展望・総説・総論

粉体成膜プロセス研究のハイスループット化のための データ駆動型粉体プロセス・インフォマティクス

長藤 圭介 *・井上 元 †・荒木 拓人 ‡・黒田 義之 §・辻口 拓也 ⁑・長 隆之 ⸸

1. はじめに

日本のものづくり産業は，高度成長期後も競争力を増してきた．しかし，そのうち，大規模集積回路（LSI）をはじめとするデジタルエレクトロニクス産業は，2000 年を境に，国産製造装置と共に他のアジア各国へ技術が流れ，今では一部を除いて当時の産業競争力はない．一方で，蓄電池や燃料電池など，材料・粉体プロセスからなる「アナログ産業」は，その複雑現象から技術流出がされにくく，今でも世界をけん引する産業が残っている．材料開発は，所望の電気特性，化学特性，磁気特性，熱特性，機械特性などを満足する元素の組合せパターンが膨大で，複雑メカニズムのため，論理的設計が困難で，試行錯誤で新材料を見つけ出しているのが現状である．さらに，それらの材料を微細構造にし，適切な性能を発揮させる粉体成膜プロセスも，その現象メカニズムは複雑で，プロセス設計者と現場作業者の勘コツ・すり合わせに依存している．これらのアナログ産業は，日本人の勘コツ・すり合わせの文化が助けて，いまだに競争力がある[1]．しかし，米国の 2011 年のマテリアルズ・ゲノム・イニシアティブ（MGI）[2] に始まり，マテリアルズ・インフォマティクス（MI）[1),3)] の流れで，材料科学と計算科学を融合させた手法で，世界の材料開発スピードが加速している．データ駆動型の材料開発という意味で，日本は米国に後れをとっているだけでなく，絨毯爆撃と AI の組合せで中国にも追い越されつつあり，その猛追は，粉体プロセスにも及びつつある．とはいえ，計算科学では表現できない開発現場での勘コツ・すり合わせの必要性はなくならず，材料開発と粉体プロセスの一貫開発が必要な製品にはまだ競争力があるが，勘コツ・すり合わせの技術にデータ科学を融合させた日本流の手法が今こそ必要と考える．

そこで，この競争力を維持・加速するために，「粉体プロセス・インフォマティクス」というコンセプトを提案した．日本の勘コツ・すり合わせに，間接センシングによる中間データ取得および機械学習という武器を持たせることで，プロセス開発のスループットを大幅に向上し開発期間を短縮，さらにこれまでにない性能の膜の開発，材料探索へのフィードバックに貢献する．このコンセプトは，著者らが科学技術振興機構未来社会創造事業「共通基盤」領域令和元年度に採択された課題である「粉体成膜プロセス研究のハイスループット化のためのデータ駆動型粉体プロセス・インフォマティクス」として推進している[4]．**図 1** に燃料電池を例とした概要を示す．出発材料（カーボン粒子，触媒粒子）を微構造化（燃料電池電極）し，量産（燃料電池スタック）する間には，粉体成膜プロセスレシピ作成とそのサンプルの評価，さらにフィードバック・レシピ改善を繰り返す工程がある．燃料電池は，材料自体の性能だけでなく，粉体成膜プロセスに基づく微構造のつくり込みが欠かせない．本プロジェクトの特徴は，プロセス中の物理現象をなるべく象徴する「中間データ」を取得し，プロセスパラメータと性能（評価データ）を紐づける機械学習システムのための試行実験の組合せを最小限にするところにある．材料開発と量産の間のプロ

* 東京大学 大学院工学系研究科 機械工学専攻
（Keisuke Nagato）
† 九州大学 大学院工学研究院 化学工学部門
（Gen Inoue）
‡ 横浜国立大学 大学院工学研究院 システムの創生部門
（Takuto Araki）
§ 横浜国立大学 大学院工学研究院 機能の創生部門
（Yoshiyuki Kuroda）
⁑ 金沢大学 理工研究域機械工学系
（Takuya Tsujiguchi）
⸸ 九州工業大学 生命体工学研究科人間知能システム工学専攻
（Takayuki Osa）

