

技術報告

# マテリアルキュレーション<sup>®</sup>支援システムの開発

吉武 道子,<sup>1,\*</sup> 佐藤 文孝,<sup>2</sup> 河野 洋行<sup>2</sup>

<sup>1</sup> 物質・材料研究機構, 〒305-0003 茨城県つくば市並木1-1

<sup>2</sup> 富士通総研, 〒105-0022 東京都港区海岸1-16-1 ニューピア竹芝サウスタワー

\* [yoshitake.michiko@nims.go.jp](mailto:yoshitake.michiko@nims.go.jp)

(2020年3月18日受理; 2020年4月30日掲載決定)

著者は、一見別々の事象を科学法則に則って結び付け、既存のデータを駆使して物性値の予測や材料探索指針の発見を行おうという、「マテリアルキュレーション<sup>®</sup>」を提唱している。これは、機械学習的手法を用いるのに十分な実験値や計算値を得るのが困難という材料特有の問題を克服するものである。分野横断的に科学法則を利用するために、物性間の関係性をネットワーク型の知識データベースとし、物性間の関係性を探索できる、マテリアルキュレーション<sup>®</sup>支援システムのプロトタイプ開発を行った。物性間の関係性の例や関係性データベースとその探索の概念、システム構築について報告する。

## Developing a Materials Curation<sup>®</sup> Support System

Michiko Yoshitake,<sup>1,\*</sup> Fumitaka Sato<sup>2</sup> and Hiroyuki Kawano<sup>2</sup>

<sup>1</sup> National Institute for Materials Science, 1-1, Namiki, Tsukuba, Ibaraki 305-0003 Japan

<sup>2</sup> Fujitsu Research Institute, 1-16-1, Kaigan, Minato-ku, Tokyo 105-0022 Japan<sup>3</sup> Aichi Center for Industry and Science Technology

\* [yoshitake.michiko@nims.go.jp](mailto:yoshitake.michiko@nims.go.jp)

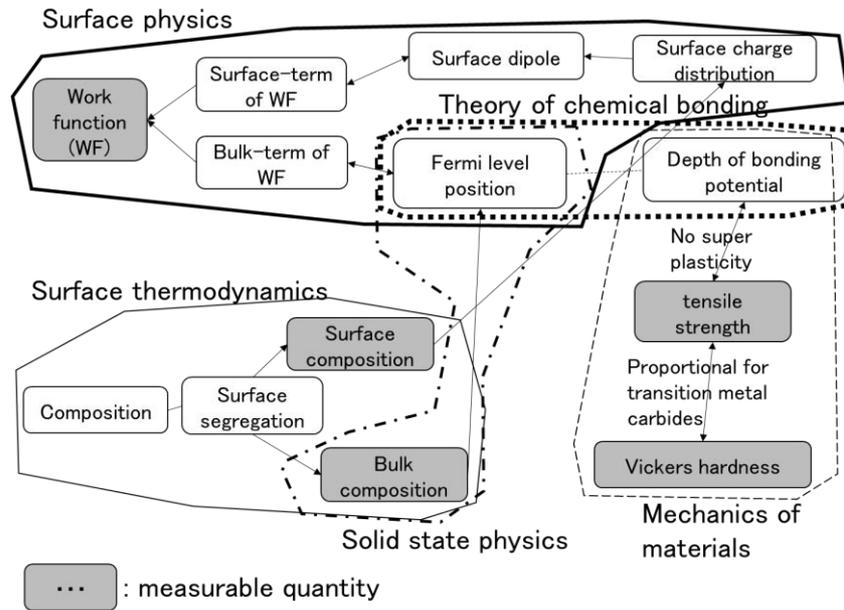
(Received: March 18, 2020; Accepted: April 30, 2020)

Materials Curation<sup>®</sup> is a method to search materials by combining scientific principles in multi-disciplinary way proposed by one of the authors. This method utilizes relations among various material properties to overcome a problem that experimental or calculation data on materials in an issue are not enough at all for big-data type analysis in machine learning. For the interdisciplinary use of scientific principles, the authors have developed a prototype of materials curation support system. Examples of relations among material properties, the concept of database of relations and searching relations, and the construction of the system including the database and searching program are described.

### 1. はじめに

物質・材料分野において、物質・材料のデータに対し 情報技術を駆使することによって材料探索をしようという、マテリアルズインフォマティクスが最近盛んになっている。ただ、実験によりデータを取得するのが必ずしも容易では無い場合も多く、第一原理計算やその他のシミュレーションから求めた物性値をデータとして用いた研究が大勢である。一方、データ数が限られていても、理論に基づいて別のデータから補間が可能な場合も多々考えられる。Fig. 1 に示したのは、電子デバイスの特性を左右す

る「仕事関数」がある種材料群に対しては、「ビッカース硬度」から予測できるということを著者が見出した[1]と時の関係性の図である。科学の専門分野が細分化された結果、各専門分野で一般的に知られている関係性の範囲は限られているが、その境界をいくつもの分野に亘って繋いでいくことで、一見無関係に見える物性量同士の間に関係性が有ることがわかる。Fig. 1 には、それぞれの関係性が知られている分野も示してあり、分野横断的に関係性を利用して、「仕事関数」と「ビッカース硬度」が繋がっていることがわかる。

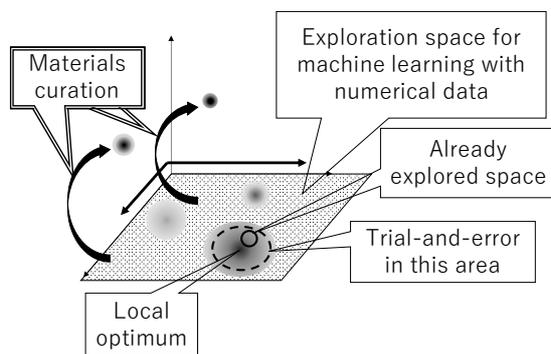


**Fig. 1** Relations among various materials quantities that connect work function and hardness. The specific fields of science where the relations are well known are surrounded by lines.

