

先端材料開発に向けた，AI 先端計測技術の 多角的視点からの考察（第2回）

岡本和也* 杉山昌章** 武藤俊介***
青柳里果**** 富谷茂隆*****

第2回のはじめに

先端材料研究において重要な先端的な検査・計測・分析技術(本稿では「先端計測技術」と略す)において，AI/ML (Machine Learning：機械学習)技術の視点が大きく取り込まれようとしている．ここにはデータ駆動という新しい概念が含まれ，従来のような単なるハードウェア開発を主軸においた領域だけではなく，この思想を巧みに活用することで先端計測分野の装置開発や原理に関わる研究への新しい糸口になりえる．前回は，このAI先端計測技術に関わる背景と概要を述べ，さらに電子顕微鏡分野における「マルチモーダル計測」を取り上げた(前回3.1節)．これは，多角的な測定情報に対し異なる技術の組み合わせ・複合化によるデータの記述座標軸を増やし，かつ変数相互の相関関係を巧みに取り入れ，高効率かつ高精度に物理量を統合解析する概念である．さて，第2回はこれを時間軸に展開した「オペランド計測」について3.2節から解説を始める．次に，スペクトルにおけるピークの多様性を数値化する「情報エントロピー計測(3.3節)」について述べる．さらに，先端計測分野の産業動向を知的財産情報等の視座から捉えた社会科学思想についても言及する．

3.2 オペランド計測

「Operando」はラテン語の動名詞であり，Working, Oper-

atingの意味を持つ．計測・分析領域にオペランドという用語を最初に用いたスペインの触媒研究者 Bañares らは，「オペランド計測とは触媒動作中に分光学的評価と触媒活性測定を同時に行う手法である」と定義している⁽¹⁾．このようにオペランド計測は触媒研究から端を発したが，その後，電池や半導体デバイスなど他の領域に適応され，実際の使用環境下で反応やデバイスを動作させながら，変化する現象を観察することで測定対象と機能との相関を見出す手法として広がっている．しかし，実際の使用動作環境を分析装置内で再現することは容易ではなく，そのための技術開発課題は多々あり，その意味でもAI/MLを組み合わせた新たな取組みが必要な領域である．一方，Operando分析を包含する，より広い意味で*in-situ*分析という用語を用いられることもあるが，この文字通りの意味はon site, in position(本来の場所)であり，必ずしも時間の概念が十分に含まれていない点で区別できる．ここでは，「適切に材料・デバイスに光・電気などの外場を与え，その機能の動的変化を時間発展的に観察・計測する手法」をオペランド計測の定義とし，時間発展という視点から様々な事例を取り上げ，AI/ML技術の活用について今後の展望について述べる．以下では，取り扱う現象の時間スケールを短時間(フェムト秒から秒)と長時間(時間から年)とに大別して考える．

3.2.1 短時間の領域で考える現象

光・電子デバイスでは電子などのキャリア挙動，触媒や電池など化学デバイスでは化学反応，また，構造材料の場合は

* 山口大学大学院；教授・副研究科長(〒755-8611 宇部市常盤台2-16-1)，大阪大学エマージングサイエンスデザインR³センター；招聘教授

** 大阪大学大学院；特任教授 *** 名古屋大学未来材料・システム研究所；教授 **** 成蹊大学；教授 ***** ソニーグループ株式会社R&Dセンター；主幹研究員

Advanced Measurement and Analysis Systems Using AI/ML for Next Generation Materials Development—Multifaceted View 2—; Kazuya Okamoto*, Masaaki Sugiyama**, Shunsuke Muto***, Satoka Aoyagi****, Shigetaka Tomiya***** (*Graduate School of Innovation and Technology Management, Yamaguchi University, Ube. **Graduate School of Engineering, Osaka University, Suita. ***Institute of Materials and Systems for Sustainability, Nagoya University, Nagoya. ****Faculty of Science and Technology, Seikei University, Tokyo. *****R&D Center, Sony Group Corporation, Atsugi)

Keywords: measurement, metrology, electron microscopy, AI (artificial intelligence), machine learning, multimodal, operando, information entropy, innovation, intellectual property

2022年3月30日受理[doi:10.2320/materia.61.579]

瞬間的な破壊挙動などが短時間で生じる現象が対象となる。図1に半導体における光励起ダイナミクスの流れを示す。光を吸収してから様々な素過程を経て、おおよそフェムト秒(fs)からナノ秒(ns)の範囲で現象が完了する。フォトリソの重合反応の場合、その化学反応機構により異なるが、時間スケールはおおよそマイクロ秒(μ s)から秒(s)の範囲である。このような現象に対するオペランド計測では、目的とする機能が発現できるような環境を構築もしくは加工することが必須となる。例えば、実際のLiイオン二次電池構造のままでは、電池内部の活物質の充放電時における電子化学反応に伴う電子状態の変化をX線吸収微細構造(XAFS: X-ray Absorption Fine Structure)により把握することはできないため、図2(a)に示すような実電池と同様に充放電が進行する特殊な電気化学セルを工夫して作製する必要がある⁽²⁾。図2(b)にはこのセルにより得られたLiCoO₂のCo-K吸収端XANES(X-ray Absorption Near Edge Structure)スペクトルの充放電に伴う変化の実例を示す。オペランド計測の成否はこのように測定環境の構築やデバイス化設計・製作が鍵となる。超高速現象のオペランド計測手法として、レーザー光を用いたポンプ-プローブ法がよく用いられる。短いレーザーパルス光を試料に照射・励起することで非平衡状態を作り、プローブ光で光学定数の変化を検出するなど光物性評価によく用いられる。この手法を放射光計測に適用することで、光触媒反応における構造変化を100 ps という高い時間分解能での観察が可能となっている⁽³⁾。このように高速な現象を捉えるには、①現象の時間速度に応じた光源・プローブ、②短時間で大量な計測データに対応する保管・転送する技術、③制御のためのデータ読み出し技術、④効率的に実験を行う方法論などの多角的視点からの検討が望まれる。④においては、放

射光などの共用設備を利用する場合など、限られた使用時間内で目的を達成するために逐次ベイズ推定法を活用することにより、次に行うべき実験条件をリアルタイムに推挙してくるシステムの構築が必要となる。この超高速現象の計測はSEMやTEMの分野でも精力的に研究されている。図3にSEMへの応用であるがポンプ-プローブ法を用いた2次元イメージングの例としての走査超高速電子顕微鏡法(SUEM: Scanning Ultrafast Electron Microscopy)を模式的に示す。短パルスレーザーで試料を光励起すると同時に短パルス電子線をプローブとして試料上に走査し2次元電子像を得る。例えば、pn接合を有する半導体試料に電子線が照射されると内蔵電位(拡散電位)が発生し、p型表面とn型表面でコントラスト差が生じるため、SUEMでの2次元電子像コントラストの時間変化をトレースすることによりpn接合近傍のキャリア再結合などキャリアダイナミクスをピコ秒オーダーで時間変化を可視化できる⁽⁴⁾。

