

## 科学研究費助成事業 研究成果報告書

令和 4 年 6 月 1 日現在

機関番号：12601  
研究種目：基盤研究(C) (一般)  
研究期間：2019～2021  
課題番号：19K04999  
研究課題名(和文) 第一原理計算と実験値インフォマティクスによるハイスループット新規熱電材料探索  
  
研究課題名(英文) High-throughput searches for new thermoelectric materials by first-principles calculations and experimental materials informatics  
  
研究代表者  
桂 ゆかり (Katsura, Yukari)  
  
東京大学・大学院新領域創成科学研究科・特任助教  
  
研究者番号：00553760  
交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：7800論文以上から集めた38000試料以上のデータを含む、世界最大規模の熱電材料の実験値のデータセットのInverse Jonkerプロット解析を行い、熱電特性の材料系依存性を可視化した。熱電特性の1変数関数表示により、最大ZTやパワーファクターを少ないデータから予測する技術を開発した。ボルツマン輸送方程式に基づく熱電特性の第一原理計算結果と合わせた解析により、これまでの第一原理計算に基づく熱電特性予測を大幅に改善できる電子緩和時間補正の理論を構築した。また、大規模実験データ解析技術として、化学組成からの自動材料系判定プログラムと、タグツリーに基づく柔軟なデータ整理技術の開発に成功した。

## 研究成果の学術的意義や社会的意義

これまで高性能な新規熱電材料の探索は難しかった。候補物質が多い上に、相互に絡みあう多数のパラメータに支配されているために母物質依存性と試料依存性が分離されずに研究されており、本質的な特性改善指針を見出すことが難しいためである。そこで本研究では、本研究者が開発したStarrydata webシステムで、過去に出版された論文から熱電特性の温度依存性の実験データを大規模収集に取り組み、本研究においてデータの大幅な追加を行った。第一原理計算と実験値Materials informatic(MI)を組み合わせ、母物質の選定から不純物元素とそのドーパ量の選択、合成条件の最適化までを効率化できた。

研究成果の概要(英文)：Using our world's largest data set of experimental thermoelectric properties, from 38,000 samples published in more than 7,800 papers, we carried out inverse Jonker plot analysis to visualize the material system dependence of thermoelectric properties. We predicted maximum ZT and power factor from small data by single parameter expression of thermoelectric properties. By analyzing the results with first-principles calculations of thermoelectric properties based on the Boltzmann transport equations, we constructed a theory of electronic relaxation time correction to improve the prediction of thermoelectric properties based on conventional first-principles calculations. We also succeeded in developing an automatic material system determination program based on chemical composition and a flexible data organization technique based on a tag tree.

研究分野：無機材料科学

キーワード：熱電材料 材料インフォマティクス データ科学 データベース 第一原理計算

### 1. 研究開始当初の背景

熱と電気を相互変換できる熱電材料は冷却素子や発電装置としての応用が進んでおり、普及には高性能化が必要である。だが熱電特性は相互に絡みあう多数のパラメータに支配されており、母物質依存性と試料依存性が分離されずに研究されてきたため、本質的な特性改善指針を見出すことが困難であった。独自開発した Starrydata2 web システムを用いて、過去に出版された論文から熱電特性の温度依存性の実験データの大規模収集に取り組んだ。その結果、6000 本以上の論文から集めた 25000 試料以上の熱電特性データを含む、世界最大規模の実験値のデータセットを得ることができた。

### 2. 研究の目的

Materials informatic (MI)を用いて、母物質の選定から不純物元素とそのドーパ量の選択、合成条件の最適化までを効率化することで、新規熱電材料探索をハイスループットに行う方法論を確立すること。

### 3. 研究の方法

本研究者が開発した Starrydata web システムで、過去に出版された論文から熱電特性の温度依存性の実験データを大規模収集に取り組んでおり、本研究においてデータの大幅な追加を行った。第一原理計算と実験値 Materials informatic(MI)を組み合わせ、母物質の選定から不純物元素とそのドーパ量の選択、合成条件の最適化までを効率化した。

### 4. 研究成果

本研究により約 1800 論文からのデータ収集を行い、7800 論文以上から集めた 38000 試料以上のデータを含む、世界最大規模の熱電材料の実験値のデータセットを得た。このデータを Figure 1 に示す。このような 1 万試料を超えるデータの同時プロットによる熱電特性分布の可視化は本研究が初めてである。