0368-5713/20/¥500/1 論文/JCOPY

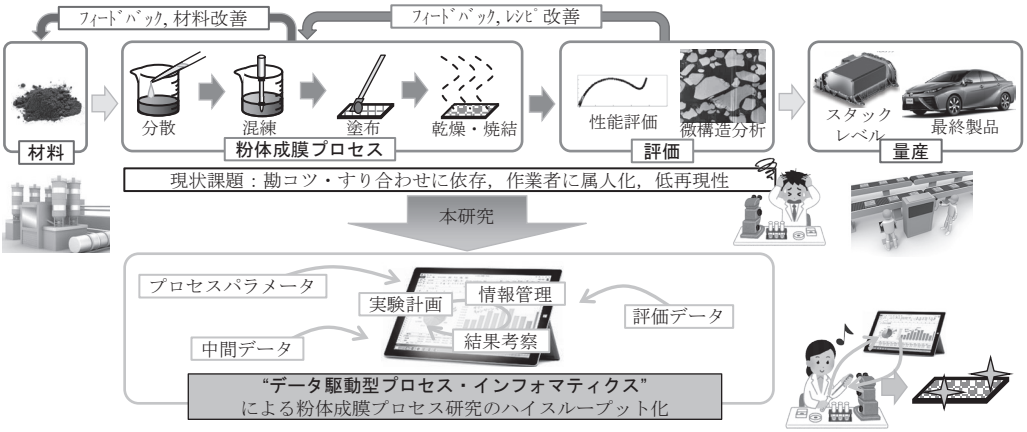


図 1　データ駆動型粉体プロセス・インフォマティクスの概要

セスと評価の部分を，プロセスパラメータ・評価データ・中間データを機械学習システムで紐づけをし，実験計画・情報管理・結果考察をオンラインでおこなうシステムにすることで，粉体成膜プロセス研究のハイスループット化する．

2. データ駆動型研究とは

ここでは，2012 年頃に流行が始まった「機械学習」[5] をきっかけに研究のやり方自体が変わってきているので，その変遷と今回のプロジェクトとの関係を説明する．近代の機械学習の研究そのものは 1950 年代からあり，その礎となる統計や確率の理論は 18 世紀に端を発するものであり長い歴史をもつものだが，近年の流行は深層学習の発展によるところが大きい．そもそも研究においては，背景・目的・仮説・実験・結果・考察・結論のストーリで論文を書いて査読し，シェアするというのが学術界の常套手段である．すなわち，実験前に仮説を立てない研究はないとされてきた（仮に実験後に気づいたとしても，論文中ではそのように記述せねばならない）．しかし，計算機の性能の向上，機械学習による自律的な最適化技術の発展により，データ解析の結果から前もって仮説することができない理論がうまれるという研究が続々

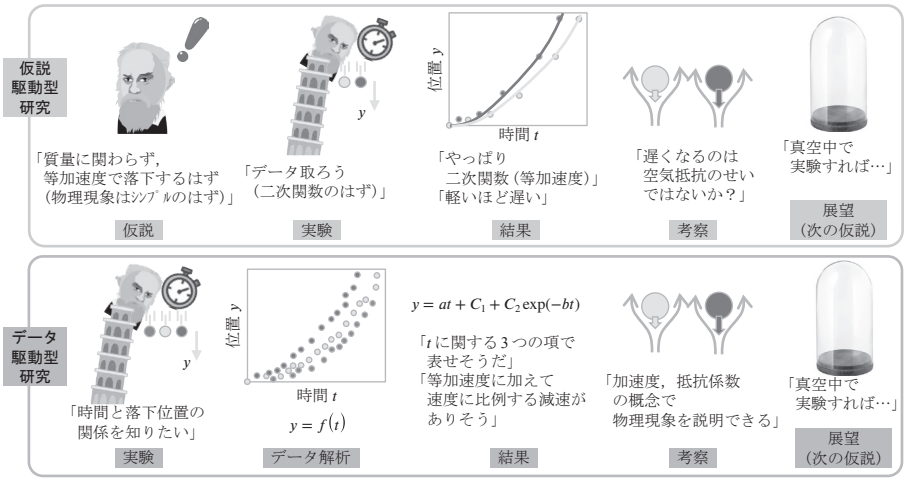


図 2　仮説駆動型研究とデータ駆動型研究

と現れている．

図 2 に，自由落下の物理現象を仮想例として，従来の「仮説駆動型研究」と近年の「データ駆動型研究」の進め方の違いを示す．仮説駆動型研究の論じ方は，実験の前にその仮説を述べる必要がある．高さと時間が 2 次関数の関係にあるはずだからそれを証明しようとする．実験の結果を仮説と照らし合わせ，そのズレを考察し，たとえば空気抵抗といった次の仮説に展開し，結論づける．一方，データ駆動型研究では，目的を単純に「高さと時間の関係を知りたい」と設定し，粛々と実験をし，得られたデータで相関関係を解析し，考察する．データの量にもよるし，解析する際のパラメータの設定に依存するが，重力の概念と空気抵抗の概念を同時に見出す可能性がある．実際の研究では，「まずやってみよう」という要素は少なからずあるため，結果的に仮説駆動型研究論文に仕上げたとしても，実験プロセスではデータ駆動型要素はあるはずである．逆に，データ駆動型研究論文だとしても，あらかじめ仮説をしていないと，設定するパラメータ（図 2 の例の場合，大きさや落下高さ）や実験回数を設計できないはずである．