近年では3次元イメージング可視化のニーズが高まっている。鉄鋼材料を始めとする構造用金属材料には析出物・粒界といった微細構造が存在し、その結晶粒径や結晶方位に依存する材料の機械的性質はその多結晶材料の3次元構造に大きく依存する⁽⁵⁾。また、半導体デバイスでも2次元構造での微細化の限界からGAA(Gate All Around)など複雑な3次元トランジスタ構造が検討されており、3次元的な構造形状やその界面近傍における不純物元素分布の解析など、従来のドーパント分布解析とは異なる新たな3次元可視化の要求が強い。3次元微細構造の観察にはX線CT法(X-ray CT: X-ray Computed Tomography)や透過電子線トモグラフィ(TEMT: Transmission Electron MicroTomography)などの手法がある。X線CT法の空間分解能は1~2 μ m、TEMTのそれは1 nm程度であり、目的に応じて使い分けられる。但し、これらの手法は光源・検出器もしくは試料を回転させながら投影像を取得し時分割で再構成する手法であるため、投影像データセットの取得時間が時間分解能となる。現状では数秒(s)~数百秒(s)にとどまってお⁽⁶⁾⁽⁷⁾、より速い現象の変化を捉えるには不十分である。後述するが、試料回転時のデータ欠如部分や画像処理技術へのAI/MLの

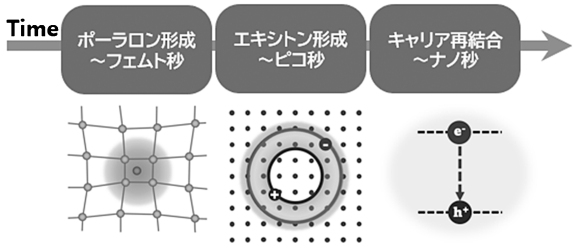


図1 半導体材料における光励起ダイナミクスの流れ。

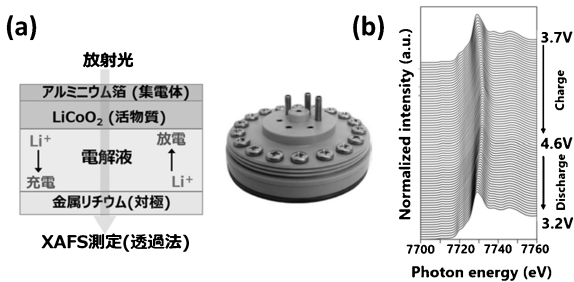


図2 (a) オペランド計測用評価電池セル。(b) LiCoO₂のCo-K吸収端XANESスペクトルの充放電に伴う変化。

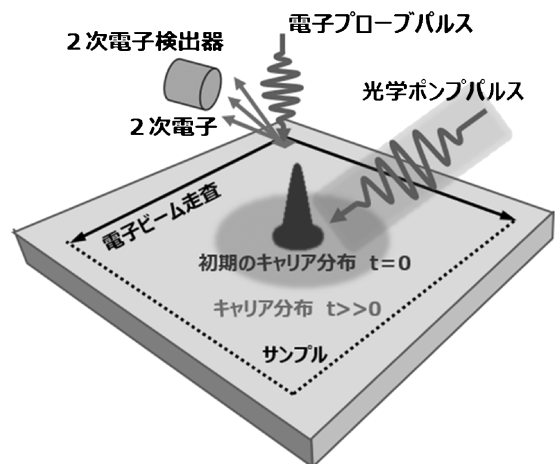


図3 走査超高速電子顕微鏡の像形成メカニズム。

活用が進む。超高速での現象解明には、上述の SUEM と共に UTEM (Ultrafast TEM) がこの10年間、精力的に世界中で研究されている。

3.2.2 長時間の領域で考える現象

先端材料やデバイスの劣化機構の解明や寿命の予測は、安心・安全なものづくりに必須である。このため、素材やデバイスの摩耗劣化や材料の疲労・腐食など何らかの特性劣化現象を捉えるには、(加速寿命試験を行う場合においても)長時間にわたる計測は必須となる。

半導体デバイスの例で考えると、LSI の基本構成要素である MOS トランジスタにおいてゲート酸化膜の絶縁破壊がしばしば問題となる。酸化膜の絶縁破壊強度よりも低い実使用電界強度で時間の経過とともに破壊が起こる現象は、経時的絶縁破壊 (TDDB : Time Dependent Dielectric Breakdown) 故障と呼ばれる。この TDDB のメカニズムとして、電流注入による電子トラップが発生・増殖し、電流パスが形成されるというパーコレーションモデルが広く知られており、ゲート酸化膜の膜厚が薄くなるにつれ TDDB 故障は顕在化する⁽⁸⁾。半導体微細化の進展に伴いゲート酸化膜の材料変革と薄層化は継続しているため、TDDB 故障は引き続き重要な信頼性課題である。例えば、薄膜化を回避するため High-k 材料など高誘電率の代替材料などが用いられるが、この場合、多量の電子トラップが内在する課題が生じている。このように TDDB 故障は、膜厚、膜密度、膜組成や不純物など複数の要因が複雑に絡み合う。これらの破壊現象に対しては、静的な分析・解析結果のみから正しく寿命・劣化予測を行うことは困難であるため、寿命加速試験を行いワイブル分布に従って予測する。しかしながら、複雑で複数の要因が絡みあうため単純な物理モデルでは予測できない場合も生じ、様々なモデルが提案されている。さらに近年、単一のモデルのみでは整合せず複数のモデルを組み合わせるとよという報告がなされている⁽⁹⁾。また、電子トラップ密度の時間変化を追いながら、複数モデルを融合するアンサンブル学習などの AI/ML モデルを活用した寿命予測が考えられる。

電池材料など機能部材をデバイスメーカーに提供する材料メーカーでは、デバイスに実装された状態で材料の信頼性評価を行う必要がある。それはデバイス特性との相関性を理解することで、より良い製品の提供につながるからである。例えば、サイクリックボルタンメトリー (CV : Cyclic Voltammetry) や交流インピーダンスなどのデバイス特性評価と X 線回折、XAFS 測定など物理解析を同時に行える評価デバイスを作製する。この評価デバイスによるオペランド分析から材料の劣化度を予測し、劣化モデルを導く。さらにデータを蓄積し、オペランド分析結果から電極材料の化学・物理構造、電極材料製造などの最適化条件を導き出すマルチタスク深層学習モデルを構築し、より良いデバイスの作製指針を見出すことができることが期待される。これらの一連の流れを概念図として図 4 に示す。