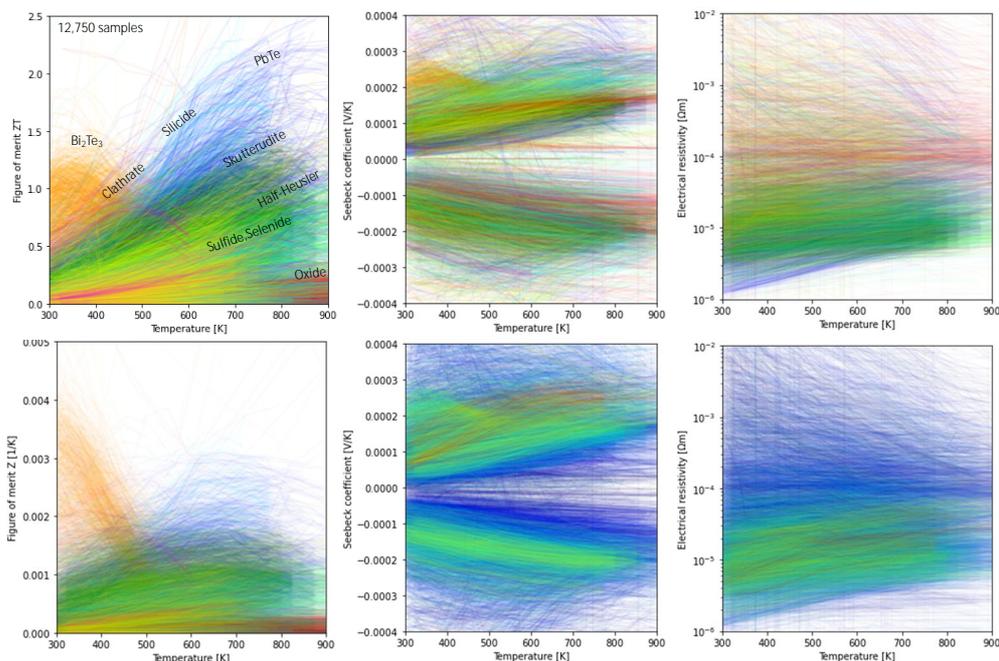


Figure 1. Starrydata web システムで収集した熱電特性の大規模実験データの例。

続いて、これらのデータの Inverse Jonker プロット解析により、熱電特性の材料系依存性を初めて可視化した。Figure 2 は、Starrydata に収録した熱電材料試料の 400 K における熱電特性をプロットした inverse Jonker plot である。\$<0\$ では右上がりの直線、\$>0\$ では左上がりの直線となっており、全体として山なりの形になっており、この傾きや切片、\$S\$ の変化幅が母物質によって異なっていることが観察できた。

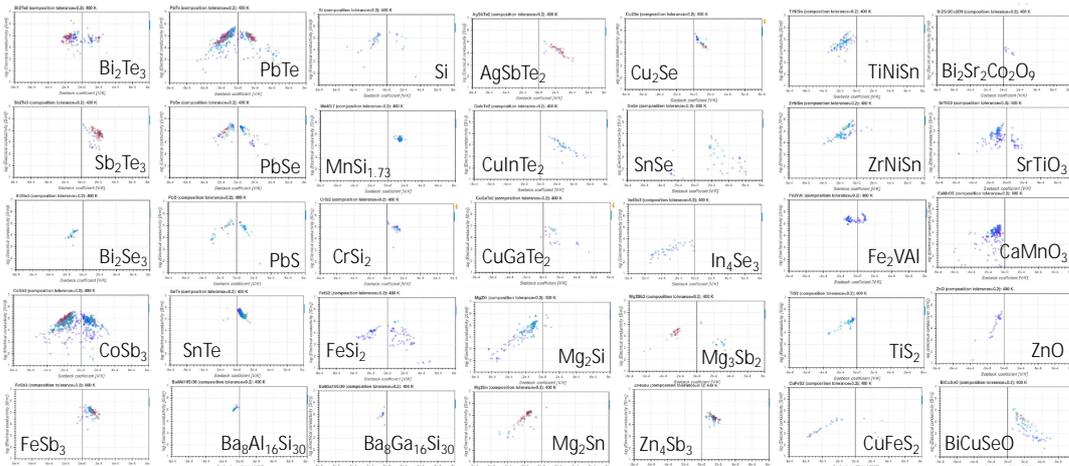


Figure 2. 400 K における熱電特性の inverse Jonker plot.