近年の相関解析技術すなわち，機械学習の技術を全面に利用すると，多くの研究は「やってみたらこうなった」というデータ駆動型研究論文になってしまい，歴史の長い仮説駆動型研究を否定することになるが，データ駆動型の有用性が増してきている中で，それを受け入れなければならない時代になっている．今回のプロセス・インフォマティクスでは，仮説駆動で，原理に則った「中間データ」をセンシング，取得し，データ駆動で「機械学習」をする．複雑現象を理解したいという思いは捨てず，可能な限りそれを象徴するデータを取得し，最後の最後で相関解析を実行するというものである．ここで，暗黙知で仮説駆動型研究をおこなえる日本人の勘コツ・すり合わせ能力の出番というわけである．

3. 粉体プロセスの現状

粉体プロセスは，複雑現象からなると前述したが，その内容を説明する[6]．**図 3** に示すように，大きく，分散・混錬・塗布・乾燥の 4 工程からなる．(a) 分散工程では，粒子と粒子が表面自由エネルギーで集合している二次粒子が，溶媒によって崩壊し一次粒子になる現象，(b) 混錬工程では，粉表面に分散剤（いわゆる界面活性剤）がコーティングされ，粒子同士の相互作用が変化したり，粒子と溶媒との相互作用が増粘剤で変化したりするなどがある．(c) 塗布工程では，たとえばスキージで成膜する方法では，工具と粉との相互作用，粒子と基板との相互作用がある．(d) 乾燥焼結工程では，溶媒自体が蒸発する現象，それにともなっ

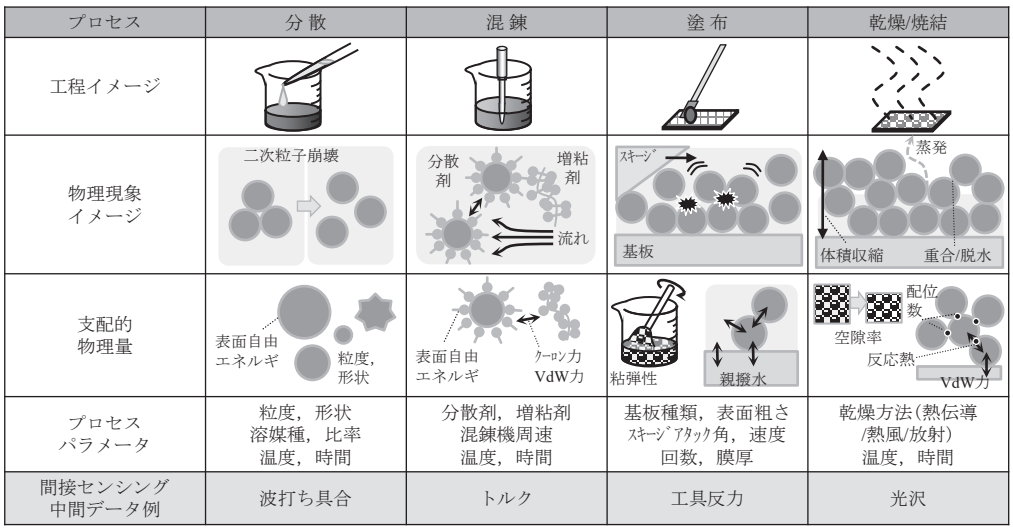


図 3 粉体プロセスの物理現象

て粒子間相互作用の変化や再配列，粒子同士の界面の結晶が移動し融合する拡散現象などがある．これらの現象について，大量の粒子の，形，数，配置などを正確にモデルする解析方法は困難である．もちろん，個々の粒子に着目したモデルに対しては解析する手法はある．外力が加わり粒子自体が変形する材料力学モデル，ぶつかった粒子が運動する機械力学モデル，溶媒が流れるが流体力学モデル，粒子の隙間の溶媒が蒸発する熱力学モデル，これらのモデルを機械系四力学とよぶが，概念が異なる物理モデルを連成させること（マルチフィジックス）だけでも難しく，さらに，多くの粒子を大量に同時に解くこと（マルチスケール）はさらに大規模解析となる．ゆえに，ボトムアップ型の物理現象解明は困難である．

