

どのような文が一般社会とも共有されるのか？ — 文埋め込みを用いた人工知能関連テキストの分析 —

陳宇鍬 ホドシチェク ボル
大阪大学 人文学研究科
u572689a@ecs.osaka-u.ac.jp
hodoscek.bor.hmt@osaka-u.ac.jp

概要

人工知能技術が一般社会に浸透しつつある現在、関連テキストにおいてどのような文が一般向けの新聞とも共有され、一般社会を繋げているのかという問いに回答することが重要になっている。本研究では、文埋め込みを用いて専門家コミュニティの人工知能関連文を、新聞文と比較しながら高・低類似度に分けた。その後、BERTopicで高・低類似度文の内容差異を分析した。結果として、人工知能の産業・社会連関および人工知能自体の製品・進化・リスクを記述する文が一般社会とも共有されやすいと発見した。

1 はじめに

本研究の目的は、学術雑誌など専門家コミュニティ内の人工知能関連テキストではどのような文が一般向けの新聞とも共有され、一般社会を繋げているのかに回答することにある。また、専門家コミュニティのテキストがどのように一般社会まで繋がりを、さらにインパクトを与えうるのかというノウハウを言語分析の側面から提示することが本研究の価値である。

近年、人工知能の発展とともに、関連技術と製品は一般社会に浸透し続けている。特に、ChatGPTを代表とした人工知能技術と製品の登場につれて、関連情報は新聞にも取り上げられ、さらに広い一般社会に伝達されている¹⁾。このように、人工知能の技術は、閉ざされた専門家コミュニティという枠を超え、一般人の日常生活へと入り込みながら、一般社会にもインパクトを与える勢いを示している。

日本語研究の側面から見れば、人工知能関連テキストの分析は、科学技術日本語の視点から捉えるこ

とができる。科学技術日本語は、科学技術分野の内容を表す全ての日本語を含んでいる [1]。これまで、科学技術日本語は主に学術論文など専門家コミュニティ内部の日本語 [2, 3] として捉えられている。

一方、専門家コミュニティの枠を超えてみれば、より広い一般社会の中で科学技術情報を普及させるという一般向けの科学技術日本語も存在する。関連分析も散見される [4, 5]。しかし、人工知能などの新技術がラボから一般社会へ浸透し続ける現在、科学技術日本語のどの部分が一般社会と繋がりやすいのか管見の限りまだ十分に検討されていない。

本研究の特色は、専門家コミュニティの科学技術日本語だけにこだわらず、一般向けとの繋がりにも関心を向けることにある。具体的には、文埋め込みの技術を用いて、専門家コミュニティの人工知能関連テキストにおいてどのような文が新聞と共有されるのか、またどのような文が共有されず、ただ専門家コミュニティ内に残されるのかという問いを解明する。図1は本研究の図示である。

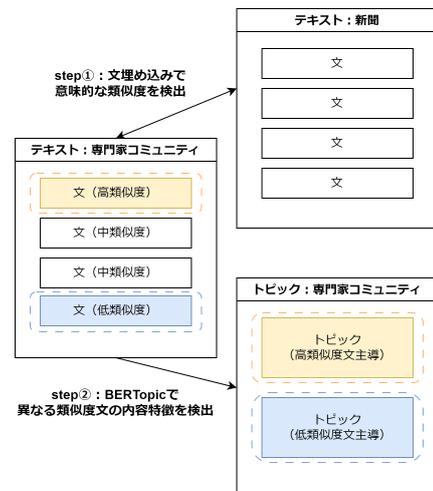


図1 本研究の図示

1) <https://yomidas.yomiuri.co.jp/yomiuri/articles/8727767> (ChatGPT 関連記事 2025 年 11 月 6 日最終アクセス)

2 関連研究

埋め込み (Embeddings) とは、自然言語における単語、フレーズ、文、あるいはより長いテキストを特定の空間における数値として表すことであり、その数値がテキストの意味を保持するよう設計されたものである [6]. 埋め込みの一種として、文埋め込み (Sentence Embeddings) は、文の意味をベクトル化し、コサイン類似度で意味的な類似文を探す手法を提供した [7].

