

科学からSDGsへの道程を 導出する方法の提案

A Method to Plan a Path to the SDGs from Sciences

武田 晴夫 | Takeda Haruo

科学の全体を近似すると考えられるデータから、SDGsにより早く至る可能性がある道程を、客観的な計算機アルゴリズムによって導出する方法を提案する。本研究では、科学の全体を、過去公表された学術論文の全体であると考え、その普遍的な全体構造を近似すると考えられる論文群の部分集合で表現する。本稿ではその最初の近似として、現時点から50年前以前に発刊され、現在も定期的に刊行されている（したがって50年以上持続している）英文論文誌約200を選び、その2019年9月時点での最新号に収録されている全論文約5,000件を、提案方法の説明のための例として用いる。本研究では、導出する道程は仮想的論文の列で表現するものとするが、本稿では科学とゴールの各組について、既存論文のタイトルに異分野の技術用語を加えたときにSDGsに最も近づく単語10個をその最初の近似として求める。具体的には：

第一に上記論文集合のクラスタリングを行う。科学のクラスタを行とし、SDGsの17ゴールを列とする二次元行列を、正方行列にするために、科学クラスタ数を定数17に設定する。本研究では同一論文誌の同一号に掲載されている論文はすべて同一の科学クラスタに属していると仮定し、上記最新号に収録されている全論文につき、その英文論文タイトルおよび雑誌名称を構成する単語の集合をクラスタリング要素の「文献」とする。文献情報のクラスタリングアルゴリズムは「k-means」、そのために各文献を固定次元ベクトルに変換・表現する手法は「LDA」、そのときに用いる2ベクトル間の距離は「コサイン類似度」を、本稿では使用した。以上いずれも既存技術である。LDAのベクトル空間次元数(トピック数)は、k-meansによって得た各クラスタの要素数の分散が最小になる数を最適化計算により求めた。

第二にSDGsの17ゴールについて上記17科学と同様のベクトル化を行う。ベクトル化に用いる手法と距離の定義は科学に用いた手法と同一とする。SDGsの各ゴールは、ゴール自身の記述に、169ターゲットの当該ゴールに関する記述を加えた記述文を構成する英単語の集合で定義する。そのように求めたSDGsの17の各ゴールのベクトルについて、17の各科学クラスタの重心ベクトルとの距離を計算する。

第三に、各科学クラスタと各SDGsゴールの17×17の距離行列を目的関数、元の論文集合を初期点、SDGs記述文を目的点とする経路探索を行う。特に本稿では、科学とSDGsの17×17の各組につき、科学の既存論文のタイトルに新たな単語を加えることによって上記SDGsへの距離が最も近づくものをその道程と定義し導出した。このとき、SDGsの記述文自身が論文タイトルとして導出される自明道程を回避するために、加える単語は元の約5,000科学文献タイトルに2回以上現れたものの4,343語中から選択した。

本実験結果の範囲で科学各クラスタとSDGs各ゴールの関係について考察を加えた。過去に公表された学術論文の全体は数千万件以上とされ、本稿の約5,000件での近似は1/10,000以下の標本からの推定のレベルであり、有意なデータ量になっていない可能性がある。ただし本実験によるクラスタリング結果、科学とSDGsの距離行列、導出された科学からSDGsへの道程に筆者の主観に合致する結果もみられることから、本提案方法が科学の今後の応用に何らかの示唆を与える可能性があることが確認でき、大規模データの実験に進む動機が得られた。

1. はじめに

SDGs (Sustainable Development Goals) は、人類世界の持続可能な開発・発展のための2016年から2030年までの世界目標であり、2015年の国連サミットで採択されたものである。全体は17のゴール、169のターゲット、232のインジケータなどから構成される¹⁾。この目標を達成するために科学・技術・イノベーション (STI: Science, Technology and Innovation) への期待も大きいことから、「STI for SDGs」の名の下に各種の活動が世界で営まれている。日本では工学アカデミーにSDGsプロジェクトが設けられた²⁾。筆者はそのリーダーを務め、AAAS (American Association for the Advancement of Science: 米国科学振興協会) の年次大会に設けられたSDGsワークショップや、ニューヨーク国連本部で開催されたSTI Forumの一ワークショップなどで講演を行った³⁾。その後SDGsに向けたSTIのロードマップの策定も開始されている⁴⁾。

