判例検索用キーワードの十分性に係る AI を利用した検討方法について

一 自然言語処理ライブラリ GiNZA を用いて —

関 本 大 樹

目次

はじめに

- 1 AI を活用した自然言語処理ライブラリ GiNZA について
- 2 調査目的・調査方法の概要
- 3 調査・検討に用いたパソコン・システム等
- 4 固定的ルール方式による抽出結果
- 5 抽出モデル方式による抽出結果
- 6 トレーニング回数とバッチサイズの抽象化への影響
- 7 まとめ

おわりに

【参考資料1】固定的ルール方式と抽出モデル方式との比較

【参考資料2】5分割分析表

はじめに

パソコン等の計算能力が昨今劇的に高まったことによって人工知能 (AI) の応用も種々の分野に広がってきている。AI の特徴は、そもそも デジタルであるコンピュータにおいて、我々人類と同様に多くの個別具体 的な経験等から一般化され、抽象化されたルール (いわゆる「経験則」や「勘」などに相当するもの) に基づいて、完璧ではないものの、柔軟かつ 比較的妥当な、いわばアナログ的な判断を可能とする点にあるといえよう。すなわち、実際には経験したことのないような事柄や黒白を付けにくい事柄についても、AI は、それまでの類似事例や経験則等に基づく柔軟

な類推により、一定の許容可能な精度で妥当な判断を行うことが可能なわけである。

課税訴訟や不服申立てにおいても、基本的に前例とは異なる新たな事実や他の事例とは異なる個別的事情を背景としているが故に、改めて司法的判断が求められるわけであり、租税判例や裁決など(以下「租税判例等」という。)において示される判断についても、原理的には法令や判例、経験則等に基づく裁判官等の総合的かつ合理的な判断に基づいて行われる(自由心証主義)ことから、AI 技術の更なる進化に伴って、今後、何らかの形でその適用範囲となる可能性は、十分にあるものと考えられる(1)。

ところで、上記のような AI の活用を図っていく上でも、AI の前述のようなアバウトな性質のため、最終的な判断については我々人類が行う必要があり、その意味でたとえ今後 AI の判断の精度が更に高まったとしても、我々人類に求められる当該最終的な判断に係る責任が軽減されることはないであろう。したがって、我々人類は、AI の力を最大限に活用しつつ、万一その判断に問題があったような場合であっても、そのことを我々がカバーできるような十分に的確かつ柔軟な情報収集手段を確保しておく必要があるわけである。

租税判例等を検索する場合にも上記の事情等は同様と考えられ、そのため当該検索結果を絞り込むためのキーワードの設定も大変プリミティブな手段ではあるものの、今後、特に AI が選定した判例等の妥当性等をチェックする上で、益々重要性が増加しこそすれ、減少することはないものといえよう。したがって、検索対象となる判例等に係るキーワード等の設定の十分性ないし適切性の確保が今後とも重要であるわけである。

しかるに、ある判例等について特定のキーワード等の設定を十分かつ適

⁽¹⁾ 例えば、現状でも、契約書の内容の法的なチェックなどについて、AIを用いて類似の契約書等を自動的に抽出するシステムなど、法務事務の効率化を図るシステムなどが実用化されている模様である。したがって、判決文や裁決文の起案についても、それと同様な AI の応用であれば、その効果の水準はさておき、少なくとも技術的には現状でも可能であろう。BUSINESS LAWYERS「法務におけるテクノロジー活用の実態に迫る『Legal Innovation Conference 法務 DX のいま』公演レポート」(令和5年2月26日現在) https://www.businesslawyers.jp/articles/1258参照。

切に行うことは、当該判例等が将来どのようなキーワードで検索されるか を想定して行わねばならず、それほど容易なものとまではいえないであろ う。今後、判例等が更に増加していくことを考慮すれば、そのようなキー ワード等の設定の十分性ないし適切性を検証し、一定の設定水準を確保す ることは、更に難しくなるものと考えられる。例えば、日本税理士連合会 の運営する和税判例検索システムである「税理士情報ネットワークシステ ム」(以下「TAINS」という。) でも検索対象とされる租税判例等が後述 のとおり2万件ほどあるが、それらに対して個々に設定・管理されている 検索用キーワードである「TAINS キーワード」(2)の件数は、後述のとお り合計で9万5千件ほどである。つまり、新たな租税判例等を一つ登録す る際には平均的にいえば4~5件ほどのキーワードを新たに設定・管理す る必要があるといえよう。

ところで、このような多量のキーワードを統一的に付与し一貫して管理 することを長期間にわたってマンパワーで行うことは必ずしも容易とはい えず、むしろ ICT の利用の必要性が高いといえよう。そこで本稿で注目 したのが、AI 技術を応用した自然言語処理システムにおける固有表現抽 出機能の応用である。ここで固有表現抽出 (named entity recognition) とは、例えば、固有名詞や数詞などについては、それらを構成している品 詞などの個々の構成要素(「形態素」、「トークン」(token) などと呼ばれ る。) に更に分解して認識するのではなく、むしろ特定の固有物を表すた めに複数のトークンが連結されたもの(以下「トークン列」といい、その ように一体的に認識すべきトークン列を「固有表現」(named entity) と いう。)として認識し、それらを文章から自動的に抽出する機能をいう (3)。

^{(2)「}TAINS キーワード」については、例えば、日税ジャーナルオンライン 「生まれ変わった『TAINS6』検索機能の改善、コンテンツの充実 etc.」(令 和5年5月5日現在) https://nichizei-journal.com/interview/%E7%94%9 F%E3%81%BE%E3%82%8C%E5%A4%89%E3%82%8F%E3%81%A3%E3%81%9F% E3%80%8Ctains%EF%BC%96%E3%80%8D%E3%80%80%E6%A4%9C%E7%B4% A 2%E6% A 9%9F%E8%83%BD%E3%81% A E%E6%94%B9%E5%96%84%E3%80%81 %E3%82%B3%E3%83%B3/参照。

^{(3)「}固有表現抽出」の詳細については、例えば、「BERT」と呼ばれるニュー ラル言語モデルにおけるものについて、近江崇宏ほか「BERT による自然 言語処理入門—Transformers を使った実践プログラミング」(オーム社・

そして、当該機能の具体例としては、2019年4月に公開された日本語自然言語処理オープンソースライブラリ「GiNZA」(ギンザ)において、二つの形態の固有表現抽出機能 (4) が利用可能である。すなわち、①「固有表現ルール」によって従来からの定義語と同様に定型的に指定する方法(以下、ここでは「固定的ルール方式」という。)と②「固有情報抽出モデル」によって当該用語の実際の使用状況を学習して柔軟に抽出する方法(以下、ここでは「抽出モデル方式」という。)の二つである。

つまり、固定的ルール方式は、個々の固有表現について文脈等には依存せず、固定的な抽出ルールに基づいて定型的に固有表現と認識して抽出する方法であり、これまでと同様に当該抽出ルール以外の固有表現が認識されることはない。他方、抽出モデル方式では、各固有表現がどのような文脈で使用されているか多数の事例に基づいて機械学習が行われ、当該学習による経験値に基づき固有表現であるか否かが判定され、抽出される方法であり、当該事例において用いられていた固有表現だけではなく、当該固有表現と同様な文脈で利用されているようなトークン列であれば、たとえ個別に指定されていなくとも、類推して固有表現として認識され、抽出される可能性がある。ただし、当該認識の精度は、一般に機械学習を相当程度繰り返し行わない限り、向上することはなく、また、たとえ機械学習を十分に繰り返したとしても、必ずしも学習した事例で示されたTAINSキーワードが抽出されるとも限らない。

²⁰²¹年)105~142頁参照。なお、同書において、「固有表現抽出は、与えられた文章をトークン化し、それぞれのトークン[が固有表現を構成している場合にそのカテゴリーを表すため]のラベルを予測する分類問題として扱うことができ」るとされている(同書119頁参照)。つまり、TAINS キーワードの場合でいえば、「TAINS キーワード」というカテゴリーを想定し、TAINS キーワードの実際の利用状況を機械学習させることにより、それらの使用状況のパターンに類似したトークン列を同カテゴリーに分類・抽出することが原理的には可能であるものと考えられる。

⁽⁴⁾なお、GiNZA ライブラリは、多言語で共用できる、最先端の機械学習技術を取り入れた自然言語処理ライブラリ「spaCy」を基盤となるフレームワークとして利用している。リクルート株式会社「リクルートの AI 研究機関、国立国語研究所との共同研究成果を用いた日本語の自然言語処理ライブラリ「GiNZA」を公開」(2019年)(令和5年2月26日現在)https://www.recruit.co.jp/newsroom/pdf/20190402_01.pdf 参照。

以上の点を踏まえ、本稿では、各TAINSキーワードを固有表現とし て取り扱った場合に、TAINS に収録されている租税判例等の要旨・判旨 などの要約情報について、固定的ルール方式によって抽出される TAINS キーワードと抽出モデル方式によって抽出される TAINS キーワードを実 際に比較してみることとした。その結果、大部分の TAINS キーワードに ついては、上記の二つの方法で共通して抽出対象となるものの、比較的 少数の TAINS キーワードが固定的ルール方式でのみ抽出対象となり、他 方、相当数の TAINS キーワードが抽出モデル方式でのみ抽出対象となる ことが確認できた。そして、機械学習の深度を高めることによって、抽 出モデル方式でのみ抽出が行われる TAINS キーワードは、減少してい くものの、それがなくなるような傾向にはなく、さらに、同方式では、 TAINS キーワードではないトークン列(以下「非 TAINS キーワード」 という。)が相当数、認識されることが確認できた。

抽出モデル方式における上記の非 TAINS キーワードが、新規事案に関 連する TAINS キーワードの十分性ないし適切性を検討する上で参考にな ることや、さらに、新たな TAINS キーワードの自動抽出処理に繋がるこ とを筆者は大いに期待している。そして、その当否については、別稿にて 更に追究したいと考えているが、本稿では、取り敢えず、筆者の検討状況 の現状について前広にご紹介することにより、興味のある読者の参考に供 することとしたい。

AI を活用した自然言語処理ライブラリ GiNZA について 1

今回の調査・検討で採用した GiNZA ライブラリは、多言語対応型 の自然言語処理ライブラリ「spaCy」をその基盤としているが、その spaCy 自体も、昨今一躍脚光を浴びている ChatGPT などでも用いら れているニューラルネットワーク型 AI を応用した自然言語処理モデル 「Transformers」(5)を取り入れた自然言語処理フレームワークである。な

⁽⁵⁾ Transformers は、機械翻訳等に用いるために開発された AI 技術のフレー

お、GiNZA ライブラリの主な特長としては、次の点が挙げられている (6):

- ①高度な自然言語処理をワンステップで導入完了
- ②高速・高精度な解析処理と依存構造解析レベルの国際化に対応
- ③国立国語研究所との共同研究成果の学習モデルを提供

そのため、幅広い分野に適応可能なモデルとして構築されていると紹介 されている。

なお、GiNZA ライブラリには、処理効率に優れた公開当初からの「ja_ginza」モデル(以下「従来型モデル」という。)とその進化形であり、基盤となる spaCy 自体が Transformers を導入したことから解析方法の精度が向上した「ja_ginza_electra」モデル(以下「electra 版モデル」という。)がある。ただし、後者を用いるためには、テンソル(行列)計算等相当規模の計数処理を行う必要があることから、処理時間を並列処理により実用的なものとするために、画像処理用ボード(GPU ボード)が利用できる環境をローカルないしネット環境で準備することが前提とされている($^{(7)}$ 。そこで、今回の調査・検討においては、処理効率よりは、むしろ精度に期待して electra 版モデルを用いることとした。

ムワークであり、多数の用例に基づいて深層学習を行い、一般的ではない個別的・例外的な事例についても精度よく対応が可能な性質がある。なお、GiNZA への適用に当たっても、日本語20億文以上を用いて、深層学習を行ったとされる。Megagon Labs「GiNZA の公開ページ」(令和5年2月26日現在)https://megagonlabs.github.io/ginza/参照。

⁽⁶⁾ 前揭注4、同資料参照。

⁽⁷⁾ なお、今回の検討のようにリアルタイム性がそれほど重要ではないような場合には、spaCyから GPUボードが利用可能ではない状態であるという趣旨の注意メッセージが表示されるものの、GPUボードなしで electra 版モデルを利用することは可能である。ただし、その場合には後述のとおり、機械学習が必要な場面などでは、処理効率が数分の1に低下する場合がある。

2 調査目的・調査方法の概要

(1)調査目的の概要

今回の調査の目的は、① TAINS キーワードを固有表現として取り扱っ た場合に、上記「はじめに」で述べた固定的ルール方式と抽出モデル方式 との間で処理結果に具体的にどのような違いが発生するか確認すること、 そして、②当該結果から TAINS キーワードを付与・管理する上での有用 な情報を得ることが可能か検討することの2点である。

(2)調査方法の概要

和税判例等は、一般に個別性が高いために、そもそも固有名詞等の取 扱いをどのようにするかが課題といえるが、それらの影響を出来るだけ 捨象できるような資料を調査対象とすることにより、固有表現としての TAINS キーワードの取り扱われ方を検討し易くすることとした。すなわ ち、租税判例等の本文自体ではなく、それらの内容が一般化され、ある程 度抽象化された判決要旨、判示事項、裁決の要旨など(以下「概要情報」 という。)を調査対象とすることとした。

具体的には、収集日時点(令和5年2月12日現在)でTAINSに掲載さ れている、当該判決日等が昭和24年以降令和5年までの判例・裁決(合計 19,202件)に係る概要情報を調査対象として、次の手順で分析・検討して みることとした⁽⁸⁾。すなわち、

- ① TAINS のサイトの TAINS キーワードの検索画面から全ての TAINS キーワードを取得する。ただし、50音別で「ん」行と「英数字」 の分類に属するものは除外する。
- ②上記①で取得した TAINS キーワードに基づき上記概要情報における TAINSキーワードの所在を個々の概要情報ごとに網羅的に把握する。
- (8) なお、TAINS からの情報に係る利用規約については、同 HP「コンテン ツご利用にあたっての注意事項」(令和5年2月25日現在) https://www. tains.org/legalnotice/参照。

それと並行して取得したキーワードに基づき固有表現ルールを作成する。

- ③上記②で把握した概要情報中のTAINSキーワードの利用状況を機械学習用データとして用いて固有表現抽出モデルを作成する。その際、機械学習の実施方法と抽出精度との関係を調査するため、(i) 1回の当該学習で学習対象とされるデータの単位である「ミニバッチ」については、当該サイズを512件、1,024件、2,048件、4,096件及び8,192件の計5通りとする。また、当該学習は、同じ機械学習用データを用いて繰り返し行われるが、全てのミニバッチについて当該機械学習を漏れなく1回行うことを「エポック」と呼ぶ⁽⁹⁾。そして、当該エポックの回数と抽出精度との関係を調査するため、(ii) 実施するエポックの回数については、それぞれのミニバッチサイズごとに50回、100回、200回、500回及び1,000回の計5通りとする。
- ④上記③で作成した25タイプの固有表現抽出モデルに基づき抽出モデル方式によって TAINS キーワードの抽出を行い、それらの結果と固定的ルール方式によって TAINS キーワードの抽出を行った結果と比較・検討する。

なお、TAINSからの概要情報の取得については、いわゆるスクレイピング手法(10)を用いて自動的に行うこととした。ちなみに、収集された租

^{(9) 1}回のエポックにおいては、ミニバッチごとに機械学習が行われるため、全ての学習データの件数をミニバッチのサイズで割って小数点以下を切り上げた回数、個々の機械学習が行われることになる。そして、次のエポックに入る前に、ランダムにミニバッチは再編成される。そのため、エポックが繰り返されることにより、特定の学習データがより多くの異なった学習データと組み合わされて機械学習されることになるわけである。したがって、ミニバッチのサイズをより大きくすることによって、より少ないエポック回数によって、機械学習用データ間の同程度の組み合わせによる学習が実現できるものと考えられる。ちなみに、機械学習について並列処理が行われず、逐次処理される場合には、ミニバッチのサイズを小さくしてもエポックで実行される機械学習の回数が逆に増えるため、全体としての処理効率は、さほど変わらない傾向がある。

⁽¹⁰⁾ 具体的なスクレイピングの方法については、国税庁の HP から基本通達の情報をスクレイピングした方法について解説した拙著「国税関係基本通達情

税判例等(19.202件)の、それぞれの年分ごとの件数は、図表1「調査対 象となった各年分の租税判例等の件数」のとおりである。

昭和24年分	2 件	昭和42年分	104 件	昭和61年分	240 件	平成17年分	599 件
昭和25年分	1 件	昭和43年分	126 件	昭和62年分	236 件	平成18年分	604 件
昭和26年分	5件	昭和44年分	105 件	昭和63年分	258 件	平成19年分	442 件
昭和27年分	7 件	昭和45年分	173 件	平成1年分	303 件	平成20年分	478件
昭和28年分	8件	昭和46年分	189 件	平成2年分	287 件	平成21年分	478 件
昭和29年分	7件	昭和47年分	195 件	平成3年分	328 件	平成22年分	554件
昭和30年分	10 件	昭和48年分	227 件	平成4年分	306 件	平成23年分	534 件
昭和31年分	24 件	昭和49年分	255 件	平成5年分	257 件	平成24年分	600 件
昭和32年分	18 件	昭和50年分	268 件	平成6年分	259 件	平成25年分	517 件
昭和33年分	22 件	昭和51年分	267 件	平成7年分	294 件	平成26年分	414 件
昭和34年分	30件	昭和52年分	208 件	平成8年分	289 件	平成27年分	451 件
昭和35年分	25 件	昭和53年分	235 件	平成9年分	349 件	平成28年分	563 件
昭和36年分	39 件	昭和54年分	266 件	平成10年分	391 件	平成29年分	542 件
昭和37年分	35 件	昭和55年分	269 件	平成11年分	405 件	平成30年分	570件
昭和38年分	29 件	昭和56年分	289 件	平成12年分	433 件	令和1年分	363 件
昭和39年分	35 件	昭和57年分	278 件	平成13年分	446 件	令和2年分	89 件
昭和40年分	36 件	昭和58年分	228 件	平成14年分	430 件	令和3年分	133 件
昭和41年分	118件	昭和59年分	224 件	平成15年分	486 件	令和4年分	47 件
昭和42年分	104 件	昭和60年分	300 件	平成16年分	569 件	令和5年分	1件

図表 1 調査対象となった各年分の租税判例等の件数

調査・検討に用いたパソコン・システム等 3

今回の調査・検討に用いたパソコン・システムは、下記(1)から(3) に示すとおり処理能力の異なった複数のパソコンで構成されている。

その主な理由は、処理能力を段階的に増強したためである。すなわち、 本稿の作成にも使用しているメインのパソコン(以下「メイン・パソコン」 という。)にはスペース的に GPU ボードを内部に増設できなかったため、 可能な限りコンパクトなパソコン・システムを目指して、当初、eGPU 方 式 (11) の採用を試みた。しかし、メイン・パソコンには eGPU 方式に必要

報のスクレイピング処理の実際─通達情報の XML 化の必要性について」久 留米大学法学第85号37頁(令和5年2月26日現在)https://kurume.repo. nii.ac.jp/?action=repository_uri&item id=1602など参照。

^{(11)「}eGPU」(external GPU) は、高速な thunderbolt 接続を用いてパソコ ン本体以外に GPU ボードを設置するための仕組みである。現状では、パソ コン本体内部に GPU ボードを増設する場合に比べ、性能等が制限されるた