社会インフラを担う構造材料は多様な環境下で使用されるため、定期的な外観検査で劣化具合を判断し、保守・修繕時期を推定している。しかし、コスト削減には現地・現場に赴

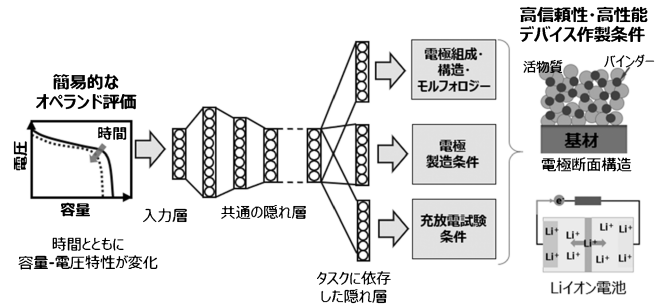


図 4 マルチタスク学習を用いた高信頼性・高機能電池デバイス作製条件探索法の概念図。

いて実施する外観検査の回数を減らしたい一方、安心・安全のためには検査頻度を減らすわけにはいかない。この相反に対峙する施策としての取り組みも、時間軸の長いオペランド計測と考えると興味深い。最近では、腐食データなどの蓄積や気象環境データまでも併せて AI/ML により統合化し、インフラ維持管理の効率化のプロジェクトが国土交通省などを中心に展開している。このような長期現象よりももう少し時間軸は短いものの、製品を上市する前のオペランド加速試験により何らかの特徴指標を抽出することも重要である。例えば、組織学的因子 (粒径、結晶方位変化など)、電気的因子 (電位など) や機械的因子 (残留応力、塑性歪) など複数のデータセットから、マルチモーダル学習など AI/ML 技術により、機械的特性に紐づいた劣化指標を導き出す研究もおこなわれている。ただそのためには種々の特徴因子と劣化指標との相関を見出すことがまず必要であり、最近では、微小クラックの形状から機械学習により破壊との相関を見出そうとする研究なども始まっている⁽¹⁰⁾。

また近年、文化財保護制度の見直しが行われるなど、文化財の劣化に対する対策にも社会的ニーズが高まっている。文化庁では木造建築物の劣化を AI で点検するシステムの試験運用を始めるという動きも始まっている。文化財の保存にはカビの発生が大敵である。微生物による損傷例などの研究に基づき、カビ臭のイオン移動度スペクトルやガスクロマトグラフィ質量分析 (GC/MS) データベースを構築し、真菌種同定ソフトウェアの開発に伴い未知の真菌種の特定などに役立てられている⁽¹¹⁾。現在では、温湿度や空気中の質 (CO₂ や有機物など) をリモートでモニタリングし、マルチモーダル学習によりカビの発生を予測することも検討されており、AI と IoT 技術を活用したリモートオペランド計測による劣化予測は文化財保護にも期待がかかる。

3.2.3 オペランド計測に必要な AI/ML 技術

上述の通り、オペランド計測に関し時間スケールを 2 つに大別し、具体例をあげながら今後の展望や課題などを述べてきた。特に、AI/ML 技術の活用が必要な事項を以下に要約する。

- (1) 逐次ベイズ推定法などにより、次に行うべき実験条件をリアルタイムに情報を抽出する技術。
- (2) 少ないデータを補間しつつ、物理的に有意な情報を抽出する技術。

- (3) アンサンブル学習法などの AI/ML 技術から、複数の物理モデルに対応した結果を予測する技術。
- (4) 簡易計測のみからマルチタスク学習など AI/ML 技術で最適化特性/条件を見出し、結果を予測する技術。
- (5) 複数の計測手法から得られるマルチモーダル学習などの AI/ML 技術から結果を予測する技術。
- (6) AI/ML 技術と IoT 技術を活用したリモート計測で結果を予測し、所望の状態を維持する技術。

上記の項目(1)から(3)は従来型の個別の研究により技術進化させていくことが可能である。特に項目(2)に関してはオペランド計測では多量のデータが得られるものの、その時々で得られる(信号強度、エネルギー分解能や空間分解能などの)情報量が静的な計測に比べて少ないことにある。それゆえ、少ないデータから多くの意味ある情報を復元し、抽出することが重要である。しかし、ここで注意すべきは完全なデータ(真値)からどこまでデータが削減されていたら、科学的に妥当な結果が担保されているかを見極めておくことである。項目(4)はデバイスメーカーと材料メーカーとの連携、さらには AI/ML 技術の専門家集団との連携、項目(5)は異なる先端計測技術の専門家の融合が必要であり、垣根を越えた研究連携の必要性が問われる。項目(6)のリモート計測の領域では、情報管理における社会的ルール構築も必要であり、真に未来社会を構築するという前提での大きな議論が必要である。このようにオペランド計測という本当に望まれる環境下での計測技術を深化させていくためには、その研究開発体制や情報交換の進め方においても新しい挑戦的仕組みと取り組んでいかないといけない。このような議論が学会や産官学連携の場でも必要になるであろう。

3.3 情報エントロピー計測

20世紀中頃に Claude Shannon が情報量を表す指標として、確率に基づいて計算する情報エントロピー(Shannon エントロピー)⁽¹²⁾を提唱してから70年以上経った現在、情報エントロピーの応用先は様々に広がっている。統計力学と情報科学に基づいた量子情報科学分野は、物理学と情報科学の両分野に刺激を与えている。特に、シグナル検出が関係する計測分野にとって、量子情報科学は今後大きな影響を及ぼす可能性がある。ただし、本稿では現状の多くの計測におけるシグナル検出は古典物理学で説明されるものとして、量子情報科学に深くは触れない。本節で用いる情報エントロピーは、次式(1)の Shannon エントロピー S を指す。

$$S = - \sum_{i=1}^n p_i \log p_i \quad (1)$$

ここで、 p は各事象が起こる確率、 n は事象の総数を示す。情報エントロピーは、例えば文章の情報量を使用した文字の多様性で表す⁽¹²⁾ことができるが、同様の考え方をスペクトルデータやイメージデータに適用することができる⁽¹³⁾⁽¹⁴⁾。具体的には、検出される各情報を 0 から 1 の間に規格化すれば、式(1)を用い様々なデータの情報エントロピーが計算できる。例えば、スペクトル上のピークの強度が出現確率に対応すると考えると、情報エントロピーを用い

て、スペクトルにおけるピークの多様性を数値化できる。

3.3.1 情報エントロピーのスペクトルへの応用

スペクトルデータを数値化する際の代表的な課題として、ピークとノイズの区別、重なり合ったピークの分離などを実施するピーク認識が挙げられる。実際の計測では、ノイズと明確に区別できる高強度で形状の整ったピークが必ずしも得られるわけではない。ここではピーク認識手法については論じず、何らかの方法でスペクトル上のはほぼ全てのピーク強度を数値化したものと仮定し、情報エントロピーがどのように応用できるかを説明する。

データ例として飛行時間型二次イオン質量分析(TOF-SIMS)を取り上げる。TOF-SIMS では一般に質量電荷比 m/z : 0~2000 程度の範囲でスペクトルが得られ、一般的な有機物の分析では1000本程度の質量ピークがスペクトル上に存在する。