一方近年、学術文献データの計算機による解析の活動が、情報技術や人工知能の発展により進展している。その原点の一つに1955年に発表されたcitation indexの提案があり⁵⁾、これも含めて近年に至るまでのこの分野の研究のサーベイも発表されている⁶⁾。既発表論文の分析のみならず、分析結果を今後の研究の方向の指針に活用しようとする試みも行われており、例えば論文の引用、被引用の関係のネットワーク (citation networks) の構造分析から新興研究領域の発見をめざした研究⁷⁾ などがある。特にSDGsについて、「sustainabl*」等のキーワードで絞り込まれた学術論文の集合について解析を行った結果が最近発表されている⁸⁾。

これらの従来研究に対して、本稿の第一の目的は、学術文献データの客観的な計算機アルゴリズムによる解析によって、科学からSDGsの実現に向けての指針を、何らかの形で得られる可能性があるかを確認することにある。第二の目的は、計算機による近年の言語解析能力を人間直観による評価が有効に機能すると思われる科学技術全体のマクロ分析で検証することにある。第三の目的は、筆者が本誌冒頭の別稿³⁾で述べた俯瞰的研究開発の新たな方法論について実例を示すことにある。

2. 科学の再分類

本研究では、科学の全体を、過去公表された学術論文の

全体であると考え、その普遍的な全体構造を近似すると考えられる論文群の部分集合で表現する。科学の大分類、大分野ごとにSDGsへの道程を俯瞰的に導出することをめざすために、最初にその論文部分集合のクラスタリングを行う。

図書全体の人間による分類法としては、1876年に刊行されたデュイの十進分類法⁹⁾がある。その後、これを基に1905年に国際十進分類法¹⁰⁾が開発され、世界の図書館等で広く使われている。日本では1929年、日本十進分類法¹¹⁾が開発され、日本の図書館等で現在も広く使われている。いずれの分類法も今日も改訂が続けられている。一方、技術論文等を主対象とした二次情報データベースが世界に複数存在しており、それぞれの収録記事について、その作成機関が独自に科学分類を作成している。主要な技術論文データベースと分類、その特徴について、サーベイも発表されている¹²⁾。本稿は、これらの既存の分類法に対して、内容に関しての先見知識を極力使わずに、客観ルールのみによって、特に過去から最新の論文も含めた科学の全体について、計算機アルゴリズムでクラスタリングを試みるものである。

最初に学術論文の全体の構造を近似すると考えられる論文群の部分集合を定める。本稿では原理実験のための最初の近似として、現時点から50年前以前に発刊され、現在も定期的に発行されている（したがって50年以上持続している）英文論文誌約200を選び、その2019年9月時点での最新号に収録されている全論文約5,000件を、提案方法の説明のための例として用いる。次にその論文部分集合のクラスタリングを行う。科学のクラスタを行とし、SDGsの17目標を列とする二次元行列を、正方行列にするために、クラスタ数を定数17に設定する。本稿では同一論文誌の最新号に収録されている全論文は同一の科学クラスタに所属していると一旦仮定し、同一論文誌の同一の号に収録される全論文につき、英文タイトルを構成する単語の集合をクラスタリング要素の「文献」とする。

まず各文献を固定次元多次元ベクトルに変換・表現する。一般に文章情報のベクトル化のための手法として、ニューラルネットワークにより、注目単語の周囲に現れる単語を用いて単語を特徴付けベクトル化する「Word2Vec」と呼ばれる手法¹³⁾や、これを複数文章に拡張した「Doc2Vec」と呼ばれる手法¹⁴⁾、文書は潜在的な「トピック」の組から確率的に生成されるとのモデルに基づく「LDA」と呼ばれる手法¹⁵⁾、文書中に出現する各単語の重要度を用いる「tf-idf」と呼ばれる手法¹⁶⁾などが提案され、近年広く実問題で活用されている。本稿では先見知識や与えるパラメータへの依存などの観点から、各文献を固定次元ベクトルに変換・表現する手法は「LDA」を使用した。なおLDAのベクトル空間次元数 (トピッ