続いて、熱電特性の 1 変数関数表示により、最大 \$ZT\$ やパワーファクターを少ないデータからでも予測できる技術を開発した。これまで、\$ZT\$ を構成する 3 つのパラメータである \$S\$、 $\sigma$ 、 $\kappa$  はキャリア濃度 \$n\$ の関数として表されてきた。これを式で表すと以下の形になる。

$$ZT(n) = \frac{S^2(n)\sigma(n)T}{\kappa(n)}$$

だが、\$n\$ の測定にはホール測定装置など特殊な実験装置が必要であるため、論文に掲載されている多くの熱電材料試料については、\$n\$ の値が分からない。このため、各パラメータの関数形を実験データから知ることが難しい。ところで、\$S\$ と \$\log n\$ の関係をプロットすると、直線に近い曲線が得られることは古くから知られていて、Jonker plot と呼ばれている。仮にこの直線の傾きを \$a\$、切片を \$S\_0\$ とおくと、\$S\$ は \$S\$ のみの関数として表せる。\$\sigma\$ も Wiedemann-Franz 則を用いることで、\$\kappa\$ と \$\sigma\$ の関数として表せる。

この方法に基づき、\$ZT\$ を以下のように \$S\$ のみの関数として表すことを試みた。ここで、\$S\_0\$ は \$S=0\$ における \$S\$、\$L\_0\$ はローレンツ数を示す。

$$zT(S) = \frac{S^2\sigma T}{\kappa(\sigma)} = \frac{S^2\sigma(S)T}{\kappa(\sigma(S))} = \frac{S^2\sigma_0 T \exp(-aS)}{L_{\text{eff}}T\sigma_0 \exp(-aS) + \kappa_1}$$

この曲線とともに収集したデータをプロットしたのが Figure 3 となる。このように、理論的な分布とよく一致していることがわかり、この曲線を決める変数を少数のデータから確定すれば、最適化後の \$ZT\$ (ピーク値) を求められることがわかった。

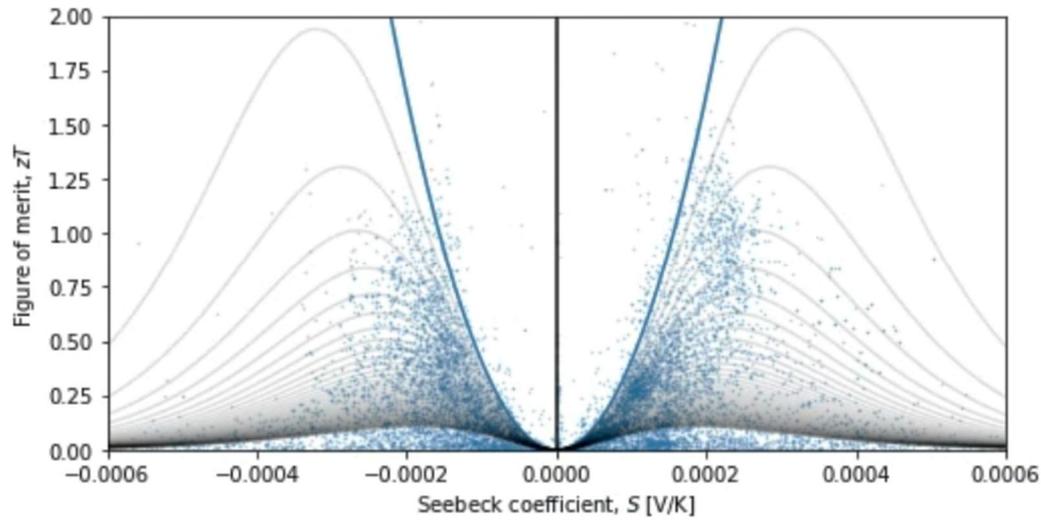


Figure 3. パラメータを変化させた  $zT$  の 1 変数表示の理論曲線と実験データの分布。

Inverse Jonker plot の傾きを第一原理計算から求めることを試みた。電子熱伝導率の第 2 項を入れる補正を施した BoltzTraP によって、ボルツマン輸送方程式に基づく熱電特性の第一原理計算を行い、Inverse Jonker plot を作成した。多くの熱電特性予測論文で行われているように、電子緩和時間を一律に  $10^{-14}$  s として計算した。機械学習から提案された候補化合物に対して第一原理計算を行い比較したところ、いくつかの材料系では Si を超える高い熱電特性が予測され、有望な熱電材料の候補だと考えられた。