著者らは、このような様々な物性の中に存在している科学法則に基づく関係性を専門分野横断的にたどり、既存のデータを駆使して物性値の予測や材料探索指針の発見を行う手法「マテリアルキュレーション<sup>®</sup>」を提案している[2-4]。材料研究開発において、現状の改良・改善を超えた画期的な材料を探索する際に、往々にして常識にとらわれがちである。ビッグデータ利用のように既に存在しているデータを集めて解析した場合に得られるのは、従来の枠内での最適化 (Fig. 2 の局所最適点) で、枠の外にあ

るかもしれない最適点を探索することはできない。「マテリアルキュレーション<sup>®</sup>」では、まず常識を忘れて科学原理をさかのぼることで、俯瞰的・分野横断的に材料探索を可能にする。それにより、Fig. 2 のハッチされた平行四辺形で囲んで示した、実験/計算により得られた数値データが存在する探索空間 (実際には多次元空間) の外に最適点の探索空間を広げ、最適点が存在する領域を見つける (最適点を見つけるわけではない)。

現在、「マテリアルキュレーション<sup>®</sup>」手法を多くの研究者・開発者が利用するためのツールとして、各種物性量間の関係性を探索するシステム (=マテリアルキュレーション<sup>®</sup>支援システム) を開発している。その基本設計と探索のフレーム作成についてはすでに発表している[5]が、実用化に向けた最大の課題は、広範囲な材料科学に関係する分野の書物から、原理に基づいて説明のなされた様々な物性間の関係性の記述を抽出することである。本稿の以下で示す例のような、著者や個々の研究者による手作業抽出では広範囲の大量な書物の処理ができない。したがって、関係性のコンピュータによる (半) 自動抽出が不可欠である。今回、共著者である企業との共同研究により関係性のコンピュータによる (半) 自動抽出を試み、マテリアルキュレーション支援システムのプロトタイプを開発したので報告する。



**Fig. 2** Schematic illustration of material search space. Hatched parallelogram shows a space where numerical data are prepared (those data which are considered influential and available) and suitable for machine learning to find optimum within the hatched space. Materials curation enables to explore outside of the hatched space without numerical data.

## 2. 物性間の関係性の例

まず、本システムの物性間の関係性のデータには物性値の数値データは含まれていないことに留意いただきたい。このシステムは、ツイッターから誰と誰が繋がっているか(友人か)という人的ネットワークを示すシステムと同様に、つながりに関してのデータである。このようなつながりデータは「知識グラフ」として人工知能の分野では様々に利用されている。なお、ここで扱う物性間の関係性データには数値データは含まれていないが、つながっている物性間の定量的な関係性(数式)は含まれる。

### 2.1 生成エンタルピーとバンドギャップ

材料は原子から構成され、原子が複数結合して分子や固体が形成される。その際の熱化学的物性量である、分子や固体の生成エンタルピーは、分子や固体の電気的・光学的物性と密接に関係している。Fig. 3 に示したのはその関係性が生じる原理を模式的に示したものである。まず左端の2つの原子 A, O が結合して分子 AO になり、それが集まって固体 AO が生成する場合を考える。結合に関与する原子 A と原子 O の電子の原子軌道間にエネルギー差  $\alpha$  が存在するとする。そのような原子同士が結合すると、単純な分子軌道論によれば、エネルギーの低い原子(この場合原子 O)の軌道よりもさらにエネルギーが  $\beta$  だけ低い結合性軌道(HOMO)と、エネルギーの高い原子(この場合原子 A)の軌道よりもさらにエネルギーが  $\beta$  だけ高い反結合性軌道(LUMO)が形成され、結合性軌道に A, O 両原子からの電子が入ってエネルギーがより安定な分子 AO が形成される。なお、各原子軌道に電子が2個ずつ入っていて合計

4個の場合(閉殻構造)に、分子軌道に4個電子が入った場合とエネルギーが保存される、という条件から結合性軌道に対する反結合性軌道のエネルギー位置は決まっている。分子の生成エンタルピーは、原子 A の電子のエネルギー安定化、 $(\alpha+\beta)$ 、と原子 O の電子のエネルギー安定化、 $\beta$  を足して、 $\alpha+2\beta$  となる。一方、分子の HOMO-LUMO ギャップは Fig. 3 から明らかなように  $\alpha+2\beta$  である。すなわち、分子の生成エンタルピーと HOMO-LUMO ギャップは等しい。

分子が多数集まって固体になると、HOMO, LUMO のそれぞれ多数から成る価電子帯と伝導帯が形成され、価電子帯の中で最も高いエネルギーと伝導帯の中で最も低いエネルギーとの差が、バンドギャップとなる。Fig. 3 から明らかなように、バンドギャップは HOMO-LUMO ギャップより小さいものの、HOMO-LUMO ギャップが大きければバンドギャップも大きいという関係が成立することがわかる。すなわち、固体の生成エンタルピーとバンドギャップは線形に近い(傾きは1より小さい)関係にあると推測される。Fig. 4(a)に示したのは酸化物の生成エンタルピーとバンドギャップとの関係である。一見関係性がなくばらついて分布しているように見えるが、Fig. 4(b)に示したように、酸化物の生成エンタルピーを酸素一原子モル当たりで換算してプロットすると、両者が線形に近い関係にあることがわかる。なお、図中二重丸で示したのは、チタン酸バリウム(BTO)とチタン酸ストロンチウム(STO)の値であり、複合酸化物においてもこの関係性が成立していることがわかる。また、1 eV はほぼ 96.5 kJ/mol なので、傾きは約 0.64 であり、1 より小さいことがわかる。

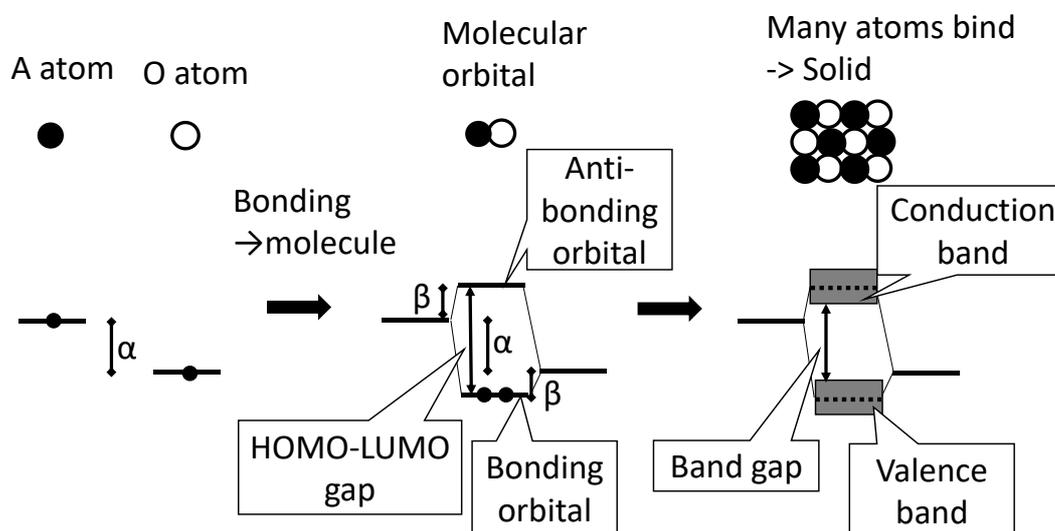


Fig. 3 Schematic illustration of formation of solids from atoms and the energetic relations.