ク数)は、ここでは変数とし、後述のクラスタリングによる各クラスタの要素数の分散を目的関数とする最適化を行い、最終的にその分散が最小となる値を採用した。ベクトル化に先立って、動詞の時制、名詞の単数／複数等から語幹を抽出するSTEMMING処理と、前置詞他、多くの文章に頻繁に現れ、文章の特徴を表すベクトルに関係が薄いと思われるストップワードの除去を前処理として行った。

多次元空間での標本のクラスタリングの一般手法として、各標本を初期クラスタに割り当て、各クラスタ重心を計算してそれに最も近い標本を新クラスタに割り当てる処理を繰り返す方法が提案され¹⁷⁾、「k-means」法と命名されている¹⁸⁾。また標本とクラスタの関係に確率分布を仮定する方法¹⁹⁾も、特に近年の計算機性能の向上によりGMM (Gaussian mixture model)²⁰⁾と呼ばれる方法を中心に広く使われるようになっている。本稿では、クラスタ数指定のために、前者のk-means法を用いた。初期クラスタは乱数のseedを固定することにより、実験の再現性に留意した。またこのとき、二つのベクトル間の多次元空間での距離は「コサイン類似度」(2ベクトル間の角度の余弦)を用いた。

表1に実験結果の一部を示す。列「学術論文誌」の誌名および最新号各論文タイトルに対して、上記のクラスタリングを行った結果の17クラスタの番号を列「クラスタ」に、そのクラスタに所属する論文誌数を列「要素数」に示す。「クラスタ」の番号は識別子であり、数字に量的、順序の意味はない。各クラスタについて、便宜上筆者の主観でラベルを付し列「ラベル」に記した。このときラベルは、ESI (Essential Science Indicators)¹²⁾で用いられている大分類の数が22と、本クラスタ数の17に比較的近いため、これを参考に付した。

本表に示す比較的に筆者の主観に合致した分類以外に、明らかに主観や既存の科学分類からは違和感あるクラスタと論文誌の組の関係も存在した。これについては、選択する論文誌および論文の数を拡大することにより、少なくとも本研究の目的である各科学クラスタからSDGsへの道程の導出には無視できる程度に小さくできる可能性があると考えている。

3. 科学とSDGsの距離

次に前章で求めた17科学の各クラスタとSDGsの各ゴールとの距離を求める。まず、SDGsの17の各ゴールについて、そのゴール自身の英文記述に、169ターゲットのうち当該ゴールに関する英文記述の文章を加えた記述文を構成する英単語の集合を当該ゴールの文献と定義する。そのベクトルを、前章で述べた手法と同じ距離定義、アルゴリズムおよび同じパラメータを用いて求める。次に各科学クラスタについて、当該クラスタに所属する文献のベクトルの重心を求める。SDGsの17ゴール文の各ベクトルについて、科学の17クラスタの各文献ベクトルの重心との距離を、上記と同様にコサイン類似度を用いて計算して、17×17行列として求める。

表2に前章で記述の近似データについて処理した結果を示す。本表は前章で定義した科学1～17の各文献ベクトルの重心に対して、SDGsのゴール1～17の各文献(ゴール記述および169のターゲット記述のうち対応するゴールのもの)とのベクトルの間の距離の17×17行列である。距離の定義は2ベクトル間の角度の余弦であるコサイン類似度であり、1が距離最小(最近接)、0が最大(最遠隔)である。特にSDGsの各ゴールに対して、距離最近接(コサイン類似度最大)と

表1| 科学の再分類

クラスタ	学術論文誌	要素数	ラベル
1	Journal of Algebra	11	数学他
1	Journal of Number Theory		
1	Linear Algebra and its Application		
1	Journal of Mathematical Analysis and Applications		
1	Progress in Aerospace Sciences		
1	:	17	材料他
2	Materials Research Bulletin		
2	Journal of Colloid and Interface Science		
2	Chemical Engineering Science		
2	:	11	環境他
3	Environmental Research		
3	Sedimental Geology		
3	Experimental Parasitology		
3	:		
:	:	:	:

なるセルの数値を表中太字で示した。そのような科学とSDGsは1:1対応に一般にはならないが、本実験では17ゴールに対して最大貢献する科学が17科学中8科学、さらに貢献度上位3科学をとると13科学となり、特定科学分野だけでなく科学全体がSDGs全体に有意な貢献をする可能性があるとする一般論を定量的に説明しているともいえる。

この実験結果において、表1で数学他とラベル付けされた科学クラスタ1は、ゴール15の陸上資源、ゴール2の飢餓、ゴール6の水・衛生に比較的近い関係にはあるが、いずれのゴールにおいても他にさらに近い科学が存在し、特定のゴールに最近接した関係は本実験では観測されていない。逆にゴール4の教育と最も遠い関係となっているが、これはSDGsのすべての人への教育に向けた記述と、数学他の学会最先端の研究論文の記述の距離を示す妥当な結果とみられることもできる。ただし次章での実験のとおり、各科学クラスタの各論文にそれぞれ特定の技術用語を加えることによって距離が縮小する程度が、他の科学より大きい結果となっている。距離が大きい科学ほど、異分野科学との融合によりSDGsにより近づく可能性を唆しているとも考えられる。材料他とラベル付けされた科学クラスタ2は、17中15のゴールで数学他とラベル付けされた科学クラスタ1よりSDGsに近接した結果となっている。特に、ゴール6の水・衛生、ゴール11の都市、ゴール10の不平等などで顕著に近い関係が示されている。環境他とラベル付けされた科学クラスタ3は、ゴール3の保健に続き、ゴール16の平和、ゴール17の実施手段等に近い関係が示されている。

他方、SDGsのゴールから科学をみると、例えばSDGsのゴール7のエネルギーやゴール9のイノベーションなど、科学の貢献可能性が比較的高いと筆者には思われる領域において、科学とSDGsが特に近接していると示されない結果となっている。このような分野においては上記の基礎科学で述べたと同様に、経路計画によってSDGsへの距離が近づく観点から本手法がより有効に働くともみられる一方、本稿の最初の近似データにおける論文誌および論文の数が限界以下であり有意な結果を生むに至っていない可能性も否定できない。

4. 科学からSDGsへの道程

前章で求めた科学各クラスタとSDGs各ゴールの17×17の距離行列を目的関数とし、元の論文集合を初期点、SDGs記述文を目的点とする経路探索を行う。

一般に多次元ベクトル空間において初期点から目的点への経路計画については、目的関数や、障害物存在などによる不能解を回避するための制約条件など、問題特質に応じて、特に次元数が大きい場合のヒューリスティクスの加え方を中心に、種々のアルゴリズムが提案されている²¹⁾。特に経路上各点での自己位置推定の不確かさ（uncertainty）が定量化できる場合の最適経路計画も開発されている²²⁾。

本研究で導出する道程は仮想的論文の列で表現するものとするが、本稿では特に科学とSDGsの17×17の各組につ

表2| 科学とSDGsの距離

距離定義はコサイン類似度で、1が最近接、0が最遠隔を意味する。

		SDGs																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
科学	1	.33	.61	.37	.26	.33	.50	.36	.41	.39	.43	.42	.34	.39	.46	.67	.38	.33
	2	.70	.70	.62	.69	.71	.82	.51	.70	.33	.76	.77	.63	.71	.67	.63	.54	.63
	3	.47	.43	.69	.47	.52	.53	.51	.60	.61	.57	.58	.61	.59	.55	.56	.66	.65
	4	.68	.68	.65	.62	.66	.74	.52	.66	.43	.71	.77	.59	.73	.76	.60	.44	.62
	5	.52	.36	.49	.48	.42	.57	.40	.47	.46	.61	.60	.45	.51	.49	.59	.61	.44
	6	.54	.59	.68	.64	.64	.56	.77	.70	.54	.62	.61	.80	.71	.57	.61	.76	.83
	7	.62	.61	.83	.70	.62	.51	.62	.61	.55	.60	.68	.54	.64	.56	.57	.89	.50
	8	.56	.64	.55	.55	.55	.65	.51	.59	.53	.56	.63	.58	.62	.63	.76	.60	.56
	9	.67	.48	.73	.74	.65	.49	.77	.70	.49	.56	.73	.74	.75	.78	.63	.75	.65
	10	.38	.79	.49	.42	.44	.46	.49	.41	.60	.51	.43	.43	.50	.40	.67	.49	.59
	11	.82	.67	.75	.86	.82	.88	.72	.85	.35	.71	.86	.88	.86	.78	.68	.57	.80
	12	.77	.70	.83	.80	.89	.79	.66	.87	.54	.91	.86	.77	.80	.72	.64	.69	.79
	13	.31	.42	.41	.30	.27	.48	.55	.38	.40	.43	.36	.47	.44	.39	.45	.40	.60
	14	.61	.40	.60	.54	.56	.56	.55	.59	.71	.56	.63	.61	.62	.53	.66	.51	.61
	15	.36	.38	.36	.41	.27	.35	.31	.27	.26	.53	.40	.18	.38	.37	.34	.49	.27
	16	.36	.47	.33	.43	.38	.38	.60	.42	.31	.38	.43	.47	.39	.47	.65	.57	.35
	17	.50	.50	.58	.51	.48	.56	.48	.54	.60	.55	.53	.64	.67	.47	.51	.57	.80