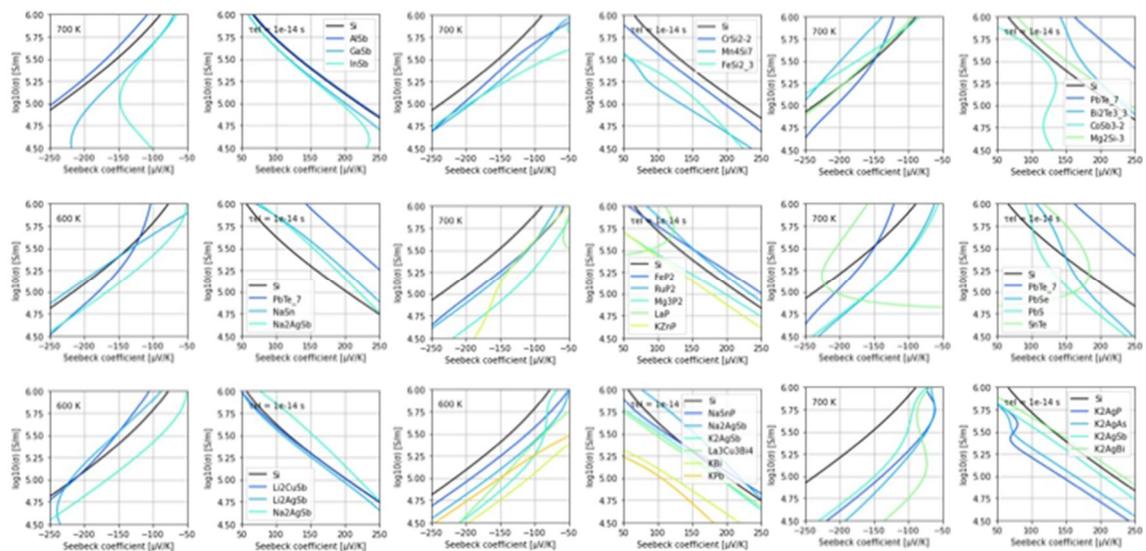


Figure 4. 第一原理計算とボルツマン輸送方程式を用いて作成した熱電材料候補物質の Inverse Jonker plot の例。

続いて、これらの予測が実際の実験データと合うのかどうか検証を試みた。Figure 5 は、さまざまな熱電材料の Inverse Jonker plot を、本研究で計算した熱電特性の第一原理計算結果と比較したものである。色は温度を表し、青が 300 K, 赤が 800 K を示す。すると、計算データと実験データの分布に大きな乖離が見られることがわかった。このうち上下方向のシフトのみでデータが一致する系は、第一原理計算結果が実験データと概ね一致していて、緩和時間のみの補正で予測が成立するものである。多くの材料系においてこのような傾向が見られていることから、材料系ごとの緩和時間の基本的な値を定義することで、概ね再現できることがわかつ

た。一方、PbTe 系や  $\text{Mg}_3\text{Sb}_2$  など一部の材料系では、Inverse Jonker plot の傾き自体が変化している様子が観察された。この原因として、温度変化に伴う熱膨張やドーピングによる格子定数変化がもたらすバンド構造の変化(Band convergence: バンド縮重)が関わっている可能性もあり、高い熱電特性と関係している可能性もあるため、今後の研究による検証が期待される。