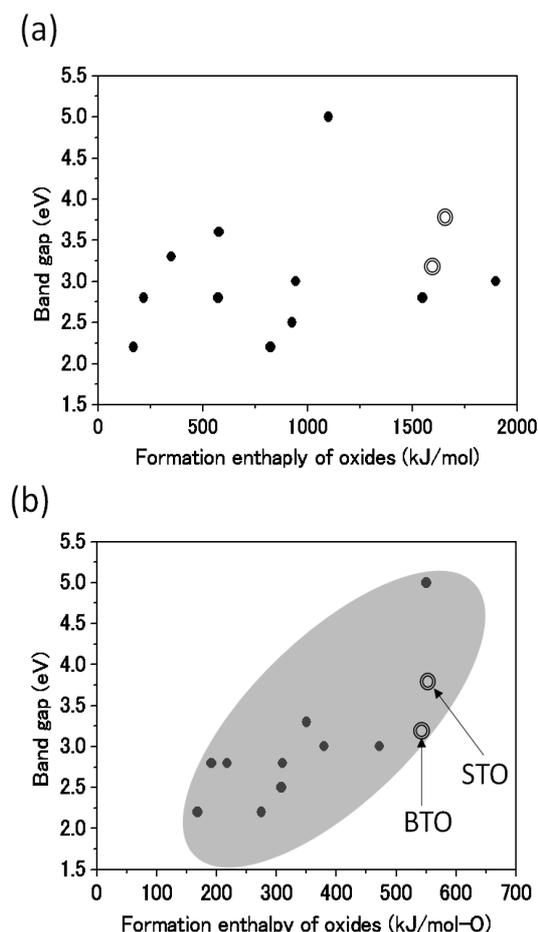


Fig. 4 Examples of relation between formation enthalpy of oxides ( (a): kJ/mol-oxide unit, (b): kJ/mol-O) and band gap energy.

## 2.2 生成エンタルピーと化学吸着エネルギー

2.1 の例と似ているが, 利用分野の異なる例として, 酸化物の生成エンタルピーと酸素吸着エネルギー (化学吸着) の関係を挙げる. Fig. 5(a)に模式的に示したように, 酸素吸着により表面に形成される結合は, 酸化物中の結合の一部という見方ができ, 実験値が入手できる系に限られる酸素吸着エネルギーが, 豊富な実験・計算値が入手できる酸化物生成エンタルピーと関連付けられると推測される. 実際, 酸素吸着エネルギーの実験値が入手できる系で, 酸化物生成エンタルピーと酸素吸着エネルギーの関係をプロットすると Fig. 5(b)に示したように高い線形関係にある. ただし, ここでも模式図からわかるように, 原子の個数と価数に対して規格化を行うことが重要である.

## 2.3 生成エンタルピーと酸化還元電位

電気化学反応において, 反応の標準ギブスエネルギー変化  $\Delta G^\circ$  とその反応の標準酸化還元電位  $E^\circ$  との間に下記の式(1)で表される関係があることは基礎的な教科書に書かれている.

$$E^\circ = -\Delta G^\circ/nF \quad (1)$$

なお, ここで  $n$  は電気反応の際にやり取りされる電子の価数,  $F$  はファラデー定数 ( $9.65 \times 10^4$  C/mol) である.

ここで, Fig. 6(a)に模式的に示したように, 金属 A と A に元素 B を混ぜた合金 A(B)を水溶液に浸した場合に, どちらから A が溶け出すかを考える. これ

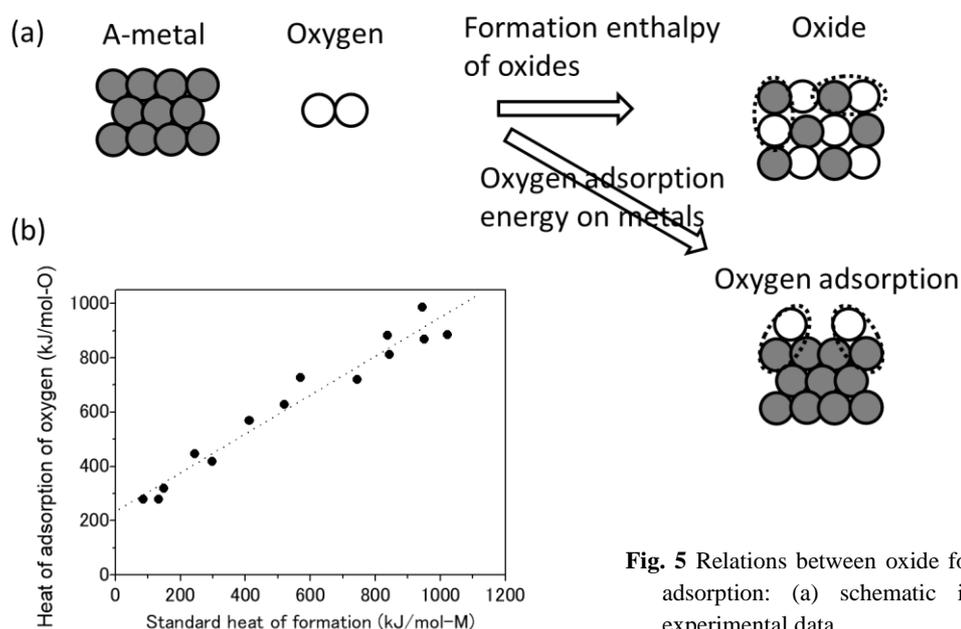
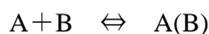


Fig. 5 Relations between oxide formation and oxygen adsorption: (a) schematic illustration and (b) experimental data.