情報エントロピーにおける確率が TOF-SIMS スペクトルの何にあたるか考えると、値として 0~1 の間に収まるデータと捉えればピーク強度を 1 以下に規格化すれば良いことになる。最小値を 0 に最大値を 1 とする規格化(正規化, min-max scaling, min-max normalization)を適応させれば、ピーク強度を確率 p として情報エントロピーは式(1)で計算できる。もしくは、各ピークのスペクトル上での存在確率を考えると物理的意味が捉えやすいかもしれない。その場合、全てのピーク強度の合計(総二次イオン強度)で割ったピーク強度(相対強度)を確率として、式(1)に代入することができる。どちらの求め方でも、スペクトル上に一本だけピークが存在する場合は情報エントロピー $S = 1 * \log 1 = 0$ と最小値になる。スペクトル上に多数のピークがほぼ同強度で存在する場合、情報エントロピーは最大値に近い。例えばスペクトル上に10本のみピークが存在し、全てのピーク強度が 0.1 であるとし、この強度をそのまま p として計算すると、情報エントロピー S は式(2)で求められる。

$$S = (0.1 \log 0.1) \times 10 = 1 \quad (2)$$

したがって、スペクトルの情報エントロピーはスペクトル上のピークの多様性を表す指標と考えることもできる。ピークの多様性が高いということは一般的には測定部の物質の多様性(物質の種類および配向・混合状態などの種類が多様であること)が高いと考えられる。つまり、情報エントロピーは異なる物質が接する界面や多くの物質が混合した部分で高い値を示す傾向があると言える。モデル試料データとして、ガスクラスターイオンビーム(GCIB)を用いた TOF-SIMS 測定におけるマトリックス効果について評価した例を示す。これは2014年の VAMAS プロジェクト⁽¹⁵⁾で用いられた 2 種類の有機物多層膜の TOF-SIMS データ⁽¹⁶⁾である。マトリックス効果は二次イオン質量分析(SIMS)の場合、共存物質などの影響によって二次イオン収率が変化することを指し、マトリックス効果の影響が大きいと二次イオン強度が由来する物質濃度に応答せずに増大したり抑制されたりするため定量分析が難しくなる。マトリックス効果の程度は物質の組み合わせや試料の種類などによって大きく異なるため、未知物質も含む多成分系の測定が多い TOF-SIMS では大きな課題の

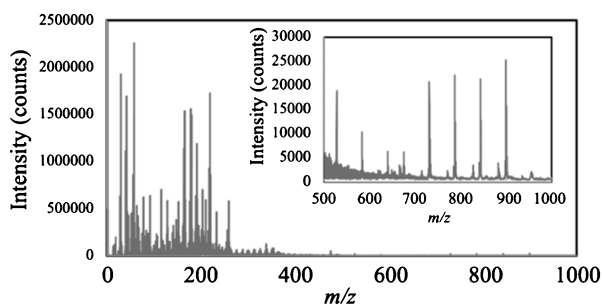


図5 有機物積層試料 (*Irganox 1010 & **Fmoc-PFLPA)の正二次イオンスペクトルの例.

*Irganox1010: Pentaerythritol tetrakis (3-(3,5-di-tert-butyl-4-hydroxyphenyl) propionate)

**Fmoc-PFLPA: Fmoc-pentafluoro-L-phenylalanine

一つである。ここで紹介するデータは、スパッタリングにArクラスターイオンビーム、分析にBiクラスターイオンビームを用い、有機物の深さ方向分析で推奨されているように分析領域より十分に大きなスパッタ領域を確保した。GCIBによるスパッタリングで試料の最表面層数nm程度を除去したのち、Biクラスターイオンで測定しTOF-SIMSスペクトルと二次イオンイメージデータを取得する過程を試料の最表面から基板に到達するまで繰り返した。TOF-SIMSスペクトルの例⁽¹⁶⁾を図5に示す。試料に含まれる物質が既知であり、その物質から発生する質量ピークについての知見がある場合、注目するピークの強度を各深さについてプロットすることで(図6(b)),各深さでの試料の状態が分かる。しかし、未知試料や注目すべき質量ピークに関する知見がない場合、スペクトルを解釈し重要な質量ピークを探さなければならない。未知試料のTOF-SIMSスペクトルの解釈は一般に容易ではないため、スペクトル解釈を待たずにスペクトルの情報エントロピーから試料の状態の概要を知ることができれば解析にとって有用である。図6(a)に示すように、スペクトルの情報エントロピーの値は最表面層(第一層:図6(a)A層)と第二層(図6(a)B層)の間、および第二層と第三層(図6(a)C層)の間のような純成分層の界面では上部層と下部層の物質両方に由来する質量ピークが検出されスペクトル上のピークの多様性が高まるため、結果として情報エントロピー値が上昇する⁽¹³⁾。また、有機物膜と基板との界面(スパッタ時間180~190s近傍)で情報エントロピー値の上昇が顕著に現れている。ここでは有機物層由来とSi基板由来の質量ピークが検出されることに加え、Si基板上の有機物膜が最薄となることで基板の影響を受け有機物由来の質量ピーク強度が増大する。すなわちスペクトル上でほとんど無視されていた低強度ピークも十分な影響力を示す強度に増大することにより、結果としてピーク多様性がより増大したと考えられる⁽¹³⁾。さらにスパッタリングが進むと有機物膜が完全に除去されSi基板の露出する部分が増え、質量ピークの多様性が減少し最終的にはSi基板由来のみの質量ピークだけ検出され情報エントロピーは最小値となる。このように、各深さ(1次元データ)でのスペクトルの情報エントロピーを求める

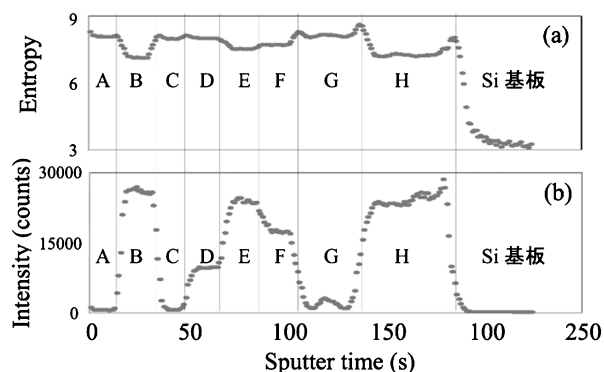


図6 (a) Fmoc-PFLPA (Fmoc)と Irg1010(Irg)混合試料における正二次イオンスペクトル情報エントロピー(縦軸)の各深さ(スパッタ時間)の関係, (b) Irganox1010由来の二次イオン m/z 163(縦軸が二次イオン強度)の深さ方向プロファイル.

A: Fmoc 100%, B: Irg 100%, C: Fmoc-PFLA 100%, D: Fmoc : Irg = 20 : 80, E: Fmoc : Irg = 80 : 20, F: Fmoc : Irg = 50 : 50, G: Fmoc 100% (中央に3nmのIrg層), H: Irg 100%(中央に3nmのFmoc層).