文埋め込みは人文社会科学の課題、特に間テキスト性 (Intertextuality) または思想と知識伝達のルートを分析する課題に、新たなツールを与えた. このツールを利用した実証分析も既に行われている. Duan et al. は文埋め込みを利用し、中国歴史における異なる哲学流派のテキスト間の相互影響を実証した [8]. 哲学思想以外に、科学知識と観念の伝播分析にもこの技術が利用されている. Li は『種の起源』における思想伝播を追跡する課題に対して、文埋め込みで異なる学問分野への影響を分析した [9]. 上記の分析対象は日本語テキストではないが、日本語にも文埋め込みを利用する可能性を示唆した.

科学技術のテキスト分析についても、文埋め込みは意味的な要素を取り入れ、類似文を探すタスクに新たな可能性を開拓した. 例えば、Wright et al. は科学論文、ニュースおよびソーシャルメディアにおける科学知見パラフレーズのデータセット開発に文埋め込みを利用した [10]. Filippi-Mazzola et al. は文埋め込みで特許テキストの類似度を調査し、特許引用の特徴を分析した [11].

これらの関連研究を踏まえ、科学技術日本語の分析に文埋め込みというツールを導入する可能性が示唆された. そこで、本研究は、従来の学術論文など専門文書における特定表現の分析を中心とする科学技術日本語研究の枠組みを少し拡張し、文埋め込みの導入で一般向けの科学技術日本語および意味的な側面を研究射程に組み込む.

3 データ収集と前処理

本研究は、人工知能関連の理工・学際科学系専門文書を収集し、これを専門家コミュニティ内のテキストと見なしている. 一方、読売新聞と朝日新聞の記事を一般向けのテキストとして収集した. また、人工知能のトピックが盛んになった時期を考え、収集したテキストはすべて 2023 年刊行のものである²⁾.

具体的には、専門家コミュニティ内のテキストに対して、J-STAGE からタイトルまたはキーワードに「人工知能、AI、大規模言語モデル、LLM、深層学習、ディープラーニング、ニューラルネットワーク」のいずれかの単語が含まれる理工・学際科学系学術雑誌、会議論文・要旨集、研究報告・技術報告の日本語文書 (PDF ファイル) を収集し、テキストに変換した. また、一般向けのテキストに対して、読売新聞と朝日新聞の 2023 年記事 (学術研究用データベース) を利用している.

また、専門家コミュニティのテキストに対して、数式の部分を [MathExpr] に置換した. 文レベルでの分析を行うために、収集したすべてのテキストを文に分割した. なお、日本語文字を含まない文、および 10 文字以下の極端に短い文を除去した.

最終に収集した文書数と文数は下記の表 1 に示している.

表 1 データ集計

	文書数	文数
専門家コミュニティ	1,012	94,023
一般向けの新聞	314,757	3,921,236

なお、専門家コミュニティの PDF ファイルをテキストへ変換する過程において、文中に改行と画像キャプションなどが割り込み、文が分断される現象が存在する. この問題に対処するため、gpt-5-mini (ver. 2025-08-07) を用いて付録 A のプロンプトで処理を実施した.

4 データ分析

データ分析は、以下の 3 ステップから構成される.

- (1) 専門家コミュニティ文に対して、新聞文との類似度を検出する;
- (2) 専門家コミュニティ文を、高類似度文と低類似度文に分割する;
- (3) BERTopic で高類似度文と低類似度文の内容特徴の違いを分析する.