な thunderbolt 接続を追加できなかったため、同接続が可能なできるだけコンパクトなミニパソコンを導入することとした(以下、当初導入したミニパソコンを「AI-1」、その後で処理能力増強のために追加導入したミニパソコンを「AI-2」という。)。しかし、結局のところ、これらのミニパソコンでは別途調達した GPU ボードを利用する上で GPU メモリが不足すること $^{(12)}$ から、残念ながら eGPU 方式を諦め、ミニタワー型のパソコンを追加調達することとした(以下、当該パソコンを「AI-3」といい、これら GPU ボード利用のために段階的に調達した三つのパソコンを合わせて「AI 用パソコン」という。)。なお、AI 用パソコンは、通常、リモート・デスクトップ機能を用いてメイン・パソコンから LAN 経由で操作等を行うこととしている $^{(13)}$ 。

(1) AI-1の構成等

- ①パソコン: MINISFOLUM 社製 TH50 (intel core i5-11320H 16.0GB)
 - (注) electra 版モデルをローカルに利用する場合には、パソコンのメモリは、16GB 以上が推奨されている (14)。
- ②GPU ボード: NVIDIA RTX A4000 (GPU メモリ16.0GB・thunderbolt 接続による eGPU として使用) ドライバ ver. 528.49

(2) AI-2の構成等

- ①パソコン: Beelink 社製 SEi 12 pro (intel core i7-1260p 64.0GB)
- ② GPU ボード: AI-3 導入までは上記(1)の GPU ボードを必要に応じて差し替えて利用

め、ノート・パソコンなどのコンパクトなパソコンの場合に用いられること が多い。

⁽¹²⁾ より正確には、パソコン側の CPU メモリを GPU ボードと共有する仕組 みである「共有 GPU メモリ」が thunderbolt 接続では利用できないことが 主な理由であった。

⁽¹³⁾ ちなみに、リモート・デスクトップを利用するメリットとしては、AI 用パソコンのオンボード GPU (Intel Iris Xe Graphics など) を無効化することができるため、eGPU との機器接続上の競合を避けることができる点がある。

⁽¹⁴⁾ 前揭注 5、同 HP 参照。

(3) AI-3の構成等

- ①パソコン: APPLIED 社製 Barikata (AMD Ryzen 9-7900 128.0GB)
- ② GPU ボード: AI-3 導入後は上記(1)の GPU ボードを AI-3 で専用

(4) 使用した主なソフトウェア

- ① OS: Windows 11 Pro
- ②プログラム言語: Pvthon 3.10
- ③ GPU コンピューティング開発環境 (GPGPU): CUDA 11.5
- ④ GPU 用数値計算ライブラリ: cupy-cuda115
- ⑤自然言語処理ライブラリ: ginza 5.1.2
- (注)上記②から⑤については、相互依存関係が強いため、互いに適合 するようにバージョン等を調整する必要がある。

(5) GiNZA ライブラリのインストール方法

GiNZA ライブラリのインストール方法については、公式の「GiNZA の公開ページ」に解説されている (15) ほか、インターネット上にある他の解説資料 (16) も参考になる。なお、GPU ボードの並列処理機能を利用するための CUDA モジュールについては、利用者の責任で事前に利用可能な状態にしておく必要がある。

4 固定的ルール方式による抽出結果

(1) 概要

固定的ルール方式では、個々のTAINSキーワードの文字列を当該固有表現の種別を表すための適宜の識別子(今回は一律的に「TAINS_keyword」とした。)と組み合わせて指定すればよい。したがって、当該

⁽¹⁵⁾ 前掲注 5、同 HP 参照。

⁽¹⁶⁾ 例えば、@BUU-SAN「【Python】GiNZA: 日本語自然言語処理オープンソースライブラリ」(2022年)(令和5年2月27日現在)https://qiita.com/BUU-SAN/items/a511981df820e07e015bなど参照。

TAINS キーワードがどのような文脈で用いられているかは一切考慮されないこととなる。また、electra 版モデルで既定状態(デフォルト)として認識される固有表現よりも優先して認識されることになる⁽¹⁷⁾。

ここで、electra 版モデルがデフォルトにおいて、どのように固有表現を認識するのか具体例を紹介しておくこととしたい。すなわち、図表 2 「固有表現抽出例」は、最判昭和42年11月8日(TAINS Z999-9099)の概要情報(判決要旨)について、デフォルトで認識される固有表現をその識別子と共に着色表示したものである。同図表により、例えば、各条文番号が固有表現である「Ordinal Number」としてほぼ正しく認識されていることなどが分かるであろう。

(2) TAINS キーワード

前述したとおり、TAINSシステムでは、収録されている判例等の各種情報について、利用者が当該情報を容易に検索・抽出できるようにするための「TAINSキーワード」と呼ばれるタグが付加されている。例えば、大島訴訟最高裁判決(TAINS Z 1 4 4 - 5 5 0 7)に付加されている TAINSキーワードは、「大島訴訟」、「概算控除」、「給与所得」、「給与所得課税制度」、「給与所得控除」など47件であり、その総数は、平成5年3月7日現在で95,032件にも及んでいる。一方、今回調査・検討対象とした概要情報自体に含まれているトークンの種類は、数詞を入れても標準形で3万2千語ほどであるため、その3倍ほどのTAINSキーワードが特定のトークンが組み合わされたトークン列として登録・管理されていることになる。

なお、各 TAINS キーワードが各種情報に付加されている回数も整理されており、それらについて筆者が集計した結果は、図表 3 「TAINS キーワードの登録状況」に示すとおりである。ただし、「具体例」欄にある

⁽¹⁷⁾ ちなみに、固定的ルール方式の固有表現の認識処理については、デフォルトの固有表現の認識処理の前処理として行うか、あるいは、後処理として行うかを指定することができるが、いずれの場合でも固定的ルール方式の固有表現の認識が優先ないし優越する。



図表 2 固有表現抽出例

付加登録	該当キー	 具体例
回数	ワード件数	(A)
0回	705件	整理手続の開始、収入利子相当額、日本不動産債券、クリーニング取次業など
1 🗆	42,243 件	適正転嫁対策推進本部、福祉施設給付金、市民 活動団体、労働者数、ドライアイスなど
2回	14,228 件	重点的取組、段階税率課税、上場基準、省エネ ビルシステム、偽りその他不正など
3回	7,580 件	軌道業用、会計顧問、失火、遺産共有、内税方 式、朝鮮民主主義人民共和国など
4回	4,599 件	株主総会決議の日、支給実績、法人税法違反事 件、実額算定、税務相談事務など
5回以上	25,677 件	審査請求理由、ATM利用明細、コーポレート ガバナンスコード、220差押えなど
計	95,032 件	

図表3 TAINS キーワードの登録状況

TAINS キーワードついては、適宜ピックアップしたものである。

同図表から、9万5千件のうちの約半数(42,948件)は、付加された回数が1回以下であることから、それらについては、TAINSキーワードのうちでも、特に個別性の高いキーワードであるということができよう。

なお、 $9 \, \pi \, 5$ 千件の先頭文字として用いられている文字の字種は、1,888 字種であり、したがって、先頭文字が同じキーワードが平均で 1 字種当たり 50.3件あることになる。ちなみに、先頭文字種のベスト 3 は、1,708件の「特」、1,161件の「東」、1,092件の「1」であった (18)。

(3) TAINS キーワードの認識件数

固定的ルール方式によって今回の調査・検討対象である概要情報を全て 検索したところ、実際に抽出された TAINS キーワードは、37,907種類で あり、固定的ルール方式による抽出対象としてルール化された95,032件の TAINS キーワードのうち約4割が判決等の本文ではない概要情報のレベ ルでも認識されたことになる。

(4) 固定的ルール方式による TAINS キーワードの抽出例

今回の調査・検討では、固定的ルール方式と抽出モデル方式との機能等の比較をし易いように、トークン間の関係性や固有表現等が一切登録されていない初期状態の言語モデルに両方式による固有表現抽出機能を追加したモデルを作成し、それらに基づいて比較検討することとした。そのうちTAINSキーワードに基づいて固定的ルール方式による固有表現抽出機能を設定した言語モデルを「ja_get_keywords_by_rules」と呼ぶこととする。ちなみに、当該名称の先頭の「ja」は、多言語対応のspaCyライブラリの慣例として日本語処理のための言語モデルであることを示している。

上記言語モデルによって固有表現としての TAINS キーワードを抽出 処理した具体例を【参考資料1】「固定的ルール方式と抽出モデル方式と

⁽¹⁸⁾ なお、「東」で始まる TAINS キーワードの大部分が東京国税不服審判所 担当裁決を表す「東裁」で始まる裁決番号である。

の比較」の左側の欄に示す。なお、当該処理の対象となったのは、最判 令和5年3月6日(TAINS Z888-2481)の判示事項であり、上 記2方式による言語モデルを作成する際に行ったTAINSキーワードの抽 出処理や言語モデルの機械学習のためのトレーニングの対象とはならな かった新件である。

また、同参考資料の左側の欄において灰色に着色されている TAINS キーワードは、上記の2方式で共通して抽出されている固有表現であり、 白地の黒枠で囲われている固有表現は、ja get keywords by rules モデ ルでのみ抽出されたものであるが、全て TAINS キーワードである。この 点で、固定的ルール方式によれば、TAINS キーワード以外の固有表現が 抽出されることはなく、安定的かつ固定的に TAINS キーワードの特定が 可能であることが分かる。

抽出モデル方式による抽出結果 5

(1) 概要

抽出モデル方式における固有表現の指定は、固定的ルール方式の場合と は異なり、文脈に依存した形式で行われる。具体的には、認識対象とした い固有表現が用いられている具体的な文章例に対して、当該固有表現の当 該文章例中の位置情報と当該固有表現の種類を表す適宜な識別子(今回 は固定的ルール方式と同様に一律的に「TAINS keyword」としている。) を付加して指定すればよい。

したがって、抽出モデル方式では同一のトークン列について、特定の文 脈でのみ TAINS キーワードと認識するように指定することも原理的には 可能である。ただし、今回の調査・検討では、固定的ルール方式との比較 を単純化するため、そのような指定は行わず、次項で示すような一律的か つ悉皆的な指定方法を採用した。

(2) TAINS キーワードによる用語の抽出

TAINS に収録されている租税判例等の各概要情報には、前述した当該情報に付加されている TAIN キーワード以外にも、その他多くの TAINS キーワードが用いられているが、具体的にどの部分が TAINS キーワードに該当するか判定するために今回採用した抽出方法は、大要次のとおり文字ではなく、むしろトークンを基準としたものである:

- ① TAINS キーワードのうち先頭文字が等しいものをグループ分け (1,888グループ) し、各グループ内では、文字数の多い方から少ない 方に順に並べておく
- ②判定対象となる文章を electra 版モデルで各トークンに分解したのち、 先頭のトークンから順に、その最初の文字から始まるような TAINS キーワードが上記①で登録されているか、文字数の多い TAINS キー ワードから少ない方へ順に比較し判定する。すなわち、いわゆる「最 長一致」ベースである
- ③もし上記②の結果、登録されているキーワードと同じ一連の文字列があれば TAINS キーワードであると認識し、なければ1トークン分後ろにずらして改めてチェックを行う。なお、認識された TAINS キーワードがトークンの区切りで終わっていない場合には、次のトークンの区切りまでの文字は単純にスキップする
- ④上記②及び③の手順を当該文章の終わりまで繰り返す

上記の抽出処理を図表 2 「固有表現抽出例」の概要情報について行った 結果を図示すると図表 4 「TAINS キーワードとトークンとの対応関係」 のとおりとなる。なお、斜線は、各トークンの区切りを示し、一重下線部 分は、一つのトークンで TAINS キーワードに対応していることを示して いる。また、二重下線部分は、複数のトークンがトークン列として TAINS キーワードに対応していることを示している。したがって、特に 条文番号など、たとえ本来的な固有表現として一体的なトークン列であっ

ても、分割されてしまう場合がある⁽¹⁹⁾。さらに、二重下線がトークンの 途中で終っているところ⁽²⁰⁾も散見されるが、今回の TAINS キーワード の抽出法が単純であるため、うまく機能しない場合があることを示してい る。おって、図表中の「<EOS>」は、各段落の終わりを表している。

旧/物品/税法/(/昭和/37/年/法律/第/48/号/に/よる/改正前/の/もの/)/<EOS>

第/18/条/第/1/項/第/2/号/に/いう/「/詐偽/その/他/不正/の/行為/」/の/意義/<EOS>

判/ /決/ /要/ /旨/<EOS>

旧/物品/税法/(/昭和/37/年/法律/第/48/号/に/よる/改正前/の/もの/)/<EOS>

第/18/条/第/1/項/第/2/号/に/いう/「/詐欺/その/他/不正/の/行為/」/と/は/、/逋脱/の/意図/を/もつ /て/、/その/手段/と/し/て/税/の/賦課/徴収/を/不能/もしくは/著しく/困難/なら/しめる/よう/な/なん/ ら/か/の/偽計/その/他/の/工作/を/行なう/こと/を/いう/もの/と/解する/の/を/相当/と/する/。/<EOS> 【/参照/】/旧/所得税法/(/昭和/25/年/法律/第/69/号/に/よる/改正前/の/もの/)/<EOS>

第/69/条/第/1/項/ / 詐偽/その/他/不正/の/行為/に/より/第/26/条/第/1/項/第/1/号/に/規定/す る/所得税額/の/全部/又/は/一部/に/つき/所得税/を/免れ/た/者/は/、/これ/を/3/年/以下/の/懲役/又/ は/その/免れ/た/税金/の/5/倍/以下/に/相当/する/罰金/若しくは/科料/に/処する/。/詐偽/その/他/不 正/の/行為/に/より/第/37/条/第/1/項/、/第/38/条/第/1/項/又/は/第/40/条/乃至/第/42/条/の/ 規定/に/より/徴収/せ/らる/べき/所得税/を/免れ/た/者/も/、/また/同様/と/する/。/<EOS>

同法/(/昭和/29/年/法律/第/52/号/に/よる/改正前/の/もの/)/<EOS>

第/69/条/第/1/項/ /詐偽/その/他/不正/の/行為/に/より/、/第/26/条/第/1/項/第/3/号/乃至/第/ <u>5/号</u>/に/規定/する/所得税額/に/つき/所得税/を/免れ/又/は/<u>第/3</u>6/条/<u>第/3</u>/頃/(/同条/第/5/項/に/ おい/て/準用/する/場合/を/含む/。/) /第/3 6/条/の/2/第/1/項/ (/同条/第/3/項/に/おい/て/準用/す る/第/3 6/条/第/5/項/に/おい/て/準用/する/場合/を/含む/。/) /<EOS>

若しくは/第/36/条/の/3/第/1/項/(/同条/第/3/項/に/おい/て/準用/する/第/36/条/第/5/項/に/お い/て/準用/する/場合/を/含む/。/) /の/規定/に/よる/所得税額/の/還付/を/受け/た/者/は/、/これ/を/3 /年/以下/の/懲役/若しくは/5百万/円/以下/の/罰金/に/処し/又/は/これ/を/併用/する/。/<EOS> 旧/物品/税法/(/昭和/24/年/法律/第/286/号/に/よる/改正前/の/もの/)/<EOS>

第/18/条/第/1/項/ / 許偽/其/ノ/他/不正/ノ/行為/二/依/リ/物品税/ヲ/逋脱/シ/又/ハ/逋脱/セントシ タル/者/ハ/其/ノ/逋脱/シ/又/ハ/逋脱/セントシタル/物品税/5/倍/ニ/相当/スル/罰金/ニ/処/ス/同法/(/ 昭和/37/年/法律/第/48/号/に/よる/改正前/の/もの/) /<EOS>

第/18/条/第/1/項/第/2/号/ /左/ノ/各号/ノ/一/二/該当/スル/者/ハ/5/年/以下/ノ/懲役/若/ハ/50 ヲ/浦脱/シ/又/ハ/其/ノ/浦脱/ヲ/図/リタル/者/<EOS>

図表 4 TAINS キーワードとトークンとの対応関係

- (19) そのような場合については、今回は条文番号等の取り扱いには余り注目し なかったため採用しなかったが、GiNZA ライブラリの本来の固有表現抽出 機能を活用して、前処理を行い、抽出対象としないこととする方法も考えら れよう。
- (20) このような場合については、そもそも TAINS キーワードではないと判定 して、抽出対象としないことも考えられよう。

(3) 固有表現抽出モデルによる対応

そもそも固有表現の抽出処理は、当該固有表現が置かれている前後の 文脈等も把握しながらバランスよく行う必要があり、単純には実現でき ない。そのため、GiNZA ライブラリでは、そのような柔軟な判定を可能 とする高度な固有表現抽出処理を機械学習により実現しているわけであ る。特に、GiNZA ライブラリのような汎用性を追求した言語処理システ ムでは、今回の調査・検討の対象である TAINS キーワードの場合とは異 なり、必ずしも事前に特定できないような固有名詞などの固有表現を柔軟 に取り扱うことが必要であり、機械学習による抽象的ないし類推的な判 定が必須といえる。そして、GiNZA ライブラリが基盤とした spaCy ライ ブラリでは、利用者が独自に個別的分野において、同ライブラリと同様な 機械学習を用いた固有表現を取り扱う手段を提供しており、その仕組みは 「固有表現抽出モデル」(Named Entity Recognition model)と呼ばれる。 そして、今回の調査・検討で採用した TAINS キーワードに係る抽出モデ ル方式も、当該モデルを用いて実現しているわけである。

しかるに、TAINSキーワードのような事前に特定可能な固有表現を敢えて抽象化を伴う固有表現抽出モデルで取り扱うメリットとしては、少なくとも次の2点が挙げられよう。すなわち、

①使用されている文脈によって同じようなトークン列であっても TAINS キーワードであったり、なかったりするため、文脈に応じた 柔軟な取扱いが必要であるが、固定的ルール方式では困難であること (例えば、固定的ルール方式では、「中古不動産売買」について、「中古不動産」が TAINS キーワードではないことから、「中古」と「不動産売買」という二つの TAINS キーワードで構成されているものと 誤認される可能性が高い。確かにこの例では「中古不動産」を新たに 登録すれば、改善され得るが、他方、抽出モデル方式では、「中古不動産」としての意味的な繋がりが、TAINS キーワードである「不動産」としての意味的な繋がりが、TAINS キーワードである「不動産売買」よりも強いため、たとえ「中古不動産」が TAINS キーワー

される場合がある。なお、詳しくは別稿にて紹介することとしたい。)

② TAINS キーワードとしての判定が抽象化されることによって、全く 新規の概要情報等からでも用法がこれまでのものと類似するトークン 列を抽出することが可能となり、その点で TAINS キーワードの新規 選定の支援手段や既存キーワードの妥当性に係る検討手段とすること が期待できること

の二つである。

そこで、まず、上記の目的を理解する上での参考として、上記(2)で述べた方法で抽出された TAINS キーワードについて、その具体的な使用状況に係る上記(1)で述べたような機械学習用データを用いてトレーニングを行った固有表現抽出モデル「ja_for_tains_keywords3_50_8192」について紹介したい。すなわち、同モデルによって固有表現としてのTAINS キーワードを抽出処理した具体例を【参考資料1】「固定的ルール方式と抽出モデル方式との比較」の右側の欄に示す。なお、抽出処理の対象となった概要情報は左右の各欄で全く同じものである。ちなみに、モデル名中の「3_50_8192」の部分は、「AI-2パソコンを用いて1バッチ8,192件(全3ミニバッチ)で50エポック」機械学習のためのトレーニングを行ったモデルであることを表している。

上記参考資料の左右の欄において TAINS キーワードとして認識されたトークン列で相違する部分(白地で黒枠の部分)を見比べると、抽出モデル方式によって、固定的ルール方式では抽出されなかったキーワード候補が抽出されていることが分かるであろう。ただし、左欄とは異なり、右欄で抽出されているトークン列は、必ずしも TAINS キーワードではない可能性がある。

具体的には、図表中の項目番号「1」で抽出モデル方式においては、 TAINSキーワードではない「買取再販売」が抽出されているのに対して、 固定的ルール方式では、既存の TAINS キーワードである「再販売」のみが抽出されている。しかるに、当該事案が転売目的の不動産に係る消費税の仕入税額控除に係る事案であることを考慮すれば、「買取再販売」を新たに TAINS キーワードとして登録する意味が十分にあるものと考えられる。