は、言い換えると A と A(B)のどちらが耐食性が高いかという問いである。ただし、単体の B は A よりも溶け出しにくいとする。まず、A と A(B)との間のエネルギー関係を考える。



なる反応において、合金 A(B)が熱力学的に安定 (= 状態図上に存在する) ならば、上式の右側への反応のギブスエネルギーは負である。そうすると(1)式で A(B)から A が溶け出す (=イオン化する)  $E^\circ$  は、A から A が溶け出す  $E^\circ$  よりも大きくなる (Fig. 6(b)参照)。すなわち、A(B)の方が A が溶け出しにくい、耐食性が高いということになる。これは、酸化還元電位のデータが無いが状態図データは存在している多種多様な合金・化合物において、非常に簡単に耐食性の判断を可能にする。また、電気化学反応において気体の発生が無い場合、エントロピー変化  $\Delta S^\circ$  のギブスエネルギー変化  $\Delta G^\circ$  への寄与は、エンタルピー変化  $\Delta H^\circ$  の寄与に比べて小さいため、ギブスエネルギー変化をエンタルピー変化 (=生成熱) で近似できることが多い。化合物形成の生成熱のデータはかなり豊富に存在しており、(1)式の  $\Delta G^\circ$  に  $\Delta H^\circ$  の値を代入することで、化合物にしたことによる酸化還元電位の変化も大まかに推定することができる。

### 3. 関係性の探索

2.で取り上げた例を元にいくつかの物性量の間関係性をつなげたネットワーク形式の図を Fig. 7 に示した。このように、様々な物性量の間には、原理から考えられる様々な関係性が存在する。その関係性は、定量的に厳密に表現されるものから、近似式で表現されるもの、定性的に表現されるものなど色々な種類が存在する。関係性のネットワークに対

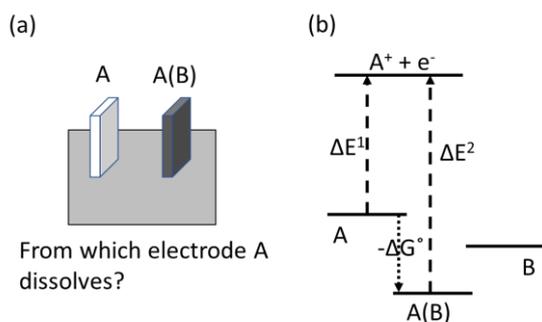


Fig. 6 (a) Schematic illustration of dissolution of A from metal-A and alloy A(B) in solution, (b) energy diagram of ionization of A in A, B and A(B).

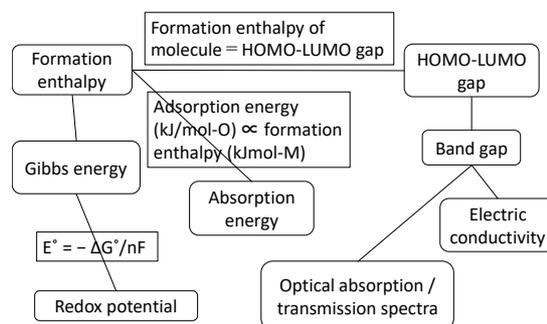


Fig. 7 Network type diagram that shows relations between material properties (no numerical data of material properties included).

- Two basic search
- Relation between property-A and property-B
  - Surrounding properties around property-M
- Choose search type
- ① path between (A) and (B)
    - shortest
    - within (n) connections
    - all paths
  - ② search properties around (M) within (n) connections

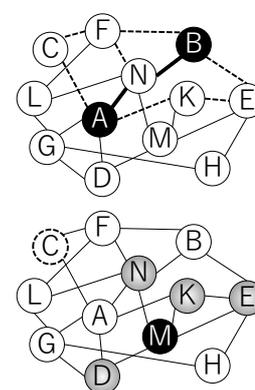
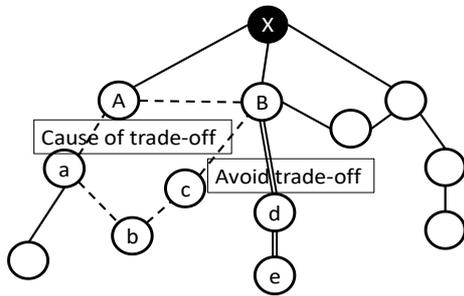


Fig. 8 Two basic types for searching relations among properties.

しては、数学的に「グラフ理論」と呼ばれる分野において様々な解析が可能である。グラフ理論は、電車路線においてA駅からB駅までの最短乗り換え経路を見つける、荷物の配送において物流拠点をどこに設けるとよいかを見つけるなどの問題に利用されている。物性間の関係性の解析のうち最も基本的な2つとして、Fig. 8に示した、①物性Aから物性Bへ至る経路(関係性の連鎖)を見つける(上の図)、②ある物性Mとつながっている物性としてどのようなものがあるかを見つける(下の図)がある(Fig. 8の左側上部の説明)。①の探索には、Fig. 8の上の図で、太線で示したNを経由する最短経路、破線で示した物性と物性を繋ぐ関係が3つ以内にある経路(A-N-F-B, A-C-F-B, A-K-E-B)、全経路(AからBに至る途中にD, M, Gなど遠回りする経路も含む)などのような探索が可能である。②の探索では、下の図でグレイになっているD, E, K, NがMと直接つながっている物性、白抜き実線囲いなのがMと2つの関係性を介して繋がっている物性であり、破線囲いのCはMと3つの関係性を介し



**Fig. 9** Schematic network diagram for finding a route to avoid trade-off when A and B have trade-off relations for improving quantity X.

