials*. 2019, vol.20, p.1055-1072.
- 29) Sakurai J. et al. 機械学習による鉄鋼材料のクリープ破断寿命. 日本金属学会 2018 年秋期講演大会. 2018.
- 30) Liu, C. et al. Machine learning to predict quasicrystals from chemical compositions. *Nature Portfolio*. 2021, doi:10.21203/rs.3.rs-240290/v1.
- 31) Lambard, G.; Gracheva, E. SMILES-X: autonomous molecular compounds characterization for small datasets without descriptors. *Mach. Learn. Sci. Technology*. 2020, vol.1, 025004.
- 32) Yamada, H. et al. Predicting Materials Properties with Little Data Using Shotgun Transfer Learning. *ACS Central Sci*. 2019, vol.5, p.1717-1730.
- 33) Shinotsuka, H. et al. Automated information compression of XPS spectrum using information criteria. *J. Electron*

- Spectrosc. 2020, vol.239, 146903.
- 34) Murakami, R. et al. Development of multiple core-level XPS spectra decomposition method based on the Bayesian information criterion. *J. Electron Spectrosc.* 2020, vol.245, 147003.
- 35) 谷藤幹子；吉川秀樹. 材料データプラットフォームシステム DICE における研究データフローの構築—実践と課題. Accepted in *Digital Practice*. 2021.
- 36) M-Dac. <https://dice.nims.go.jp/services/M-DaC/>, (accessed 2021-04-07)
- 37) Oka, H. et al. *Science and Technology of Advanced Materials* in press.
- 38) DICE. <https://dice.nims.go.jp/>, (accessed 2021-04-07)
- 39) 出村雅彦. データ時代の材料・プロセス研究. *スマートプロセス学会誌*. 2021, in press.

**Special feature:** X-Informatics. Data-driven materials research and materials data platforms – An example from NIMS. Masahiko DEMURA (Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan)

**Abstract:** There are high expectations for data-driven research in the materials field. Recent activities in data-driven materials development are overviewed. Although materials data are not necessarily abundant, data-driven exploration of materials and processes has been expanding its application fields from metals through polymers to inorganic compounds, and has begun to be linked with automated experiments and measurements. Advanced applications such as the exploration of mechanisms and the development of original tools are also in progress. The materials data platform is expected to expand its scope into nationwide in the future.

**Keywords:** Data driven researches / Materials / Microstructures / Materials data platform