一方で，粉体プロセスで作られている製品は多く，日本の主要な製品とその粒子種を例に挙げると，リチウムイオン電池・固体高分子形燃料電池のカーボン粒子，酸素センサ・ディーゼルパーティキュレートフィルタ・プリンタヘッドのセラミクス粒子，セラミクスコンデンサの電極層のニッケル粒子，レーザプリンタのトナー，インクジェットプリンタのインクなどがある．今後成長余地のある分野としては，ウェアラブルデバイス用のフレキシブル電極の金属インク，熱電素子のセラミクス粒子，金属3Dプリンタの金属粒子などがある．これらは，複雑現象からなる粉体プロセスであるが，所望の微構造を得るためには，粒子や溶媒の条件だけでなく，分散～乾燥・焼結に至るまでの各工程における温度や工具などの条件（これらをプロセスパラメータという）の組合せを最適にする必要があるが，細かい物理現象が解明されていないにも関わらず，実際に実用化している．その大きな理由は，物理現象を勘とコツで把握しながら，工程をまたいですり合わせ，プロセスパラメータを見つけ出す現場主義のエンジニアの成果である．言い換えると，日本人は暗黙知に則った効率的な仮説駆動型研究が得意である．対照的に欧米は，エンジニアとテクニシャンの分業文化が強く，前述した計算科学を駆使したデータ駆動型研究が先行しており，中国をはじめとする国は，多くの試行錯誤に則った絨毯爆撃実験が文化的であるが，最近のデータ駆動型研究手法も取り入れつつあり，日本の勘コツに頼ったプロセス開発には限界がある．

ここで，粉体プロセスを料理に例えてみる．**図4**は固体高分子形燃料電池とお好み焼きの生産プロセスを比較したものである．固体高分子形燃料電池では，出発材料としてのカーボン粉体材料に分散・混錬・塗布・乾燥焼結を施して微細構造を形成し，インピーダンス測定（電池の電極の性能）などで電池特性が評価される．それに対してお好み焼きでは，出発材料としての小麦粉を生地化・他の材料と混錬・鉄板に塗布・加熱により調理され，試食・おいしさの分析をおこなう．ほとんどの日本人はお好み焼きを自作したことがあると思われ