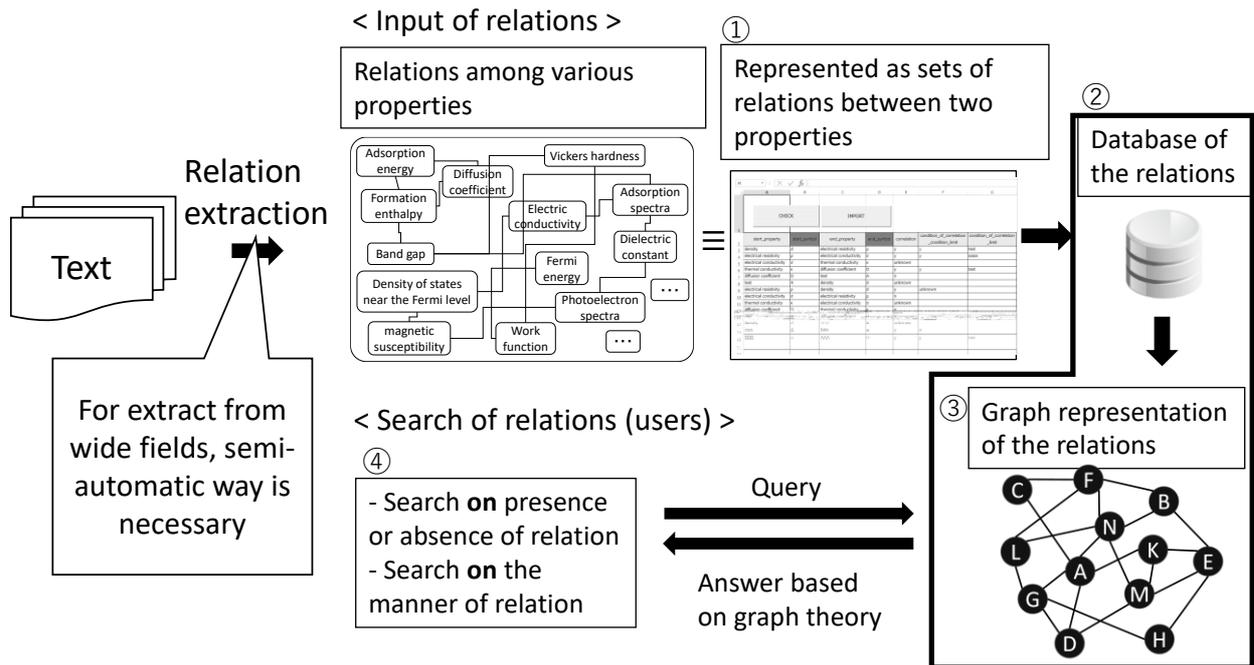
て繋がっている。

物性間の関係性をネットワークの形のデータベースとして蓄積し、グラフ理論に基づく解析が可能になると、一般的にトレードオフと言われている2つの物性間に対し、トレードオフを避ける方法があるかどうかを調べるなどができる。Fig. 9 に示したのはその模式図で、X という物性を決めている A と B の物性間にトレードオフ関係がある場合に、そのトレードオフは破線で示した複数の物性間の関係性に由来する。その場合、破線でつながっている物性 a, b, c の値を変化させても（その値が従来材料と異なる材料を探す）A と B との間のトレードオ

フ関係から逃れることはできない。一方、B が破線とは別の二重線でつながった物性 d, e と関係していることがわかると、d, e の値を変化させる（その値が従来材料と異なる材料を探す）ことにより、A と B との間のトレードオフ関係から逃れて B の値を制御することができる。すなわち、トレードオフを逃れた材料が見つかる可能性が示される。なお、二重線でつながった物性 d, e のようなトレードオフを逃れる経路が見つかったとしても、必ずしもトレードオフを逃れた材料が見つかる訳ではない。

#### 4. システムの構築

各種物性量の間関係性をネットワーク型データベースとして蓄積し、物性間の関係性を探索するシステムを研究者・技術者が利用できるようになると、材料開発の世界に革新をもたらすと期待される。このようなシステムの基本設計と探索のフレーム作成について著者の一人が既に発表している[5]が、実際に稼働するシステムにするには、コンピュータによる物性量の間関係性の（半）自動抽出が不可欠である。そこで、今回マテリアルキュレーション®支援システムとして、コンピュータによる関係性の抽出を備えたシステムの構築を試みた。その概要を Fig. 10 に示した。



**Fig. 10** Search system of relations among materials properties prototype demonstrated in March 2019 at JJSP.

#### 4.1 関係性のコンピュータによる抽出

現時点では、9冊の教科書の書籍[6-14]から関係性抽出を行ったデータが入っており、教科書分野としては一般的な材料科学に加え、電子デバイス材料、固体電気化学の書籍を用いている。関係性は、基本的に自然言語処理技術の構文解析と照応解析を用い、係り受けを基に抽出している。この点は、現在多く見られる、共起関係を用いた統計的な自然言語処理技術の手法とは異なる。構文解析・照応解析のアルゴリズムの背後には統計的手法が用いられているが、関係性の抽出そのものには統計的手法は用いていない。

#### 4.2 プロトタイプのシステム

上記のように抽出した関係性を入力したデータベースに対し、Fig. 8 に示した2つの基本的な関係性の探索ができる GUI (Graphic User Interface) を備えたシステムを開発した。Fig. 11 には、始点と終点を指定して探索した結果の一例を示す。図中、(1)(2)と記した部分が、Fig. 8 の上または下のどちらかの探索をしたいかを選んで、その探索条件を入力する部分である。例では、始点として(1)の横長画面の from の部分に polarizability という電気的性質を表す物性を、to の部分に yield strength という機械的(力学的)性質を表す物性を入力して、これらの物性を最短で

結ぶ経路を表示させたものである。この図では見づらいが、表示されたノード(物性の名前が書かれた円)をドラッグすることで、どの物性とどの物性がつながっているのかを見やすくすることができる。

Fig. 12 に示したのは、Fig. 11 の(2)の探索方式で keyword として tensile strength を選択して、この物性と直接つながっている他の物性を探索した結果である。例では、thermal conductivity という熱電的性質が繋がっていると表示されている。そのつながり線(エッジ)にカーソルを当ててクリックすると、そのつながりが記述されている書籍名と、コンピュータがそのつながりを見つけた文章が表示されるようになっている。

#### 5. 最後に

教科書のような科学法則が記述されている文書から、ある物性がどの物性と関係しているか、どの物性により制御されているかというような物性間の関係性に関する情報を、①コンピュータにより抽出し、②データベース化した上で、③物性間の関係性を探索する、マテリアルキュレーション支援システムの構築を行った。まだプロトタイプの段階で、将来的にはより広範囲の材料科学分野の文書を入力し、より精度の高い関係性抽出と定量的な関係性の抽出、より複雑な関係性探索や2つの物性の値のトレード