また、図表中の項目番号「3」で抽出モデル方式では TAINS キーワードではない「関係機関」が抽出されているのに対して、固定的ルール方式では、それぞれ TAINS キーワードである「関係」と「機関」の二つのキーワードとして抽出されている。

さらに、同項目番号において、抽出モデル方式ではTAINSキーワードではない「課税対応課税仕入れ」が抽出されているのに対して、固定的ルール方式では、「課税対応」と「課税仕入れ」の二つの既存のTAINSキーワードの組合せとして認識されている。しかるに、これも形容語である「課税対応」、「共通対応」、「非課税対応」などと被修飾語である「課税仕入れ」とを区分して認識するよりも、むしろ「課税対応課税仕入れ」、「共通対応課税仕入れ」、「非課税対応課税仕入れ」などと一体のキーワードとして認識すべきとも考えられよう。ちなみに、「課税対応」及び「共通対応」は、TAINSキーワードであるものの、「非課税対応」は、TAINSキーワードとしては未登録である(本稿執筆時点、以下同じ。)。

おって、図表中の項目番号「5」で抽出モデル方式では、「第1審」が抽出されているものの、固定的ルール方式では、「第1」部分のみが TAINS キーワードとして抽出されている。ちなみに、「控訴審」、「上告審」は TAINS キーワードとして登録されているものの、「第一審」ないし「第1審」は未登録である。

(4)固有表現抽出モデルの特徴

それでは、つぎに上記(3)の②で述べた目的、すなわち、抽出モデル 方式を TAINS キーワードの新規選定の支援手段や既存キーワードの妥当 性に係る検討手段とすることの可能性について検討してみることとしたい。

一般にミニバッチ方式における機械学習の一般的傾向として、学習デー

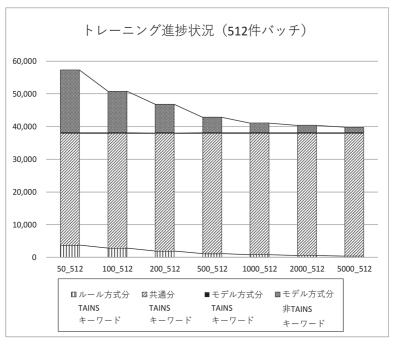
タに基づくトレーニングを繰り返し行えば行うほど、学習水準が高まり、 AI の判定精度が向上することが挙げられる。例えていえば、トレーニン グの初期では、AIは、ややピンボケの判定を行いがちであるものの、ト レーニングを繰り返すことにより、判定の精度が上がり、ピントが合うよ うになるといえよう。

今回の調査・検討においても、上記のようなトレーニングの効果を端的 に示しているのが図表5「トレーニング進捗状況」である。

同図表は、バッチサイズを512件に固定して固有表現抽出モデルに対し てトレーニングを50エポック、100エポック、200エポック、500エポック、 1000エポック、2000エポック及び5000エポック、それぞれ実施したのち にトレーニングに用いた学習データについて、① ja get keywords by rules モデルを使用した固定的ルール方式及び②各トレーニング後の固有 表現抽出モデルを使用した抽出モデル方式の両方式により固有表現抽出処 理を実施した場合において、各モデルが抽出した固有表現がどの程度正解 の TAINS キーワードであったのかを表した図表である。

同図表で、まず注目して頂きたいのは、「ルール方式分非 TAINS キー ワード 欄の数値が常に 0 件となっていることである。すなわち、固定的 ルール方式では、何度抽出作業を行っても非 TAINS キーワードが抽出さ れることはないわけである。つぎに、「モデル方式分非 TAINS キーワー ド」欄の数値に注目して頂きたい。これは、各エポック数のトレーニング 後の固有表現抽出モデルで固有表現と認識されたトークン列のうち実際に はTAINSキーワードではなかったものの件数である。50エポック版固有 表現抽出モデルでは、19,147件であったが、5000エポック版では、1,627 件まで低下している。これは、トレーニングによって、TAINS キーワー ドを抽出する精度が向上したためであると考えられるが、これを逆説的に 捉えれば、各モデルの水準で既存の TAINS キーワードに類似したパター ンのトークン列であると類推(誤認)する比率は、トレーニングが少ない ほど多いといえるわけである。したがって、TAINS キーワードを精度よ く抽出するためには、できるだけ多数回のエポックによりトレーニング

	固定的ルール方式		両方式共通	抽出モデル方式	
固有表現 抽出モデル	ルール方式分 非TAINS キーワード	ルール方式分 TAINS キーワード	共通分 TAINS キーワード	モデル方式分 TAINS キーワード	モデル方式分 非TAINS キーワード
50_512	0	3,709	34,198	234	19,147
100_512	0	2,810	35,097	220	12,667
200_512	0	1,892	36,015	208	8,649
500_512	0	1,113	36,794	222	4,687
1000_512	0	796	37,111	220	2,933
2000_512	0	584	37,323	212	2,237
5000_512	0	410	37,497	216	1,627



図表 5 トレーニング進捗状況

されたモデルを用いるべきであるものの、他方、今回のように、既存の TAINS キーワードに類似したトークン列に注目したい場合には、むしろ 少数回のエポックでトレーニングされたモデルを用いることにメリットが あるといえるのではなかろうか。

そこで、上記の仮説を確認するため、今回の調査・検討で機械学習用 データとした裁判例等以降にTAINSに収録された上記(3)で紹介した 【参考資料1】の事例について TAINS キーワード抽出処理を試みたわけ である。

(5) 新規事案に抽出モデル方式を適用した事例

会和5年5月7日現在でTAINSに収録されていた判決日等が会 和5年1月1日以降の裁判例等は5件(21)あったが、そのうち、今回の調 香・検討で既に機械学習用データとして収集されていた 1件(TAINS Z 888-2474) 以外の4件の中から、最判令和5年3月6日 (TAINS Z888-2481) を適宜選定した。

上記裁判例は、住宅用賃貸部分を含む中古建物の用途区分に関連した消 費税の更正処分に係る過少申告加算税について、その免除要件である「正 当な理由」の有無が争われた事例であり、下級審については、機械学習用 データに含まれているものではあった。

上記事案に係る固定的ルール方式と抽出モデル方式との固有表現の抽出 結果の主な相違点については、上記(3)に紹介したとおりであるが、特 に「買取再販売」というトークン列は、TAINS キーワードとしては「買取」 部分が TAINS キーワードとして未登録であることから、AI による類推 によって抽出されたものと考えられる。ただし、「買取り」は TAINS キー ワードとして登録済みである。

確かに、上記で行ったような単純な抽出・比較手順は、概要情報中の多 くの TAINS キーワードの中から将来的にも検索目的で重要度が高いと想

⁽²¹⁾ 具体的には、TAINS Z 9 9 9 - 7 2 2 6、TAINS Z 9 9 9 - 5 4 6 9、 TAINS Z 8 8 8 - 2 4 7 4, TAINS Z 8 8 8 - 2 4 8 0, TAINS Z 888-2481の5件であった。

定される付加対象のものを選別したり、そもそも概要情報中には存在しない別の TAINS キーワードを当該概要情報に付加したりすることに比べれば低レベルのものではあるものの、少なくとも、当該手順には、上記のような高度な知的作業を支援したり、見落としを減らしたり、あるいは、検討のための着眼点を提供したりする効果が期待できるのではあるまいか。

6 トレーニング回数とバッチサイズの抽象化への影響

上記5の(4)ではトレーニング回数の増加によって固有表現抽出のピントがよりシャープになることについて触れたが、バッチサイズによってもピントが変動する傾向がある。具体的には、ミニバッチのサイズを大きくすると同じトレーニング回数でもピントが甘くなる(抽出される固有表現が実際にTAINSキーワードである確率が減少する)傾向が認められ、その関係を図示したのが図表6「トレーニング回数とバッチサイズの相互関係」である。

すなわち、同図表は、トレーニング回数とバッチサイズによって固有表現抽出モデルによって抽出される固有表現について TAINS キーワードと非 TAINS キーワードの件数等がどのように変化するか一覧表にしたものである。具体的には、その①では、そもそも抽出されるキーワードの総数、その②では、上記①のうちで固定的ルール方式と共通の TAINS キーワードの件数、その③では、固定的ルール方式では抽出されなかったキーワードの件数、その④では、上記③のうち TAINS キーワードである件数、そして、その⑤では、上記③のうち非 TAINS キーワードである件数をそれぞれ表示している。

しかるに、上記の各表のうち特に⑤の表に注目して頂きたい。同表において、最大の非 TAINS キーワード件数は、トレーニング回数が50エポックで、バッチサイズが8,192件のとき(以下、当該固有表現抽出モデルを「50_8192モデル」と略称し、他のモデルについても同様に略称する。)の38.749件であり、最小の非 TAINS キーワード件数は、トレーニング回数

①抽出モデル方式認識キーワード総数

トレーニ	バッチサイズ				
ング回数	512	1,024	2,048	4,096	8,192
50	53,579	55,961	62,916	69,967	68,932
100	47,984	50,758	54,704	62,372	65,966
200	44,872	45,560	48,838	55,211	59,566
500	41,703	42,126	44,914	46,432	49,618
1000	40,264	40,653	41,527	43,022	44,039

②上記①のうち固定的ルール方式と共通なTAINSキーワード件数

トレーニ	バッチサイズ					
ング回数	512	1,024	2,048	4,096	8,192	
50	34,198	33,421	32,594	31,603	29,859	
100	35,097	34,551	33,728	33,056	31,422	
200	36,015	35,632	35,002	34,332	33,136	
500	36,794	36,584	36,375	35,726	34,974	
1000	37,111	37,045	36,895	36,545	36,050	

③上記①のうち上記②を除いた認識キーワード件数

トレーニ	バッチサイズ					
ング回数	512	1,024	2,048	4,096	8,192	
50	19,381	22,540	30,322	38,364	39,073	
100	12,887	16,207	20,976	29,316	34,544	
200	8,857	9,928	13,836	20,879	26,430	
500	4,909	5,542	8,539	10,706	14,644	
1000	3,153	3,608	4,632	6,477	7,989	

④上記③のうちTAINSキーワード件数(固定的ルール方式では認識不能)

トレーニ	バッチサイズ				
ング回数	512	1,024	2,048	4,096	8,192
50	234	237	271	253	324
100	220	239	226	272	276
200	208	216	219	244	262
500	222	207	217	220	223
1000	220	214	216	221	211

⑤上記③のうち非TAINSキーワード件数

トレーニ	バッチサイズ					
ング回数	512	1,024	2,048	4,096	8,192	
50	19,147	22,303	30,051	38,111	38,749	
100	12,667	15,968	20,750	29,044	34,268	
200	8,649	9,712	13,617	20,635	26,168	
500	4,687	5,335	8,322	10,486	14,421	
1000	2,933	3,394	4,416	6,256	7,778	

図表 6 トレーニング回数とバッチサイズの相互関係

が1000エポックで、バッチサイズが512件であるとき(1000_512 モデル)の2,933件であり、13倍ほどの開きがある。

そこで、トレーニング回数とバッチサイズの変化によって抽出される非 TAINSキーワードが具体的にどのように変化していくか、その全般的な 傾向を把握し易くするために、【参考資料2】「5分割分析表」を作成した。

なお、当該分析表は、次のような手順で作成したものである。すなわち、①固有表現抽出モデルを非 TAINS キーワードが最小となる「1000_512モデル」から「500_1024モデル」、「200_2048モデル」、「100_4096 モデル」、そして、最後に「50_8192モデル」へと、最大の変化率が得られるように固有表現抽出モデルを変化させながら、それぞれのモデルで把握される非 TAINS キーワードを抽出する。その際、重複を避けるために②それまでの抽出処理で既に把握されている非 TAINS キーワードは除外する。そして、③認識回数が多い順(ただし認識回数が同じものについては更に文字数の多い順)に並べて、5等分したのち、上位のグループから第1階級ないし第5階級と呼んで、各階級の上位のものを約1000文字分ずつ例示した。

つまり、上記のように固有表現抽出モデルで5区分、さらに、それぞれのモデルごとに5階級に区分することによって、TAINSキーワードに格上げする必要性のより高い非TAINSキーワードが、抽出のピントのより強い、すなわち、トレーニング回数がより多く、バッチサイズがより小さい固有表現抽出モデルの、必要性ないし重要度の指標である認識回数のより多い第1階級に集められることになるものと考えられる。

そして、上記のように固有表現抽出モデルの抽出のピントの強さで5区分、さらに、それぞれのモデルごとに認識頻度の高さで5階級に区分することによれば、既存のTAINSキーワードとの類似性の強さとTAINSキーワードとして認識すべき必要性を大雑把に整理することが可能となる。

例えば、今回の試行では、前出の「買取再販売」というトークン列は、 【参考資料2】では記載が省略されているものの、抽出のピントのより強い1000 512モデル、500_1024モデルでは認識されず、中間的な200_2048モ

デルで初めて非TAINSキーワードとして2件の概要情報(22)で3回認識 されている(第1階級)。すなわち、当該トークン列は、既存のTAINS キーワードとの類似性が中程度であるが、TAINS キーワードとして採用 された場合の効果は、比較的高いと整理できるわけである。

なお、抽出のピントのより強い1000 512モデルで採用効果の高い第1階 級の非 TAINS キーワードについてみると、それらに対応する【参考資 料2】の先頭の欄内で下線を付した「日付け」、「大蔵省令」、「揮発油税法」、 「売上値引き」、「不動産貸付業務」、「延滞税額」、「対価額」、「事業場所」、 「報酬金額」、「調査手続き」、「不動産取得税額」、「重加算税額」、「先物取 引に係る雑所得の金額」、「地価税法」、「譲渡所得税額」、「慰謝料額」、「電 力料金1、「不納付加算税額1、「土地重課税1、「自動車重量税法1、「申告所 得税額」などは、認識回数が比較的多いにも関わらず TAINS キーワード としては未登録であることが分かる。ただし、これらの用語の大部分につ いては、TAINS キーワードである文字列が含まれており、検索用として は、敢えて登録をする意味が乏しかったものとも考えられる。

他方、50 8192モデルの第1階級では、例えば、同図表中で下線を付し た「更正請求」の認識回数が372回と比較的高いにも関わらず登録されて いないが、TAINS キーワードである本来の「更正の請求」を補足するも のとして追加登録する価値はあろう。また、同じく下線を付した「当初更 正処分」、「損金計上」、「国賠法」、「寄付金」、「同業者比率法」、「原処分調 査」なども認識回数が比較的高いことから、キーワード検索の精度を上げ るためにも、追加登録する価値はあるのではなかろうか。ちなみに、未登 録である「損金計上」の対照語である「益金計上」については既に登録済 みである。さらに、同じく二重下線を付した「資産の譲渡代金」、「特定の 事業用資産」、「株式の譲渡」、「土地の取得費」、「不動産の譲渡」、「譲渡所 得の特例」、「相続財産の評価」、「相続財産の価額」、「税負担の公平」、「民

⁽²²⁾ 具体的には、ムゲンエステート事件関連の TAINS Z 2 6 9 - 1 3 3 2 5 及び TAINS Z 2 7 1 - 1 3 5 5 1 の各判示事項である。ただし、「買取再販 住宅」というトークン列については、課税庁からの開示情報に係るTAINS キーワードとして既に登録されている(TAINS 所得事例東京局R 0 4 1 2)。

法上の組合契約」、「原処分の取消し」、「納税者間の公平」、「更正の理由付記」、「労務の対価性」、「資産の取得価額」などについては、助詞の「の」で連結された二つの用語のペアがあたかも単独のキーワードとして定型的に用いられる傾向にあることが分かる。以上の具体例からも、やはり抽出のピントのより甘い50_8192モデルでは、ピントが明瞭な1000_512モデルよりもより広い視野ないし視界で、既存のTAINSキーワードに類似する用語等が抽出される傾向にあることを見て取れよう。

以上のような分析結果から、抽出のピントからいえば中程度の類似性が認められる200_2048モデルにおいて有用な新規 TAINS キーワードが抽出されている可能性が高いといえそうである。そこで、試みに【参考資料2】にある200_2048モデルの第1階級の欄の非 TAINS キーワードを眺めてみると、例えば、下線を付した「更正すべき理由がない旨の通知」(14件の概要情報⁽²³⁾ で計15回認識)及び「更正すべき理由のない旨の通知処分」(12件の概要情報⁽²⁴⁾ で計12回認識)があるが、これは、TAINS キーワードである「更正すべき理由がない旨の通知処分」(98件の概要情報で計164回認識)と同義ないしほぼ同義であると考えられ、また、下線を付した「引渡しのあった日」(10件の概要情報⁽²⁵⁾ で計10回認識)も TAINSキーワードである「引渡しがあった日」(36件の概要情報で計65回認識)

⁽²³⁾ 具体的には、TAINS Z 1 0 3 - 4 2 7 0、TAINS Z 1 1 4 - 4 6 6 5、TAINS Z 1 8 3 - 6 7 2 7、TAINS Z 1 9 3 - 7 0 5 3、TAINS Z 1 8 9 - 6 9 3 0、TAINS Z 1 9 4 - 7 0 8 6、TAINS Z 1 9 9 - 7 2 5 4、TAINS Z 2 2 0 - 7 7 6 7、TAINS Z 2 5 3 - 9 4 3 6、TAINS Z 2 5 5 - 1 0 2 2 0、TAINS Z 2 5 9 - 1 1 2 3 4、TAINS Z 2 6 2 - 1 1 9 5 9、TAINS Z 2 6 4 - 1 2 4 6 7、TAINS Z 8 8 8 - 2 4 3 4 の各判示事項である。

⁽²⁴⁾ 具体的には、TAINS Z 1 7 4 - 6 3 8 8、TAINS Z 1 8 0 - 6 5 6 2、TAINS Z 2 1 4 - 7 6 1 1、TAINS Z 2 2 1 - 7 8 0 3、TAINS Z 2 2 1 - 7 8 0 9、TAINS Z 2 2 3 - 7 9 3 4、TAINS Z 2 4 9 - 8 7 9 4、TAINS J 6 0 - 1 - 0 3、TAINS Z 2 5 1 - 8 9 9 7、TAINS Z 2 5 8 - 1 1 0 8 6、TAINS Z 2 5 8 - 1 0 8 6 1、TAINS Z 2 5 8 - 1 0 9 6 6 σ 8判示事項等である。

⁽²⁵⁾ 具体的には、TAINS F 0 - 2 - 2 0 1、TAINS F 0 - 5 - 0 9 8、TAINS J 8 1 - 6 - 1 5、TAINS F 0 - 5 - 2 1 4、TAINS F 0 - 5 - 1 3 4、TAINS F 0 - 5 - 2 0 8、TAINS F 0 - 5 - 2 0 9、TAINS F 0 - 5 - 2 3 9、TAINS F 0 - 5 - 2 3 4、TAINS F 0 - 5 - 2 3 3 の各裁決の要旨である。

と同義であると考えられる。したがって、これらの非 TAINS キーワード も区別せずに、同義の TAINS キーワードとして取り扱うことが便宜であ ろう。さらに、下線を付した「取得した日」については、20件の概要情報⁽²⁶⁾ で計30回認識されているにも関わらず、筆者収集時点ではTAINSキー ワードとして登録されていなかったものであるが、最近の3件の相談事 例⁽²⁷⁾ に対する TAINS キーワードとして現状では既に登録されている。 おって、下線を付した「総所得」については、それを包含する「総所得 金額」がTAINSキーワードではあるものの、「総所得」単体としては未 登録である。なお、「総所得」というトークン列は、6件の概要情報⁽²⁸⁾で 計8回認識されているが、それらは必ずしも「総所得金額」を表現したも のとはいえないものと考えられる。したがって、少なくとも以上に紹介し たこれらの非 TAINS キーワードについては、TAINS キーワードとして 登録するか、あるいは、同義の TAINS キーワードとして一体的に取り扱 う一定の妥当性がありそうである。