き、科学の既存論文のタイトルに異分野の技術用語を新たに加えたときに当該SDGsゴールへの距離が最も近づくもの上位10個をその最初の近似として定義した。ただし、SDGsの定義文章自身が論文タイトルなどとして導出される自明解は、科学技術論文としての主観検証が必要なことから、これを回避するために、加える単語は元の約5,000科学文献タイトルのいずれかに現れたものの中から選択するものとした。本実験においては、2章で記載のステミング処理およびストップワードの除去処理後の重複を許さない全単語10,054語のうち、2度以上現れていた4,343語について、全数探索を行った。

表3に実験処理結果の一部を示す。本表において、科学1, 2, 3は17科学中それぞれ数学他, 材料他, 環境他と表1でラベル付けした科学である。「ゴール距離」は、現在の科学からSDGsまでの距離であり、表2の当該行を対応する列にそのまま転記したものである。「道程」は、論文タイトルに他の科学分野の論文タイトルに現れた技術用語を加えたときに、ゴール距離が縮減する程度が大きいものから上位10個を求め、基本的にはその最上位用語を示した。本表では3×17=51中42単語がこれに相当する。最上位単語が筆者の主観において説明困難であった3個について次点のものに置換した。下記のように最上位が同一の単語となる重複が多数発生したが、その中で次点には特徴的な単語が存在するケースが3例あり、この場合その3単語に置換した。他に当該科学、当該ゴールを筆者の主観においてよく表していると思われる次点単語が存在するケースが3例あり、これも置換した。以上いずれのケースにおいても置換は、最上位単語と置換単語の距離削減効果の差異が十分に小さい場合に

限った。「移動距離」は各単語を加えたときに変化するゴールまでの変化距離から、元の「ゴール距離」を単純に減じた値である。これが大きいほど、研究戦略によりSDGsへの貢献可能性が高まる分野と仮定する。

科学1の数学他は、前章記載のように全体にSDGsへの距離が大きい³、特定の技術用語を異分野から加えることによって、その距離が大きく縮まる傾向がみられる。例えばゴール6の水・衛生、ゴール4の教育、ゴール5のジェンダー、ゴール12の生産・消費、ゴール1の貧困において、0.20以上と、本表の他科学においては存在しない大きな距離削減がみられる。科学2の材料他は、前章記載のように全体にSDGsへの距離が小さい³、何らかの技術用語1語を異分野から加えても、その距離がほとんど縮まらない傾向がみられる。ただし距離縮減の度合いが、その中でも大きい分野はゴール9のイノベーション、ゴール7のエネルギー、ゴール12の生産・消費、ゴール15の陸上資源となっており、筆者の主観におけるSDGsへの材料科学貢献期待が高い分野によく符合した結果となっている。科学3の環境他は、全体に科学2の材料他と類似の傾向がある。ただし距離縮減の度合いが大きい分野は、ゴール14の海洋資源、ゴール6の水・衛生、ゴール11の都市、ゴール4の教育、ゴール1の貧困などとなっており、上記材料他とはまったく異なる分野への貢献可能性が示唆されており、かつそれらの分野が筆者の主観にもよく符合した結果となっている。