ことにより、注目すべき二次イオンが分からない試料についても、物質の変化や混合状態変化について概要が得られることになる。

3.3.2 未知試料からのスペクトルへの応用

化学イメージングが得られる計測データでは、スペクトル上のそれぞれのピークに対応する分布図がある。一般的には、注目する物質に関連するピークについて分布図を描くことによって解析を進めるが、注目するピークが分からない場合、スペクトル上の全てのピークの分布図を確認するのは現実的ではない。多変量解析やAI/MLによる解析によって重要なピーク候補が示されることも多いが、そのような解析結果を得るためには時間と労力が求められる。そこで、データの概要が分かる解析法があれば未知試料についてどのような解析を進めるべきかの指標が得られ、数値解析をする際にも有用である。スペクトルの情報エントロピーはスペクトル上の全ての質量ピーク情報に基づいて試料の概要を示す手法であり、このような目的には有効と考えられる。実際に、この有機物多層膜試料の断面(2次元データ)について、各ピクセルでのスペクトルの情報エントロピーを計算すると、図7(a)に示すように各位置におけるスペクトルの情報エントロピー値を用いたイメージングが可能である。スパッタリング時間は試料の深さに対応し、この試料の場合は情報エントロピー値が上昇するスパッタリング時間(破線で示す)がほぼ界面に対応している。情報エントロピーは物質の変化によって値が変化するため総二次イオン像(図7(b))よりも情報エントロピー像の方が物質変化が分かりやすい。総二次イオン像と情報エントロピー像の両方を比較することにより、試料の概要が理解しやすくなることが期待される。

このように、数値化した測定データ(デジタルデータ)は測定手法が示す物理・化学信号だけではなく、データそのものが持つ情報量を手掛かりに解析することが可能であり、さらにそのように得られたデジタルデータの情報から物理化学的

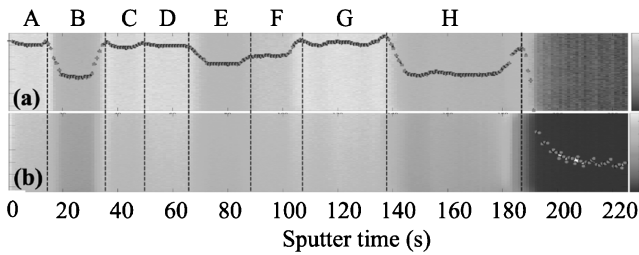


図7 Fmoc-PFPLA と Irg1010混合試料における断面図。(a) 各ピクセルにおける正二次イオンスペクトルの情報エントロピー強度図(図中のプロットは数値強度の変化を示す), (b) 総二次イオン像(正二次イオン)

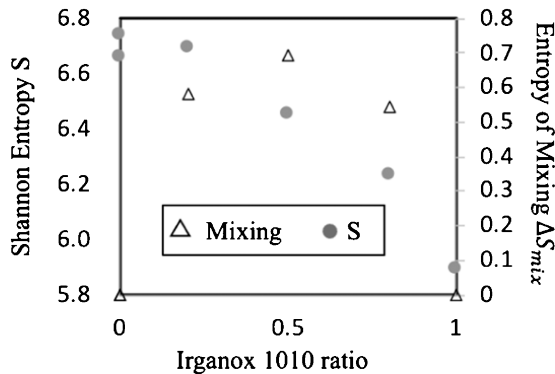


図8 Fmoc-PFPLA と Irg1010混合試料 TOF-SIMS 正二次イオンスペクトルの情報エントロピー S と混合のエントロピー ΔS_{mix} の関係。

な知見も得られる。例えば、図8にモル数に基づいて求めた混合のエントロピーと TOF-SIMS スペクトルの情報エントロピーの関係を示す。同図の混合のエントロピーは式(3)で求めた。

$$\Delta S_{mix} = -n_A R \ln \frac{n_A}{n_A + n_B} - n_B R \ln \frac{n_B}{n_A + n_B} \quad (3)$$

n_A と n_B はそれぞれ物質 A と B のモル数を示し、 R は気体定数である。ただし、混合によって体積変化および熱変化がないものとして、二種類の物質 A と B の混合によるエントロピー変化を求めた。スペクトル上の質量ピークは、試料中の物質の混合状態を反映しているが、各物質の濃度に線形応答するとは限らない。二種類の物質の混合比とピークパターンが線形応答する場合は、スペクトルから求めた情報エントロピーは、二種類の物質が等分ずつ存在する(濃度0.5)で最大値を取るはずである。しかし、マトリックス効果によって TOF-SIMS データの濃度応答性が崩れたため、スペクトルから求めた情報エントロピーの最大値は Irganox 1010 低濃度域で得られている。一方で混合のエントロピーは Irganox 1010 と Fmoc-PFLPA が同濃度存在する0.5が最大値となる。マトリックス効果の影響が小さい TOF-SIMS データの場合は実際に情報エントロピーの最大値は混合のエントロピーと等しく0.5に近づく⁽¹³⁾。

3.3.3 熱・統計力学でのエントロピーとの関連

本節の最後に、情報エントロピーと熱・統計力学のエントロピーの関連について簡単に述べる。情報エントロピーと

熱・統計力学のエントロピー S は異なるものではあるが、両方を考慮する重要性は以前から指摘⁽¹⁷⁾⁽¹⁹⁾されていた。熱力学第二法則 $\Delta S \geq 0$ に情報量 I を含めて $(\Delta S + \Delta I) \geq 0$ に拡張した場合、情報の変化量 ΔI は情報エントロピーの減少分 $-\Delta S_{info}$ と等しくなるため、 $\Delta I = -\Delta S_{info}$ であり、

$$\Delta S - \Delta S_{info} \geq 0 \quad (4)$$

$$\therefore \Delta S \geq \Delta S_{info} \quad (5)$$

となる。これは物理系を測定することにより情報を ΔI 得ると、それ以上に測定装置も含めた全体のエントロピーが増加することを意味する⁽¹⁷⁾⁽¹⁸⁾。ただし、情報エントロピーの変化量は熱・統計力学のエントロピーよりもはるかに小さく、一般に無視できる。現在では、先端計測機器の技術が発展し1原子・1分子(場合によってはそれ以下)の計測が可能になり、極めて高い空間分解能で物質の物理化学情報(結晶構造・化学構造など)を複数の視点で得られるようになった。さらに、このような高度な計測手法から得られた複雑な計測データを AI/ML 技術によってより高度に数値解析できるようになった現在、情報量と物理量の関係を改めて見直すことに意義があると考えられる。実際に、2010年に「マクスウェルの悪魔」と同様の状況の実験的再現⁽¹⁹⁾が報告され、フィードバックする情報をシラードのエンジン(分子の存在を観測し、その結果に応じて操作を変えることでエネルギーを取り出すこと)のように表現することにより、熱・統計力学のエントロピーに変換できることが明確に示された⁽²⁰⁾。情報エントロピーが熱・統計力学のエントロピーに変換できる系があることから、条件を整えばスペクトルの情報エントロピーから熱・統計力学のエントロピーを導ける可能性もあることが分かる。つまり、試料の化学情報が得られるスペクトルを測定することにより、試料のエントロピー、自由エネルギーなどが予測できる可能性がある。この点については慎重な検討が必要であるが、先端計測技術から得られる情報が大いに広がることが期待できる。