いずれにしても、適切な固有表現抽出モデルを用いることにより、少な くとも用語間の用法上の類似性に基づいて非 TAINS キーワードについて 一定の整理ができるようになったといえるのではなかろうか。また、今後 の当面の課題としては、【参考資料2】にリストアップされた非 TAINS キーワードのうち、TAINS キーワードとする価値の乏しいものを効率的

⁽²⁶⁾ 具体的には、TAINS F 0-1-199、TAINS J 19-4-07、 TAINS J 49-2-15, TAINS J 55-3-19, TAINS J 57-2-19, TAINS F 0-1-214, TAINS J 68-2-10, TAINS J 70-5-20, TAINS F 0-2-311, TAINS Z 999-5278, TAINS J 8 5 - 2 - 1 1, TAINS J 8 8 - 2 - 0 2, TAINS J 8 8 -2-0.2, TAINS J 9.4-2-0.7, TAINS F 0-5-1.8.8, TAINS F 0 - 1 - 8 4 9 TAINS I 1 0 6 - 2 - 0 4 TAINS Z 2 6 8 -1 3 1 1 7, TAINS F 0 - 5 - 2 4 8, TAINS Z 2 6 9 - 1 3 2 5 3 σ 各判示事項等である。

⁽²⁷⁾ 具体的には、TAINS 所得事例東地会 0 2 0 1 5 9、TAINS 源泉事例H P0659、TAINS 法人事例700508の各相談事例である。

⁽²⁸⁾ 具体的には、TAINS Z 2 5 5 - 1 0 2 5 8、TAINS Z 2 5 6 -1 0 5 8 3 TAINS Z 2 6 2 - 1 1 9 2 5 TAINS Z 2 6 4 -1 2 5 3 8 TAINS F 0 - 1 - 1 1 8 3 TAINS F 0 - 1 - 1 1 8 2 0 各判示事項等である。

かつ効果的に除外する方法について検討することが挙げられよう。おって、抽出モデル方式の類似性検出能力を活用する上で、そのための抽出モデルをそもそもどのようにトレーニングするかも課題であろう。基本的に自然言語処理モデルのトレーニングは、飽くまでも当該モデルの出力を何らかの正解(本稿でいえば「TAINSキーワード」)にいかに近付けるかを本来の目的とするものであって、必ずしも当該モデルについて正解と有意に類似するもの(本稿でいえば TAINSキーワードとすべき「非 TAINSキーワード」)を抽出ないし選別するための能力を高めるためのものとはいえないであろう。また、当該トレーニング自体も乱数を用いた不確定のものであって、その結果も再現性を伴わないものである。そこで、差し当たり、そのような非再現性を踏まえた統計的な TAINSキーワード候補の選定方法について、別稿でより具体的に検討してみたいと考えている。

おわりに

今回の調査・検討を実施する前の段階では、果たして本調査で用いたような軽装の処理系(上記3参照)で必要な処理が能力的に可能か心配されたものであったが、結果として、今回の検討内容程度であれば十分に実用的なものであった。比較的手軽に入手可能なパソコンやGPUボードの計算能力の高さをまさしく実感させられた次第である。ただし、今回の調査・検討では当初eGPU方式によりGPUボードをパソコンの外部に置いて利用した場合には、大きめのバッチサイズのトレーニングについてはGPUボードをうまく利用できなかったことから、少なくとも現状では、パソコン内部にGPUボードを設置する利用方法が無難といえよう。また、たとえGPUボードを利用してもトレーニングには、相当の時間を要する。具体的には、「ja_for_tains_keywords3_5000_512_GPU」のトレーニングには AI-1で12日間以上、AI-2でも9日間以上を要した⁽²⁹⁾。

⁽²⁹⁾ ただし、この場合には、仮に GPU ボードを使用しなければ、数倍の期間 を要することになるが、モデルが更に大きくなると能力面で GPU ボードが

いずれにしても、ICTの急速な進歩により、以前では想定もできなかっ たような膨大かつ強力な計算資源を、今回のようにプライベートに独占す ることが劇的に容易になったといえよう。このことは、今後の我々にとっ て大きな福音であろうし、また、そうすべきでもあろうとも思われる。

今回の調査・検討は、いわば租税法分野への AI 応用のための単なる入 り口における大変基礎的かつ雑駁なものではあったものの、AIの抽象化 能力に基づく類推機能の活用を目指した点で読者の今後の参考になれば、 当方にとって頗る幸いである。

使用できなくなることもある。また、デスクトップ・パソコンをトレーニン グ等のために長期間安定的に稼働させるためには、電力の瞬断対策が必要に なってくるし、Windows Update 等による無用な再起動を予防することも 必須である。

【参考資料1】固定的ルール方式と抽出モデル方式との比較

ja_get_keywords_by_rules	ja_for_tains_keywords3 _50_8192_GPU
利示事項 1 本件は、不動産 TAINS,keyword の買取 再販売 TAINS,keyword 等を行う 株式会社 TAINS,keyword で、全部 TAINS,keyword 又は一部 TAINS,keyword が 住宅 TAINS,keyword として 賃貸 TAINS,keyword されている 建物 TAINS,keyword の 購入 TAINS,keyword をし、これに係る 消費税額 TAINS,keyword として 申告 TAINS,keyword を 控除対象仕入税額 TAINS,keyword として 申告 TAINS,keyword を とたところ、日本橋 税務署長 TAINS,keyword から、その 全額 TAINS,keyword を 接除 TAINS,keyword を でないとして 東正処分 TAINS,keyword の 既課決定処分 TAINS,keyword を 受けるなどしたことから、 上告 TAINS,keyword 人を相手に、各 東正処分 TAINS,keyword の 取講決定処分 TAINS,keyword を 要にないとした まままままます。 「本語 TAINS,keyword を 要にないとしたことから、 上告 TAINS,keyword の 取講決定処分 TAINS,keyword を 要にならいたことから、 上告 TAINS,keyword を 要に処分 TAINS,keyword を を 要に処分 TAINS,keyword の 取講決定処分 TAINS,keyword を 要にならいたことから、 上告 TAINS,keyword を 要に処分 TAINS,keyword の 取消し TAINS,keyword 等を求める 事業 TAINS,keyword であ	判示事項 TAINS_keyword 等を行う 株式会社 TAINS_keyword である被 上告 TAINS_keyword 人が、 転売目的 TAINS_keyword で、 全部 TAINS_keyword 又は 一部 TAINS_keyword が 住宅 TAINS_keyword として 賃貸 TAINS_keyword されている 建物 TAINS_keyword の 購入 TAINS_keyword されている 建物 TAINS_keyword の 費力 TAINS_keyword を としたところ、日本橋 税務署長 TAINS_keyword から、その 全額 TAINS_keyword を としたところ、日本橋 税務署長 TAINS_keyword から、その 全額 TAINS_keyword を 控除 TAINS_keyword から、その 全額 TAINS_keyword を 控除 TAINS_keyword から、その 全額 TAINS_keyword を で表したところ、日本橋 税務署長 TAINS_keyword から、その 全額 TAINS_keyword を 受けるなどしたことから、 上告 TAINS_keyword 人を相手に、各 更正処分 TAINS_keyword のうち 申告額 TAINS_keyword を を 受けるなどしたことから、 上告 TAINS_keyword 人を相手に、各 更正処分 TAINS_keyword のうち 申告額 TAINS_keyword を を である 事業
を は	R書は、各 建物 TAINS keyword は 転売 TAINS keyword まで 住宅 TAINS keyword として 賃貸 TAINS keyword されることが見込まれていたから、各 課税仕入れ TAINS keyword は、個別対応方式 TAINS keyword による 用途区分 TAINS keyword において 共通対応課税 TAINS keyword 仕入れに 区分 TAINS keyword されるべきであり、各 更正処分 TAINS keyword は 適法 TAINS keyword であるどとした上で、各 賦課決定処分 TAINS keyword は 違法 TAINS keyword であるとして、その 取消請求 TAINS keyword を認容 TAINS keyword した。しかしながら、原審の上記 判断 TAINS keyword は 是認 TAINS keyword することができない。その 理由 TAINS keyword は、次のとおりである。
3 税務 TAINS_keyword 当局は、運ぐとも 平成17年 TAINS_keyword 以降、本件と同様の 課税仕入れ TAINS_keyword	3 税務 TAINS_keyword 当局は、運ぐとも 平成17年 TAINS_keyword 以降、本件と同様の 課税仕入れ TAINS_keyword

を、建物 TAINS_keyword が 住宅 TAINS_keyword として 賃貸	を、 建物 TAINS_keyword が 住宅 TAINS_keyword として 賃
TAINS_keyword されること(その他の資産の譲渡等 TAINS_keyword	貸 TAINS_keyword されること(その他の TAINS_keyword) 資産
に 対応 TAINS_keyword すること)に着目して 共通対応	の譲渡 TAINS_keyword 等に 対応 TAINS_keyword すること)に着
TAINS_keyword 課税仕入れ TAINS_keyword に 区分	目して 共通対応課税 TAINS_keyword 仕入れに 区分
TAINS_keyword すべきであるとの 見解 TAINS_keyword を採って	TAINS_keyword すべきであるとの 見解 TAINS_keyword を採って
おり TAINS_keyword 、そのことは、各 申告 TAINS_keyword 当	おり TAINS_keyword 、そのことは、各 申告 TAINS_keyword 当
時、 税務 TAINS_keyword 当局の 職員 TAINS_keyword が 執	時、 税務 TAINS_keyword 当局の 職員 TAINS_keyword が 執
筆 TAINS_keyword した 公刊物 TAINS_keyword や、 公表	筆 TAINS_keyword した 公刊物 TAINS_keyword や、 公表
TAINS_keyword されている 裁決例 TAINS_keyword 及び下級審の	TAINS_keyword されている 裁決例 TAINS_keyword 及び下級審の
裁判例 TAINS_keyword を通じて、 一般 TAINS_keyword の 納	裁判例 TAINS_keyword を通じて、 一般 TAINS_keyword の
税者 TAINS_keyword も知り得たものということができる。他方、それ	納税者 TAINS_keyword も知り得たものということができる。他方、そ
以前 TAINS_keyword に 税務 TAINS_keyword 当局が 作成	れ 以前 TAINS_keyword に 税務 TAINS_keyword 当局が 作
TAINS_keyword した 部内資料 TAINS_keyword や 税務	成 TAINS_keyword した 部内資料 TAINS_keyword や 税務
TAINS_keyword 当局 関係者 TAINS_keyword が編者である 公刊	TAINS_keyword 当局 関係者 TAINS_keyword が編者である 公
物 TAINS_keyword 及び平成7 年 TAINS_keyword 頃の 関係	刊物 TAINS_keyword 及び平成7 年 TAINS_keyword 頃の 関係
TAINS_keyword 機関 TAINS_keyword からの 照会	機関 TAINS_keyword からの 照会 TAINS_keyword に対する 回
TAINS_keyword に対する 回答 TAINS_keyword には、 事業者	答 TAINS_keyword には、 事業者 TAINS_keyword の 目的
TAINS_keyword の 目的 TAINS_keyword に着目して 用途区分	TAINS_keyword に着目して 用途区分 TAINS_keyword を 判定
TAINS_keyword を 判定 TAINS_keyword していたとも 理解	TAINS_keyword していたとも 理解 TAINS_keyword され得る 記
TAINS_keyword され得る 記載 TAINS_keyword 等があるものの、こ	載 TAINS_keyword 等があるものの、これらは、本件と同様の 課税
れらは、本件と同様の 課税仕入れ TAINS_keyword に 直接	仕入れ TAINS_keyword に 直接 TAINS_keyword 言及するもので
TAINS_keyword 言及するものでなく、その 趣旨 TAINS_keyword や	なく、その 趣旨 TAINS_keyword や 前提 TAINS_keyword となる
前提 TAINS_keyword となる 事実関係 TAINS_keyword が明らか	事実関係 TAINS_keyword が明らかでないなど、必ずしも上記 見
でないなど、必ずしも上記 見解 TAINS_keyword と 矛盾	解 TAINS_keyword と 矛盾 TAINS_keyword するものとはいえな
TAINS_keyword するものとはいえない。また、 税務 TAINS_keyword	い。また、 税務 TAINS_keyword 当局は、平成9 年
当局は、平成9 年 TAINS_keyword 頃、 関係 TAINS_keyword	TAINS_keyword 頃、 関係機関 TAINS_keyword からの 照会
機関 TAINS_keyword からの 照会 TAINS_keyword に対し、本件	TAINS_keyword に対し、本件と同様の 課税仕入れ TAINS_keyword
と同様の 課税仕入れ TAINS_keyword を 課税対応 TAINS_keyword	を 課税対応課税仕入れ TAINS_keyword に 区分 TAINS_keyword
課税仕入れ TAINS_keyword に 区分 TAINS_keyword すべき旨の	すべき旨の 回答 TAINS_keyword をしているが、このことから、 直
回答 TAINS_keyword をしているが、このことから、 直ちに	ちに TAINS_keyword 、 税務 TAINS_keyword 当局が 一般的
TAINS_keyword 、 税務 TAINS_keyword 当局が 一般的	TAINS_keyword に当該 課税仕入れ TAINS_keyword を 事業者
TAINS_keyword に当該 課税仕入れ TAINS_keyword を 事業者	TAINS_keyword の 目的 TAINS_keyword に着目して 課税対応
TAINS_keyword の 目的 TAINS_keyword に着目して 課税対応	課税 TAINS_keyword 仕入れに 区分 TAINS_keyword する 取扱
TAINS_keyword に 区分	い TAINS_keyword をしていたものということはできないし、上記 回答
TAINS_keyword する 取扱い TAINS_keyword をしていたものというこ	TAINS_keyword が 公表 TAINS_keyword されるなどしたとの 事
とはできないし、上記 回答 TAINS_keyword が 公表 TAINS_keyword	情 TAINS_keyword もうかがわれない。
されるなどしたとの 事情 TAINS_keyword もうかがわれない。	

```
平成17年 TAINS_keyword 以降、 税務 TAINS_keyword 当局
                                               平成17年 TAINS_keyword 以降、 税務 TAINS_keyword 当局
が、本件と同様の 課税仕入れ TAINS_keyword を 建物
                                              が、本件と同様の 課税仕入れ TAINS_keyword を 建物
TAINS_keyword が 住宅 TAINS_keyword として 賃貸
                                              TAINS_keyword が 住宅 TAINS_keyword として 賃貸
TAINS_keyword されることに着目して 共通対応 TAINS_keyword
                                              TAINS_keyword されることに着目して 共通対応課税
                                              TAINS_keyword 仕入れに 区分 TAINS_keyword する 取扱い
課税仕入れ TAINS_keyword に 区分 TAINS_keyword する 取
扱い TAINS_keyword を 周知 TAINS_keyword するなどの 積極的
                                              TAINS_keyword を 周知 TAINS_keyword するなどの 積極的
TAINS_keyword な 措置 TAINS_keyword を講じていないとしても、
                                              TAINS_keyword な 措置 TAINS_keyword を講じていないとしても、
事業者 TAINS_keyword としては、上記 取扱い TAINS_keyword
                                              事業者 TAINS_keyword としては、上記 取扱い TAINS_keyword
がされる 可能性 TAINS_keyword を 認識 TAINS_keyword してし
                                              がされる 可能性 TAINS_keyword を 認識 TAINS_keyword してし
かるべきであったということができる。そして、上記 取扱い
                                              かるべきであったということができる。そして、上記 取扱い
TAINS_keyword は 消費税法 TAINS_keyword 30条2項1 号
                                              TAINS_keyword は 消費税法 TAINS_keyword 30条2項1 号
TAINS_keyword の 文理 TAINS_keyword 等に照らして自然であると
                                              TAINS_keyword の 文理 TAINS_keyword 等に照らして自然である
いえ、各 申告 TAINS_keyword 当時、本件と同様の 課税仕入れ
                                              といえ、各 申告 TAINS_keyword 当時、本件と同様の 課税仕入れ
TAINS_keyword を 事業者 TAINS_keyword の 目的
                                              TAINS_keyword を 事業者 TAINS_keyword の 目的
TAINS_keyword に着目して 課税対応 TAINS_keyword 課税仕入
                                              TAINS_keyword に着目して 課税対応課税 TAINS_keyword 仕入
れ TAINS_keyword に 区分 TAINS_keyword すべきものとした 裁
                                              れに 区分 TAINS_keyword すべきものとした 裁判例
判例 TAINS_keyword 等があったともうかがわれないこと等をも 考慮
                                              TAINS_keyword 等があったともうかがわれないこと等をも 考慮
TAINS_keyword すれば、被 上告 TAINS_keyword 人が各 課税仕
                                              TAINS_keyword すれば、被 上告 TAINS_keyword 人が各 課税
入れ TAINS keyword を 課税対応 TAINS keyword 課税仕入れ
                                              仕入れ TAINS_keyword を 課税対応課税仕入れ TAINS_keyword
TAINS_keyword して 控除対象仕入税額
                                             に 区分 TAINS_keyword して 控除対象仕入税額 TAINS_keyword
TAINS_keyword の 計算 TAINS_keyword をしたことにつき、真に
                                              の 計算 TAINS_keyword をしたことにつき、真に 納税者
納税者 TAINS_keyword の責めに帰することのできない 客観的な事
                                              TAINS_keyword の責めに帰することのできない 客観的な事情
情 TAINS_keyword があり、 過少申告加算税 TAINS_keyword の
                                              TAINS_keyword があり、 過少申告加算税 TAINS_keyword の
趣旨 TAINS_keyword に照らしてもなお 納税者 TAINS_keyword に
                                              趣旨 TAINS_keyword に照らしてもなお 納税者 TAINS_keyword
過少申告加算税 TAINS_keyword を 賦課 TAINS_keyword するこ
                                             に 過少申告加算税 TAINS_keyword を 賦課 TAINS_keyword す
とが 不当 TAINS_keyword 又は酷になるということはできない。
                                              ることが 不当 TAINS_keyword 又は酷になるということはできない。
以上によれば、被 上告 TAINS_keyword 人が各 課税仕入れ
                                               以上によれば、被 上告 TAINS_keyword 人が各 課税仕入れ
TAINS_keyword に係る 消費税額 TAINS_keyword の 全額
                                              TAINS_keyword に係る 消費税額 TAINS_keyword の 全額
TAINS_keyword を 控除 TAINS_keyword したことにつき、 国税通
                                              TAINS_keyword を 控除 TAINS_keyword したことにつき、 国税
則法 TAINS_keyword 65条4項にいう「 正当な理由 TAINS_keyword
                                              通則法 TAINS_keyword 65条4項にいう 正当な理由
」があると認めることはできない。以上と異なる原審の 判断
                                              TAINS_keyword 」があると認めることはできない。以上と異なる原審の
TAINS_keyword には、 判決 TAINS_keyword に 影響
                                              判断 TAINS_keyword には、 判決 TAINS_keyword に 影響
TAINS_keyword を及ぼすことが明らかな 法令 TAINS_keyword の
                                              TAINS_keyword を及ぼすことが明らかな 法令 TAINS_keyword の
違反 TAINS_keyword がある。論旨は 理由 TAINS_keyword があ
                                              違反 TAINS_keyword がある。論旨は 理由 TAINS_keyword が
```