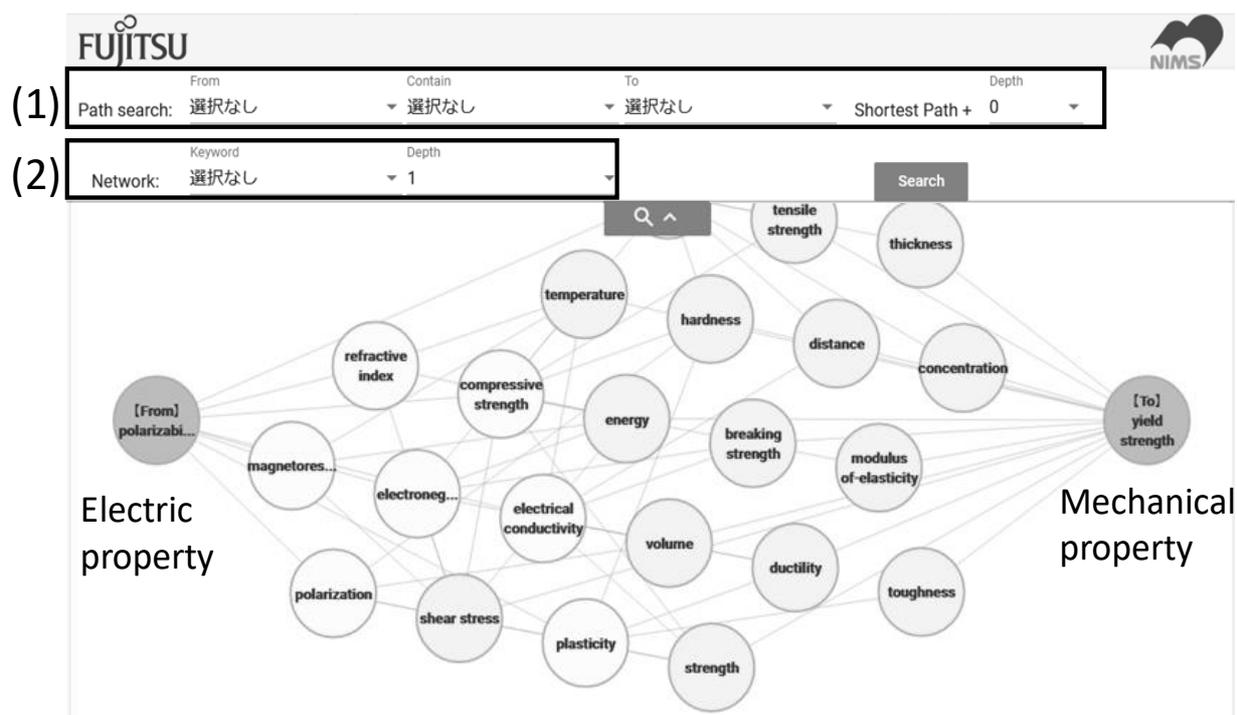


Fig. 11 Materials curation support system start window and an example of path search (1) result.

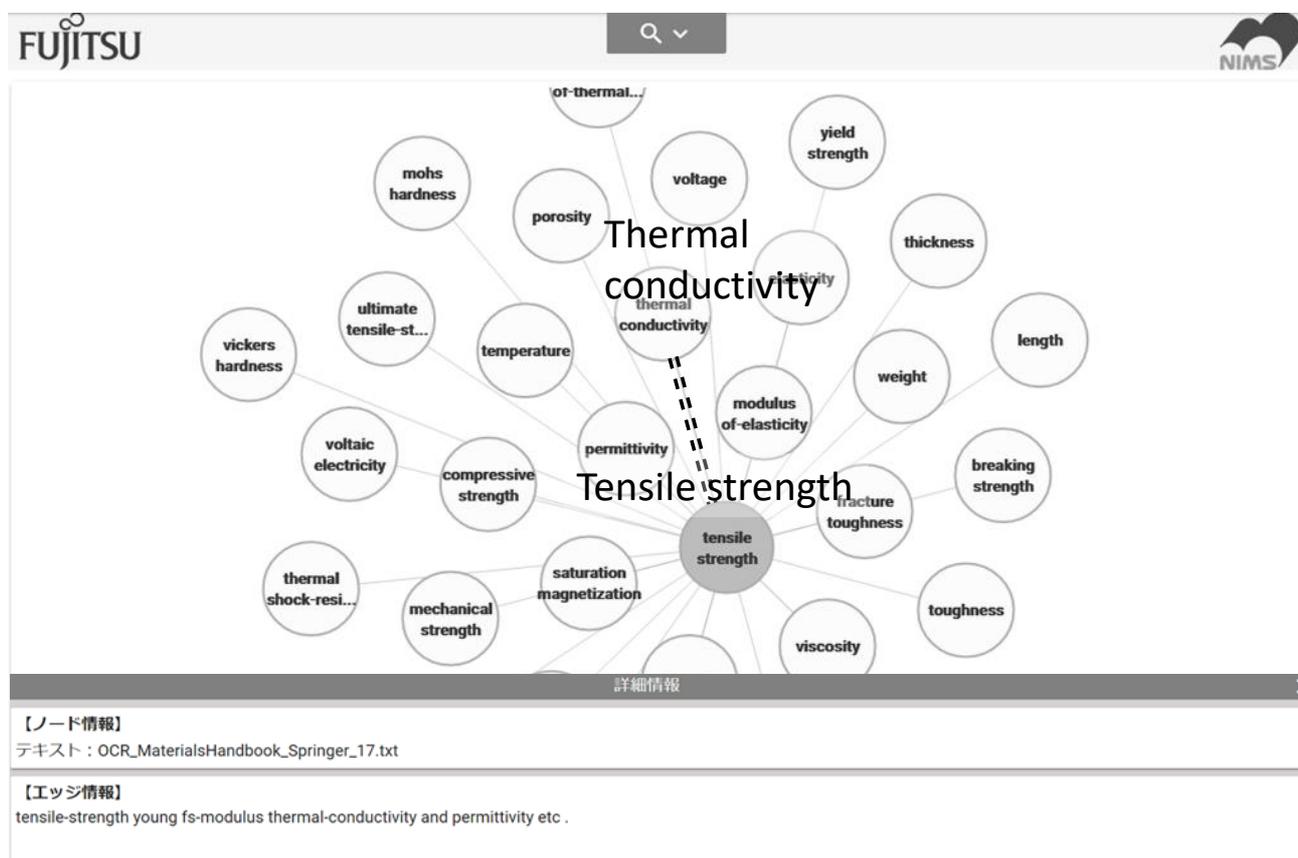


Fig. 12 An example of network (2) (see Fig. 11) result and the information on knowledge source at the edge of the network.

オフ回避の可能性判断など, を実装する予定である.

また, 構想としては既に, この知識データベースで関係性を抽出している物性の値を測定するための物性測定法と各物性との関係性をデータベース化して, 探索するシステムへの拡張も考えられている [15]. 例えば, 「バンドギャップ」という物性値は, ①「可視・紫外吸収スペクトル法」という測定法を用いて, その吸収端の値から求めたり, ②「電気伝導度測定」という測定法を用いて, 電圧・電流曲線 (式がある) の値から求めたりできることが分かるようなシステムである.

その他にも様々な拡張が構想されているが, それらの実現は, コンピュータが教科書的書籍を如何に専門家に近い形で読解 (情報抽出) できるかにかかっている. コンピュータによるテキストの読解技術は自然言語処理と呼ばれる技術分野で, 現在非常な速さで発展しており, その技術を随時取り入れて実用システムを開発する予定である.

## 6. 文献

[ 1 ] M. Yoshitake: *J. Vac. Sci. Technol. A* **32** (2014)

061403(1-6).