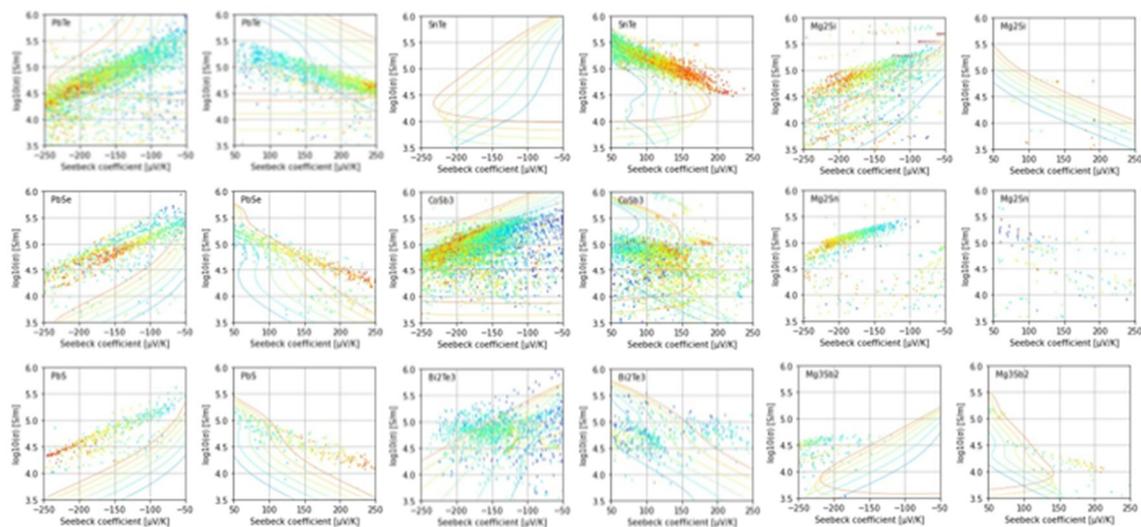


Figure 5. 第一原理計算による熱電材料の実験データの分布と第一原理計算結果の比較。

以上本研究では、第一原理計算による熱電特性予測と、論文から収集した大規模熱電材料データの実験値 Materials Informatics を組み合わせることにより、有望な熱電材料の探索と、熱電特性の予測技術の開発に取り組んだ。この結果、Inverse Jonker plot や ZT の 1 変数表示を用いた解析が有望であることを見出し、データに基づく最大 ZT の予測技術を開発した。これまでの緩和時間一定近似の第一原理計算結果は、直接実験データを再現することは困難であることがわかった。この予測結果は、材料系ごとの緩和時間の値を実験データを用いて補正することと、バンド縮重効果の考慮によって改善できることを見出した。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Katsura Yukari, Kumagai Masaya, Kodani Takushi, Kaneshige Mitsunori, Ando Yuki, Gunji Sakiko, Imai Yoji, Ouchi Hideyasu, Tobita Kazuki, Kimura Kaoru, Tsuda Koji	4. 巻 20
2. 論文標題 Data-driven analysis of electron relaxation times in PbTe-type thermoelectric materials	5. 発行年 2019年
3. 雑誌名 Science and Technology of Advanced Materials	6. 最初と最後の頁 511 ~ 520
掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/14686996.2019.1603885	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計17件（うち招待講演 8件／うち国際学会 7件）

1. 発表者名 Yukari Katsura, Masaya Kumagai, Takushi Kodani, Kaoru Kimura, Koji Tsuda
2. 発表標題 Collective analysis of experimental thermoelectric properties of the published plot data in Starrrydata web system
3. 学会等名 5th Asian Conference on Thermoelectrics（招待講演）（国際学会）
4. 発表年 2020年

1. 発表者名 桂 ゆかり, 熊谷 将也, 小谷 拓史, 佐藤 陸, 木村 薫, 津田 宏治
2. 発表標題 Starrydata webシステムによる熱電材料のMaterials Informatics
3. 学会等名 第81回応用物理学会秋季学術講演会
4. 発表年 2020年

1. 発表者名 桂 ゆかり, 熊谷 将也, 小谷 拓史, 木村 薫, 津田 宏治
2. 発表標題 Starrydata の熱電特性データのInverse Jonker プロット
3. 学会等名 2021年春季 第68回 応用物理学関係連合講演会
4. 発表年 2021年

1. 発表者名 Yukari Katsura, Masaya Kumagai
2. 発表標題 Starrydata: an open database of experimental data from published plot images
3. 学会等名 Open Databases Integration for Materials Design (招待講演) (国際学会)
4. 発表年 2020年

1. 発表者名 Yukari Katsura
2. 発表標題 Starrydata: an open database for materials informatics of experimental plot data
3. 学会等名 NSF-JST joint workshop (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 桂 ゆかり
2. 発表標題 Starrydata: 論文中のグラフから集めた材料実験データのデータベース
3. 学会等名 物性科学におけるデータ科学の今と未来 (招待講演)
4. 発表年 2021年

1. 発表者名 桂ゆかり
2. 発表標題 熱電材料の大規模実験データのMaterials Informatics
3. 学会等名 名古屋工業技術協会研究会
4. 発表年 2020年