【参考資料2】5分割分析表

言語モデル 5分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
1000_512 第 1 階級	586	$\frac{B-(1)}{B-(1)}$ (503) (旧法人 (99) / 大蔵省令 (65) / 揮発油税法 (43) / 国税に関する (39) / 第4 7 (38) / 売上値引き (37) / 第 1 1 (36) / 課稅仕入れに係る (35) / 対価として (32) / 適周を受けようとする (31) / 不知又 (27) / 裁決があったこと (25) / 譲渡所得の (24) / 私人間 (24) / 株式等に係る (23) / 過納付 (23) / 時機に後れた攻撃 (22) / は脱し (22) / 不動産貸付業
1000_512 第 2 階級	586	電話代金 ⁽²⁾ /青果物類 ⁽²⁾ /飲食店舗 ⁽²⁾ /駐車場所 ⁽²⁾ /不正当 ⁽²⁾ /不正義 ⁽²⁾ /事業を ⁽²⁾ /再開院 ⁽²⁾ /利用日 ⁽²⁾ /割賦払 ⁽²⁾ /労働費 ⁽²⁾ /同一の ⁽²⁾ /地域間 ⁽²⁾ /契約料 ⁽²⁾ /学会費 ⁽²⁾ /実親子 ⁽²⁾ /弁護料 ⁽²⁾ /支出日 ⁽²⁾ /支給前 ⁽²⁾ /有利息 ⁽²⁾ /期間の ⁽²⁾ /業とし ⁽²⁾ /注文日 ⁽²⁾ /拾殊の ⁽²⁾ /第13 ⁽²⁾ /第23 ⁽²⁾ /第38 ⁽²⁾ /管理業 ⁽²⁾ /終利益 ⁽²⁾ /課稅後 ⁽²⁾ /譲渡書 ⁽²⁾ /受愛権 ⁽²⁾ /会額表 ⁽²⁾ /長寿会 ⁽²⁾ /入一方。 ⁽²⁾ /大理 ⁽²⁾ /官名 ⁽²⁾ /所府 ⁽²⁾ /事業の用に供されることなく取り壊されているの ⁽¹⁾ /人的役務の提供を主たる内容とする事業を行う者 ⁽¹⁾ /人的役務の提供を主たる内容とする事業の範囲 ⁽¹⁾ /公益を目的とする事業の用に供していない場合 ⁽¹⁾ /その相続の開始のあったことを知った日 ⁽¹⁾ /支出した日の属する事業年度の終了の日 ⁽¹⁾ /研究開発費及びソフトウエアの会計処理 ⁽¹⁾ /指定介護予防特定施設入居者生活介護 ⁽¹⁾ /特定上場株式等に係る譲渡所得等の ⁽¹⁾ /納税者の責めに帰すことのできない ⁽¹⁾ /財産に関する帳簿書類その他の物件 ⁽¹⁾ /先行取得資産に係る買換えの特例 ⁽¹⁾ /収入を得るために直接要した金額 ⁽¹⁾ /固定資産課税台帳登録事項証明書 ⁽¹⁾ /敷地の用に供されていることから ⁽¹⁾ /相続開始があったことを知った日 ⁽¹⁾ /調査について必要があると認めて ⁽¹⁾ /通常取引されると認められる株式 ⁽¹⁾ /非末造家屋経年減点補正率基準表 ⁽¹⁾ /事業の用に供さなくなった理由 ⁽¹⁾ /処分があったことを知ったこと ⁽¹⁾ /労務その他の役務の提供の対価 ⁽¹⁾ /収入を得るために支出した費用 ⁽¹⁾ /居住の用に供さなくなったとき ⁽¹⁾ /最初連結親法人事業年度の開始 ⁽¹⁾ /以入を得るために支出した費用 ⁽¹⁾ /給与所得において業務独占資格 ⁽¹⁾ /通貨オブション取引事前確認書 ⁽¹⁾ /その他の資産の譲渡等の対価 ⁽¹⁾ /社入税額控除に関する明細書 ⁽¹⁾ /順位を失ったことにより損失 ⁽¹⁾ /での他の資産の譲渡等の対価 ⁽¹⁾ /社入税額控除に関する明細書 ⁽¹⁾ /収入を得るために必要な費用 ⁽¹⁾ /収入を得るために必要な費用 ⁽¹⁾ /収入を得るために必要な費用 ⁽¹⁾ /収入を得るために用いられる ⁽¹⁾ /土地の無償返還に関する届出 ⁽¹⁾ /小規模多機能型居宅介護施設 ⁽¹⁾ /【他508件】

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの(ただし、同じ認識回数では文字数のより多い順)
1000_512 第 3 階級	586	新借地権の設定(1)/新設分割の効力(1)/旧市街地改造法(1)/旧耐用年数通達(1)/時価総額の評価(1)/更正があった日(1)/更正及び再更正(1)/書面による契約(1)/期限内申告書を(1)/未納稅法人稅額(1)/本人確認記錄票(1)/本来の相続稅額(1)/本来の職務執行(1)/条件変更承諾料(1)/架空業務委託費(1)/株券電子化会社(1)/株式を市場価格(1)/株式価額の低下(1)/株式等売買損益(1)/株式譲渡損失額(1)/業務委託費差額(1)/権利義務の確定(1)/残余財産請求権(1)/残存期間の定め(1)/法人がある資産(1)/法定相続人全員(1)/海運代理店業務(1)/消費稅還付請求(1)/消費者金融会社(1)/清算分割請求権(1)/無人時間貸駐車(1)/無(1)/無形資産の使用(1)/照会事項(エ)(1)/物納財産の収納(1)/特例物納許可額(1)/特別譲渡予定地(1)/無比資産の使用(1)/期金事項(エ)(1)/物納財産の収納(1)/特例物納許可額(1)/特別譲渡予定地(1)/特定商品取引法(1)/特定寄附金の額(1)/特定市街化区域(1)/特許権侵害訴訟(1)/組占的支配権能(1)/独立企業の原則(1)/生命保険受取人(1)/生命保険契約者(1)/生命保険積立金(1)/申出があった日(1)/直接審查請求人(1)/相当の地代の額(1)/相当の期間内で(1)/相続稅申告書類(1)/相続稅路線価額(1)/相続開始の前日(1)/相続開始の前日(1)/相当の地代の額(1)/相当の期間内で(1)/相稅稅申告書類(1)/相続稅路線価額(1)/相続開始の前日(1)/相続開始の時期(1)/知り得べく状態(1)/砂利採取区域内(1)/確定した対価額(1)/確定申告の承認(1)/確定申告還付金(1)/私法上の法制度(1)/私生活上の利益(1)/組稅法上の行為(1)/稅務会計事務所(1)/稅理土補助業務(1)/立退移転補價金(1)/第三者に立会い(1)/簿外現金出納帳(1)/精神障害者福祉(1)/納稅申告の勧奨(1)/累積損金不算入(1)/組織法上の効果(1)/経済的対応関係(1)/精神障害者福祉(1)/納稅申告の勧奨(1)/累積損金不算入(1)/組織法上の効果(1)/経済的対応関係(1)/結与事務担当者(1)/給料支払請求権(1)/臨時時別措置法(1)/臨時駐車場収入(1)/自走式二段立体(1)/被相続人と住所(1)/被相続人と生計(1)/被相続人と規族(1)/被相続人の居住(1)/記載しない帳簿(1)/許可があった日(1)/許偽その他不正(1)/課代表低限度額(1)/課稅決定通知書(1)/課稅関係の終了(1)/調查時提出書類(1)/【他 495件】
1000_512 第 4 階級	586	工事の終了 ⁽¹⁾ /工事指導料 ⁽¹⁾ /工事監理費 ⁽¹⁾ /工事着手金 ⁽¹⁾ /差押処分庁 ⁽¹⁾ /市街地価格 ⁽¹⁾ /建物用敷地 ⁽¹⁾ /建物譲渡損 ⁽¹⁾ /建築正事費 ⁽¹⁾ /建築施行業 ⁽¹⁾ /建築計画表 ⁽¹⁾ /升政連関連 ⁽¹⁾ /升済の猶予 ⁽¹⁾ /升護士名簿 ⁽¹⁾ /引当不足額 ⁽¹⁾ /引渡しない ⁽¹⁾ / 当期未処理 ⁽¹⁾ / 当然加入制 ⁽¹⁾ / 形得稅債務 ⁽¹⁾ / 所得稅確定 ⁽¹⁾ / 所得稍償金 ⁽¹⁾ / 于数料方式 ⁽¹⁾ / 大養配偶者 ⁽¹⁾ / 大按術使用權 ⁽¹⁾ / 大核的判断 ⁽¹⁾ / 投資信託 ⁽¹⁾ / 投資受益權 ⁽¹⁾ / 担保權確保 ⁽¹⁾ / 招集手続き ⁽¹⁾ / 控除保険料 ⁽¹⁾ / 定除外国稅 ⁽¹⁾ / 支払予定額 ⁽¹⁾ / 支払事務所 ⁽¹⁾ / 支払可能額 ⁽¹⁾ / 支給取消し ⁽¹⁾ / 支給年月日 ⁽¹⁾ / 改正稅法の ⁽¹⁾ / 放棄の効力 ⁽¹⁾ / 整理手続き ⁽¹⁾ / 新たな事由 ⁽¹⁾ / 新たな取得 ⁽¹⁾ / 新东 全費用 ⁽¹⁾ / 新幹線代金 ⁽¹⁾ / 施行日の後 ⁽¹⁾ / 施行日以前 ⁽¹⁾ / 旅館用建物 ⁽¹⁾ / 上本銀行法 ⁽¹⁾ / 更新請求權 ⁽¹⁾ / 更正の利益 ⁽¹⁾ / 更正の権限 ⁽¹⁾ / 書面による ⁽¹⁾ / 最高価申込 ⁽¹⁾ / 期間満了前 ⁽¹⁾ / 未納付税額 ⁽¹⁾ / 未計上利益 ⁽¹⁾ / 東京都心部 ⁽¹⁾ / 株式投資家 ⁽¹⁾ / 機而重整 ⁽¹⁾ / (機成工業会 ⁽¹⁾ / 生工程度 ⁽¹⁾ / (業務用機械 ⁽¹⁾ / 業務関連費 ⁽¹⁾ / (規設告書 ⁽¹⁾ / 構造計算書 ⁽¹⁾ / 標準住宅地 ⁽¹⁾ / 機械、装置 ⁽¹⁾ / (機械工業会 ⁽¹⁾ / 正当な注意 ⁽¹⁾ / 歯科医療用 ⁽¹⁾ / 比準経費率 ⁽¹⁾ / (民事則定法 ⁽¹⁾ / 民事更生法 ⁽¹⁾ / 人民。 其位 ⁽¹⁾ / (民事 建 全 ⁽¹⁾ / (民事 重 ⁽¹⁾ / (表) 有助 (本 統) ⁽¹⁾ / (持) 被収金 ⁽¹⁾ / 特別裁判所 ⁽¹⁾ / (特計権利者 ⁽¹⁾ / 独自の権利 ⁽¹⁾ / 现受付決済 ⁽¹⁾ / 生計の維持 ⁽¹⁾ / 特別 你

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
1000_512 第 5 階級	589	投資時期 ⁽¹⁾ /投資財産 ⁽¹⁾ /抽出資料 ⁽¹⁾ /担保価格 ⁽¹⁾ /担保取引 ⁽¹⁾ /拒否事由 ⁽¹⁾ /指定管理 ⁽¹⁾ /控除漏れ ⁽¹⁾ /推計基準 ⁽¹⁾ /提供行為 ⁽¹⁾ /提出主義 ⁽¹⁾ /换算所得 ⁽¹⁾ /换算金額 ⁽¹⁾ /損失概念 ⁽¹⁾ /損失者間 ⁽¹⁾ /損益処理 ⁽¹⁾ /支払う者 ⁽¹⁾ /支払主体 ⁽¹⁾ /支払割合 ⁽¹⁾ /支援手当 ⁽¹⁾ /支給ごと ⁽¹⁾ /支給の額 ⁽¹⁾ /支給內容 ⁽¹⁾ /支給控除 ⁽¹⁾ /支給時間 ⁽¹⁾ /支配実態 ⁽¹⁾ /大正使取理 ⁽¹⁾ /所有上,如
500_1024 第 1 階級	695	第35 (222)/第14 (146)/第12 (94)/第28 (72)/第19 (22)/旧法人税法(12)/称制(12)/氏続(11)/裁決があったことを知ったもの(10)/相当の期間内(10)/第27 (10)/第12項(9)/1口当たり(8)/船食業者(8)/買取り等の申出のあった日(7)/価額に相当する価額(7)/取得した時期(7)/取立金額(7)/税務に関する専門的知識(6)/居住の用に供するため(6)/外貨建借入金相当(6)/本来納付すべき(6)/管理委託料(6)/滞納金額(6)/貸付金又(6)/費差額(6)/素(3)表決があったことを知った(5)/納税義務の適正な履行(5)/適用を受けるべき(5)/滞納者以外の者(5)/取引委託契約(5)/第23条2項(5)/医療事業(5)/収入時期(6)/第19項(5)/追納付(5)/青色(5)/滞納者以外の者(6)/取引委託契約(5)/第23条2項(5)/医療事業(5)/収入時期(6)/第19項(5)/追納付(5)/青色(5)/滞納者以外の者(6)/取引委託契約(6)/生地の賃借権の設定(4)/給与所得の源泉徴収(4)/固定資産の譲渡等(4)/貸付金返還請求権(4)/分割方法の指定(4)/年金給付の総額(4)/業務についての(4)/特に有利な価額(4)/損害賠償訴訟(4)/支払った時点(4)/相続人が承継(4)/資産の取得に(4)/取得の時期(4)/合理的資料(4)/本来の用法(4)/経理保険料(4)/給料手当で(4)/被保険車両(4)/固入資金(4)/償還時点(4)/前橋市長(4)/取引会社(4)/対象事項(4)/対象金額(4)/年度ごと(4)/第18項(4)/第18(4)/相続の開始があったことを知ったもの(3)/民法の規定による相続分の割合(3)/所得税の確定申告書の提出(3)/退職により支払われたもの(3)/適用を受けようとする場合(3)/本来の納税義務者以外(3)/譲渡の時における価額(3)/適用を受けるべきもの(3)/課稅仕入れ等に係る(3)/本来の納税義務者以外(3)/組続税の申告納付(3)/組続税の申告新付(3)/間与税の申告手続(3)/選択不適用届出書(3)/遺留分減役請求書(3)/不動産賃貸契約(3)/交付を受けた者(5)/使用者以外の者(3)/原則としながら(3)/国に対する請求(3)/実質所得者課稅(3)/課稅売上高割合(3)/不動産鑑定業(3)/事実上の合意(3)/債権差押通知(3)/合理的な事情(3)/売上げに係る(3)/委任契約締結(3)/実質的当事者(3)/【他601件】

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
500_1024 第 2 階級	695	中小企業緊急雇用安定助成金 ⁽¹⁾ /仕入れに係る消費稅額の特例 ⁽¹⁾ /外国為替取引口座開設申込書 ⁽¹⁾ /外部からうかがい得るような ⁽¹⁾ /宅地建物取引主任者資格試験 ⁽¹⁾ /居住の用に供されなくなった ⁽¹⁾ /履行の確実と認められる債務 ⁽¹⁾ /所有権移転登記請求権仮登記 ⁽¹⁾ /指定特定施設入居者生活介護 ⁽¹⁾ /控除を受けることはできない ⁽¹⁾ /欠損金額の繰戻しによる還付 ⁽¹⁾ /納付すべきこととされるもの ⁽¹⁾ /自己の居住の用に供している ⁽¹⁾ /裁決があったことを知ったと ⁽¹⁾ /貸倒れに係る消費稅額の控除 ⁽¹⁾ /通増定期保険特勢付定期保険 ⁽¹⁾ /非未造家屋経年減点補正率表 ⁽¹⁾ /上場株式等の配当等の金額 ⁽¹⁾ /事業の用に供している割合 ⁽¹⁾ /事業の用に供している資産 ⁽¹⁾ /先行取得資産に係る買換え ⁽¹⁾ /公法上の不当利得返還請求 ⁽¹⁾ /公益を目的とする事業の用 ⁽¹⁾ /効力を失うことになるもの ⁽¹⁾ /収入を得るために支出した ⁽¹⁾ /収用等に係る資産の買取り ⁽¹⁾ /取得予定資産明細書の添付 ⁽¹⁾ /引き続き居住の用に供する ⁽¹⁾ /時効の利益を受けるべき者 ⁽¹⁾ /決定があるべきことを十分 ⁽¹⁾ /生活の保護に寄与すること ⁽¹⁾ /申告期限後3年以内の分割 ⁽¹⁾ /相続開始時の相続稅評価額 ⁽¹⁾ /確定判決の効力が生じた時 ⁽¹⁾ /確定申告書等作成システム ⁽¹⁾ /算定の基礎となるべき事実 ⁽¹⁾ /買取りの申出があったこと ⁽¹⁾ /貸倒引当金繰入限度超過額 ⁽¹⁾ /適用に関する明細書の添付 ⁽¹⁾ /インセンティブ再配賦額 ⁽¹⁾ /インセンティブ資本勘定 ⁽¹⁾ /インターネット接続料金 ⁽¹⁾ /ジャスダック証券取引所 ⁽¹⁾ /ベトナム社会主義共和国 ⁽¹⁾ /上場有価証券等以外株式 ⁽¹⁾ /不動産長期譲渡所得金額 ⁽¹⁾ /不相当に高額な部分の額 ⁽¹⁾ /事前通知を要しない理由 ⁽¹⁾ /保有個人情報開示請求書 ⁽¹⁾ /慎うたとができない損害 ⁽¹⁾ /取りために要した費用 ⁽¹⁾ /上場有価証券等以外株式 ⁽¹⁾ /不動産長期譲渡所得金額 ⁽¹⁾ /不相当に高額な部分の額 ⁽¹⁾ /事前通知を要しない理由 ⁽¹⁾ /保有個人情報開示請求書 ⁽¹⁾ /債うたどができない損害が入り、類別の表別の表別の表別の表別の表別の表別の表別の表別の表別の表別の表別の表別の表別
500_1024 第 3 階級	695	保険料領収書 ⁽¹⁾ /保険金支払額 ⁽¹⁾ /信用情報機関 ⁽¹⁾ /信託契約締結 ⁽¹⁾ /修正前契約書 ⁽¹⁾ /個人所得課稅 ⁽¹⁾ /個人間の贈与 ⁽¹⁾ /個別的必要性 ⁽¹⁾ /個別鑑定評価 ⁽¹⁾ /借地権の移転 ⁽¹⁾ /債務整理費用 ⁽¹⁾ /債務者適格性 ⁽¹⁾ / 債権放棄契約 ⁽¹⁾ /債権譲渡取引 ⁽¹⁾ /債権譲渡承諾 ⁽¹⁾ /成元为象財産 ⁽¹⁾ /次害防止計画 ⁽¹⁾ /公正ナル価格 ⁽¹⁾ /公法上の契約 ⁽¹⁾ /共同事業者間 ⁽¹⁾ /共同使用部分 ⁽¹⁾ /共同施行方式 ⁽¹⁾ /共同生活期間 ⁽¹⁾ /共同社員旅行 ⁽¹⁾ /共同開発契約 ⁽¹⁾ /内装工事費用 ⁽¹⁾ /再分割協議書 ⁽¹⁾ /再委託の禁止 ⁽¹⁾ /再建計画期間 ⁽¹⁾ /再生計画変更 ⁽¹⁾ /処分代金債権 ⁽¹⁾ /処分要件事実 ⁽¹⁾ /処理委託契約 ⁽¹⁾ /出納等明細書 ⁽¹⁾ /出資持分放棄 ⁽¹⁾ /分割対価の額 ⁽¹⁾ /分離先物取引 ⁽¹⁾ /刑事事件書類 ⁽¹⁾ /判断基準時点 ⁽¹⁾ /別表第一10 ⁽¹⁾ /利子稅相当額 ⁽¹⁾ /利用実態調查 ⁽¹⁾ /利用許諾要素 ⁽¹⁾ /利益積立金又 ⁽¹⁾ /利益配分割合 ⁽¹⁾ /制限納稅義務 ⁽¹⁾ /剩余金の配分 ⁽¹⁾ /労働者の募集 ⁽¹⁾ /医療費負担者 ⁽¹⁾ /匿名組合事業 ⁽¹⁾ /単なる案内役 ⁽¹⁾ /厚生労働省医 ⁽¹⁾ /収入金額割合 ⁽¹⁾ /収益対応基準 ⁽¹⁾ /取付工事費用 ⁽¹⁾ /取引運送収入 ⁽¹⁾ /受任者の死亡 ⁽¹⁾ /受取利息収入 ⁽¹⁾ /受益者の意思 ⁽¹⁾ /受領事業年度 ⁽¹⁾ /司法書士資格 ⁽¹⁾ /各相続人名義 ⁽¹⁾ /与理的な計算 ⁽¹⁾ /同一の委託者 ⁽¹⁾ /同意する旨の ⁽¹⁾ /可業者比準法 ⁽¹⁾ /同業者比準率 ⁽¹⁾ /高品代金債権 ⁽¹⁾ /高品販売債権 ⁽¹⁾ /商業協同組合 ⁽¹⁾ /固定資産稅法 ⁽¹⁾ /固定資産稅法 ⁽¹⁾ /固定資産稅法 ⁽¹⁾ /国产資産課稅 ⁽¹⁾ /国内関連企業 ⁽¹⁾ /国家賠償債務 ⁽¹⁾ /国家賠償養務 ⁽¹⁾ /国庫補助金又 ⁽¹⁾ /国民健康条例 ⁽¹⁾ /土地使用貸借 ⁽¹⁾ /土地売買予約 ⁽¹⁾ /地代収受割合 ⁽¹⁾ /地方稅の還付 ⁽¹⁾ /執行対象財産 ⁽¹⁾ /基準人口比率 ⁽¹⁾ /基準年度ごと ⁽¹⁾ /增差所得稅額 ⁽¹⁾ /赤上記帳割合 ⁽¹⁾ /未予設置認可 ⁽¹⁾ /【他 595件】