表中に道程として示した単語例については、数学他および環境他におけるaustenit、材料他におけるmpa（メガパスカル）、環境他におけるofferなど、多くのゴールに対して効果

表3 | 科学からSDGsへの道程

SDGs	科学1 (数学他)			科学2 (材料他)			科学3 (環境他)		
	ゴール距離	道程	移動距離	ゴール距離	道程	移動距離	ゴール距離	道程	移動距離
1 (貧困)	.33	austenit	.20	.70	lentivirus	.03	.47	offer	.11
2 (飢餓)	.61	phenomenon	.17	.70	pyrolysi	.03	.43	lymph	.08
3 (保健)	.37	inclin	.16	.62	mwcnts	.02	.69	offer	.10
4 (教育)	.26	austenit	.25	.69	nanoarray	.03	.47	offer	.11
5 (ジェンダー)	.33	austenit	.21	.71	nanoarray	.03	.52	austenit	.01
6 (水・衛生)	.50	austenit	.27	.82	disc	.02	.53	austenit	.15
7 (エネルギー)	.36	preferenti	.16	.51	mpa	.05	.51	greenhous	.06
8 (成長・雇用)	.41	austenit	.18	.70	nanoarray	.03	.60	austenit	.10
9 (イノベーション)	.39	glioblastoma	.18	.33	glioblastoma	.06	.61	plus	.05
10 (不平等)	.43	inclin	.14	.76	modular	.03	.57	austenit	.08
11 (都市)	.42	austenit	.18	.77	patent	.02	.58	austenit	.11
12 (生産・消費)	.34	marriag	.20	.63	mpa	.05	.61	austenit	.08
13 (気候変動)	.39	austenit	.16	.71	patent	.03	.59	offer	.09
14 (海洋資源)	.46	chute	.13	.67	mpa	.04	.55	offer	.09
15 (陸上資源)	.67	preferenti	.12	.63	mpa	.05	.56	distinguish	.01
16 (平和)	.38	sensori	.08	.54	mpa	.03	.66	reward	.01
17 (実施手段)	.33	partner	.16	.63	clearanc	.04	.65	disc	.05

があるとするもの、ゴール9のイノベーションにおいて3科学共通に特徴的に上位に現れるglioblastomaなどがある（同語は科学3においてもplusの次点であった）。一方ゴール1の貧困、ゴール2の飢餓に向けての材料他に対するlentivirusやpyrolysiなど、ゴール3の保健、ゴール4の教育、ゴール5のジェンダーに向けての材料他に対するMWCNTs (multi-walled carbon nanotubes) やnanoarray、ゴール7のエネルギーに向けての環境他のgreenhous、ゴール16の平和やゴール17の実施手段に向けての数学他に対するsensoriやpartnerなど、特徴的かつ説明がある程度可能と思われる単語もみることができる。これらの分析についても、本稿の最初の近似データにおける論文誌および論文の数が限界以下であり、有意な結果を生むに至っていない可能性も否定できないが、十分なデータ量での実験がさまざまな分析と研究方向性の示唆を与える可能性が得られたものとする。

5. おわりに

科学からSDGsへの道程を導出する方法として、過去の学術論文の全体を近似すると考えられる文献の集合を17の分野に計算機アルゴリズムでクラスタリングし、その科学17分野とSDGs 17ゴールの間の距離の17×17行列を計算し、これを目的関数として今後SDGsにより早く向かう可能性がある論文の経路を探索する方法を提案した。方法の確認のために、約5,000件の論文で実験を行い、実験結果に対して筆者の主観による考察を加えた。

過去に公表された学術論文の全体は数千万件以上とされ、本稿の約5,000件での最初の近似は1/10,000以下の標本からの推定のレベルであり、有意な結果を生むデータ量になっていない可能性がある。ただし本実験の結果にみられる、クラスタリング結果、科学とSDGsの距離行列、導出された科学からSDGsへの道程に筆者主観に合致する結果も少なからずみられることから、本提案方法が実応用で何らかの示唆を与える可能性があることが確認でき、大規模データの実験に進む動機が得られた。