- [ 2 ] 吉武道子, *機能材料*, **33**, 48 (2013)..
- [ 3 ] 吉武道子, *MITS2013 材料データベースシンポジウム予稿集*(2013) p. 40.
- [ 4 ] 吉武道子, *日本金属学会誌*, **80**, 603 (2016).
- [ 5 ] 吉武道子, 桑島功, 柳生進二郎, 知京豊裕, *表面と真空*, **61**, 200 (2018).
- [ 6 ] C. Barry Carter, M. Grant Norton, *Ceramic Materials, Science and Engineering*, Springer (2007), ISBN-10: 0-387-46270-8 e-ISBN-10: 0-387-46271-6, ISBN-13: 978-0-387-46270-7 e-ISBN-13: 978-0-387-46271-4.
- [ 7 ] Waldfried Plieth, *Electrochemistry for Materials Science*, Elsevier (2008), ISBN:978-0-444-52792-9
- [ 8 ] Rolf E. Hummel, *Electronic Properties of Materials*, Fourth Edition, Springer (2011), ISBN 978-1-4419-8163-9 e-ISBN 978-1-4419-8164-6
- [ 9 ] Clifton G. Fonstad, *Microelectronic Devices and Circuits*, Published under Creative Commons License 2.5 (2006), which is available at <http://creativecommons.org/licenses/by-nc-nd/2.5/>

- [10] Eric J. Mittemeijer, *Fundamentals of Materials Science*, Springer (2010), ISBN 978-3-642-10499-2 e-ISBN 978-3-642-10500-5
- [11] François Cardarelli, *Materials Handbook*, Springer (2008), ISBN 978-1-84628-668-1 e-ISBN 978-1-84628-669-8
- [12] S. M. Sze and Kwok K. Ng, *Physics of Semiconductor Devices*, 3rd Edn., John Wiley & Sons (2007), ISBN-I 3: 978-0-471-14323-9, ISBN-10: 0-471-14323-5
- [13] Vladislav V. Kharton (ed.), *Solid State Electrochemistry I, Fundamentals, Materials and their Applications*, WILEY-VCH Verlag (2009).
- [14] Rolf E. Hummel, *Understanding Materials Science*, 2nd Edn., Springer (2004), ISBN 0-387-20939-5
- [15] 「探索システムおよび探索方法」特願 2019-026191

## 査読コメント, 質疑応答

### 査読者 1. 鈴木昇 (宇都宮大学)

本稿は, 科学技術分野における種々の探索方法に関する全く新しい方法論を提案し, その有効性について記述した技術報告である. 内容は大変興味深く, 今後の発展が大いに期待され, 材料開発などの諸分野で利用されることで, 画期的な技術開発に繋がるものと考えられ, 掲載を推奨する. ただし, 以下の点で修正等を検討されたい.

#### [査読者 1-1] 1. はじめに

1 ページ目, 左カラム, 下から 3 行目: 「理屈」より「理論」が好ましいと考えます. ご検討ください.

#### [著者]

まず, 本稿の内容に非常に興味を持っていただきありがとうございます. また, 丁寧に読んでくださり, 細かなミスなども見つけてくださりありがとうございます. 以下, ご指摘に従い修正した内容を記述しますので, よろしくお願いいたします.

ご指摘をありがとうございました. 「理論」に修正しました.

#### [査読者 1-2] 2.2 生成エンタルピーと吸着エネルギー

新鮮面 (表面酸化物等がない状態, fresh surface or nascent surface) における酸素の吸着は化学吸着であり, 酸化反応と考えるべきでしょう. その面からの検討は如何でしょうか. また, 横軸を standard heat of formation (kJ/mol-O) とするとどうなりますでしょうか. さらに, 「原子の個数と価数に対して規格化を行う」と記載されていますが, どの様に規格化したのでしょうか.

#### [著者]

ご指摘のように, 新鮮面における酸素の吸着は化学吸着です. その場合においても, 著者が関わっている分野では「吸着エネルギー」という言葉を使うのが一般的ですが, 混乱を避けるために, 「化学吸着エネルギー」としました. 横軸を standard heat of formation (kJ/mol-O) とすると, プロットはもっとばらついて,  $M_xO_y$  の x:y 比が同じ酸化物が直線上に並ぶ感じになります. 「原子の個数と価数に対して規格化を行う」とは,  $A_xO_y$  なる組成の酸化物の生成エ

ンタルピーと酸素吸着エネルギーの関係を見るには、生成エンタルピーを  $\text{kJ/mol-M}$  に規格化するという意味です。例えば  $\text{MO}_2$  という組成の酸化物ではバンドギャップのようにバルクの性質を考えると  $\text{M-O}$  結合 2 モルに対して生成エンタルピーが  $\text{kJ/mol}$  で与えられており、 $\text{M-O}$  結合 1 モル当たりには換算して用いるということです。“酸素吸着エネルギー”の場合は、金属  $\text{M}$  に酸素が吸着する際に酸化物の価数に関係なく結合が形成されるので、 $\text{M}$  原子 1 モル当たりには換算して用います。なお余談ですが、最も安定な酸化物だけでなく、例えば  $\text{TiO}_2$  だけでなく  $\text{TiO}$  や  $\text{Ti}_2\text{O}_3$  などの生成エンタルピーを ( $\text{kJ/mol-O}$ )、( $\text{kJ/mol-M}$ ) で眺めるといろいろ面白いことが分かります(大学院の特別講義で使ったことがあります)。

### [査読者 1-3] 3. 関係性の探索

最終行：「**B** の値が好ましい材料を発見できる可能性があることがわかる。」この文章が理解しづらいので、もう少し深く説明してください。

### [著者]

その前の文章も含めて、文章を修正しました。

### 査読者 2. 名越正泰 (JFE テクノリサーチ)

材料の物性間の関係性を狭い範囲、特定の分野にとらわれることなく見出して、材料設計に資する支援システムを構築する試みと現状がわかりやすく紹介されています。このシステムにより想像していなかった物性間の関係性を知ることができれば、材料設計や製造プロセス設計に携わっている技術者、研究者を刺激し発想の回転を加速することが期待できると感じました。JSA 誌の多く読者も、直接的あるいは間接的に材料やプロセスの設計にかかわっていますので、本記事は業務の幅を広げ結果の価値を高めるために参考になると思います。また、2 節で紹介されているいくつかの事例における考え方は、表面分析に携わる方にとって教育的な価値もあると思いました。以上により本記事は JSA に掲載する価値が高いと判断します。このままの掲載でも良いですが、読者の理解や期待をさらに高めることにつながればと思い、質問とコメントをさせていただきます。