1. 発表者名 Yukari Katsura, Masaya Kumagai, Mitsunori Kaneshige, Yuki Ando, Sakiko Gunji, Riku Sato, Takushi Kodani, Ruonan Ni, Yoji Imai, Hideyasu Ouchi, Kaoru Kimura, Koji Tsuda
2. 発表標題 Starrydata: an open database for experimental data extracted from published plot images
3. 学会等名 The 38th International Conference of Thermoelectrics / 4th Asian Conference on Thermoelectrics_ (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Masaya Kumagai, Yukari Katsura, Riku Sato, Ruonan Ni, Takushi Kodani, Hideyasu Ouchi, Sakiko Gunji, Yuki Ando, Yoji Imai, Kaoru Kimura, Koji Tsuda
2. 発表標題 Thermoelectric Materials Design by Materials Informatics using Enormous Experimental Data on Starrydata
3. 学会等名 The 38th International Conference of Thermoelectrics / 4th Asian Conference on Thermoelectrics_ (国際学会)
4. 発表年 2019年

1. 発表者名 桂 ゆかり, 熊谷 将也, 佐藤 陸, 安藤 有希, 郡司 咲子, 今井 庸二, 金重 光則, 小谷 拓史, 木村 薫, 津田 宏治
2. 発表標題 Starrydata熱電特性データベースの材料インフォマティクスへの応用
3. 学会等名 第15回日本熱電学会学術講演会
4. 発表年 2019年

1. 発表者名 熊谷 将也, 桂ゆかり, 佐藤 陸, 安藤 有希, 郡司 咲子, 今井 庸二, 小谷 拓史, 木村 薫, 津田 宏治
2. 発表標題 大規模論文データの機械学習による実験的熱電特性の予測
3. 学会等名 第15回日本熱電学会学術講演会
4. 発表年 2019年

1. 発表者名 桂 ゆかり
2. 発表標題 新規熱電材料探索の道のり～実験・計算・データ科学への挑戦～
3. 学会等名 第15回日本熱電学会学術講演会（招待講演）
4. 発表年 2019年

1. 発表者名 桂 ゆかり, 熊谷 将也, 佐藤 陸, 安藤 有希, 郡司 咲子, 今井 庸二, 金重 光則, 小谷 拓史, 木村 薫, 津田 宏治
2. 発表標題 論文からの大規模実験データの機械学習による熱電特性の直接予測
3. 学会等名 第80回応用物理学会秋季学術講演会（招待講演）
4. 発表年 2019年

1. 発表者名 桂 ゆかり, 熊谷 将也, 金重 光則, 安藤 有希, 郡司 咲子, 今井 庸二, 佐藤 陸, 小谷 拓史, 木村 薫, 津田 宏治
2. 発表標題 Starrydata: an open database of published experimental data
3. 学会等名 NIMS WEEK 2019
4. 発表年 2019年～2020年

1. 発表者名 Yukari Katsura, Masaya Kumagai, Mitsunori Kaneshige, Yuki Ando, Sakiko Gunji, Yoji Imai, Riku Sato, Takushi Kodani, Kaoru Kimura, Koji Tsuda
2. 発表標題 Starrydata: a plot-mined database for published experimental data
3. 学会等名 The 3rd_Forum of Materials Genome Engineering（招待講演）（国際学会）
4. 発表年 2019年

1. 発表者名 Yukari Katsura, Masaya Kumagai, Mitsunori Kaneshige, Yuki Ando, Sakiko Gunji, Yoji Imai, Riku Sato, Takushi Kodani, Kaoru Kimura, Koji Tsuda
2. 発表標題 Starrydata: an open database of published plot data for materials in formatics
3. 学会等名 Materials Research Meetings 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 桂 ゆかり, 熊谷 将也, 小谷 拓史, 佐藤 陸, 木村 薫, 津田 宏治
2. 発表標題 大規模実験データの機械学習による熱電特性予測
3. 学会等名 2020年春季 第回 応用物理学関係連合講演会
4. 発表年 2020年

〔図書〕 計2件

1. 著者名 木村薫, 桂ゆかり	4. 発行年 2020年
2. 出版社 シーエムシー出版	5. 総ページ数 20
3. 書名 次世代熱電変換材料・モジュールの開発ー熱電発電の黎明ー 第I編 第1章	

1. 著者名 桂ゆかり	4. 発行年 2020年
2. 出版社 シーエムシー出版	5. 総ページ数 19
3. 書名 次世代熱電変換材料・モジュールの開発ー熱電発電の黎明ー 第II編 第3章第6節	

〔産業財産権〕

〔その他〕

Starrydata  
https://starrydata.sakura.ne.jp/wp/  
Starrydata2 web system  
https://www.starrydata2.org/  
Starrydata  
https://www.starrydata2.org  
Starrydata  
https://starrydata.wordpress.com  
Starrydata project  
https://starrydataproject.wordpress.com

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------