4. AI 先端計測の社会科学的考察

本章では先端計測技術の産業を俯瞰し、知的財産情報を中心にした社会科学的な考察を行う。各産業における市場の競争状態(集中度)を表す指標の一つとして、寡占度指数であるハーフィンダール・ハーシュマン指数 (HHI : Herfindahl-Hirschman Index)があり、式(6)のように当該産業に属する全ての企業の市場占有率の2乗和で定義される⁽²¹⁾。

$$HHI = \sum_{i=1}^N S_i^2 \quad (6)$$

ここで、 S_i は市場占有率、 N は当該産業の企業数である。本稿の対象である先端計測分野18領域(35項目)の国内市場⁽²²⁾における HHI の年次推移を1992年から30年間を計算した結果、分野により年次変動は見られるものの、いずれの機器分野においても $HHI > 2500$ にあり⁽²³⁾、高位寡占市場⁽²¹⁾に該当することがわかった。つまり、どの先端計測機器も平均30%程度以上の市場占有率を有する数社独占で市場は形成されており、この前提をもとに考察を展開する。

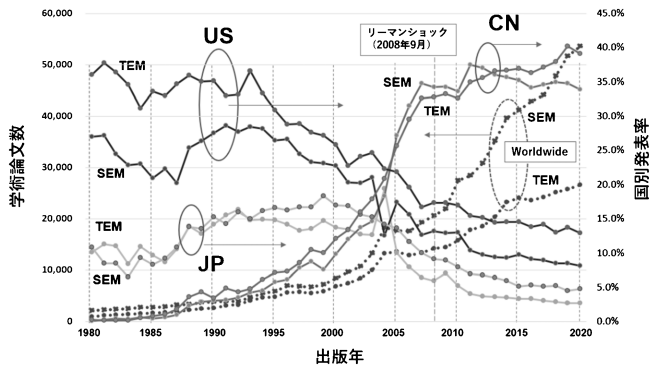


図9 電子顕微鏡に関する学術論文数(破線)と国別発表率(実線)の推移。(US:米国, CN:中国, JP:日本)。

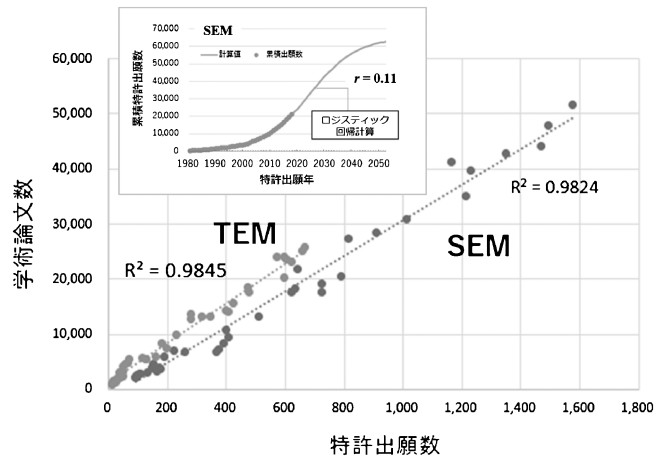


図10 特許出願数と論文数との相関性およびSEMの累積出願数の特許出願数年次推移。

4.1 学術性と産業性に関する考察

学術的価値を示す代表的指標が学術論文である。一方、産業的価値の当該指標として知的財産権の一つである特許権の行使に至る特許を本稿では提起する。ここでは、その出願状況から技術と産業へ視野を広げ、代表的な先端計測技術および機器の動向を把握すると共に、AI/ML活用との親和性を考察した。誌面の都合上、委細は省略するが、特許出願数から見る研究開発投資効率⁽²⁴⁾は先端計測産業全般として低下の傾向にある。その中で調査研究⁽²⁵⁾の結果として得られた日本での産業競争力の強い分野の一つはSEMであり、これを中心にTEMや他の技術との比較を行い学術性と産業性の双方から考察する。

まず、学術性について、図9に両者の全世界の学術論文の年次数(破線: worldwide)及び国別発表率(実線)の推移を示す。学術論文数(破線)は両者とも増加の傾向にあるが、その伸長率は2000年ごろから急伸するものの、2005年以降、SEMとTEMの年平均成長率(CAGR: Compound Annual Growth Rate)には差異が拡大している。また、国別の傾向からは中国の急伸と2008年のリーマン・ショック時期からは発表率が一定化する傾向が確認されるものの、論文数は35%以上の伸長を示している。日本と米国はSEM、TEMともに発表率は減少しているが、米国に比して日本の減少率は小さく国際競争力の担保に繋がっている。これは半導体技術への先端計測技術適用の重要性とSEMやTEMがその主軸であることに起因するものと考えられる。しかし、AI/ML活用の概念の中で今一度、日本の先端計測技術の底上げが必要と考える。

次に、産業性について特許出願動向から技術の掘り下げを試みる。各機器の特許出願動向(全世界)について、1980年からの代表的な先端計測技術の全世界の累積出願総数(INPADOC: International Patent Documentation Center Family)の時系列変化からいずれの機器分野も増加傾向にあり、産業的価値の意義は深化している。ここでは、SEMおよびTEM領域における過去40年の特許出願動向および学術論文発表から当該技術を確認した。なお、各装置本体(物)に絡む内容とその方法を抽出できるようにクエリを工夫した。

SEMについては、1980年から2020年現在までの全世界ファミリーで年間出願数およびElsevier・Scopusデータベースに抄録された当該領域の単年論文数を抽出し、両者の相関性を図10に示す。特徴的なのは1980年において高い論文数に比して特許出願数は少なく(SEM: 90件, TEM: 13件)、電子顕微鏡分野は学術指向で推進され、その後に産業化に至った分野と考えられる(2018年国内市場はSEM: 131億円, TEM: 72億円)。一方、論文数と特許数には強い相関性(決定変数 $R^2 > 0.98$)を有し、両者には類似した伸長傾向を有している点が見出された。さらに、両者の特許出願数の増加傾向を確認してみる。ここでは、全世界のINPADOCファミリーでの累積出願数およびロジスティック回帰にて近似計算した。1980年には90件の出願に留まっていたが、2015年には1,577件と20倍弱に至り、累積出願数から計算される成長係数 r は0.11とゆるやかな延伸にあると試算され、特許出願の視点からはSEM領域は産業分野では「萌芽から成長期」といえる。同様にTEM領域についても、1980年には13件の出願が2015年には572件と40倍強になっており、SEMよりはやや急峻な成長係数0.14と延伸にあり「成長期」と判断された(比較として、スマートフォンの成長係数 r は0.68である)。すなわち、電子顕微鏡分野の歴史は古いものの、新しい分野創成が可能な期待領域と考えられる。