### [査読者 2-1] 関連付けた結果の評価

構築しているシステムでは、抽出された関係の信

頼性、強弱、および関係性(どのような関係か? 単純ではない場合が多いと思いますが)を評価できるようになりますでしょうか? 例えば、Fig. 12 において tensile strength につながる様々な物性との間を結ぶ線の太さのようなもの、Fig. 11 において polarizability と yield strength をむすぶ経路においてつながりの強い順番(最短ルートではない可能性あり)、などです。Fig. 9 においては、材料設計では A と B の反する特性を制御する事例が述べられていますが、線で結ばれた物性間の強さ関係がわかれば D や e を変化させることが有効かどうかを検討するのに役立つように思いました。まずは、教科書から物性間の関連を抽出した際にその関係を評価できれば良いのかもしれませんが。

### [著者]

有益なコメント、当該技術に対する高い関心、それに由来する質問をありがとうございました。また、「2 節で紹介されているいくつかの事例における考え方は、表面分析に携わる方にとって教育的な価値もあると思いました。」に関して、実はこの支援システムの教育・研修用バージョンも構想しており(2020 年の 3 月の、中止されたが予稿集の発行をもって発表済み扱いになった応用物理学会で発表)、可能なら別に JSA 誌に投稿させていただければと存じます。

以下、回答いたしますので、よろしくお願いいたします。

抽出された関係の信頼性、強弱、および関係性の種類の評価に関してですが、「評価」には種類があり、その種類ごとに説明をさせていただきたく存じます。

### 関係の抽出の信頼性、強弱の抽出の信頼性、関係性の種類の抽出の信頼性

こちらは、「教科書に記載されている通りにコンピュータが抽出できているか」という自然言語処理の技術的に関わる評価です。自然言語処理では、基本、どのシステムでも人がラベル付けした学習データに対して、どれだけコンピュータアルゴリズムがラベルを再現できるかを評価します(猫の画像に対して「猫」、犬の画像に対して「犬」のように、ラベル付けをする)。抽出の信頼性に関しては現在、抽出のアルゴリズムを向上させつつ評価しておりますが、まだ商品になるだけの信頼性は得られていま

せん。強弱や関係性の種類については、数式の抽出が必須で、現状ではプロトタイプには組み込まれていません。数式の抽出については2020年の3月の応用物理学会（中止だが、予稿集の発行をもって発表済み扱い）で第一弾を発表しました。機会があればお話ししたいのですが、プロトタイプで用いている自然言語処理とは全く異なる技術を用います。

### 関係、関係の強弱、関係性の種類自体の評価

2つの物性が線で結ばれるのが妥当かどうか、線で結ばれた物性間の強さ関係（比例定数など）や関係性の種類（線形か、指数的か、など）の妥当性に関しては、本システムでは評価しません。あくまでも処理した教科書本文の記述の通りを採用します。そのため、Fig. 12に示したように、関係性がどの書籍のどのフレーズから抽出されているかを表示する機能を搭載しています。関係の妥当性に疑義がある場合は、情報を与えてユーザーに判断していただくという方針を取っています。なお、経路探索をする際には、必ずしも関係性の強弱で判断するとは限らず（トレードオフを避けるためにある物性の制御を遠回りな方法で行うなど）、関係性の種類に関して材料の置かれた条件によって変化するので、それらを考慮した探索が可能なシステムを構想しています（NIMS 単独特許出願済み）が、実装にはまだまだ技術課題が山積しています。

### [査読者 2-2] プロトタイプについて

イントロダクションの後半で、「書物から、原理に基づいて説明のなされた様々な物性間の関係性の記述を抽出すること」というシステム構築の最大課題に対して、プロトタイプでは「関係性のコンピュータによる（半）自動抽出を試み」、「自然言語処理技術の構文解析と照応解析を用い（4.1）」で抽出したと記載されています。今回試みた方法の結果をどう評価していらっしゃるでしょうか？ 課題はございましたでしょうか？ また、関係性を抽出する際に1) 対象とする物性値の選定はどのように行われたか、2) 教科書中の数式もこの方法で抽出できたかどうか。以上の点、知りたいと思いました。

### [著者]

コンピュータによる書物からの関係性抽出は、手で抽出した数千の関係性に対し、コンピュータで抽出した結果を比較して正解率を評価しております。自然言語処理は、最先端の研究領域でも発展途上で

問題は山積しており、専門家が読むようにコンピュータに読解させるにはまだまだ程遠いです。1) 関係性の抽出は、まず基本的教科書を選定して、その内容をコンピュータで構文解析・照応解析して、文法的に関係性を記述しているフレーズを抽出し、それが物性間の関係性であるかを吉武が判別し、物性を表すフレーズのデータベースを作成し、そのフレーズの文脈的な現れ方からコンピュータが物性を表すフレーズを抽出する、というプロセスを踏んでいます。2) 構文解析・照応解析はあくまでも自然言語の解析法なので、人工的な言語の一種である数式を抽出するのは不可能で、別の方法が必要です。こちらも現在、吉武単独で開発中ですが、プロトタイプには含まれていません。

### [査読者 2-3] 物性の値を測定するための物性測定法

すでに考えていらっしゃると思いますが、表面分析技術で得られるパラメータも関係づけていただくと面白いと思いました。例えば、XPSで評価する内殻順位の Binding energy が何かの物性に結びつかないか？ 意外な物性を測定できるかもしれない、・・・と考えることを想像します。もし有益な関係が見つかる。Binding energy をつかって極表面の思ってもみなかった物性評価、やその空間分布がわかるようになるかもしれません。

### [著者]

はい、おっしゃる通りです。「5. 最後に」に少し述べていますが、物性値の測定方法との関連付けも吉武の構想の中には入っており、NIMS 単独特許出願済みです。そのほか、教科書に記載されていない関係性も思いつきますが、そのような関係性は、基幹システムに各ユーザーが付加できる構造を考慮しており、例えば名越さんが考えて見出した教科書に書かれていない関係性を名越データベースとして有償で外部に利用できるようにする、といったシステムが考えられます。これらは吉武の構想であり、共同研究・開発している富士通グループが商用化時に採用するかどうかは、潜在的ユーザーの要望とビジネスとして成立するかの問題かと存じます。SASJから講演を依頼されておりましたので、研究会の開催の目途が立ったら将来構想（2020年の3月の、中止されたが予稿集の発行をもって発表済み扱いになった応用物理学会で発表）も含めて話をさせていただければと存じます。

**[査読者 2-4]**

要旨：英文のほうが本原稿の内容紹介として情報量が多いと思います。文字数が許せば和文を英文に近づけてはいかがでしょうか。

**[著者]**

和文要旨を英文に近い形に情報量を増やしました。