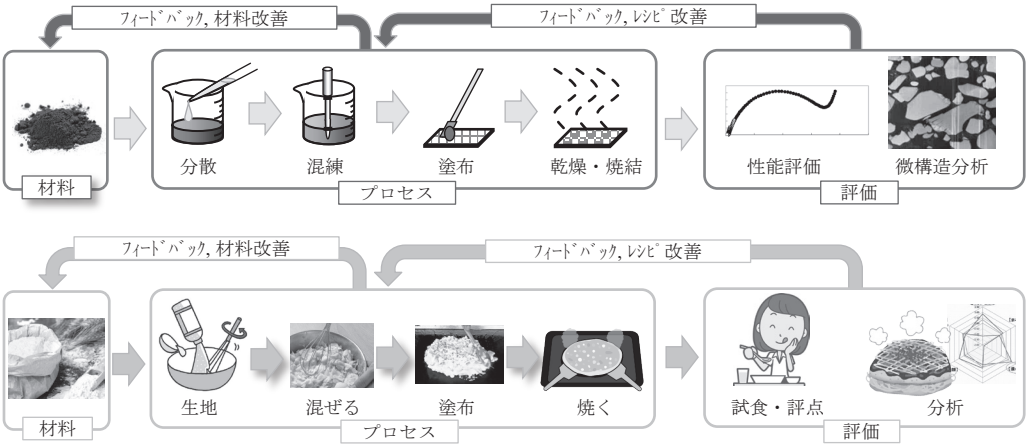
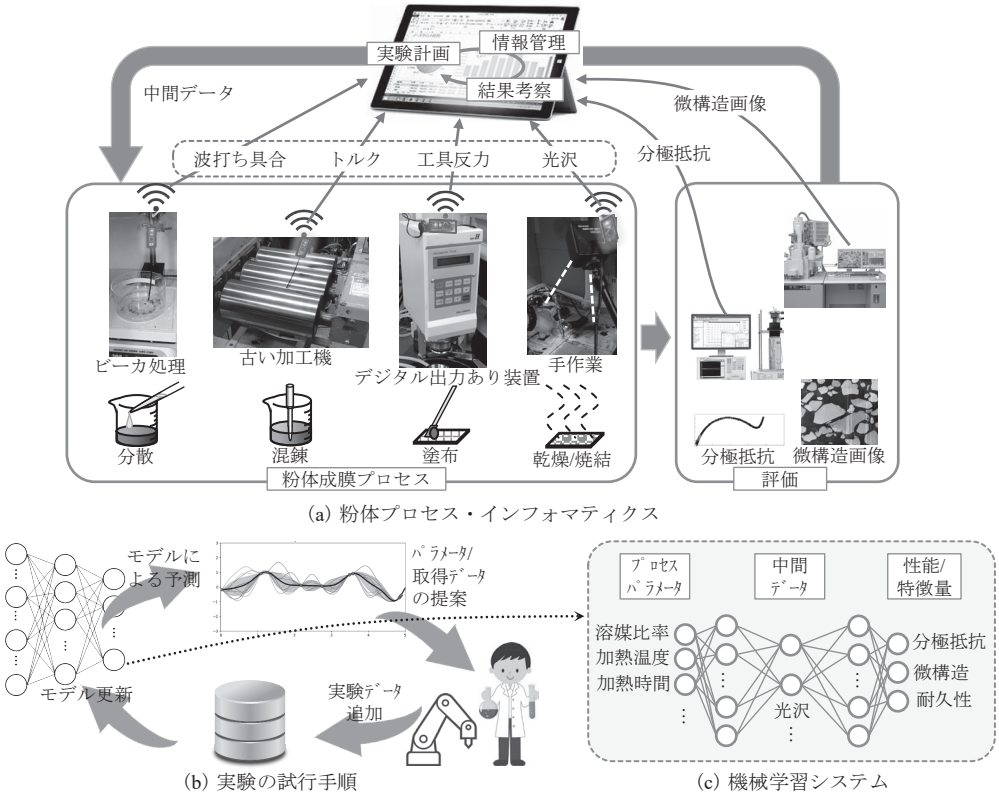


図4　工業製品の粉体プロセスと料理の粉体プロセス

るが，「よい小麦粉」が出発材料だったとしても，適切なプロセスパラメータ（レシピ）を適用しない限り，「おいしいお好み焼き」にならないのと同じで，燃料電池においても「よい特性の材料」が開発されたとしても，その能力を存分に発揮するプロセスパラメータを開発する必要がある．お好み焼きのレシピを数値化するのが難しいのはいうまでもない．たとえば，加熱温度・加熱時間は，微妙な生地の材料条件に依存するし，用いる鉄板の種類によって設定温度と，生地が感じる温度が一対一対応するとは限らない．ここで重要になるのが，本プロジェクトの特徴の 1 つである「中間データ」である．図 3 の下段に示したように，たとえば，乾燥工程時の溶媒の乾燥状態や粒子の配置を，画像から光沢を間接的にセンシングすることが考えられる．溶媒の乾燥状態や粒子の配置を直接センシングするのは難しいし，仮にナノスケールの分解能の X 線 CT で観察可能だったとしても，実工程に実装するのは難しい．光沢の変化を特徴づける数値データと性能データを紐づけられれば，再現性を高めるだけでなく，より高性能な膜を作り出すプロセスパラメータ探索のハイスループット化の可能性がある．中間データを「粉体プロセス・インフォマティクス」の中で，どのように利用するかは次節で説明する．

4. 中間データ取得と機械学習システム

プロセス・インフォマティクスのコンセプトは，化学工学の分野で，化学プラントにおける複数の外部センシングデータから，温度や濃度，反応状態をソフトウェア的に予測し（ソフトセンサとよぶ），化学精製の歩留まりや動的な運転制御を最適にする相関解析の手法に類する[7]．それに対して本プロジェクトでは，材料合成という意味での化学プロセスよりもさらに下流で，より多様な粉体プロセスに着目している．**図 5** (a) に，実験室でのセンシングイメージを示す．分散工程でのビーカ処理，混錬工程での三本ロール，塗布工程の粘