4.2 AI関連特許についての考察

AI関連の国内特許出願状況について、特許識別子FI(File Index)のサブクラスG06Nで見ると2015年から2019年までの4年間はCAGR~26.1%であり急伸性が理解できる。しかし、AI/MLを活用した電子顕微鏡分野については、2020年現在でも出願率は0.1%以下と低く萌芽期にあることが判明した。その出願内容について全世界での特許出願データより抽出されたINPADOCファミリー151件について、IPC(International Patent Classification)サブグループを用い、共起グラフからの考察を試みた。結果を図11に示す。自然言語処理の分野では、任意の文書や文においてある文字列とある文字列が同時に出現することを「共起」という。ここでは、各特許に付与されたIPCを文字列と認識し、その共起性を見

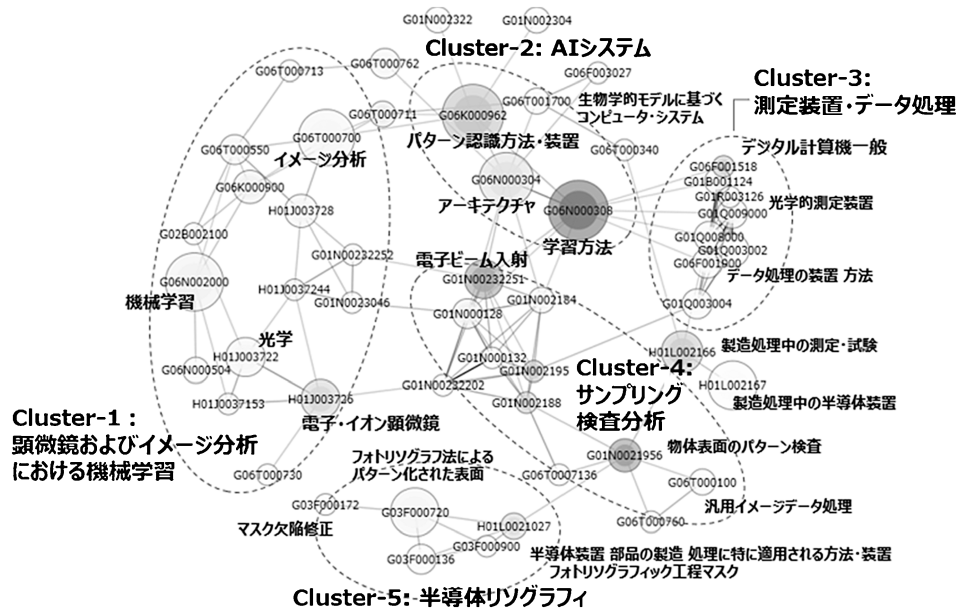


図11 AIに絡むSEM(装置及び方法)：特許出願のIPCコード共起グラフ。

出すことで特許技術の関連性を視覚化した。同図で、○の大きさは出現頻度、○の濃度が高いほど他のクラスタとの媒介中心性が高いことを示す。なお、出現頻度が多く媒介中心性の高いIPCのみその内容を記述した。ここでは、5つのクラスタ(Cluster)にわけられた。さらに、第1回で述べたように、AI先端計測を2つの分類、すなわち、分類(I)：機器にAIを活用し機器の性能・Usabilityを高める自律型、分類(II)：単体もしくはつながる様々な機器から得られた情報にAIを有効活用し新しい領域を発見する複合型、を含めて各クラスタの考察を深めた。

Cluster-1：顕微鏡およびイメージ分析に機械学習を適用するクラスタ：分類(II)が主体。

Cluster-2：AIシステムに関する：分類(I)が主体。AI技術の基本となるIPC：識別子G06Nであり、Neural Networkを主軸にCluster1とつながる。

Cluster-3：測定装置・データ処理：AIシステムにつながる測定装置に関する。分類(II)が主体。

Cluster-4：サンプリング・検査分析に関する。分類(II)が主体。

Cluster-5：半導体リソグラフィに関する：この分野のマスク検査、パターン検査等に繋がっている。分類(II)が主体。先端材料向けの適用ではなく半導体の検査計測に関わる出願が主体である。特許出願人を見ると、蘭ASML、米AMD、東京エレクトロン、米Intel、米KLAなど半導体デバイス、半導体製造装置および検査装置の各製造業が圧倒的である。結論として、SEMについては半導体製造における測定検査への応用が圧倒的に多く(いわゆる、Virtual Metrology領域)、得られたデータへ積極的にAI/MLを適用する分類(II)が2021年現在では主流である。最近では、セマンティック分類(画像の意味的分類)に関わるもの、インフォマティクス応用も含まれる。また、装置の自動校正等の分類(I)に関する出願も確認された。

4.3 AI先端計測におけるデータの取扱い

AI先端計測におけるデータの取扱いは機微な状況にある。先述のとおり、AI/ML活用のためには優れた基礎データが必要であり、数多く質の高いデータを有する日本はAIの進化において国際的競争優位に立つ格好の機会を得ている。そこでは、データの価値とあり方を再認識すべきである。データの共有化(Data Sharing)においては、Open/Closed戦略のありかたをグローバルに進めるべきであり、研究であっても戦略立案に絡む技術マーケティング手法の導入が必要であろう。以下、3点を提起する。

(1) データオープン化の困難性とその対応

一般的に企業の分析データがオープンになれば、クラウドデータ化の時代においては情報拡散による機会損失の恐れが指摘される。オープン化による価値(特に経済的、産業競争力の視点)向上について、具体的なイメージを示しつつ、期待効果とリスクの両側面から検討される場が真に設定されることが重要である。

(2) オープン化の可能事例

マテリアルズ・インフォマティクス(MI)は材料開発効率をあげる上で重要で、その展開においては膨大なデータベースが必要になる。公知の論文や特許情報から、MI戦略として基本的なデータベースが国内では物質・材料研究機構を中心に集約されているが、産業界単位、もしくは大企業の場合は企業単位で、何らかの形で分析データにおけるデータベースの保有活動が行われているであろう。それらの相互活用で新たな競争が起きることが想定される。

(3) オープン化に向けた提言

様々な分野において、商品上市までのTAT(Turn Around Time)短縮と一層の品質向上に向け、先端計測と制御が一体となった判断の高速化(自律化・インテグレーション)が求められる時代が来るときには、オープン化戦略は必

須と思われる。ユーザーと装置製造業との個別技術開発は非開示と考えられるが、顧客製品に依存しない重要データを含め、相互に発展できるモデル(本質的に相互に付加価値が低く非効率な領域を共通化することにより効率化・一般化できるモデルを含む共通言語的なもの)およびインフラを整備していくことが出来ればオープン化の可能性があり、それにより今後その戦略において利益を得てくる企業群が現れるであろう。また、共有化することで価値が高まるデータであれば、その新たな仕組みが求められる。そのために暗号化や一部情報のブラックボックス化のような、解析は可能だが個々のデータ内容が判らないような仕組みづくりが出来れば良い。例えば、「暗号化したま AI/ML を行うシステム」が構築されれば、機密保持しつつ AI/ML 結果の恩恵を享受できるという点で、有望な仕組みとなるであろう。