4.1 文埋め込みを用いた類似度検出

文埋め込みの計算には、cl-nagoya/ruri-v3-310m を使用した [12]. これは日本語テキストに特化した埋

2) <https://www.jiyu.co.jp/singo/index.php?eid=00040> (2023 年流行語大賞に「生成 AI」がランクイン 2025 年 12 月 30 日最終アクセス)

め込みモデルであり、JMTEB というベンチマーク上の Avg. 数値が 77.24 であり³⁾、同類の日本語埋め込みモデルでは高い性能を示している。

上記のモデルを用いて文埋め込みの数値を得た後、専門家コミュニティと一般向けの文埋め込みに対してコサイン類似度を算出した。

4.2 高低類似度文への分割

この段階では、専門家コミュニティの 94,023 文に対して、新聞文とのコサイン類似度が最も高い文ペア (Top-1) を取り出した。この手順により、専門家コミュニティ文、およびそれと対応する類似新聞文の文ペアを作成した。

また、上から得た類似文ペアを、類似度の四分位数で分類した後、上位 25 % と下位 25 % の文をそれぞれ、高類似度と低類似度の文ペアに分割した。次頁の表 2 と表 3 では、高類似度と低類似度における文ペアの例を示した。

4.3 トピックモデリング

高類似度文と低類似度文における内容特徴を分析するため、本節では Grootendorst が提案した BERTopic^[13] を採用し、文埋め込みに基づくトピックモデリングを実施した。

BERTopic に使う埋め込みでは前述処理で獲得した専門家コミュニティにおける高類似度文と低類似度文の文埋め込み表現を使用した。ハイパーパラメータの設定に関して、UMAP では `n_components` を 5、`min_dist` を 0.0、`n_neighbors` を 10 に設定し、HDBSCAN では、`min_cluster_size` を 40 に設定した。モデル実行後、`coherence` と `diversity` の指標がそれぞれ 0.65 と 0.84 であり、トピック内の語彙一致性とトピック間の多様性が保持されている。

また、BERTopic によるトピックモデリングから 147 トピックを獲得した。さらに、この中から高類似度文 (Q1) と低類似度文 (Q4) の含有率の差が最も大きい 20 トピックを抽出し、c-TF-IDF で得た各トピックの代表的な 10 単語を付録 B に載せた。これら 20 トピックは全トピックの 47.53 % を占めている。具体的な含有率分布を図 2 に示した。

図 2 から、0, 2, 4, 5, 6, 8, 9, 11, 12, 18, 22 番のトピックでは新聞と共有されやすい高類似度文 (Q1) が主導的役割を果たしていることが確認で

きる。一方、1, 3, 13, 16, 17, 23, 27, 28, 31 番のトピックでは、新聞と共有されにくい低類似度文が主導的である。また、Q1 と Q4 含有率分布の差から見れば、とりわけ 0, 1, 2, 3 番の差が大きく、それぞれ 0.14, 0.14, 0.08, 0.08 の差が存在している。なお、上記 20 トピックの例文を付録 B に載せた。