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
500_1024 第 4 階級	695	慰謝料債務 ⁽¹⁾ (慶弔禍福金 ⁽¹⁾)所得の実態 ⁽¹⁾ /所得の配当 ⁽¹⁾ /所得申告書 ⁽¹⁾ /所得税計算 ⁽¹⁾ /所有権帰属 ⁽¹⁾ / 手形債権者 ⁽¹⁾ /承諾期間中 ⁽¹⁾ /抵当債権者 ⁽¹⁾ /担保の方法 ⁽¹⁾ /担保抹消料 ⁽¹⁾ /指定相談官 ⁽¹⁾ /指定製作業 ⁽¹⁾ /指定通知書 ⁽¹⁾ /接待內訳書 ⁽¹⁾ /支缺分《大 ⁽¹⁾ /支払任告書 ⁽¹⁾ /规策的規定 ⁽¹⁾ /教育事務所 ⁽¹⁾ /教育事務所 ⁽¹⁾ /教育事務所 ⁽¹⁾ /教育描導科 ⁽¹⁾ /敷地権割合 ⁽¹⁾ /敷金保証金 ⁽¹⁾ /文書収受印 ⁽¹⁾ /文書所持人 ⁽¹⁾ /新規不動産 ⁽¹⁾ /新規從業員 ⁽¹⁾ /施設預託金 ⁽¹⁾ /敷地権割合 ⁽¹⁾ /數金保証金 ⁽¹⁾ /文書収受印 ⁽¹⁾ /文書所持人 ⁽¹⁾ /前規不動産 ⁽¹⁾ /新規從業員 ⁽¹⁾ / 施設預託金 ⁽¹⁾ / 旅费精算書 ⁽¹⁾ /旧税理土法 ⁽¹⁾ /明示的指定 ⁽¹⁾ /明細ノート ⁽¹⁾ /時価の時価 ⁽¹⁾ /時価認定類 ⁽¹⁾ /時間的割合 ⁽¹⁾ /書與作成費 ⁽¹⁾ /書與作成費 ⁽¹⁾ /最終加工業 ⁽¹⁾ /最終返済日 ⁽¹⁾ /期末在庫品 ⁽¹⁾ /期間商了日 ⁽¹⁾ /期間経過分 ⁽¹⁾ /未分割申告 ⁽¹⁾ /未払交通費 ⁽¹⁾ /未払仲介料 ⁽¹⁾ /未払報酬金 ⁽¹⁾ /架空売上高 ⁽¹⁾ /株主の帰属 ⁽¹⁾ /株式優待金 ⁽¹⁾ /株式取得額 ⁽¹⁾ /根抵当債務 ⁽¹⁾ /検収報告書 ⁽¹⁾ /業務の禁止 ⁽¹⁾ /業務代理人 ⁽¹⁾ /業務外死亡 ⁽¹⁾ /業務对価性 ⁽¹⁾ /業務用車両 ⁽¹⁾ /業者登録票 ⁽¹⁾ /概算控除額 ⁽¹⁾ /構成指導科 ⁽¹⁾ /構築物ごと ⁽¹⁾ /標準地価格 ⁽¹⁾ /標準評点表 ⁽¹⁾ /推利の成立 ⁽¹⁾ /推利簽務者 ⁽¹⁾ /権利金支払 ⁽¹⁾ /投損全控除 ⁽¹⁾ /正当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /工三当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /正当な時価 ⁽¹⁾ /注律適科医師業 ⁽¹⁾ /協科矯正医 ⁽¹⁾ /表介財産空間 ⁽¹⁾ /上の範囲 ⁽¹⁾ /法分の範囲 ⁽¹⁾ /法人の事理 ⁽¹⁾ /法人の彰囲 ⁽¹⁾ /法人等事業 ⁽¹⁾ /法令の効力 ⁽¹⁾ /法定申告期 ⁽¹⁾ /法律的効力 ⁽¹⁾ /法律適合性 ⁽¹⁾ /法人の範囲 ⁽¹⁾ /法人の前間 ⁽¹⁾ /消滅の範囲 ⁽¹⁾ /消費稅債務 ⁽¹⁾ /清算の開始 ⁽¹⁾ /減額請求権 ⁽¹⁾ /測量調查費 ⁽¹⁾ /源泉徵収者 ⁽¹⁾ /半ずる処分 ⁽¹⁾ /準備期間中 ⁽¹⁾ /準備的工事 ⁽¹⁾ /滞納所得稅 ⁽¹⁾ /【他 583件】
500_1024 第 5 階級	697	新破産法 ⁽¹⁾ /新規資産 ⁽¹⁾ /新規購入 ⁽¹⁾ /施工費用 ⁽¹⁾ /施行決定 ⁽¹⁾ /施術料金 ⁽¹⁾ /既存船舶 ⁽¹⁾ /日常業務 ⁽¹⁾ /早期納付 ⁽¹⁾ /旭川市民 ⁽¹⁾ /旭川市民 ⁽¹⁾ /普通契約 ⁽¹⁾ /更正決議 ⁽¹⁾ /最終期限 ⁽¹⁾ 人最終株価 ⁽¹⁾ (最高額法 ⁽¹⁾ /育價委任 ⁽¹⁾ /期末一括 ⁽¹⁾ /期間中断 ⁽¹⁾ /未収措置 ⁽¹⁾ /未払期間 ⁽¹⁾ /未納期間 ⁽¹⁾ /根迎事由 ⁽¹⁾ /根拠基準 ⁽¹⁾ /棚卸過少 ⁽¹⁾ /検査基準 ⁽¹⁾ /株主法人 ⁽¹⁾ /株価比準 ⁽¹⁾ /株式代金 ⁽¹⁾ /株式会員 ⁽¹⁾ /根拠事由 ⁽¹⁾ /根拠基準 ⁽¹⁾ /棚卸過少 ⁽¹⁾ /検査基準 ⁽¹⁾ /業務徒事 ⁽¹⁾ /業務決定 ⁽¹⁾ /業務部門 ⁽¹⁾ /概算請求 ⁽¹⁾ /構成機器 ⁽¹⁾ /構成機器 ⁽¹⁾ /構造ごと ⁽¹⁾ /構造基準 ⁽¹⁾ /推刊登記 ⁽¹⁾ /機能価値 ⁽¹⁾ /機能停止 ⁽¹⁾ /欠損企業 ⁽¹⁾ /欠損申告 ⁽¹⁾ /水产時又 ⁽¹⁾ /残余金額 ⁽¹⁾ /残補償金 ⁽¹⁾ /比半山林 ⁽¹⁾ /比率年度 ⁽¹⁾ /比較資料 ⁽¹⁾ /民事局長 ⁽¹⁾ /水產製品 ⁽¹⁾ /水資源局 ⁽¹⁾ /求人広告 ⁽¹⁾ /決算表示 ⁽¹⁾ /法介金額 ⁽¹⁾ /法人企業 ⁽¹⁾ /法人格者 ⁽¹⁾ /法人経営 ⁽¹⁾ /法人資産 ⁽¹⁾ /法定多数 ⁽¹⁾ /法律制度 ⁽¹⁾ /法律条項 ⁽¹⁾ /法往根拠 ⁽¹⁾ /法)人格者 ⁽¹⁾ /法入经营 ⁽¹⁾ /法入前算金的 ⁽¹⁾ /法律根地 ⁽¹⁾ /法的資格 ⁽¹⁾ /消算金的 ⁽¹⁾ /消算金的 ⁽¹⁾ /消算金的 ⁽¹⁾ /消算金的 ⁽¹⁾ /海向寿者 ⁽¹⁾ /海向寿者 ⁽¹⁾ /特別免許 ⁽¹⁾ /特別利用 ⁽¹⁾ /特別収益 ⁽¹⁾ /特別立法 ⁽¹⁾ /特定区域 ⁽¹⁾ /特定区域 ⁽¹⁾ /特定店轴 ⁽¹⁾ /特定路線 ⁽¹⁾ /特別利用 ⁽¹⁾ /特別収益 ⁽¹⁾ /特別立法 ⁽¹⁾ /特定の額 ⁽¹⁾ /特定区域 ⁽¹⁾ /特定店轴 ⁽¹⁾ /特定路線 ⁽¹⁾ /状況確認 ⁽¹⁾ /状況確認 ⁽¹⁾ /状況調查 ⁽¹⁾ /指予国税 ⁽¹⁾ /現の取得 ⁽¹⁾ /现金配当 ⁽¹⁾ /共产路線 ⁽¹⁾ /共产活相談 ⁽¹⁾ /申告ごと ⁽¹⁾ /申告名義 ⁽¹⁾ /期の取得 ⁽¹⁾ /现金収受 ⁽¹⁾ /现金配当 ⁽¹⁾ /集合配当 ⁽¹⁾ /整行名義 ⁽¹⁾ /発行到由 ⁽¹⁾ /整記経過 ⁽¹⁾ /登録处分 ⁽¹⁾ /登録手続 ⁽¹⁾ /登録面積 ⁽¹⁾ /监查要件 ⁽¹⁾ /【他 572件】

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
200_2048 第1階級	1,834	第 3 7 ⁽³³⁴⁾ /不可能又 ⁽¹²⁹⁾ /第 2 2 ⁽⁶⁸⁾ /総収入金額に算入 ⁽³⁴⁾ /第 2 0 ⁽³²⁾ /取得した日 ⁽³⁰⁾ /第 2 4 ⁽³⁰⁾ /一般人 ⁽²³⁾ /支払金 ⁽²³⁾ /申告書の提出 ⁽²²⁾ /第 1 6 ⁽¹⁹⁾ /外国産原 ⁽¹⁸⁾ /法規定 ⁽¹⁸⁾ /確定した時 ⁽¹⁸⁾ /課税仕入れに係る 支払対価 ⁽¹⁷⁾ /譲渡所得の基因 ⁽¹⁷⁾ /原則とし、 ⁽¹⁶⁾ /所轄税務署管内 ⁽¹⁶⁾ /会社財産 ⁽¹⁵⁾ /国税通則法 2 3 条 4 項 ⁽¹⁵⁾ /売買契約締結時 ⁽¹⁵⁾ /工事収入金額 ⁽¹⁵⁾ /更正すべき理由がない旨の通知 ⁽¹⁵⁾ /収入の原因 ⁽¹³⁾ /同様な事情 ⁽¹⁵⁾ /比得額 ⁽¹⁴⁾ /販売費、一般管理費 ⁽¹⁴⁾ /これらに類する ⁽¹³⁾ /处分通知書 ⁽¹³⁾ /収入の原因 ⁽¹³⁾ /同様な事情 ⁽¹³⁾ /土地価額 ⁽¹³⁾ /月末日 ⁽¹³⁾ /業務上の必要性 ⁽¹³⁾ /区分所有法 ⁽¹²⁾ /更正すべき理由のない旨の通知処分 ⁽¹²⁾ /業務譲渡 ⁽¹²⁾ /事実上の根拠 ⁽¹¹⁾ /契象物 ⁽¹¹⁾ /支払代金 ⁽¹¹⁾ /月平均 ⁽¹¹⁾ /本来所得 ⁽¹¹⁾ /限議審理 ⁽¹¹⁾ /税法の規定 ⁽¹¹⁾ /継続的に行う事業 ⁽¹¹⁾ /評価差額に対する法人税額 ⁽¹¹⁾ /課税資産の譲渡等の時期 ⁽¹¹⁾ /取引益 ⁽¹⁰⁾ /国税債権者 ⁽¹⁰⁾ /引渡 上のあった日 ⁽¹⁰⁾ /特定事業用資産の買換え ⁽¹⁰⁾ /納付所得税額 ⁽¹⁰⁾ /組合法 ⁽¹⁰⁾ /被相続人以外 ⁽¹⁰⁾ /贈与金 ⁽¹⁰⁾ /適正退職給与 ⁽¹⁰⁾ /その他の不動産 ⁽⁹⁾ /不動産鑑定書 ⁽⁹⁾ /事業資産 ⁽⁹⁾ /保管管理 ⁽⁹⁾ /保険料経理額 ⁽⁹⁾ /借入契約 ⁽⁹⁾ /慎権部分 ⁽⁹⁾ /取得不動産 ⁽⁹⁾ /死亡時点 ⁽⁹⁾ /私法上の効果 ⁽⁹⁾ /納稅方法 ⁽⁹⁾ /解約時 ⁽⁹⁾ /評価時 ⁽⁹⁾ /財産の額 ⁽⁹⁾ /違法理由 ⁽⁹⁾ /還付を受けるべき者 ⁽⁹⁾ /鑑定評価基準 ⁽⁹⁾ /事実的 ⁽⁸⁾ /参別評価 ⁽⁸⁾ /国内に住所を有する者 ⁽⁸⁾ /契約証書 ⁽⁸⁾ /所得税の納付すべき税額 ⁽⁸⁾ /担保目的 ⁽⁸⁾ /旧氏統 ⁽⁸⁾ /業 第上の ⁽⁸⁾ /登記請求権 ⁽⁸⁾ /税額変更 ⁽⁹⁾ /算入すべき費用 ⁽⁸⁾ / 総所得 ⁽⁸⁾ / (18) 選年分 ⁽⁸⁾ / (18) 財産 関係 ⁽⁸⁾ / (19) 開与株式 ⁽⁸⁾ / 選定条件 ⁽⁸⁾ /配当利益 ⁽⁸⁾ /金銭としての性質 ⁽⁸⁾ /預託金請求権 ⁽⁸⁾ /不動産評価 ⁽⁷⁾ /で付者 ⁽⁷⁾ /代表者給与 ⁽⁷⁾ /、会計書類 ⁽⁷⁾ /(価格調查 ⁽⁷⁾ /保険金の額 ⁽⁷⁾ /信地権の対価 ⁽⁷⁾ /信地権取引 ⁽⁷⁾ /(債務免除日 ⁽⁷⁾ /【他 1,733件】
200_2048 第 2 階級	1,834	売上総利益算入(2)/売上総利益額(2)/売却契約書(2)/売却条件(2)/売却物件(2)/売却行為(2)/売買事実(2)/売買原価(2)/売買実例価格方式(2)/売買市場(2)/売買残代金相当額(2)/変更制度(2)/変更認可申請(2)/外国法人立法(2)/外国稅額控除規定(2)/外国親会社株式(2)/外注帳(2)/外注生産(2)/外注請求書(2)/大宴会(2)/大工事(2)/大規模土地(2)/大規模所有者(2)/契約の実行(2)/契約上の制限(2)/契約変更権(2)/契約目的(2)/契約行為(2)/契約開始(2)/委任事項(2)/委託基準(2)/委託基準(2)/委託販売取引(2)/婚姻の効力(2)/婚姻法秩序維持(2)/子供室(2)/学校敷地(2)/宅地交換契約(2)/宅地利用(2)/宅地建物(2)/宅地用(2)/宗全保護義務違反(2)/完了した株式(2)/完全親会社株式(2)/完成見学会(2)/完成費用(2)/宗教法人以外(2)/宗教目的活動(2)/定価表(2)/定年後(2)/実体的効果(2)/実成見学会(2)/完成費用(2)/宗教士人以外(2)/宗教目的活動(2)/定価表(2)/実理前第位(2)/実與時期(2)/実現義務(2)/実行可能性(2)/実質的不利益(2)/実質帰属者課稅の原則(2)/実質的な同一性(2)/実質的に同一性(2)/実質的に権利(2)/実質的不利益(2)/実際の取得費(2)/実額算出(2)/客観的行動(2)/客観的行動(2)/客観的な理由(2)/客観的に勘案(2)/客観的に権利(2)/客観的に動制(2)/客観的行動(2)/家報的行動(2)/家報的作用(2)/家報的行動(2)/家事費額(2)/家事関連的(2)/家庭としての効用(2)/家屋の用途(2)/家庭生活(2)/家族労働(2)/家族単位(2)/家族従業員数(2)/家族法(2)/家族経営体(2)/家質地代収入(2)/家賃支払(2)/寄附を受ける者(2)/寄附金の交付(2)/寄附金等の額(2)/寄附金除外(2)/審判所比準価格(2)/審判所開発法(2)/審查制度(2)/審查書まだ(2)/審查書制度(2)/審查書表(2)/新企等の額(2)/新研金除外(2)/審判所比準価格(2)/需判所開発法(2)/審查制度(2)/審查書表(2)/審查書書表(2)/審查書表(2)/对価の(2)/対価の決定(2)/対価支払(2)/対応特計(2)/【他 1,724件】