(a) 粉体プロセス・インフォマティクス

(b) 実験の試行手順

(c) 機械学習システム

図 5　プロセス・インフォマティクスにおける機械学習システムのイメージ

度計，乾燥・焼結工程での手作業など，使用する装置もさまざまである．それぞれの工程から出てくる波打ち具合，トルク，工具反力，光沢などの計測値（インライン計測値）を中間データとして取得する．そのプロセスで作製されたサンプルの評価データ，すなわち分極抵抗や微構造画像を，プロセスパラメータと共に管理,その結果の考察，次の実験計画をオンラインでおこなうことを目指す.センシングデータを無線で集めるモジュール，データ処理を効率的におこなうクラウドウェブサービスなどは，最近の製品やサービスで安価におこなうことができる．図5（b, c）にそれぞれ，実験の試行手順と機械学習システムのイメージを示す．人またはロボットが実験をし，そのデータを追加することで機械学習モデルを更新，そのモデルによってベイズ最適化[8]などを用いて検証すべきプロセスパラメータを予測，次の試行実験プロセスを提案する，この繰り返しによって，プロセスパラメータの最適化をおこなう．

このコンセプトは，一般に能動学習とよぶ．機械学習システムでは，インプットにプロセスパラメータを，アウトプットに性能や特徴量を，その紐づけをニューラルネットワークなどでおこなうが，前述したようにプロセスパラメータは膨大にあるため，試行プロセス数が膨大になる可能性がある．また，装置依存性によるバラツキなどを相殺するためにも，試行プロセス数が多くなる．学習効率を高めるため，または律速条件を効率的に探索するために，「中間データ」を用いる．「中間データ」は，プロセス中であるという意味のほかに，機械学習のインプットとアウトプットの中間層に用いるという意味がある．プロセスパラメータから性能や特徴量を推定する過程で中間データは，製造過程で得られる観測を最大限活用するとともに，研究者やユーザが理解しやすいモデルにすることを目指している．人間の知見をモデルの構造に盛り込み，学習を効率化させる，というアプローチはinductive biasなどとよばれ，機械学習の分野でも近年取り組まれている．2節で複雑な物理現象のモデリングが困難であると述べたが，プロセス設計者と現場作業者はその物理現象を「想像しながら」条件出しをしているわけで，その想像を数値化する間接センシングおよびそこから取得される中間データが，物理現象を象徴しているほど，効率的な機械学習が実現できる．「よい」中間データを取得するためには,日本人の勘コツ・すり合わせ能力が欠かせない．

このように,本プロジェクトの「粉体プロセス・インフォマティクス」では，なるべくプロセスの物理現象を象徴する中間データの取得方法の開発と,それを利用した機械学習システムの開発が2つの開発ポイントである．

5. おわりに

本報告では，日本の強みである粉体成膜プロセス開発のさらなるハイスループット化のために，プロセス中の複雑物理現象を象徴する「中間データ」を間接センシングし，それを中間層に用いることで，効率的な機械学習システムを作り上げるコンセプト，データ駆動型「粉体プロセス・インフォマティクス」を紹介した．これまで，粉体プロセス開発をけん引してきた日本人の勘コツ・すり合わせの能力を最大限に活用して取得された中間データで，実験回数すなわち取得すべきデータ数を最小限にする日本流インフォマティクスが期待される．

参考文献

1）国立研究開発法人 新エネルギー・産業技術総合開発機構（NEDO）技術開発センターレポート Vol.2「機能性材料分野の技術戦略策定に向けて」2015年10月；https://www.nedo.go.jp/content/100762561.pdf

2）National Science and Technology Council, “Materials Genome Initiative for Global Competitiveness” June 2011；https://www.mgi.gov/sites/default/files/documents/materials_genome_initiative-final.pdf（https://www.mgi.gov/）.

3）R. Potyrailo, K. Rajan, K. Stoewe, I. Takeuchi, B. Chrisholm and H. Lam : “Combinatorial and High-Throughput Screening of Materials Libraries: Review of State of the Art”, ACS Comb. Sci., 13, 6（2011）pp.579–633.

4）国立研究開発法人 科学技術振興機構 未来社会創造事業「共通基盤」領域；https://www.jst.go.jp/mirai/jp/program/core/index.html

5）T.J. Sejnowski : “The Deep Learning Revolution”, 桐谷賢治，藤崎百合訳「ディープラーニング革命（日本語）」, Newton Press（2019）.

6）粉体工学会 編:粉体工学ハンドブック, 朝倉書店（2014）.

7）船津公人:「特集「マテリアルズ・インフォマティクス」の企画にあたって」, 日本化学会情報化学部会誌, 34, 4（2016）p.95.

8）Y, Chen et al. : Bayesian Optimization in AlphaGo, arXiv（2018）.