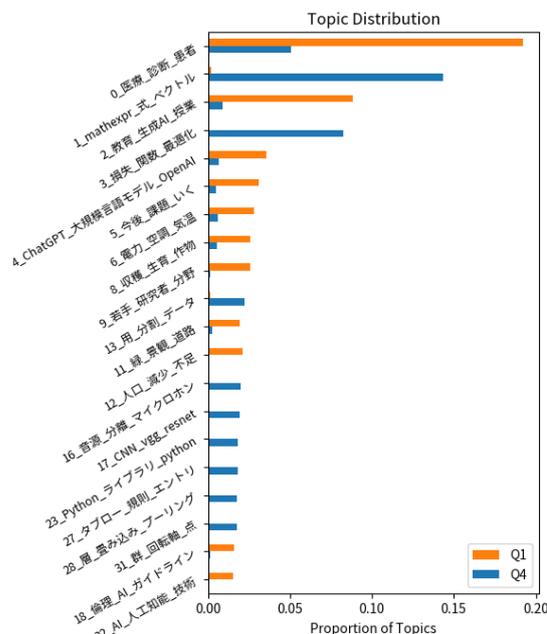


図 2 含有率差が最も大きい 20 トピック

また、トピックと文 (Q1/Q4) の対応分析を図 3 に載せた。

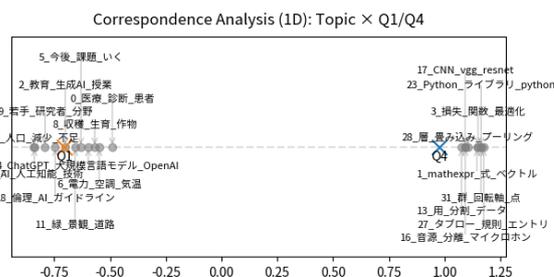


図 3 トピックと文 (Q1/Q4) の対応分析

5 分析結果

全体傾向として、新聞と共有されやすい文は、(1) 人工知能の産業・社会連関、(2) 人工知能自体の製品・進化・リスクを記述する内容に大別される。一方、新聞と共有されにくい文は、数式、モデル訓練過程、アーキテクチャの記述という類の文に分類できる。

このように、専門家コミュニティ内のテキストでは、人工知能の産業社会連関、技術進化とリスクに

3) <https://huggingface.co/cl-nagoya/ruri-v3-310m> (2025 年 12 月 31 日最終アクセス)

表2 高類似度（上位25%）文ペアの例（コサイン類似度 Top-1）

類似度	専門家コミュニティ	一般向けの新聞
0.98	電気は需要と供給の量を一致させないと周波数が乱れ大規模な停電事故につながる。〔出典:10.14936/ieiej.43.258〕	電気は需要と供給のバランスがとれていないと周波数が乱れ、大規模な停電につながる恐れがある。〔出典:読売新聞「東電 再エネ制御を検討 受け入れ 5月連休中に可能性」〕
0.97	しかしながら、実際のところ、AIが感情を持つことは可能なのだろうか。〔出典:10.11517/pjsai.JSAI2023.0_1K40S11a05〕	AIは感情を持つようになるのか。〔出典:朝日新聞「AIも訓練すれば「共感力」 リスクへの対策、国際協調を ヒントン氏インタビュー」〕
0.96	SF作家であり、生化学者でもあったアイザック・アシモフは、1950年に発表した『I Robot』（『われはロボット』）の冒頭で、「人間に危害を加えてはならない」、「人間の命令に従わなければならない」、「自己を守らなければならない」といった3原則を提示している。〔出典:10.20664/jabes.30.0_117〕	思い出すのは米国の生化学者・SF作家のアイザック・アシモフが1950年の作品の中で掲げた「ロボット工学三原則」だ。〔出典:朝日新聞「(声) AI開発も「ロボット三原則」で」〕

表3 低類似度（下位25%）文ペアの例（コサイン類似度 Top-1）

類似度	専門家コミュニティ	一般向けの新聞
0.83	この出力テンソルの点数方向に対して Max-Pooling により特徴量を集約し、256次元のベクトルに変換する。〔出典:10.11532/jsceiii.4.3_442〕	生産能力は2・5倍に向上するという。〔出典:読売新聞「半導体レーザー 生産棟を増設へ 浜松ホトニクス=静岡」〕
0.82	最後に、学習時の損失関数については、正規化した Focal Loss を用いることで、ハードネガティブサンプル対策を行う。〔出典:10.11517/pjsai.JSAI2023.0_3Xin401〕	それを防ぐための練習だ。〔出典:朝日新聞「侍の生命線、走塁を磨け 二塁から本塁へ、短期決戦へ決める徹底 野球・WBC」〕
0.81	連続時間力学系は、時間 t と状態 u に対して、主に常微分方程式 ODE[MathExpr] で記述される。〔出典:10.11517/pjsai.JSAI2023.0_2M6GS1006〕	記述は答えを求める過程を式や言葉で説明させるもの。〔出典:読売新聞「全日制3万4533人受験 公立高入試初日 追検査は来月1日=千葉」〕

関する記述が窓口として一般社会を繋げていると本研究によって明らかになった。また、日本の社会課題と関連する 12_人口_減少_不足 というトピックの文も一般向けの新聞に繋がりがやすいと確認できる。これらの結果は、産業革新や社会課題に関する内容が、専門家コミュニティと一般社会の接点形成に寄与しうることを示唆している。