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの(ただし、同じ認識回数では文字数のより多い順)
200_2048 第 3 階級	1,834	事業税課税 ⁽¹⁾ /事業規模判断 ⁽¹⁾ /事業費の総額 ⁽¹⁾ /事業資金捻出 ⁽¹⁾ /事業追加 ⁽¹⁾ /事業開始時 ⁽¹⁾ /事業関連支出 ⁽¹⁾ /事業類型ごと ⁽¹⁾ /事由が生じたことを知った場合 ⁽¹⁾ /事由が生じた旨 ⁽¹⁾ /二会計年度 ⁽¹⁾ /ご時点 ⁽¹⁾ /二期 ⁽¹⁾ /二次 ⁽¹⁾ /二面地 ⁽¹⁾ /二重の制裁 ⁽¹⁾ /三重の損失 ⁽¹⁾ /二重売上 ⁽¹⁾ /互助的 ⁽¹⁾ /交付された日 ⁽¹⁾ /交付に要した費用 ⁽¹⁾ /交付を受ける権利 ⁽¹⁾ /交検率 ⁽¹⁾ /交渉価額 ⁽¹⁾ /交沙請求書 ⁽¹⁾ /交通相談 ⁽¹⁾ /交通 推費 ⁽¹⁾ /交際費損金不算入制度 ⁽¹⁾ /京都店 ⁽¹⁾ /京都府八幡市 ⁽¹⁾ /人事課職員 ⁽¹⁾ /人件費支給総額 ⁽¹⁾ /人口座 ⁽¹⁾ /人員管理 ⁽¹⁾ /人外,大派遣費 ⁽¹⁾ /人格権的権利 ⁽¹⁾ /人格的な利益 ⁽¹⁾ /人作費支給総額 ⁽¹⁾ /人為的な行為 ⁽¹⁾ /人的信頼関係 ⁽¹⁾ /人的役務の提供の対価 ⁽¹⁾ /小商標下的支援事業契約 ⁽¹⁾ /介護体制 ⁽¹⁾ /介護保護 ⁽¹⁾ /介護保険料相当額 ⁽¹⁾ /介護を受ける権利 ⁽¹⁾ /介護予防支援事業契約 ⁽¹⁾ /介護保定 ⁽¹⁾ /位身小売業 ⁽¹⁾ /仕上げ業 ⁽¹⁾ /仕事の終了 ⁽¹⁾ /仕入れに係る記載 ⁽¹⁾ /仕入れに係る金額 ⁽¹⁾ /仕入れに対する対価の返還 ⁽¹⁾ /仕入れに関しての ⁽¹⁾ /仕入れに係る記載 ⁽¹⁾ /仕入取に係る金額 ⁽¹⁾ /仕入和に関しての ⁽¹⁾ /仕入が ⁽¹⁾ /位外 ⁽¹⁾ /位入が ⁽¹⁾ /位入 ⁽¹⁾ /付入 ⁽¹⁾ /位入 ⁽¹⁾ /(位入 ⁽¹⁾ /(心入 ⁽¹⁾ /(心入 ⁽¹⁾ /(心入 ⁽¹⁾ /(心入 ⁽¹⁾ /(心入 ⁽¹⁾ /(心入
200_2048 第 4 階級	1,834	工事前受金 ⁽¹⁾ /工事前渡金 ⁽¹⁾ /工事協同組合 ⁽¹⁾ /工事対価 ⁽¹⁾ /工事日程 ⁽¹⁾ /工事検査通知書 ⁽¹⁾ /工事監督 ⁽¹⁾ /工事自的 ⁽¹⁾ /工事金受取証明 ⁽¹⁾ /工業用原 ⁽¹⁾ /左官工事 ⁽¹⁾ /差入れ保証手形 ⁽¹⁾ /差引減少 ⁽¹⁾ /差引金額 ⁽¹⁾ /差押えの日 ⁽¹⁾ /差押之処分 ⁽¹⁾ /差押え通知書 ⁽¹⁾ /差押处分取消訴訟 ⁽¹⁾ /差押対象債権 ⁽¹⁾ /差金の額 ⁽¹⁾ /差額開稅 ⁽¹⁾ /市場の価格 ⁽¹⁾ /市場の管定 ⁽¹⁾ /市場流通性減価 ⁽¹⁾ /市場状況 ⁽¹⁾ /市場用建物 (1) ⁽¹⁾ /市場開設 (1) ⁽¹⁾ /市場所係者 (1) ⁽¹⁾ /市街化区域農業地 (1) ⁽¹⁾ /市街地の宅地 (1) ⁽¹⁾ /市道正事 (1) ⁽¹⁾ /布施回向料 (1) ⁽¹⁾ /条薄性 (1) ⁽¹⁾ /等的 (1) ⁽¹⁾ /等等公司。 (1) ⁽¹⁾ /等。 (1) ⁽¹⁾

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
200_2048 第 5 階級	1,835	税務上の損金 $^{(1)}$ 税務上の計算 $^{(1)}$ 税務代 $^{(1)}$ /稅務会計業務 $^{(1)}$ /稅務協定 $^{(1)}$ /稅務契約書 $^{(1)}$ /稅務百里課 $^{(1)}$ /稅務署整理 $^{(1)}$ /稅務國務資料 $^{(1)}$ /稅務的區定 $^{(1)}$ /稅務管理課 $^{(1)}$ /稅務署整理 $^{(1)}$ /稅務國務資料 $^{(1)}$ /稅回避 $^{(1)}$ /稅間的当期 $^{(1)}$ /稅別後 $^{(1)}$ /稅及查面額 $^{(1)}$ /稅課工業承継 $^{(1)}$ /稅率の変更 $^{(1)}$ /稅的 $^{(1)}$ /稅退付 $^{(1)}$ /稅之納付義務 $^{(1)}$ /稅理工業承継 $^{(1)}$ 代稅里土法人でない者 $^{(1)}$ /稅的 $^{(1)}$ /稅退付 $^{(1)}$ /稅金約付義務 $^{(1)}$ /稅租土業承继 $^{(1)}$ /稅限工法人でない者 $^{(1)}$ /稅的 $^{(1)}$ /稅國和正如分 $^{(1)}$ /稅額算定方式 $^{(1)}$ /稅金負担額 $^{(1)}$ /稅額負担 $^{(1)}$ /稅國制的 $^{(1)}$ /稼働可能年数 $^{(1)}$ /稼働更整 $^{(1)}$ /積回时 $^{(1)}$ /積值的な評価 $^{(1)}$ /積面負担金 $^{(1)}$ /積立於。 $^{(1)}$ /積算過程 $^{(1)}$ /家回对吃 $^{(1)}$ /窓口引受時間 $^{(1)}$ /立合せ $^{(1)}$ /立始人 $^{(1)}$ 立替人 $^{(1)}$ 公营材、 $^{(1)}$ 公营材、 $^{(1)}$ 或证。 $^{(1)}$ 少或正风 $^{(1)}$ 9第二码 $^{(1)}$ 9第一码市街地再開発事業 $^{(1)}$ 9第二債務 $^{(1)}$ 9第二人。 $^{(1)}$ 9第二人。 $^{(1)}$ 9第二者以外の者 $^{(1)}$ 9第三者的需查裁决 $^{(1)}$ 9第三者公会、 $^{(1)}$ 9第二者以外の者 $^{(1)}$ 9第三者对抗力 $^{(1)}$ 19第三者的需查裁决 $^{(1)}$ 9第三者立会、 $^{(1)}$ 9第二者以外の者 $^{(1)}$ 9第三者对抗力 $^{(1)}$ 19第三者的需查裁决 $^{(1)}$ 9第三者立会、 $^{(1)}$ 9第1230月第1230号 $^{(1)}$ 9第1243号 $^{(1)}$ 9第13号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第13号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第13日3号 $^{(1)}$ 9第12日9号 $^{(1)}$ 9第12日9 $^{(1)}$ 9第13日 $^{(1)}$ 9第1日
100_4096 第1階級	3,533	第39 ⁽²¹⁹⁾ /所得税の申告 ⁽¹³⁵⁾ /計算の基礎となる ⁽¹²⁴⁾ /事業年度の所得 ⁽¹¹⁸⁾ /土地の所有者 ⁽¹⁰⁹⁾ /居住の用に供していた ⁽⁹⁵⁾ /所得額 ⁽⁸²⁾ /消費税額等 ⁽⁷⁶⁾ /相続税の ⁽⁷⁶⁾ /平成16年3月 ⁽⁶³⁾ /国税通則法23条 ⁽⁶²⁾ /加算税賦深決定処分 ⁽⁶⁰⁾ /生計を一にして ⁽⁶⁰⁾ /相続税の負担 ⁽⁵⁵⁾ /不相当に高額な部分 ⁽⁵³⁾ /給与所得の金額 ⁽⁵¹⁾ /相続税の納税義務 ⁽⁵⁰⁾ /課税银 ⁽⁵⁰⁾ /課稅化入れに係る消費稅額 ⁽⁴⁶⁾ /平成28年分 ⁽⁴⁶⁾ /これに類する ⁽⁴⁴⁾ /贈与税の負担 ⁽⁴⁴⁾ /申告稅額 ⁽⁴⁴⁾ /启当額 ⁽⁴⁴⁾ /合理的な説明 ⁽⁴³⁾ /土地の所有権移転登記 ⁽⁴⁰⁾ /資金利息 ⁽⁴⁰⁾ /適用を受ける旨の ⁽³⁹⁾ /平成16年6月 ⁽³⁹⁾ /所得の特別控除 ⁽³⁷⁾ /事業用資産の ⁽³⁷⁾ /不動産の評価 ⁽³⁶⁾ /土地取得 ⁽³⁶⁾ /不当な ⁽³⁶⁾ /平成19年1月 ⁽³⁴⁾ /独立した ⁽³⁴⁾ /第二次 ⁽³⁴⁾ /贈与証 ⁽³⁴⁾ /平成19年2月 ⁽³³⁾ /平成20年2月 ⁽³²⁾ /契約当事者間 ⁽³²⁾ /計上額 ⁽³²⁾ /事業の財に供していた ⁽³¹⁾ /課税処分取消しの訴え ⁽³⁰⁾ /平成20年5月 ⁽³⁰⁾ /当事者の主張 ⁽³⁰⁾ /相続の開始前3年以内 ⁽²⁹⁾ /事業必遂行 ⁽²⁹⁾ /更正処分庁 ⁽²⁹⁾ /販売事業 ⁽²⁹⁾ /利益金 ⁽²⁹⁾ /処分取消しの訴え ⁽²⁸⁾ /所得稅法基本通達 ⁽²⁸⁾ /平成 ⁽²⁸⁾ /評稅仕入れに係る消費稅額空控 ⁽²⁷⁾ /納付すべき消費稅等 ⁽²⁶⁾ /平成 18年2月 ⁽²⁶⁾ /帳簿の記載 ⁽²⁶⁾ /三額稅 (27)/平成 18年8月 ⁽²⁶⁾ /平成 29年分 ⁽²⁵⁾ /平正の理由附記の程度 ⁽²⁶⁾ /平成 18年2月 ⁽²⁶⁾ /帳簿の記載 ⁽²⁶⁾ /三額稅 (26) (国稅に関する法律の規定 ⁽²⁷⁾ /納付すべき消費稅等 ⁽²⁶⁾ /平成 18年2月 ⁽²⁶⁾ /平成 29年分 ⁽²⁵⁾ /平成 18年8月 ⁽²⁵⁾ /平成 29年分 ⁽²⁵⁾ /平成 18年8月 ⁽²⁵⁾ /平成 16年9月 ⁽²⁴⁾ /株式等の ⁽²⁴⁾ /正拠関係 ⁽²⁵⁾ /上告会社 ⁽²⁵⁾ /互制税関係 ⁽²⁵⁾ /一成積 ⁽²⁶⁾ /一級積 ⁽²⁶⁾ /平成 16年9月 ⁽²⁶⁾ /株式等の ⁽²⁴⁾ /正拠関係 ⁽²⁶⁾ /中成時 ⁽²⁶⁾ /小規模宅地等の ⁽²³⁾ /譲渡資産の譲渡 ⁽²³⁾ /所得の計算上 ⁽²³⁾ /修正申告額 ⁽²³⁾ /市成時 ⁽²³⁾ /作成時 ⁽²³⁾ /決定的 ⁽²³⁾ /相続後 ⁽²³⁾ /1棟 ⁽²³⁾ /異議決定取消しの 訴え ⁽²²⁾ /【他 3,443件】

言語モデル 5分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
100_4096第2階級	3,533	銀行債務 ⁽³⁾ /銀行届出 ⁽³⁾ /開差本数 ⁽³⁾ /開業支援 ⁽³⁾ /関係強化 ⁽³⁾ /集合郵便 ⁽³⁾ /青色更正 ⁽³⁾ /預入資金 ⁽³⁾ /風俗事業 ⁽³⁾ /養母名義 ⁽³⁾ /駐車可能 ⁽³⁾ /黑服従業 ⁽³⁾ /お渡し ⁽³⁾ /一回目 ⁽³⁾ /一貫氏 ⁽³⁾ /中央区 ⁽³⁾ /事実の ⁽³⁾ /二重払 ⁽³⁾ /交換用 ⁽³⁾ /低廉性 ⁽³⁾ /住基法 ⁽³⁾ /作成前 ⁽³⁾ /使途先 ⁽³⁾ /参正の ⁽³⁾ /制定時 ⁽³⁾ /割合的 ⁽³⁾ /别類 ⁽³⁾ /加工用 ⁽³⁾ /迎額 (3)/平设面 ⁽³⁾ /印盤紙 ⁽³⁾ /以用時 ⁽³⁾ /以為源 ⁽³⁾ /以初时第 ⁽³⁾ /取处分 ⁽³⁾ /取立額 ⁽³⁾ /受入額 ⁽³⁾ /受入者 ⁽³⁾ /房进者 ⁽³⁾ /和原時 ⁽³⁾ /常原業 ⁽³⁾ /土地局 ⁽³⁾ /在籍中 ⁽³⁾ /基本法 ⁽³⁾ /大津市 ⁽³⁾ /等人義 ⁽³⁾ /资清图的 ⁽³⁾ /平均酒 ⁽³⁾ /年金中 ⁽³⁾ /床仕上 ⁽³⁾ /皮渡法 ⁽³⁾ /旋内神 ⁽³⁾ /延利金 ⁽³⁾ /引揚げ ⁽³⁾ /第名義 ⁽³⁾ /後背地 ⁽³⁾ /微収後 ⁽³⁾ /感覚的 ⁽³⁾ /慣行的 ⁽³⁾ /手引書 ⁽³⁾ /投票法 ⁽³⁾ /规出時 ⁽³⁾ /推計值 ⁽³⁾ /提起時 ⁽³⁾ /激炒外外 ⁽³⁾ /方式化 ⁽³⁾ /施设の ⁽³⁾ /施設の ⁽³⁾ /施設内 ⁽³⁾ /旧水路 ⁽³⁾ /用波时 ⁽³⁾ /期細の ⁽³⁾ /曳航船 ⁽³⁾ /最高限 ⁽³⁾ /有形力 ⁽³⁾ /清直性 ⁽³⁾ /未决定 ⁽³⁾ /条件下 ⁽³⁾ /構造上 ⁽³⁾ /正常な ⁽³⁾ /正当な ⁽³⁾ /水平的 ⁽³⁾ /派遣時 ⁽³⁾ /清算時 ⁽³⁾ /清了時 ⁽³⁾ /無意義 ⁽³⁾ /提却場 ⁽³⁾ /理解力 ⁽³⁾ /申告の ⁽³⁾ /申古時 ⁽³⁾ /第三力 ⁽³⁾ /的中者 ⁽³⁾ /県三条 ⁽³⁾ /真実の ⁽³⁾ /为代金 ⁽³⁾ /就知策 ⁽³⁾ /第32 ⁽³⁾ /第4次 ⁽³⁾ /前便的 ⁽³⁾ /组金率 ⁽³⁾ /終結時 ⁽³⁾ /接段的 ⁽³⁾ /義於製的 ⁽³⁾ /義成取 ⁽³⁾ /正常数 ⁽³⁾ /或上,一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个
100_4096 第 3 階級	3,533	日給額 ⁽²⁾ /旧債務 ⁽²⁾ /旧債権 ⁽²⁾ /旧基準 ⁽²⁾ /旧条項 ⁽²⁾ /時效既 ⁽²⁾ /時給額 ⁽²⁾ /居年中 ⁽²⁾ /暫定払 ⁽²⁾ /更正の ⁽²⁾ /書類中 ⁽²⁾ /最支重变 ⁽²⁾ /最長期 ⁽²⁾ /最大能位 ⁽²⁾ /素於位 ⁽²⁾ /素於位 ⁽²⁾ /表於位 ⁽²⁾ /表於位 ⁽²⁾ /表於位 ⁽²⁾ /表於位 ⁽²⁾ /表於位 ⁽²⁾ /表於也 ⁽²⁾ /持計等 ⁽²⁾ /格於位 ⁽²⁾ /申込時 ⁽²⁾ /申述べ ⁽²⁾ /異議等 ⁽²⁾ /発展的 ⁽²⁾ /的見解 ⁽²⁾ /前接客 ⁽²⁾ /相正会 ⁽²⁾ /相当の ⁽²⁾ /相に性 ⁽²⁾ /相及類 ⁽²⁾ /社内称 ⁽²⁾ /提回市 ⁽²⁾ /规可下位 ⁽²⁾ /新正型 ⁽²⁾ /研磨材 ⁽²⁾ /或於時 ⁽²⁾ /在生文 ⁽²⁾ /社內外 ⁽²⁾ /社內外 ⁽²⁾ /社同化 ⁽²⁾ /折祷料 ⁽²⁾ /祝日割 ⁽²⁾ /福井県 ⁽²⁾ /秘書室 ⁽²⁾ /整出時 ⁽²⁾ /移植後 ⁽²⁾ /松政策 ⁽²⁾ /立会葬 ⁽²⁾ /立法時 ⁽²⁾ /競争用 ⁽²⁾ /第四次 ⁽²⁾ /等質性 ⁽²⁾ /算定の ⁽²⁾ /算定書 ⁽²⁾ /来国産 ⁽²⁾ /条附接 ⁽²⁾ /新定额 ⁽²⁾ /新的方 ⁽²⁾ /新明分 ⁽²⁾ /納明分 ⁽²⁾ /納骨箱 ⁽²⁾ /累積数 ⁽²⁾ /案進性 ⁽²⁾ /疑理の ⁽²⁾ /統計法 ⁽²⁾ /総会員 ⁽²⁾ /総括的 ⁽²⁾ /総額的 ⁽²⁾ /総食数 ⁽²⁾ /繁華性 ⁽²⁾ /的体的 ⁽²⁾ /自作地 ⁽²⁾ /自家製 ⁽²⁾ /自然的 ⁽²⁾ /的之 ⁽²⁾ /前空便 ⁽²⁾ /表於死 ⁽²⁾ /表於死 ⁽²⁾ /見以之 ⁽²⁾ /就是於的 ⁽²⁾ /期与的 ⁽²⁾ /就是於的 ⁽²⁾ /期停案 ⁽²⁾ /提降的 ⁽²⁾ /期子之 ⁽²⁾ /最味的 ⁽²⁾ /買以表 ⁽²⁾ /置於的 ⁽²⁾ /表述大 ⁽²⁾ /数述大 ⁽²⁾ /题与の ⁽²⁾ /贈与の ⁽²⁾ /贈号来 ⁽²⁾ /提晓的 ⁽²⁾ /期停案 ⁽²⁾ /提陈的 ⁽²⁾ /新成 ⁽²⁾ /更以为的 ⁽²⁾ /表述大 ⁽²⁾ /数述为 ⁽²⁾ /表述大 ⁽²⁾ /数述为 ⁽²⁾ /表述大 ⁽²⁾ /数述大 ⁽²⁾ /数述为 ⁽²⁾ /表述大 ⁽²⁾ /表述大 ⁽²⁾ /数本大 ⁽²⁾ /表述大 ⁽