5. 総 論

本解説では、社会基盤を支える先端材料研究において鍵となる、先端計測分野に焦点をおき、AI 活用による技術と産業動向について解説した。研究開発の進め方そのものに革新を迫る第3次 AI ブームにおいて、各種データからの情報抽出によるインフォマティクス研究が活況にある。本稿では、AI/ML の歴史から紐解き、先端計測技術への適用性に踏み込んだ。ここでは、データ駆動という新しい概念を含んでおり、従来のような単なるハードウェア開発やソフトウェア開発を主軸においた領域だけではなく、この概念を巧みに活用することで分析装置間の複合化が進むだけでなく、装置開発や原理に関わる研究においても新しい糸口になり得る可能性を秘めている。その具体例として、電子顕微鏡のマルチモーダル計測、時間軸を加味したオペランド計測、二次イオン質量分析を主体とした情報エントロピー計測の3点に注目した。上記、マルチモーダル計測処理に関わり、多種多様な種類の膨大なデータに対し多角的にアプローチしつつ統一的な解を導くために必要とされる工夫や仕組みについて、次の3点が考察される。第一が、共通座標軸の設定(3.1節で述べた STEM スペクトラムイメージデータでは、電子ビーム走査の各標本点で同時に異種分光データを収集するため空間座標が共通な多次元データ構造となる)、第二がデータフォーマットの共通化である。第三として、スキームの構築段階での各スクリーニングステップにおける各計測手法の専門家、情報・材料科学者の洞察による総合的判断が重要であることを掲げる。そのためには、各分野の専門家が混然一体となり議論し正しい解釈を進めるといった場の形成が求められる。分析技術・機器が産業的に高位寡占状態にある今、上記 AI 活用計測の可能性を自然科学および社会科学の双方の広い視野まで拡張し考察した。社会科学思想においては、産業の固定化を見出し日本の産業競争力の高い SEM に着目した。学術的価値として学術論文、産業的価値として出願特許を指標として提起し、両者に高い相関性を見出した。また、同分野の AI 活用は圧倒的に半導体産業への適用性が高いものの、2020年時点においても特許出願率は0.1%にも満たない萌芽

期であり、新しい産業創成の機会が期待され、いくつかの興味深い事案の出願特許も見出された。

今後はデータ駆動の特性を活かした新しい領域への期待がかかる。ESG(環境, 社会, ガバナンス)・SDGs 投資が加速する中、これは一企業で成り立つものでなく、無駄を排除する全体最適の視点でのサーキュラーエコノミーの時代において、AI 先端計測においてもその重要性から全く異次元の Startup 企業の創発など新たな風が起きる気配を感じる。

最後に、DX 人材育成が加速している今、AI 先端計測においても、情報系技術と分析系技術および社会全体を俯瞰できる人材が求められる。その基盤は多種多様な断片的な情報をその人なりに組み合わせて(統合)、自分の次の行動や他の人に提案するための新しい知識を作り上げていくリテラシー、すなわち「トランスリテラシー人材の育成」にあることを強調し、産官学連携のさらなる重要性を提起し結論としたい。

本解説にあたり、日本学術振興会研究開発専門委員会「自律型・複合型 AI 先端計測の新しい価値創造」の委員各位に深く感謝する。

文 献

- (1) M. A. Bañares: *Catalysis Today* **100**, **71**(2005), 71-77.
- (2) 後藤習志, 稲葉雄大, 越谷直樹, 細井 慎, 工藤喜弘: サンビーム年報・成果集, **6**(2016), 59-61.
- (3) S. Nozawa, *et al.*: *J. Synchrotron Rad.*, **14**(2007), 313-319.
- (4) B. Liao and E. Najafi: *Materials Today*, **2**(2017), 46-53.
- (5) 波多 聡, 趙 一方, 井原史朗, 齊藤 光, 光原昌寿, 村山光宏: *まてりあ*, **61**(2022), 84-88.
- (6) 鳴海大翔, 河原崎琢也, 加藤藤一, 森下浩平, 安田秀幸: *軽金属*, **70**(2020), 339-346.
- (7) W. Albrecht and S. Bals: *J. Phys. Chem. C*, **124**(2020), 27276-27286.
- (8) J. W. McPherson: *Microelectron. Reliability*, **52**(2012), 1753-1760.
- (9) K. Park, K. Park, S. Im, SE. Hong, K. Son and J. Jeon: *Electronics*, **8**(2019), 942.
- (10) R. Meya, C. F. Kusche, C. Lobbé, T. Al-Samman, S. Korte-Kerzeland and A. E. Tekkaya: *Metals*, **9**(2019), 319.
- (11) T. Takeuchi, T. Kimura, H. Tanaka, S. Kaneko, S. Ichii, M. Kiuchi and T. Suzuki: *Surf. Interface Anal.*, **44**(2012), 694-698.
- (12) クロード・E・シャノン, ワレン・ウィーバー著, 植松友彦訳: *通信の数学的理論*, ちくま学芸文庫(1947).
- (13) S. Aoyagi, K. Mizomichi, K. Kamochi and A. Miisho: *Surf. Interface Anal.*, (in press). <https://doi.org/10.1002/sia.7047>.
- (14) R. M. T. Mадiona, D. L. J. Alexander, D. A. Winkler, B. W. Muir and P. J. Pigram: *Appl. Surf. Sci.*, **493**(2019), 1067-1074.
- (15) A. G. Shard, *et al.*: *J. Phys. Chem. B*, **119**(2015), 10784-97.
- (16) K. Takahashi, S. Aoyagi and T. Kawashima: *Surf. Interface Anal.*, **49**(2017), 721-727.
- (17) L. Brillouin: *Science and Information Theory: Second Edition*, Dover Publications, (2013, 初版1962年).
- (18) 甘利俊一: *情報理論*, ちくま学芸文庫, (2011, 初出1970年).
- (19) S. Toyabe, T. Sagawa, M. Ueda, E. Muneyukil and M. Sano: *Nature Physics*, **6**(2010), 988-992.
- (20) S. Ito and T. Sagawa: *Phys. Rev. Lett.*, **111**(2013), 180603.
- (21) <https://www.justice.gov/atr/horizontal-merger-guidelines-08192010>
- (22) 科学機器マーケット総覧2020(アールアンドディ, 2020.3.20).
- (23) K. Okamoto, M. Sugiyama and S. Mabu: *e-J. Surface Sci. Nanotechnol.*, **18**(2020), 214-222.
- (24) 岡本和也, 福代和宏: *MOT 研究開発マネジメント入門*, 朝倉書店, (2020), 103.
- (25) <https://jsps-ai-imat.mot.yamaguchi-u.ac.jp/>