一方、数式、モデル訓練過程、アーキテクチャの部分は、一般社会との共有傾向が低いことが示唆された。この傾向は、専門家コミュニティ内のテキストにおける専門性または閉鎖性を反映している。

6 おわりに

本稿は、文埋め込みを用いて、専門家コミュニティ内の人工知能関連テキストではどのような文が一般向けの新聞とも共有され、一般社会を繋げているのかという問いを解明した。その結果、人工知能の産業と社会連関、人工知能の製品、進化およびリスクを記述する文は一般向けの新聞に入りやすい一方、モデル訓練過程などの文は専門家コミュニティに残されることが確認された。

新聞に入りやすいということは、広い一般社会にとっても関心があり、馴染みやすいことを意味すると解釈できる。これを踏まえれば、人工知能の専門家が広い一般社会とコミュニケーションを行う際、産業連関、社会課題解決および現段階の性能・リスクというトピックをめぐって説明するほうが円滑なコミュニケーションを達成しやすいと考えられる。

また、今後の発展として、文埋め込みを利用し、一般向けのサイエンスコミュニケーション支援を目的とした研究と応用の方向が期待される。特に、専門家コミュニティではどのような文が一般向けの場面にも関心があり、または馴染みやすいのかを提示する実用的なツールやガイダンス開発も予想される。

謝辞

本研究は、JST 次世代研究者挑戦的研究プログラム JPMJSP2138 の支援を受けたものです。

参考文献

- [1] 畝田谷桂子. 科学技術日本語の提案. 留学生センター報告書 鹿児島大学, Vol. 2001, pp. 11–19, March 2001.
- [2] 仁科喜久子, 武田明子. 助詞相当句を中心にみた科学技術日本語テキストの分析-理工系大学院留学生の専門書読解のために-. 東京工業大学人文論叢, Vol. 19, pp. 139–152, March 1993.
- [3] 因京子, 村岡貴子, 米田由喜代, 仁科喜久子, 深尾百合子. 日本語専門文書作成支援の方向: 理系専門日本語教育の観点から. 専門日本語教育研究, No. 9, pp. 55–60, 2007.
- [4] 西條美紀, 野原佳代子, 日下部治. 談話研究から見た科学技術コミュニケーションの意義と実践. **Journal of JSEE**, Vol. 55, No. 1, pp. 59–65, 2007.
- [5] 野原佳代子. 科学技術コミュニケーションの言語と機能—原発事故情報を翻訳理論で読む試み—. 専門日本語教育研究, Vol. 13, pp. 19–24, 2011.
- [6] Ruiqi Li, Xiang Zhao, and Marie-Francine Moens. A Brief Overview of Universal Sentence Representation Methods: A Linguistic View. **ACM Comput. Surv.**, Vol. 55, No. 3, pp. 56:1–56:42, March 2022.
- [7] Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. <https://arxiv.org/abs/1908.10084v1>, August 2019.
- [8] Siyu Duan, Jun Wang, Hao Yang, and Qi Su. Disentangling the cultural evolution of ancient China: A digital humanities perspective. **Humanities and Social Sciences Communications**, Vol. 10, No. 1, p. 310, June 2023.
- [9] Lucian Li. Tracing the Genealogies of Ideas with Sentence Embeddings. In Mika Hämmäläinen, Emily Öhman, So Miyagawa, Khalid Alnajjar, and Yuri Bizzoni, editors, **Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities**, pp. 9–16, Miami, USA, January 2024. Association for Computational Linguistics.
- [10] Dustin Wright, Jiaxin Pei, David Jurgens, and Isabelle Augenstein. Modeling Information Change in Science Communication with Semantically Matched Paraphrases. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, **Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing**, pp. 1783–1807, Abu Dhabi, United Arab Emirates, February 2022. Association for Computational Linguistics.
- [11] Edoardo Filippi-Mazzola, Federica Bianchi, and Ernst C. Wit. Drivers of the decrease of patent similarities from 1976 to 2021. **PLOS ONE**, Vol. 18, No. 3, p. e0283247, March 2023.
- [12] Hayato Tsukagoshi and Ryohei Sasano. Ruri: Japanese General Text Embeddings, 2024.
- [13] Maarten Grootendorst. BERTopic: Neural topic modeling with a class-based TF-IDF procedure, March 2022.