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの(ただし、同じ認識回数では文字数のより多い順)
100_4096 第 4 階級	3,533	社団設立行為 ⁽¹⁾ /福祉施設工事 ⁽¹⁾ /私法上の事実 ⁽¹⁾ /私法上の解釈 ⁽¹⁾ /租税ほ脱の罪 ⁽¹⁾ /租税債権侵害 ⁽¹⁾ /租稅公課関係 ⁽¹⁾ /租税判例年報 ⁽¹⁾ /租稅徵収事務 ⁽¹⁾ /租稅徵収費用 ⁽¹⁾ /租稅帶納処分 ⁽¹⁾ /租稅養務違反 ⁽¹⁾ /租稅軽減措置 ⁽¹⁾ /稅務署資産稅 ⁽¹⁾ /稅務署長名義 ⁽¹⁾ /稅務関係通達 ⁽¹⁾ /稅引後手取額 ⁽¹⁾ /稅理士の利益 ⁽¹⁾ /稅理士の展院 ⁽¹⁾ /稅理士の展院 ⁽¹⁾ /稅理士の展院 ⁽¹⁾ /稅理士の展院 ⁽¹⁾ /稅理士の展院 ⁽¹⁾ /稅理士の通院 ⁽¹⁾ /稅理士の選院 ⁽¹⁾ /稅理士の選院 ⁽¹⁾ /稅理士の通院 ⁽¹⁾ /稅理士の選院 ⁽¹⁾ /稅期全更处分 ⁽¹⁾ /稅額変更处分 ⁽¹⁾ /稅額確定手統 ⁽¹⁾ /養医-書亦於「人養を請求權 ⁽¹⁾ /競三者立会下(1)/第三者納付分 ⁽¹⁾ /第三本納付前 ⁽¹⁾ /第三者に高額 ⁽¹⁾ /第三者の債務 ⁽¹⁾ /第二本の計算 ⁽¹⁾ /第三者立会下(1)/第三者納付分 ⁽¹⁾ /第三者納付前 ⁽¹⁾ /第三者納付後 ⁽¹⁾ /第二次的權利 ⁽¹⁾ /第127号 ⁽¹⁾ /第107号 ⁽¹⁾ /第110号 ⁽¹⁾ /第1137号 ⁽¹⁾ /第1207号 ⁽¹⁾ /第127号 ⁽¹⁾ /第128号号 ⁽¹⁾ /第110号 ⁽¹⁾ /第11号号 ⁽¹⁾ /第127号 ⁽¹⁾ /第128号
100_4096 第 5 階級	3,534	当初価格 ⁽¹⁾ /当初手形 ⁽¹⁾ /当初決算 ⁽¹⁾ /当初渡さ ⁽¹⁾ /当初適用 ⁽¹⁾ /当然利息 ⁽¹⁾ /当然調查 ⁽¹⁾ /当該他の ⁽¹⁾ /当面資金 ⁽¹⁾ /形成状況 ⁽¹⁾ /役務の提 ⁽¹⁾ /役員たる ⁽¹⁾ /役員体制 ⁽¹⁾ /役員改選 ⁽¹⁾ /待機場所 ⁽¹⁾ /待機時間 ⁽¹⁾ /後原処分 ⁽¹⁾ /後妻名義 ⁽¹⁾ /従事員中 ⁽¹⁾ /侵命內容 ⁽¹⁾ /復高內容 ⁽¹⁾ /復活割含 ⁽¹⁾ /微税整理 ⁽¹⁾ /心理状態 ⁽¹⁾ /必修科目 ⁽¹⁾ /必要 + 分 ⁽¹⁾ /必要要素 ⁽¹⁾ /心募学生 ⁽¹⁾ /心能稅的 ⁽¹⁾ /念書記載 ⁽¹⁾ /性質目的 ⁽¹⁾ /恣意性の ⁽¹⁾ /患者專用 ⁽¹⁾ /恶質重大 ⁽¹⁾ /情勢変化 ⁽¹⁾ /情報誰 ⁽¹⁾ /忧形之的 ⁽¹⁾ /意義內容 ⁽¹⁾ /意見对立 ⁽¹⁾ /態度方法 ⁽¹⁾ /慰安目的 ⁽¹⁾ /惠遗的 ⁽¹⁾ /的懲役前科 ⁽¹⁾ /成り行き ⁽¹⁾ /成形製品 ⁽¹⁾ /成鴻到数 ⁽¹⁾ /户建住居 ⁽¹⁾ /户建往宅 ⁽¹⁾ /户数契约 ⁽¹⁾ /所內調查 ⁽¹⁾ /所存制的 ⁽¹⁾ /所得制的 ⁽¹⁾ /所得合算 ⁽¹⁾ /所得減少 ⁽¹⁾ /所得稅の ⁽¹⁾ /所得額の ⁽¹⁾ /所有世 ⁽¹⁾ /所有山林 ⁽¹⁾ /所有概念 ⁽¹⁾ /所要経費 ⁽¹⁾ /所長代理 ⁽¹⁾ /手取収入 ⁽¹⁾ /手形 ⁽¹⁾ /手形取立 ⁽¹⁾ /手形面上 ⁽¹⁾ /手能代行 ⁽¹⁾ /手。保留 ⁽¹⁾ /手形内容 ⁽¹⁾ /手形面 ⁽¹⁾ /手。依代行 ⁽¹⁾ /手。保留 ⁽¹⁾ /手。形成 ⁽¹⁾ /手。形成 ⁽¹⁾ /手。形成 ⁽¹⁾ /于,表述 ⁽¹⁾ /并。 ⁽¹⁾ /本证 一次 ⁽¹⁾ /并定 中的 ⁽¹⁾ /并定 中的 ⁽¹⁾ /并定 中的 ⁽¹⁾ /持定 中的 ⁽¹⁾ /指定 中的 ⁽¹⁾ /指定 中的 ⁽¹⁾ /指定 中的 ⁽¹⁾ /指示管理 ⁽¹⁾ /指输代金 ⁽¹⁾ /持续 ⁽¹⁾ /指读場所 ⁽¹⁾ /指读場所 ⁽¹⁾ /指读場所 ⁽¹⁾ /指读等 ⁽¹⁾ /指示管理 ⁽¹⁾ /指输代金 ⁽¹⁾ /指读 ⁽¹⁾ /注读 ⁽¹⁾ /

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
50_8192 第 1 階級	3,545	事実上 ⁽⁵⁸⁵⁾ /更正請求 ⁽³⁷²⁾ /第1需 ⁽²⁸⁸⁾ /相続税の申告書 ⁽²⁶⁹⁾ /基本的 ⁽²⁶⁵⁾ /納税者主張 ⁽²³²⁾ /平成2 ⁽²⁰³⁾ / <u>資産の譲渡代金</u> ⁽¹⁸⁹⁾ /法律上 ⁽¹⁷⁸⁾ /その他これ ⁽¹⁷³⁾ /特 <u>たの事業用資産</u> ⁽¹⁷⁰⁾ /課税上 ⁽¹⁶⁹⁾ /第22条 ⁽¹⁶⁴⁾ /第14条 ⁽¹⁴⁶⁾ /業 ⁽¹⁴¹⁾ /時間的 ⁽¹³³⁾ /主位的 ⁽¹¹⁹⁾ / <u>株式の譲渡</u> ⁽¹¹⁷⁾ /旧所得税法 ⁽¹¹⁰⁾ /当初更正処分 ⁽¹⁰⁸⁾ /納税者間 ⁽¹⁰²⁾ /納税告間 ⁽¹⁰²⁾ /納税告間 ⁽¹⁰²⁾ /土地の取得費 ⁽⁶⁶⁾ /一団の ⁽⁹⁴⁾ /平成13年分 ⁽⁹⁰⁾ /第13条 ⁽⁹⁰⁾ /第13条 ⁽⁹⁰⁾ /平成12年分 ⁽⁸⁷⁾ / <u>不動産の譲渡</u> ⁽⁸⁵⁾ /第26条 ⁽⁸²⁾ /第28条 ⁽⁸⁰⁾ /総務 ⁽⁸⁰⁾ /第19条 ⁽⁷³⁾ /土地上 ⁽⁷⁸⁾ /計算すべき ⁽⁷⁶⁾ /譲渡所得の <u>特例</u> ⁽⁷⁵⁾ /第12条 ⁽⁷⁵⁾ /具体化 ⁽⁷⁴⁾ /社会的 ⁽⁷³⁾ /不作為の ⁽⁷²⁾ /贈与等 ⁽⁷¹⁾ /帳簿上 ⁽⁷⁰⁾ /相続財産の評価 ⁽⁶⁹⁾ /担続財産の価額 ⁽⁸⁸⁾ /納税者の訴え ⁽⁶⁸⁾ /損金計上 ⁽⁶⁸⁾ /差別的 ⁽⁶⁷⁾ /第27条 ⁽⁶⁵⁾ /事業から対価 ⁽⁶³⁾ /実質上 ⁽⁶²⁾ /税 <u>負担の公平</u> ⁽⁶¹⁾ /著しく地積 ⁽⁶¹⁾ /第1条 ⁽⁶¹⁾ / <u>居法上の組合契約</u> ⁽⁶⁰⁾ /当初の ⁽⁵⁸⁾ /第17条 ⁽⁶⁷⁾ /国 <u>開法</u> ⁽⁵⁶⁾ /
50_8192 第 2 階級	3,545	著しくかい難 ⁽³⁾ /著しく収益性 ⁽³⁾ /虚偽の申立て ⁽³⁾ /行政上の決定 ⁽³⁾ /製本加工業者 ⁽³⁾ /見積原価計算 ⁽³⁾ /計上漏れ割合 ⁽³⁾ /計算すべき旨 ⁽³⁾ /計算上の価額 ⁽³⁾ /計算上の分配 ⁽³⁾ /課税上の便宜 ⁽³⁾ /課税上の効果 ⁽³⁾ /課税売上対応 ⁽³⁾ /調査上の必要 ⁽³⁾ /調査担当係官 ⁽³⁾ /貸付業務の用 ⁽³⁾ /賃貸業務の用 ⁽³⁾ /資産の償却費 ⁽³⁾ /近代化補助金 ⁽³⁾ /退職給与支給 ⁽³⁾ /通常貯金口座 ⁽³⁾ /連鎖販売業者 ⁽³⁾ /道路敷設地積 ⁽³⁾ /適正利率超過 ⁽³⁾ /遺言書の効力 ⁽³⁾ /酒販免許制度 ⁽³⁾ /金銭の無利息 ⁽³⁾ /鑑定評価事例 ⁽³⁾ /無察時負担金 ⁽³⁾ /便連法人名義 ⁽³⁾ /雇用契約関係 ⁽³⁾ /電気メツキ業 ⁽³⁾ /音の他转別 ⁽³⁾ /ま公知の事項 ⁽³⁾ /その他医療 ⁽³⁾ /その他建物 ⁽³⁾ /その他方 ⁽³⁾ /その他日常 ⁽³⁾ /その他特別 ⁽³⁾ /その他責任 ⁽³⁾ /ウィンドウ ⁽³⁾ /エックス線 ⁽³⁾ /エンジニア ⁽³⁾ /ガレージ代 ⁽³⁾ /クーボン債 ⁽³⁾ /バー事業者 ⁽³⁾ /フラッシュ ⁽³⁾ /ブライバシ ⁽³⁾ /ホステス数 ⁽³⁾ /下応合理性 ⁽³⁾ /上毛補償科 ⁽³⁾ /不動産差押 ⁽³⁾ /不当な意別 ⁽³⁾ /不当な差別 ⁽³⁾ /不当に利益 ⁽³⁾ /不販審査会 ⁽³⁾ 世帯主一人 ⁽³⁾ /主に代表者 ⁽³⁾ /主に請求人 ⁽³⁾ /争点主義的 ⁽³⁾ /事業認定時 ⁽³⁾ /交付決定額 ⁽³⁾ /交際費勘定 ⁽³⁾ /人別支給額 ⁽³⁾ /他の年度分 ⁽³⁾ /他人の債務 ⁽³⁾ /仮定的主張 ⁽³⁾ /企業会計上 ⁽³⁾ /会社分割の ⁽³⁾ /情家権消滅 ⁽³⁾ /債務の免脱 ⁽³⁾ /傾斜地崩壊 ⁽³⁾ /微力の過機が ⁽³⁾ /で成定的主張 ⁽³⁾ /企業会計上 ⁽³⁾ /会社分割の ⁽³⁾ /情家権消滅 ⁽³⁾ /債務の免脱 ⁽³⁾ /傾斜地崩壊 ⁽³⁾ /微力の認及 ⁽³⁾ /元本損害額 ⁽³⁾ /公正の確保 ⁽³⁾ /共同の事業 ⁽³⁾ /円控除特例 ⁽³⁾ /再び処分庁 ⁽³⁾ /出コンベヤ ⁽³⁾ /分割の遡及 ⁽³⁾ /利益の減少 ⁽³⁾ /前年の申告 ⁽³⁾ /分力の程度 ⁽³⁾ /用り意義 ⁽³⁾ /取引の理由 ⁽³⁾ /司法警察員 ⁽³⁾ /同等の設備 ⁽³⁾ /名義請求書 ⁽³⁾ /名義部品費 ⁽³⁾ /商工会会費 ⁽³⁾ /器具小売業 ⁽³⁾ /加可四理由 ⁽³⁾ /一末税務署 ⁽³⁾ /大阪支店長 ⁽³⁾ /定期預貯金 ⁽³⁾ /実質的改変 ⁽³⁾ /実額近似値 ⁽³⁾ /就労の事実 ⁽³⁾ /【他 3,437件】

言語モデル 5分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
50_8192 第 3 階級	3,545	靴小売商 ⁽²⁾ 音楽部門 ⁽²⁾ 順次移転 ⁽²⁾ 預かつて ⁽²⁾ 預入状況 ⁽²⁾ 預金出金 ⁽²⁾ 領事任務 ⁽²⁾ 食品製造 ⁽²⁾ 飲食料率 ⁽²⁾ 飼育農業 ⁽²⁾ 駅前共同 ⁽²⁾ 駆け引き ⁽²⁾ 鳴門市長 ⁽²⁾ うえ他 ⁽²⁾ お願い ⁽²⁾ くる子 ⁽²⁾ に脱の ⁽²⁾ とちろん ⁽²⁾ アウト ⁽²⁾ エキス ⁽²⁾ エンド ⁽²⁾ グラム ⁽²⁾ フンバ ⁽²⁾ シアン ⁽²⁾ スセオ ⁽²⁾ ダイヤ ⁽²⁾ グラム ⁽²⁾ カンバ ⁽²⁾ カンが、では、では、では、では、では、では、では、では、では、では、では、では、では、
50_8192 第 4 階級	3,545	必要経費程度 ⁽¹⁾ /情報要請行為 ⁽¹⁾ /意思決定方法 ⁽¹⁾ /意思決定過程 ⁽¹⁾ /所得の原則的 ⁽¹⁾ /所得合算制度 ⁽¹⁾ /所得字額立証 ⁽¹⁾ /所得申告行為 ⁽¹⁾ /所得発生原因 ⁽¹⁾ /所得発生時点 ⁽¹⁾ /所得税の金額 ⁽¹⁾ /所得算出方法 ⁽¹⁾ / 所有名義移転 ⁽¹⁾ /所管課長通知 ⁽¹⁾ /所謂発生主義 ⁽¹⁾ /手当たり次第 ⁽¹⁾ /手形回収業務 ⁽¹⁾ /手続ご指定日 ⁽¹⁾ /抱える大企業 ⁽¹⁾ /抵当権 / 劝力 ⁽¹⁾ /抹消請求訴訟 ⁽¹⁾ /抽出判断過程 ⁽¹⁾ /抽象的法規範 ⁽¹⁾ /担保権の効力 ⁽¹⁾ /拡張型心筋症 ⁽¹⁾ /持6回り方式 ⁽¹⁾ /指名入札業者 ⁽¹⁾ /挟むべき事情 ⁽¹⁾ /挟むべき理由 ⁽¹⁾ /振込関連行為 ⁽¹⁾ /推計課税期間 ⁽¹⁾ /提供取りやめ ⁽¹⁾ /接除前所得率 ⁽¹⁾ /推し量る指標 ⁽¹⁾ /推計基礎期間 ⁽¹⁾ /推計所得金額 ⁽¹⁾ /推計課税期間 ⁽¹⁾ /提供取りやめ ⁽¹⁾ /揚物小売業者 ⁽¹⁾ /提り許可申請 ⁽¹⁾ /損害金請求権 ⁽¹⁾ /損金性の評価 ⁽¹⁾ /推建設置工事 ⁽¹⁾ /支出の対価性 ⁽¹⁾ /支払いながら ⁽¹⁾ /支払い交際費 ⁽¹⁾ /支払い請求権 ⁽¹⁾ /支払い過払金 ⁽¹⁾ /支払う和解金 ⁽¹⁾ /支払う手数料 ⁽¹⁾ /支払う標準的 ⁽¹⁾ /支払う見舞金 ⁽¹⁾ /支払書受付印 ⁽¹⁾ /支払訴求権金 ⁽¹⁾ /支援事業契約 ⁽¹⁾ /改正法人税法 ⁽¹⁾ /心良費設備費 ⁽¹⁾ /支払う見舞金 ⁽¹⁾ /支払者受付印 ⁽¹⁾ /支払請求権金 ⁽¹⁾ /支提出命令 ⁽¹⁾ /新しい退職金 ⁽¹⁾ /新寿務所開設 ⁽¹⁾ /新株の払込金 ⁽¹⁾ /新株予約券付 ⁽¹⁾ /新聞等原価率 ⁽¹⁾ /下章提出命令 ⁽¹⁾ /新規投入資金 ⁽¹⁾ /日付付キック ⁽¹⁾ /日付け改正前 ⁽¹⁾ /日付け聴取書 ⁽¹⁾ /日付収用裁決 ⁽¹⁾ /日付委託契約 ⁽¹⁾ /日付消費税法 ⁽¹⁾ /日付付き理運営 ⁽¹⁾ /日付け改正前 ⁽¹⁾ /日本アマゾン ⁽¹⁾ /日本バスプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日付消費税法 ⁽¹⁾ /日付管理運営 ⁽¹⁾ /日払キャスト ⁽¹⁾ /日本アマゾン ⁽¹⁾ /日本バスプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日付消費税法 ⁽¹⁾ /日付管理運営 ⁽¹⁾ /日払キャスト ⁽¹⁾ /日本アマゾン ⁽¹⁾ /日本バスプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日付消費税法 ⁽¹⁾ /日付管理運営 ⁽¹⁾ /日払キャスト ⁽¹⁾ /日本アマゾン ⁽¹⁾ /日本バスプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日付消費税法 ⁽¹⁾ /日付管理運営 ⁽¹⁾ /日払キャスト ⁽¹⁾ /日本アマゾン ⁽¹⁾ /日本バスプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日付消費税法 ⁽¹⁾ /日本電信電話 ⁽¹⁾ /日の近務書長 ⁽¹⁾ /日本の大スプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日本派スプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日本銀行、「日本の大スプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日本の大スプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日本の大スプロ ⁽¹⁾ /日本国内業務 ⁽¹⁾ /日本の大スプロ ⁽¹⁾ /日本

98 研究ノート

言語モデル 5 分位階級	件数	当該階級のうち認識回数の多いもの (ただし、同じ認識回数では文字数のより多い順)
50_8192 第 5 階級	3,545	引落金類 ⁽¹⁾ /当初内定 ⁽¹⁾ /当初原告 ⁽¹⁾ /当初売却 ⁽¹⁾ /当初亳和 ⁽¹⁾ /当初亳和 ⁽¹⁾ /当初亳和 ⁽¹⁾ /当初克给 ⁽¹⁾ /当加克给 ⁽¹⁾ /当然常聚 ⁽¹⁾ /当然育用 ⁽¹⁾ /当面保有 ⁽¹⁾ /当面限统 ⁽¹⁾ /当面来字 ⁽¹⁾ /影響調查 ⁽¹⁾ /役員在任 ⁽¹⁾ /役員辞任 ⁽¹⁾ /待合室用 ⁽¹⁾ /後半部分 ⁽¹⁾ /说前主張 ⁽¹⁾ /從前勤務 ⁽¹⁾ /從前車集 ⁽¹⁾ /從前取得 ⁽¹⁾ /從前同樣 ⁽¹⁾ /從前建築 ⁽¹⁾ /從前控訴 ⁽¹⁾ /從前提出 ⁽¹⁾ /從前立 ⁽¹⁾ /從前日本 ⁽¹⁾ /從前課稅 ⁽¹⁾ /從前贈与 ⁽¹⁾ /從業員每 ⁽¹⁾ /御坊地区 ⁽¹⁾ /微収更員 ⁽¹⁾ /微取担当 ⁽¹⁾ /微税強化 ⁽¹⁾ /心当たり ⁽¹⁾ /必要やむ ⁽¹⁾ /必要月間 ⁽¹⁾ /忘年会兼 ⁽¹⁾ /応援演說 ⁽¹⁾ /応札状況 ⁽¹⁾ /急ぐ事情 ⁽¹⁾ /性格付け ⁽¹⁾ /恣意専斯 ⁽¹⁾ /意向確認 ⁽¹⁾ /意向