参考文献

- 1) The 2030 agenda for sustainable development, The United Nations, A/RES/70/1 (2015)
- 2) 武田晴夫：SDGsへの日本工学アカデミーの取り組み、学術の動向, Vol.23, No.1, p.60~63 (2018)
- 3) 武田晴夫：SDGsと社会イノベーション、日立評論, 101, 6, 634~635 (2019)
- 4) A guidebook for the preparation of STI for SDGs roadmap, draft for consultation, The Fourth STI Forum, The United Nations, New York (2019)

- 5) E. Garfield: Citation indexes for science: a new dimension in documentation through association of ideas, Science, Vol.122, pp.108-111 (1955)
- 6) S. Fortunato et al.: Science of Science, Science, No.359 (2018)
- 7) N. Shibata: Detecting emerging research fronts based on topological measures in citation networks of scientific publications, Technovation No.28, pp.758-775 (2008)
- 8) K. Asatani et al.: Scientific research on sustainability: A meta-analysis of academic papers about sustainability and the Sustainable Development Goals, The 14th Conference on Sustainable Development of Energy, Water and Environment Systems, Croatia (2019)
- 9) M. Dewey: Classification and subject index for cataloguing and arranging the books and pamphlets of a library, Project Gutenberg eBook (1876)
- 10) Manuel du répertoire bibliographique universel, Institut International de Bibliographie (1905)
- 11) 森清：日本十進分類法 和漢洋書共用分類表及索引、間宮商店 (1929)
- 12) 小野寺夏生：論文データベースにおける主題分類、情報の科学と技術, Vol.66, No.6, p.272~276 (2016)
- 13) T. Mikolov, et al.: Distributed representations of words and phrases and their compositionality, Neural Information Processing Systems Conference, Lake Tahoe, NV (2013)
- 14) L. Quoc et al.: Distributed representations of sentences and documents, Proceedings of the 31st International Conference on Machine Learning, Beijing, China (2014)
- 15) J. Pritchard, et al.: Inference of population structure using multilocus genotype data, Genetics, Vol.155, No.2, pp.945-959 (2000)
- 16) K. Jones.: A statistical interpretation of term specificity and its application in retrieval, Journal of Documentation, Vol.28, No.1, pp.11-21 (1972)
- 17) H. Steinhaus: Sur la division des corps matériels en parties, Bulletin de l'Académie Polonaise des Sciences, Classe 3, pp.801-804 (1957)
- 18) J. MacQueen: Some methods for classification and analysis of multivariate observations, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol.1, pp.281-297 (1967)
- 19) G. McLachlan et al.: Finite mixture models, Jhon Wiley & Sons, Inc. (2000)
- 20) C. Rasmussen: The infinite Gaussian Mixture Model, Advances in Neural Information Processing Systems 12, MIT Press, MIT Press, pp.554-560 (2000)
- 21) J.-C. Latombe: Robot motion planning, Springer Science & Business Media (2012)
- 22) H. Takeda et al.: Planning the motions of a mobile robot in a sensory uncertainty field, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.16, No.10, pp.1002-1017 (1994)

執筆者紹介



武田 晴夫

日立製作所 研究開発グループ 所属

1980年東京大学工学部計数工学科数理工学専修コース卒業。在学中ニューラルネットワークの機械学習プログラムを囲碁を対象に作成。日立製作所入社以降、25年間一貫してICT／知的情報処理の研究開発に従事。2005年より同社研究戦略統括センター長、基礎研究所所長、技術戦略室長などを経て現在技師長としてAI研究に従事。以上の間、スタンフォード大学ロボティクス(AI)研究所客員研究員(2年間)、東大客員教授、NEDO次世代人工知能・ロボット中核技術開発次世代人工知能技術分野先端研究審査採択委員長、JST研究主幹(多国間共同研究プログラム 先端技術融合分野／インテリジェント社会インフラ分野)、国内外大学および国立研究機関アドバイザーボードメンバー、日本工学アカデミーSDGsプロジェクトリーダー、日本電機工業会総合政策委員長、IECボードメンバー(MSB)、関東工学会教育協会会長など、また内閣府、総務省、外務省、文科省、農水省、経産省、金融庁などにおいて諸委員会等々で有識者、構成員、委員、座長等を務める。博士(工学)。