付録 A: 前処理に使用したプロンプト

システムプロンプト	あなたは与えられたルールに厳密に従い、日本語のテキストを機械的に処理する AI だ。
ユーザプロンプト	<p>下のルール及び厳守事項に従い、日本語学術雑誌のテキストを処理せよ（下記ルールで定めたもの以外の変更はすべて禁止）。</p> <p>【ルール】 文の途中で改行（又は改行及び当該文に属さないテキスト）が挿入されている場合、その改行（又は改行及び当該文に属さないテキスト）を削除し、文として正しく繋げる。</p> <p>【厳守事項】 1. 判断に迷う場合は、原文を変更せずそのまま出力する。 2. 内容の追加、要約、言い換えは禁止する。 3. 出力は処理後のテキストのみとする。</p> <p>上記のルールに従い、以下のすべてのテキストを処理せよ。</p>

付録 B: 高・低類似度文の含有率差が最も大きい 20 トピック

トピック	代表的な単語	例文	文分類
0	医療, 診断, 患者, 検査, 内視鏡, 疾患, 治療, AI, 診療, 臨床	AI 内視鏡画像診断研究の結果は、診断の精度管理の向上、内視鏡検診の 2 次読影へ応用される と考える。〔出典:10.50917/jga.7.1_5〕	Q1
2	教育, 生成 AI, 授業, 学校, 学生, 図書館, リテラシー, 生徒, 活用, 大学	その結果, 生成 AI に関する技術的・社会的情報を収集したのちに学生と教員が協働して指針 を作成する方針とした。〔出典:10.11307/mededjapan.54.4_416〕	Q1
4	ChatGPT, 大規模言語モデル, OpenAI, LLM, GPT, Language, large, 生成, 文章, チャット	近年, 大規模言語モデル (LLM)、特に OpenAI が開発した ChatGPT の出現は多岐にわたる領 域での応用が行われている [1,2]。〔出典:10.11517/jsaisigtwo.2023.FIN-031_81〕	Q1
5	今後, 課題, いく, 予定, 改善, たい, さらなる, 検討, 検証, 改良	また検知だけでなく, 予知が可能かについても検討を行う予定である。〔出典:10.14864/ fss.39.0_888〕	Q1
9	若手, 研究者, 分野, 人材, 技術, 新しい, いく, 技術者, 仕事, クラス	このことから, 新規分野への挑戦や開拓が期待される業務には, 若手技術者を積極的に登 用することとしている。〔出典:10.11408/jjsidre.91.10_723〕	Q1
6	電力, 空調, 気温, 外気, 放水, 消費, 温度, 湿度, 運転, ファン	今後はこの空調システムの消費電力とサーバ消費電力を合わせて最小化出来るよう AI にて 運用させることで全体の消費電力の削減を目指す。〔出典:10.18948/shasetaikai.2023. 03.0_57〕	Q1
8	収穫, 生育, 作物, 栽培, 農業, 生産, 収量, 果実, 選別, 選果	局所管理を実現するために, 作物の生育量や収量, 土壌の状態を把握するためのセンシング技 術が開発されてきた。〔出典:10.7210/jrsj.41.692〕	Q1
12	人口, 減少, 不足, 労働力, 高齢化, 年, 人, カーボンニュートラル, 燃料, 炭	その結果, 1 つめの観点については, 少子高齢化で労働力が減少し続けている日本において, AI を導入するポジティブな面として労働力不足が深刻な業種における就労者の過重労働の支 援として強く期待されていることが明らかになった。〔出典:10.32189/jjsts.32.0_31〕	Q1
11	緑, 景観, 道路, CG, km/h, 速度, 樹木, 車両, 街路, 走行	④では左端の沿道施設に反応しており, 道路空間と沿道空間, 緑の一体性が Lingerability を 高めることを示唆している。〔出典:10.14954/jste.9.2_A_213〕	Q1
22	AI, 人工知能, 技術, 近年, はじめ, artificial intelligence, 発展, 進む, 進化, つつ	近年, 人工知能 (artificial intelligence:AI) 技術がさまざまな分野で用いられるようになってきた。 〔出典:10.11499/sicejl.62.176〕	Q1
18	倫理, AI, ガイドライン, 倫理的, 原則, 規制, 国際的, 議論, ガバナンス, 策定	上記の AI と倫理に関する国際的なガイドラインの潮流は, AI の開発を抑止, 減速させるも のではなく, 開発を促進することが目的であった。〔出典:10.20664/jabes.30.0_117〕	Q1
1	mathexpr, 式, ベクトル, 表す, 層, 行列, 集合, ここ, 関数, 出力	ここで, [MathExpr] はベクトル x の 1 から n 番目の要素を取り出して並べたベクトルを表す。 〔出典:10.11517/pjsai.JSAI2023.0_3J1GS105〕	Q4
3	損失, 関数, 最適化, 誤差, 学習, パラメータ, 勾配, エポック, Adam, 重み	損失関数は交差エントロピー誤差, 最適化手法は Adam, バッチサイズは 256 とした。〔出 典:10.11522/pscjspe.2023A.0_482〕	Q4
13	用, 分割, データ, データセット, テストデータ, 学習, 検証, validation, 検証用, 評価	データセットは 8:2 の割合で学習用画像と評価用画像に分割し, モデルの学習と評価を行う。 〔出典:10.14864/fss.39.0_465〕	Q4
16	音源, 分離, マイクロホン, 信号, 音, 周波数, 騒音, フィルタ, 時間, ビーム	ここまで示した音源分離アルゴリズムでは各時間のマイクロホン信号をサンプリングする analog/digital(A/D) 変換器がマイクロホン間で同期されていることを前提としていた。〔出 典:10.1587/essfr.16.4_257〕	Q4
17	CNN, vgg, resnet, 畳み込み, ネットワーク, NET, モデル, Network, 16, imagenet	ネットワークは, 独自に構築した単純な構造の CNN に加え, 高性能な画像分類ネットワー クとして一般利用向けに公開されている VGG[10], ResNet[11], EfficientNet[12] を用いた。 〔出典:10.3811/jjmf.2023.015〕	Q4
23	Python, ライブラリ, python, GPU, pytorch, NVIDIA, Gb, インストール, Intel, Version	種判別モデルの構築には, Python(3.8.11version) の機械学習ライブラリ Pytorch(1.8.1version) をもちい, 学習は ubuntu OS18.04.5LTS にて GPU(NVIDIA TITAN RTX Quadro P400,24GB)with CUDA11.1 で行なった。〔出典:10.11211/birdresearch.19.A31〕	Q4
27	タブロー, 規則, エントリ, m , v , 閉じる, nli, 論理, 含意, 閉鎖	タブロー中に, タブロー規則が適用可能なエントリ e が存在するとき, そのエントリが支 配するすべての葉に, それらの子として新たな m 個の経路 [MathExpr] を追加する。〔出 典:10.5715/jnlp.30.632〕	Q4
31	群, 回転軸, 点, メッシュ, 法線, 方向, 推定, 源線, 不確か	本研究では, 点群の法線から点群の回転軸の方向を推定する深層学習モデルとして図 3 に示 すモデルを作成した。〔出典:10.11522/pscjspe.2023S.0_579〕	Q4
28	層, 畳み込み, プーリング, 特徴, 畳み込む, マップ, エンコーダ, フィルタ, CNN, デコーダ	CNN は特徴を抽出する畳み込み層と, 特徴を維持しながらサイズを小さくするプーリング層 の組み合わせから構成される。〔出典:10.4144/rpsj.69.142〕	Q4